

GeoCLEF: the CLEF 2005 cross-language geographic information retrieval track

Fredric Gey¹, Ray Larson¹, Mark Sanderson², Hideo Joho², Paul Clough² and Vivien Petras¹

¹University of California, Berkeley, CA, USA

²Department of Information Studies, University of Sheffield, Sheffield, UK

Overview

- GeoCLEF run as a pilot track in 2005
- Aim: to evaluate retrieval of multilingual documents with an emphasis on geographic search (GIR)
 - “find me news stories about riots near Dublin”
- Many documents contains geo-references expressed in multiple languages
- Standard IR systems (and evaluations) pay little attention to spatial aspects of queries and documents

GeoCLEF

- Document languages: English and German
- Topic languages: English (En), German (De), Portuguese and Spanish
- Two ad-hoc retrieval tasks
 - Monolingual (En \rightarrow En; De \rightarrow De)
 - Bilingual (X \rightarrow English; X \rightarrow De)
- 12 research groups submitted 117 runs

Document collections

- Existing English and German CLEF collections
 - variety of topics and geographical regions from news stories between 1994/95
- Documents in English (169,477)
 - Glasgow Herald (British) 1995
 - LA Times (American) 1994
- Documents in German (294,809)
 - Der Spiegel 1994/95
 - Frankfurter Rundschau 1994/95
 - Swiss news agency SDA 1994/95

Search tasks

- 25 topics created for GeoCLEF
 - 10 extended from past CLEF topics
 - 7 created based on articles of LA Times
 - 8 created based on Glasgow Herald
- Varied geographic granularity
- GeoCLEF topic structure
 - <concept><spatial relation><location>

Example topic (1)

```
<top>
<num> GC001 </num>
<orignum> C084 </orignum>
<EN-title>Shark Attacks off Australia and California</EN-title>
<EN-desc> Documents will report any information relating to shark
attacks on humans. </EN-desc>
<EN-narr> Identify instances where a human was attacked by a
shark, including where the attack took place and the circumstances
surrounding the attack. Only documents concerning specific attacks
are relevant; unconfirmed shark attacks or suspected bites are not
relevant. </EN-narr>
<!-- NOTE: This topic has added tags for GeoCLEF -->
<EN-concept> Shark attacks </EN-concept>
<EN-spatialrelation>near</EN-spatialrelation>
<EN-location> Australia </EN-location>
<EN-location> California </EN-location>
</top>
```

Example topic (2)

```
<top>  
<num>GC020</num>  
<EN-title>Wind power in the Scottish Islands</EN-title>  
<EN-desc>Find documents on electrical power generation using  
wind power in the islands of Scotland.</EN-desc>  
<EN-narr>A relevant document will describe wind power-based  
electricity generation schemes providing electricity for the islands of  
Scotland.</EN-narr>  
<EN-concept>Wind power</EN-concept>  
<EN-spatialrelation>in</EN-spatialrelation>  
<EN-location>the Scottish Islands</EN-location>  
</top>
```

Relevance assessments

- Shared between sites
 - Sheffield (topics 18-25 English)
 - Berkeley (topics 1-17 English; 1-25 German)
- Berkeley used conventional pool approach
 - Top 60 documents used from all submissions
- Sheffield used ISJ
 - 17 students and the authors performed searches
- Qrels created from overlap between sites

ISJ system used in GeoCLEF

Interactive Searching and Judging - Microsoft Internet Explorer

File Edit View » Address <http://spirit.shef.ac.uk/isj/isj.cgi?q=Walking+Scotland&username=mark&collect>

Username: **mark** | Topic: **GC018** ([Topic description](#)) | [Help](#) | [Back to Topic List](#) | [Logout](#)

Tip: Try as many different queries as you can.

New! [National Farm Walk Week](#)
ID: GH950422-000103
NATIONAL Farm Walk Week will be run from June 2 -- 11 when farmers will be asked to open their gates and show local communities in Scotland how they are managing their land. Scottish development officer for the Farming and Wildlife Advisory ...
Relevance: A whole or part of document is Relevant Not relevant Not judged

New! [Company in the country](#)
ID: GH950504-000017
Rennie McOwan finds he need never walk alone THE rambler in a formal sense is a fairly new type of outdoor person in Scotland and is on the increase as the Ramblers Association grows in size and influence. ...
Relevance: A whole or part of document is Relevant Not relevant Not judged

New! [Walking lab](#)
ID: GH951109-000195
THE Anderson Gait-Analysis Laboratory, which studies walking, was opened by former Scotland rugby captain Gavin Hastings at Princess Margaret Rose Hospital, Edinburgh, yesterday. Most

[View my judged documents](#)

