

11th European Summer School in Information Retrieval (ESSIR 2017)

Harrison Scells
Queensland University of Technology
harrison.scells@qut.hdr.edu.au

Abstract

The 11th European Summer School in Information Retrieval (ESSIR 2017) was held in Barcelona, Spain from the 4th to the 8th of September 2017 at the Universitat Pompeu Fabra. ESSIR consisted of a week of lectures and seminars delivered by top invited experts in the field of Information Retrieval. This year the program of ESSIR included 13 lectures, one symposium (FDIA), 13 student presentations, one poster session, one industry session, and two social events. This report is an overview of the successful summer school which attracted a total of 46 students from around the world.

1 Introduction

The European Summer School in Information Retrieval (ESSIR) is a scientific, biannual summer school focused on advanced Information Retrieval (IR) and search technology topics. Founded in 1990, ESSIR has since given rise to a number of summer schools that provide high quality teachings to an audience of research students and early stage researchers. ESSIR is often a week-long series of lectures and talks from academic and industry experts. This year, in addition to the traditional tracks, there was an emphasis on fields that are also applicable to Information Retrieval (i.e. Machine Learning), as well as other appealing tracks such as Music Information Retrieval, Semantic Search for Finance, and Neural Information Retrieval. The 11th ESSIR would not have been possible without the organisation by Ana Freire and Ricardo Baeza-Yates. The FDIA symposium, held in conjunction with ESSIR, was successfully chaired by Haiming Liu and Ingo Frommholz.

The 11th ESSIR was held in Barcelona, Spain from the 4th to the 8th of September, hosted by the Universitat Pompeu Fabra. Barcelona is the capital city of Catalonia, and the largest city in this region of Spain home to vibrant art culture, a bustling business district, world-renowned education institutions, and approximately 1.5 million people. Barcelona is also home to perhaps some of the most well known architectural buildings and monuments in Europe which have since been placed on the UNESCO World Heritage Sites list such as the *Sagrada Família*, *Casa Batlló*, and *Park Güel* (Figure 1).¹ Other notable sights of Barcelona include *The Arc de Triomf*, the *Santa Maria del Mar*, and the *Fabra Observatory* (Figure 1). The last of which was the location of the gala dinner social event, as part of ESSIR.

¹works by prominent 19th and 20th century architect *Antoni Gaudí*.



Figure 1: Sagrada Família, The Arc de Triomf, Santa Maria del Mar (inside), Fabra Observatory.

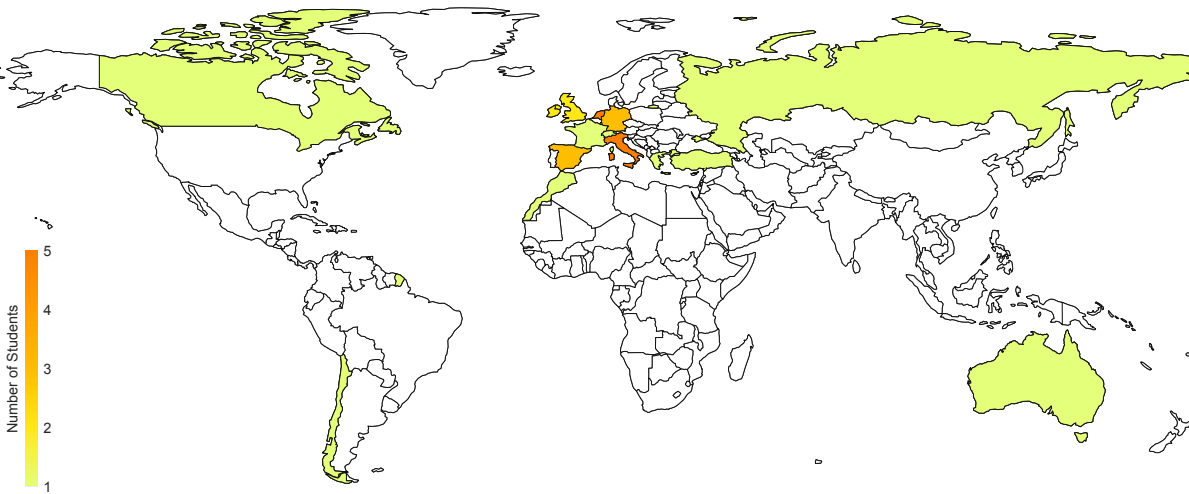


Figure 2: Country affiliation of the student attendees of ESSIR 2017.

