

Geographische Bedingungen im Information Retrieval:

Neue Ansätze in Systementwicklung und Evaluierung*

*Kerstin Bischoff, Ralph Kölle,
Thomas Mandl & Christa Womser-Hacker*

Informationswissenschaft
Universität Hildesheim
Marienburger Platz 22, D-31141 Hildesheim
mandl@uni-hildesheim.de

Zusammenfassung

Anfragen an Information Retrieval Systeme beinhalten häufig geographischen Einschränkungen. Solche geographischen Angaben wie Städte- oder Ländernamen ermöglichen Suchmaschinen die Optimierung ihrer Ergebnisse jenseits von klassischen Indexierungstechniken. Beispielsweise lassen sich geographische Entitäten in Ontologien einordnen und liefern somit gute Hinweise für Erweiterungsterme. Dieser Beitrag stellt den Stand der Forschung dar, analysiert aktuelle Evaluierungsergebnisse und leitet daraus Optimierungsansätze für die Systementwicklung ab. Experimente und deren Auswertung zeigen, dass flexible Anwendungen von Relevanz Feedback zu verbesserten Ergebnisse führen können.

Abstract

Queries with a geographic scope are often posted to information retrieval systems. Geographic search terms conditions like city or country names allow search engines to optimize their results beyond classic indexing techniques. Geographic entities can be identified within ontologies and be properly extended, for example. This paper briefly reviews the state of the art, analyses current evaluation results and presents new optimization techniques. Experiments show that the flexible application of relevance feedback can improve the retrieval results.

* Veröffentlicht in: OSSWALD, Achim; STEMPFHUBER, Maximilian; WOLFF, Christian (2007). Open Innovation. Proc. 10. Internationales Symposium für Informationswissenschaft. Konstanz: UVK, 15-26.

I Einleitung

Internet-Suchmaschinen gehören für viele Menschen bereits zum Alltag und bieten den wichtigsten Zugang zum Wissen im Internet. So werden nach Schätzungen täglich mehr als 500 Millionen Anfragen an Suchmaschinen gestellt¹, um unter Milliarden von Seiten im Internet die relevantesten auszusuchen. Die grundlegenden Verfahren für Suchmaschinen liefert das Information Retrieval, das heute eine der Schlüsseltechnologien in der Wissensgesellschaft darstellt. Information Retrieval behandelt die Repräsentation des Inhalts der Dokumente und den Abgleich zwischen Anfragen des Benutzers und den gespeicherten Repräsentationen.

Zahlreiche Anfragen an Suchmaschinen enthalten geographische Begriffe. Informationen müssen nicht nur inhaltlich relevant sein, sondern sollen sich auch konkret auf den angegebenen Ort beziehen. Diese geographischen Anfragen lassen sich vielfältig klassifizieren. Bei Eingaben wie „*Apotheke mit Nachtdienst in Nandlstadt*“ wünschen Benutzer einzelne Seiten mit konkreten Informationen, während „*Investitionen in Hastenbeck*“ auf eine vorab nicht festlegbare Menge von Dokumenten zu einem bestimmten Thema abzielt. Erstere werden im Web-Retrieval häufig als *navigational queries (known item search)* und die zweite Gruppe als *informational queries* (klassisches ad-hoc Retrieval) bezeichnet [Broder 2002].

Geographische Eigennamen können wie andere Eigennamen das Retrieval verbessern. Einige Suchmaschinen liefern teilweise bereits Ergebnisdokumente, welche zu Landkarten führen. Darüber hinaus bieten die geographischen Begriffe weitere Möglichkeiten zu formalen Schlüssen. Liegen Ontologien vor, so kann beispielsweise automatisch mit geeigneten Unterbegriffen erweitert werden (z. B. *Oberpfalz* -> *Nabburg, Wackersdorf, ...*).

