

\*\*\*\*\*

**5<sup>th</sup> INTERNATIONAL WINTER SCHOOL ON BIG DATA**

**BigDat 2019**

**Cambridge, United Kingdom**

**January 7-11, 2019**

Co-organized by:

Cambridge Big Data Initiative, University of Cambridge

Institute for Research Development, Training and Advice (IRDTA)  
Brussels / London

<http://bigdat2019.irdta.eu/>

\*\*\*\*\*

--- Early registration deadline: November 11, 2018 ---

\*\*\*\*\*

**SCOPE:**

BigDat 2019 will be a research training event with a global scope aiming at updating participants on the most recent advances in the critical and fast developing area of big data, which covers a large spectrum of current exciting research and industrial innovation with an extraordinary potential for a huge impact on scientific discoveries, medicine, engineering, business models, and society itself. Renowned academics and industry pioneers will lecture and share their views with the audience.

Most big data subareas will be displayed, namely foundations, infrastructure, management, search and mining, security and privacy, and applications (to biological and health sciences, to business, finance and transportation, to online social networks, etc.). Major challenges of analytics, management and storage of big data will be identified through 2 keynote lectures, 24 four-hour courses, and 1 round table, which will tackle the most active and promising topics. The organizers are convinced that outstanding speakers will attract the brightest and most motivated students. Interaction will be a main component of the event.

An open session will give participants the opportunity to present their own work in progress in 5 minutes. Moreover, there will be two special sessions with industrial and recruitment profiles.

**ADDRESSED TO:**

Master's students, PhD students, postdocs, and industry practitioners will be typical profiles of participants. However, there are no formal pre-requisites for attendance in terms of academic degrees. Since there will be a variety of levels, specific knowledge background may be assumed for some of the courses. Overall, BigDat 2019 is addressed to students, researchers and practitioners who want to keep themselves updated about recent developments and future trends. All will surely find it fruitful to listen and discuss with major researchers, industry leaders and innovators.

### **STRUCTURE:**

3 courses will run in parallel during the whole event. Participants will be able to freely choose the courses they wish to attend as well as to move from one to another.

### **VENUE:**

BigDat 2019 will take place in Cambridge, a city home of a world-renowned university. The venue will be:

University of Cambridge  
Department of Engineering  
Trumpington Street  
Cambridge CB2 1PZ

### **KEYNOTE SPEAKERS:**

tba

### **PROFESSORS AND COURSES:**

Thomas Bäck (Leiden University), [introductory/intermediate], Data Driven Modeling and Optimization for Industrial Applications

Richard Bonneau (New York University), [introductory] Large Scale Machine Learning Methods for Integrating Protein Sequence and Structure to Predict Gene Function

Altan Cakir (Istanbul Technical University), [introductory/intermediate] Processing Big Data with Apache Spark: From Science to Industrial Applications

Jiannong Cao (Hong Kong Polytechnic University), [introductory/intermediate] Cross-domain Multi-source Big Data Fusion and Analytics

Nitesh Chawla (University of Notre Dame), [intermediate/advanced] Network Science: Representation Learning and Higher Order Networks

Nello Cristianini (University of Bristol), [introductory] The Interface between Big Data and Society

Geoffrey C. Fox (Indiana University, Bloomington), [intermediate] High Performance Big Data Computing

David Gerbing (Portland State University), [introductory] Data Visualization with R

Craig Knoblock (University of Southern California), [intermediate/advanced] Building Knowledge Graphs

Geoff McLachlan (University of Queensland), [intermediate/advanced] Applying Finite Mixture Models to Big Data

Folker Meyer (Argonne National Laboratory), [intermediate] Skyport2: A Multi Cloud Framework for Executing Scientific Workflows

Wladek Minor (University of Virginia), [introductory/advanced] Big Data in Biomedical Sciences

Soumya Mohanty (University of Texas Rio Grande Valley), [introductory/intermediate] Swarm Intelligence Methods for Statistical Regression

