

# Autumn School for Information Retrieval and Information Foraging 2019

Felix Engl  
University of Bamberg  
*felix.engl@uni-bamberg.de*

Robin Jegan  
University of Bamberg  
*robin.jegan@uni-bamberg.de*

Leon Martin  
University of Bamberg  
*leon.martin@uni-bamberg.de*

## Abstract

The Autumn School for Information Retrieval and Information Foraging (ASIRF) 2019 took place at Schloss Dagstuhl in Germany from September 22nd to 27th. The event featured eight lectures and tutorials from information retrieval experts and stood out due to the diversity of the participants, both regarding their cultural background and research. A varied social program complemented the scientific exchange.

## 1 Introduction

The *Autumn School for Information Retrieval and Foraging* (ASIRF) is a repeating event that brings together Information Retrieval (IR) experts and students from around the world. The goal of the event is to share insights and discuss both fundamental and specialized topics related to IR. ASIRF 2019 took place at the *Schloss Dagstuhl – Leibniz Center for Informatics* which is located in Saarland, a federal state in the west of Germany, from September 22nd to 27th. Schloss Dagstuhl is well known for its excellent environment, catering, and facilities that encourage working on scientific topics. Overall, 21 students, mostly PhD candidates, from six countries participated (see Fig. 1). The event featured eight lectures and tutorials from IR experts over the course of six days while providing enough opportunities for group work, networking, and social activities. This report summarizes the lectures and tutorials, and gives a general impression of the time at Schloss Dagstuhl.

## 2 Lectures and Tutorials

This section summarizes the eight lectures and tutorials of ASIRF 2019. Each lecture took about three hours with a coffee break for intermediate discussions in between.



Figure 1: The lecturers David Elsweiler, Ingo Frommholz and Ralf Schenkel and the participants of ASIRF 2019 in front of the chapel of Schloss Dagstuhl

### **Foundations of IR (Ingo Frommholz, University of Bedfordshire)**

The lecture of Ingo Frommholz provided an overview of fundamental aspects of IR. After a summary of the timeline of IR models, TF-IDF-based ranking functions, the probability of relevance framework, binary independency retrieval, relevance feedback, and language modelling have been addressed in particular. One of the main takeaway messages of the lecture was to be careful with the selection of IR models and ranking functions for an application by inspecting their properties thoroughly.

### **Interactive IR (Norbert Fuhr, University of Duisburg-Essen)**

Norbert Fuhr introduced the audience to interactive IR, a field that combines Human-Computer Interaction and IR. One main topic of the lecture was the Interactive Probability Ranking Principle (IPRP). The IPRP widens the scope of traditional probability-based ranking systems by adding user-specific parameters like activity costs, changes in information need, and search phases to the formula. The second part of the lecture addressed information seeking behavior, information searching, and strategic measures to support users with these tasks.

### **Task-Based Evaluation (David Elsweiler, University of Regensburg)**

The evaluation of information retrieval systems was one of the main topics of ASIRF 2019, which is why both David Elsweiler and Thomas Mandl addressed this subject in their lectures. The lecture *Task-Based Evaluation* by David Elsweiler focused on the components,

---

measurements, and goals of an evaluation as well as a discussion about empirical studies. In the latter part, different types of errors that typically occur during testing were presented in addition to various types of experimental designs and their applicability for IR studies.

### **Semantic Annotations for IR (Klaus Berberich, University of Applied Sciences Saarbrücken)**

Natural language processing for the analysis of text was the central topic of Klaus Berberich's lecture on semantic annotations. Additional information on textual resources were introduced such as named entities, temporal expressions, and geographic references. The collection of large quantities of contextual information can be organized using knowledge graphs to understand the relationship between entities and enable structured queries on these concepts. A specific application of semantic annotations was displayed in the form of ad-hoc retrieval. The lecture concluded with a presentation of tools and data sets that can be used for this kind of IR.

### **IR for the Digital Humanities (Andreas Henrich, University of Bamberg)**

Digital Humanities are a large and comparatively young field of research, combining humanities with new approaches from computer science. The lecture *IR for the Digital Humanities* by Andreas Henrich addressed the possibilities and necessity of information retrieval techniques and concepts in modern humanities, such as the use of federated collections in the DARIAH-DE project. Further topics were a search system for finding appropriate paper-journal combinations and neural network based emblem recognition with automatic annotation of segments for image retrieval.

