Organizers

Chair CAS – **Wenqing Shen**
Member of CAS
Former Vice President of NSFC
Chair of Shanghai Institute for Advanced Studies, CAS
Professor of Shanghai Institute for Applied Physics

Co-Chair CAS – **Xu Zhang**
Vice President of Shanghai Branch, CAS
Director of Shanghai Institute for Advanced Studies, CAS
Professor of Institute of Neuroscience, SIBS, CAS

Chair MPG – **Hans Wolfgang Spiess**
Director (em.) Max Planck Institute for Polymer Research

Steering Committee CAS
Chair: Huadong Guo, Member of CAS,
Institute of Remote Sensing and Digital Earth, CAS
Runsheng Chen, Member of CAS, Institute of Biophysics, CAS
Jianjun Sun, President, School of Information Management, Nanjing University
Zhiwei Xu, Institute of Computing Technology, CAS

Steering Committee MPG
Chair: Thomas Lengauer,
Managing Director of the Max Planck Institute for Informatics
Claudia Draxl, Physics Department and IRIS Adlershof, Humboldt Universität zu Berlin; Max Planck Fellow at the Fritz Haber Institute of the Max-Planck-Gesellschaft
Klaus-Robert Müller, Head of Group Machine Learning, Technical University of Berlin
Jürgen Renn, Director at the Max Planck Institute for the History of Science
Matthias Scheffler, Director at the Fritz Haber Institute of the Max-Planck-Gesellschaft
Martin Vingron, Managing Director of the Max Planck Institute for Molecular Genetics, Director at the CAS-MPG Partner Institute for Computational Biology

Fellows CAS
Lizhe Wang, Institute of Remote Sensing and Digital Earth, CAS
Jianjun Luo, Institute of Biophysics, CAS
Huawei Shen, Institute of Computing Technology, CAS
Dongxiao Gu, School of Information Management, Hefei University of Technology
Dong Liang, Institute of Remote Sensing and Digital Earth, CAS

Fellows MPG
Marcel Schulz, Cluster of Excellence Multimodal Computing and Interaction, Saarland University / Max Planck Institute for Informatics
Luca Ghiringhelli, Fritz Haber Institute of the Max-Planck-Gesellschaft
Grégoire Montavon, Technical University of Berlin
Florian Schmaltz, Max Planck Institute for the History of Science
Robert Schöpflin, Max Planck Institute for Molecular Genetics

CAS Headquarters Beijing
Tieniu Tan, Bureau of International Co-operation, CAS
Jiaofeng Pan, Bureau of Development and Planning, CAS
Feng Zhang, Bureau of Development and Planning, CAS
Fang Jiang, Bureau of Development and Planning, CAS
Wang Dongyao, Bureau of International Co-operation, CAS

CAS Shanghai
Wenjun Zhang, Shanghai Institute for Advanced Studies (SIAS)
Xiaolong Teng, Shanghai Institute for Advanced Studies (SIAS)
Fang Xue, Shanghai Institute for Advanced Studies (SIAS)

MPG Administrative Headquarters Munich
Christoph Ettl, Presidential Division, Scientific Coordination
Christiane Walch-Solimena, Presidential Division, Scientific Coordination
Sabine Panglung, Division for International Relations

The ERTC is funded by the CAS and the MPG

Max-Planck-Gesellschaft
Hofgartenstraße 8
80539 Munich, Germany
P.O. Box 10 10 62
800084 Munich, Germany

www.mpg.de/en

CHINESE ACADEMY OF SCIENCES
52, Sanlihe Road
Beijing 100864, China

www.cas.cn

VENUE

Shanghai Institute for Advanced Studies
Chinese Academy of Sciences
Building 1, 319 Yue Yang Rd.
Shanghai 200031, China

CONTACT

Ms Fang Xue
(Conference Office) ..... xuef@shb.ac.cn
Phone: +86 21 6433 7978
Fax: +86 21 6433 7927
Ms Sabine Panglung ..... panglung@gv.mpg.de
Dr. Christoph Ettl ..... ettl@gv.mpg.de
Dr. Christiane Walch-Solimena ..... walch-solimena@gv.mpg.de

CHINESE ACADEMY OF SCIENCES — MAX-PLANCK-GESELLSCHAFT
SHANGHAI INSTITUTE FOR ADVANCED STUDIES
EXPLORATORY ROUND TABLE CONFERENCES

6th Exploratory Round Table Conference

Big Data in the Natural Sciences and Humanities

Shanghai, November 19th to 21st, 2015

SIAS House

Exploratory Round Table Conferences of the Chinese Academy of Sciences and the Max-Planck-Gesellschaft

Exploratory Round Table Conferences or ERTC are a joint activity of the Chinese Academy of the Sciences **(CAS)** and the Max-Planck-Gesellschaft **(MPG)** under the auspices of the Shanghai Institute of Advanced Studies **(SIAS)**.

ERTC are intended to provide a platform for scientists of both **MPG** and **CAS** to exchange ideas and reflect on opportunities of newly emerging research areas together with the respective international key players at an early stage of these evolving fields. The main objective of the project is to act as a seed towards establishing new topical areas as part of a priority-setting process at the leading edge of science in the supporting organisations **CAS** and **MPG**. Moreover, the reports of the **ERTC** will be widely communicated to both science policy makers as well as to the general scientific community. **ERTC** are to be held at Shanghai in the premises of **SIAS** at least once a year.

MPG and CAS have maintained an exclusive partnership for over 30 years. The SIAS is an Institute of the CAS Shanghai Branch and was founded in 2001 with the support of the MPG as a hub for interdisciplinary and international dialogue.

