

DOI: 10.5281/zenodo.21652873

A BIBLIOMETRIC ANALYSIS OF GLOBAL RESEARCH ON DATA USE IN SOCIAL SERVICES (1950–2026): EVIDENCE FROM THE WEB OF SCIENCE DATABASE USING VOSVIEWER

Kholud S. Makkawi^{1*}, Mohammed A. Alkohaiz²

¹PhD Candidate, Department of Social Studies, King Saud University, Riyadh, Saudi Arabia,
kshmakkawi@uqu.edu.sa, <https://orcid.org/0009-0006-6203-3993>

²Department of Social Studies, King Saud University, Riyadh, Saudi Arabia, malkohaiz@ksu.edu.sa,
<https://orcid.org/0000-0003-2210-1884>

Received: 25/06/2026
Accepted: 13/07/2026

Corresponding Author: Kholud S. Makkawi
(kshmakkawi@uqu.edu.sa)

ABSTRACT

This study investigates the global scholarly production on the use of data in social services using a bibliometric approach based on the Web of Science database. In the context of rapid digital transformation and the increasing integration of data-driven approaches in social work, there is a growing need to systematically map and analyze the intellectual structure and evolution of this interdisciplinary field. The study aims to (1) identify the most scientifically productive countries in the field of data use in social services, (2) analyze the temporal evolution of scientific production, (3) explore the intellectual structure and research clusters, and (4) identify emerging future research directions in the field. A total of (512) publications were retrieved and analyzed using VOSviewer to perform co-occurrence, network, and visualization analyses, in order to identify the knowledge structure of scientific production in the field of data use for improving social services. The results show that China, the United States, and the United Kingdom are the leading contributors in terms of scientific productivity. Temporal analysis revealed four developmental stages in the evolution of research on data use in social services, with a marked acceleration in scientific output between 2020 and 2026. The intellectual structure of the field was organized into five major clusters, focusing on data privacy and security, advanced analytics techniques, data mining and social networks, big data and machine learning applications, and social media-driven data analytics. Emerging research directions include computational social science, socio-technical imaginaries, digital twins, the Social Internet of Things (SIoT), social genomics, and Social Representation Theory. The study concludes that data-driven approaches are increasingly reshaping social services and social work practice, highlighting the need for stronger interdisciplinary collaboration and capacity-building in data analytics within the social sector

KEYWORDS: Bibliometric Analysis; Social Work; Social Services; Data Science; Big Data; Vosviewer; Web Of Science; Research Trends; Knowledge Mapping.

1. INTRODUCTION

Global scientific output has experienced rapid growth over the past decades, as the number of published research articles increased from approximately (2) million papers in 2010 to more than (3.3) million papers in 2022—an increase of nearly (65%) within a single decade (Adams, 2026). This accelerated growth has led to a massive influx of dispersed scientific knowledge across various disciplines. Consequently, the importance of employing advanced quantitative approaches has emerged, as they can analyze scientific production, revealing its intellectual structure and emerging knowledge trends, rather than relying solely on traditional systematic reviews. In this context, bibliometric analysis has emerged as one of the most important methodologies used in evaluating scientific production and mapping knowledge structures in interdisciplinary studies. It enables systematic analysis of large-scale bibliographic data to identify patterns of scientific publishing, relationships among researchers, institutions, and research topics, as well as their temporal evolution and emerging trends. This makes it a key tool for supporting research planning and guiding future studies (Passas, 2024).

In light of rapid digital transformation, data-related topics have become central fields that have fundamentally changed decision-making mechanisms, resource management, and service development across various sectors. The effects of this transformation have extended to social services and social work, where data derived from electronic records of social, health, and educational services, open data platforms, social media, online search activities, and geolocation devices provide extensive opportunities to better understand social problems, analyze beneficiaries' needs, and develop more accurate and effective professional interventions (Coulton et al., 2015). Moreover, the use of data and its tools contributes to supporting planning and decision-making processes, evaluating social programs, and improving resource allocation, in addition to enhancing the capacity of social institutions to respond to changing societal needs based on empirical data (Cariceo et al., 2018).

