

Additional Workshops Co-located with CIKM 2017

Workshop on Data and Text Mining in Biomedical Informatics (DTMBio 2017)

While CIKM presents state-of-the-art research in informatics with the primary focus on data and text mining, the main focus of DTMBio is on biomedical and healthcare informatics. Delegates to the 11th edition of DTMBio will present and discuss interesting applications of up-to-date informatics in the context of biomedical research.

Organizers: **Doheon Lee**, KAIST, KOREA | **Mark Stevenson**, University of Sheffield, United Kingdom

Workshop on Data and Algorithm Bias (DAB 2017)

More and more, we as members of society are becoming subject to socio-economic and political decisions made using statistical models trained on enormous amounts of cross-referenced data. This data may originate from many different sources, including governments (e.g. census data), industry (e.g. telephone or credit card transactions) and even ourselves (e.g. our use of online social networks).

However, even the cleanest of datasets, those generated with the utmost care, using careful phrasing of survey questions and careful sampling, may contain bias. Data sets often reflect historical bias of gender, age or ethnicity that can be extremely subtle and deep-rooted. In addition, these “smaller” subtle biases can be further amplified algorithmically into full-blown discriminatory profiling of certain groups. It is therefore imperative to study scientifically the causes and effects of bias in the era of big data and propose palliative measures.

The aim of this workshop is to gather researchers in industry and academia working on algorithmic and data bias in all areas of society: health care, finance, education and other that can help to design discrimination-free algorithms and fairness-aware data mining.

Organizers: **Ricardo Baeza-Yates**, NTENT, USA; U Pompeu Fabra, Spain & University of Chile | **Loreto Bravo**, Data Science Institute, Universidad del Desarrollo, and Telefónica R&D, Chile | **Ciro Cattuto**, ISI Foundation Torino, Italy | **Leo Ferres**, Data Science Institute, Universidad del Desarrollo, and Telefónica R&D, Chile | **Jeanna Matthews**, Clarkson University, USA | **Daniela Paolotti**, ISI Foundation Torino, Italy

Workshop on Big Data Analytics for Enhancing Public Transport (BigTransport17)

Public transport is a critical component of a smart city. In a dense urban city, public transport system is the preferred means to move people around. As the most sustainable and scalable solution, public transport now needs innovation to respond to new challenges brought by improving commuting experiences. These challenges include:

- Increased expectation of service quality, comfort and efficiency from commuters;
- Influx of new commuters working or visiting cities;
- Imbalanced supply and demand; and
- Last mile commuting gaps.

Meanwhile, public commuters today generate a massive amount of data traces. These include:

- (i) sensor, image, and video data collected by the existing public transport infrastructure;
- (ii) train and bus trips recorded by electronic farecard systems;
- (iii) taxi bookings and taxi trips recorded by mobile apps;
- (iv) bicycle rental and biking trips recorded by bike sharing apps; and
- (v) social media posts on public transport events.

These rich data traces offer new opportunities for research in information retrieval, database, data mining and machine learning to enhance commuting experience enhancement, namely:

- Identifying areas for public transport service improvement;
- Discovering regular travel patterns of commuters;
- Modelling and monitoring of commuting experience;
- Personalizing public transport services to improve individual commuting experience; and
- Engaging commuters in crowdsourcing resources to address unmet demand.

Organizers: **Baihua Zheng**, Singapore Management University | **Chih-Chieh Hung**, Tamkang University, Taiwan | **Wang-Chien Lee**, Penn State University, USA | **Ee-Peng Lim**, Singapore Management University