Advanced Research Techniques: Theories, Methods, and Practices (Volume-2)

Editor

Dr. R. Sundar

Associate Professor of Commerce [Retired],

Government Arts College, Udumalpet 642128

&

Co-Editor

Dr. N. Maliqjan

Associate Professor in the department of commerce,

Government Arts College

2025

First Edition: 2025

ISBN: 978-93-49028-02-9

Published by:

AG Volumes (an imprint of AG Publishing House)

58, Priyadarshini Phase-3, near Meenakshi Planet City, Shri Rameshwaram, Bagmugaliya, Bhopal, Madhya Pradesh 462043

Website: https://www.agphbooks.com

About the Book

Advanced Research Techniques: Theories, Methods, and Practices is a comprehensive and insightful guide that bridges the gap between foundational research knowledge and contemporary research innovations. This book is tailored for students, academics, and professionals across disciplines who seek to elevate their understanding of research methodologies in an ever-evolving academic and technological landscape.

The book covers a wide range of topics, from traditional research theories and frameworks to advanced methods, including big data analytics, artificial intelligence in research, computer-assisted qualitative analysis, and mixed-methods design. Each chapter is designed to offer a clear conceptual foundation, followed by practical applications and real-world examples, enabling readers to implement these techniques effectively in their own work.

Emphasis is placed on ethical research practices, data transparency, and methodological rigor. The book also addresses the growing importance of interdisciplinary research and equips readers with tools to navigate complex research questions using both qualitative and quantitative approaches.

Advanced Research Techniques is more than just a textbook—it is a strategic resource for developing the critical thinking, technical skills, and ethical awareness necessary for conducting high-impact research. Whether used as a course companion or a reference guide, this book empowers readers to explore, question, and innovate in their research endeavors.

Preface

In an era marked by rapid technological advancements and an ever-expanding body of knowledge, research has evolved beyond traditional paradigms. Advanced Research Techniques: Theories, Methods, and Practices is conceived as a comprehensive guide for students, scholars, and professionals seeking to deepen their understanding of cutting-edge research methodologies across diverse disciplines. This book serves as a bridge between foundational research principles and modern innovations, offering both theoretical grounding and practical insights into advanced methods of inquiry.

The need for rigorous, innovative, and ethical research practices has never been more pressing. From the rise of big data and artificial intelligence to the growing emphasis on interdisciplinary approaches, the landscape of research is undergoing a profound transformation. This book addresses these changes by introducing readers to contemporary tools, techniques, and frameworks that are reshaping the way we conduct, interpret, and apply research findings.

Each chapter blends theoretical exploration with practical application, providing step-by-step guidance and real-world case studies that exemplify the implementation of advanced research strategies. Topics include data analytics, mixed-methods approaches, digital tools in qualitative research, experimental design, and ethical considerations in modern contexts.

This book is the result of extensive collaboration and reflection, aiming to empower researchers to critically engage with emerging challenges and opportunities in their fields. Whether you are conducting your first major research project or refining your methodological toolkit, we hope this work will inspire a thoughtful, responsible, and innovative approach to research that contributes meaningfully to knowledge and society.

CONTENT

AG Publishing House

Advanced Research Techniques: Theories, Methods, And Practices (Volume-2)

Volume 2
Year: 2025

AG PH Books

New Techniques and Methods of Data Collection

Bachchu Singh Azad [1*]

[1]*Dept.of Zoology, D.A.V. College, Kanpur*

Abstract

Data is simply the information required to examine a research issue after a well-designed study. Data collection is essential because without the precise information obtained, research cannot be carried out. Review the many researchers' studies on data gathering strategies and procedures in this page. It concluded that the selection of appropriate data collection methods is crucial for enhancing research quality and efficiency. Different techniques have their strengths and limitations depending on factors such as participant demographics and research objectives. The rise of online platforms has further expanded data collection possibilities, necessitating adherence to ethical guidelines. Utilizing advanced tools and software optimizes data management and analysis, leading to insightful outcomes. Combining various analytical techniques ensures comprehensive interpretation and evidence-based decision-making. The practical application of a minimum dataset aids in systematic data planning, while pre-collected routine data remains valuable for ongoing research efforts.

Keywords: Data collection and gathering, Data collection techniques, Data collection methods, Techniques and methods, Tools and software, Data management and analysis, Decision-making, etc.

1 Introduction

Following the definition of the research problem and the development of the research design/plan, the duty for data collection is assigned. Information regarding circumstances, particular problems, or any other phenomenon is to be collected using two primary methods. On occasion, the essential information is already accessible and only requires extraction [1]. The researcher will be required to determine the type of data that will be utilised in their research and, as a result, will need to choose one of the available data acquisition methods. Data collection is defined as the process of acquiring, assessing, and analysing

[*] ISBN No. - 978-93-49028-02-9

exact perceptions for research using conventional authenticated methods [2]. The hypothesis can be evaluated by a researcher using the data that has been collected. Data acquisition is the most critical phase of research in nearly all instances, irrespective of the field of study [3]. The information that is essential in the various disciplines of study determines the data collection method. Data collection's main objective is to safeguard the reliable, information-rich data that is gathered for statistical analysis, which then helps researchers make data-driven choices [4].

A. Data collection

The act of data collection or data gathering involves the accumulation and measurement of information on specific variables within a predetermined system. This process allows for the evaluation of outcomes and the resolution of pertinent enquiries. Data collecting is a component of research in all fields of study, including business, the humanities, and the physical and social sciences [5]. The emphasis on the accumulation of accurate and honest data remains consistent, despite the fact that the methodologies used within each discipline may differ. Finding evidence that allows data analysis to support the creation of reliable answers to the questions addressed is the aim of all data collecting [6]. In order to preserve the integrity of research, it is imperative to acquire precise data, regardless of the field or preference for data definition (qualitative or numeric). The choice of appropriate data gathering tools (pre-existing, modified, or newly built) and the supply of explicit guidelines for their appropriate usage reduce errors [7].

B. Methods and techniques of data collection

In the workplace, there are numerous methods of data collection that may be implemented, such as:

Observation: The primary objective of observational techniques is to examine items and collect data about them. One way to do this would be to observe certain animals or people in their natural habitat. Avoiding direct interaction among researchers and the subjects of their studies may provide more trustworthy results [8].

Survey: The main focus of survey methods is gathering written or multiple-choice answers from respondents on a variety of subjects. People usually interact with these questions online, and there is often little to no interaction between survey distributors and respondents. Companies might utilise them to quickly get internal or external feedback.

Focus group: Focus group strategies aim to directly obtain information from participants. Instead than concentrating on numbers, this approach often emphasises sentiments, views, or emotions. Focus groups may be used by businesses to learn more about their customers.

Interview: Techniques for conducting interviews may be more personal when participants and the researcher speak face-to-face. Researchers may provide participants the questions before the interview so they have time to decide whether or not they feel comfortable participating. This strategy can include obtaining consent documents for video or audio recordings [9].

Design thinking: In order to produce original ideas or solutions, design thinking techniques may emphasise brainstorming among participants. Businesses may use this if they want to solve the challenges that consumers have while using their goods. Digital or in-person sessions may be used, depending on the participants' and researchers' locations.

User testing: Businesses often test their products or services either during or after they are created. In the event that customers choose to utilise the product during development, it might be used to pinpoint regions that they find challenging to use. Even after launching a product or service, people could use it to make improvements if they're interested [10].

Web Scrapping: Web scraping is a method of gathering data that involves mechanically extracting information from dynamic websites. You may get up-to-date data on rivals and market trends using tools like BeautifulSoup or Scrapy. You must be aware of data privacy regulations and terms of service infractions, however.

Log Files: Detailed records created by servers, apps, or devices are called log files. They record a chronology of interactions, transactions, and events that take place inside a system. Log files may be analysed and visualised using Splunk or ELK Stack to provide information on system performance, web traffic, and security events.

API Integration: By enabling information interchange between two software programs or systems, application programming interfaces, or APIs, automate data collecting. It offers a unified picture of the data and is scalable. This method is used, for instance, to extract datasets from cloud services or social media networks [11].

