

University of Groningen at GeoCLEF 2007

G. Andogah G. Bouma

Computational Linguistics Group
Center for Language and Cognition Groningen (CLCG)
University of Groningen
Groningen, The Netherlands

CLEF 2007 Workshop, 19-23 September, Budapest, Hungary

Outline

- 1 Introduction
 - Motivation
 - Resources

- 2 Indexing and searching
 - Indexing
 - Similarity measures
 - Searching

- 3 Evaluation
 - Result
 - Discussion

Outline

- 1 Introduction
 - Motivation
 - Resources
- 2 Indexing and searching
 - Indexing
 - Similarity measures
 - Searching
- 3 Evaluation
 - Result
 - Discussion

Outline

- 1 Introduction
 - Motivation
 - Resources
- 2 Indexing and searching
 - Indexing
 - Similarity measures
 - Searching
- 3 Evaluation
 - Result
 - Discussion

The Motivation

- Test geographic scope (geo-scope) based relevance ranking for geographic information retrieval (GIR).
- We participated in monolingual English task.
- Our evaluation results show no significant improvement for geo-scope based relevance ranking approach.

The Motivation

- Test geographic scope (geo-scope) based relevance ranking for geographic information retrieval (GIR).
- We participated in monolingual English task.
- Our evaluation results show no significant improvement for geo-scope based relevance ranking approach.

The Motivation

- Test geographic scope (geo-scope) based relevance ranking for geographic information retrieval (GIR).
- We participated in monolingual English task.
- Our evaluation results show no significant improvement for geo-scope based relevance ranking approach.

Resources Used

- GKB – World Gazetteer, Geonames, Wikipedia and Wordnet
- GeoTagger – Alias-i Lingpipe used to detect place names, geo-concepts, geo-relations, geo-adjectives
- Indexing/search engine – Apache Lucene

Resources Used

- GKB – World Gazetteer, Geonames, Wikipedia and Wordnet
- GeoTagger – Alias-i Lingpipe used to detect place names, geo-concepts, geo-relations, geo-adjectives
- Indexing/search engine – Apache Lucene

Resources Used

- GKB – World Gazetteer, Geonames, Wikipedia and Wordnet
- GeoTagger – Alias-i Lingpipe used to detect place names, geo-concepts, geo-relations, geo-adjectives
- Indexing/search engine – Apache Lucene

Document Indexing

- Documents were processed with the Porter stemmer and the default Lucene English stopwords list.
- Index structure shown in Table below.

Table: Reference document index structure

Field	Lucene Type	Description
docid	Field.Keyword	Document unique identification/ number
content	Field.Unstored	Combination of HEADLINE and TEXT tag content

Document Indexing

- Documents were processed with the Porter stemmer and the default Lucene English stopwords list.
- Index structure shown in Table below.

Table: Reference document index structure

Field	Lucene Type	Description
docid	Field.Keyword	Document unique identification/ number
content	Field.Unstored	Combination of HEADLINE and TEXT tag content

Non-Geographic Similarity Measure

The Lucene similarity score formula:

$$\text{NonSim}(q, d) = \sum_{t \text{ in } q} \text{tf}(t \text{ in } d) \cdot \text{idf}(t) \cdot \text{bst} \cdot \text{IN}(t.\text{field in } d) \quad (1)$$

where,

- $\text{tf}(t \text{ in } d)$ is the term frequency factor for term t in document d ,
- $\text{idf}(t)$ is the inverse document frequency of term t ,
- bst is the field boost set during indexing,
- $\text{IN}(t.\text{field in } d)$ is the normalization value of a field given the number of terms in the field.

Non-Geographic Similarity Measure

The Lucene similarity score formula:

$$\text{NonSim}(q, d) = \sum_{t \text{ in } q} \text{tf}(t \text{ in } d) \cdot \text{idf}(t) \cdot \text{bst} \cdot \text{IN}(t.\text{field in } d) \quad (1)$$

where,

- $\text{tf}(t \text{ in } d)$ is the term frequency factor for term t in document d ,
- $\text{idf}(t)$ is the inverse document frequency of term t ,
- bst is the field boost set during indexing,
- $\text{IN}(t.\text{field in } d)$ is the normalization value of a field given the number of terms in the field.

