Guest editorial

Data for better health

The wide adoption of electronic health records (EHRs), increasing use of digital health devices and applications, and the recent upsurge of wearable equipment all have contributed to the vast amount of health data generated daily. Overall our health is being transformed from experience-based to data-driven (van Panhuis *et al.*, 2018; Sharma *et al.*, 2018). The availability of huge volumes of data is setting the ground for evidence-based practice, precision medicine and predictive analytics (Devara *et al.*, 2013; Wu *et al.*, 2016; Cohen *et al.*, 2014). In addition, health data have gone beyond healthcare facilities and been extended to the home, workplace and communities (Patil *et al.*, 2016). For better health outcomes, there is a call to focus on the connections between data, technology and people, which demands an interdisciplinary effort (Fichman *et al.*, 2011).

Health data used to be intensively investigated by medical or health informatics communities only, such as the *International Association of Medical Informatics*, *the American Association for Health Information Management* and the *American Society for Medical Informatics*. However, in recent years, many professional communities outside the healthcare domain have developed special interests in collecting, using, and analyzing health data, including the iSchool consortium and the Association for Computing Machinery.

Therefore, this special issue – Data for Better Health – in the field of Library and Information Science, is to present the latest studies on health information technology and health data, calling for interdisciplinary participation to promote the connections between data, technology and people for better health outcomes. Thus, theories, methods, models or concepts from different disciplines are encouraged to address the big data challenge in health.

Manuscripts have been sought to achieve three aims: (1) to advance interdisciplinary collaboration for better health outcomes and healthcare quality by calling researchers from multiple fields to work together in terms of integrating theories, methods, models or concepts for competence and synergy; (2) to explore new service areas and models for libraries and information organizations by promoting thinking and planning regarding health data and information technology services; (3) to educate the current and future information professionals by disseminating the latest evidence, experience, innovative methods and best practices related to health information technologies and health data.

This special issue of *Library Hi Tech* presents a total of 11 selected articles which are contributed by 31 researchers across multiple disciplines (e.g. library and information science, public health, medicine, nursing, computer science and economics) in four countries (i.e. China, the United States, South Korea and Iran). The included studies demonstrate the following patterns in promoting the connections between data, technology, and people and achieving the three aims.

First, researchers adopt not only methodologies specialized in library and information science (e.g. bibliometrics, altmetrics and scientometrics) but also the latest data analytics techniques (e.g. text mining and natural language processing) to investigate research literature related to complex health data issues. Liu *et al.* (2020b) take a bibliometric approach to systematically examine 1,953 eligible publications that used Medical Expenditure Panel Survey (MEPS) data, delineating how MEPS data have supported scientific discoveries in biomedical and health sciences research. Moradi and Dokhani (2020) use altmetrics indicators together with content analysis and scientometrics to assess the impact of health science research in D8 countries (i.e. Iran, Turkey, Pakistan, Bangladesh, Indonesia, Malaysia, Egypt and Nigeria) and identify research gaps. Zhu *et al.* (2020) propose a literature-based drug

Library Hi Tech
Vol. 38 No. 4, 2020
pp. 701-703

0737-8831
DOI 10.1108/LHT-11-2020-280

repurposing strategy for Parkinson's disease. They adopt natural language processing and network science to analyze unstructured text data and produce actionable knowledge for drug repurposing.

Second, researchers systematically review scientific publications to discover and synthesize the existing research evidence related to health data collection, human-data interaction and multimodal data integration for clinical insight. Kumar and Mostafa (2020) scrutinize articles retrieved from four scientific databases and the gray literature, which are recommended by 17 international experts. Their literature review discloses a research landscape regarding the EHRs development and usage in the low- and middle-income countries. Zou *et al.* (2020) thoroughly survey the studies classified as Healthcare-related Internet of Things (H-IoT) and closely examine how people engage with the H-IoT data. Yu and Lu (2020) conduct a comprehensive analysis using research papers related to human gliomas, which are retrieved from both PubMed and CNKI databases. Key factors affecting survival and prognosis are extracted and screened, and clinical evidence and guidance for gliomas treatment are summarized accordingly.

Third, researchers apply state-of-the-art machine learning approaches to clinical data modeling for disease detection and diagnosis. Two studies use health data analytics to facilitate autism diagnosis. Zhao and Lu (2020) adopt VGG19 and ResNet18 algorithm models of deep convolution neural network to analyze facial expression data in order to improve the accuracy of autism diagnosis. Chen *et al.* (2020a) propose a novel diagnostic approach to autism spectrum disorder by integrating, classifying and stratifying multiple types of data (i.e. clinical assessment, neuroimaging, gene mutation and expression, and response signal data). Liu *et al.* (2020a) suggest a majority voting feature selection method to evaluate the most important clinical and demographic indicators for diabetes diagnosis. This method combines Lasso Regression with Akaike information criterion (LR-AIC), Lasso Regression with Bayesian information criterion (LR-BIC) and Random Forest, and selects the only indicators with excellent predictive performance to build the tree-based ensemble models.