Transactional Tracking: Data collection about your customers' purchases is part of transactional tracking. Transactions done via websites, third-party services, or in-store point-of-sale systems may be monitored to get data on product combinations ordered, delivery locations, and other details. By analysing this information, you may target the best client categories and improve your marketing tactics.

Mobile Data Collection: Through applications, polls, and GPS monitoring, real-time data is directly gathered from users via mobile devices. Mobiles and tablets are widely used, which makes them perfect for collecting data while on the move.

Social Media Monitoring: Numerous social media sites provide data analytics tools that let you monitor engagement numbers, demographic data, and other information about your target audience. Information on consumer mood and new trends may be found using tools like Hootsuite or Brandwatch.

C. Common challenges in data collection

Some of the challenges that are often faced while collecting data include the following:

Data quality issues: Inaccuracies, inconsistencies, and other issues are common with raw data. Data gathering methods should ideally be created to prevent or lessen these issues. But generally speaking,

that isn't infallible. Consequently, gathered data often has to undergo data cleaning to address problems and data profiling to find them.

Finding relevant data: Because they must navigate a range of systems, gathering data for analysis may be a difficult task for data scientists and other users within an organisation. Data curation techniques may help make data easier to find and access. For instance, such may include developing searchable indexes and a data catalogue.

Deciding which data to collect: These are fundamental issues that pertain to both the initial raw information collection and the data collection by consumers for analytics applications. Time and expense are increased as a result of the process being complicated by an excessive amount of data acquisition. On the other hand, the exclusion of valuable data may have an impact on the results of analytics and diminish the economic value of a data set.

Dealing with big data: Large amounts of semistructured, unstructured, and structured data are often found in big data contexts. As a result, the earliest phases of data collecting and processing become increasingly intricate. In addition, data scientists frequently must filter collections of unprocessed data stored in a data lake for specific analytics applications.

Low response and other research issues: The validity of the data gathered in research projects is called into doubt when there are no consenting participants or answers. To ensure the accuracy of the data, it is necessary to devise appropriate quality assurance methods and train personnel to collect it. These are additional research obstacles.

2 Literature Review

(IYANUOLUWA Blessing et al., 2024)[12] Gives a summary of the key methods for collecting and analysing data. Qualitative, descriptive, inferential, and predictive analyses are the different categories of analysis approaches. The efficiency and quality of data management are improved by a variety of tools and software, including statistical programs and platforms for data visualisation. Ensuring appropriate data management requires careful consideration of ethical factors, such as informed permission, data privacy, and integrity. New techniques and instruments are always being developed as technology advances, influencing how data collecting and analysis will develop in the future. This abstract emphasises the significance of using the right methods to extract trustworthy and useful insights from data.

(Cheong et al., 2023) [13] Lays forth a novel, methodical approach to secondary qualitative investigation using publicly accessible, online interview data. By using a procedural approach, the research community may extend its datasets, boost rigour, and actively address and minimise possible hazards. To address the predetermined research questions, the data analysis approach consists of theme discourse analysis and content analysis for dataset classification. The technique also discusses the legal and ethical issues surrounding the use of secondary web data and the publication of research conclusions derived

from it. We illustrate how the technique gives secondary qualitative research structure by using the topic of forced migration as an example.

(Kwok et al., 2022) [14] The importance of data analytics in guiding healthcare choices has generated a lot of interest in the field. The purpose of this article is to introduce the idea of a minimal dataset for any variable. The value of a variable represents the most fundamental level of data collecting. The way a variable's values changed over time is described by its time course. To prevent erroneous results, it's also critical to consider the validity or correctness of a variable's values. To sum up, the minimal dataset is a paradigm that may be used to research design and evaluation. The minimal dataset framework for each variable does not need to be fully taken into account for every data, but if more in-depth findings are sought, the framework could be crucial.

(Qiu et al., 2022) [15] Concentrate on researching ways to collect data for assessing socially assistive systems (SAS). The findings showed that the most common targets of the SASs that were in place were youngsters with autism, elderly folks, and those who had visual impairments. Most often, observation data, questionnaires, and interviews were used to evaluate SASs. The interview studies consisted of target users in approximately half of the cases, while the other half contained "stakeholders or secondary users". Surveys were primarily utilised to gauge the social interaction, emotional state, and system use of older folks and those with visual impairments. Most observational studies included users in certain age groups, including elderly individuals and autistic children. As a result, we provide a summary of the approaches used to collect data from distinct SAS target users. In order to guide future research and development, pertinent insights are retrieved.

(Newman et al., 2021) [11] Researchers are using online tools like Amazon Mechanical Turk (MTurk) more and more to gather survey and experimental data. However, these platforms can represent a turbulent environment for reviewers and researchers alike. In order to ensure data quality, ethically encourage participant involvement, preserve openness, and collect representative samples from online participant cohorts, researchers must manage many challenges. The challenges of assessing how well such data gathering and execution initiatives address significant research concerns, however, must be handled by reviewers. This article offers a number of suggestions for researchers and reviewers on how to properly conduct and assess data gathering using online platforms in order to shed light on these problems.

(Syeda Ayeman Mazhar et al., 2021) [10] In essence, data is the information required to investigate a research topic that has been properly designed. The significance of data collection stems from the fact that research cannot be carried out without the precise information being obtained. It might be original or secondary data. Primary data collection methods in the behavioural sciences often use questionnaires, interviews, database access, and observational methods. Secondary data sources include "unpublished biographies and autobiographies, books, magazines, journals", and other earlier editions. Consequently, data collecting is the primary research instrument as it is necessary to complete the research process. The several approaches and strategies of obtaining data for doing research are thoroughly reviewed in this text.

(Flanagan et al., 2015) [16] The usefulness of the data gathering techniques used while doing research with children, teens, and young adults is summarised. There is a discussion of the advantages and disadvantages of the variety of data gathering techniques that are used, including whether certain techniques are suitable for certain age groups or which approach is best to use when examining a particularly delicate subject. Numerous data gathering techniques are used while doing research with children, adolescents, and young adults. The evidence that is currently available is included in this review, together with an outline of the advantages and disadvantages of the data gathering techniques used.

3 Conclusion

The advancement of data collection techniques has significantly enhanced research methodologies across various disciplines. The selection of appropriate methods depends on factors such as participant demographics, research objectives, and ethical considerations. Arts-based methods may be more suitable for younger participants, while focus groups and digital technologies are often preferred for adolescents. The rise of online platforms has provided researchers with efficient ways to gather data, though ethical guidelines must be strictly followed. Primary data collection methods, including observations, interviews, and questionnaires, remain fundamental in behavioral sciences, while secondary data sources such as books, journals, and biographies offer valuable supplementary insights. The integration of advanced tools and software has streamlined data collection and analysis, improving research efficiency and collaboration. Employing a combination of descriptive, inferential, predictive, and qualitative techniques allows for a deeper understanding of data, enabling evidence-based decision-making. The concept of a minimum dataset plays a crucial role in defining data collection strategies, particularly in routine data scenarios where information is pre-existing. By leveraging modern data collection techniques, researchers can enhance the reliability, accuracy, and impact of their findings, ultimately contributing to knowledge advancement and informed decision-making across diverse fields.

References

[1] E. Paradis, B. O'Brien, L. Nimmon, G. Bandiera, and M. A. T. Martimianakis, "Design: Selection of Data Collection Methods," J. Grad. Med. Educ., vol. 8, no. 2, pp. 263–264, 2016, doi: 10.4300/JGME-D-16-00098.1.

[2] I. Ali, I. Ahmedy, A. Gani, M. Talha, M. A. Raza, and M. H. Anisi, "Data collection in sensor-cloud: A systematic literature review," IEEE Access, vol. 8, pp. 184664–184687, 2020, doi: 10.1109/ACCESS.2020.3029597.

[3] H. Taherdoost, "Data Collection Methods and Tools for Research; a Step-by-Step Guide to Choose Data Collection Technique for Academic and Business Research Projects Hamed Taherdoost. Data Collection Methods and Tools for Research; A Step-by-Step Guide to Choose Data Coll," Int. J. Acad. Res. Manag., vol. 2021, no. 1, pp. 10–38, 2021, [Online]. Available: https://hal.science/hal-03741847

[4] E. Gilbert, “New technologies and innovative methods in data collection: scoping review,” 2020, [Online]. Available: www.cls.ucl.ac.uk.

