

9th European Summer School in Information Retrieval

Granada, September 2-6, 2013



- Final report -

October 21, 2013

The Summer School

The European Summer School in Information Retrieval (ESSIR – <http://ims.dei.unipd.it/websites/essir>) is a scientific event with temporal regularity founded in 1990, which gave rise to a series of Summer Schools held to provide high quality teaching of information retrieval and advanced information retrieval topics to a mostly European audience of researchers and research students. ESSIR is typically a week-long event consisting of a series of lectures and seminars delivered from invited lecturers recognized as experts in the field.

This 9th edition of the ESSIR (<http://www.ugr.es/~essir2013>) was held in Granada, Spain, in September 2-6, 2013, organized by the Spanish IR Society and co-organized the MUMIA COST Action. In this training school, 15 lecturers were invited to give a talk about IR foundations and hot topics. There were a total of 75 attendants from Europe, Africa and South America.

The school was held at the Technical School of Informatics and Telecommunications (Escuela Técnica Superior de Ingenierías Informática y de Telecomunicación - ETSIIT) at the University of Granada. The school venue was located in the Aynadamar Campus of the University, which is in walking distance from the city centre (approximately 20 minutes on foot), and also connected by several bus links from the suggested hotels and residences.

Organizers

The general Co-chairs of ESSIR 2013 were Juan Manuel Fernández-Luna and Juan Huete, from the University of Granada. The rest of the Organization Committee includes the following members:

- Pablo Castells (Universidad Autónoma de Madrid, Spain) – MUMIA Link.
- Mike Salampanis (Alexander Technology Educational Institute (ATEI) of Thessaloniki) – MUMIA Link.
- Fidel Cacheda (Universidad de A Coruña, Spain).
- Luis M. de Campos (Universidad de Granada, Spain).
- Irene Díaz (Universidad de Oviedo, Spain).
- Juan Antonio Martínez Comeche (Universidad Complutense de Madrid, Spain).
- Jesús Vilares (Universidad de A Coruña, Spain).

In addition, There were a group of six students from the ETSIIT who were acting as ESSIR volunteers on site.

Financial Support

The summer school was financially supported by the following institutions and companies:

- Google.
- The British Computer Society Information Retrieval Specialist Group.
- Yahoo! Labs (providing several scholarships),



- European Science Foundation by means of ELIAS.
- MUMIA COST Action, paying directly the cost of 8 lecturers.
- FINDWISE.

And finally the University of Granada as institution, by means of the Computer Science School, the Faculty of Information Science and the Computer Science and Artificial Department, providing the corresponding venues for the sessions and several activities.

Scholarships

- There were a good number of scholarships for Master or PhD students to attend the summer school provided by MUMIA (31). This was very important for the success of the ESSIR in terms of this high number of attendants but also with this financial support lots of students who could not have attended the conference because of not having enough money were able to do it. The grants (600€ per scholarship) covered some of the students' expenses. All the logistics needed for this task was supplied by the COST Action.
- Yahoo! Labs also offered a total of 2000€ for students grants. In this case, 10 scholarships were given to students coming from the furthest places who did not get any MUMIA grant.
- The benefits obtained from the ESSIR 2013 (3000€) are being shared to all the students who did not get one of the scholarships previously referred.

Lecturers and lectures

The following 17 lecturers, with their corresponding lecture titles and summaries, were participated at the ESSIR:

