



národní  
úložiště  
šedé  
literatury

## **European Summer School in Information Retrieval ESSIR 2005**

Linková, Zdeňka  
2006

Dostupný z <http://www.nusl.cz/ntk/nusl-35578>

Dílo je chráněno podle autorského zákona č. 121/2000 Sb.

Tento dokument byl stažen z Národního úložiště šedé literatury (NUŠL).

Datum stažení: 03.05.2024

Další dokumenty můžete najít prostřednictvím vyhledávacího rozhraní [nusl.cz](http://www.nusl.cz).



**Institute of Computer Science**  
**Academy of Sciences of the Czech Republic**

## **European Summer School in Information Retrieval ESSIR 2005**

Zdeňka Linková

Technical report No. 949

February 2006



**Institute of Computer Science**  
**Academy of Sciences of the Czech Republic**

## **European Summer School in Information Retrieval ESSIR 2005**

Zdeňka Linková<sup>1</sup>

Technical report No. 949

February 2006

### Abstract:

Information Retrieval (IR) as a process of searching relevant information is a significant discipline of a data processing field. European Summer School in Information Retrieval ESSIR provides students, academic and industrial researchers and developers a grounding in the core objects of IR (models, architectures, algorithms), as well as covering some current topics, e.g. information retrieval from the Web. We have participated its 5th year that was held at Dublin City University in Dublin, Ireland.

### Keywords:

Summer school, Information Retrieval.

---

<sup>1</sup>Institute of Computer Science, Academy of Sciences of the Czech Republic, 182 07 Prague 8, Czech Republic

# Contents

1	INTRODUCTION.....	2
2	EUROPEAN SUMMER SCHOOL IN INFORMATION RETRIEVAL.....	2
3	COURSE OVERVIEW.....	3
4	LECTURERS .....	4
5	SCHOOL BACKGROUND.....	7

# **1 Introduction**

Information retrieval is a science of searching information. It can be described as a process of identification of relevant documents or other information sources that is usually proceed as an action to users query. The term information retrieval includes processes, methods and procedures used for searching of information in data sources.

Automated information retrieval systems were originally used to manage information in libraries. With the information explosion enabled by a progress in computer science field, information retrieval systems spread also to other data processing areas.

Information retrieval is a broad interdisciplinary field that draws on many other disciplines. It stands at the junction of many established fields, and draws upon cognitive psychology, information architecture, information design, human information behavior, linguistics, information science, computer science, librarianship and statistics. Web search engines (such as Google) are amongst the most visible applications of information retrieval research.

## **2 European Summer School in Information Retrieval**

The European Summer School in Information Retrieval ESSIR 2005 [1] was held 5-9 September 2005 in Dublin, Ireland. It was hosted by Centre for Digital Video Processing, part of the Adaptive Information Cluster at Dublin City University.

ESSIR 2005 was the 5th European Summer School in Information Retrieval. The first one, ESSIR'90 was organized by the University of Padova (Prof. Maristella Agosti) and was held in Bressanone, Italy, in 1990. The 2nd ESSIR (ESSIR'95) was organized by the University of Glasgow (Prof. Keith van Rijsbergen) and held in Glasgow, Scotland, in 1995, in conjunction with the BCS Colloquium in Information Retrieval, an IRIDES project meeting, and the final MIRO workshop. The 3rd ESSIR (ESSIR 2000) was organized by Dr. Gabriela Pasi, Prof. Maristella Agosti and Prof. Fabio Crestani, and took place in Varenna, Italy, in 2000. The 4th ESSIR (ESSIR 2003) was held in Aussois, France and organized by Prof. Yves Chiaramella and Prof. Catherine Berrut.

The target audience of the Summer School were advanced undergraduate students, PhD students, post-doctoral researchers and academic and industrial researchers and developers. The Summer School lasted one week and the program included lectures from leading researchers in their fields, demonstrations and panels.

### 3 Course overview

This five-day school gave participants a grounding in the core subjects of IR - architectures, algorithms, formal theoretical models of IR, and evaluation - as well as covering some of the current topics such as retrieving from multimedia document collections, and retrieving over wide area networks such as the Internet and the Web.