[5] S. R. Niraula, “A review of research process, data collection and analysis,” Insights Biol. Med., vol. 3, no. 1, pp. 001–006, 2019.

[6] P. Pandey and M. M. Pandey, RESEARCH METHODOLOGY: TOOLS AND TECHNIQUES. 2015. doi: 10.1007/978-981-16-6281-2_5.

[7] Y. Zhang, “Book Review: Data Collection Research Methods in Applied Linguistics,” Front. Psychol., vol. 12, no. March, pp. 1–2, 2021, doi: 10.3389/fpsyg.2021.668712.

[8] M. Arrthi, “TOOLS & TECHNIQUES OF DATA COLLECTION IN RESEARCH METHODOLOGY,” vol. 3, no. 4, pp. 555–573.

[9] O. P. Aborisade, “Data collection and new technology,” Int. J. Emerg. Technol. Learn., vol. 8, no. 2, pp. 48–52, 2013, doi: 10.3991/ijet.v8i2.2157.

[10] Syeda Ayeman Mazhar, R. Anjum, A. I. Anwar, and A. A. Khan, “Methods of Data Collection: A Fundamental Tool of Research,” J. Integr. Community Heal. vol. 10, no. 01, pp. 6–10, 2021, doi: 10.24321/2319.9113.202101.

[11] A. Newman, Y. L. Bavik, M. Mount, and B. Shao, “Data Collection via Online Platforms: Challenges and Recommendations for Future Research,” Appl. Psychol., vol. 70, no. 3, pp. 1380–1402, 2021, doi: 10.1111/apps.12302.

[12] A. IYANUOLUWA Blessing, Z. Henry, and D. Micheal, “Data Collection andAnalysis Techniques,” ResearchGate, no. December, 2024.

[13] H. I. Cheong, A. Lyons, R. Houghton, and A. Majumdar, “Secondary Qualitative Research Methodology Using Online Data within the Context of Social Sciences,” Int. J. Qual. Methods, vol. 22, pp. 1–19, 2023, doi: 10.1177/16094069231180160.

[14] C. S. Kwok, E. A. Muntean, C. D. Mallen, and J. A. Borovac, “Data Collection Theory in Healthcare Research: The Minimum Dataset in Quantitative Studies,” Clin. Pract., vol. 12, no. 6, pp. 832–844, 2022, doi: 10.3390/clinpract12060088.

[15] S. Qiu, P. An, K. Kang, J. Hu, T. Han, and M. Rauterberg, “A review of data gathering methods for evaluating socially assistive systems,” Sensors, vol. 22, no. 1, pp. 1–31, 2022, doi: 10.3390/s22010082.

[16] S. M. Flanagan, S. Greenfield, J. Coad, and S. Neilson, “An exploration of the data collection methods utilised with children, teenagers and young people (CTYPs),” BMC Res. Notes, vol. 8, no. 1, 2015, doi: 10.1186/s13104-015-1018-y.

AG Publishing House

Advanced Research Techniques: Theories, Methods, And Practices (Volume-2)

Volume 2
Year: 2025

AG PH Books

Advanced Techniques in Environmental Biotechnology: Addressing Global Challenges

Dr. V. Sailaja [1*]

[1]*Assistant Professor*

Abstract

Ecological biotechnology is a contemporary method of environmental protection and preservation that effectively integrates the advancements of microbiology, biochemistry, genetic engineering, and chemical technologies. This article surveys the diverse literature on the challenges and advancements in environmental biotechnology. It concluded that, biotechnology is revolutionizing environmental sustainability by offering innovative solutions to global challenges through advanced waste processing, agricultural enhancements, and the creation of sustainable industrial cycles. Its expanding role in everyday life and economic development highlights its transformative potential. Despite ongoing ethical, legal, and environmental concerns, extensive research supports the safe and beneficial use of biotechnological innovations, such as GM crops. Continued progress in this field requires responsible development, balancing scientific advancement with societal and ecological considerations to ensure a sustainable future for all.

Keywords: Ecological or environmental biotechnology, Ethical, and legal concerns, Waste management, Genetically modified, Disease, etc.

1 Introduction

Biotechnology is a vast topic that encompasses the modification and enhancement of certain traits of microbes, plants, and animals to enable new products, methods, and organic entities intended to benefit human health. This multifaceted discipline is important to many fields, such as environmental biotechnology, food technology, genetic technology, medicine, and agriculture [1]. Utilizing some of the

* ISBN No. - 978-93-49028-02-9

most advanced biotechnological techniques, such as gene modification, chimeric DNA technology, and life sciences, to create a product that benefits people and does not have negative environmental effects also provides social, economic, and environmental benefits [2]. One of the key areas where the most amazing adventures take place is the environment. Some of the major issues the world is now dealing with include population growth, resource scarcity, biodiversity loss, rising pollution, and climate change [3], [4]. Man's progress in industry, transportation, agriculture, housing, and other sectors has led to greater pollution of the air, water, and land. Various instruments, methods, and strategies are used to address these significant issues. Environmental biotechnology has shown to be a useful tool for limiting or resolving these issues while also promoting sustainable development. Its contributions to environmental protection, conservation, and preservation have been substantial [5], [6].

A. Environmental biotechnology

Applying and studying biotechnology to the natural world is known as environmental biotechnology. Environmental biotechnology may also mean that biological processes are being tried to be exploited and used for profit [7]. According to the International Society for Environmental Biotechnology, environmental biotechnology is the creation, application, and control of biological systems for the cleanup of polluted areas (land, air, and water) and for environmentally friendly procedures (sustainable development and green manufacturing technologies) [8]. Simply put, environmental biotechnology is the best possible use of nature—plants, animals, bacteria, fungi, and algae—to create food, nutrients, and renewable energy in a profit-making cycle of integrated, synergistic processes where waste from one process is used as feedstock for another [9], [10].

B. Applications of environmental biotechnology

Four primary categories of environmental biotechnology applications exist. They consist of the following: Blood marker Energy derived from biomass Remediation through biological processes [11]. The process of biological transformation

Biomarker: The chemical response to this type of environmental biotechnology application is used to quantify the extent of injury, hazardous exposure, or contamination effect. For example, biomarkers are also known as biological markers. The primary function of these applications is to establish a connection between the oils and their sources.

Bioenergy: Bioenergy refers to the common purpose of biogas, biomass, fuels, and hydrogen. This environmental biotechnology application is employed in the industrial, domestic, and space sectors. It has been determined that the current urgent need is for pure energy from these fuels and alternative methods of obtaining clean energy.

Bioremediation: The procedure known as bioremediation is used to remove harmful materials and replace them with non-toxic molecules. The main use for this procedure is any technological cleanup that makes use of natural microbes.

Biotransformation: The term "biotransformation process" refers to the biological changes that occur in the environment, such as the conversion of complex compounds from simple, non-toxic to poisonous, or vice versa. The manufacturing industry uses it to turn hazardous materials into byproducts.

C. Challenges in environmental biotechnology

Ethical and sociocultural challenges: Biotechnology presents intricate ethical dilemmas concerning the moral implications of human embryos, cloning, and organismal modification. The regulation and assurance of the safety of biotechnology use necessitate the establishment of effective ethical standards and laws [12].

Safety and environmental risks: Environmental consequences may be unpredictable as a result of the application of biotechnology. In order to mitigate and prevent potential adverse environmental consequences, ample consideration must be given to research [13].

Regulation and legislation: The accelerated advancement of biotechnology is not always consistent with the current legislation. In order to guarantee the safety and effective administration of biotechnological processes, it is imperative to establish a suitable legal framework.

Technological challenges and innovation: The advancement of new biotechnology technologies and methodologies holds immense potential for the enhancement of the quality of life, the increase in agricultural productivity, and the alleviation of disease. Development of novel methodologies and enhancements to extant technologies ought to be the primary objectives of research and innovation.

