SciencePG
Science Publishing Group

Big Data and Data Science Education in Traditional Chinese Medicine Informatics: Challenges and Opportunities

Zhixian Yi[1, *], Hong Nie[2], Cao Dong[3], Jianhua Jiang[4], Fu Su[5], Zhenghui Wang[6], Hui Wang[7], Hekun Zeng[2]

[1]School of Information and Communication Studies, Faculty of Arts and Education, Charles Sturt University, Wagga, Australia

[2]Guangdong Province Key Laboratory of Pharmacodynamics Constituents of TCM and New Drugs Research, College of Pharmacy, Jinan University, Guangzhou, China

[3]School of Medical Information Engineering, Guangzhou University of Chinese Medicine, Guangzhou, China

[4]Department of Data Science, Jilin University of Finance and Economics, Changchun, China

[5]Department of Management Science, School of Management, Guizhou University of Commerce, Guiyang, China

[6]School of Literature and Law, East China University of Technology, Nanchang, China

[7]Foreversafe Cloud Technology Co., Ltd, Shenzhen, China

Email address:
gyi@csu.edu.au (Zhixian Yi)
[*]Corresponding author

Abstract: With the rapid changes of scientific technologies, the explosion of big data and the current COVID-19 pandemic, the internal and external environments of traditional Chinese medicine (TCM) informatics are changing rapidly and the development of TCM informatics faces numerous challenges such as digital transformation. To meet the challenges, there is a strong need to apply big data and big data analytics and technologies to TCM informatics and to incorporate data science education into the programs of TCM informatics. The purpose of this paper is to identify the key challenges facing programs of TCM informatics, and to discuss the opportunities in data science education in TCM programs. The key identified challenges are how to enable students to have a good mastery of knowledge, skills and competencies of big data analytics and technologies and how to incorporate big data analytics and technologies and the key aspects of data science education into the curricula of TCM informatics. The main opportunities are that big data and data science education can assist in meeting the challenges facing TCM informatics and offer the analytics, technologies and applications needed to improve the learning, teaching and research outcomes and standards of TCM informatics and to review, revise and constantly update the teaching contents and curriculum knowledge system of TCM informatics. This paper will be useful and helpful to the programs of TCM informatics in their curriculum design as well as the application of big data analytics and technologies to teaching, learning and research, and to faculty, students, researchers and practitioners.

Keywords: Big Data, Big Data Analytics, Data Science, Data Science Education, Traditional Chinese Medicine Informatics, Challenges, Opportunities

1. Introduction

With the rapid changes of scientific technologies, the explosion of big data and the current COVID-19 pandemic, the internal and external environments of traditional Chinese medicine (TCM) informatics are changing rapidly and the development of TCM informatics faces numerous challenges such as digital transformation. To meet the challenges, there is a strong need to apply big data and big data analytics and technologies to TCM informatics and to incorporate data science education into the programs of TCM informatics. The

purpose of this paper is to identify the key challenges facing programs of TCM informatics, and to discuss the opportunities in data science education in the TCM programs.

2. Traditional Chinese Medicine Informatics

What is traditional Chinese medicine? It is "a complex and holistic system of medical practice with its own philosophy, diagnosis, treatment systems and pharmacology", takes into account the human body regarding "its own natural, physical and social environment", and "involves physical therapy (nonmedication) using acupuncture, moxibustion and related disciplines" which refer to "Tuina Massage and Qi Gong and chemical therapy using Chinese medicinal materials (CMM) of animal, mineral and plant origin in the form of decoctions of combined CMM or related proprietary products" in practical ways [1]. With the rapid developments of science and technology in the 21st century and information science, and especially "the integration of life sciences, information science, complexity science, and systems sciences", many new disciplines such as TCM informatics have come into being [2].

What is TCM informatics? It is "a cross discipline of TCM and informatics" and concentrates on "TCM information", "the motion law of TCM information" and "TCM informatics methodology" [3]. TCM informatics is a combination of information science and traditional Chinese medicine which enables traditional Chinese medicine to use modern information technologies "to extract data, generate information, establish relationships, form knowledge, make full use of the knowledge of traditional Chinese medicine, and quicken the steps of promoting traditional Chinese medicine to the whole world" [4].