### **IR Evaluation (Thomas Mandl, University of Hildesheim)**

Thomas Mandl elaborated on evaluation in IR by explaining its significance and describing detailed and objective standards for the comparison of IR systems. A critical discussion on relevance and the distinction of subjective and objective relevance as well as bias and subjectivity were the principal topics in this lecture. Some metrics for IR methods were shown, while also talking about the pros and cons for the respective use case. The final part of the talk addressed bias in user studies, user satisfaction, and group design.

### **Deep Learning for IR (Ahmet Aker, University of Sheffield)**

*Deep Learning for IR* was a tutorial for various deep learning concepts with focus on neural networks. The topics of the lecture part addressed the basic concept of a neuron, binary classifiers, word embeddings, recurrent neural networks, and convolutional neural networks. Using Google Classroom and Google Colab, the practical part of the course comprised programming exercises with Python and the open-source neural network library Keras. In particular, the use of Jupyter Notebooks contributed significantly to the easy and intuitive access to the rather complex subject deep learning.

### **IR Systems (Ralf Schenkel, University of Trier)**

The final lecture of ASIRF 2019 by Ralf Schenkel addressed aspects regarding the design, implementation, and evaluation of IR systems. One central topic of the lecture were the

---

three classes of algorithms for query processing: term-at-a-time, document-at-a-time, and score-at-a-time. Afterwards, data sets for evaluating IR models and evaluation approaches like crowdsourcing and evaluation-as-a-service were discussed. The final lecture closed with a short summary of the popular open-source IR framework Apache Lucene and its derivations Apache Solr and Elasticsearch.

### 3 Student Presentations and Case Study

The participants were invited to present their own work in 5-minute-talks during two slots in the varied schedule of the autumn school. For this purpose, the organizers asked the participants some weeks in advance to prepare a few slides, covering one's background and field of interest or thesis research. The short presentations did not only speed up the process of getting to know each other but also facilitated the exchange of information and expertise.

Furthermore, the participants were asked to take part in a case study project that occupied various time slots throughout the week. In the first session, the participants were divided into small groups of four to six to work on a provided or self-developed research question related to IR. The given research questions covered some classic but ever-present and vital tasks like designing a ranking function, supporting information seeking processes, evaluating IR systems, and improving the user interface of an IR system. By working on these tasks in the diverse groups we were able to get to know different ways of working in other cultures.

### 4 Social Program

The autumn school offered many opportunities to interact with fellow participants as well as lecturers. As a warm-up, the participants received an introduction to the German university and research system by Ralf Schenkel on the first day, which was especially interesting for the international students. The seating arrangements during lunch and dinner were mixed and shuffled randomly with the objective of facilitating new conversations between those participants who did not have the chance to talk to each other beforehand. Every evening after the last lecture, further gatherings were held in the wine cellar of Schloss Dagstuhl, accompanied by a delicious cheese platter and drinks.

The athletically inclined participants and lecturers were also offered a huge leisure program inside and around Schloss Dagstuhl such as volleyball on a court right beside the lecture hall, table tennis, darts, and a sauna in the basement. The steep but short hike to the top of the small mountain was rewarded with a picturesque view onto the natural preserve Saar-Hunsrück as well as the ruins of the old castle built in medieval times.

Fig. 2 gives an impression of the excursion to Trier on Wednesday afternoon. It provided a nice opportunity to explore the historic city of Trier with its many churches and famous Roman monuments like the Porta Nigra. After a guided tour through the old city center of Trier, the participants visited a traditional wine cellar and took part in a wine tasting on the outskirts of the city at the Winery von Nell.



Figure 2: Market square in Trier with our guide and some participants

## 5 Conclusion

Throughout the six days of ASIRF 2019, 21 participants from all over the world took part in intercultural exchange, recreational activities, and most importantly advanced their knowledge in IR concepts, algorithms, models, and much more. The eight lecturers provided us a new understanding of IR topics by sharing their expert knowledge. In addition, they gave us valuable advice on many general topics like academic working, proper research methods, and being a scientist over the meals or while enjoying a drink in the wine cellar of Schloss Dagstuhl. Especially the casual atmosphere inside and outside the lectures made the event enjoyable and memorable.

## 6 Acknowledgements

Special thanks go to the sponsors of ASIRF 2019: University of Trier, the German Academic Exchange Service (DAAD), ACM's Special Interest Group on Information Retrieval (SIGIR), and Schloss Dagstuhl. Furthermore, we want to thank the organizers Ralf Schenkel, Ingo Frommholz, Norbert Fuhr and all other lecturers for making this a successful event.