Global collaboration: The exchange of knowledge, experience, and technology in biotechnology is only possible through strategic international cooperation. This collaboration will enable the development of innovative solutions to global challenges, including the prevention of infectious diseases and the promotion of sustainable development.

2 Literature Review

(Macwan et al., 2025)[14] The primary objective of agricultural biotechnology is to ensure food security by creating genetically modified (GM) crops that are more resistant to parasites, have higher yields, and are more adaptable to climate change. Biotechnology enhances the efficacy of industrial processes by employing microorganisms and enzymes, thereby participating in sustainable production. Eco-friendly solutions are promoted and the resilience of economies and societies is enhanced by its critical role in sustainable development. The discipline, however, poses ethical and regulatory concerns, including the equitable distribution of benefits and the risks associated with genetic engineering. Global well-being can be enhanced through responsible innovation and the optimization of biotechnology's potential through international collaboration among industries, governments, and scientists. As a transformative science, biotechnology continues to influence the future by achieving a harmonious balance between innovation, sustainability, and equity.

(Kuppan et al., 2024) [15] The review emphasizes the increasing significance of bioremediation in the reduction of industrial effluents, contaminated soils, and groundwater. It is anticipated that future developments will improve its efficiency and applicability. Exploring the potential of recent developments in IoT, AI, and biosensors to revolutionize waste management and bioremediation. AI enhances data analysis and predictive modeling, while IoT enables real-time monitoring and remote management, and biosensors contribute to precise pollutant detection and environmental monitoring. The review emphasizes the synergistic integration of these technologies, as it introduces intelligent bioremediation systems that possess adaptive capabilities and real-time feedback loops. This represents a substantial advancement in the direction of sustainable environmental management, as these technologies provide scalable solutions for the mitigation of environmental contamination.

(Verdezoto-Prado et al., 2024) [16] Due to its comprehensive applicability and cost-effectiveness, CRISPR/Cas9 has become the most widely used method for genome editing, significantly contributing to the advancement of sustainable practices in a variety of sectors. The research divides the scientific advancements into three trends: biofuel production, gene editing techniques, and agricultural advancements. The utilization of CRISPR/Cas9 in the development of fourth-generation biofuels and environmental biosensors, as well as its applications in the enhancement of genetic resilience and the control of invasive species, are among the most significant topics of discussion. The potential of CRISPR/Cas9 to promote sustainable resource management and energy generation is underscored by these innovations, which represent a critical contribution to ecological conservation and sustainability initiatives.

(Ibragimova et al., 2023) [12] The manipulation of biological organisms or their components to generate beneficial products is known as biotechnology. This interdisciplinary discipline is innovative and has a significant impact on a variety of industries, such as agriculture, veterinary medicine, medicine, pharmaceuticals, and specialty chemicals. It is one of the primary technologies for sustainable production. Nevertheless, biotechnology presents considerable obstacles and offers some promising opportunities. This chapter examines the diverse domains in which biotechnology presents challenges and the potential for it to serve as a solution in the future if applied effectively.

(Smirnova et al., 2023) [17] The article offers a comprehensive account of prospective development areas and identifies the current challenges of environmental biotechnology as a promising and evolving field that affects both humans and the environment. Ecological biotechnology has been realized as a scientific field as a result of the aspiration to enhance the quality of life through the application of cutting-edge technology. The challenges that are associated with the implementation of a variety of environmental improvement strategies are revealed in the article's definition of its goals and objectives. In this context, the perspectives for the advancement of biotechnology as a scientific discipline, which is inextricably linked to the preservation of the health of humans and animals, are examined.

(Nezhmetdinova et al., 2020) [18] This paper presents a novel classification of the risks associated with the introduction of biotechnologies in agriculture, including food, agricultural, environmental, patent, social, and ethical risks. Timely notification of the potential hazards of specific products was a challenge

for numerous national food safety systems in the past. This may account for the observed phenomenon. In numerous nations, the rejection of specific products and the manipulation of genetically modified organisms (GMOs) can be attributed to social and ethical views. Many of these conflicts are indicative of more fundamental concerns regarding the interaction between human society and nature, which must be thoroughly considered in any endeavor to facilitate social communication. The authors propose agrobioethics as a method of social governance and consensus, which is comparable to bioethics in biomedical technologies.

(Agarwal, 2016) [19] Biological technology encompasses a diverse array of specialized disciplines, including the most advanced genetic engineering techniques and the age-old fermentation process. Biotechnological methods are essential in the purification of waste water, soil, and exhaust air. In addition, the efficiency of biowaste recycling and purification facilities is enhanced by microbial and system biology. However, biotechnology is not only instrumental in the remediation of environmental harm but also in the detection of such damage. However, the environment is frequently exposed to hazards in addition to the advantages that most advanced technologies provide to humanity. Consequently, in order to ensure the safe and effective use of biotechnology, it is imperative to choose genetically engineered microorganisms that are both secure and applicable to waste treatment processes of particular environmental importance.

3 Conclusion

In conclusion, biotechnology has emerged as a transformative force in addressing global environmental challenges, significantly enhancing both the quality of human life and economic development. No longer confined to the realm of applied science, biotechnology is becoming an integral part of everyday life, offering sustainable solutions to pressing issues such as waste management, agricultural productivity, and industrial efficiency. Its expanding role in creating closed-loop systems in the chemical industry and agriculture underscores its potential in fostering environmental sustainability. The development of genetically modified (GM) crops illustrates biotechnology's capacity to tackle food security, climate change, and environmental degradation. Despite ongoing debates, extensive research affirms the safety and advantages of GMOs for human health and ecological systems. However, the advancement of biotechnological methods must be carefully navigated, considering ethical, legal, social, and environmental implications. Continued innovation, guided by responsible research and regulatory oversight, will be critical to maximizing the benefits of environmental biotechnology while minimizing associated risks.

Reference

[1] N. K. Arora and T. Fatima, “Environmental sustainability and biotechnology: opportunities and challenges,” Environ. Sustain, vol. 7, no. 2, pp. 115–119, 2024, doi: 10.1007/s42398-024-00317-9.

[2] J. Awewomom et al., “Addressing global environmental pollution using environmental control techniques: a focus on environmental policy and preventive environmental management,” Discov. Environ, vol. 2, no. 1, 2024, doi: 10.1007/s44274-024-00033-5.

[3] P. Gortares-Maroyoqui et al., “Advances in environmental biotechnology and engineering 2018,” Environ. Sci. Pollut. Res., vol. 27, no. 23, pp. 28463–28468, 2020, doi: 10.1007/s11356-020-09377-w.

[4] D. A. Sreenivas, “Ecological and Evolutionary Responses to Recent Climate Change in India: A Review,” Int. J. Innov. Sci. Eng. Manag., vol. 37, pp. 637–669, 2023, doi: 10.1146/annurev.ecolsys.37.091305.110100.

[5] A. Kaur, J. Jaiswal, and M. Sharma, “Challenges in Environmental Biotechnology,” Microb. Biotechnol. Food Ind., no. March, 2024, doi: 10.1007/978-3-031-51417-3.

[6] L. Hennen, J. Hahn, M. Ladikas, R. Lindner, W. Peissl, and R. van Est, Technology Assessment in a Globalized World: Facing the Challenges of Transnational Technology Governance. 2023. doi: 10.1007/978-3-031-10617-0.

[7] K. Pakshirajan, E. R. Rene, and A. Ramesh, “Biotechnology in environmental monitoring and pollution abatement,” Biomed Res. Int., vol. 2014, 2014, doi: 10.1155/2014/235472.

[8] J. Aminian-Dehkordi et al., “Synthetic biology tools for environmental protection,” Biotechnol. Adv., vol. 68, no. August, p. 108239, 2023, doi: 10.1016/j.biotechadv.2023.108239.

[9] A. Schwartz, “Applications of Environmental Biotechnology : Aerobic Treatment and Bioremediation,” vol. 13, no. 1000472, pp. 1–2, 2022.