Non-Geographic Similarity Measure

The Lucene similarity score formula:

$$NonSim(q, d) = \sum_{t \text{ in } q} tf(t \text{ in } d) \cdot idf(t) \cdot bst \cdot IN(t.field \text{ in } d) \quad (1)$$

where,

- $tf(t \text{ in } d)$ is the term frequency factor for term t in document d ,
- $idf(t)$ is the inverse document frequency of term t ,
- bst is the field boost set during indexing,
- $IN(t.field \text{ in } d)$ is the normalization value of a field given the number of terms in the field.

Non-Geographic Similarity Measure

The Lucene similarity score formula:

$$\text{NonSim}(q, d) = \sum_{t \text{ in } q} \text{tf}(t \text{ in } d) \cdot \text{idf}(t) \cdot \text{bst} \cdot \text{IN}(t.\text{field in } d) \quad (1)$$

where,

- $\text{tf}(t \text{ in } d)$ is the term frequency factor for term t in document d ,
- $\text{idf}(t)$ is the inverse document frequency of term t ,
- bst is the field boost set during indexing,
- $\text{IN}(t.\text{field in } d)$ is the normalization value of a field given the number of terms in the field.

Geographic Similarity Measure

Equation 2 depicts our geographic similarity measure formula between query q and document d :

$$GeoSim(q, d) = \begin{cases} SF \times WTS & \text{if } SF > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where;

$$SF = \begin{cases} \sqrt{\frac{N_{(d,q)}}{N_d + N_q + |N_d - N_q|}} & \text{if } N_{(d,q)} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

$$WTS = \sum \sqrt{wt_{(q,s)}} \times \log(1 + wt_{(d,s)}) \quad (4)$$

where;

- N_q and N_d – number of scopes in the query and document scope sets respectively.
- $N_{(d,q)}$ – number of document scopes present in query scope set.
- $wt_{(q,s)}$ – weight assigned to scope s in query q by the scope resolver.
- $wt_{(d,s)}$ – weight assigned to scope s in document d by the scope resolver.
- For a given query, N_q is invariable whilst N_d and $N_{(d,q)}$ vary per document retrieved.

where;

- N_q and N_d – number of scopes in the query and document scope sets respectively.
- $N_{(d,q)}$ – number of document scopes present in query scope set.
- $wt_{(q,s)}$ – weight assigned to scope s in query q by the scope resolver.
- $wt_{(d,s)}$ – weight assigned to scope s in document d by the scope resolver.
- For a given query, N_q is invariable whilst N_d and $N_{(d,q)}$ vary per document retrieved.

where;

- N_q and N_d – number of scopes in the query and document scope sets respectively.
- $N_{(d,q)}$ – number of document scopes present in query scope set.
- $wt_{(q,s)}$ – weight assigned to scope s in query q by the scope resolver.
- $wt_{(d,s)}$ – weight assigned to scope s in document d by the scope resolver.
- For a given query, N_q is invariable whilst N_d and $N_{(d,q)}$ vary per document retrieved.

where;

- N_q and N_d – number of scopes in the query and document scope sets respectively.
- $N_{(d,q)}$ – number of document scopes present in query scope set.
- $wt_{(q,s)}$ – weight assigned to scope s in query q by the scope resolver.
- $wt_{(d,s)}$ – weight assigned to scope s in document d by the scope resolver.
- For a given query, N_q is invariable whilst N_d and $N_{(d,q)}$ vary per document retrieved.

where;

- N_q and N_d – number of scopes in the query and document scope sets respectively.
- $N_{(d,q)}$ – number of document scopes present in query scope set.
- $wt_{(q,s)}$ – weight assigned to scope s in query q by the scope resolver.
- $wt_{(d,s)}$ – weight assigned to scope s in document d by the scope resolver.
- For a given query, N_q is invariable whilst N_d and $N_{(d,q)}$ vary per document retrieved.

Geo-IR Similarity Measure

$$Sim(q, d) = \lambda_T NonSim(q, d) + \lambda_G GeoSim(q, d) \quad (5)$$

$$\lambda_T + \lambda_G = 1 \quad (6)$$

where;

- λ_T and λ_G are the non-geographic and geographic interpolation factor respectively.
- Non-geographic and geographic relevance scores are normalized to $[0, 1]$.

