Received: 25 April 2022 | Accepted: 25 April 2022

DOI: 10.1002/sres.2878

RESEARCH PAPER

SYSTEMS and BEHAVIORAL RESEARCH SCIENCE

WILEY

Study on the interaction between big data and artificial intelligence

Jin Li[1] | **Ziwei Ye**[2] | **Caiming Zhang**[3]

[1]School of Journalism and Communication, University of Chinese Academy of Social Sciences, Beijing, China

[2]Department of Information Technology & Decision Sciences, Old Dominion University, Norfolk, Virginia, USA

[3]China University of Labor Relations, Beijing, China

Correspondence
Caiming Zhang, China University of Labor Relations, No.45, Zengguang Road, Haidian District, Beijing 100048, China.
Email: zhangcaiming@culr.edu.cn

Jin Li, School of Journalism and Communication, University of Chinese Academy of Social Sciences, Beijing 100732, China.
Email: echojinli@icloud.com

Abstract

The explosive growth of information has rapidly ushered people into the era of big data. Due to the large volume, high variety, and rapid velocity characteristics of big data, most traditional data mining methods developed for a centralized data analysis process cannot be applied directly. AI leverages machine intelligence to provide insights, automation, and new methods to interact with data, thereby promoting data literacy throughout the organization. Based on the theory of Big Data Cycle, this paper discusses the relationship between big data and AI and how they interact and influence each other. It adopts the integrative review research method to screen latest literature and summarizes the role of AI in different phases of big data cycle. We also provide an insight into the applications of big data and AI in three different areas, that is, social network, health care, and finance.

KEYWORDS

artificial intelligence, big data, integrative review, interaction effect

1 | INTRODUCTION

"Perhaps the most dramatic outcome of the digital revolution is the amount of data that's now collected and analyzed." Debanjan Saha, Google Cloud's former vice president of data analytics, exclaims at the beginning of his paper (Saha, 2020). International Data Corporation (IDC) calculates that in 2010 the world created about 2 zetta bytes (ZB) of digital information. That sounds impressive, until we consider that the World Economic Forum figures there will be 46.3 ZB collected in 2025 (Desjardins, 2019). If those data were put into 1 GB thumb drives that were then laid end to end, it would make a line that could stretch across 4232 million football fields, which reflects a significant characteristic of big data, the large volume. The specific techniques of big data date back to at least the 1990s, but the term entered business discourse through a 2001 Gartner report that defined big data as the "3Vs": volume, velocity, and variety (Laney, 2001). More recent developments suggest that big data problems are identified by the "5V": volume (quantity of data), variety (data from different categories), velocity (rapid generation of new data), veracity (quality of the data), and value (in the data) (Chen et al., 2020; Terzo et al., 2013). Therefore, new technologies are strongly needed since most of the traditional data mining methods or data analytics developed for a centralized data analysis process may not be able to be applied directly to big data (Li et al., 2007).

Artificial intelligence (AI) refers to a set of technologies that enable computers to simulate human intelligence. It uses advanced machine learning, pattern recognition, and data mining techniques to build models for information refinement, information processing, and value-added services (Kaplan & Haenlein, 2019). Popular AI techniques include machine learning methods for structured data, such as the classical support vector machine, random forests, neural network, and deep learning, as well as natural language processing for unstructured data. Practical examples of AI include speech recognition, such directing virtual assistants like Alexa to perform tasks, image recognition for identification, and autonomous driving. While in the context of big data analytics, AI leverages machine intelligence to provide insights, automation, and new methods to interact with data, thereby promoting data literacy throughout the organization. Big data, AI, cloud infrastructure, and active intelligence have contributed to the emergence of the third generation of business intelligence, in which big data and AI play the dominant parts.