By the end of 2008, there were 23 colleges and universities of traditional Chinese medicine and the majors of TCM information management and information system were offered by 6 of them [5]. Currently, "there are 26 TCM colleges and universities and 11 of them have offered the major of TCM informatics" [6]. Medicine informatics made great progress in the past decades, and despite that, there still existed major problems such as "more required courses, fewer elective courses, and single training mode" based on a preliminary analysis of the status of programs of information management and information system offered by Chinese medicine colleges and universities [5]. Then, Chinese medicine colleges and universities gradually established their majors of information management and information system of medicine and offered many TCM informatics related courses such as "Chinese medicine literature information access and use "," Chinese medicine literature information retrieval "," public health network information resources "," medical network resource utilization "," modern library use ", "medical information system ", " multimedia technology in medicine application", and "medical software design" [4]. The majors

of TCM information management and information system were offered by most Chinese medicine colleges and universities, medicine informatics course as the fundamental theory course was provided in the courses for all majors, and TCM informatics course was ignored [7]. With the rapid developments of science, technology and society, and the explosion of big data, colleges and universities of traditional Chinese medicine should put an emphasis on the application of big data and big data analytics and technologies to TCM informatics, and the provision of data science education in the programs of TCM informatics in order to address the challenges such as digital transformation that traditional Chinese medicine faces and to cultivate high-quality graduates of traditional Chinese medicine.

3. Big Data and Big Data Analytics

Big data and big data analytics have become the hot topics and how to apply big data and big data analytics and technologies to TCM informatics as well as traditional Chinese medicine has been the new research area. What is big data? There are numerous definitions of big data. Big data refers to "large volumes of high velocity, complex, and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management and analysis of the information" [8]. It is characterized as "a collection of data elements whose size, speed, type, and/or complexity require one to seek, adopt, and invent new hardware and software mechanisms in order to successfully store, analyse, and visualize the data" [9]. The features of big data are "volume, velocity, variety, value, variability and veracity", and data in healthcare includes "genomic data, clinical data, clinical notes, behavior data, patient sentiment data, health publication and clinical reference data, and other important data" [10]. The characteristics of TCM big data include "continuous circular loop, dynamic processing, overall system and value" [11]. Big data provides a variety of benefits including "disease prevention, reduced medical errors and the right care at the right time and better medical outcomes" [10]. One of the benefits of big data is "providing advantages to health informatics", and the application areas of big data include "genomics analytics, flu outbreak prediction and control, clinical outcome analytics, fraud detection and prevention, medical device design and manufacturing, personalized patient care, e-consultation and tele-diagnosis, pharmaceuticals and medicine, medical education, and smart health and wellbeing" [10].

Big data analytics is defined as "the process of collecting, organizing, analyzing large data sets to discover different patterns and other useful information", and includes four types: "descriptive analytics; diagnostic analytics; predictive analytics and prescriptive analytics" [12]. The field of big data analytics and applications is promising because it can offer new insights through the analysis of very large data sets and not only make outcomes better but also make costs lower [13]. There are many advantages of big data analytics. For example,

big data analytics as "asset of advanced technologies" can be "designed to work with large volumes of heterogeneous data leading to improved performance via data driven decision making", and the large amount of data can be analyzed and managed using the analytical tools [14].

4. Data Science Education in the TCM Informatics

What is data science? There are many definitions of data science. Data science is defined as "the extraction of actionable knowledge directly from data through a process of discovery, hypothesis, and analytical hypothesis analysis" [15]. According to Song and Zhu [16], data science is "the discipline that makes sense of big data". Data science is the study of tools, models, processes, techniques, analytics and technologies that can make data converted into knowledge, generate the value, and discover new knowledge via the collection, storage, transformation, visualization, analysis and management of data. To meet big data challenges, to adjust to the changes of social, economic and technological developments and to satisfy the needs for qualified graduates, universities in China have continuously offer degree programs and courses in data science and data science has been integrated into the programs and disciplines such as the discipline of library and information studies [17]. To cultivate qualified graduates in the age of big data, TCM colleges and universities should provide data science education in the teaching and research of TCM informatics.

5. Challenges and Opportunities

In the digital age, information and communication technologies, computing technologies and big data technologies develop very rapidly. Nowadays, the majors and curriculum design of TCM informatics are facing numerous challenges in data science education. The main challenges are as follows:

(1) How to incorporate big data analytics and technologies and the key aspects of data science education into the curricula of TCM informatics;
(2) How to enable students to have a good mastery of knowledge, skills and competencies of big data technologies and analytics;
(3) "How to educate students to create tools/value/information/knowledge from big data by utilizing data science skills" [16];
(4) "How to expand curriculum to newer topics such as intelligent augmentation and cognitive computing" [16]; and
(5) "How to put these emerging data management topics into curriculum" [16].

With regard to the field of Chinese medicine, traditional Chinese medicine, Tibetan medicine, Miao medicine and other ethnic medicine have a real big data at multiple levels, and how to make such a large amount of big data converted into knowledge, generate the value, and how to discover new knowledge via the collection, storage, transformation, visualization, analysis and management of data are huge challenges and good opportunities [18]. The era of big data has brought far-reaching social changes, placed higher demands for TCM colleges and universities to cultivate qualified graduates, and offered new development opportunities in the learning, teaching and research of TCM informatics [19]. The key opportunities are to:

(1) Review and revise the teaching contents of TCM informatics,
(2) Constantly update the curriculum knowledge system of TCM informatics [8],
(3) Apply the big data analytics and technologies and data science education to the teaching and research of TCM informatics,
(4) Incorporate big data analytics and technologies and the key aspects of data science education into the curricula of TCM informatics,
(5) "Address student needs with customised modules, assignments, feedback and learning trees in the curriculum that will promote better and richer learning" [20], and
(6) Develop the TCM as well as TCM informatics.

