

The 8th Russian Summer School in Information Retrieval (RuSSIR 2014)

Biljana Lazić

biljana.lazic@rgf.bg.ac.rs

University of Belgrade,
Faculty of Mining and Geology

Gordana Nedeljkov

gordana.nedeljkov06@gmail.com

Public Library
"Đura Jakšić", Srpska Crnja

DATE OF SUBMISSION:

12 September 2014

DATE OF ACCEPTANCE:

14 September 2014

1 Introduction

Russian Summer School in Information Retrieval has been held for eight years in a row. It is customary that every year different town in Russia is a host. This year The 8th Russian Summer School in Information Retrieval was held on August 18-22, 2014 in Nizhny Novgorod. This year's organizers were National Research University Higher School of Economics and the Russian Information Retrieval Evaluation Seminar. The mission of the school is to enable participants to learn about problems and methods in information retrieval and related disciplines, to support scientific research and collaboration in this field, and also to help with making contacts between scientists, students and professionals in information retrieval area. The school is aimed

at advanced graduate and PhD students, scholars and researchers. This year main theme of the school was visualization of information retrieval.

Professor Cvetana Krstev informed us about school and advised us to apply. Participation in the school was free of charge, and thanks to the sponsor food costs were covered. Also, one could apply for scholarship which included accommodation costs. For application it was necessary to submit CV, brief poster paper that would be presented during the school, motivation for participating and a recommendation letter. Professor Krstev provided a recommendation letter. Two weeks after applying we got a notification of acceptance and we were informed that we were expected in Nizhny Novgorod.

2 School program

This year's school was attended by about 100 participants. Some of them had already participated in the school earlier. Most of participants were from Russia and about 15 students were from abroad (Great Britain, Italy, Romania,

Bangladesh, Serbia and Ukraine). School program was very intensive. It consisted of courses, visiting professors' lectures, Young Scientists Conference and several social events organized after classes.



FIGURE 1. RuSSIR 2014 Participants

Following courses were held:

- *Online Experimentation for Information Retrieval* (lecturer Katja Hofmann, Microsoft Research, Cambridge, England)
- *Web as a Corpus: Going Beyond the n-gram* (lecturer Preslav Nakov, Qatar Computing Research Institute, Doha, Qatar)
- *Visualization & Data Mining for High Dimensional Data* (lecturers Alfred Inselberg, Computer Science and Applied Mathematics Department, Tel Aviv University, Israel, and Pei Ling Lai, Southern Taiwan University of Science and Technology, Tainan, Taiwan)
- *Introduction to Formal Concept Analysis and Its Applications in Information Retrieval and Related Fields* (lecturer Dmitry Ignatov, School of Applied Mathematics and Information Science, Higher School of Economics, Moscow, Russia)
- *Document Analysis and Retrieval in Scientific Digital Libraries: Case studies in applying Machine Learning for Information Retrieval* (lecturer Sujatha Das Gollapalli, Department of Computer Science, College of Engineering, Pennsylvania, USA)
- *Author Profiling and Plagiarism Detection* (lecturer Paolo Rosso, Natural Language Engineering Lab, Technical University of Valencia, Spain).

It was possible to attend three of offered courses, because two courses were held in the same time. All courses that we attended consisted of theoretical lectures and practical examples. The

first course, *Web as a Corpus: Going Beyond the n-gram*, was about computational linguistics and natural language processing, using Web as a corpus. Special emphasis was put on Web-based approaches that go beyond the n-gram. Course *Introduction to Formal Concept Analysis and Its Applications in Information Retrieval and Related Fields* dealt with the-

oretical and practical aspects of mathematical theory and technique that is used in information retrieval, called Formal Concept Analysis.

Course *Document Analysis and Retrieval in Scientific Digital Libraries: Case studies in applying Machine Learning for Information Retrieval* was the closest to our field. During the course we discussed techniques of machine learning. Digital libraries or repositories that contain diverse digital objects – documents, videos and pictures, were used as representative sample. Scientific Digital Library CiteSeer, a digital library portal for scientific documents in Computer Science and related areas, was used as an example. Topics such as web crawling, document classification, content analysis, metadata and page ranking algorithms were also mentioned during the course.

In addition to courses that were held every day, we had an opportunity to attend lectures given by guest speakers coming from eminent institutions. School was opened by skype-lecture by Marti Hearst named "Seeking Simplicity in Search User Interfaces". Professor Hearst spoke about user interfaces and their role in everyday Internet usage, such as electronic resources research and online learning, giving positive examples of using interfaces with simple design. Conclusion of this lecture was that it is important to take care of users' needs in interface design.

On the second day of school professor Marcel Worring gave an interesting lecture "Multimedia Analysis and Multimedia Visualization". He spoke about frequent usage of various multimedia information carriers. Furthermore, he presented semantic detectors with ability to identify motives in the picture that is part of large image collection.

Ludmila Ostroumova from Yandex gave the

lecture “Crawling of New Web Pages”. She spoke of growing number of web pages and about significance of fast crawling and indexing of those pages carried out by search engines. Predicting page’s popularity using characteristics of URL address solely on the time of their discovery was

also discussed in this lecture.

Dmitry Solovyov from the Mail.ru gave lecture on the penultimate day of the school. He talked about significance of data analysis for development of self-organizing maps for search engines and about their implementation.

3 Young Scientist Conference

On the first and the second day of school Young Scientist Conference was held. As organizers noted, the aim of the conference is to connect young scientists in the field of mathematics and computer science, as well as in the field of social and media sciences. Six candidates whose papers passed selection took part in the conference with five-minute presentations of their work, with their papers scheduled to be published in conference proceedings.

When applying for school, all candidates were obliged to submit brief description of their scientific research or project in which they had participated and which would be presented during the poster session. Poster session was conducted in an informal atmosphere with collegiate exchange of ideas and experiences.



FIGURE 2: Poster session – poster by Biljana Lazić

Biljana Lazić participated with the poster “Digital Library of Scientific Electronic Journals”. She presented method of work and functions of digital library based on TMX documents and browser that provides morphological query expansion.

Gordana Nedeljkov’s poster was titled “Libraries and education: from creative workshop to lifelong learning”. The topic was the role of libraries in supporting formal and informal education, with examples of implementing theory in practice.



FIGURE 2: Poster session – poster by Gordana Nedeljkov

4 Extracurricular activities

School organizers did their best to organize various interesting activities after school time. The event scheduled for the first night was the welcoming party where we had an opportunity to familiarize with colleagues from the school. Guided tour to Nizhny Novgorod Kremlin during

which we learnt about city's history and some of its sights was organized on the second day. After the third day's lectures participants went

to organized sport evening. On the penultimate day of our stay we enjoyed the sunset during ship cruise on rivers Oka and Volga.

3 Conclusion

Participation in the 8th Russian Summer School in Information Retrieval, visiting Russia and staying in Nizhny Novgorod, was a great experience for us. Attending summer schools such

as the RuSSIR is ideal as a form of professional development. In addition, it is a great opportunity to get to know different countries and cultures and to make new acquaintances.

Acknowledgements

Biljana Lazić is grateful to the management of the Faculty of Mining and Geology for

understanding and generous financial support for attending this school.

References

The 8th Russian Summer School in Information Retrieval (RuSSIR 2014), accessed 13.09.2014, <http://romip.ru/russir2014/>