Our research investigates the following question: *What is the relationship between big data and AI and how do they interact and influence each other?* To answer this research question, we adopt the integrative review research method recommended by Snyder (2019), using big data and AI as search terms simultaneously or respectively, extract relevant literature from mainstream databases, such as Web of Science and IEEE Xplore, and check the latest research findings in an attempt to give overview of the interaction between big data and AI. More specifically, we explore how AI technologies can help in various stages of the big data cycle and the latest research in the intersection of the two.

The remainder of this paper proceeds as follows. First, it gives some basic concepts and brief history of big data and AI to help explore the relationship between these two terms. Then we focus on the role of AI in all phases of the Big Data Cycle and novel research findings, which are used to shed light on the interplay between AI and big data. The next section summarizes the updated findings related to applications of big data and AI in various sections. The paper concludes with contributions to theory, limitations of this study, and suggestions for future work.

2 | LITERATURE REVIEW

2.1 | Relationship between big data and AI

Unlike traditional data, big data is defined by most researchers and data scientists by the following three main characters, which is called 3Vs and subsequently developed into 5Vs (Furht & Villanustre, 2016). The 5Vs are as follows:

- **Volume**: Big volumes of digital data are generated continuously from millions of devices and applications (ICTs, IoT, social networks, smart phones, sensors, logs, etc.). According to the study of McAfee et al. (2012), it is estimated that about 2.5 exa bytes were generated each day in 2012, ten years ago. This amount is doubling every 40 months approximately. While World Economic Forum predict that there will be 46.3 ZB collected in 2025 (Desjardins, 2019). This indicates that the growth of data is exponential, not linear.
- **Velocity**: Data is generated rapidly and should be processed quickly in order to extract useful information and relevant knowledge.
- **Variety:** Big data is generated from a wide range of sources and in a variety of formats (e.g., words, voices, videos, and logs). Large data sets include both structured and unstructured data, as well as data that is public or private, local or remote, disclosed or confidential, complete or incomplete, and so on.
- **Variability:** Aside from increasing data velocities and variety, data flows can be highly inconsistent with periodic peaks. For instance, the number of user comments on social media like Twitter, can peak at certain times of the day, such as in the evening after people leave work or when breaking news is generated.
- **Value:** We need to consider whether the information extracted from the large amount of data has commercial value or research significance and for which type of users. And it is also necessary to connect and correlate relationships, hierarchies, and multiple data linkages to create and enhance the value of data.

These characteristics of big data require people to use faster, efficient, and automated methods or technologies in processing and analysing data. Big data analytics is the use of processes and technologies to combine and analyse massive datasets with the goal of identifying patterns, extracting knowledge, and developing actionable insights.

AI, as the core technology of simulating human intelligence, meets these requirements. The origin of AI can be traced back to the 1950s. At the Dartmouth conference held at Dartmouth College in 1956, researchers first put forward the term AI and defined its task (Mindell & Mindell, 2002). Regarding the definition of AI, Kaplan and Haenlein (2019) argue that AI is a systematic ability to correctly interpret external data, learn from it, and adapt flexibly to accomplish specific goals and tasks through learning. Lu (2019) states that any theory, method, and technology that helps machines (especially computers) analyse, simulate, utilize, and explore human thought processes and behaviours can be considered as AI. The processes and behaviours of human thought are also presented in the form of data. So in a broad sense, data, especially big data, is the fuel that powers the evolution of AI's decision making. AI requires a massive scale of data to learn and improve the accuracy of models, as well as decision-making processes.

Big data and AI have a synergistic relationship. Big data analytics leverages AI for better data analysis. In turn, AI requires big data to boost decision-making processes. With this convergence, data analysts and researchers can more easily leverage advanced analytics capabilities such as augmented or predictive analytics, and surface actionable insights from massive data stores, both in industry and academia. Companies can empower their customers with the intuitive tools and robust technologies they need to extract high-value insights from data using big data AI powered analytics, fostering data literacy across the organization while reaping the benefits of becoming a truly data-driven organization, which definitely improves the business performance and efficiency.