- *Indexing and MapReduce* by **Enrique Alfonseca (Google Research Zurich)**. For web search engines, due to the large amount of data available in the web, it is necessary to generate the index of the crawled documents using distributed processing techniques. Amongst these, one of the most commonly used paradigms is MapReduce, which provides a robust, powerful and conceptually simple way of implemented distributed processes. This class provided an introduction of Indexing in Information Retrieval systems to MapReduce for distributed programming.
- *Diversity* By **Charles Clarke (University of Waterloo)**. For a given query, a search engine should respond with a ranked list that reflects the breadth of available information, the range of possible user needs, and any ambiguity inherent in the query. To satisfy these requirements, researchers have proposed ranking algorithms and evaluation methodologies for novelty and diversity, which we will cover in this lecture. Clarke presented some of the most successful diversification algorithms, including MRR and xQuAD, comparing implicit and explicit diversification approaches. He outlined key properties required from evaluation measures, discuss widely used measures, and review international evaluation experiments conducted at TREC and NTCIR. Finally, he suggested directions for future research.
- *IR Foundations and Formal Models I and II* by **Bruce Croft (University of Massachusetts Amherst)**. In this talk, Croft overviewed of Information Retrieval (IR), including the important concepts, terminology, relationships to other fields of Computer Science, and key research issues. He then focused on the properties of text and text documents and how structure in a text corpus can be discovered through techniques such as text analysis, link analysis, information extraction, clustering, and classification. Finally, he described the crucial role of queries and information needs and provided an overview of the retrieval models.
- *IR Evaluation I and II* by **Maarten de Rijke (University of Amsterdam)**. He presented an overview of the state-of-the-art IR evaluation techniques, offering a deep sight of each one, as well as their main properties, advantages and disadvantages.
- *Recommender Systems* by **Alexandros Karatzoglou (Telefónica Research)**. Recommender Systems have bloomed in an attempt to serve users' information needs without requiring the user to enter an explicit query. These systems are built on engines that use many of the general IR constructs, but have personalization at its core, and do not require any text processing component. The recommendation field has since evolved to include many different approaches that go from tensor factorization to personalized learning to rank models. In this session, Karatzoglou gave an overview of the Recommendation problem, describing the different basic approaches with references to the literature as well as practical examples of their use in industry.
- *Recipes for PhD* By **Milad Shokouhi (Bing at Microsoft Research Cambridge)**. In this talk, Shokouhi summarized all the steps needed for making a PhD, offering very valuable information to all the students who

are involved in such process.

- *Natural Language Processing for Information Retrieval: an informal overview* by **Hugo Zaragoza (Websays)**. Improving IR models (and search technology in general) with Natural Language Processing has proven remarkably hard. This talk offered a summary of some of the (few) successes and (numerous) failures in this area, including many examples from my own work in ranking models and web search applications.
- *Multilingual Information Retrieval* by **Paul Clough (University of Sheffield)**. In this session he considered how to adapt IR methods to multilingual settings (referred to as MLIR and CLIR). The session was structured into the following: (1) the motivation for multilingual retrieval; (2) MLIR and CLIR techniques; and (3) the challenges of providing multilingual retrieval.
- *IR and Social Media* by **Arjen de Vries (CWI)**. The first part of his talk dealt with the presentation of the concept of social media. In the second part, Arjen went inside the opportunities for information retrieval research. These social media also form a challenge for information retrieval research: the many platforms vary in functionalities, and we have only very little understanding of clearly desirable features like combining tag usage and ratings in content recommendation! In the final part of the lecture he briefly touch upon an even wider range of opportunities, where data derived from social media form a key component to enable new research and insights.
- *Interactive IR* by **Norbert Fuhr (University of Duisburg-Essen)**. This course focused on three aspects of interactive retrieval, namely quantitative modeling, cognitive models, and user interface design. For the quantitative models, the interactive probability ranking principle was introduced along with methods for estimating the required parameters and for constructing Markov models of the user's interaction with the system. Cognitive models will cover a variety of approaches describing information seeking and searching.
- *Opinion Retrieval: Looking for the Opinion in the Wild* by **Georgios Paltoglou (University of Wolverhampton)**. Opinion Retrieval combines approaches and theories from two distinct but related areas of research; Information Retrieval and Sentiment Analysis. The former is necessary to retrieve content which is relevant to the user's information need and the latter to detect and analyse any affective content within it. Within the course, the students will become familiar with all aspects relating to effectively and efficiently distilling opinions from social media with a particular focus on the current state-of-the-art approaches on sentiment analysis and social information retrieval.
- *Contextual Search: Issues and Challenges* by **Gabriella Pasi (Università degli Studi di Milano Bicocca)**. To overcome the "one size fits all" behaviour of most search engines, in recent years a great deal of research has addressed the problem of defining techniques aimed at tailoring the search outcome to the user context in order to improve the quality of search. The main idea is to produce context-dependent and user-tailored search results. Search tasks are subjective and often complex. The user-system interaction, based on keyword-based querying and on the presentation of search results as a list of web pages ordered according to their estimated relevance, is often unsatisfactory. This lecture will present an overview of the main issues related to contextual search.
- *Designing the Search Experience* by **Tony Russel-Rose (UXLabs)**. In this session he reviewed the basic principles of search usability, with a focus on practical solutions that integrate information-seeking theory with UI design best practice. He explored the fundamental concepts of user-centred design for information search and discovery and showed how to differentiate between various types of search behaviour: known-item, exploratory, lookup, learning, investigation, etc. He reviewed the primary 'dimensions' of search user experience and how to apply them to different contexts.
- *MUMIA: Integrating IR Technologies for Professional Search* by **Michael Salampanis (Alexander Technology Educational Institute (ATEI) of Thessaloniki)**. This talk presented MUMIA which is a research networking activity that brings together various facets of state-of-the-art search technology research which can contribute to the development of search tools for next generation professional search systems. Additionally, it presented a general framework which provides a useful topology for better understanding the design space of professional search systems and how different IR/NLP technologies can be integrated to enable rich information seeking environments where different tools can support specific objectives within a typically lengthy search process.
- *Patent Retrieval: A Research Challenge for Multi-media Information Retrieval Research* by **John Tait (johntait.net)**. The lecture reviewed the state of the art and point out where the key challenges are, especially for early stage researchers in Multi-media Information Retrieval.

