

BigScholar 2019: The 6th Workshop on Big Scholarly Data

Feng Xia

School of Software
Dalian University of Technology
Dalian 116620, China
f.xia@acm.org

Irwin King

Department of Computer Science and Engineering
The Chinese University of Hong Kong
Shatin, NT, Hong Kong
king@cse.cuhk.edu.hk

Huan Liu

School of Computing, Informatics, and Decision
Systems Engineering
Arizona State University
Tempe, AZ 85287-8809, USA
huan.liu@asu.edu

Kuansan Wang

Microsoft Research
Redmond, Washington, USA
kuansan.wang@microsoft.com

ABSTRACT

Recent years have witnessed the rapid growth in the number of academics and practitioners who are interested in big scholarly data as well as closely-related areas. Quite a lot of papers reporting recent advancements in this area have been published in leading conferences and journals. Both non-commercial and commercial platforms and systems have been released in recent years, which provide innovative services built upon big scholarly data to the academic community. Examples include Microsoft Academic Graph, Google Scholar, DBLP, arXiv, CiteSeerX, Web of Knowledge, Udacity, Coursera, and edX. The workshop will contribute to the birth of a community having a shared interest around big scholarly data and exploring it using knowledge discovery, data science and analytics, network science, and other appropriate technologies.

KEYWORDS

Big Scholarly Data; Educational Big Data; Data Science; Science of Science; Academic Networks

ACM Reference format:

Feng Xia, Huan Liu, Irwin King, and Kuansan Wang. 2019. BigScholar 2019: The 6th Workshop on Big Scholarly Data. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM'19), November 3-7, 2019, Beijing, China*. 2 pages. <https://doi.org/10.1145/3357384.3358806>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.
CIKM '19, November 3–7, 2019, Beijing, China
© 2019 Copyright is held by the owner/author(s).
ACM ISBN 978-1-4503-6976-3/19/11.
<https://doi.org/10.1145/3357384.3358806>

1. INTRODUCTION

The number of scholarly documents produced by academics, researchers, and practitioners worldwide is increasing at an unprecedented speed. The term big scholarly data is coined for this rapidly growing scholarly source of information. Many large collections of scholarly data including digital libraries, search engines, repositories, knowledge bases, Wikipedia, and the Web have already covered millions of journal articles, conference proceedings, degree theses, books, patents, technical reports, tutorials, course materials, etc. For instance, the Microsoft Academic Graph contains scientific publication records, citation relationships between those publications, as well as authors, institutions, journals, conferences, and fields of study. The DBLP bibliography now lists more than 5000 conference and workshop series, as well as more than 1500 journals in computer science, which involve more than 4 million publications by more than 2 million authors.

Big scholarly data bring about new opportunities and challenges with respect to knowledge discovery, data mining, science of science, and education. It is imperative and vital for scholars to drive their knowledge towards the innovative generation of values from big scholarly data. New knowledge can be extracted by analyzing and mining big scholarly data to, e.g., better understand research dynamics, scientific collaboration and success, identify new directions of research, assess the quality of science, and enable personalized teaching and learning. To achieve these goals, however, a lot of challenges facing big scholarly data acquisition, storage, management, processing and usage must be addressed.

Following the success of the previous five editions, the BigScholar 2019 workshop aims at bringing together academics and practitioners from diverse fields to share ideas and experience with management, analysis, mining, and applications of big scholarly data. The goal is to contribute to the birth of a community having a shared interest around big scholarly data

and exploring it using knowledge discovery, data science and analytics, network science, and other appropriate technologies.

The workshop will be a half-day workshop. The format of the workshop will include keynote talks, research and position paper presentations, and one discussion panel. The workshop will be held in conjunction with the 28th ACM International Conference on Information and Knowledge Management (CIKM 2019). Detailed program of the workshop will be posted on the workshop website: <http://thealphalab.org/big scholar/>.

2. SCOPE OF THE WORKSHOP

Authors are invited to submit original papers that must not have been submitted to or published in any other workshop, conference, or journal. The workshop will accept full papers describing completed work, work-in-progress papers with preliminary results, as well as position papers reporting inspiring and intriguing new ideas. All accepted papers will be published in the journal *Frontiers in Big Data* (as Article Collection on Big Scholarly Data).

Topics of interest include (but not limited to):

- New approaches to search, crawling and integration of scholarly data from various data sources
- Methods for storing, indexing, and query processing for big scholarly data
- Practices for scholarly data management and sharing
- Big scholarly data analysis, mining, and visualization
- Network science for scholarly data analytics
- Graph and text mining in big scholarly data
- Measuring the impact of publications, funding, courses, individuals, teams, etc.
- Computational behavioural sciences in research and education
- Academic social network analysis and mining
- Scholarly recommendation
- Understanding and predicting success in research and education
- Design of next generation platforms and systems for research and education
- Novel services and applications for research and education

3. PREVIOUS EDITIONS

The first edition of the BigScholar workshop (i.e. BigScholar 2014) was held in Seoul, Korea, on April 8, 2014, as a workshop of WWW 2014 (The 23rd International World Wide Web Conference).

The second edition of the BigScholar workshop (i.e. BigScholar 2015) was held in Florence, Italy, on May 18, 2015, as a workshop of WWW 2015 (The 24th International World Wide Web Conference).

The third edition of the BigScholar workshop (i.e. BigScholar 2016) was held in Montreal, Canada, on April 12, 2016, as a workshop of WWW 2016 (The 25th International World Wide Web Conference).

The fourth edition of the BigScholar workshop (i.e. BigScholar 2017) was held in Perth, Australia, on April 3, 2017, as a workshop of WWW 2017 (The 26th International World Wide Web Conference).

The five edition of the BigScholar workshop (i.e. BigScholar 2018) was held in London UK, on Aug 20, 2018, as a workshop of KDD 2018 (The 24rd ACM SIGKDD Conference on Knowledge Discovery and Data Mining).

Over 200 researchers and practitioners from around the world have participated in previous editions of the BigScholar workshop. A number of leading experts have delivered keynote speeches at the workshop, including:

- Hanghang Tong, University of Illinois at Urbana-Champaign, USA
- Kuansan Wang, Microsoft Research, USA
- C. Lee Giles, Pennsylvania State University, USA
- Jie Tang, Tsinghua University, China
- Lucy Montgomery, Curtin University, Australia
- Jianxin Li, Deakin University, Australia
- Maarten de Rijke, University of Amsterdam, The Netherlands
- Dashun Wang, Northwestern University, USA
- Roberta Sinatra, IT University of Copenhagen, Denmark
- Francesco Osborne, The Open University, UK

ACKNOWLEDGMENTS

We take the opportunity to thank the editorial office of *Frontiers in Big Data* for publishing all accepted papers of this workshop as well as the reviewers who shared their valuable expertise and time to make the workshop a success. We also thank all the authors for submitting their papers to BigScholar 2019.

The Organizing Committee of the CIKM 2019 conference is acknowledged for making local arrangement for BigScholar 2019. Special thanks go to the Workshop Co-Chairs of CIKM 2019 for their support and guidance.