Each ERTC has a total duration of 3-6 months for preparation and follow-up. The topic of the 6th meeting will be Big Data in the Natural Sciences and Humanities, including the following subthemes:

- **Big Data in Biomedicine**
- **Big Data in Physics, Chemistry and Earth Science**
- **Big Data in the Humanities and Social Sciences**
- **Technology Underlying Big Data**

ERTC on Big Data in the Natural Sciences and Humanities

Big Data has become a ubiquitous notion in recent years. Technological developments, in particular in informatics and high-throughput approaches, have revolutionized data generation in all fields of science. As a consequence, researchers in almost all areas of science face new and unforeseen challenges: Gathering data is so easy and quick that it exceeds by far the capacity to validate, analyze, visualize, store, and curate all the information. Tackling this challenge will without doubt lead to unprecedented data-driven scientific discoveries.

In the field of biomedicine, the dramatic advances in technologies that can be summed up as omics, such as high throughput DNA-sequencing, lead to vast amounts of data at dramatically plummeting costs. This revolution in biomedical research raises high expectations as to the increase of knowledge, understanding of health and disease, and eventually the development of powerful therapies to treat thus far uncurable diseases such as cancer or depression, in a personalized and precise fashion.

Such data-driven methods are revolutionizing not only drug design and drug discovery. Regarding chemistry and materials science, Big Data techniques in combination with computational modeling facilitate analyzing the vast space of yet unexplored compounds and materials - thus complementing and in several cases even replacing experiments. This high-throughput screening needs to be combined with novel big-data analytics tools, which then enables the identification of new scientific phenomena, advances materials science and engineering, and predicts materials with technologically relevant properties and functions.

Last but not least, the analysis of large volumes of data opens up new avenues of research in the field of the humanities and social sciences. The wealth of data which is already born digital as well as mass digitizing existing analog data allow to answer complex questions that were previously unanswerable. Analyzing the development and diffusion of knowledge, modeling cultural evolution, and predicting human behavior are just some of the challenges that lie ahead.

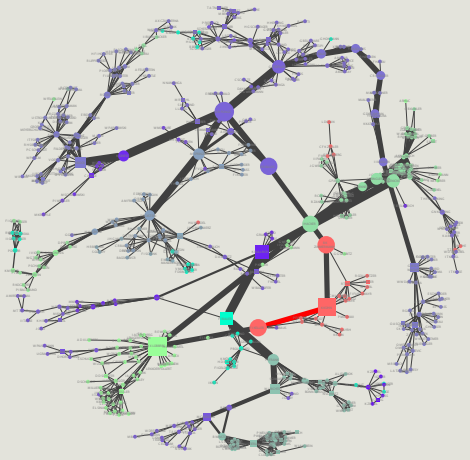
Big Data requires innovative technologies to efficiently process large quantities of data within tolerable elapsed times. Machine learning is one of today's most rapidly growing technical fields lying at the core of data science and artificial intelligence, indispensable for analyzing and classifying data. This development also provides challenges for theory-building. Whereas data mostly exhibit correlations and statistical dependencies, theory provides causal relationships. The interplay between data mining and theory building is an important issue, as Big Data continues to pervade scientific and private life. Furthermore, data capture both in the health segment and in daily life can pose a severe threat to privacy, as individuals are increasingly divulging data relating to individual behavior and performance.

This ERTC aims at elaborating a critical review of the presently existing ideas, strategies and aspirations of Big Data science. The results of the ERTC will serve as a basis for further consideration by CAS and MPG regarding research in this field.

Point cloud of the only modern building survey of the Pantheon in Rome
© *MPI for the History of Science*

Invited Speakers and Participants

Alessandro DE VITA (King's College London, UK)
Michael BACKES (Saarland University, Germany)
BI Jun (Nanjing University, China)
CHENG Xueqi (Institute of Computing Technology, CAS, China)
Linda DIVARCI (Max Planck Institute for the History of Science, Germany)
Roland EILS (University Heidelberg, Germany)
Gerd GRASSHOFF (Humboldt University of Berlin, Germany)
GUO Huadong (Institute of Remote Sensing and Digital Earth, CAS, China)
Moritz HELMSTAEDTER (Max Planck Institute for Brain Research, Germany)
HUANG Cui (Tsinghua University, China)
JIANG Hualiang (Shanghai Institute of Materia Medica, CAS, China)
Manfred LAUBICHLER (Arizona State University, USA)
Thomas LENGAUER (Max Planck Institute for Informatics, Germany)
LI Hong (Institute of Physics, CAS, China)
LI Jiang (Zhejiang University, China)
LI Yixue (Shanghai Center for Bioinformation Technology, SIBS, CAS, China)
LUO Jianjun (Institute of Biophysics, CAS, China)
Klaus Robert MÜLLER (Technical University of Berlin, Germany)
Nico PFEIFER (Max Planck Institute for Informatics, Germany)
Matthias SCHEFFLER (Fritz Haber Institute of the Max-Planck-Gesellschaft, Germany)
Matthias SCHEMMEL (Max Planck Institute for the History of Science, Germany)
WANG Jun (Beijing Genomics Institute, China)
WANG Lizhe (Institute of Remote Sensing and Digital Earth, CAS, China)
XU Zhiwei (Institute of Computing Technology, CAS, China)
ZHANG Baichun, (Institute for the History of Natural Sciences, CAS, China)



Connectivity between scholars at various research institutes, based on co-authorship in 1973 © *MPI for the History of Science*