Andererseits stellen Suchanfragen mit geographischen Eigennamen auch besondere Herausforderungen an die Anfragenverarbeitung. Die Erkennung und Klassifikation von Eigennamen gelingt heute noch nicht zu 100%. Der Gebrauch für politische als auch geographische Einheiten (*Europa*), metaphorischer Einsatz (*Tal der Ahnungslosen*), vage geographische Bezeichnungen (*Mittelmeerraum, Ostbayern, Tropen*) oder Mehrwortgruppen (*Conselheiro Lafaiete, Neunburg vorm Wald*) seien als Beispiele für die Probleme bei der Erkennung und Verarbeitung genannt. Gerade in mehrsprachigen Kontexten verschärfen sich die Schwierigkeiten für die Systeme. Manche geographischen Eigennamen müssen übersetzt oder zerlegt werden (z. B.

¹ <http://searchenginewatch.com>

Deutsch: *Donau*, Spanisch: *Danubio*, Ungarisch: *Duna*, etc.), andere müssen in ihrer Originalform bestehen bleiben (z. B. *Waldheim*).

2 Geographisches Information Retrieval

Mit den heute verfügbaren Technologien kann die Bedeutung von Texten (in Anwendungen für Massendaten) nicht zu 100% erkannt und repräsentiert werden. Die Systeme begnügen sich daher mit unsicheren Repräsentationen (*bag of words*), welche die Bedeutung nur vage widerspiegeln [Mandl & Womser-Hacker 2006]. Geographische Informationssysteme erlauben dagegen exakte Suchen, da geographische Beziehungen exakt im Sinne des Datenbank-Retrieval repräsentiert werden. Im geographischen Information Retrieval treffen diese beiden scheinbar widersprüchlichen Paradigmen zusammen. Für geographisches Information Retrieval wurden zahlreiche plausible Optimierungsansätze vorgeschlagen, wozu besonders das Hinzuziehen von Wissensressourcen und Schlussfolgerungen zählen. Das Zusammenspiel mehrerer Komponenten in einem System und ihre Effektivität in Bezug auf ein bestimmtes Korpus lassen sich jedoch nie vorhersagen. Somit kommt der Evaluierung der Effektivität auch beim geographischen Retrieval entscheidende Bedeutung bei.

Umfangreiche Evaluierungen und Systemvergleiche erfolgen heute im Rahmen großer Evaluierungsinitiativen [Mandl 2006], welche eine Infrastruktur für die Evaluierung von Systemen mit Massendaten zur Verfügung stellen. Diese Infrastruktur umfasst Dokumenten-Korpora, Aufgaben für die Systeme (*Topics*), intellektuelle Relevanz-Bewertungen sowie die statistische Auswertung. Als Trainingsdaten stehen den teilnehmenden Systemen dabei die Korpora der Vorjahre zur Verfügung, die Relevanzurteile enthalten. Den Besonderheiten des mehrsprachigen Retrieval widmet sich seit dem Jahr 2000 das Cross-Language Evaluation Forum (CLEF², cf. [Peters et al. 2006]). Die CLEF-Organisatoren bieten eine mehrsprachige Infrastruktur für alle beteiligten Sprachen. *Cross Language Information Retrieval* (CLIR) modelliert die Situation von Benutzern, die eine Fremdsprache zwar passiv beherrschen und die Relevanz von Dokumenten in dieser Sprache zumindest abschätzen können, deren aktive Sprachkenntnisse aber nicht für die Konstruktion einer guten Anfrage ausreichen. Mehrsprachiges und insbesondere sprachübergreifendes Information Retrieval erfordern weitere linguistische Komponenten.

