

ПРОГРАММА
Всероссийской молодёжной конференции
с международным участием
«10 РОССИЙСКАЯ ЛЕТНЯЯ ШКОЛА-
КОНФЕРЕНЦИЯ ПО ИНФОРМАЦИОННОМУ
ПОИСКУ»

RuSSIR 2016

22 — 26 августа 2016 г.

Саратов, 2016

ПРИГЛАШЕНИЕ

Уважаемые коллеги!

**Организационный комитет приглашает Вас принять участие во
Всероссийской молодёжной конференции с международным участием
«10 российская летняя школа-конференция по информационному поиску
“RUSSIR-2016”»**

Школа-конференция будет проходить с 22 по 26 августа 2016 г. в ФГБОУ ВО «Саратовский национальный исследовательский государственный университет имени Н. Г. Чернышевского» по адресу: 410012, г. Саратов, ул. Астраханская, 83.

Информацию о конференции можно получить на сайте <http://romip.ru/russir2016/>

Работа школы-конференции проводится в рамках научных направлений

1. Общие вопросы теории информационного поиска
2. Структуры данных и алгоритмы индексирования
3. Системы веб-поиска и веб-майнинга
4. Семантический веб и базы знаний
5. Анализ социальных сетей и соцграфа.

ГРАФИК МЕРОПРИЯТИЙ

Всероссийской молодёжной конференции с международным участием

«10 РОССИЙСКАЯ ЛЕТНЯЯ ШКОЛА-КОНФЕРЕНЦИЯ ПО ИНФОРМАЦИОННОМУ ПОИСКУ “RUSSIR-2016”»

Работа по секциям проводится:

1. В актовом зале 10 корпуса СГУ по адресу ул. Астраханская, 83
2. В 208 аудитории 11 корпуса СГУ по адресу ул. Астраханская, 83

Выступления в рамках конференции молодых учёных проводятся в актовом зале 10 корпуса СГУ по адресу ул. Астраханская, 83

Стендовая секция конференции молодых учёных проводится в атриуме 11 корпуса СГУ по адресу ул. Астраханская, 83

Кофе-брейки проводятся в атриуме 11 корпуса СГУ по адресу ул. Астраханская, 83

20 – 22 августа – Заезд участников

22 августа – Первый официальный день конференции

8:30 – 9:30 – Регистрация (ул. Астраханская, 83, 10 корпус СГУ, 2 этаж)

9:30 – 11:00 – Пленарное заседание, начинающееся с приветствия участников руководством СГУ и программным комитетом (ул. Астраханская, 83, 10 корпус СГУ, актовый зал)

11:00 – 11:30 – Кофе-брейк

11:30 – 13:00 – Работа по секциям, работа круглых столов

13:00 – 14:00 – Обед

14:00 – 15:30 – Работа по секциям, работа круглых столов

15:30 – 16:00 – Кофе-брейк

16:00 – 17:30 – Работа по секциям, работа круглых столов

17:30 – 18:00 – Кофе-брейк

18:00 – 19:30 – Конференция молодых учёных

19:30 – ... – Внепрограммные мероприятия: Фуршет

23 августа

9:30 – 11:00 – Работа по секциям, работа круглых столов

11:00 – 11:30 – Кофе-брейк

11:30 – 13:00 – Работа по секциям, работа круглых столов

13:00 – 14:00 – Обед

14:00 – 15:30 – Работа по секциям, работа круглых столов

15:30 – 16:00 – Кофе-брейк

16:00 – 17:30 – Работа по секциям, работа круглых столов

17:30 – 18:00 – Кофе-брейк

18:00 – 19:30 – Конференция молодых учёных

24 августа

- 9:30 – 11:00 – Работа по секциям, работа круглых столов
- 11:00 – 11:30 – Кофе-брейк
- 11:30 – 13:00 – Работа по секциям, работа круглых столов
- 13:00 – 14:00 – Обед
- 14:00 – 15:30 – Работа по секциям, работа круглых столов
- 15:30 – 16:00 – Кофе-брейк
- 16:00 – 17:30 – Работа по секциям, работа круглых столов
- 17:30 – 18:00 – Кофе-брейк
- 18:00 – 19:30 – Спонсорская лекция

25 августа

- 9:30 – 11:00 – Работа по секциям, работа круглых столов
- 11:00 – 11:30 – Кофе-брейк
- 11:30 – 13:00 – Работа по секциям, работа круглых столов
- 13:00 – 14:00 – Обед
- 14:00 – 15:30 – Работа по секциям, работа круглых столов
- 15:30 – 16:00 – Кофе-брейк
- 16:00 – 17:30 – Работа по секциям, работа круглых столов
- 17:30 – 18:00 – Кофе-брейк
- 18:00 – ... – Внепрограммные мероприятия: традиционный банкет «RuSSIR party»