The classes that were lectured:

- **Introduction to Information Retrieval** - by Prof. Keith van Rijsbergen
- **Information Retrieval Systems** - by Prof. Fabio Crestani
- **Web Information Retrieval [2]**- by Dr. Monika Henzinger
- **Formal Models** - by Dr. Djoerd Hiemstra
- **Evaluation in Information Retrieval [3]** - by Dr. Stefano Mizzaro
- **Users and Information Retrieval** - by Dr. Ian Ruthven
- **Multimedia Techniques** - by Dr. Noel Murphy & Dr. Noel O'Connor
- **Natural Language Processing in IR** - by Prof. Maarten de Rijke
- **Cross-Language Information Retrieval** - by Dr. Gareth Jones
- **Multimedia Information Retrieval** - by Prof. Alan Smeaton, Dr. Gareth Jones & Dr. Cathal Gurrin
- **Machine Learning** - by Dr. Fabrizio Sebastiani
- **Personalization** - by Prof. Barry Smyth
- **Structure/XML Retrieval** - by Dr. Mounia Lalmas

## 4 Lecturers

### **Prof. Fabio Crestani**

Fabio Crestani [4] holds a chair in Information Science at the Department of Computer and Information Sciences of the University of Strathclyde in Glasgow, U.K. His research interest is design, implementation, and evaluation of advanced information access systems concentrating on hypermedia, information retrieval, information filtering, and digital libraries. Previously he held research fellowship positions at the Rutherford Appleton Laboratory of the CLRC (UK), the International Computer Science Institute in Berkeley (USA), University of Glasgow (UK), and was Assistant Professor at the University of Padova (Italy). He holds a degree in Statistics from the University of Padova, and a M.Sc. and Ph.D. in Computing Science from the University of Glasgow.

### **Dr. Cathal Gurrin**

Cathal Gurrin [5] is a post-doctoral researcher working in the Centre for Digital Video Processing at Dublin City University (DCU), Ireland. He obtained his PhD from DCU in 2002 for his work on developing techniques to support searching on very large internet-sized web collections. His current research interests include Video Retrieval from large video libraries and image/photo retrieval, specifically the retrieval of personal photo collections. In addition, he is interested in large scale retrieval of WWW documents, and has recently been heavily involved in the development of a high-performance, scalable, distributed WWW search engine.

### **Dr. Monika Henzinger**

Monika Henzinger is a professor in the School of Computer and Communication Sciences, EPFL, in Lausanne, Switzerland. Until 2004 she was a director of research at Google. Before coming to Google, she was on the research staff of the Digital Systems Research Center in Palo Alto, California and on the faculty of the Computer Science departments at Cornell University and the University of Saarbruecken in Germany. She holds a Master degree in Computer Science from the University of Saarbruecken and a Ph.D. in Computer Science from Princeton University.

### **Dr. Djoerd Hiemstra**

Djoerd Hiemstra [6] is an assistant professor in the Database Group of the Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, The Netherlands. His research interests are multimedia databases, digital libraries, information retrieval models, cross-language information retrieval, and statistical language modeling.

### **Dr. Gareth Jones**

Gareth Jones [7] is a senior lecturer in the School of Computing and member of the Centre for Digital Video Processing, DCU, Ireland. From 1993-96 he was a member of the Speech, Vision and Robotics Group, Department of Engineering and Computer Laboratory, University of Cambridge, working as a Research Associate on the Video Mail Retrieval using Voice (VMR) project. From 1996-2003 he was a Lecturer in Media Computing in the Department of Computer Science at the University of Exeter. From 1997-98 he was a Toshiba Fellow and engineer at the Toshiba Corporation Research and Development Centre in Kawasaki, Japan. His research interests focus primarily on information processing issues in Human-Computer Interaction, with the topics including spoken document retrieval, Asian

language retrieval, cross-language information retrieval, speech recognition, natural language modeling, information visualization, mobile computing, interaction with virtual worlds, and affective computing.

**Dr. Mounia Lalmas**

Mounia Lalmas [8] is a reader in the Queen Mary Information Retrieval (QMIR) research group, Queen Mary University of London, U.K. Prior to this, she was a Research Scientist at the University of Dortmund in 1998, a Lecturer from 1995 to 1997 and a Research Fellow from 1997 to 1998 at the University of Glasgow, where she received her PhD in 1996. Her research focuses on the development and evaluation of intelligent access to interactive heterogeneous and complex information repositories. She has applied uncertainty theories, like the Dempster-Shafer theory of evidence, for the modeling of information retrieval tasks in an expressive and unified manner and covering a range of domains such as Web, XML, and MPEG-7. She is also co-leader of the international evaluation initiative for content-oriented XML retrieval (INEX), a large-scale project with over 50 participating organizations worldwide.