² <http://www.clef-campaign.org>

Im Rahmen von CLEF wurde das geographische Retrieval als wichtige Herausforderung erkannt und in die Evaluierungsinitiative aufgenommen. Ein *Pilot Task* 2005 ging in einen regulären Task GeoCLEF 2006 über³. Die Gestaltung der Evaluierungsinfrastruktur fällt bei geographischem Retrieval schwer. Was sind für die bereitstehende Kollektion natürliche Anfragen mit geographischem Bezug? Wie fein sollen die geographischen Einheiten modelliert werden? Sollen explizite geographische Bezüge angegeben werden oder sollen Systeme diese selbst erkennen? In GeoCLEF 2005 gaben explizite geoTags im Markup der Anfrage klare Hinweise, während in GeoCLEF 2006 diese Zusatzhilfe wegfiel, nicht zuletzt weil sie wenig genutzt wurde. Die Verarbeitungsstrategie eines Systems hängt stark davon ab, ob einfache Suchwortübereinstimmungen geprüft oder aufwendige geographische Verarbeitungsschritte eingesetzt werden.

```
<top><num>GC036</num>
<DE-title>Automobilindustrie rund um das Japanische Meer</DE-title>
  <DE-desc>Küstenstädte am Japanischen Meer mit Automobilindustrie oder -werken</DE-desc>
  <DE-narr>Relevante Dokumente berichten von Automobilindustrie oder -werken in Städten an der Küste des Japanischen Meeres (auch Ostmeer (von Korea) genannt), einschließlich wirtschaftlicher oder sozialer Ereignisse wie geplante Joint Ventures oder Streiks. Neben Japan grenzen auch die Länder Nordkorea, Südkorea und Russland an das Japanische Meer.</DE-narr></top>
```

Abbildung 1: Beispiel für ein Topic aus GeoCLEF 2006

GeoCLEF übernahm das in CLEF entwickelte, mehrsprachige, aus Zeitungsartikeln bestehende Korpus. Die Topics werden dabei parallel für unterschiedliche Sprachen entwickelt, indem möglichst realistische Benutzeranfragen modelliert, recherchiert und dann von Muttersprachlern in die jeweilige Sprache übersetzt werden. Im Jahre 2005 bot GeoCLEF 25 Topics in den Ausgangssprachen Deutsch, Englisch, Spanisch und Portugiesisch an. Die Testkollektionen standen in deutsch und englisch zur Verfügung, so dass monolinguale und bilinguale Experimente möglich waren. 2006 waren auch monolinguale Versuche (*runs*) gegen Kollektionen in Spanisch und Portugiesisch möglich, hinzu kamen Topics in Japanisch.

³ <http://ir.shef.ac.uk/geoclef/>

3 Ergebnisse von GeoCLEF 2006

An GeoCLEF nahmen 17 Gruppen teil und lieferten 149 vollständige Experimente für alle 25 Topics ab [Gey et al. 2007]. Nicht zuletzt die Beteiligung von Microsoft, das beweist die hohe Relevanz der Thematik.

Für diese Untersuchung wurden die in Tabelle 1 angegebenen Runs ausgewertet. Als Vergleichswert bietet Tabelle 2 einige Ergebnisse von ad-hoc Runs.

Run Typ	Anzahl Teilnehmer	Anzahl Runs	MAP top System	StAbw. Topics (absolut)	StAbw. Systeme (abs.)
Monolingual DE	4	16	22,29%	16,07%	5,56%
Monolingual EN	16	73	30,34%	15,47%	7,84%
Bilingual X->DE	3	11	16,82%	18,00%	2,98%
Bilingual X->EN	3	12	27,07%	23,98%	4,18%

Tabelle 1: Ausgewertete Experimente GeoCLEF 2006 mit Mean Average Precision (MAP) und Standard-Abweichung (StAbw) der MAP zwischen den Topics und Systemen

Run Typ	CLEF Track	Anzahl Runs	MAP top System	StdAbw. Topics (absolut)	StdAbw. Systeme (absolut)
Bilingual X->EN	CLEF 2002 Ad-hoc	51	49,35%	24,98%	9,11%
Monolingual DE	CLEF 2002 Ad-hoc	21	52,34%	24,44%	11,45%

Tabelle 2: Vergleichsexperimente CLEF 2002 Ad-hoc Retrieval

In beiden Tabellen zeigt sich deutlich, dass die Topics für die höchste Varianz in den Ergebnissen verantwortlich sind. Dem gegenüber ist der Einfluss der Systeme eher gering. Im Vergleich zu den ad-hoc Experimenten sind die durchschnittlichen Precision-Werte (*mean average precision* = MAP) bei geographischem Retrieval weitaus niedriger. Dies kann bedeuten, dass für die Systeme geographisches Retrieval noch schwieriger ist oder dass die Anfragen aus GeoCLEF noch ungleich schwieriger sind als die ad-hoc Topics. Zwar gibt es keine generell akzeptierte Definition von einem schwierigen Topic [Grivolla & Mori 2005], innerhalb dieses Beitrags wird die Schwierigkeit als umgekehrt proportional zu der Systemleistung (MAP) verstanden.