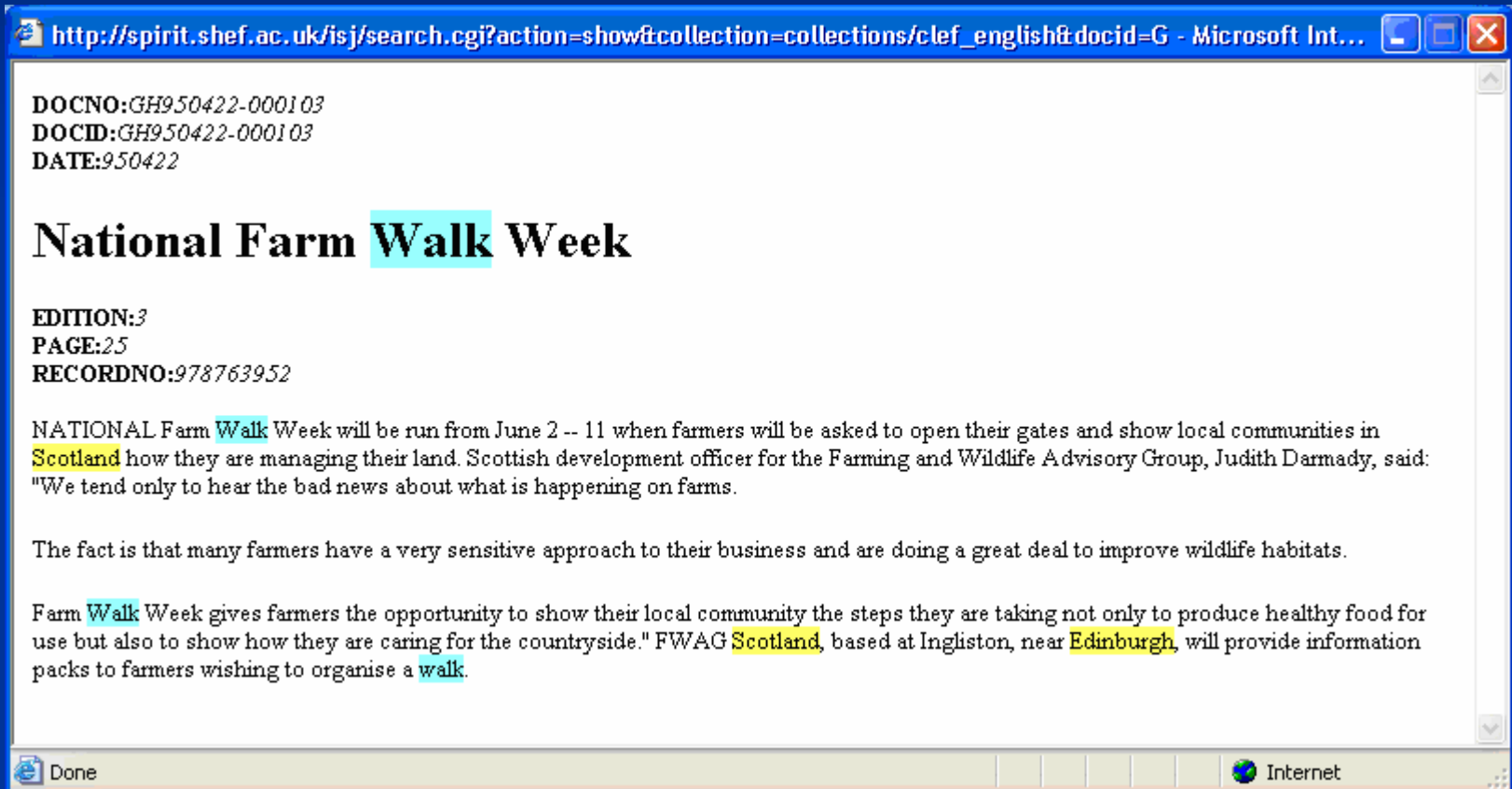
Search history

Recent query first

- 1: **Walking Scotland**
- 2: **trekking Scotland**
- 3: **hiking Scotland**
- 4: **Walking Scotland**
- 5: **Walking holidays in Scotland**

Internet

ISJ system used in GeoCLEF



http://spirit.shef.ac.uk/isj/search.cgi?action=show&collection=collections/clef_english&docid=G - Microsoft Int...

DOCNO:GH950422-000103
DOCID:GH950422-000103
DATE:950422

National Farm Walk Week

EDITION:3
PAGE:25
RECORDNO:978763952

NATIONAL Farm Walk Week will be run from June 2 -- 11 when farmers will be asked to open their gates and show local communities in Scotland how they are managing their land. Scottish development officer for the Farming and Wildlife Advisory Group, Judith Darmady, said: "We tend only to hear the bad news about what is happening on farms.

The fact is that many farmers have a very sensitive approach to their business and are doing a great deal to improve wildlife habitats.

Farm Walk Week gives farmers the opportunity to show their local community the steps they are taking not only to produce healthy food for use but also to show how they are caring for the countryside." FWAG Scotland, based at Ingliston, near Edinburgh, will provide information packs to farmers wishing to organise a walk.

Done Internet