

CIKM 2018 Workshops: Preface

Alfredo Cuzzocrea
University of Trieste and ICAR-CNR
Trieste, Italy
alfredo.cuzzocrea@dia.units.it

Francesco Bonchi
ISI Foundation
Turin, Italy
francesco.bonchi@isi.it

Dimitris Gunopulos
University of Athens
Athens, Greece
dg@di.uoa.gr

1 Introduction

This paper briefly introduces the workshops that have been co-located with the *27th ACM International Conference on Information and Knowledge Management* (CIKM 2018), held during October 22-26, 2018 in Turin, Italy.

CIKM 2018 selected 9 workshops, which have been focused on several topics, all concerned with information and knowledge management in the big data era. The CIKM 2018 workshop are listed in the following:

- DAB 2018 – *2018 Workshop on Data and Algorithms Bias*;
- DTMBio 2018 – *12th International Workshop on Data and Text Mining in Biomedical Informatics*;
- EYRE 2018 – *1st International Workshop on Entity REtrieval*;
- GLARE 2018 – *1st International Workshop on Generalization in InformAtion REtrieval*;
- KDAH 2018 – *2018 International Workshop on Knowledge-Driven Analytics Impacting Human Quality of Life*;
- LeDAM 2018 – *2018 International Workshop on Legal DAta Mining*;
- INRA 2018 – *6th International Workshop on News Recommendation and Analytics*;
- RDSM 2018 – *2nd International Workshop on Rumours and Deception in Social Media*;
- SIR 2018 – *2018 International Workshop on Social Interaction-based Recommendation*.

“*From Big Data and Big Information to Big Knowledge*” has been the main theme of both CIKM 2018 main conference and CIKM 2018 workshops. Indeed, nowadays big data management and analytics (e.g., [ZE11, CCS12]), including advanced topics like knowledge and information extraction from big data sources (e.g., [BCC⁺14, WYC⁺13]) and privacy-preserving big data management (e.g., [CB11]), play a critical role, and they are attracting the attention from the community more and more. Papers presented at these workshops have been focused on these topics mainly, with relevant proposals that span from theory to practice and high-impact research project results.

The workshops have finally originated in some relevant questions that correspond to some open problems in big data management and analytics research. In the following, we report on some among the most noticeable ones.

Open Big Data Open big data define the management, mining and analytics of big data in the so-called *open environments* (e.g., industry 4.0 settings), where (big) data can be exchanged among heterogeneous systems, and analytics processes must be designed as to capture the peculiarities and the (business) goals of the different environments, simultaneously.

Big Data Quality Big data quality refers to the issue of understanding the *quality* of big data sources. This problem is strictly related to the so-called *provenance problem* (e.g., [CD18]). Indeed, when we process a big data source, usually we do not know the origin of such data, but simply use them. The challenge is thus determining models, techniques and algorithms to check about the quality of the (big) data source, in an automatic or semi-automatic manner.

Big Data in Emerging Domains The usage, processing and management of big data sources in emerging domains, such as bio-medical systems, social networks, sensor systems, and so forth, is, without doubts, one of the most relevant challenges of future years, as

practical issues pose the basis for relevant innovation in science and technology.

We firmly hope that papers of the CIKM 2018 workshops will represent a milestone in the exciting context of big data research.

References

- [BCC⁺14] Peter Braun, Juan J. Cameron, Alfredo Cuzzocrea, Fan Jiang, and Carson Kai-Sang Leung. Effectively and efficiently mining frequent patterns from dense graph streams on disk. In *18th International Conference in Knowledge Based and Intelligent Information and Engineering Systems, KES 2014*, pages 338–347, 2014.
- [CB11] Alfredo Cuzzocrea and Elisa Bertino. Privacy preserving OLAP over distributed XML data: A theoretically-sound secure-multiparty-computation approach. *J. Comput. Syst. Sci.*, 77(6):965–987, 2011.
- [CCS12] Hsinchun Chen, Roger H L Chiang, and Veda C. Storey. Business intelligence and analytics: From big data to big impact. *MIS Quarterly: Management Information Systems*, 36(4):1165–1188, 2012.
- [CD18] Alfredo Cuzzocrea and Ernesto Damiani. Pedigree-ing your big data: Data-driven big data privacy in distributed environments. In *18th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, CCGRID 2018, Washington, DC, USA, May 1-4, 2018*, pages 675–681, 2018.
- [WYC⁺13] Zhiang Wu, Wenpeng Yin, Jie Cao, Guandong Xu, and Alfredo Cuzzocrea. Community detection in multi-relational social networks. In *Web Information Systems Engineering - WISE 2013 - 14th International Conference, Nanjing, China, October 13-15, 2013, Proceedings, Part II*, pages 43–56, 2013.
- [ZE11] Paul Zikopoulos and Chris Eaton. *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*. McGraw-Hill Osborne Media, 1st edition, 2011.