Despite the significant potential offered by data science and its tools, the literature indicates that the field of social services and social work still faces challenges in systematically integrating data and analytics into professional practice. These challenges stem from the complexity of social data, variability in its quality, and limited technical and analytical capacities in some social service institutions (Dimas

et al., 2023). Furthermore, although many institutions adopt positive orientations toward data use and the development of data-driven strategies, the actual implementation of these approaches in improving services and decision-making remains limited. This reflects a gap between the intention to use data and the actual practice of data analytics and utilization (Nonprofit Technology Network, 2025).

Accordingly, this study aims to identify the knowledge structure linking the field of data and the field of social work, as well as how research topics have evolved over time. It represents a comprehensive study of the intellectual structure in this domain through four main objectives: identifying the most influential countries in this field; examining the temporal evolution of scientific production in data and social services; identifying the knowledge structure and main research clusters within the field; and exploring future research directions in data and social work.

2. THEORETICAL FRAMEWORK OF THE STUDY:

2.1. The Use of Data in Improving Social Services

Within the context of this study, data science refers to the process of identifying, acquiring, cleaning, and using data to determine relationships and extract value from transforming raw data into valuable assets. This is achieved by working with different types of data to guide actions and support rational decision-making. It is an interdisciplinary field that intersects with other disciplines, most notably statistics and mathematics (Cariceo et al., 2018).

The term social services refers to safety-net programs provided to children, families, and individuals, sometimes referred to as social protection programs. These include services such as housing, food assistance, rehabilitation programs, and child protection services. Such services are delivered by governmental, private, or non-profit organizations (Dimas et al., 2023).

Despite the critical importance of data use across various disciplines and the potential to employ its tools in support of social services, those interested in applying data science in social contexts face several challenges. The most prominent of these include algorithmic fairness and the small size and limited nature of data in some organizations, which may obscure individual circumstances and differences among clients (Dimas et al., 2023).

According to the report of the Nonprofit

Technology Network (NTEN) for 2025, which was based on a survey of 220 organizations, the main challenges in data collection and use include: a lack of qualified personnel and the absence of internal staff capable of adopting advanced data analytics; low employee awareness of the value of data and its strategic use; limited resources and funding; insufficient tools and systems within social organizations and their poor integration with workflows; and organizational fragmentation and weak interdepartmental coordination, which hinder data sharing and integrated use within organizations.

Several experts in the use of data in the social field emphasize the importance of collaborative partnerships between social work professionals and data science specialists to ensure fair and ethical use of data. There is also a growing need to collect more contextualized client data. This is particularly important given that data held by most social organizations is often limited in variables and focused on current objectives or funder priorities. Therefore, time is needed to accumulate richer datasets that can support predictive and future-oriented analysis (Dimas et al., 2023, p. 3).

At the planning level, organizations must balance a variety of considerations when designing data collection efforts (Nonprofit Technology Network Report, 2025, p. 9). This includes educating and training professionals in data use, including social work students, and developing academic programs and diplomas that enhance knowledge and skills in client data usage, while ensuring data security and privacy. It also requires reliance on advanced data analytics, training practitioners and service providers to improve their skills in data collection and analysis using artificial intelligence techniques or advanced statistical methods and raising awareness of the importance of using data to improve services and support evidence-based decision-making (Casola, 2020).

2.2. Bibliometric Analysis Approach

The measurement of scientific output has become a dominant trend in academic research in recent years. The term “bibliometrics” first appeared in the 1930s through the Belgian documental researcher Paul Otlet, and it was later reformulated and widely popularized by Pritchard in 1969. The importance of bibliometrics lies in its applicability to handling large volumes of scientific data and its significant contribution to enhancing research impact. Its growing relevance has been further strengthened by the development of scientific productivity

measurement tools such as R and VOSviewer, their accessibility and availability, as well as the emergence of scientific databases such as Google Scholar, Scopus, and Web of Science.

Researchers use bibliometric analysis for multiple purposes, including identifying emerging trends in scientific production, exploring the intellectual structure of a specific field within the literature, mapping cumulative scientific knowledge, and understanding it objectively, as well as capturing the evolutionary dynamics of established disciplines. Accordingly, well-conducted bibliometric studies can provide a solid foundation across various fields (Passas, 2024).