The two social events of ESSIR this year enabled the student attendees to socialise and network with prominent experts in the field, a unique opportunity. The 16 experts were invited to give lectures and industry talks to 46 students (29 graduates, 11 industry, 6 academic). The students attending ESSIR were primarily from Europe, however there were also students from outside of Europe (Figure 2). The generous sponsorship of the ACM Special Interest Group on Information Retrieval (ACM SIGIR) and ThoughtWorks enabled many students to attend ESSIR that would have otherwise been unable to. In total 66 participants including students, lecturers, industry partners, and organisers attended ESSIR 2017.

2 Lectures

Lectures in this year’s ESSIR covered topics ranging from Information Retrieval fundamentals to Machine Learning, to applications of Information Retrieval such as Music IR and Semantic Search for Finance. This section provides an overview of the lectures for the 11th ESSIR.²

Foundations of Information Retrieval, *Maarten De Rijke (University of Amsterdam,*

²Slides for these lectures are available at <http://eventum.upf.edu/go/essir2017>.

Netherlands): This lecture opened the summer school to an overview of the foundations of Information Retrieval from a different perspective to the traditional “Computer Science stack” (i.e. indexing, evaluation, retrieval models, etc.). Instead, the picture presented in this lecture was that of sense, model, and act. Here, Information Retrieval is categorised into four areas: Front door (i.e. where the user and search engine interact), Evaluation (i.e. assessing the quality of a search engine’s responses), Online (i.e. generating responses to users’ queries), and Offline (i.e. preparing to address information needs). Each area was addressed with examples of what is currently state of the art and future directions.

An Introduction to System-Oriented Evaluation in Information Retrieval, *Mounia Lalmas (Yahoo Research @ Oath, United Kingdom)*: This lecture covered the need for system-oriented evaluation in Information Retrieval. System-oriented evaluation is a method of comparing the effectiveness of one system to one or more other systems. To perform this evaluation, the Information Retrieval system is measured in terms of effectiveness (the ability for the system to retrieve documents that are relevant to a users information need). This lecture provided an in-depth exploration of different measures for evaluating Information Retrieval systems.

Introduction to Machine Learning for Information Retrieval, *Hugo Zaragoza (Amazon, Spain)*: This lecture offered a gentle introduction to Machine Learning (ML) for Information Retrieval. Two approaches to ML for Information Retrieval were discussed: Statistical ML and Probabilistic ML. Many of the common pitfalls to ML in this field were described, for example features and separability, and the curse of dimensionality. The lecture concluded with an application of ML in Information Retrieval by exploring the currently hot topic of learning to rank.

Neural Information Retrieval, *Joemon Jose (University of Glasgow, United Kingdom)*: This technical lecture provided a low-level exploration of the application of neural networks for Information Retrieval. Specifically, this lecture explored neural networks in general and their learning functions, and word representations. For each of the topics covered, the issues faced by researchers and future directions for researchers were discussed.

Challenges of Conversational Search, *Enrique Alfonseca (Google, Switzerland)*: This lecture discussed the challenges that conversational search faces in providing information to users. Specifically, the lecture explored language generation and internationalisation techniques (and the structured and unstructured sources that form answers), user interface challenges for voice search, and sentence compression and summarisation. The lecture was contextualised to production systems and improving the user experiences for question answering.

Contextual Search, *Gabriela Pasi (University of Milano-Bicocca, Italy)*: In this lecture, the topic of information need via interaction and context was discussed. This lecture explored the various limitations of Information Retrieval systems that do not integrate models of context, and emphasised the importance of the user. Next, the lecture explored personalisation in search in order to exploit the user to improve search quality. Finally, current and future trends in mobile and social search were discussed and how to exploit context in these areas.

Deep Learning for Recommender Systems, *Alexandros Karatzoglou (Telefónica Research, Spain)*: This lecture covered several deep learning approaches used in recommender systems. After providing an explanation of deep learning in general, the focus of the lecture turned to applications. Here, the lecture covered word embeddings (e.g. word2vec, paragraph2vec, doc2vec, etc.) and neural network models (i.e. convolutional neural net-

works and recurrent neural networks). The lecture was concluded by outlining the future directions for deep learning in recommender systems.