ESSIR Final Programme

| Monday September 2nd 9:00h – 9:45h | Tuesday September 3rd 9:30h – 11:00h | Wednesday September 4th 9:00h – 10:45h | Thursday September 5th 9:00h – 10:45h | Friday September 6th 9:00h – 10:45h |
|--|--|--|--|---|
| <i>ESSIR 2013 Opening Session</i> | <i>FDIA Welcome and Introduction (Lecture hall). Group Activity 1 (Faculty of Arts).</i> <i>Liad Kelly and Leif Azzopardi</i> | <i>IR and Social Media Arjen de Vries</i> | <i>Interactive IR Norbert Fuhr</i> | <i>Multilingual IR Paul Clough</i> |
| 9:45h – 11:15h | 11:00h – 11:30h | 10:45h – 11:15h | 10:45h – 11:15h | 10:45h – 11:15h |
| <i>IR Foundations and Formal Models I</i> <i>Bruce Croft</i> | Coffee Break | Coffee Break | Coffee Break | Coffee Break |
| 11:15h – 11:30h | 11:30h – 13:30h | 11:15h – 13:00h | 11:15h – 13:00h | 11:15h – 13:00h |
| Coffee Break | <i>FDIA Paper Session (Lecture Hall)</i> | <i>Natural Language Processing for Information Retrieval: an informal overview Hugo Zaragoza</i> | <i>Contextual search: Issues and Challenges Gabriella Pasi</i> | <i>Recommender Systems Alexandros Karatzoglou</i> |
| 11:30h – 13:00h | 13:30h – 15:00h | 13:00h – 14:15h | 13:00h – 14:15h | 13:00h – 14:15h |
| <i>IR Foundations and Formal Models II</i> <i>Bruce Croft</i> | Lunch | Lunch | Lunch | Lunch |
| 13:00h – 14:30h | 15:00h – 16:00h | 14:15 – 16:00h | 14:15 – 16:00h | 14:15 – 16:00h |
| Lunch | <i>FDIA Keynote Speech (Lecture Hall) Recipes for PhD Milad Shokouhi</i> | <i>A Research Challenge for Multi-media Information Retrieval Research John Tait</i> | <i>Opinion Retrieval: Looking for the Opinion in the Wild Georgios Paltoglou</i> | <i>Diversity Charles Clarke</i> |
| 14:30h – 16:00h | 16:00h – 17:00h | 16:00h – 16:30h | 16:00h – 16:30h | 16:00h – 16:30h |
| <i>IR Evaluation I</i> | <i>FDIA</i> | Coffee Break | Coffee Break | Coffee Break |