Sankar K. Pal (Indian Statistical Institute), [introductory/advanced] Machine Intelligence and Soft Granular Mining: Features, Applications and Challenges

Lior Rokach (Ben-Gurion University of the Negev), [introductory/advanced] Ensemble Learning

Michael Rosenblum (University of Potsdam), [introductory/intermediate] Synchronization Approach to Time Series Analysis

Hanan Samet (University of Maryland), [introductory/intermediate] Sorting in Space: Multidimensional, Spatial, and Metric Data Structures for Applications in Spatial and Spatio-textual Databases, Geographic Information Systems (GIS), and Location-based Services

Rory Smith (Monash University), [intermediate/advanced] Statistical Inference: Optimal Methods for Learning from Signals in Noise

Jaideep Srivastava (University of Minnesota), [intermediate] Social Computing – Concepts and Applications

Mayte Suárez-Fariñas (Icahn School of Medicine at Mount Sinai), [intermediate] A Practical Guide to the Analysis of Longitudinal Data Using R

Jeffrey Ullman (Stanford University), [introductory] Big-data Algorithms That Aren't Machine Learning

Andrey Ustyuzhanin (National Research University Higher School of Economics), [intermediate/advanced] Surrogate Modelling for Fun and Profit

Wil van der Aalst (RWTH Aachen University), [introductory/intermediate] Process Mining: Data Science in Action

Zhongfei Zhang (Binghamton University), [introductory/advanced] Relational and Multimedia Data Learning

### **OPEN SESSION**

An open session will collect 5-minute voluntary presentations of work in progress by participants. They should submit a half-page abstract containing title, authors, and summary of the research to david@irdta.eu by December 30, 2018.

### **INDUSTRIAL SESSION:**

A session will be devoted to 10-minute demonstrations of practical applications of big data in industry. Companies interested in contributing are welcome to submit a 1-page abstract containing the program of the demonstration and the logistics needed. At least one of the people participating in the demonstration must register for the event. Expressions of interest have to be submitted to david@irdta.eu by December 30, 2018.

### **EMPLOYER SESSION:**

Firms searching for personnel well skilled in big data will have a space reserved for one-to-one contacts. It is recommended to produce a 1-page .pdf leaflet with a brief description of the company and the profiles looked for, to be circulated among the participants prior to the event. At least one of the people in charge of the search must register for the event. Expressions of interest have to be submitted to david@irdta.eu by December 30, 2018.

### **ORGANIZING COMMITTEE:**

Chiara Paola Codebò (Genova)  
Sara Morales (Brussels)  
Manuel J. Parra-Royón (Granada)  
Jeffrey Salmond (Cambridge)  
David Silva (London, co-chair)  
Filippo Spiga (Cambridge, co-chair)  
Richard E. Turner (Cambridge)

### **REGISTRATION:**

It has to be done at

<http://bigdat2019.irdta.eu/registration/>

The selection of up to 8 courses requested in the registration template is only tentative and non-binding. For the sake of organization, it will be helpful to have an approximation of the respective demand for each course. During the event, participants will be free to attend the courses they wish.

Since the capacity of the venue is limited, registration requests will be processed on a first come first served basis. The registration period will be closed and the on-line registration facility disabled when the capacity of the venue is exhausted. It is highly recommended to register prior to the event.

**FEES:**

Fees comprise access to all courses and lunches. There are several early registration deadlines. Fees depend on the registration deadline.

**ACCOMMODATION:**

Suggestions for accommodation are available on the event website.

**CERTIFICATE:**

A certificate of successful participation in the event will be delivered indicating the number of hours of lectures.

**QUESTIONS AND FURTHER INFORMATION:**

david@irdta.eu

**ACKNOWLEDGMENTS:**

Cambridge Big Data Initiative, University of Cambridge

Institute for Research Development, Training and Advice (IRDTA) –  
Brussels/London