Geo-IR Similarity Measure

$$Sim(q, d) = \lambda_T NonSim(q, d) + \lambda_G GeoSim(q, d) \quad (5)$$

$$\lambda_T + \lambda_G = 1 \quad (6)$$

where;

- λ_T and λ_G are the non-geographic and geographic interpolation factor respectively.
- Non-geographic and geographic relevance scores are normalized to $[0, 1]$.

Topic categorization

- In [1], geographic topics are categorized into eight according to the way they depend on a place, geographic subject or geographic relation.
- We grouped GeoCLEF 2007 topics into two: (1) topics whose geographic scopes can easily be resolved to a place (GROUP1), and (2) topics whose geographic scopes cannot be resolved to a place (GROUP2).

Topic categorization

- In [1], geographic topics are categorized into eight according to the way they depend on a place, geographic subject or geographic relation.
- We grouped GeoCLEF 2007 topics into two: (1) topics whose geographic scopes can easily be resolved to a place (GROUP1), and (2) topics whose geographic scopes cannot be resolved to a place (GROUP2).

GROUP1 Topics

- We performed geographic query expansion (GQE) on the following members of GROUP1 – 51, 59, 60, 61, 63, 65, 66, 70.
- Motivation for GQE is the lack of sufficient geographic information or ambiguous geographic information provided.
- For example, topic 10.2452/59-GC: "Meetings of the Andean Community of Nations (CAN)" is expanded by adding the names of major cities in Bolivia, Columbia, Ecuador and Peru.
- GQE terms are assigned Lucene boost of 0.45F and original GQ terms left at default boost 1.0F.

GROUP1 Topics

- We performed geographic query expansion (GQE) on the following members of GROUP1 – 51, 59, 60, 61, 63, 65, 66, 70.
- Motivation for GQE is the lack of sufficient geographic information or ambiguous geographic information provided.
- For example, topic 10.2452/59-GC: "Meetings of the Andean Community of Nations (CAN)" is expanded by adding the names of major cities in Bolivia, Columbia, Ecuador and Peru.
- GQE terms are assigned Lucene boost of 0.45F and original GQ terms left at default boost 1.0F.

GROUP1 Topics

- We performed geographic query expansion (GQE) on the following members of GROUP1 – 51, 59, 60, 61, 63, 65, 66, 70.
- Motivation for GQE is the lack of sufficient geographic information or ambiguous geographic information provided.
- For example, topic 10.2452/59-GC: "Meetings of the Andean Community of Nations (CAN)" is expanded by adding the names of major cities in Bolivia, Columbia, Ecuador and Peru.
- GQE terms are assigned Lucene boost of 0.45F and original GQ terms left at default boost 1.0F.

GROUP1 Topics

- We performed geographic query expansion (GQE) on the following members of GROUP1 – 51, 59, 60, 61, 63, 65, 66, 70.
- Motivation for GQE is the lack of sufficient geographic information or ambiguous geographic information provided.
- For example, topic 10.2452/59-GC: "Meetings of the Andean Community of Nations (CAN)" is expanded by adding the names of major cities in Bolivia, Columbia, Ecuador and Peru.
- GQE terms are assigned Lucene boost of 0.45F and original GQ terms left at default boost 1.0F.

GROUP2 Topics

- Members of GROUP1 – 56, 67, 68, 72.
- These topics fall under *geographic subject with non-geographic restriction* with exception of topic 67 which is more complex.
 - Topic 10.2452/56-GC: "Lakes with monsters".
 - Topic 10.2452/67-GC: "F1 circuits where Ayrton Senna competed in 1994" is a more complex topic aggregating GNE (F1 circuits), PNE (Ayrton Senna) and DNE (1994).
- These topics are best resolved to geographic subject scopes, e.g. city, lake, beach, river, etc.

GROUP2 Topics

- Members of GROUP1 – 56, 67, 68, 72.
- These topics fall under *geographic subject with non-geographic restriction* with exception of topic 67 which is more complex.
 - Topic 10.2452/56-GC: "Lakes with monsters".
 - Topic 10.2452/67-GC: "F1 circuits where Ayrton Senna competed in 1994" is a more complex topic aggregating GNE (F1 circuits), PNE (Ayrton Senna) and DNE (1994).
- These topics are best resolved to geographic subject scopes, e.g. city, lake, beach, river, etc.