Zwar sind die absoluten Werte in der Standardabweichung zwischen den besten Systemleistungen für die Topics in beiden Tabellen vergleichbar, aber die relative Varianz ist bei den geographischen Anfragen deutlich höher. Innerhalb der geographischen Anfragen bei CLEF liegen also erhebliche Unterschiede vor. Ein genauere Blick auf die Performanz der einzelnen Systeme soll erweisen, wie dieses Phä-

nomen zustande kommt. Die Retrieval-Qualität der besten zehn Systeme für zehn Anfragen verdeutlicht Abbildung 1.

Visuell zeigen sich drei Gruppen von Topics. Für einige Topics erzielen alle Systeme konstant Werte von über 65% MAP. Für eine größere Gruppe von Topics liegen die Werte aller Systeme unter 50% MAP. Bei Betrachtung aller 25 Topics tritt bei etwa fünf bis acht Topics eine starke Abweichung zwischen den Systemen auf. Die verschiedenen Systemleistungen (MAP) für die Topics variieren entweder innerhalb der zweiten Gruppe stark oder die Werte schwanken noch stärker. Für eines der Topics mit sehr unterschiedlichen Ergebnissen liefert selbst das beste System einen Wert von Null, während das zweitbeste System als einziges eine befriedigende Lösung erbringt, die dann aber gleich bei 100% liegt. Der sechs-beste Run in der Abbildung weicht deutlich von den anderen ab und liefert fast für jedes Topic andere Werte als vergleichbare Runs.

