

## Autumn School for Information Retrieval and Foraging 2018

**Amelia W. Cole**

University of British Columbia  
*miaco@mail.ubc.ca*

**Linda Achilles**

University of Hildesheim  
*achilles@uni-hildesheim.de*

### Abstract

The Autumn School for Information Retrieval and Information Foraging (ASIRF) is a five-day intensive seminar held in mid-September at [Schloss Dagstuhl](#) in Saarland, Germany. From September 16 – 21 in 2018, 20 people from five different countries engaged in a series of lectures, tutorials, and project-based work with internationally recognized senior scholars who specialize in information retrieval and behaviour.

## 1. Introduction

The Autumn School for Information Retrieval and Foraging (ASIRF) is a week-long scientific, annual autumn school focused on the interaction between Information Retrieval (IR) and Information Seeking Behaviour (IB) and features lectures and talks from internationally recognized academics who specialize in the field of Information Science. In 2018, participants experienced nine lectures that covered the foundations of Information Retrieval and Information Foraging as well as in-depth tutorials on collaborative IR, searching and mining argumentation structure, human-centered research design methods, and practicums in machine and deep learning. ASIRF is usually held the second week of September in an 18<sup>th</sup> century castle, [Schloss Dagstuhl](#), in Saarland Germany that boasts all-inclusive meals, a remote location, and deluxe amenities to provide researchers with the optimal environment to exchange knowledge and meet new potential collaborators.

## 2. Lectures

This year's [ASIRF](#) lectures covered topics ranging from Information Retrieval fundamentals to Machine Learning. This section provides an overview of the lectures.

### **Bibliographic Database (Keynote).** *Michael Ley, University of Trier*

The keynote opened the autumn school series with a lecture on the structure, innerworkings, and challenges of [dblp](#), an open bibliographic system that currently hosts over 4 million computer science publications. The main takeaways of this lecture are that system designers must contend with difficult and abstract concepts like relevance and the practical issues of resolving misspelled names and same-named conferences and authors in addition to ensuring that all publications linked within the system meet a high-quality standard.

---

**Foundations of Information Retrieval.** *Ingo Frommholz, University of Bedfordshire*

This lecture covered the history and foundations of Information Retrieval models. Participants learned the notation system to describe the events, sets, and frequencies as well as the differences between popular heuristic models (e.g., TF-IDF) and probability-based relevance models (e.g., Probability Ranking Framework and Probability Ranking Principle). The main struggle for many IR systems is relevance, which is a subjective concept that prioritizes the order of search results based on the search query. While relevance may seem like a simple concept, it can be particularly challenging when searches are dynamic and there are several changes to the information need within the same session.

**Information Retrieval Systems.** *Ralf Schenkel, University of Trier*

This three-part lecture series detailed how to build your own information retrieval system. The series covered a range of query models from simple Boolean queries to complex processing algorithms as well as how to connect to and use existing open-source IR systems like Terrier, Indri, Apache Lucene, and more.

**Interactive Information Retrieval.** *Norbert Fuhr, University of Duisburg-Essen*

This lecture straddled the mathematical principles of relevance ranking as well as information seeking strategies and cognitive models of search. More specifically, Fuhr discussed the Interactive Probability Ranking Principles (IPRP) in the context of IIR, reporting that as people move between different situations their information needs may change, and therefore there may be different costs for different activities, which influence the order of results.

**Methodology of Interactive Information Retrieval.** *Pia Borlund, Oslo Metropolitan University*

Interactive Information Retrieval is a user-centred perspective that depends on analysing human behaviour, cognition, and affect that are performed in realistic settings using realistic work tasks. The main point of the lecture was how to write relevant work tasks for research studies that will enhance the future of IR systems. Work tasks are exceptionally complex to create because not all work tasks suit all participants. For example, giving 18-year old undergraduates a critical incident in which they need to imagine themselves researching a month-long trip to Bora Bora with their eight-year old son will provide you with unusable data as very few 18-year old undergraduates will relate to the scenario.

**Information Behaviour and User Interfaces.** *Christa Womser-Hacker, University of Hildesheim*

This lecture defined the differences between information seeking and information searching. Information seeking is when a person senses a problem, identifies a gap in their knowledge, and seeks information to resolve that gap. Information searching is the interaction between the person and the information (e.g., SERPs). There are several information seeking, sense-making, and cognitive models that can be leveraged to understand the process through which a person (or group of people) engage in information seeking and search behaviours.