2.2 | AI's role in the big data cycle

So, how do big data and AI work together? We need to focus on the role of AI in all phases of the Big Data Cycle and novel research findings, which are used to shed light on the interplay between AI and big data.

The Big Data Cycle is the standard collection of core tasks that involve the collection, storage, and application of data analytics, with which much of everyday society is associated (Surya, 2015). Big data is a field concerned with methods of processing, evaluating, and efficiently obtaining information from, and otherwise interacting with, data points that are too complex and multi-dimensional to be handled in a traditional spreadsheet software system. The procedure is a basic command and control procedure for centrally managing large amounts of data from various sources. With iteration, the cycles typically flow from left to right, that is, from data management to risk management. The phases of the big data cycle include data management, pattern management, context management, decision management, action management, goal management, and risk management.

2.2.1 | AI in data management

Data management is a process that involves collecting, verifying, storing, securing, processing, and modifying data to ensure that it is easily accessible, reliable, and timely for multiple users. Big data collection and storage are the main technical bottlenecks and active research topics in multiple communities.

AI, especially machine learning as an accurate and scalable data collection technique, can help boost the efficiency of collection process from three different perspectives (Roh et al., 2021). First, if the goal of data collection is to search and share new datasets, the data acquisition techniques such as collaborative analysis, data lake, crowdsourcing, and generative adversarial networks can be used to discover and generate datasets. For example, Kandogan et al. (2015) introduced an open, social, and collaborative data analysis platform called LabBook to reduce friction and accelerate discovery. It can assist users in leveraging each other's knowledge and experience to find the data, tools, and collaborators required to integrate, visualize, and analyse data. Crowdsourcing provides a collaborative environment to gather and pre-process big data. From the earliest, most popular Amazon Mechanical Turk (MTurk) to recent crowdsourcing platforms such as ALFRED and TurkPrime.com, AI is gradually solving the problems of single source, inefficiency and unclear process decomposition in the data collection process (Litman et al., 2017). Second, when the datasets are ready, kinds of data labelling techniques will be used to label the individual examples, such as graph-based label propagation and active learning (Ricci et al., 2015). Third, AI can also help to improve the quality of existing data by data cleaning and relabelling. Baylor et al. (2017) presented a production machine learning platform called TensorFlow Extended (TFX), a TensorFlow-based general-purpose machine learning platform implemented at Google. The platform can reduce the time to production from months to weeks while maintaining platform stability and minimizing disruptions. More recently, Wang et al. (2019) used an angle-based outlier detection method to obtain the training data of the cleaning model, which is then established through support vector machine. Leveraging AI, such as IoT networks, cloud and other latest technologies for big data storage is also becoming a new popular trend. In Yang et al.'s (2019) study, a privacy-preserving smart IoT-based healthcare big data storage system with self-adaptive access control is proposed, which supports smart cross-domain data sharing, self-adaptive access control and smart deduplication.

2.2.2 | AI in pattern management

Pattern Management refers to find frequent patterns in data, including frequent item sets, frequent subsequences, and frequent substructures. Mining frequent patterns reveals interesting associations and correlations within data. Companies can use frequent items to develop marketing strategies for different kinds of customers. For example, it is better to reposition items to put milk and bread in the same shelf, which are frequently bought together in grocery stores by many customers. All types of businesses must be on the lookout for "patterns of interest" that will alert them to potential events and, later, confirm the presence of external risks, allowing them to act independently or implement measures that have already been "identified and processed" for implementation.

The continuous optimization of AI algorithms has reduced the runtime and memory space consumed for pattern mining and overcame the difficulties caused by inefficient processes and complex data structures, making it possible to mine frequent patterns in large amounts of data. Braun et al. (2019) reviewed many serial, distributed, parallel, and MapReduce-based (Hadook-based and Spark-based) big data mining algorithms such as hyperlinked array based frequent pattern mining algorithms and bitwise frequent algorithms in their latest paper. Yoon and Kim (2020) proposed a pattern adaptive prefetcher with DRAM-based dual buffers for the memory-disk integrated system (MDIS) that enable flexible adaptation to the unpredictable and inconsistent workload patterns.