Semantic Search for Finance, *Edgar Meij (Bloomberg, United Kingdom)*: This lecture was an overview of the unique problems and solutions faced by Bloomberg. Specifically, the lecture covered different retrieval models (i.e. language modelling, entity retrieval, knowledge graphs), ranking a mixture of structured and semi-structured content using a learning to rank, and using Information Retrieval techniques for streaming and alerting (e.g. trending news, stock market).

Music Information Retrieval: Overview, Challenges and Current Trends, *Emilia Gómez (Universitat Pompeu Fabra, Spain)*: This lecture provided a survey of the application of Information Retrieval to music. The focus of the lecture was the extraction of semantically meaningful information from various sources such as audio, lyrics, web pages or microblogs. The lecture also explored the main tasks of music IR: music identification and transcription, search by similarity, query by humming, music recommendation, and playlist generation. The lecture concluded with a summary of the future directions for music IR.

Visual Retrieval and Mining, *Stefan Rueger (Open University, United Kingdom)*: This lecture discussed the challenges of Information Retrieval within the context of images and videos (visual retrieval). Various techniques for representation and retrieval were discussed, as well as problems such as query-by-image. In addition, machine learning approaches to visual IR was highlighted and explored, and the lecture was concluded with a speculative session for the future of the field.

User Experience in Information Retrieval, *Tony Russel-Rose (UXLabs, United Kingdom)*: This lecture discussed the user interfaces of Information Retrieval systems. The lecture explored introductory concepts of design for user experiences by comparing different search interfaces and provided an introduction to practical user experience design.

Scalability and Efficiency Challenges in Large-Scale Web Search Engines, *Berkant Barla Cambazoglu (NTENT, Spain)*: This lecture provided a comprehensive overview of the scalability and efficiency challenges of web search engines. The lecture detailed the many aspects of a typical web search engine, and was focused on web crawling, indexing, and query processing. For each aspect, the scalability and efficiency challenges were discussed at four levels: a single computer, a cluster of computers, a single data centre, and a multi data centre.

Query Log Mining, *Fabrizio Silvestri (Facebook, United Kingdom)*: This lecture covered the many facets of mining search engine query logs to improve the search quality. In particular, the lecture identified a taxonomy of query classes (e.g. navigational, informational, resource) and how to use the query logs to improve search (i.e. by refining the query, query expansion, query recommendation). These techniques were contextualised to large, distributed systems.

3 Industry Talks

In addition to the excellent lectures on Information Retrieval, there were three industry talks. These talks explored Information Retrieval in industry and provided an insight into what challenges companies face and how they are approached. Each talk was approximately half an hour long and allowed the students of the summer school to see how Information Retrieval is used in industry. The first talk shared how **NTENT** transforms large structured

and unstructured content, and how the information is used to predict and deliver relevant information to clients. In the second talk, **Schibsted Media Group** discussed the challenges they face managing, storing, and searching for millions of products across Europe and Asia. The final talk by **ThoughtWorks** discussed their community and culture of like-minded engineers working on problems to make the world a better place. These three talks provided an interesting, industry-focused perspective for the attendees of the summer school.

4 Future Directions in Information Access

The Future Directions in Information Access (FDIA) Symposium was held in conjunction with the summer school. It provided a forum for PhD students, researchers new to the field, and post-doctoral students of the summer school to share their research and interact with senior researchers in an informal and relaxed atmosphere. The objectives of FDIA are to (i) provide an accessible forum for early researchers (particularly PhD students, and researchers new to the field) to share and discuss their research, (ii) create and foster formative and tentative research ideas, and (iii) encourage discussion and debate about new future directions.

The 7th FDIA programme committee was chaired by Haiming Liu and Ingo Frommholz from the University of Bedfordshire and Leif Azzopardi from the University of Strathclyde. This year, there were 13 accepted submissions (87% acceptance rate) to the symposium within a wide range of novel and emerging topics including: health and medical information retrieval, collaborative search, personalization and geographical information search, emergency and trend detection, prediction of public reaction and search behaviour in microblogs, and trend mining using bibliographic data. The symposium comprised a series of fast-paced five minute talks from each of the participants, followed by a poster session. The 7th FDIA papers are listed below:

