

## Editor's introduction to the special issue of the 6th Biomedical Linked Annotation Hackathon (BLAH6)

Jin-Dong Kim<sup>1\*</sup>, Kevin Bretonnel Cohen<sup>2</sup>, Fabio Rinaldi<sup>3</sup>, Zhiyong Lu<sup>4</sup>, Nigel Collier<sup>5</sup>, Hyun-Seok Park<sup>6</sup>

<sup>1</sup>Database Center for Life Science (DBCLS), Research Organization of Information and Systems (ROIS), Kashiwa, Chiba 277-0871, Japan

<sup>2</sup>School of Medicine, University of Colorado, Aurora, CO 80045, USA

<sup>3</sup>Dalle Molle Institute for Artificial Intelligence Research (IDSIA), 6928 Manno, Switzerland

<sup>4</sup>National Center for Biotechnology Information (NCBI), National Institutes of Health (NIH), Bethesda, MD 20894, USA

<sup>5</sup>Faculty of Modern & Medieval Languages, University of Cambridge, Cambridge CB3 9DP, UK

<sup>6</sup>Center for Convergence Research of Advanced Technologies, Ewha Womans University, Seoul 03760, Korea

As data science gains in importance and popularity, the need for accessing data in scientific literature is rapidly increasing. While structured databases are supposed to supply readily machine-readable data, unstructured contents, particularly scientific literature, are recognized as a biggest source of data with comprehensive details, e.g., experimental environments and actual observations.

Since the importance of scientific literature for data science has been widely recognized, several groups have invested to develop various text mining resources. While many of them are publicly available, interoperability of them remains a critical issue, hindering efficient use or reuse of them, particularly in mix with others.

The Biomedical Linked Annotation Hackathon (BLAH) series is annually organized to join forces of biomedical text mining for the goal to promote interoperability among text mining resources. The sixth edition of it was held in Tokyo, February 4–7, 2020, with 52 participants from 9 countries. The first day was held as a symposium to exchange and publicise the activities and ideas of the participants, and the following three days was held as a hackathon: the participants worked on implementing their ideas with collaboration with other participants.

While the main theme of the event was improving interoperability of biomedical literature mining, which include annotation datasets, tools, platforms, terminology resources, and so on, this year, “social media mining” was also explored as a special theme. Social media is recognized as a good source of raw signals on how people are thinking about what is going on in the world, which are largely missing in scientific literature. Therefore, social media mining is expected to complement literature mining.

This special issue is a collection of the reports on achievements from the hackathon, which address various issues of biomedical literature and social media mining, including document collection, automatic annotation, manual annotation, annotation platform, translation, terminology, ontology, and so on. Note that, except a few, many of the works began just before or even during the hackathon, and due to the limited time for work,

they are often small-sized works, which are expected to benefit from collaboration with other participants. Readers will find that many of the articles have co-authorship with, or acknowledgment of other participants, which is a typical nature of hackathon-oriented publications.

We hope that this will be an opportunity for the readers of the journal *Genomics & Informatics* to get aware of the state-of-the-art activities regarding interoperability of biomedical text mining, and at the same time to observe activities of hackathons like BLAH.

## ORCID

Jin-Dong Kim: <https://orcid.org/0000-0002-8877-3248>

Kevin Bretonnel Cohen: <https://orcid.org/0000-0003-1749-8290>

Fabio Rinaldi: <https://orcid.org/0000-0001-5718-5462>

Zhiyong Lu: <https://orcid.org/0000-0002-8301-9553>

Nigel Collier: <https://orcid.org/0000-0002-7230-4164>

Hyun-Seok Park: <https://orcid.org/0000-0002-6617-2740>

## Acknowledgments

The 6th Biomedical Linked Annotation Hackathon was held with financial support of National Bioscience Database Center (NBDC) of Japan Science and Technology Agency (JST) and Research Organization of Information and Systems (ROIS).