26 августа

- 11:30 – 13:00 – Работа по секциям, работа круглых столов
- 13:00 – 14:00 – Обед
- 14:00 – 15:30 – Работа по секциям, работа круглых столов
- 15:30 – 16:00 – Кофе-брейк
- 16:00 – 17:30 – Работа по секциям, работа круглых столов
- 17:30 – 18:00 – Кофе-брейк
- 18:00 – 19:30 – Закрытие конференции

26 – 27 августа – Отъезд участников

Расписание

	22 августа, Понедельник		23 августа, Вторник		24 августа, Среда		25 августа, Четверг		26 августа, Пятница							
9:30 – 11:00	Приветствие ректора и президента СГУ (10 к., актовый зал)		Демартини (10 к., акт. зал)		Люпу (11 к., 208)		Бэлог (10 к., акт. зал)		Миркин (11 к., 208)		Люпу (11 к., 208)		Марков, Чуклин, Де Райке (10 к, акт. зал)			
11:30 – 13:00	Суханек (11 к., 208)	Вебер (10 к., акт. зал)	Джи (10 к., акт. зал)	Вебер (11 к., 208)	Котов (11 к., 208)	Вебер (10 к., акт. зал)	Котов (11 к., 208)	Вебер (10 к., акт. зал)	Люпу (10 к., акт. зал)	Берковски (11 к., 208)						
14:00 – 15:30	Суханек (11 к., 208)	Марков, Чуклин, Де Райке (10 к., акт. зал)	Джи (10 к., акт. зал)	Марков, Чуклин, Де Райке (11 к., 208)	Люпу (11 к., 208)	Марков, Чуклин, Де Райке (10 к., акт. зал)	Миркин (10 к., акт. зал)	Берковски (11 к., 208)	Демартини (10 к., акт. зал)	Берковски (11 к., 208)						
16:00 – 17:30	Демартини (11 к., 208)	Вебер (10 к., акт. зал)	Бэлог (10 к., акт. зал)	Марков, Чуклин, Де Райке (11 к., 208)	Демартини (10 к., акт. зал)	Берковски (11 к., 208)	Демартини (10 к., акт. зал)	Берковски (11 к., 208)	Люпу (10 к., акт. зал)							
18:00 – 19:30	Конференция молодых учёных (11 к., атриум и 208)		Конференция молодых учёных (11 к., атриум и 208)		Спонсорская лекция: Георгий Иванов, Mail.RU (10 к., актовый зал)		Традиционное мероприятие «RuSSIR party»									
	Фуршет по поводу открытия															

ΠΡΟΓΡΑΜΜΑ ΚΥΡΣΟΒ

Ontological Information Extraction

Instructor: Dr. Fabian Suchanek (Associate Professor, Telecom ParisTech University, France)

Course description: In this course, we will take a detailed overview of Information Extraction techniques. Information extraction is the process of deriving structured information (such as alive (Elvis)) from digital text (such as the sentence “Elvis is alive”). We will focus on factual and semantic information extraction, i.e., we will cover named entity recognition, entity disambiguation, instance extraction, fact extraction, and ontological information extraction. We will also touch upon applications of Information Extraction, such as Google’s knowledge graph/vault, and IBM’s Watson question answering system, as well as academic projects such as YAGO, DBpedia, and NELL. In particular, we will cover the following topics: Knowledge Representation on the Semantic Web, Named Entity Recognition, Named Entity Annotation, Instance Extraction and Information Extraction from Unstructured Sources.

Computational Social Science: Theories, Methods and Data

Instructor: Dr. Ingmar Weber (Senior Scientist, Qatar Computing Research Institute, Qatar)

This course is supported by ACM’s Distinguished Speaker Program.

Course description: Due to the increasing availability of large-scale data on human behavior collected on the social web, as well as advances in analyzing larger and larger data sets, interest in applying computer science methods to address research questions in the social sciences continues to grow. “Big Data” researchers and “Data Scientists” entering the interdisciplinary field of Computational Social Science (CSS) often lack background in theories and methods in sociology, whereas sociologists are often not aware of data collection and analysis techniques in computer science. This tutorial helps to bridge this gap by providing an introduction (i) to social theories and models that help to understand the process that generated the data, as well as (ii) to statistical and computational methods that are useful for addressing social science research questions with observational data as found on the Web. The goal of this tutorial is to give participants a rich repertoire of methods that help to answer not only interesting “how” questions but also more fundamental “why” questions. This includes both a general knowledge of important sociological theories and how to derive verifiable hypotheses from them, as well as a good grasp of methods for causal inference from observational data to go beyond mere correlations. To maximize the learning outcome, there will be a set of short practical examples provided in the form of IPython notebooks. Furthermore, there will be a small competition where students are invited to submit their own research proposals. This course is an extended version of a tutorial given at WWW’16 together with Markus Strohmaier, Claudia Wagner, and Luca Aiello.