- **Supporting Complex Information Needs via Large-Scale Collaborative Search** - *Felipe Moraes* (Delft University of Technology)
- **Reducing Workload of Systematic Review Searching and Screening Processes** - *Harrisen Scells* (Queensland University of Technology)
- **Web application for development of domain information space for thematic information retrieval and reuse** - *Sarkisova Elena* (Volgograd State Technical University)
- **Early Prediction of Public Reactions to News Events Using Microblogs** - *Cagri Toraman* (Bilkent University)
- **Personalization in Geographical Information Search** - *Noemi Mauro* (University of Turin)
- **Detection of Trending Topic Communities: Bridging Content Creators and Distributors** - *Lorena Recalde* (Universitat Pompeu Fabra)
- **Detecting Emergency Situations by Inferring Locations in Twitter** - *Hernan Sarmiento* (University of Chile)
- **Exploring Search Behaviour in Microblogs** - *Anirban Chakraborty* (Trinity College Dublin)
- **Word Embedding based Approaches for Information Retrieval** - *Dwaiipayan Roy* (Indian Statistical Institute)



Figure 3: Setting for the welcome reception at Nova Icaria Beach.



Figure 4: View of Barcelona from the Fabra Observatory.

- **Developing Positive Emotion through Affective Design for Interactive Information Seeking** - *Sehrish Sher Khan* (University Of Bedfordshire)
- **Trend Mining on Bibliographic Data** - *Christin Katharina Kreutz* (Trier University)
- **Supporting Collaborative Information Seeking Through Shared Annotations** - *Samuel Dodson* (University of British Columbia)
- **User Information Needs Through Query Reformulation** - *Asad Ullah* (University of Bedfordshire)

The proceedings of the 7th FDIA Symposium have been published by the BCS Electronic Workshops in Computing (eWiC) series.³

5 Social Programme

The social programme of ESSIR allows the students of the school to network and socialise with each other and prominent figures in the Information Retrieval field. These unique

³<https://ewic.bcs.org/category/19215>



Figure 5: Moments from ESSIR 2017.

opportunities let the attendees relax and enjoy the beautiful setting of Barcelona. The social programme provided by the summer school included:

- Welcome reception at Mokaï Beach Bar (Figure 3)⁴, located in an exclusive environment at Nova Icaria Beach. Attendees enjoyed the sunset while drinking cocktails and eating appetisers.
- Gala dinner at the Fabra Observatory⁵ overlooking the city of Barcelona (Figure 4). Attendees enjoyed fine cuisine before exploring the interior of the Observatory where they interacted with one of the largest and oldest (1904) telescopes still in use in Europe.

⁴<http://anishgroup.com/en/mokai-beach-bar/>

⁵<https://www.sternalia.com/en/dinners-with-stars.html>

6 Conclusions

Overall, the 11th European Summer School for Information Retrieval was a success that brought many participants from around the world together for a week long intensive of lectures, talks, and social events. This school allowed students a unique opportunity to hear from experts and prominent figures in the field of Information Retrieval, and allowed participants of the Future Directions of Information Retrieval Symposium to receive highly valuable feedback about their research from the same experts and their peers. Social gatherings allowed for all attendees to enjoy the beauty of Barcelona and to form lasting connections that extend beyond the duration of the school. The consensus of the students (the author of this report being one of them) was that the 11th ESSIR was an overwhelming success with excellent lectures and talks in a historic and beautiful city.

The ESSIR 2017 website⁶ contains the slides and information about the 11th ESSIR, in addition to links to the 7th FDIA proceedings. Photos (Figure 5) and commentary from the 11th ESSIR can be found on the schools official twitter page.⁷

7 Acknowledgements

Thank you to all of the lecturers, the industry panel, the FDIA chairs, and all of the participants who attended the school. A special thank you to the volunteers from the Universitat Pompeu Fabra for assisting the organisers to help the school run smoothly during the entire week, and to the events organisation team at Universitat Pompeu Fabra.

Thank you most sincerely to the sponsors of ESSIR 2017: the Universitat Pompeu Fabra, the Maria de Maeztu Strategic Research Program, the Association for Computing Machinery Special Interest Group on Information Retrieval (ACM SIGIR) Facebook, ThoughtWorks, Brainly, Shibsted Media Group, Google, and the Information Retrieval Specialist Group of the British Computing Society (BCS-IRSG) who supported the publishing of the FDIA proceedings in the BCS eWiC series.⁸

⁶eventum.upf.edu/go/essir2017

⁷<https://twitter.com/essir2017>

⁸<https://ewic.bcs.org/category/19215>