[10] D. K. Kumari, “RECENT TRENDS IN MUNICIPAL SOLID WASTE MANAGEMENT IN PATNA, A REVIEW,” Int. J. Innov. Sci. Eng. Manag. OPEN, pp. 146–151, 2024.

[11] D. K. Martin et al., “A brief overview of global biotechnology,” Biotechnol. Biotechnol. Equip., vol. 35, no. 1, pp. 354–363, 2021, doi: 10.1080/13102818.2021.1878933.

[12] Z. M. Ibragimova, E. F. Amirova, and O. N. Anokhina, “Challenges and future prospects for biotechnology,” BIO Web Conf., vol. 76, 2023, doi: 10.1051/bioconf/20237607007.

[13] A. Krupa, V. Nikitenko, and V. Voronkova, “THE ROLE OF DIGITAL TECHNOLOGIES IN ADDRESSING TODAY’S GLOBAL CHALLENGES,” Balt. J. Econ. Stud., vol. 10, no. 3, pp. 193–199, 2024.

[14] A. Macwan, D. Desai, N. Vandra, K. K. Mishra, and Y. Jha, “Biotechnology for Global Sustainability : Innovations, Applications, and Challenges,” Int. Res. J. Adv. Eng. Manag., pp. 59–65, 2025.

[15] N. Kuppan, M. Padman, M. Mahadeva, S. Srinivasan, and R. Devarajan, “A comprehensive review of sustainable bioremediation techniques: Eco friendly solutions for waste and pollution management,” Waste Manag. Bull., vol. 2, no. 3, pp. 154–171, 2024, doi: 10.1016/j.wmb.2024.07.005.

[16] J. Verdezoto-Prado et al., "Advances in Environmental Biotechnology with CRISPR/Cas9: Bibliometric Review and Cutting-Edge Applications," Preprints.Org, 2024, doi: 10.20944/preprints202405.0100.v1.

[17] E. Smirnova, N. Pimenov, R. Ivannikova, A. Laishevtsev, and A. Konovalov, "Problems and the future of environmental biotechnology," E3S Web Conf., vol. 463, 2023, doi: 10.1051/e3sconf/202346302013.

[18] F. T. Nezhmetdinova, M. E. Guryleva, N. K. Sharypova, R. I. Zinurova, and A. R. Tuzikov, "Risks of modern biotechnologies and legal aspects of their implementation in agriculture," BIO Web Conf., vol. 17, 2020, doi: 10.1051/bioconf/20201700227.

[19] P. Agarwal, "Biotechnology in Environmental Protection," Biochem. Cell. Arch., vol. 16, no. January, pp. 74–79, 2016.

AG Publishing House

Advanced Research Techniques: Theories, Methods, And Practices (Volume-2)

Volume 2
Year: 2025

AG PH Books

A Review of Sampling Techniques in Social Science Research

Dr. Anviti Rawat [1*]

[1]*Associate Professor,University School Of Education,Guru Gobind Singh Indraprastha University.*

Abstract

The sampling procedure has a significant impact on a study's accuracy. Because they have the power to determine the validity of results, choosing the right sample size and selection procedure for a given research problem is essential. Examine the many studies on sampling methods in social science research that have been published in the literature. This review highlights the critical role of sampling techniques in social science research, emphasizing the trade-offs between probability and non-probability methods. While probability sampling ensures representativeness and statistical reliability, non-probability sampling offers practical advantages in terms of cost, speed, and feasibility, especially in digital research contexts like Twitter. However, limitations such as lack of generalizability and inability to compute confidence intervals must be acknowledged. Ultimately, the choice of sampling technique should align with the research objectives, available resources, and population characteristics to ensure valid, reliable, and meaningful findings in social science inquiry.

Keywords; Sampling techniques, Cluster sampling, Probability and non-probability methods, Sample size, Population size, Convenience, and snowball sampling, etc.

1 Introduction

Finding solutions to problems, gaining more knowledge about them, or devising a better approach to them are all considered forms of research. For research to accomplish its objective, it must always be planned and executed methodically. The use of sampling procedures is crucial in “quantitative research” [1]. A representative sample is obtained by selecting only a few of units or examples from a much broader

* ISBN No. 978-93-49028-02-9

set or population. This is the primary goal of sampling. From the findings of the smaller sample, the researcher can make reliable generalisations about the larger group [2]. The techniques that would provide highly similar samples are the main focus of the researchers. They make use of probability sampling, a kind of sampling that is based on mathematical ideas about the likelihood that certain occurrences would occur. The sample data allows for the drawing of conclusions about the whole population [3]. A sufficiently representative probability sample is necessary to enable the study population to make inferences. The lack of funds and time would make it difficult to conduct a comprehensive examination of the whole population in any kind of research. A sample, as used in statistics, is the portion of the population from which inferences about the population may be made [4]. For the researcher, gathering data with an appropriate sample strategy is a difficulty. The researchers won't be able to collect data from every case, thus the present techniques for answering the data gathering questions are problematic. To choose which of the many sampling techniques and processes available is best for the specific topic being examined, the researcher must be aware of the differences between them [5]. The sample size must be enough to allow for significant conclusions to be drawn. This means that the lowest sample size needed to estimate the true population percentage with the specified degree of confidence and margin of error is this one [6]. Consequently, a typical problem in statistical analysis is determining the appropriate sample size. Its equation may be found using the normal distribution's crucial parameter, population size. It serves as the foundation for all research designs and is also a cost-effective and time-saving approach [7], [8].

A. Sampling

Sampling is the intentional process of selecting a subset or individuals from a group with the intention of predicting the characteristics of the entire population and deriving statistical conclusions. Given the expense, effort, and often impossibility of studying the whole population, it offers a practical and practical way to examine its traits. Using a range of sampling methodologies, market researchers collect samples from a large population to provide relevant insights. A study's goal, the time and money available, and the research hypothesis are some of the factors that influence the optimal sampling approach [9], [10].

B. Sampling techniques

Study participants are selected using two different methods—probability sampling and non-probability sampling—to differentiate between the two. Everyone has pros and cons of its own. There is a known, nonzero possibility of selection for every member of the population when using probability sampling. This design's randomisation reduces selection bias and makes the sample representative of the whole population [11]. With probability sampling, the researcher may use the sample to make inferences about the population and predict the error of sampling with a certain level of confidence. This category includes techniques including basic random "sampling, stratified sampling, and cluster sampling". Probability sampling is precise but sometimes time-consuming and difficult since it relies on a thorough inventory of the population and choices must be made via a complicated process, especially when the target population is vast or scattered [12].

C. Probability sampling methods

The sampling technique in question is characterised by an equal chance of selecting each unit or part of the population for the final sample. Random sampling emphasises the element of sample selection that is random and has a non-zero probability. By ensuring a more impartial and representative sample, this sampling method permits reliable conclusions about the whole population [9].

Simple random sampling: In spite of its simplicity, the random sample ensures that each member of the population has a fair opportunity of being selected. Following the definition and creation of a comprehensive list of population members, the sample size is determined, and then the real random selection process—either by drawing lots or using a random number generator—begins. Though logistically challenging, particularly when the population is large or scattered, this is highly prized since it lessens selection bias.

Stratified sampling: Using consistent criteria like age, sex, or education, stratified sampling would separate the population into discrete groups, or strata. From each stratum, random samples would then be selected to provide a sample representative of the entire spectrum of categories. Given that each subgroup is well-represented, this will improve the estimates' accuracy and make the technique more appropriate for populations with high levels of variability in important features. But without a thorough understanding of the population, stratification is hard to establish in a meaningful way.

Cluster sampling: Cluster sampling is a method that simplifies the data collection process by performing a random selection of individuals for analysis from a sample that has been divided into clusters, which may represent institutions or geographic regions. The sample will then include each member of the selected clusters. The compilation of a comprehensive inventory of the population would be nearly impossible in large-scale research. Consequently, this will be extremely beneficial. The cost of cluster sampling is low; however, biases may be introduced if the clusters selected are not representative of the entire population.

Systematic sampling: The systematic sampling method involves the continuous selection of each Kth person from a population list, beginning at a random point. The method is both straightforward and effective when applied over a consistent sample period. However, systematic sampling may be biassed if the population exhibits a natural pattern that corresponds to the sample interval.