Click Models for Web Search

Instructors: Dr. Ilya Markov (Postdoctoral Researcher, University of Amsterdam, Netherlands), Aleksandr Chuklin (Software Engineer, Google, Switzerland; Researcher, University of Amsterdam, Netherlands), Dr. Maarten de Rijke (Full Professor, University of Amsterdam, Netherlands)

Course description: Click models, probabilistic models of the interaction behavior of search engine users, have been studied extensively by the information retrieval community in recent years. We now have a handful of click models, parameter estimation methods, evaluation principles and applications of click models, that form the building blocks of ongoing research efforts in the area. Moreover, click models appear in many areas of IR, such as ranking, evaluation, user simulation, etc. This course is based on a recently published book on click models (<http://clickmodels.weebly.com/the-book.html>) and covers a wide range of topics: from basic to advanced and neural click models and from click model estimation and evaluation techniques to applications of click models. Most topics are augmented with live demos, where the participants can try the presented material in practice. Also, the course features two practical sessions, where the participants have a chance to implement a basic and an advanced click models using open-source tools and publicly available datasets with click logs. The course will be useful as an overview for anyone starting research work in IR as well as for practitioners seeking concrete recipes. More details about the course are available at <http://clickmodels.weebly.com/russir-2016-course.html>. The setup for following live demos and practical sessions is outlined in <http://clickmodels.weebly.com/russir-2016-setup.html>.

Crowdsourcing and Human Computation for Semantic Search

Instructor: Dr. Gianluca Demartini (Senior Lecturer, University of Sheffield, UK)

Course description: The availability of high-quality large-scale datasets like, for example, knowledge graphs is key to enable effective semantic search solutions. In this course we will present crowdsourcing and human computation approaches that are used to build better semantic search systems. We will introduce students to the human computation domain using successful examples and popular crowdsourcing platforms. We will then present hybrid human-machine systems that combine the scalability of machine-based data processing with the ability of human intelligence to interpret content and data semantics. Next, we will discuss in detail quality and efficiency challenges of such systems. Finally, we will present current research trends in the area of crowdsourcing and human computation relevant to semantic search and knowledge graphs.

Knowledge Base Population

Instructor: Dr. Heng Ji (Associate Professor, Rensselaer Polytechnic Institute, USA)

Course description: This course will introduce state-of-the-art Information Extraction (IE) and Knowledge Base Population (KBP) techniques. In the first lecture, we will focus on the quality issue of IE and KBP. We will give a comprehensive overview about the successful methods for each task. We will review where we have been (the most successful methods in literature), and where we are going (the remaining challenges, and novel methods to tackle these challenges). In the second lecture, we will focus on the portability issue, namely how to rapidly build a new IE/KBP system for a new language, domain, genre within a short time with low cost. We will introduce a brand new “Liberal” Information Extraction (IE) paradigm to combine the merits of traditional IE (high quality and fine granularity) and Open IE (high scalability). Liberal IE aims to discover schemas and extract facts from any input corpus, without any annotated training data or predefined schema.

Domain Specific Semantic Search

Instructor: Dr. Mihai Lupu (Postdoctoral Researcher, Vienna University of Technology, Austria)

Course description: Domain-specific search engines only index documents relevant to a specific domain, such as health information or intellectual property information. As prior knowledge is available about the domain of interest, such search engines can be adapted to take advantage of this knowledge for improving search results. Furthermore, the users of these search engines often have specific requirements as to how the search engines should function. This course begins with a general introduction to domain-specific search and the methods used, such as adapting vector space weighting and using specialized vocabularies. This is followed by more detailed coverage of two domains. In web search for health information, the quality and trustworthiness of the information is extremely important, as is the readability level will this information be understood by a layperson or only by a medical professional? In the intellectual property area, professional patent searchers are faced with the task of finding all patents related to a query (high recall), even though patent documents are sometimes not written to be found easily. Horizontally to these and other domains, the issue of credibility in Information Retrieval will also be addressed in detail. Finally, after an overview of evaluation, evaluation campaigns and their results in the area of medical information search and intellectual property search will be discussed.

