Exploring the cross-disciplinary collaboration: a scientometric analysis of social science research related to Artificial Intelligence and Big Data application

Han-Teng Liao¹, Zijia Wang* and Yu Liu¹

¹ New Media Research Center, Sun Yat-sen University Nanfang College, Guangzhou, Guangdong, 510970, China

*Corresponding author’s e-mail: wang.zijia@outlook.com

Abstract. Given the significant impacts of Big Data and Artificial Intelligence (AI) on the planet and people, it is important for researchers from all disciplines to pay attention to the social, economic and policy needs and challenges with its applications. With the aim to provide a systematic review of social science literatures on the topic, this article has summarized the annual trends since 1962, main research disciplines and publication outlets based on a scientometric analysis of 10,229 articles collected from the WoS database. The research findings show that disciplines/areas such as management and psychology have been constantly engaging the Big Data and AI, whereas new emerging disciplines/areas such as interdisciplinary social sciences and geography have experienced exponential growth in recent years. Both the application and critiques of Big Data and AI have been published in various new journals such as Big Data & Society, Value in Health, etc. The article therefore provides essential groundwork for understanding and contextualizing the Big Data and AI applications, so that researchers, funding bodies, policy-makers and industry professionals can enhance cross-disciplinary collaboration.

1. Introduction

The social, economic, and political implications of the advancement and adoption of Big Data and Artificial Intelligence (AI) application have been in the forefront of high-level dialogues in recent years, discussing its promises and potential dangers[1]. AI will have “good, bad, transformative and plain weird effects” on societies; thus, major organizations such as Google, Microsoft, Facebook, UNICEF, etc., founded the Partnership on Artificial Intelligence to Benefit People and Society in 2016 [2,3]. For another, the ITU has established the “AI for Good” Global Summit with ACM and other UN organizations as the UN “platform for global and inclusive dialogue on AI”[4]. UN Chief Information Technology Officer also argued that we need to reflect on both positive and negative ramifications and we need to “transform ourselves as innovators”[5]. In addition to AI, various partnerships have been built as part of “Data Revolution”, leveraging Big Data for sustainable development[6]. The United Nations Statistical Commission also created the Global Working Group (GWG) on Big Data for Official Statistics to investigate its benefits and challenges[7]. Acknowledging the significant impacts of Big Data and AI on the planet and people, these examples demonstrate the strong demand for responsible AI and Big Data for a positive future.

Thus, any Big Data and AI application must pay attention to the social, economic and policy needs and challenges. Some literature review efforts have been made to explore the application of Big Data or AI on specific areas of social sciences [e.g. 8–11], but none of them are comprehensive enough in terms
of the scope of research areas and time periods to provide an overview picture. It is then vital to map out, in relation to AI and Big Data, how the social science literatures and their sources evolve. To provide an overview picture, the article aims to provide a systematic literature review to contextualize such latest research concerns and developments within social sciences in relation to the Big Data and AI. Such a map will be useful for tracing and exploring the cross-disciplinary collaboration across time, thereby providing groundwork for contextualizing and understanding the application of AI and Big Data in relation to social sciences.

2. Data and methods

To explore the dynamics of social sciences literatures in tackling Big Data and AI related issues and applications, this article starts with the query design that covers the social sciences domain.

The Thomson Reuters Web of Science (WoS) database has 25 WoS “research areas” under the main area of social sciences, ranging from archaeology to women’s studies. Major research indexes such as SCI-EXPANDED, SSCI, A&HCI, and ESCI are included in the scope of this research.

The below “Advanced Search” query shows how our data query design incorporates “Artificial Intelligence”, “machine learning” and Big Data” as topics (coded by the WoS Field Tag as TS) and “Social Sciences” as research areas (coded by the WoS Field Tag as SU).

- TS = ("Artificial Intelligence" OR "machine learning" OR "Big Data") AND SU="Social Sciences"

Together with other narrower 25 social science areas, total 26 queries executed in November 2019, resulted in 10,229 articles. VOSviewer and Python visualization packages were used for mapping.

3. Research mapping results

Given the scope and time span of collected data, it may be helpful to identify several time periods and then focus the historical, publication outlets, and disciplinary dynamics to frame the data sets.