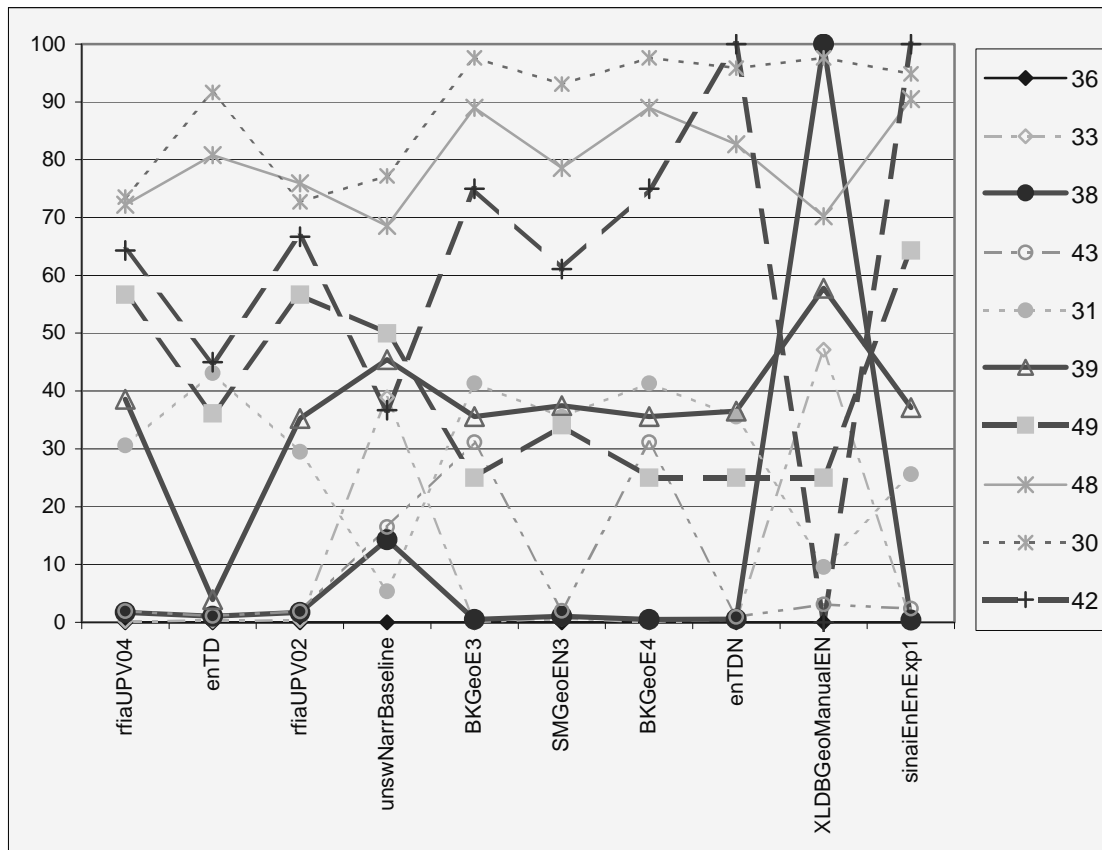


Abbildung 1: Unterschiedliche Leistung der besten zehn Systeme für zehn Topics

Um zu prüfen, welche Faktoren zu dieser hohen Varianz führen, erfolgte eine Analyse der Topics. Dazu wurde intellektuell analysiert, wie viele geographische Eigennamen jedes Topic enthielt und ob es sich um Länder- oder Kontinent-Be-

zeichnungen handelt. Die Anzahl der übrigen inhaltstragenden Begriffe ging ebenfalls in die Analyse mit ein. Für die Anzahl der geographischen Begriffe ergab sich keine klare Tendenz hinsichtlich der Auswirkung auf die Systemleistung. Alle Korrelationen zwischen der Performanz der Systeme für ein Topic und der Anzahl der untersuchten Begriffe im Topic (Länder-Namen, Kontinent-Namen und andere, nicht geographische, inhaltstragende Begriffe) zeigt Tabelle 3. Die Performanz wurde einmal zum besten System für jedes Topic (top) und einmal zum Durchschnitt aller Systeme für ein Topic (ave) berechnet.

	<i>Länder-Namen</i>		<i>Kontinent-Namen</i>		<i>Inhaltstragende Begriffe</i>	
	Top	Ave	Top	Ave	Top	Ave
Monolingual DE	0,23	0,08	-0,22	-0,18	0,17	0,18
Monolingual EN	0,14	0,27	0,05	-0,21	0,29	0,05
Bilingual X-> DE	0,28	0,16	-0,06	-0,09	0,15	0,10
Bilingual X-> EN	0,11	0,06	0,12	-0,11	0,27	0,10

Tabelle 3: Korrelationen zwischen MAP und in den Topics vorkommenden Begriffe für analysierte Experimente aus GeoCLEF 2006 für Deutsch (DE) und Englisch (EN)

Sämtliche Korrelationen sind sehr schwach. Lediglich für inhaltstragende Begriffe sowie für Ländernamen ergeben sich konsistente Tendenzen für alle vier Tracks. Allerdings stehen bisher nur 25 Topics für die Auswertung zur Verfügung, so dass keine weitreichenden Folgerungen erlaubt sind. Es zeigt sich aber, dass die Rolle von Eigennamen im Retrieval vor dem Hintergrund geographischer Eigennamen neu bewertet werden muss. Bisher war bekannt, dass Eigennamen sich positiv auf die Retrieval-Qualität auswirken [Mandl & Womser-Hacker 2005]. Jedoch könnten die Beiträge unterschiedlicher Typen von Eigennamen stark voneinander abweichen.

4 Experimente mit einem geographischen Information Retrieval System

Das System der Informationswissenschaft der Universität Hildesheim hat sich seit mehreren Jahren im Rahmen des Cross Language Evaluation Forum bewährt. Das System baut auf der *open source* Suchmaschine Lucene auf, optimiert dieses durch Mehrwertkomponenten und implementiert die Funktionalität für mehrsprachiges Retrieval [Hackl et al. 2006]. Für die Experimente im Rahmen von geoCLEF wurde es weiter adaptiert.

Ein Systemgestalter eines geographischen Information Retrieval Systems steht derzeit vor zwei Alternativen. Zum einen bietet sich an, traditionelles ad-hoc Retrieval zu betreiben und die geographischen Begriffe als normale Suchterme betrachten. Dies führt nach heutigem Wissensstand oft zu einer guten Leistung. Zum anderen kann ein System geographisches Wissen in starkem Maße formalisieren und darauf Schlussfolgerungen aufbauen.