2.2.3 | AI in context management

The value of the collected data also varies depending on the context in which it is interpreted and the results to which this data may be applied. Hellerstein et al. (2017) introduce three key sources of information—the ABCs of Data Context—to emphasize the conceptual shifts from traditional data to data context and as a complement to the "3Vs" of big data. The core information that describes how raw bits are interpreted for use is known as application context, which ranges from basic data description, that is, ontologies, encodings, and schemas, to statistical models and parameters, to user annotations. Behaviour of data context refers to information about how data was created and used over time, including logs of usage and data lineage, that is, upstream lineage (the data sets and code that led to the creation of a data object) and downstream lineage (products derived from this data object). Change of data context indicates information about the version history of data, code, and related information. By checking historical record of context change, data scientists could easily debug and enable auditing and counterfactual analysis.

AI plays a significant role in data context management by facilitating the complex computer processes and providing state-of-the-art data management systems. For instance, under the context of on-going pandemic, a mathematical model susceptible, infected and recovered (SIR) was implemented for classifying COVID-19 cases and neural network (RNN) with long-term short memory is enacted to predict the COVID-19 disease (Magesh et al., 2020). AI models can also learn subtle changes in data and context-specific distinctions to track the development of interpretations of data in different environments, which deals exactly with the problem of data variability in big data analysis.

2.2.4 | AI in decision management

Decision management contains the processes of designing, developing, and maintaining the automatic decision-making mechanisms, which is typically characterized by uncertainty, complexity, and equivocality.

AI technologies can assist humans with different aspects of decision making, especially in the context of massive data. Addressing uncertainty in the decision-making process, AI provides access to real-time information for reference, such as anomaly detection in finance industry. For instance, fuzzy system is considered as an effective tool with access to uncertain information by fuzzy representations. Son et al. (2020) proposed a new representation of intuitionistic fuzzy systems based on complex numbers (IFS-C) in the polar form to overcome the restriction of order relations of datasets. Complex decision management situations demand the processing of massive information at a fast speed, which exceeds the speed required for human decision making. Combined with big data, AI algorithmic decision making provides new opportunities for tackling with complexity and gives more effective ways of comprehensive data analytics. For example, through causal loops, AI enables reduce the complexity of a problem by identifying causal connections and implying the acceptable cause of action among many possibilities (Marwala, 2015). Meanwhile, AI can provide some utilities that help decision-makers overcome equivocality and address relevant conflicting interests between stakeholders, customers, and policy makers.

2.2.5 | AI in action management

Action management aims to design project plans and orchestrate work activities to maximize the business value or satisfy leaders' requirements. In the process, action information of all humans, applications, machines, and technologies are tracked to enrich the big data pools for any further study or utility.

AI can track and evaluate the action records, then pretest future adapted behaviour prior to implementation, in conjunction with algorithms and models, providing quantitative assurance for the ideal outcome. To be specific, AI tools first analyse work activities and then try to associate different activity dots with the previous decision steps and AI could be merged into every measures or specific tasks that are conducted in the executive actions. Let us consider how marketers can leverage AI to smooth the action pace towards good marketing performance. In the stage of planning direction, objectives, and marketing support, the most common application of AI is in customer service (Campbell et al., 2020). AI use in customer service has expected to increase by 143% between 2019 and 2021, including applications such as chatbots and text and voice analysis (Salesforce, 2019). Chatbots are applied in customer service to deal with simple queries from customers, which can reduce service costs and improve customer satisfaction by immediate reply. Customers' questions are also recorded by the system for subsequent analysis and treatment or to provide precise marketing.