Multi-stage sampling: In multi-stage sampling, a population is divided into numerous stages or levels. Cluster sampling, for example, is a sophisticated form of multi-stage sampling that involves the division of populations into substantial clusters (such as areas or organisations), from which additional random samples are obtained in a succession of stages.

For example, a random sample of individuals from each area may be selected when geographical areas are selected. By concentrating on broad or dispersed populations, stepwise stratification aids in gradually reducing the sample size in large-scale studies. Even while this might save money and time, if the sample is not representative at every step of the selection process, there is a chance that sampling error will rise.

D. Non-probability sampling methods

Another sampling method is non-probability sampling, which essentially involves drawing data by choosing individuals non-randomly based on preset criteria. Although the final sample may not be a perfect representation of the total population, this provides a simple method of collecting data [9].

Convenience sampling: Among the several data collecting techniques, convenience sampling entails selecting samples from the population segment that the researcher finds most accessible. When resources or time are scarce, this is an expeditious and cost-effective method of acquiring data. Although the sample may not be representative, this approach often increases bias since it makes it impossible to extrapolate the results to the whole population.

Purposive sampling: The researcher uses purposive sampling, often referred to as judgemental or expert sampling, to choose participants based on their potential value in providing the needed data. Research that calls for specific individuals with certain traits or specialities often uses this. Although the approach may provide very valuable insights, it bears a significant risk of researcher bias and may not be representative of the whole community.

Snowball sampling: When examining populations that are hidden or difficult to reach, snowball sampling is used. In this instance, participants must create a network of referrals by recommending others who meet the requirements for research participation. In exploratory research, this method is helpful. But since social networks drive samples to become homogenous over time, it has been argued that this results in biassed samples.

Quota sampling: In order to satisfy specific quotas within each segment, quota sampling is a method of selecting participants from a group that has been divided into mutually exclusive subgroups. The sample is biassed and less representative due to the fact that selection within the subgroup is not random, despite the fact that quotas ensure that the sample has specific demographic features.

2 Literature Review

(Ahmed, 2024) [13] Thoroughly revised the rules governing sample size computation and sampling techniques, providing sufficient proof to help researchers boost the validity and statistical strength of their work. In exploratory settings, non-probability sampling is helpful, but only probability can guarantee generalisability. A further critical procedure is the determination of the optimal sample size, which necessitates taking into account of several factors: the confidence level, the margin of error, the effect size, the size of the entire population, and the statistical power. The study provides researchers with the theoretical direction and useful tools they need to choose suitable sampling procedures and validate methodologically sound sample size estimations.

(Ali & Hatef, 2024) [14] It is the objective of this methodological review to underscore the importance of sampling and sample size estimation methodologies and to offer suggestions for the precise definition of sample sizes. The sample size for research on health is determined using methods that are elucidated with examples, depending on the type of study design. Researchers need to determine acceptable

accuracy levels, study power, the confidence level to be utilised, and the statistical analysis that will be used in order to estimate sample size. This method review provides a comprehensive and straightforward explanation of "the sample size formula and sampling technique", which will assist in the demystification of the complexities associated with statistical equations for sample size estimates in social health care and science research.

(Adeoye, 2023) [15] Provides a comprehensive outline of the steps involved in producing articles for publication in a variety of academic subjects, including data collection methods and the use of sample methodologies. Probability and non-probability sampling are two types of sampling procedures used in statistics. One advantage of probability sampling is that it guarantees the sample's representativeness of the body. Among the several types of probability sampling techniques are "stratified, cluster, systematic, and basic random sampling". In non-probability sampling, the sample is selected using personal judgement. Since non-probability sampling incorporates the representation of every topic in the population, it is seen to be the optimum method. Numerous non-probability sampling techniques exist, such as convenience sampling, quota sampling, judgemental sampling, snowball sampling, and sequential sampling.

(Makwana et al., 2023) [16] A study's sampling procedure has a big impact on its accuracy. The many sampling techniques used in research are summarised in the article. Probability sampling techniques and non-probability sampling techniques are the two basic categories into which these methods may be divided. Simple, systematic, and stratified random sampling are examples of probability sampling procedures. Non-probability sampling methods, on the other hand, include purposive, convenience, snowball, self-selection, and quota sampling.

(Mishra & Lavater, 2023) [7] Sampling is a vital part of every research project. Choosing the right sample size and selection method for your specific research problem is essential since they may determine the reliability of your results. In order to get reliable and accurate data, this article will go over the optimal sample size, some of the most popular sampling methods, and how they could be used in practice. This study looks at the fundamentals of probability sampling as well as the benefits and drawbacks of non-probability sampling and various probability sampling methods. Social science scholars will find this study useful in choosing the most suitable sampling method.

(Vicente, 2023) [17] Since Twitter has a vast user base and vast amounts of data, sampling is usually necessary when utilising it for study. Despite its importance for legitimate and trustworthy research results, there is currently a lack of expertise about how to choose representative samples of Twitterverse users. This report presents a thorough quantitative literature assessment of sample strategies developed and used in social science research on Twitter. The review covers the following topics: (1) target population definition; (2) sampling frames used to support sample selection; (3) sampling techniques used to get Twitter user samples; (4) gathering of information from Twitter users; (5) “sample size; and (6) research validity”. This review may be used as methodological advice for scholars and professionals who want to do social science research with Twitter users and the Twitterverse.

(Zickar & Keith, 2023) [18] Examine various sampling strategies, such as snowball, probability-based, convenience, and purposeful sampling. To assist organisational researchers in selecting the best sample strategies for their research issues, we point out the advantages and disadvantages of each strategy. In order to enhance the quality of online sampling, we analyse screening methods and highlight recommended practices that researchers may apply to improve the quality of their samples. In order to draw conclusions about the lack of methodological and sample diversity in organisational research, the over-reliance on a small number of sampling techniques, the necessity of disclosing important aspects of sampling, and participant quality concerns, we lastly reviewed the sampling practices of all empirical research articles published in the Journal of Applied Psychology during the previous five years.

(Raifman et al., 2022) [19] Gave instructions on how to use one of the most often used strategies, respondent-driven sampling (RDS), and a summary of sampling techniques for groups that are difficult to reach. Estimates derived from convenience or non-probability-based sample data may be skewed or not applicable to the intended audience. In an attempt to produce representative samples, inclusion-influencing variables may be evaluated and taken into consideration in RDS and time-location sampling (TLS). RDS is especially well-suited to connect with the most obscure members of communities that are difficult to reach. In addition to more generalisable estimates of population characteristics, TLS, RDS, or a combination of the two may provide a rigorous approach to find and enrol samples from communities that are difficult to reach. These techniques should be added to the toolkits of researchers that are interested in sampling populations that are difficult to reach.

(Rahman et al., 2022) [20] It might be difficult for a researcher to get data using a suitable sampling strategy. The study's research questions cannot be addressed in their present form as the researchers will not be able to gather data from every scenario. To choose the best sampling technique or method for the particular study being considered, the researcher has to be aware of the distinctions between the vast array of possible sampling techniques and procedures. In this regard, the research also examines the fundamental ideas of probability sampling, including the many types of probability sampling methods and their benefits and drawbacks. This study will help social science researchers choose the best probability sampling approach or techniques to use in order to conduct their research efficiently and effectively.

(Cornesse et al., 2020) [21] There is ongoing debate in the survey research literature on the reliability of estimates of a larger population using "probability and nonprobability sample surveys". As a consequence of survey design, statistical theory supports confidence in probability sampling, but findings from nonprobability sampling are entirely dependent on models for validity. This article reviews the controversy surrounding "probability and nonprobability sample surveys". Theoretically, nonprobability sample surveys may provide reliable results, and we discuss empirical evidence on which sample types offer the most accuracy in reality. We derive best-practice suggestions and suggest directions for future study from these theoretical and empirical concerns.

