Editorial for the Special Issue on "CDEC: Cross-Disciplinary Data Exchange and Collaboration"

Teruaki Hayashi * and **Yukio Ohsawa**

Department of Systems Innovation, School of Engineering, The University of Tokyo, Tokyo 113-8656, Japan; ohsawa@sys.t.u-tokyo.ac.jp

* Correspondence: teru-h.884@nifty.com

Due to recent developments in big data and artificial intelligence (AI), the importance of data and data mining is increasing. Beyond these expectations, there are externalizations of interdisciplinary issues. Many papers regarding data mining have been published, and the processes of analyzing data have been widely shared. However, there are only a limited number of studies on the process of cross-disciplinary data exchange and collaboration and its ecosystem. As this process encompasses various activities of different stakeholders, it is difficult to quantitatively evaluate the patterns or processes. Moreover, both big data and small data are necessary for decision-making. Thus, it is essential to discuss the dynamics of such a network of heterogeneous data in different fields.

To address these gaps, we have established and conducted an international workshop series on Cross-disciplinary Data Exchange and Collaboration (CDEC) since 2018. The scope of CDEC includes practical fields where topics are analyzed using data, solutions for challenging social issues, and cross-disciplinary data collaboration and its processes. CDEC considers not only cleanly formatted single data, but also heterogeneous data that affect human behavior, thoughts, and intentions in different domains. We also focus on discussions to obtain tacit knowledge of data mining through analysis and synthesis. In addition to these research fields, the cognitive approach for observing the processes of knowledge discovery and data exchange is also included in our workshop series.

As part of the International Conference on Data Mining (ICDM) workshops, we successfully conducted the 1st and 2nd International Workshop on CDEC in Singapore and Beijing (2018 and 2019, respectively). Seven distinguished papers were accepted after a triple blind review process (acceptance rate: 43%) and presented to an audience of approximately 30 participants in 2018. In 2019, 10 papers were accepted and presented (acceptance rate: 48%) after a triple blind review process. Thanks to the editors of MDPI, a Special Issue on CDEC was published in *Information* (ISSN 2078-2489) after the conclusion of the workshop under the section "Information and Communications Technology".

Based on the outcomes of these past workshops, the six papers below were accepted and published for this Special Issue. All submitted papers were processed via the standard peer-review process of MDPI, where reviewers were fairly selected with careful consideration by the MDPI editorial team to avoid the influence of any personal interest from the guest editors. Let us introduce the papers and the shared mission of CDEC.

In the paper "Feature Extraction of Laser Machining Data by Using Deep Multi-Task Learning" [1], Zhang et al. proposed a framework of deep feature extraction and applied it to the data of laser processing images. Their method performed better than a deeper model and could also meet real-time requirements due to lower computational costs. Interestingly, this research was realized by exchanging knowledge and data by discovering research subjects from communication between AI researchers and physicists. It is expected that the number of collaborations through data and technology will increase greatly in the future.

Similarly, the article "Topic Jerk Detector: Detection of Tweet Bursts Related to the Fukushima Daiichi Nuclear Disaster" [2] described a study that was realized through communication and data exchange in various fields. Nagaya et al. proposed a new model, the topic jerk detector, which accurately detects the timing of bursts of topics. They applied the topic jerk detector to Twitter data related to the Fukushima Daiichi nuclear power plant accidents and received expert feedback on the validity of the results and the practicality of the methodology. In addition to research methods, evaluation methods and perspectives vary greatly depending on the field. This study is interesting because it introduces a qualitative and quantitative evaluation of this cross-disciplinary issue.

In the study "Standardization Procedure for Data Exchange" [3], Fukami discussed the development and diffusion of common technological data specifications via a case analysis of the update of the HTML5 standard. To promote data sharing and trading among various organizations and individuals, data must be generated according to common specifications such as ontology, vocabulary, and syntax. However, standards within industries can interfere with the development and migration for standards across industries. This paper provides important insights into data interoperability and collaboration for innovations.

The article "How Will Sense of Values and Preference Change during Art Appreciation?" [4] introduced an unusual perspective in this Special Issue. Abe et al. discussed how the value and preference of art will change according to offered information and determined the influencing factors. In data exchange and collaboration, evaluating the value of data has been one of the biggest issues. This study considers the important question of whether we can exchange reliable information about our data and our partners before the actual data exchange and collaboration.

In the paper "Forecasting Net Income Estimate and Stock Price Using Text Mining from Economic Reports", Suzuki et al. proposed a methodology of forecasting movements in analysts' net income estimates and those of stock prices using natural language processing and neural networks in the context of analyst reports [5]. The significance of the obtained results and discussion was reported, and interesting findings were conveyed: In forecasting estimated net incomes, the distinction between opinion and non-opinion sentences was effective; in contrast, in forecasting stock price movements, the distinction was not effective. There is no unified solution for distinguishing or combining data; thus, it remains a debated topic. This paper provides important insights.

In the study "TEEDA: An Interactive Platform for Matching Data Providers and Users in the Data Marketplace" [6], Hayashi and Ohsawa proposed and discussed description items for sharing users' calls for data as data requests in the data marketplace and developed an interactive platform, named "treasuring every encounter of data affairs (TEEDA)". TEEDA facilitates matching and interactions between data providers and users, and functions as the communication platform for data exchange. Although there are numerous data exchange services and platforms on the Web, they only provide unilateral information from data providers. A platform to share calls for data by users can be expected to promote more data-oriented ventures across different fields.

The abovementioned studies are excellent in that they follow the concept of CDEC, and go beyond data mining, focusing on combining and reinterpreting diverse data and phenomena, and discovering and creating new values of data. The scope of CDEC can be expanded further, and we will continue to welcome new research and deeper discussions. This year, we are planning to extend CDEC to marketing, as well as present, share, and discuss the entire process from data design to analysis by setting the theme of "Designing Data Exchange Ecosystem for Human-AI Collaborative Society" (http://www.panda.sys.t.u-tokyo.ac.jp/CDEC/2020/index.html). Please consider submitting your paper. We hope our activities and the presented studies can contribute to the development of cross-disciplinary data exchange and collaboration worldwide in the near future.

References

1. Zhang, Q.; Wang, Z.; Wang, B.; Ohsawa, Y.; Hayashi, T. Feature Extraction of Laser Machining Data by Using Deep Multi-Task Learning. *Information* **2020**, *11*, 378. [CrossRef]
2. Nagaya, H.; Hayashi, T.; A. Torii, H.; Ohsawa, Y. Topic Jerk Detector: Detection of Tweet Bursts Related to the Fukushima Daiichi Nuclear Disaster. *Information* **2020**, *11*, 368. [CrossRef]
3. Fukami, Y. Standardization Procedure for Data Exchange. *Information* **2020**, *11*, 339. [CrossRef]
4. Abe, A.; Fukushima, K.; Kawada, R. How Will Sense of Values and Preference Change during Art Appreciation? *Information* **2020**, *11*, 328. [CrossRef]
5. Suzuki, M.; Sakaji, H.; Izumi, K.; Matsushima, H.; Ishikawa, Y. Forecasting Net Income Estimate and Stock Price Using Text Mining from Economic Reports. *Information* **2020**, *11*, 292. [CrossRef]
6. Hayashi, T.; Ohsawa, Y. TEEDA: An Interactive Platform for Matching Data Providers and Users in the Data Marketplace. *Information* **2020**, *11*, 218. [CrossRef]