Bibliometric analysis is defined as the process of extracting diverse scientific outputs through the study of publications within a specific field or academic journal, using quantitative and statistical analyses of bibliometric indicators. These include the number of annual publications, the most productive universities, leading journals in the field, the most prolific authors, citation counts, and keywords. This approach relies on more objective and precise quantitative methods, in contrast to systematic literature reviews, which often depend on qualitative approaches that may be influenced by researchers’ interpretive biases. Therefore, bibliometric analysis helps reduce such biases (Ozturk et al., 2024).

The main steps in bibliometric analysis include precisely defining the research objectives, selecting the database and the tools or software to be used for analysis, collecting data by designing a search query that defines the scope of the study, reviewing the retrieved data to ensure its relevance to the research topic, and finally conducting bibliometric analysis. This includes performance analysis, which examines research components such as institutions, countries, authors, and journals, as well as science mapping to identify the knowledge structure of the field through network visualization and data representation (Donthu et al., 2021).

2.3. Study Methodology

This study employed bibliometric analysis of the literature retrieved from the Web of Science database and utilized scientific mapping through VOSviewer. The bibliometric analysis process was conducted in three stages.

First: Data Collection

Data was collected from the database using the following search query

(("social work" OR "social service" OR "social welfare" OR "social care") AND ("data science" OR

"data analytics" OR "data-driven" OR "data mining" OR "data quality" OR "data use"))

The initial search results yielded (822) scientific publications, including journal articles, theses, and

book chapters. From these, (512) publications were selected according to the criteria presented in the following table.

Table 1. Bibliometric Data Screening and Eligibility Criteria.

Description	Eligibility Criterion
The initial search using the predefined search query retrieved a total of (822) publications. After applying the first eligibility criterion, which included only English-language publications, the number of eligible documents was reduced to (792) publications.	Language of Publication
After applying the document type criterion, which included scientific articles, review articles, early access publications, and dissertation theses, the number of eligible documents was reduced to (512) publications.	Document Type
The study covered the period from 1950 to 2026, with 2026 data limited up to July 8, 2026, as it includes early access publications released during the same year. This allows inclusion of the most recent research developments.	Publication Period

Second: Data Processing

All titles and abstracts of publications included within the study scope were reviewed to ensure their relevance to the use of data in social services and social work. All data were exported in Text format, and using VOSviewer software, a total of (2,050) terms were extracted from the initial dataset. In order to improve the quality of the analysis and focus on the most frequent and highly connected terms, a minimum occurrence threshold of (5) was applied. This reduced the number of terms to (39).

Third: Data Analysis

Descriptive analysis was performed using Microsoft Excel to address the first and second research objectives. VOSviewer was used to conduct network analysis and generate visualizations for the third and fourth research objectives.

4. STUDY RESULTS

1. Identify the most scientifically productive countries in the field of data use in social services.

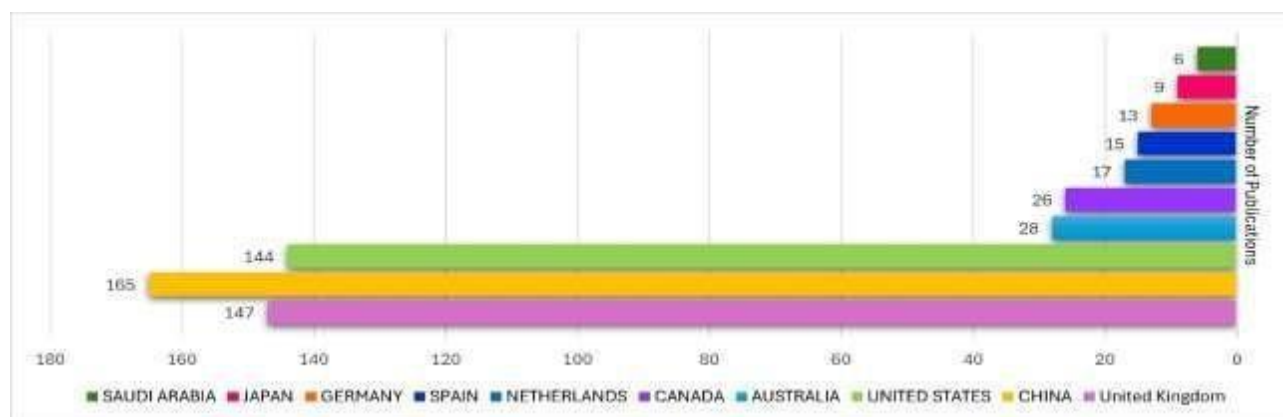


Figure 1. Distribution of Publications by Country.