3.1. Annual trends

In order to identify time periods, we conducted linear regressions between the variable of years and that of number of articles (in log values), and three time periods were found through several iterations.

As shown in Figure 1, the number of publications has continued to rise exponentially, except for the time period between 1995 and 2004. Also, Figure 1(c) shows that the growth rate after 2005 is also faster than the time period before 1995, based on the slope value of regression lines.

![Figure 1. Annual Publications Trend](image)

(a) 1962-1994
(b) 1962-2004
(c) all years since 1962

3.2. Top WoS categories and research areas

Table 1 lists the top WoS categories across three time periods. While management has been at the top category throughout, research categories such as economics, business, geography and communication have emerged after 2005, taking the place of those such as operations research and psychology. Offering slightly different disciplinary categorization, Table 2 indicates that the social sciences research around the world not only focus on the classic disciplines such as psychology, business, operation research and education, but also increasingly cover “Other Topics” of Social Sciences.
Table 1. Top WoS categories (WoS tag: WC)

| 1962-1994                          | 1995-2004                          | 2005-2020                          |
|-------------------------------------|-------------------------------------|-------------------------------------|
| Management                          | Management                          | Management                          |
| Psychology, Multidisciplinary       | Operations Research & Management Science | Economics                          |
| Operations Research & Management Science | Psychology, Multidisciplinary       | Business                            |
| Psychology                          | Education & Educational Research    | Geography, Physical                 |
| Computer Science                    |                                    | Communication                        |

Table 2. Top research areas (WoS tag: SC)

| 1962-1994                          | 1995-2004                          | 2005-2020                          |
|-------------------------------------|-------------------------------------|-------------------------------------|
| Psychology                          | Business & Economics                | Business & Economics                |
| Business & Economics                | Psychology                          | Psychology                          |
| Operations Research & Management Science | Operations Research & Management Science | Social Sciences - Other Topics         |
| Education & Educational Research    | Computer Science                    | Education & Educational Research    |
| Computer Science                    | Social Sciences - Other Topics      | Physical Geography                  |

3.3. Top publication outlets

While some publication outlets may be representatives of a discipline or research area’s work, some may serve as sites of collaboration across disciplines. Table 3 lists the top publication outlets across different time periods, indicating similar (or corresponding) patterns found above: For instance, the journal *Big Data and Society* belongs to the WoS research area of “Interdisciplinary Social Sciences”; the journal *Value in Health* belongs to that of “Economics; Health Care Sciences & Services; Health Policy and Services”. Both of which indicates significant presence of cross-disciplinary collaboration.

Table 3. Top publication outlets

| 1962-1994                          | 1995-2004                          | 2005-2020                          |
|-------------------------------------|-------------------------------------|-------------------------------------|
| European Journal of Operational Research | European Journal of Operational Research | Big Data & Society                  |
| Contemporary Psychology             | Journal of the Operational Research Society | ISPRS International Journal of Geo-information |
| Interfaces                          | International Journal of Human-Computer Studies | European Journal of Operational Research |
| Journal of the Operational Research Society | Computational Linguistics      | Value in Health                      |
| Behavioral and brain sciences       | Behavioral and Brain sciences      | IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing |

3.4. Bibliographic coupling network and clusters

Since citation-based bibliographic coupling can be used to identify research fronts and intellectual bases [12], we conducted network mapping based on bibliographic coupling of publication outlets. As shown in Figure 2, four clusters have been identified and visualized for two time periods: (a) from 1962 to 1994 and (b) from 2005 to 2018. The operation/management cluster (in red color, with *European Journal of Operational Research* at its prominent core) and the psychology cluster (in green color) have been consistently prominent. In contrast, the interdisciplinary social sciences cluster (in blue color) and the geography cluster (in yellow color) have only become much more prominent in the last decade. Indeed,
as also indicated by the top publication outlets in Table 3, the journals *Big Data and Society* and *ISPRS International Journal of Geo-information* correspond to these two emerging clusters of interdisciplinary social sciences and geography respectively.