The features of traditional Chinese medicine information are similar to those of big data, and the application of big data and big data analytics and technologies will help with the learning, teaching and research on TCM information, and it is necessary to make full use of the opportunities to quicken the development steps of TCM [21]. For instance, there exist the similarities between characteristics of big data and TCM informatics, and these similarities reflect in the four distinct characteristics of TCM information: "epistemological information, phenomenon information, overall information and time information" which "overlap with the three characteristics of 'big data', namely integrity data, fuzzy data and correlation data, so the advent of the information age characterized by 'big data' is bound to create good opportunities for the development of TCM informatics" [3].

In the age of big data, "the traditional data processing mode has been unable to adapt to the fast pace of big data", and therefore it is time to utilize big data analytics, technologies and applications to discover new valuable insights and knowledge of traditional Chinese medicine and to find the best solutions to the problems of TCM, and the combination of big data analytics, technologies and applications with the fields of traditional Chinese medicine and TCM informatics will successfully promote the development of Chinese medicine industry as well as TCM informatics [22]. The combination of traditional Chinese medicine with big data and big data analytics and technologies will inevitably promote the development of Chinese medicine and the application of big data and big data analytics and technologies to the fields of Chinese medicine includes "TCM informatics, TCM prevention diseases, Chinese medicine diagnosis of diseases, Chinese medicine treatment of diseases, Chinese medicine knowledge discovery, and development of Chinese medicine" [23].

In terms of the curricula and research of TCM informatics, there are many problems that need to be investigated and solved. For example, how to incorporate big data analytics, technologies and applications, and the key aspects of data science education into the curricula of TCM informatics in a successful way is the key problem that needs to be solved. According to Liu [24], "graduate courses should be set up to attract professional researchers from the fields of TCM and data study to cultivate interdisciplinary talents, who not only understand TCM but are also well versed in clinical practice, data collection, and analysis". As far as the research on TCM informatics is concerned, it needs "to solve the bottleneck problem impeding the development of traditional Chinese medicine, promote the inheritance and innovation development of theory and methodology of TCM, and promote the improvement of the clinical efficacy by improving the ability of utilization of TCM information" [25]. Research conducted in TCM informatics can focus on "data acquisition, retrieval, storage, analytics employing data mining technologies, and so on" [26], and "the determinants of the TCM utilizations" [27].

6. Conclusion

With the current explosion of big data and rapid developments of big data and big data analytics, technologies and applications, TCM informatics has met numerous challenges in data science education and good opportunities. The key identified challenges are how to enable students to have a good mastery of knowledge, skills and competencies of big data analytics and technologies and how to incorporate big data analytics, technologies and applications, and the key aspects of data science education into the curricula of TCM informatics. The main opportunities are that big data and data science education can assist in meeting the challenges facing TCM informatics and offer the analytics, technologies and applications needed to improve the learning, teaching and research outcomes and standards of TCM informatics and to review, revise and constantly update the teaching contents and curriculum knowledge system of TCM informatics. This paper will be useful and helpful to the programs of TCM informatics in their curriculum design as well as the application of big data analytics and technologies to teaching, learning and research, and to faculty, students, researchers and practitioners. Future research will focus on the exploration of feasible strategies for addressing the identified challenges such as how to enable students to master knowledge, skills and competencies of big data analytics and technologies in a successful way, and the application of "new technologies such as deep learning in the field of Chinese medicine data" and medical data mining technology [28, 29].

References

[1] Chan, K. (2005). Chinese Medicinal Materials and Their Interface with Western Medical Concepts. Journal of Ethnopharmacology, 96, 1–18.

[2] Xie, Q., Cui, M., Wu, Z. and Zhao, H. (2010). Traditional Chinese Medicine Information Digitalization Discussion. The Journal of Alternative and Complementary Medicine, 16, 1207-1209.

[3] Cui, M., Li, H. and Hu, X. (2014). Similarities between "big data" and traditional Chinese medicine information. Journal of Traditional Chinese Medicine, 34, 518-522.

[4] Liu, J. (2011). A study of the teaching contents of traditional Chinese medicine informatics. Journal of Liaoning University of TCM, 13, 22-23.

[5] Li, X. (2009). A study of the curriculum design of information management specialization in Chinese medicine colleges and universities. Northwest Medical Education, 17, 473, 490.