The study found that China, the United Kingdom and the United States were the leading countries in scientific production on the use of data to improve social services. China ranked first followed by the United Kingdom in second place and the United

States in third. Other countries including Australia and Canada also contributed to the scientific output although their contributions were relatively limited, with no more than (30) publications each. Among the top ten most productive countries presented in the

following figure, Saudi Arabia ranked tenth (last), with (6) publications during the study period.

The leading position of China, the United Kingdom and the United States in scientific output related to the use of data in improving social services may be attributed to international research cooperation, which contributes to enhancing research productivity, exchanging knowledge among researchers with different cultures and knowledge backgrounds, exploiting the different resources available to researchers, obtaining broader funding opportunities, and thus increasing scientific output and achieving a greater impact for countries.(Adams& Szomszor,2024)

Note. Due to the full counting method used by Web of Science, internationally co-authored publications are counted for each contributing

country. Therefore, the sum of country-level publications exceeds the total number of publications included in the analysis.

2. Temporal Analysis of Scientific Production in the Field of Data Use in Social Services

The temporal analysis results showed that scientific production remained relatively low until 2015, with annual publications not exceeding (5) papers per year. Starting from 2016, a gradual upward trend in scientific output began to emerge, reflecting a growing interest in the use of data to improve social services. This growth peaked between 2022 and 2026, indicating a clear shift in academic attention toward the use of data in enhancing and developing social services.

Table 2. Annual Distribution of Publications.

Number of publications	Years of publication
44	2026
93	2025
50	2024
45	2023
50	2022
42	2021
30	2020
51	2019
23	2018
17	2017
13	2016
7	2015
4	2014
5	2013
6	2012
4	2011
3	2010
4	2009
2	2008
2	2006
4	2005
2	2004
3	2003
1	2002
1	2001
2	1998
2	1995
1	1973
1	1950

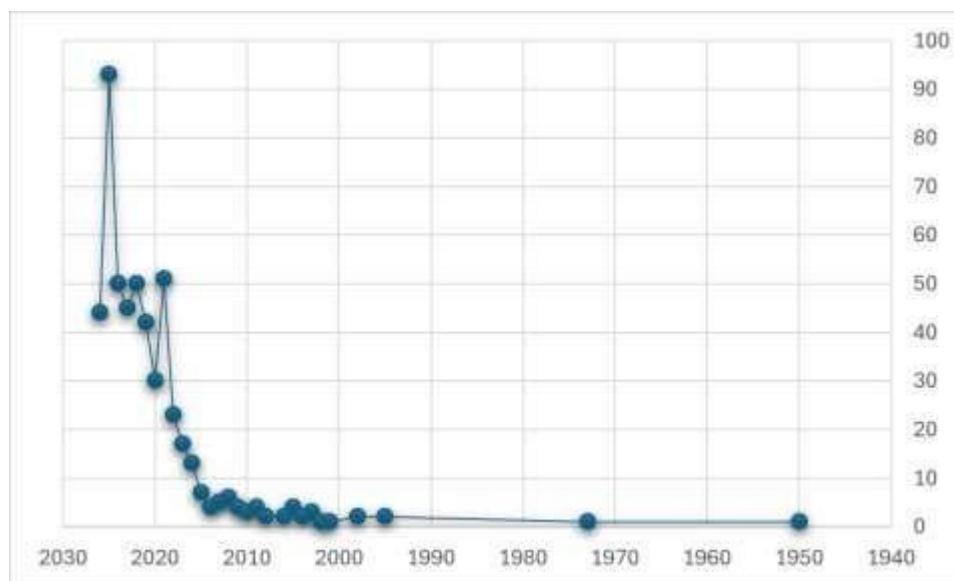


Figure 2. Temporal evolution of scientific production.

Interpretation of the Temporal Evolution of Scientific Production

Based on the observed publication trend, the evolution of research on the use of data to improve social services can be divided into four distinct periods.

