



## Special Issue Editorial

Mohammad A. Hasan<sup>1</sup>  · Faizan Javed<sup>2</sup> · Ioana E. Marinescu<sup>3</sup>

Published online: 24 September 2018  
© The Author(s) 2018

This issue of Data Science and Engineering (DSE) contains a collection of four extended papers from the following Workshop: “The First International Workshop on Data Science for Human Capital Management (DSHCM).” Besides these four papers, this DSE issue also has one invited paper and one regular research paper.

The first DSHCM workshop was held on November 18, 2017, in New Orleans, LA, USA, collocated with the 17th Edition of IEEE International Conference on Data Mining (ICDM). The second DSHCM workshop was held on September 14, 2018, with ECML (European Conference of Machine Learning) conference in Dublin, Ireland. The theme of DSHCM workshop is human capital management, which refers to the set of practices and systems that facilitate talent acquisition and management in labor market. Specifically, this workshop solicits research works that discuss application of data science and machine learning methodologies in all areas of human capital management, both pre-hire and post-hire, including talent and labor market analytics, job advertising and distribution, professional social networks, candidate sourcing, tracking, onboarding, employee training, benefits administration and compliance.

There are many recent successful applications of data mining and data science techniques to problems in the HCM domain. For example, text classification techniques are used for job posting classification; sequence labeling and statistical modeling approaches find application in resume and job parsing; near-deduplication algorithms in concert with big data pipelines power many job aggregators; predictive analytics models employee flight risk; ontology mining techniques help build knowledge graphs of human capital entities; personalized search and semantic search help job

seekers by demystifying searchers’ intent and contextual meaning of terms in the recruitment domain; and recommender systems have been used for expertise search and job recommendations. To this date, DSHCM is the only venue that provides a platform for publishing research works for solving problems in HCM domain by applying data science and engineering methodologies.

The 2017 DSHCM workshop was very successful. About 17 papers were submitted in this workshop, of which 11 were selected for publications in the workshop proceedings. The selected papers have been published and archived by IEEE digital library [1]. DSE Guest editors, Mohammad Hasan, Faizan Javed and Ioana E. Marinescu, have selected four papers from the accepted papers based on the relevance of the journal and the reviews of the workshop version of the papers. The authors were asked to revise their workshop paper for journal publication and in accordance with customary practice to add 30% new materials. The revised paper again went through the normal journal-style review process and was finally presented to the readers in the present form. We appreciate the willingness of the authors to help in organizing this special issue.

The four papers in this special issue cover both pre-hire and post-hire data analytics. The first paper titled, “Talent Flow Analytics in Online Professional Networks,” belongs to pre-hire data analytics. In this paper, the authors study job preferences and career progression of an individual by using data from online professional networks. Authors also perform connectivity analysis at job and organization levels to derive insights on talent flow as well as on job and organizational competitiveness. The second paper, “Automatically Detecting Errors in Employer Industry Classification using Job Postings,” also belongs to pre-hire data analytics. This work is important for online job portals, such as careerbuilder.com and monster.com. In this work, authors show how job posting data can be used for the classification of employer industry so that the jobs can be organized and displayed based on the employer industry. Such classification also helps effective job recommendation to potential job seekers. The remaining two papers belong to post-hire

---

✉ Mohammad A. Hasan  
alhasan@iupui.edu

<sup>1</sup> Indiana University Purdue University Indianapolis, Indianapolis, IN, USA

<sup>2</sup> CareerBuilder, Atlanta, GA, USA

<sup>3</sup> University of Pennsylvania, Philadelphia, PA, USA

data analytics. One of these papers is titled “What’s Next? A Recommendation System for Industry Training.” In this paper, the authors discuss how an employer can develop a system for recommending its employees the correct training module that they (the employees) should consider for improving their skill set. Authors present several unsupervised sequence mining algorithms to mine the past training data from the organization for building personalized recommendation system. Authors also show that their proposed recommendation system is far superior than the traditional recommendation engines. The final paper is titled, “Estimating fungibility between skills by combining skill-similarities obtained from multiple data sources.” In this paper, the authors show different methods for combining skill similarities from multiple data sources to obtain a single skill-similarity matrix which can be used for finding substituted skills of any given skill. Estimating skill fungibility is useful for industry as they can use this information for selecting candidates (for training) when a relatively new skill is hard to find within the existing employees.

We thank the DSE editor-in-chief, Dr. Bertino, and Dr. Wang, for facilitating the publication of extended version of DSHCM papers in this issue of DSE journal. We hope that the readers enjoy the special issue.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## Reference

1. Gottumukkala R, Ning X, Dong G, Raghavan V, Aluru S, Karypis G, Miele L, Wu X (eds) (2017) In: IEEE international conference on data mining workshops (ICDMW), ICDM 2017, IEEE computer society, 18–21 Nov 2017, New Orleans, LA, USA