3 Conclusion

In conclusion, this review underscores the critical role of sampling techniques in shaping the validity and reliability of social science research. Even though statistical rigour and representativeness are provided by probability sampling techniques such "simple random, stratified, cluster, and systematic sampling", they often need substantial resources and population access. In contrast, non-probability methods, including convenience and snowball sampling, provide pragmatic alternatives, especially in exploratory studies or when dealing with platforms like Twitter. However, these methods come with limitations such as the inability to calculate confidence intervals or margins of error. Despite these drawbacks, the rising cost and complexity of data collection have led many researchers to adopt non-probability approaches. Importantly, researchers must carefully align their sampling strategy with study objectives, population characteristics, and resource constraints. By acknowledging the strengths and limitations of each method, social science researchers can make informed methodological choices, enhancing the credibility and interpretability of their findings.

References

[1] P. Cash, O. Isaksson, A. Maier, and J. Summers, “Sampling in design research: Eight key considerations,” Des. Stud., vol. 78, p. 101077, 2022, doi: 10.1016/j.destud.2021.101077.

[2] A. E. Berndt, “Sampling Methods,” J. Hum. Lact, vol. 36, no. 2, pp. 224–226, 2020, doi: 10.1177/0890334420906850.

[3] M. Elfil and A. Negida, “Sampling methods in clinical research; an educational review,” Arch. Acad. Emerg. Med., vol. 7, no. 1, pp. 3–5, 2017.

[4] M. R. W. Hiebl, “Sample Selection in Systematic Literature Reviews of Management Research,” Organ. Res. Methods, vol. 26, no. 2, pp. 229–261, 2023, doi: 10.1177/1094428120986851.

[5] S. Noor and J. Golzar, “Simple random sampling,” Handb. Stat., vol. 6, no. November, pp. 97–109, 2022, doi: 10.1016/S0169-7161(88)06006-7.

[6] M. Memon, H. Ting, J.-H. Cheah, R. Thurasamy, F. Chuah, and T. Huei Cham, “SAMPLE SIZE FOR SURVEY RESEARCH: REVIEW AND RECOMMENDATIONS,” J. Appl. Struct. Equ. Model, vol. 4, no. 2, pp. 2590–4221, 2020.

[7] S. K. Mishra and M. Lavater, “Sampling Methods in Social Science Research,” no. July, 2023, [Online]. Available: https://www.researchgate.net/publication/371987239

[8] N. Saxena, “A systematic Survey of Natural Language Processing (NLP) Models & its Application,” Int. J. Innov. Sci. Eng. Manag., pp. 25–31, 2022.

[9] S. J. Stratton, “Population Sampling: Probability and Non-Probability Techniques,” Prehosp. Disaster Med., vol. 38, no. 2, pp. 147–148, 2023, doi: 10.1017/S1049023X23000304.

[10] S. Dehariya and D. Sharma, “Study On Highway Roadside Safety Prabhat Chauraha to Bajrang Chauraha Raisen Road Bhopal,” Int. J. Innov. Sci. Eng. Manag., pp. 44–50, 2024.

[11] Hamed Taherdoost, "Sampling Methods in Research Methodology; How to Choose a Sampling Technique for Research," Int. J. Acad. Res. Manag., vol. 5, no. 2, pp. 18–27, 2016.

[12] M. Haque, "Sampling Methods in Social Research," Glob. Res. Methodol., pp. 1–6, [Online]. Available: http://www.pansoi.com/grmgrlaranya.org/Journals/SAMPLING METHODS IN SOCIAL RESEARCH.pdf

[13] S. K. Ahmed, "How to choose a sampling technique and determine sample size for research: A simplified guide for researchers," Oral Oncol. Reports, vol. 12, no. September, p. 100662, 2024, doi: 10.1016/j.oor.2024.100662.

[14] M. D. Ali and E. A. J. Al Hatef, "Types of Sampling and Sample Size Determination in Health and Social Science Research," J. Young Pharm., vol. 16, no. 2, pp. 204–215, 2024, doi: 10.5530/jyp.2024.16.27.

[15] M. A. Adeoye, "Review of sampling techniques for education," ASEAN J. Sci. Educ., vol. 2, no. 2, pp. 87–94, 2023, [Online]. Available: https://ejournal.bumipublikasinusantara.id/index.php/ajsed

[16] D. Makwana, P. Engineer, A. Dabhi, and H. Chudasama, "Sampling methods in research: A review," Int. J. Trend Sci. Res. Dev., vol. 7, no. 3, pp. 762–768, 2023, [Online]. Available: https://www.researchgate.net/publication/371985656

[17] P. Vicente, Sampling Twitter users for social science research: evidence from a systematic review of the literature, vol. 57, no. 6. Springer Netherlands, 2023. doi: 10.1007/s11135-023-01615-w.

[18] M. J. Zickar and M. G. Keith, "Innovations in Sampling: Improving the Appropriateness and Quality of Samples in Organizational Research," Annu. Rev. Organ. Psychol. Organ. Behav., vol. 10, pp. 315–337, 2023, doi: 10.1146/annurev-orgpsych-120920-052946.

[19] S. Raifman, M. A. DeVost, J. C. Digitale, Y.-H. Chen, and M. D. Morris, "Respondent-Driven Sampling: a Sampling Method for Hard-to-Reach Populations and Beyond," Curr. Epidemiol. Reports, vol. 9, no. 1, pp. 38–47, 2022, doi: 10.1007/s40471-022-00287-8.

[20] M. M. Rahman, M. I. Tabash, A. Salamzadeh, S. Abduli, and M. S. Rahaman, "Sampling Techniques (Probability) for Quantitative Social Science Researchers: A Conceptual Guidelines with Examples," SEEU Rev., vol. 17, no. 1, pp. 42–51, 2022, doi: 10.2478/seeur-2022-0023.

[21] C. Cornesse et al., "A review of conceptual approaches and empirical evidence on probability and nonprobability sample survey research," J. Surv. Stat. Methodol., vol. 8, no. 1, pp. 4–36, 2020, doi: 10.1093/jssam/smz041.

[22] M. Medhi, "a Study on the Green Marketing Practices Adopted By Various Companies in India," vol. 6, no. 3, pp. 83–88, 2015, [Online]. Available: http://www.iaeme.com/ijmhrm.asp83http://http://www.iaeme.com/currentissue.asp?JType=IJMHRM&VType=6&IType=3http://www.iaeme.com/currentissue.asp?JType=IJMHRM&VType=6&IType=3

AG Publishing House

Advanced Research Techniques: Theories, Methods, And Practices (Volume-2)

Volume 2
Year: 2025

AG PH Books

Ethical Dilemmas in Research: A Systematic Review

Mahendra Pratap Swain [1*]

[1]*Principal Research Scientist, Leiutis Pharmaceuticals LLP, Hyderabad, Telangana and NDF-SRF-AICTE, Department of Pharmaceutical Sciences and Technology, Birla Institute of Technology, Mesra, Ranchi.*

Abstract

Ethics must be respected at every stage of scientific research, including the dissemination stage. Adverse ethical situations are frequently encountered as a result of the dynamism of qualitative studies. This article provides a comprehensive review of the diverse literature that has investigated ethical dilemmas in the field of research. This review underscores the critical importance of ethical mindfulness in research involving human participants. The integrity and reliability of results depend on adherence to values including "informed consent, privacy, and voluntary" involvement. The evolving research landscape—marked by digitalization, systemic injustices, and shifting cultural norms—presents new ethical challenges, particularly in qualitative and medical research. Addressing these dilemmas requires a shift toward quality over quantity, reflexivity, and continuous ethical engagement. Fostering awareness, peer collaboration, and education in research ethics is vital for safeguarding participants' dignity and advancing responsible scientific inquiry.

Keywords: Ethical dilemmas, Scientific research, Ethical challenges, Qualitative research and studies, Research ethics committees (RECs), Anonymity, Privacy, Confidentiality, etc.