**Dr. Stefano Mizzaro**

Stefano Mizzaro [9] is an assistant professor at the Department of Mathematics and Computer Science in University of Udine, Italy. His research interests are the theoretical and practical aspects of information retrieval, digital libraries and scholarly publishing, mobile and wireless devices, and natural language processing.

**Dr. Noel Murphy**

Noel Murphy [10] is a Senior Lecturer in the School of Electronic Engineering, DCU, Ireland, where he has fifteen years experience of teaching and program development and management, at undergraduate and postgraduate levels. He is a member of the Centre for Digital Video Processing, DCU. His primary degree was in theoretical physics, and his postgraduate work was on Computer Vision. His research interests include image and video coding, 3-D imaging and visual perception. During the last decade he contributed to the development of the ITU-T H.263 and ISO MPEG-4 video-coding standards.

**Dr. Noel E. O'Connor**

Noel E. O'Connor [11] is a lecturer in the School of Electronic Engineering in DCU, Ireland, and is the Program Chair of the B.Eng. in Digital Media Engineering undergraduate degree program in the School. He is a member of the Centre for Digital Video Processing, DCU. His research interests are combining low-/mid- and high-level AV object features for enhanced video content indexing, AV scene-level classification for summarization and event detection, object segmentation and tracking in generic video, pedestrian tracking and monitoring using 3-D vision techniques, traffic monitoring using computer audition techniques, multi-modal video analysis for safety/security applications, and low-power hardware for advanced video processing operations on mobile devices.

**Prof. Maarten de Rijke**

Maarten de Rijke [12] is a full professor in the University of Amsterdam, The Netherlands, and the leader of the Language and Inference Technology group within the university. His core research theme is intelligent information access to which he brings to bear ideas from information retrieval, applied natural language processing, and knowledge representation and reasoning.

### **Dr. Ian Ruthven**

Ian Ruthven [13] is a lecturer in the Department of Computer and Information Sciences, University of Strathclyde, U.K. His research interests are Interactive Information Retrieval, formal methods for information-seeking, interfaces to information retrieval systems, interfaces to digital libraries, document summarization, affective computing and information-seeking behaviors.

### **Dr. Fabrizio Sebastiani**

Fabrizio Sebastiani [14] is an associate professor in the Department of Pure and Applied Mathematics, University of Padova, Italy. His research interests are information retrieval, machine learning, human language technology, with particular emphasis on text categorization and its applications. Prior he worked as research scientist at the Department of Linguistics in the University of Pisa, and Institute of Information Science and Technologies (ISTI) in the Italian National Research Council (CNR).

### **Prof. Alan F. Smeaton**

Prof. Alan F. Smeaton [15] is a professor of Computing at Dublin City University, Ireland, where he is Director of the Centre for Digital Video Processing. Previously he was Dean of the Faculty of Computing and Mathematical Sciences and Head of the School of Computing in DCU. His early research interests covered the application of natural language processing techniques to information retrieval (text) but this has broadened to cover the indexing and content-based retrieval of information in all media, text, image, audio and especially digital video. He holds the B.Sc., M.Sc. and PhD degrees in Computer Science from the National University of Ireland.

### **Prof. Barry Smyth**

Barry Smyth [16] is a professor and Head of the Department of Computer Science, University College Dublin, Ireland. He holds the Digital Chair of Computer Science in UCD and also a co-founder, director, and Chief Technical Officer of ChangingWorlds Ltd. His research covers a broad set of topics within artificial intelligence including case-based reasoning, machine learning, user modeling and planning with particular focus on so-called personalization techniques, which looks at ways of combining ideas from these areas to develop information systems that automatically learn about, and adapt to, the needs of individual users.