2.2.6 | AI in goal management

Goal management is the process of creating and tracking objectives in order to provide feedback and guidance, as well as to analyse factors of success and make recommendations to all assets for better performance. The interest in self-directed goal attainment is growing as more organizations turn to empowerment software, dynamic network web bots, and edge computing. In the research of Schwabe (2020), an AI powered web application demo was proposed, which aims to assist students with creation a homework and study schedule. The system uses key principles from Self-Regulated Learning and Cognitive Load Theory to translate the large, abstract problem of "creating a study schedule from scratch" into a structured, repeatable set of review tasks, then uses a constraint satisfaction AI agent to recommend a weekly schedule of activities that supports the student to achieve the specified goals.

2.2.7 | AI in risk management

Risks—defined as probable deviations from a planned condition—are fundamental components of all entrepreneurial activities and, as a result, a key aspect in business decision-making processes. Risk awareness, as well as risk assessment metrics and risk control and governance approaches, are crucial to an organization's performance and are commonly referred to as risk management.

To meet the challenges of risks deriving from massive various data, AI may benefit organization by identifying and responding to circumstances timely. Important trends and inconsistencies can be discovered in occurrences, trends, system logs, and personal feedback (such as social networking sites) for possible or impending risks. In the field of fintech, a research study showed that explainable AI model can be used in measuring the risks that arise when credit is borrowed employing peer to peer lending platforms (Bussmann et al., 2020). The obtained empirical evidence in the study turned out that the interpretability (explainability) of the results could be improved while increasing the predictive accuracy with respect to a standard logistic regression model.

As discussed above, AI plays a significant role in each phase of Big Data Cycle. Using natural language processing, AI can distinguish data categories, find possible links between datasets, and recognize expertise knowledge. It can be used to automate and speed up data preparation operations, such as data model building, as well as aid data exploration. It can recognize and correct common human error patterns, detecting and addressing potential data issues. It can also learn by observing how a user interacts with an analytics software, allowing it to quickly uncover unexpected insights from large datasets. In order to help users better interpret quantitative data sources, AI can learn minor distinctions in meaning or context-specific nuances. It can also notify users of abnormalities or unexpected trends in data, as well as actively monitor events and identify potential risks from system logs or social networking data.

2.3 | Big data and AI application

Big data and AI have numerous applications in various industries. Identifying the current applications and effects of the technologies is the best way to see the current and future state of big data and AI. We chose three hot research areas with rapid data growth, complex data structure forms, and significant business value to discuss the impact of the interaction between big data and AI, especially how AI can help organizations create benefits in the context of big data.

2.3.1 | Social network

With the rapid development of network technology, people have become connected through social networking sites like as Facebook, Instagram, LinkedIn, Twitter, and Tiktok in recent years. Many users regularly contribute posts or tweets to these social networking sites to share their activity with their friends or contacts (Li & Xu, 2020). This results in a large amount of social network data. AI techniques can extract intelligence from posted messages, qualify the user behaviours, and identify the social structure, which addresses the exploitation difficult due to the quantity of information. In Gadek et al.'s

(2018) work, they illustrate how to exploit AI on a very peculiar social network, named Ga-laxy2, by proposing an analysis of 1000 days of activity using NLP techniques to find the most interesting topics and to discover key actors. Then they proceed with a machine learning-based profiling of the user behaviours and introduce influence and cohesion scores for groups of users, which help characterization and evaluation. To address the problem that social network data often contains a large number of missing values and therefore cannot be predicted accurately, Lam and Hsiao (2019) proposed a neural network-based approach and adapted the generative adversarial network concept to generate the missing values. The generator could generate possible missing value values, whereas the discriminator could determine whether or not the Generator's output values are suitable for further training. Based on the theory of big public opinion generation and communication, Li (2022) c-onstructs the public opinion system and its main functions of big data employees based on data, including big data collection, data processing, data storage, data analysis, data visualization, and employee public opinion early warning.