Entity Linking

Instructor: Dr. Krisztian Balog (Associate Professor, University of Stavanger, Norway)

Course description: Recognizing and disambiguating entities is a key step towards understanding the meaning of a text. Entity linking refers to the task of identifying entity mentions in text and linking them to the corresponding entries in a reference knowledge base. Entity annotations allow readers of a document to acquire contextual or background information with a single click. They can also be used in downstream processing to improve retrieval performance or to facilitate better user interaction with documents or search results, e.g., by providing easy access to related entities. Most entity linking approaches can be seen as a pipeline of three main steps: (i) mention detection, to determine “linkable” phrases, (ii) link generation, to rank and/or selects candidate entities for each mention, and (iii) disambiguation, to select, with the help of context, a single entity (or none) for each mention. The course consists of two lectures. The first lecture introduces theory and methods. The second lecture provides practical guidance with hands-on examples and exercises.

Knowledge Graph Entity Representation and Retrieval

Instructor: Dr. Alexander Kotov (Assistant Professor, Wayne State University, USA)

Course description: Recent studies indicate that more than 75% of queries issued to Web search engines aim at finding information about entities, which could be material objects or concepts that exist in the real world or fiction (e.g. people, organizations, locations, products, etc.). Most common information needs underlying this type of queries include finding a certain entity (e.g. “Einstein relativity theory”), a particular attribute or property of an entity (e.g. “Who founded Intel?”) or a list of entities satisfying a certain criteria (e.g. “Formula 1 drivers that won the Monaco Grand Prix”). These information needs can be efficiently addressed by presenting structured information about the target entity or a list of entities retrieved for these queries from a knowledge graph either directly as search results or in addition to the ranked list of documents. This course provides a summary of the latest research in knowledge graph entity representation methods and retrieval models.

In the first part of this course, I will introduce different methods for entity representation: from multi-fielded documents with flat and hierarchical structure to latent dimensional representations based on tensor factorization. In the second part of this course, I will discuss recent developments in entity retrieval models, including Mixture of Language Models (MLM), Probabilistic Retrieval Model for Semi-structured Data (PRMS), Fielded Sequential Dependence Model (FSDM) and its parametric extension (PFSDM) as well as learning-to-rank methods.

Model- and experiment-driven recommendations for haunting issues in clustering

Instructor: Prof. Dr. Boris Mirkin (Professor, National Research University Higher School of Economics, Russia & Professor (Emeritus), Birkbeck, University of London, UK)

Course description: Clustering is a well-established area of data and text analysis, oriented at segmentation of the phenomenon in question into homogeneous fragments. Yet there are a number of haunting issues and concerns in clustering, of which probably the most imperative are: (a) Is there any cluster structure at all? (b) If yes, how many clusters? (c) What object-to-object similarity measure to choose? (d) Which features are useful for clustering, and which are not? (e) Can a mixed, categorical and numerical, feature space be used for clustering? (f) How can one reconcile clustering solutions if they do differ from each other? The talk brings in some novel and not-so-novel developments in clustering, that are useful in addressing these and similar concerns. Specifically, an SVD-like modeling will be described to underlie approaches such as k-means clustering, Ward divisive clustering, one-cluster clustering, consensus clustering, spectral clustering, network community detection. The first lecture will concentrate on the case of object-to-feature data, and the second, on object-to-object (dis)similarity data. A number of examples of application will be presented as well.

Social Personalization and Recommender Systems

Instructor: Dr. Shlomo Berkovsky (Senior Research Scientist, The Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia)

Course description: The quantity of accessible information has been growing rapidly and far exceeded human processing capabilities. The sheer abundance of information often prevents users from discovering the desired information or aggravates making informed choices. This highlights the pressing need for intelligent personalized applications that simplify information access and discovery and provide adaptive services, which take into account the preferences and needs of their users. One type of personalized application that has recently become tremendously popular in research and industry is recommender systems. These provide to users personalized recommendations about information and products they may be interested to examine or purchase. This is often achieved by exploiting social methods, which amalgamate past experiences of other users in order to identify most valuable information and products. Extensive research into recommender systems over the last decade has yielded a wide variety of techniques, which have been published at a range of reputable venues and subsequently adopted by numerous Web-sites and services. This course will provide the participants with broad overview and thorough understanding of algorithms and practically deployed Web and mobile applications of personalized technologies.

ДЛЯ ЗАМЕТОК

Yandex

golden sponsor



golden sponsor



bronze sponsor



EXITUS ACTA PROBAT
London Stock Exchange Group

bronze sponsor



partner



partner

ROMIP
SEMINAR