GROUP2 Topics

- Members of GROUP1 – 56, 67, 68, 72.
- These topics fall under *geographic subject with non-geographic restriction* with exception of topic 67 which is more complex.
 - Topic 10.2452/56-GC: "Lakes with monsters".
 - Topic 10.2452/67-GC: "F1 circuits where Ayrton Senna competed in 1994" is a more complex topic aggregating GNE (F1 circuits), PNE (Ayrton Senna) and DNE (1994).
- These topics are best resolved to geographic subject scopes, e.g. city, lake, beach, river, etc.

GROUP2 Topics

- Members of GROUP1 – 56, 67, 68, 72.
- These topics fall under *geographic subject with non-geographic restriction* with exception of topic 67 which is more complex.
 - Topic 10.2452/56-GC: "Lakes with monsters".
 - Topic 10.2452/67-GC: "F1 circuits where Ayrton Senna competed in 1994" is a more complex topic aggregating GNE (F1 circuits), PNE (Ayrton Senna) and DNE (1994).
- These topics are best resolved to geographic subject scopes, e.g. city, lake, beach, river, etc.

GROUP2 Topics

- Members of GROUP1 – 56, 67, 68, 72.
- These topics fall under *geographic subject with non-geographic restriction* with exception of topic 67 which is more complex.
 - Topic 10.2452/56-GC: "Lakes with monsters".
 - Topic 10.2452/67-GC: "F1 circuits where Ayrton Senna competed in 1994" is a more complex topic aggregating GNE (F1 circuits), PNE (Ayrton Senna) and DNE (1994).
- These topics are best resolved to geographic subject scopes, e.g. city, lake, beach, river, etc.

CLCGGeoEE{T00, TD00, TDN00, TDN00P}

- GeoCLEF 2007: we participated in English monolingual task submitting 5 Runs.
- T00, TD00, TDN00 formulated by content of topic TITLE (T), TITLE-DESC (TD) and TITLE-DESC-NARR (TDN) respectively.
- Both GROUP1 and GROUP2 topics are ranked according to Eq 1 (i.e. default Lucene similarity rank).
- TDN00 erroneously used topic TITLE content. TDN00P corrected version using topic TDN contents.

CLCGGeoEE{T00, TD00, TDN00, TDN00P}

- GeoCLEF 2007: we participated in English monolingual task submitting 5 Runs.
- T00, TD00, TDN00 formulated by content of topic TITLE (T), TITLE-DESC (TD) and TITLE-DESC-NARR (TDN) respectively.
- Both GROUP1 and GROUP2 topics are ranked according to Eq 1 (i.e. default Lucene similarity rank).
- TDN00 erroneously used topic TITLE content. TDN00P corrected version using topic TDN contents.

CLCGGeoEE{T00, TD00, TDN00, TDN00P}

- GeoCLEF 2007: we participated in English monolingual task submitting 5 Runs.
- T00, TD00, TDN00 formulated by content of topic TITLE (T), TITLE-DESC (TD) and TITLE-DESC-NARR (TDN) respectively.
- Both GROUP1 and GROUP2 topics are ranked according to Eq 1 (i.e. default Lucene similarity rank).
- TDN00 erroneously used topic TITLE content. TDN00P corrected version using topic TDN contents.

CLCGGeoEE{T00, TD00, TDN00, TDN00P}

- GeoCLEF 2007: we participated in English monolingual task submitting 5 Runs.
- T00, TD00, TDN00 formulated by content of topic TITLE (T), TITLE-DESC (TD) and TITLE-DESC-NARR (TDN) respectively.
- Both GROUP1 and GROUP2 topics are ranked according to Eq 1 (i.e. default Lucene similarity rank).
- TDN00 erroneously used topic TITLE content. TDN00P corrected version using topic TDN contents.

CLCGGeoEE{TDN01, TDN01P}

- TDN01 erroneously used topic TITLE content. TDN01P corrected version using topic TDN contents.
- GROUP1 topics ranked according to Eq 5 with $\lambda_T = 0.85$ and $\lambda_G = 0.15$. And GROUP2 topics ranked according to Eq 1.
- λ_T and λ_G values decided through experiment on GeoCLEF 2006 Monolingual English task.