2.3.2 | Healthcare

AI has also brought a paradigm shift in healthcare domain, powered by growing availability of healthcare data and rapid advancement of analytics techniques. Healthcare data is different form big data in other fields existed in but not limited to the form of demographics, medical notes, physical examinations, clinical laboratory and images, varying form structured data to unstructured data, which poses great challenges for data analysis. In healthcare applications, machine learning algorithms try to cluster patients' symptoms and features or deduce the probability of disease outcomes by analysing structured data such as imaging, genetic and electrophysiological (EP) data. While natural language processing (NLP) methods target at extracting information from unstructured data like medical notes, then convert text into structured medical data, which may be analysed by machine learning techniques. Major disease areas that use AI tools include cancer, neurology, and cardiology. Xie et al. (2021) proposed an AI-based computed tomography processing framework with an image compression module, an image denoising module, and an image segmentation module based on computed tomography images for surgical telementoring of Congenital Heart Disease (CHD), which may tackle the problem of high resource consumption on medical data transmission and storage.

2.3.3 | Finance

Finance big data (FBD) is quickly becoming one of the most promising areas of financial management and analysis. FBD has significantly changed the prediction models in financial firms. Many scholars argue that big data is accelerating the change of finance and business in ways that are difficult to quantify currently. With the help of AI, a new field of research is forming to investigate quantitative models and econometric methodologies for financial studies that can bridge the gap between empirical finance and data science. Applications that combine big data and AI mainly include fraud prevention, financial forecasting like AI stocks, and credit-souring and rating. Take credit card fraud prevention as an example. Fraud detection approaches that use supervised learning techniques rely on the notion that fraudulent tendencies may be learned from a study of historical transactions. However, when the sample size exceeds a certain range, that is, when the variability of the data increases, customer behaviour and fraudster capabilities evolve new fraud patterns, leading to a decrease in the accuracy of the prediction model. In this context, unsupervised learning techniques can help the fraud detection systems to find anomalies. Carcillo et al. (2021) provided a hybrid strategy for improving fraud detection accuracy by combining supervised and unsupervised techniques. In their experiments, unsupervised outlier scores, computed at different levels of granularity, are compared and tested on a real, annotated, credit card fraud detection dataset. Experimental results show that the combination is efficient and does indeed improve the accuracy of the detection.

3 | CONCLUSION

The interaction of big data and AI in academy and industry has always been a hot topic. In the context of big data, AI leverages machine intelligence to provide insights, automation, and new methods to interact with data, thereby promoting data literacy throughout the organization. First, this paper gives some basic concepts and brief history of big data and AI to help explore the relationship between these two terms. We found that big data and AI have a synergistic relationship. Big data analytics leverages AI for better data analysis. In turn, AI requires big data to boost decision-making processes. Second, we focus on the role of AI in all phases of the Big Data Cycle and novel research findings, which are used to shed light on the interplay between AI and big data. In data management phase, AI technology can help break through bottlenecks in the collection and storage of large amounts of data. And continuous optimization of AI algorithms has reduced the runtime and memory space consumed for pattern mining. From the aspect of context management, information systems empowered by AI can track and record all relevant information and operation traces of the behaviour for subsequent analysis and processing. AI technologies can also help address the uncertainty, complexity, and equivocality in decision management. Through tracked action data, AI helps pretest future adapted behaviour prior to implementation and provide quantitative assurance for the ideal outcome. To meet the challenges of risks deriving from all phases of big data cycle, AI may benefit organization by identifying and responding to circumstances timely. Third, this paper provides latest research findings related to applications of big data and AI in three different areas. We have used concrete examples to show how advances in AI and big data technologies have contributed to the development of different industries, providing ideas for future research.

The paper focus on revealing the interplay between AI and big data through the role AI plays in each stage of the big data cycle. Therefore, we do not present a comprehensive overview of all AI technology categories and their applications, and the argumentative cases are focused on literature published in the last 5 years, which is one of the limitations of this paper. Future research can expand the scope of the literature review or explore the application of each type of technology in big data analytics from the perspective of AI technology classification.

FUNDING INFORMATION

No funding available.