Der Ansatz der Informationswissenschaft der Universität Hildesheim kombiniert beide Ansätze, wobei jedoch der Schwerpunkt auf Stichwortsuche und Standard-Ansätzen liegt. Im Rahmen eines Standard-Ansatzes werden während der Indexierung geographische Begriffe innerhalb der Dokumente erkannt und in einem gesonderten Index gesammelt. Als klassisches und bewährtes Optimierungsinstrument wurde blindes Relevanz Feedback (BRF) angewandt. Dabei geht das System davon aus, dass die im ersten Suchlauf gefundenen Dokumente relevant sind und extrahiert daraus weitere Suchterme, die zu einer neuen Anfrage führen, welche die endgültigen Ergebnisse liefert. In der Phase des blinden Relevanz Feedbacks nutzt der neue Ansatz nun aber die geographischen Terme sehr spezifisch. Das Relevanz Feedback wirkt sowohl auf den normalen Index als auch auf den geographischen Index, kann jedoch unterschiedlich gewichtet werden (geoBRF). So können durch spezielle Gewichtung von geographischen Termen in den Top-Dokumenten des ersten Durchlaufs mehr geographische Begriffe in die modifizierte Suchanfrage eingehen [Bischoff et al. 2007].

Ein ähnlicher Ansatz zur Optimierung durch flexibles Relevanz Feedback wurde auch erfolgreich in einem System mit kontrolliertem Vokabular eingesetzt, das mehrere Index-Felder umfasste [Berghaus et al. 2006]. Die Experimente basieren auf den Retrievalfunktionen des Lucene-Pakets, dem Lucene-Stemmer für das Deutsche und dem Snowball-Stemmer für das Englische.

Für die Erkennung von Eigennamen wird aufgrund von Evaluierungsergebnissen [Mandl et al. 2005] das maschinelle Lerntool LingPipe eingesetzt, welches anhand eines trainierten Modells Eigennamen identifiziert und in die Kategorien PERSON, LOCATION, ORGANISATION und MISC klassifiziert. Als Modell diente für das Englische das mitgelieferte, an einem englischen Nachrichten-Korpus trainierte News-Modell und für das Deutsche ein selbst entwickeltes Modell.

Die Experimente griffen auf die 25 GeoCLEF Anfragen von 2006 zurück, wobei nur die Felder *title* und *description* genutzt wurden. Eine Testmenge stand nicht zur Verfügung, so dass die gemessenen Werte den Ergebnissen in der Trainingsmenge entsprechen. Tabellen 4 und 5 zeigen die Resultate.

Dokumente für Standard BRF	Terme aus Standard BRF	Dokumente für geoBRF	Terme aus geoBRF	MAP
5	100	5	5	0,2082
5	100	15	15	0,2109
5	100	10	10	0,2085
5	100	10	20	0,2092
30	5	15	3	0,1559
10	5	10	3	0,1507
5	10	5	10	0,1718
5	40	0	0	0,1885
5	100	0	0	0,2077

Tabelle 4: Ergebnisse für mono-lingual Englisch

Dokumente für Standard BRF	Terme aus Standard BRF	Dokumente für geoBRF	Terme aus geoBRF	MAP
5	50	15	15	0,2696
10	50	0	0	0,2652
10	10	10	10	0,2807
0	0	0	0	0,256
5	10	10	10	0,2453
15	10	10	10	0,2841
20	10	10	10	0,2625
30	10	10	10	0,2578
5	50	15	5	0,2696
10	50	15	10	0,2652
10	10	0	0	0,2528

Tabelle 5: Ergebnisse für mono-lingual Deutsch

Für beide Sprachen wirkt sich das flexible Relevanz Feedback auf die Retrieval-Qualität aus, wobei die Varianz für das Deutsche geringer ist. Die besten Ergebnisse konnten mit unterschiedlichen Parametersetzungen für das Relevanz Feedback hinsichtlich der Anzahl der berücksichtigten Dokumente und der einbezogenen Terme sowohl für allgemeine als auch geographische Begriffe erzielt werden.