4. Conclusion

The purpose of this literature review was to provide an overall understanding of how social science literatures and their publication outlets evolve, in relation to AI and Big Data. Three time periods and then four prominent clusters (operation/management, psychology, interdisciplinary social sciences, and geography) have been identified to show both the constant and emerging research fronts and intellectual bases of AI and Big Data related social science work. Although the data source (the WoS database) may be limited in scope, and some of the articles may not be considered as social science work, the findings...
here nonetheless provide essential groundwork for anyone to begin to explore the application of AI and Big Data in relation to social sciences.

The cross-disciplinary collaboration identified in the article is mostly limited to the level of publication outlets and disciplines. Future research can examine, in greater or and with specific research concerns, how collaboration takes place across disciplines at the level of articles and authors. Notwithstanding its limitations, the article offers a roadmap for researchers, funding bodies, policy-makers and industry professionals to explore and learn from cross-disciplinary collaboration fruitful areas and publication outlets. Also, such a map should help us pay attention to the social, economic and policy needs and challenges facing the adoption of Big Data and AI.

Acknowledgments

The research is funded by the Application-oriented Curriculum Development Projects of “API, Machine Learning and Artificial Intelligence” (NFU 02-40248) and “Big Data” (NFU 02-40249), under the Guangdong Province Department of Education 2018 "Innovation-Strengthening Higher Education Program” Grants, and partly by a project of HCI for Education Development (2018WQNCX283), under the 2018 Major Projects of Guangdong Province Department of Education "Young Innovative Talents” Grants. Han-Teng Liao and Zijia Wang have made equal contributions to the paper, including the data collection and analysis work and have given the approval of the current version for publication.

References

[1] UN News (2019) Artificial intelligence summit focuses on fighting hunger, climate crisis and transition to ‘smart sustainable cities.’ UN News
[2] PAI (2019) The Partnership on AI. Partnersh. AI PAI
[3] Butler, D. (2017) AI summit aims to help world’s poorest. Nat. News 546: 196.
[4] ITU (2019) AI for Good Global Summit, Geneva, Switzerland. AI Good Glob. Summit
[5] Patterson, D. (2016) United Nations CITO: Artificial intelligence will be humanity’s final innovation. TechRepublic
[6] UN (2017) Big Data for Sustainable Development. U. N.
[7] GWG (2019) UN Global Working Group (GWG) for Big Data. Big Data UN Glob. Work. Group
[8] Mountrakis, G., Stefanidis, A. (2008) Special issue: Artificial Intelligence in remote sensing - Foreword. Photogramm Eng Remote Sens 74.0: 1199.
[9] Coetzee, S., Rautenbach, V. (2017) A Design Pattern Approach to Cartography with Big Geospatial Data. Cartogr J 54.0: 301.
[10] Akter, S., Wamba, S.F. (2016) Big data analytics in E-commerce: a systematic review and agenda for future research. Electron Mark 26: 173.0.
[11] Guenther, W.A., Mehrizi, M.H.R., Huysman, M., Feldberg, F. (2017) Debating big data: A literature review on realizing value from big data. J Strateg Inf Syst 26: 191.0.
[12] Persson, O. (1994) The intellectual base and research fronts of JASIS 1986–1990. J. Am. Soc. Inf. Sci. 45: 31–8.
[13] boyd, danah, Crawford, K. (2012) Critical Questions for Big Data. Inf. Commun. Soc. 15: 662–79.
[14] Loebbecke, C., Picot, A. (2015) Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. J Strateg Inf Syst 24: 149.0.
[15] Mittelstadt, B. (2017) Ethics of the health-related internet of things: a narrative review. Ethics Inf Technol 19: 157.
[16] Cui, J. (2018) Dynamic Migration Algorithm of Marine Big Data in Cloud Computing Environment. J Coast Res 706.
[17] Sung, Y.-T., Chen, J.-L., Cha, J.-H., Tseng, H.-C., Chang, T.-H., Chang, K.-E. (2015) Constructing and validating readability models: the method of integrating multilevel linguistic features with machine learning. Behav Res Methods 47.0: 340.

[18] Cowls, J., Schroeder, R. (2015) Causation, Correlation, and Big Data in Social Science Research. Policy Internet 7: 447–72.