The first period (1950–2005) represents the initial emergence of scholarly interest in the use of data within the social services field. One of the earliest influential studies was conducted by Gooden and Pollock (1998), which aimed to integrate health and social service data in an urban area of London to identify the characteristics of service users and support integrated care planning. The findings highlighted the value of linking databases to improve understanding of service utilization patterns while also revealing challenges related to data quality and accuracy. These findings reflect the primary research focus of this early period, which emphasized data collection and integration rather than the advanced analytical methods and predictive models that emerged in later years.

Toward the end of this period, May and Roberts (2005) published a study emphasizing the importance of high-quality data for evidence-based policymaking, particularly in the development of poverty reduction policies. The study stressed the growing responsibility of national statistical agencies to provide reliable data that governments could use to strengthen public policies and address social challenges.

The second period (2006–2012) was characterized by a substantial increase in data-driven practices and more advanced research applications. For example, Greble and Cenan (2010) proposed a client data-

based architecture designed to improve patient services, facilitate collaboration among healthcare professionals, and enhance access to electronic records. During this period, data also began to be used proactively to identify potential client risks and reduce pressure on care organizations by supporting home-based care. This shift is illustrated by the study of McNaul et al. (2012), which explored the use of smart living systems and data technologies to support the care of older adults.

The third period (2013–2019) witnessed a rapid expansion of interest in data, particularly digital data, driven by the increasing volume and continuous generation of information. Researchers began to focus on concepts such as real-time data, as demonstrated by Mori et al. (2013). By 2019, publication activity reached its highest level within this period, and topics related to big data became increasingly prominent in social work research.

Studies such as Zetino and Mendoza (2019) highlighted the growing importance of big data in informing social work practice and decision-making.

The current period (2020–2026) is characterized by interdisciplinary research and the application of data science as an independent discipline with its own methodologies and analytical tools in the social services domain. For example, Jung (2020) demonstrated the integration of data science approaches into social work research. In 2025, scientific production reached its highest level within the study sample, with 93 publications, reflecting the increasing maturity of the field. During this period, researchers increasingly incorporated advanced technical concepts into both social work practice and care research. For instance, Nollet and Trubey (2025)

examined the limitations of using low-fidelity synthetic data for scientific research. This period continues to the present, with the emergence of new methodological approaches, such as data care, which employs ethnographic data to investigate how data infrastructures operate and support the delivery of digital social care services.

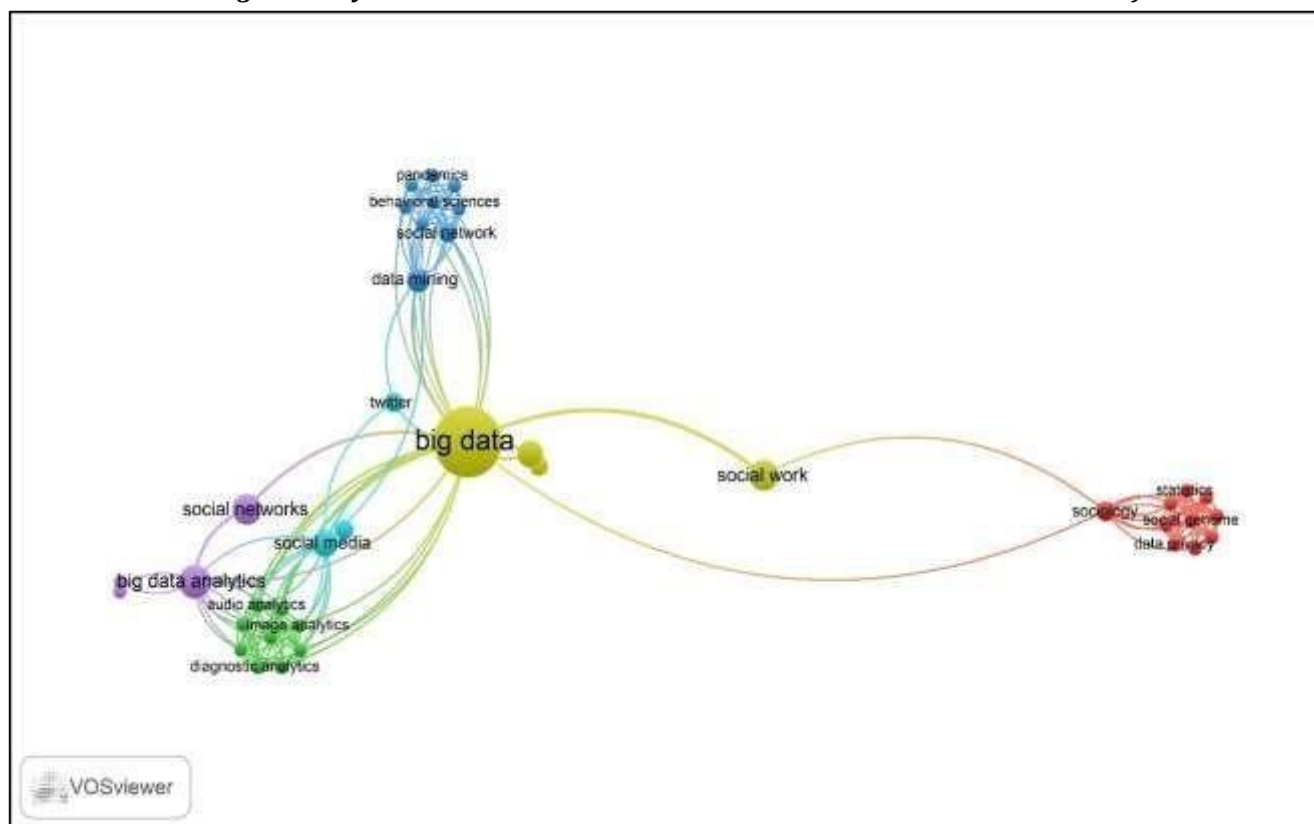
3. Intellectual Structure of Research Trends in the Use of Data in Social Services