Nachdem in den hier vorliegenden Experimenten geographische Ressourcen mit traditionellen Information Retrieval-Optimierungsverfahren kombiniert wurden, steht für weitere Systemverbesserungen im Rahmen von GeoCLEF 2007 die Integration weiterer Ressourcen im Vordergrund. Unter anderem werden *Wikipedia* zur Erkennung optimaler Übersetzungen von geographischen Eigennamen und die Dienste *Meinestadt.de* und *OpenGeoDB* für geographische Schlussfolgerungen eingesetzt.

5 Ausblick: GeoCLEF 2007

GeoCLEF wird auch im Rahmen von CLEF 2007 weitergeführt und umfasst nun zwei Aufgaben. Das in diesem beschriebene ad-hoc Retrieval mit geographischen Einschränkungen ist bisher nicht ausreichend erforscht und der Korpus muss weiter entwickelt werden. Die Aufgabentypen bedürfen aus Evaluierungsperspektive noch einer genaueren Analysen. Die Systementwickler werden noch weitere Ressourcen aufbauen und mit deren Eingliederung in den Information Retrieval Prozess experimentieren.

Des Weiteren führt GeoCLEF 2007 erstmals einen Track zur Klassifizierung von Anfragen ein. Microsoft stellt dafür eine Kollektion von realen Anfragen bereit. Die Teilnehmer sollen die geographischen Anfragen und darin die geographischen Entitäten bestimmen. Ebenso sollen teilnehmende Systeme die gewünschte geographische Beziehung identifizieren (*in, südlich von, entlang, in einer bestimmten Entfernung*). Die Entfernung von Stoppwörtern als Standard-Verarbeitung würde bereits zahlreiche Hinweise auf solche Relationen löschen.

Für weitere Tracks werden Möglichkeiten zum geographischen Retrieval von Fotografien analysiert. Dort liegt der geographische Bezug auf der Hand. Bereits im imageCLEF Track vorhandene Kollektionen und Tauschbörsen wie *Flickr* könnten die dafür nötigen Daten liefern. Inhaltsanalyse und Retrieval auf der Basis der Bildunterschriften bzw. Tags sind dann denkbar.

6 Literaturverzeichnis

- [Berghaus et al. 2006] Berghaus, Benjamin; Kluck, Michael; Mandl, Thomas (2006): Dynamic Entry Vocabulary Module Implementation and Evaluation. In: Schaaf, Martin; Althoff, Klaus-Dieter (Hrsg): LWA 2006 – Lernen – Wissensentdeckung – Adaptivität Proceedings Workshopwoche GI. 9.-11. Oktober 2006, Universität Hildesheim. Workshop Information Retrieval 2006 [Hildesheimer Informatik-Berichte] S. 94-101.
- [Bischoff et al. 2007] Bischoff, Kerstin; Mandl, Thomas; Womser-Hacker, Christa (2007): Blind Relevance Feedback and Named Entity based Query Expansion for Geographic Retrieval at GeoCLEF 2006. In: Peters, Carol et al. (Eds.): 7th Workshop of the Cross-Language Evaluation Forum, CLEF 2005, Alicante, Spain, Revised Selected Papers. Berlin et al.: Springer [Lecture Notes in Computer Science] erscheint. Vorab in: Nardi, Alessandro; Peters, Carol; Vicedo, José Luis (Eds.): CLEF 2006 Working Notes. http://www.clef-campaign.org/2006/working_notes/workingnotes2006/bischoffCLEF2006.pdf
- [Broder 2002] Broder, Andrei (2002): A taxonomy of web search. In: ACM SIGIR Forum vol. 36(2) S. 3 – 10.