[6] Huang, L., & Zou, Y. (2020). The development path and system construction of traditional Chinese medicine informatics. Journal of Medical Informatics, 41 (12), 7-9.

[7] Du, X., Qing, W., & Liu, J. (2014). The study of the construction of knowledge system of traditional Chinese medicine informatics. Lishizhen Medicine and Materia Medica Research, 25, 2259-2260.

[8] Institute for Health Technology Transformation. (2013). Transforming Health Care through Big Data Strategies for Leveraging Big Data in the Health Care Industry. Retrieved from http://ihealthtran.com/wordpress/2013/03/iht%C2%B2-releases-bigdata-research-reportdownload-today/.

[9] Belle, A., Thiagarajan, R., Reza Soroushmehr, S. M., Navidi, F., Beard, D. A. and Najarian, K. (2015). Big data analytics in healthcare. BioMed Research International, 1-16.

[10] Wang, L., and Alexander, C. A. (2015). Big data in medical applications and health care. Current Research in Medicine, 6, 1-8.

[11] Cheng, X., Wen, C., Xu, Q., Hu, Y., & Zhang, X. (2020). Development demand of artificial intelligence based on traditional Chinese medicine to study on the typical of the big data of traditional Chinese medicine. Modernization of Traditional Chinese Medicine and Materia Medica-World Science and Technology, 22 (4), 1243-1248.

[12] Riahi, Y., and Riahi, S. (2018). Big data and big data analytics: Concepts, types and technologies. International Journal of Research and Engineering, 5 (9), 524-528.

[13] Raghupathi, W. and Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. Health Information Science and Systems, 2, 1-10.

[14] Gill, J., & Singh, S. (2015). Enormous possibilities in big data: Trends and applications. Asian Journal of Computer Science and Technology, 4, 23-26.

[15] Fox, G. (2014). Data science curriculum. Retrieved from https://rd-alliance.org/system/files/filedepot/240/1-RDA-Data Science-21Sept-2014.pdf.

[16] Song, I. and Zhu, Y. (2017). Big data and data science: Opportunities and challenges of iSchools. Journal of Data and Information Science, 2, 1-18.

[17] Zhang, J., Fu, A., Wang, H., & Yin, S. (2017). The development of data scienc education in China from the LIS perspective. International Journal of Librarianship, 2 (2), 3-17.

[18] Pan, W., Cheng, T., Niu, C. and Zhao, X. (2014). Application of traditional Chinese medicine information in the age of big data. Chinese Journal of Library and Information Science for Traditional Chinese Medicine, 38, 2-4.

[19] Chang, K., Ma, H., Sun, J. and Deng, W. (2015). The cultivation of professionals of traditional Chinese medicine in big data age. Journal of Traditional Chinese Medicine Management, 23, 27-29.

[20] Daniel, B. (2014). Big data and analytics in higher education: Opportunities and challenges. British Journal of Educational Technology, 46, 1-17.

[21] Wu, J. (2015). The similarities between big data and the characteristics of traditional Chinese medicine informatics. Journal of Shangdong University of TCM, 39, 406-407.

[22] Xie, Y. (2015). Opportunity and challenge for traditional Chinese medicine in the era of big data. Chinese Journal of Medical Library and Information Science, 24, 33-35.

[23] Wu, Q., Hu, Y., Cha, Y., Zhu, J., Xue, H., & Mao, X. (2016). The application status of big data in the field of Chinese medicine. Journal of Guangxi University of Chinese Medicine, 19, 153-156.

[24] Liu, B. (2014). Utilizing big data to build personalized technology and system of diagnosis and treatment in traditional Chinese medicine. Frontier of Medicine, 8, 272–278.

[25] Yu, Q., Cui, M., Li, H., Hu, X., Zhang, H., Yang, S. and Gao, B. (2014). Discussion about characteristics of information science of traditional Chinese medicine from the perspective of big data. China Digital Medicine, 9, 32-34.

[26] Herland, M., Khoshgoftaar, T. M. and Wald, R. (2014). A review of data mining using big data in health informatics. Journal of Big Data, 1, 1-35.

[27] Xu, J. and Yang, Y. (2009). Traditional Chinese medicine in the Chinese health care system. Health Policy, 90, 133–139.

[28] Li, Y., Xia, Y., Zhang, J., Zhang, Y., Zhao, L., Dou, Z., & Han, D. (2021). Research progress and current status on application of TCM big data in real world. CJTCMP, 36 (6), 3471-3474.

[29] Yang, J., Li, Y., Liu, Q., Li, L., Feng, A., Wang, T., Zheng, S., Xu, A., & Lyu, J. (2020). Brief introduction of medical database and data mining technology in big data era. Journal of Evidence-Based Medicine, 13, 57–69.