Last but not least, researchers explore consumers' online health information searching behavior. The study by Chen *et al.* (2020b) users' search sequences, query reformulation and conceptual changes while they are seeking health information on the Internet. Zhao *et al.* (2020) select one of the Chinese online health communities to investigate the patient-generated content on social media in terms of health information needs, attitudes and psychological reactions to hypertension as well as social support exchange.

In summary, this special issue reflects the synergy of interdisciplinary collaboration and expertise, contributing to both library and information science and health sciences. Notably, this special issue serves as a unique example of enriching and accelerating multidisciplinary science by integrating various perspectives, theories, models and concepts. For library and information professionals, several potential service areas are indicated, such as systematic literature searching and screening, citation data analytics, health-related data curation and online health data literacy. We are excited to deliver this special issue to the *Library Hi Tech* communities and believe a broad spectrum of audience will value the included original and innovative studies.

Dan Wu
School of Information Management, Wuhan University, Wuhan, China, and
Fei Yu
University of North Carolina at Chapel Hill, Carrboro, North Carolina, USA

References

Chen, T., Froehlich, T., Li, T. and Lu, L. (2020a), "Big data approaches to develop a comprehensive and accurate tool aimed at improving autism spectrum disorder diagnosis and subtype stratification", *Library Hi Tech*, Vol. 38 No. 2.

Chen, Y., Zhao, Y. and Wang, Z. (2020b), "Understanding online health information consumers' search as a learning process: search platform sequences, query reformulation and conceptual changes", *Library Hi Tech*, Vol. 38 No. 2.

Cohen, I.G., Amarasingham, R., Shah, A., Xie, B. and Lo, B. (2014), "The legal and ethical concerns that arise from using complex predictive analytics in health care", *Health Affairs*, Vol. 33 No. 7, pp. 1139-1147.

Devaraj, S., Ow, T.T. and Kohli, R. (2013), "Examining the impact of information technology and patient flow on healthcare performance: a theory of swift and even flow (TSEF) perspective", *Journal of Operations Management*, Vol. 31 No. 4, pp. 181-192.

Fichman, R.G., Kohli, R., Krishnan, R. and Kane, G.C. (2011), "Editorial overview—the role of information systems in healthcare: current research and future trends", *Information Systems Research*, Vol. 22 No. 3, pp. 419-428.

Kumar, M. and Mostafa, J. (2020), "Electronic health records for better health in the lower- and middle-income countries: a landscape study", *Library Hi Tech*, Vol. 38 No. 2.

Liu, J., Wang, L., Zhang, L., Zhang, Z. and Zhang, S. (2020a), "Predictive analytics for blood glucose concentration: an empirical study using tree-based ensemble approach", *Library Hi Tech*, Vol. 38 No. 2.

Liu, J., Yu, F. and Song, L. (2020b), "A systematic investigation on the research publications that used the medical expenditure panel survey (MEPS) data through a bibliometrics approach", *Library Hi Tech*, Vol. 38 No. 2.

Moradi, S. and Dokhani, F. (2020), "Using the quadruple helix model for evaluation of health science researches: case study of D8 countries", *Library Hi Tech*, Vol. 38 No. 2.

Patil, S., Lu, H., Saunders, C.L., Potoglou, D. and Robinson, N. (2016), "Public preferences for electronic health data storage, access, and sharing—evidence from a pan-European survey", *Journal of the American Medical Informatics Association*, Vol. 23 No. 6, pp. 1096-1106.

Sharma, A., Harrington, R.A., McClellan, M.B., Turakhia, M.P., Eapen, Z.J. and etc (2018), "Using digital health technology to better generate evidence and deliver evidence-based care", *Journal of the American College of Cardiology*, Vol. 71 No. 23, pp. 2680-2690.

van Panhuis, W.G., Cross, A. and Burke, D.S. (2018), "Project Tycho 2.0: a repository to improve the integration and reuse of data for global population health", *Journal of the American Medical Informatics Association*, Vol. 25 No. 12, pp. 1608-1617.

Wu, P.Y., Cheng, C.W., Kaddi, C.D., Venugopalan, J., Hoffman, R. and Wang, M.D. (2016), "Omic and electronic health record big data analytics for precision medicine", *IEEE Transactions on Biomedical Engineering*, Vol. 64 No. 2, pp. 263-273.

Yu, Z. and Lu, L. (2020), "Comprehensive analysis of potential prognostic biomarker in gliomas", *Library Hi Tech*, Vol. 38 No. 2.

Zhao, W. and Lu, L. (2020), "Research and development of autism diagnosis information system based on deep convolution neural network and facial expression data", *Library Hi Tech*, Vol. 38 No. 2.

Zhao, D., Zhang, Q. and Ma, F. (2020), "Communication that changes lives: an exploratory research on a Chinese online hypertension community", *Library Hi Tech*, Vol. 38 No. 2.

Zhu, Y., Jung, W., Wang, F. and Che, C. (2020), "Drug repurposing against Parkinson's disease by text mining scientific literature", *Library Hi Tech*, Vol. 38 No. 2.

Zou, N., Liang, S. and He, D. (2020), "Issues and challenges of user and data interaction in healthcare-related IoT: a systematic review", *Library Hi Tech*, Vol. 38 No. 2.