DATA AVAILABILITY STATEMENT

No datasets were generated or analysed during the current literature review study.

ORCID

Caiming Zhang https://orcid.org/0000-0002-6365-6221

REFERENCES

Baylor, D., Breck, E., Cheng, H.-T., Fiedel, N., Foo, C. Y., Haque, Z., Haykal, S., Ispir, M., Jain, V., & Koc, L. (2017). Tfx: A tensorflow-based production-scale machine learning platform. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining

Braun, P., Cuzzocrea, A., Leung, C. K., Pazdor, A. G. M., Souza, J., & Tanbeer, S. K. (2019). Pattern Mining from big IoT Data with fog Computing: Models, Issues, and Research Perspectives.

Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2020). Explainable AI in fintech risk management. *Frontiers in Artificial Intelligence*, *3*, 26. https://doi.org/10.3389/frai.2020.00026

Campbell, C., Sands, S., Ferraro, C., Tsao, H.-Y. J., & Mavrommatis, A. (2020). From data to action: How marketers can leverage AI. *Business Horizons*, *63*(2), 227–243. https://doi.org/10.1016/j.bushor.2019.12.002

Carcillo, F., Le Borgne, Y.-A., Caelen, O., Kessaci, Y., Oblé, F., & Bontempi, G. (2021). Combining unsupervised and supervised learning in credit card fraud detection. *Information Sciences*, *557*, 317–331. https://doi.org/10.1016/j.ins.2019.05.042

Chen, H., Li, L., & Chen, Y. (2020). Explore success factors that impact artificial intelligence adoption on telecom industry in China. *Journal of Management Analytics*, *8*(1), 36–68. https://doi.org/10.1080/23270012.2020.1852895.

Desjardins, J. (2019). How much data is generated each day?. World Economic Forum. https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/

Furht, B., & Villanustre, F. (2016). Introduction to big data. In *Big data technologies and applications* (pp. 3–11). Springer. 10.1007/978-3-319-44550-2_1

Gadek, G., Pauchet, A., Brunessaux, S., Khelif, K., & Grilhères, B. (2018). AI techniques to analyse a social network on text, user and group level: application on Galaxy2. 4ème conférence sur les Applications Pratiques de l'Intelligence Artificielle APIA2018.

Hellerstein, J. M., Sreekanti, V., Gonzalez, J. E., Dalton, J., Dey, A., Nag, S., Ramachandran, K., Arora, S., Bhattacharyya, A., & Das, S. (2017). Ground: A data context service. CIDR.

Kandogan, E., Roth, M., Schwarz, P., Hui, J., Terrizzano, I., Christodoulakis, C., & Miller, R. J. (2015). Labbook: Metadata-driven social collaborative data analysis. 2015 IEEE International Conference on Big Data (Big Data).

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Whos the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, *62*(1), 15–25. https://doi.org/10.1016/j.bushor.2018.08.004

Lam, L. T., & Hsiao, S.-W. (2019). AI-based online P2P lending risk assessment on social network data with missing value. 2019 IEEE International Conference on Big Data (Big Data).

Laney, D. (2001). *3D data management: Controlling data volume, velocity and variety* (pp. 1–4). Meta Group.

Li, J. (2022). Intelligent real-time news communication platform in education. *Mobile Networks and Applications*, *3*, 1–10. https://doi.org/10.1007/s11036-022-01948-0

Li, J., & Xu, X. (2020). A study of big data-based employees' public opinion system construction. *Journal of Industrial Integration and Management*, *5*(2), 225–233.