CLCGGeoEE{TDN01, TDN01P}

- TDN01 erroneously used topic TITLE content. TDN01P corrected version using topic TDN contents.
- GROUP1 topics ranked according to Eq 5 with $\lambda_T = 0.85$ and $\lambda_G = 0.15$. And GROUP2 topics ranked according to Eq 1.
- λ_T and λ_G values decided through experiment on GeoCLEF 2006 Monolingual English task.

CLCGGeoEE{TDN01, TDN01P}

- TDN01 erroneously used topic TITLE content. TDN01P corrected version using topic TDN contents.
- GROUP1 topics ranked according to Eq 5 with $\lambda_T = 0.85$ and $\lambda_G = 0.15$. And GROUP2 topics ranked according to Eq 1.
- λ_T and λ_G values decided through experiment on GeoCLEF 2006 Monolingual English task.

CLCGGeoEE{TDN01B, TDN01BP}

- TDN01B erroneously used topic TITLE content. TDN01BP corrected version using topic TDN contents.
- GROUP1 topics ranked according to Eq 5 with $\lambda_T = 0.85$ and $\lambda_G = 0.15$.
- For GROUP2 topics,
 - extract geo-types of each returned document.
 - assign a weight of $[0,1]$ to geo-types according to frequency of occurrence.
 - scores of documents containing query geo-types are argumented with geo-type weights and re-ranked.

CLCGGeoEE{TDN01B, TDN01BP}

- TDN01B erroneously used topic TITLE content. TDN01BP corrected version using topic TDN contents.
- GROUP1 topics ranked according to Eq 5 with $\lambda_T = 0.85$ and $\lambda_G = 0.15$.
- For GROUP2 topics,
 - extract geo-types of each returned document.
 - assign a weight of $[0,1]$ to geo-types according to frequency of occurrence.
 - scores of documents containing query geo-types are argumented with geo-type weights and re-ranked.

CLCGGeoEE{TDN01B, TDN01BP}

- TDN01B erroneously used topic TITLE content. TDN01BP corrected version using topic TDN contents.
- GROUP1 topics ranked according to Eq 5 with $\lambda_T = 0.85$ and $\lambda_G = 0.15$.
- For GROUP2 topics,
 - extract geo-types of each returned document.
 - assign a weight of [0,1] to geo-types according to frequency of occurrence.
 - scores of documents containing query geo-types are argued with geo-type weights and re-ranked.

CLCGGeoEE{TDN01B, TDN01BP}

- TDN01B erroneously used topic TITLE content. TDN01BP corrected version using topic TDN contents.
- GROUP1 topics ranked according to Eq 5 with $\lambda_T = 0.85$ and $\lambda_G = 0.15$.
- For GROUP2 topics,
 - extract geo-types of each returned document.
 - assign a weight of $[0,1]$ to geo-types according to frequency of occurrence.
 - scores of documents containing query geo-types are argued with geo-type weights and re-ranked.

CLCGGeoEE{TDN01B, TDN01BP}

- TDN01B erroneously used topic TITLE content. TDN01BP corrected version using topic TDN contents.
- GROUP1 topics ranked according to Eq 5 with $\lambda_T = 0.85$ and $\lambda_G = 0.15$.
- For GROUP2 topics,
 - extract geo-types of each returned document.
 - assign a weight of $[0,1]$ to geo-types according to frequency of occurrence.
 - scores of documents containing query geo-types are argued with geo-type weights and re-ranked.

CLCGGeoEE{TDN01B, TDN01BP}

- TDN01B erroneously used topic TITLE content. TDN01BP corrected version using topic TDN contents.
- GROUP1 topics ranked according to Eq 5 with $\lambda_T = 0.85$ and $\lambda_G = 0.15$.
- For GROUP2 topics,
 - extract geo-types of each returned document.
 - assign a weight of $[0,1]$ to geo-types according to frequency of occurrence.
 - scores of documents containing query geo-types are argued with geo-type weights and re-ranked.

TDN01B: GROUP2 Topic formula

GROUP2 topics are processed according to:

$$SimScore = NonSim(q, d) + \frac{1.0}{\sqrt{1 + \frac{N_{qFClass} - N_{qFType}}{N_{qFClass}}}} \quad (7)$$

- For example, let the document $NonSim(q, d) = 0.453$, feature types (FType) recognized in document are PPL [8 times] and PPLX [2 times] and required feature class is P [8 + 2 = 10].
- And therefore, document $SimScore = 1.36587$ and this is performed for all documents. The scores are normalized to [0,1] and re-ranked in descending order.