### **Prof. Keith van Rijsbergen**

Keith van Rijsbergen [17] is a professor and leader of the Information Retrieval Group in the Department of Computing Science at the University of Glasgow, U.K. Since about 1969 his research has been devoted to information retrieval, covering both theoretical and experimental aspects. He has specified several theoretical models for information retrieval and seen some of them from the specification and prototype stage through to production. His current research is concerned with the design of appropriate logics to model the flow of information and the application of Hilbert Space theory to content-based information retrieval.

## **5 School Background**

ESSIR 2005 was held in Dublin City University (DCU) in Ireland, hosted by Centre for Digital Video Processing. Dublin City University is a young university, situated on an 85 acre campus three miles north of the River Liffey in the city centre and just a 15-minute drive from Dublin airport. It has all the usual facilities and amenities of a modern university including computers labs with Internet access for students and lecturers, library, sports centre and swimming pool, residences, supermarket, bars, cafes, and restaurants. With the city just a 10-minute bus drive away, students of DCU have the best of both worlds; the social and cultural benefits of city life, but with the security and vibrancy of a university campus built very much for today.

### **Registration**

Registration details could be found at school's website, however, on-line registration was not offered; a registration form had to be faxed. This made no difficulties; in case that any occurs, contact address was offered. Registration fees were distinguished according to the date of registration, and also student reduced fees were available.

Registration fees included a copy of the lecture slides, stationery, map, and a day trip in social program. Although, I would appreciate to get or to have a possibility to buy also lecture notes or lecture proceedings.

### **Accommodation**

For Summer school participants, some rooms in Dublin City University Campus Residences were reserved. For these, who did not want to stay on campus, a list of guest houses, Bed & Breakfast and hotels nearby was available.

### **Social program**

In a social program, a day trip on Wednesday afternoon was organized. All trip taxes (transport, admission, pub grub, evening meal & entertainment) were already included in the registration tax. The trip was taken to famous Glendalough, the old village with an ancient monastery. Then a dinner in an original Irish pub was taken, and the evening was ended with dancing and singing Irish songs.

## Acknowledgements

I participated in the European Summer School in Information Retrieval ESSIR 2005 with my colleagues Martin Řimnáč and Roman Špánek. Our participation was supported by the project 1ET100300419 of the Program Information Society (of the Thematic Program II of the National Research Program of the Czech Republic) “Intelligent Models, Algorithms, Methods and Tools for the Semantic Web Realization”.

## References

- [1] ESSIR 2005 homepage: <http://www.cdvdp.dcu.ie/ESSIR2005/index.html>.
- [2] Henziger M.: “Hyperlink Analysis on the World Wide Web”, <http://www.cdvdp.dcu.ie/ESSIR2005/p-1.pdf>
- [3] Miyaro S.: “Lecture Slides + example excel files”, <http://www.cdvdp.dcu.ie/ESSIR2005/Mizzaro-ESSIR05.zip>
- [4] Fabio Crestani’s homepage: <http://www.cis.strath.ac.uk/~fabioc/>.
- [5] Cathal Gurrin’s homepage: <http://www.computing.dcu.ie/~cgurrin/>.
- [6] Djoerd Hiemstra’s homepage: <http://wwwhome.cs.utwente.nl/~hiemstra/>.
- [7] Gareth Jones’s homepage: <http://www.computing.dcu.ie/~gjones/>.
- [8] Mounia Lalmas’s homepage: <http://www.dcs.qmul.ac.uk/~mounia/>.
- [9] Stefano Mizzaro’s homepage: <http://www.dimi.uniud.it/mizzaro/>.
- [10] Noel Murphy’s homepage: <http://www.eeng.dcu.ie/~murphyn/home.html>.
- [11] Noel E. O'Connor’s homepage: <http://www.eeng.dcu.ie/~oconnorn/>.
- [12] Maarten de Rijke’s homepage: <http://staff.science.uva.nl/~mdr/page8/page8.html>.
- [13] Ian Ruthven’s homepage: <http://www.cis.strath.ac.uk/~ir/>.
- [14] Fabrizio Sebastiani’s homepage: <http://nmis.isti.cnr.it/sebastiani/>.
- [15] Alan F. Smeaton’s homepage: <http://www.computing.dcu.ie/~asmeaton/>.
- [16] Barry Smyth’s homepage: <http://dejavu3.ucd.ie/>.
- [17] Keith van Rijsbergen’s homepage: <http://www.dcs.gla.ac.uk/~keith/>.