Bibliometric analysis based on keywords and an examination of the titles and abstracts of the retrieved publications identified (2,050) terms from the primary dataset. To enhance the quality of the analysis and focus on the most frequent and relevant terms, a minimum occurrence threshold of (5) was applied, reducing the number of terms to (39). These terms were categorized into (5) clusters representing

the main research themes and the scientific knowledge structure of this field.

The first cluster focuses on Big Data, which represents the most influential and dominant theme in the field. The second cluster addresses the relationship between data processing topics and issues related to privacy and secure data access. The third cluster includes research on advanced data analytics tools, such as voice analysis, image analysis, predictive and descriptive analytics, and text analysis. The fourth cluster focuses on data mining and its relationship with social media and social networks, highlighting the application of data mining in humanitarian contexts, such as social crises, and emphasizing the importance of proactive data utilization. The fifth cluster represents social network analytics, particularly through the collection and utilization of customer data from social media platforms.

Figure 3. Keyword co-occurrence network and thematic clusters in the research field.



4. Future Research Directions in the Use of Data in Social Services

The findings of the study revealed a range of emerging topics that have the potential to contribute

to the development of social services and the advancement of social work research. Among the most prominent of these topics are socio-technical imaginaries, digital twin technology, computational social science, the Social Internet of Things (SIoT), Social Representation Theory, and the social genome.