**Information Retrieval for the Digital Humanities.** *Andreas Henrich, University of Bamberg*

---

This lecture combined the fields of Information Retrieval with the Digital Humanities by unpacking the use of federated databases, using [DARIAH-DE](#) as the core example. The lecture also covered the challenges industry organizations (e.g., Accenture) face with building and maintaining internal search systems that support employees in finding the information they need when they need it.

**Argumentation Retrieval and Analysis.** *Henning Wachsmuth, University of Paderborn*

This lecture unpacked computational argumentation retrieval and analysis, showcasing the exemplary work, [args.me](#). The field of computational argumentation seeks to automatically identify the premises of arguments to enhance our decision-making process; this research may eventually influence the future of how we analyse and predict political outcomes, speak with intelligent agents, design decision-making tools, and support people with their writing.

**Fake News Detection.** *Ahmed Aker, University of Duisburg-Essen*

This lecture focused on how we are using machines to combat the presence of fake news online. First, we used human assessment practices to determine the legitimacy of a news article, and then moved to the effectiveness of machine learning algorithms in the evaluation of news articles using stance detection and rumour verification. Learners experienced a hands-on workshop creating their own deep learning algorithms using a custom practice file.

**Collaborative Information Retrieval.** *Stefanie Elbeshausen, University of Hildesheim*

This lecture challenged the preconceived notions that all information retrieval is performed by an individual. Many information tasks are done as a group either in a co-located or remote scenario. One of the primary takeaways was the definition of collaboration. Collaboration is a very specific activity in which all participants share the same goal and perceive the meaning of the goal in similar ways. Once meaning and goals become different for participants, it is no longer considered collaboration.

### 3. Participant Presentations and Projects

The autumn school included engagement from the participants through presentations and group case studies throughout the week. Participants were asked in advance to create a five-minute presentation on their work-in-progress thesis topic, area of interest, or recent research study to invite conversation between people who share a common interest. The participant topics presented ranged from search as learning to computational argumentation; showcasing the true interdisciplinary nature of ASIRF. Mid-week, teams were formed to unpack a predefined case study using a bibliographic database as the only restriction. The participants could either create a ranking system or design a user interface. The teams set out on their own in the evenings to work independently on their projects, which usually included a glass of wine, beer, or soda pop.

### 4. Social Programme

The social programme for ASIRF was interwoven into the day's schedule. At each main meal, the wonderful Schloss Dagstuhl staff randomly shuffled the name tags around the tables, inviting participants to get to know each other at the event. And every evening at 8pm, the same staff produced a delightful selection of wonderful cheeses and fruits in the wine cellar for everyone to

---

---

enjoy. Mid-week all participants were rewarded with an afternoon break. We climbed into a luxurious bus that carried us to the oldest city in Germany, Trier, whose history dates back to the Romans. We all experienced a walking tour of the city; wandering through its monuments, churches, and public places. We even had the incredible luxury of participating in a winery tour and tasting at the “von Nell” winery, where the group enjoyed regional Riesling wine and a delightful German barbecue in the warm evening weather. A week of great lectures and activities created new insights, connections, and opportunities that enhance our research in the field of information science.



**Figure 1 - St. Peter's Cathedral Cloister, Trier, Germany**



**Figure 2 – A portion of the ASIRF 2018 participants in front of the ruins of Dagstuhl Castle.**

## **5. Conclusion**

The Autumn School for Information Retrieval and Foraging brought many people from around the world to participate in an intensive series of lectures, tutorials, social events, and project work. This series provided learners with a unique opportunity to hear from internationally recognized experts in the field of Information Science and understand the perspectives that are influencing the future direction of how we think about our interaction with information. And of course, the social gatherings allowed for everyone to enjoy the majesty of living inside an 18<sup>th</sup> century castle, broaden their academic network, and feel inspired by the breadth and depth of research being conducted in the field.

## **6. Acknowledgements**

A special thank you to the sponsors for making ASIRF 2018 possible: ACM's Special Interest Group on Information Retrieval (SIGIR), Gesellschaft für Informatik (GI), Schloss Dagstuhl, Centre for Informatics Research and Technology (CIRT), and the University of Trier. And a note of gratitude to this year's organizers, Ralf Schenkel, Ingo Frommholz, Norbert Fuhr, and Tobias Zeimetz who all worked incredibly hard to make ASIRF 2018 an inclusive and successful event.