Li, L., Warfield, J., Guo, S. J., Guo, W. D., & Qi, J. Y. (2007). Introduction: Advances in intelligent information processing. *Information Systems*, *32*(7), 941–943. https://doi.org/10.1016/j.is.2006.10.001

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, *49*(2), 433–442. https://doi.org/10.3758/s13428-016-0727-z

Lu, Y. (2019). Artificial intelligence: A survey on evolution, models, applications and future trends. *Journal of Management Analytics*, *6*(1), 1–29. https://doi.org/10.1080/23270012.2019.1570365

Magesh, S., Niveditha, V., Rajakumar, P., & Natrayan, L. (2020). Pervasive computing in the context of COVID-19 prediction with AI-based algorithms. *International Journal of Pervasive Computing and Communications*, *16*, 477–487.

Marwala, T. (2015). *Causality, correlation and artificial intelligence for rational decision making*. World Scientific. 10.1142/9356

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D., & Barton, D. (2012). Big data: The management revolution. *Harvard Business Review*, *90*(10), 60–68.

Mindell, D. A., & Mindell, D. A. (2002). *Between human and machine: Feedback, control, and computing before cybernetics*. JHU Press.

Ricci, F., Rokach, L., & Shapira, B. (2015). Recommender systems: Introduction and challenges. In *Recommender systems handbook* (pp. 1–34). Springer. https://doi.org/10.1007/978-1-4899-7637-6_1

Roh, Y., Heo, G., & Whang, S. E. (2021). A survey on data collection for machine learning: A big data - AI integration perspective. *IEEE Transactions on Knowledge and Data Engineering*, *33*(4), 1328–1347. https://doi.org/10.1109/tkde.2019.2946162

Saha, D. (2020). How the world became data-driven, and what's next.

Salesforce. (2019). State of service: 2019's changing customer service trends. https://www.salesforce.com/au/blog/2019/04/state-of-service–2019-s-changing-customer-service-trends.html

Schwabe, A. (2020). Demo of JARET: AI powered web app for goal review and time management. Proceedings of the Seventh ACM Conference on Learning@ Scale.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Son, L. H., Ngan, R. T., Ali, M., Fujita, H., Abdel-Basset, M., Giang, N. L., Manogaran, G., & Priyan, M. K. (2020). A new representation of intuitionistic fuzzy systems and their applications in critical decision making. *IEEE Intelligent Systems*, *35*(1), 6–17. https://doi.org/10.1109/mis.2019.2938441

Surya, L. (2015). An exploratory study of AI and big data, and it's future in the United States. International Journal of Creative Research Thoughts (IJCRT), ISSN, 2320-2882.

Terzo, O., Ruiu, P., Bucci, E., & Xhafa, F. (2013). Data as a service (DaaS) for sharing and processing of large data collections in the cloud. 2013 Seventh International Conference on Complex, Intelligent, and Software Intensive Systems.

Wang, T., Ke, H., Zheng, X., Wang, K., Sangaiah, A. K., & Liu, A. (2019). Big data cleaning based on mobile edge computing in industrial sensor-cloud. *IEEE Transactions on Industrial Informatics*, *16*(2), 1321–1329. https://doi.org/10.1109/TII.2019.2938861

Xie, W., Yao, Z., Ji, E., Qiu, H., Chen, Z., Guo, H., Zhuang, J., Jia, Q., & Huang, M. (2021). Artificial intelligence–based computed tomography processing framework for surgical telementoring of congenital heart disease. *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, *17*(4), 1–24. https://doi.org/10.1145/3457613

Yang, Y., Zheng, X., Guo, W., Liu, X., & Chang, V. (2019). Privacy-preserving smart IoT-based healthcare big data storage and self-adaptive access control system. *Information Sciences*, *479*, 567–592. https://doi.org/10.1016/j.ins.2018.02.005

Yoon, S.-K., & Kim, S.-D. (2020). Pattern analysis based data management method and memory-disk integrated system for high performance computing. *Future Generation Computer Systems*, *106*, 185–198. https://doi.org/10.1016/j.future.2020.01.013

How to cite this article: Li, J., Ye, Z., & Zhang, C. (2022). Study on the interaction between big data and artificial intelligence. *Systems Research and Behavioral Science*, *39*(3), 641–648. https://doi.org/10.1002/sres.2878