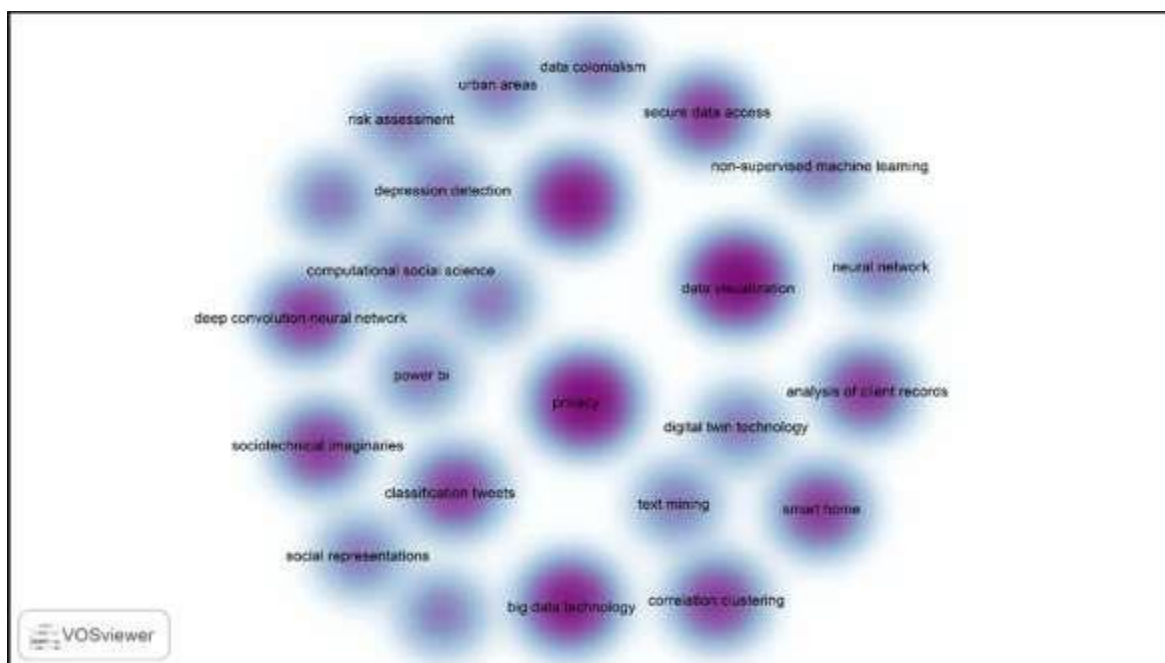


Figure 4. Knowledge Map Illustrating Future Research Directions in the Use of Data in Social Services.

5. DISCUSSION OF FINDINGS

The results of the study indicate a rapid and continuous development in scientific interest in the use of data within social services. The temporal analysis shows a clear transition of the field from an early focus on data collection and integration during the initial decades to a more advanced stage characterized by predictive analytics and data-driven proactive interventions. The notable growth in scientific output during the period (2020–2026) can be attributed to the global expansion of digital transformation, the acceleration of data flows from diverse sources, and the increasing reliance on big data and open data to support social decision-making. In addition, rapid advancements in artificial intelligence and machine learning technologies have created new opportunities for understanding and addressing social problems more efficiently.

The findings also demonstrate that the evolution of data use in social services has undergone several successive stages reflecting the development of the technical and methodological infrastructure of the field. In the early stages, attention was primarily directed toward building databases, improving data quality, and linking data across institutions. This gradually shifted toward the use of data for risk prediction and improving the quality of care and services provided to beneficiaries, culminating in the current stage characterized by the integration of data science with social work and social care disciplines through the application of artificial intelligence tools,

advanced analytics, and computational models.

With regard to the most productive countries, the dominance of China, the United States, and the United Kingdom in scientific output reflects a strong relationship between national technological advancement and the capacity to produce knowledge related to the application of data in social services. In contrast, the relatively limited scientific output from Arab countries—including Saudi Arabia—indicates a promising research opportunity to strengthen local studies in this field in alignment with digital transformation requirements and Saudi Vision 2030.

The keyword co-occurrence map further reveals that research interest is not limited to technical aspects alone but extends to ethical and regulatory issues related to data use, such as privacy, information security, and secure data access. The findings also highlight a growing emphasis on advanced analytical tools, including text, image, and speech analysis, as well as data mining and social network analysis, alongside a relative decline in reliance on traditional data sources such as social surveys and opinion polls. This reflects an increasing shift toward the use of non-traditional data sources to study social phenomena and better understand service users' needs with greater precision.

Moreover, the results confirm the strengthening relationship between big data and social work through the application of machine learning techniques and predictive analytics to improve social services. This aligns with global trends advocating

for evidence-based and data-driven professional practice. Such an approach enhances the capacity of social institutions to anticipate social problems, allocate resources to the most vulnerable groups, and improve the efficiency of professional interventions.

Regarding future research directions, emerging topics such as socio-technical imaginaries, computational social science, digital twin technology, the Social Internet of Things (SIoT), the social genome, and Social Representation Theory are

expected to contribute significantly to the development of scientific output. These themes suggest that the field of social services is increasingly moving toward the adoption of intelligent digital models and multi-source data analytics to understand social problems and design more precise and effective interventions, thereby enhancing the quality and sustainability of social services in the future.