- [Buckley & Voorhees 2005] Buckley, Chris; Voorhees, Ellen (2005). Retrieval System Evaluation. In: TREC: Experiment and Evaluation in Information Retrieval. Cambridge & London: MIT Press. S. 53-75.
- [Gey et al. 2007] Gey, Fredric; Larson, Ray; Sanderson, Mark; Bishoff, Kerstin; Mandl, Thomas; Womser-Hacker, Christa; Santos, Diana; Rocha, Paulo; Di Nunzio, Giorgio; Ferro, Nicola (2007): GeoCLEF 2006: the CLEF 2006 Cross-Language Geographic Information Retrieval Track Overview. In: Peters, Carol et al. (Eds.): 7th Workshop of the Cross-Language Evaluation Forum, CLEF 2005, Alicante, Spain, Revised Selected Papers. Berlin et al.: Springer [Lecture Notes in Computer Science] erscheint. Vorab in: Nardi, Alessandro; Peters, Carol; Vicedo, José Luis (Eds.): CLEF 2006 Working Notes. http://www.clef-campaign.org/2006/working_notes/workingnotes2006/geyOCLEF2006.pdf
- [Grivolla & Mori 2005] Grivolla, J.; Jourlin, P. and de Mori, R. (2005): Automatic Classification of Queries by Expected Retrieval Performance. In: Predicting Query Difficulty – Methods and Applications. Workshop at the 28th Annual International ACM Conference on Research and Development in Information Retrieval (SIGIR) Salvador, Bahia, Brazil, August 19, 2005. <http://www.haifa.il.ibm.com/sigiro5-qp/>
- [Hackl et al. 2006] Hackl, René; Mandl, Thomas; Womser-Hacker, Christa (2006): Ad-hoc Mono- and Multilingual Retrieval Experiments at the University of Hildesheim. In: Peters, Carol; Gey, Fredric C.; Gonzalo, Julio; Jones, Gareth J.F.; Kluck, Michael; Magnini, Bernardo; Müller, Henning; Rijke, Maarten de (Eds.). Accessing Multilingual Information Repositories: 6th Workshop of the Cross-Language Evaluation Forum, CLEF 2005, Vienna, Austria, Revised Selected Papers. Berlin et al.: Springer [Lecture Notes in Computer Science 4022] S. 37-43.
- [Mandl 2006] Mandl, Thomas (2006): Neue Entwicklungen bei den Evaluierungsinitiativen im Information Retrieval. In: Mandl, Thomas; Womser-Hacker, Christa (Hrsg.): Effektive Information Retrieval Verfahren in der Praxis: Proceedings Vierter Hildesheimer Evaluierungs- und Retrievalworkshop (HIER 2005) Hildesheim, 20.7.2005. Konstanz: Universitätsverlag [Schriften zur Informationswissenschaft 45] S. 117-128.
- [Mandl & Womser-Hacker 2005] Mandl, Thomas; Womser-Hacker, Christa (2005): The Effect of Named Entities on Effectiveness in Cross-Language Information Retrieval Evaluation. In: Applied Computing 2005: Proceedings of 2005 ACM SAC Symposium on Applied Computing (SAC). Information Access and Retrieval (IAR) Track. Santa Fe, New Mexico, USA. March 13.-17. 2005. ACM Press. S. 1059-1064.
- [Mandl et al. 2005] Mandl, Thomas; Schneider, René; Schnetzler, Pia; Womser-Hacker, Christa (2005): Evaluierung von Systemen für die Eigennamenerkennung im cross-lingualen Information Retrieval. In: Fisseni, Bernhard; Schmitz, Hans-Christian; Schröder, Bernhard; Wagner, Petra (Hrsg.): Sprachtechnologie, mobile Kommunikation und linguistische Ressourcen: Beiträge der zur GLDV Tagung 2005 in Bonn (Gesellschaft für linguistische Datenverarbeitung) 30.3.-01.04. [Sprache, Sprechen und Computer/Computer Studies in Language and Speech 8] Frankfurt a. M. et al.: Peter-Lang. S. 145-157.

- [Mandl & Womser-Hacker 2006] Mandl, Thomas; Womser-Hacker, Christa (2006):
Information Retrieval – eine Schlüsseltechnologie in der Wissensgesellschaft. In: WISU:
Das Wirtschaftsstudium. Erscheint.
- [Peters et al. 2006] Peters, Carol; Gey, Fredric; Gonzalo, Julio; Jones, Gareth J.F.; Kluck,
Michael; Magnini, Bernardo; Müller, Henning; Rijke, Maarten de (2006) (Eds.).
Accessing Multilingual Information Repositories: 6th Workshop of the Cross-Language
Evaluation Forum, CLEF 2005, Vienna, Austria, Revised Selected Papers. Berlin et al.:
Springer [Lecture Notes in Computer Science 4022]