| | | | | |
|---|--|-----------------------------|---------------------------------|------------------------|
| Maarten de Rijke | Group Activity 2 (Faculty of Arts) | | | |
| 16:00h – 16:30h | 17:00h – 19:00h | 16:30h – 18:15h | 16:30h – 18:15h | 16:30h – 18:30h |
| <i>Coffee Break</i> | <i>FDIA</i> | Panel | Designing the Search Experience | Indexing and MapReduce |
| | <i>Poster Session (CS School Hall)</i> | CEI BioTic Granada | Tony Russell-Rose | Enrique Alfonseca |
| 16:30h – 18:00h | | 21:00h – 00:00h | 21:00h – 23:00h | 18:30h – 18:45h |
| <i>IR Evaluation II</i> | | <i>Social Activity</i> | <i>Social Activity</i> | ESSIR 2013 |
| Maarten de Rijke | | Night visit to the Alhambra | School Dinner | Closing Session |
| 18:00h – 18:45h | | | | |
| <i>MUMIA: Integrating IR Technologies for Professional Search</i> | | | | |
| Michail Salamopsis | | | | |
| 21:30h – ??:??h | | | | |
| <i>Social Activity</i> | | | | |
| Flamenco show | | | | |

Social Activities

The summer school contained 3 social activities in its programme. The objective was to include some acts out of the school sessions where the attendants could spend time socializing and discovering the culture of the city of Granada:

- Flamenco Show
- Night visit to the Alhambra.
- School Dinner.

FDIA Symposium

The 5th Symposium on Future Directions in Information Access (FDIA) was held in conjunction with the European Summer School in Information Retrieval in Granada, Spain. It was organized by Leif Azzopardi and Liah Kelly. The symposium aimed to provide a forum for participants of the summer school to share their research and interaction with senior researchers in an informal and relaxed atmosphere. It took place in September 2nd, during the whole day. There was a keynote speech given by Milad Shokouhi, which was very interesting and useful for the attendants involved in their PhD process.

Financial Support

The summer school was financially supported by the following institutions and companies:



- Google.
- The British Computer Society Information Retrieval Specialist Group.
- Yahoo! Labs (providing several scholarships),
- European Science Foundation by means of ELIAS.
- MUMIA COST Action, paying directly the cost of 8 lecturers and giving 31 scholarships.
- FINDWISE.

And finally the University of Granada as institution, by means of the Computer Science School, the Faculty of Information Science and the Computer Science and Artificial Department, providing the corresponding venues for the sessions and several activities.

We would like to thank all of them their support. Considering the times we are having in Spain, we could not have organized this summer school without their financial support.

Benefits for ESSIR Attendants

This summer school has been a very interesting experience for all the trainees in terms of different aspects:

- The scientific level of the lecturers and lectures were really high, so the trainees could get valuable knowledge about both the IR foundations and hot research topics.
- There were interesting discussions during and after each lecture.
- There were individual contacts between trainees and lecturers, which has originated research collaborations and research interships.
- The trainees have known other people working on the same, similar or complementary research areas, which has also originated research collaborations.
- From the point of view of the social activities, the attendants have time to spend time out of the lecture hall and the scientific context, what originated a very good atmosphere to share and exchange experiences.
- The attendants have also discovered the organizers of the summer schools as a very active research group, with a solid trajectory on IR.
- Low cost accomodation was offered to all the attendants.

ESSIR Evaluation

A survey was given to the ESSIR attendants. After the analysis of the results provided by them, we conclude that the summer school has been a success in terms of the high rates given to the scientific programme and organization. In fact, in a scale of 1 to 5 points, the attendants rated the quality of the summer school with an average of 4.59 points.

Juan Manuel Fernández Luna
Juan Francisco Huete Guadix