REFERENCES

- Adams, A. (2026). Number of academic papers published per year. Booketic. <https://booketic.com>
- Adams, Jonathan, & Szomszor, Martin. (2024). National research impact is driven by global collaboration, not rising performance. *Scientometrics*, 129, 2883–2896. Springer Nature. An empirical analysis of individual and collective determinants of international research collaboration | *Scientometrics* | Springer Nature Link
- Casola, L. (2020). Data science postsecondary education: A compilation of meeting highlights. National Academies Press. <https://nap.nationalacademies.org>
- Cariceo, O., Nair, M., & Lytton, J. (2018). Data science for social work practice. SAGE Publications. Data science for social work practice - Oscar Cariceo, Murali Nair, Jay Lytton, 2018
- Coulton, C., Goerge, R., Haan, B., & others. (2015). Harnessing big data for social good: A grand challenge for social work. American Academy of Social Work and Social Welfare. aaswsw.org/wp-content/uploads/2015/12/WP11-with-cover.pdf
- Dimas, G., Goldkind, L., & Konrad, R. (2023). Big ideas, small data: Opportunities and challenges for data science and the social services sector. *International Journal of Care and Caring*. <https://doi.org/10.1177/20539517231171051>
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285–296. <https://doi.org/10.1016/j.jbusres.2021.04.070>
- Gooden, S., & Pollock, A. M. (1998). How to profile the population's use of health care and social care in one district. *Journal of Public Health Medicine*, 20, 175–179. [JPH_1998_Godden_ProfileServiceUse.pdf](https://www.jph.org.uk/jph_1998_Godden_ProfileServiceUse.pdf)
- Greble, H. A., & Cenan, C. O. (2010). A proposed data-driven architecture for cardiology network application. *BRAIN: Broad Research in Artificial Intelligence and Neuroscience*, 1(1), 9–20. (PDF) A Proposed Data Driven Architecture for Cardiology Network Application
- Jang, W. (2020). Using data science to strengthen the social safety net: Predicting risk for Medicare and Medicaid insurers. *Journal of Governmental Studies*, 26, 29–61. Using data science to strengthen the social safety net: Predicting risk for Medicare and Medicaid insurers | Request PDF
- May, J., & Roberts, B. (2005). Poverty diagnostics using poor data: Strengthening the evidence base for pro-poor policy making in Lesotho. *Social Indicators Research*, 74, 477–510. Poverty Diagnostics Using Poor Data: Strengthening the Evidence Base for Pro-Poor Policy Making in Lesotho | *Social Indicators Research* | Springer Nature Link
- McNaull, J., Augusto, J. C., Mulvenna, M., & McCullagh, P. (2012). Data and information quality issues in ambient assisted living systems. *Data-and-Information-Quality-Issues-in-Ambient-Assisted-Living-Systems.pdf*
- Mori, A. R., Mazzeo, M., Mercurio, G., & Verbicaro, R. (2013). Holistic health: Predicting our data future—from interoperability among systems to co-operability among people. *International Journal of Medical Informatics*, 82(1), 14–28. Holistic health: predicting our data future (from inter-operability among systems to co-operability among people) - PubMed
- Nollet, C., & others. (2025). Discussing data: A UK-wide public consultation on the use of low-fidelity synthetic data for research. *International Journal of Population Data Science*, 10(1), Article 32. <https://ijpds.org/article/view/32>
- Nonprofit Technology Network. (2025). 2025 Data Empowerment Report. NTEN. <https://www.nten.org/publications/data-empowerment-report>

- Ozturk, O., Kocaman, R., & Kanbach, D. (2024). How to design bibliometric research: An overview and a framework proposal. *Review of Managerial Science*, 18, 3333–3361. <https://doi.org/10.1007/s11846-023-00689-0>
- Passas, I. (2024). Bibliometric analysis: The main steps. *Encyclopedia*, 4, 1014–1025. *Bibliometric Analysis: The Main Steps*
- Zetino, J., & Mendoza, N. (2019). Big data and its utility in social work: Learning from the big data revolution in business and healthcare. *Social Work in Public Health*, 34(5), 409–417. *Big Data and Its Utility in Social Work: Learning from the Big Data Revolution in Business and Healthcare: Social Work in Public Health: Vol 34 , No 5 - Get Access*