TDN01B: GROUP2 Topic formula

GROUP2 topics are processed according to:

$$SimScore = NonSim(q, d) + \frac{1.0}{\sqrt{1 + \frac{N_{qFClass} - N_{qFType}}{N_{qFClass}}}} \quad (7)$$

- For example, let the document $NonSim(q, d) = 0.453$, feature types (FType) recognized in document are PPL [8 times] and PPLX [2 times] and required feature class is P [8 + 2 = 10].
- And therefore, document $SimScore = 1.36587$ and this is performed for all documents. The scores are normalized to [0,1] and re-ranked in descending order.

Official & Post Results

Table: Mean Average Precision & R-Precision

Run	MAP	R-Precision	Dif.Off.Best	Catalunya
T00	0.2023	0.2186	-19.6%	
TD00	0.2515	0.2595	0.0%	-11.8% (MAP: 0.2850)
TDN00	0.2023	0.2186	-19.6%	
TDN01	0.2053	0.2234	-18.4%	
TDN01B	0.1847	0.2019	-26.6%	
TDN00P	0.2647	0.2743	+5.25%	-7.1%
TDN01P	0.2681	0.2878	+6.6%	-5.9%
TDN01BP	0.2442	0.2579	-2.9%	

General Observation

- Geographic information may be useful in improving the performance of an IR system in answering geography restricted user information needs (e.g. compare TD00 & TDN00P).
- Scope-based relevance ranking as implemented shows no significant improvement (e.g. compare TDN00P & TDN01P)

General Observation

- Geographic information may be useful in improving the performance of an IR system in answering geography restricted user information needs (e.g. compare TD00 & TDN00P).
- Scope-based relevance ranking as implemented shows no significant improvement (e.g. compare TDN00P & TDN01P)

Future Work

Future work will focus on investigating:

- better geographic similarity measure formulae
- use of geographic scopes selected by the searcher from the returned list of documents for relevance feedback
- term-relevance feedback based on geographic terms (e.g. placenames) extracted by the searcher after examining the list of retrieved documents
- techniques to answer GROUP2 type topics i.e. geographic subject with non-geographic restriction.

Future Work

Future work will focus on investigating:

- better geographic similarity measure formulae
- use of geographic scopes selected by the searcher from the returned list of documents for relevance feedback
- term-relevance feedback based on geographic terms (e.g. placenames) extracted by the searcher after examining the list of retrieved documents
- techniques to answer GROUP2 type topics i.e. geographic subject with non-geographic restriction.

Future Work

Future work will focus on investigating:

- better geographic similarity measure formulae
- use of geographic scopes selected by the searcher from the returned list of documents for relevance feedback
- term-relevance feedback based on geographic terms (e.g. placenames) extracted by the searcher after examining the list of retrieved documents
- techniques to answer GROUP2 type topics i.e. geographic subject with non-geographic restriction.

Future Work

Future work will focus on investigating:

- better geographic similarity measure formulae
- use of geographic scopes selected by the searcher from the returned list of documents for relevance feedback
- term-relevance feedback based on geographic terms (e.g. placenames) extracted by the searcher after examining the list of retrieved documents
- techniques to answer GROUP2 type topics i.e. geographic subject with non-geographic restriction.

Summary

- Described non-geographic similarity, geographic similarity and combined similarity measures employed for GeoCLEF 2007 participation.
- Tested geographic scope based relevance ranking for GIR.
- No significant improvement with geo-scope based approach.

Summary

- Described non-geographic similarity, geographic similarity and combined similarity measures employed for GeoCLEF 2007 participation.
- Tested geographic scope based relevance ranking for GIR.
- No significant improvement with geo-scope based approach.

Summary

- Described non-geographic similarity, geographic similarity and combined similarity measures employed for GeoCLEF 2007 participation.
- Tested geographic scope based relevance ranking for GIR.
- No significant improvement with geo-scope based approach.



F. Gey, R. Larson, M. Sanderson, K. Bischoff, T. Mandl, and C. Womser-Hacker.

GeoCLEF 2006: the CLEF 2006 Cross-Language Geographic Information Retrieval Task Overview.

[Working Notes for the CLEF 2006 Workshop, 20-22 September, Alicante, Spain.](#)