



SIGIR 2007 Workshop "Improving Web retrieval for non-English queries"

Amsterdam - 27 July 2007

" iNEWS 07 " (improving Non English Web Searching)

a SIGIR 2007 Workshop

Important Dates

25 May 2007 Paper submissions due
15 June 2007 Notifications of acceptance
22 June 2007 Camera-ready copy due
27 July 2007 Workshop

<http://rea.teimes.gr/~lazarinf/ir7w/>

Call for Papers

Workshop Theme:

Over 60% of the online population are non-English speakers and it is probable the number of non-English speakers is growing faster than English speakers. Recent studies showed that non-English queries and unclassifiable queries have nearly tripled since 1997. Most search engines were originally engineered for English. They do not take full account of inflectional semantics nor, for example, diacritics or the use of capitals.

The main conclusion from the literature is that searching using non-English and non-Latin based queries results in lower success and requires additional user effort so as to achieve acceptable recall and precision. Further international search engines (like Yahoo and Google) are relatively weaker with monolingual non-English queries.

New tools and resources are needed to support researchers in non-English retrieval. New methodologies need to be proposed which will help the identification of problems in existing search engines. New teaching strategies should be formed aiding users to become more efficient in formulating their queries.

Aims and Topics:

The main objectives of the workshop are to propose techniques and to evaluate tools which improve the effectiveness of the existing search engines. The specific aims of the workshop are:

- Evaluate search engines in non-English queries and measure the additional user effort.
- Define methodologies for evaluating the effectiveness of search engines in non-English queries.
- Study the user query patterns in non-English Web retrieval.
- Identify the factors that influence utilization of search engines in a multicultural world.
- Propose extensions to the search engines to improve non-English Web retrieval.
- Propose teaching strategies for helping users improve their searching behaviour.
- Identify how standard IR techniques (Indexing, Query representation, Query reformulation, etc) can be adapted in Web retrieval for non-English languages.
- Discuss the application of natural language processing techniques for non-English Web IR.

Workshop areas of interest include, but are not limited to:

- Evaluation methodologies
- Analysis of query logs
- Localization of search engine interfaces
- Performance issues of local search engines
- NLP applications in non-English IR
- User studies
- Image and video retrieval services
- Summarization
- Teaching

Invited speaker:

Maarten de Rijke <http://staff.science.uva.nl/~mdr/>

Journal publication:

We are actively pursuing an agreement to publish extended versions of selected papers from the workshop in a special issue of a prestigious journal.

Location and Date

The workshop will be held as part of **SIGIR 2007** in Amsterdam, July 27, 2007.

Submissions

Authors are invited to submit full papers or posters in PDF format using **the ACM template page**.

The workshop accepts two types of submissions: Full papers (8 pages max) and Posters (4 pages max).

Accepted papers will be included in the print proceedings which will be distributed among the participants. Extended abstracts of the posters will also be included in the proceedings (proceedings will have an ISBN).

An electronic version of the proceedings will also be available online.

Please use the online submission system (<http://www.easychair.org/iNEWS07/>) to submit your paper.

(PDFs should not contain author information. So please exclude author names and affiliations and all other information associated with the authors.)

If you have any questions, comments, suggestions, please contact **Fotis Lazarinis** (lazarinf AT teimes.gr) and/or **Jesus Vilares** (jvilares AT udc.es).

Workshop Organisers

Workshop Chairs

Fotis Lazarinis, Technological Educational Institute, Mesolongli, Greece

Jesus Vilares Ferro, University of A Coruna, Spain

John Tait, University of Sunderland, UK

Programme committee (alphabetically)

1. Miguel Alonso (Univ. A Coruña, Spain)
2. Kuang-hua Chen (National Taiwan University, Taiwan)
3. Theodore Dalamagas (National Technical Univ. of Athens, Greece)
4. Chu-Ren Huang (Academia Sinica, Taiwan)
5. Ghassan Kanaan (Yarmouk University, Jordan)
6. Noriko Kando (National Institute of Informatics, Japan)
7. David Losada (Univ. Santiago de Compostela, Spain)
8. Alexandros Ntoulas (Microsoft Search Labs, USA)
9. Doug Oard (Univ. of Maryland, USA)
10. Gabriel Pereira (Universidade Nova de Lisboa, Portugal)
11. Carol Peters (ISTI-CNR, Italy)
12. Owen Rambow (Columbia University, USA)
13. Mark Sanderson (Univ. of Sheffield, UK)
14. Jacques Savoy (Univ. of Neuchatel, Switzerland)
15. Nasredine Semmar (LIC2M/CEA-LIST, France)

- | | |
|-----------------------|---|
| 16. Min Song | (New Jersey Institute of Technology, USA) |
| 17. Sofia Stamou | (Univ. of Patras, Greece) |
| 18. Richard Sutcliffe | (Univ. Limerick, Ireland) |
| 19. Manuel Vilares | (Univ. Vigo, Spain) |

Sponsorship

The workshop is partially funded by: "Rede Galega de Procesamento da Linguaxe e Recuperacion de Informacion (Galician Network for Language Processing and Information Retrieval), funded by **Xunta de Galicia**"