1 Introduction

Participants must be completely and openly informed about the hazards, advantages, and repercussions of the research; their agreement must be provided willingly and without compulsion or duplication; and the justification for any scientific investigation involving human beings must be relevant and unambiguous. Confidentiality must be preserved and damage must be repaired on a regular basis. The

* ISBN No. - 978-93-49028-02-9

ethical principles that motivate scientific inquiry are predicated on the guarantee of human freedom and dignity [1]. The organisations that regulate research involving human beings, known as "Research Ethics Committees (RECs)", adopt ethical standards and guidelines that express these principles. Participants are protected by these guidelines, which also ensure that the search is conducted ethically throughout. However, during qualitative research, a variety of unforeseen circumstances may occur that call for judgements that do not fall within the purview of fundamental ethical norms [2]. Due to the dynamic nature of qualitative research and the possibility of unforeseen occurrences, it is essential that the researcher anticipate potential obstacles and take steps to avoid them. Because of this, ethical standards and codes are not always enough to address issues that arise throughout the course of research [3]. Situational solutions are necessary for many emergency situations that develop in a particular environment, always maintaining the requirements of the participants as the primary consideration [4], [5].

Because the researcher is a subject placed inside the culture being studied, doing qualitative research is a comprehensive and intricate process that presents ethical dilemmas as the researcher does the research and experiences its effects. Studies involving human subjects rather than human subjects are influenced by the setting, which sometimes necessitates that the researcher modify ethical standards and guidelines to fit the circumstances, which may occasionally result in a conundrum [6]. As many researchers tend to think, ethics in study is thus more than just following rules, guidelines, or an ethics committee's permission. Its scientific excellence is inextricably linked to the researcher's duty for maintaining the research's ethical integrity. A lack of ethics renders data of no quality, and the reverse is also true [7]. Ethical considerations are present throughout the entire process, from the selection of the object to the formulation of the theoretical foundations, targets, and approaches. Structure, in addition to "the interpretation and distribution" of the outcomes to the scientific community, managers, participants in the study, or society at large. Although there is no direct danger to participants in a qualitative research, there is a chance that it will cause damage to a person's physical, psychological, moral, intellectual, social, and cultural aspects at any point during the study or after the fact [8].

A. Ethical dilemma

An ethical dilemma, also known as a moral dilemma or ethical paradox, is a situation in philosophy when an actor is faced with two or more moral imperatives that clash with one another but do not supersede one another. According to a nearly similar concept, an ethical dilemma occurs when all of the options are incorrect [9]. The term is often used more widely in everyday speech to refer to moral conundrums that may be solved, morally difficult choices, and other difficult moral problems. In a strict philosophical sense, this article addresses what are frequently referred to as genuine ethical challenges or ethical dilemmas [10]. Although several examples have been published, there is disagreement about whether they reflect actual or hypothetical ethical conundrums. At the core of the debate around them is the question of whether ethical difficulties really exist. Defenders usually point to clear-cut examples, while their opponents usually attempt to show how their presence violates fundamental ethical principles [11]. There are many different kinds of ethical quandaries. The difference between epistemic and true or

ontological difficulties is important since the latter might give the actor of an unresolvable conflict a false impression. Even if the majority of people think that there are epistemic problems, ethical conundrums are mostly ontological in nature. Philosophers have always maintained that ethical difficulties must not exist in sound moral systems. Modern philosophy has challenged this assumption, however [12].

B. Type of ethical dilemmas

A variety of ethical difficulties exist. When there are arguments about the existence of ethical difficulties, the differences between these categories are often crucial. While certain types may not be addressed by specific arguments against the fact that they exist, others could. And only a few kinds—if any—might qualify as true ethical conundrums.

Epistemic vs ontological: The agent is unsure of what to do in epistemic ethical issues because they are unable to decide which moral requirement comes first. This kind of uncertainty is present in many daily decisions, ranging from the insignificant selection between two different-packaged cans of beans at the grocery store to major career decisions. However, unresolvable epistemic conflicts may arise independently of unresolvable conflicts, and vice versa.

Self-imposed vs world-imposed: The origin of the competing demands distinguishes self-inflicted ethical issues from those imposed by the outside world. The agent is in charge of the conflict in the self-imposed scenario. Making two incompatible commitments, such as attending two events that are taking place at different locations at the same time, is a typical example in this category. The agent is thrust into the predicament in the world-imposed situation, however, and is not held accountable for its occurrence. For moral theories, the distinction between these two categories is significant [13]. In the past, most philosophers have held that ethical theories should be free of ethical challenges and that moral theories that allow or demand the existence of ethical dilemmas are flawed.

Obligation vs prohibition: An obligation is a moral need to behave in a certain way, while a prohibition is a moral responsibility to abstain from behaving in a specific way. The majority of ethical dilemma debates concentrate on obligation dilemmas, which are situations in which an actor must choose between two morally contradictory acts. Prohibition problems, however, are circumstances in which no action is permitted. Many arguments against ethical problems have been said to be effective only when they apply to obligation dilemmas; they do not work well when they apply to prohibition dilemmas.

Single-agent vs multi-agent: Ethical dilemmas are characterised by the existence of two courses of action that are both mandatory but mutually exclusive: it is impossible to execute both actions simultaneously. Typically, a single agent is faced with conflicting obligations in single-agent cases. These actions remain incompatible in multi-agent scenarios; however, the obligations pertain to distinct individuals. For instance, two competitors who had pledged to their families that they would prevail in a competition may both be obligated to achieve victory [14]. The conflicting obligations of these two individuals are due to the fact that there can only be one victor.

2 Literature Review

(Sharma et al., 2024) [15] The objective is to perform a comprehensive bibliometric analysis of research papers that address the ethical issues and concerns related to "digital mental health care". According to the study, several of the most important issues that need proper management, such as data protection, emergency response, therapist competence, and permission, are clearly addressed in the most often referenced publications on digital mental healthcare. Failure to do so may result in the client experiencing distress and questioning the reliability of “the Digital Mental Health Care system”. According to this bibliometric study, there are a number of issues that may be major standards for mental health care provided online. Mental health personnel who operate online must also be informed about the ethical issues that are associated with online mental healthcare.

(Drolet et al., 2023) [16] This study employed a descriptive phenomenological approach to document the ethical dilemmas encountered by a diverse array of "Canadian researchers, REB members, and research ethics specialists". Data was collected through sociodemographic questionnaires and individual semi-structured interviews. Following the triangulation of different perspectives (researchers, REB members, and ethics experts), the following ten units of meaning were created: "ethical distress, distributive injustices, social injustices, research integrity, conflicts of interest, respect for research participants, individualism and performance, lack of supervision and power imbalances, and ethical distress". The results of this investigation revealed a number of problematic components that could potentially facilitate the development of future solutions to the transversal ethical challenges in research that affect the academic community's diverse members.

(Lathan et al., 2023) [17] Investigate the extent to which research ethical practices were implemented and the research methodologies employed in the analysis of Facebook data in public health research. Found six categories of analytical methods to use this data: predictive model development, utility (i.e., whether Facebook is useful as a tool for public health dissemination, surveillance, or attitudes), associational studies of user behaviour and health outcomes, network analysis, and two types of content analysis (sentiment analysis and thematic analysis). Utility studies and prediction studies had the lowest likelihood of requesting IRB review (0/4, 0% and 1/4, 25%, respectively), while associational studies had the highest likelihood (5/6, 83%). It is necessary to provide more comprehensive guidance on research ethics for the use of Facebook data, particularly in relation to the use of personal identifiers.

(Nii Laryeafio & Ogbewe, 2023) [18] Explore the ideas that support the ethics of qualitative research and look at all the ethical issues that the researcher should be aware of while gathering primary data via interviews. The study's conclusions also show that virtue, utilitarianism, rights, and deontology are the main ideologies that underpin ethical concerns in research. Participant rights must be upheld in qualitative research to support the collection of trustworthy data and achieve study objectives. The researcher is guided by these and other ethical standards, including anonymity, privacy, secrecy, voluntary involvement, and the ability to opt out, in order to gather qualitative data via interviews in a way that produces reliable findings.

(Lee, 2022) [19] Clinical research involving human subjects must be truth-based, demonstrate scientific integrity, and follow ethical norms and principles in order to protect study participants. Transparency and strict authorship criteria should be maintained in the publication of clinical research. The extent of investigator malfeasance in clinical research and publication can be diminished by a comprehensive comprehension and knowledge of ethical issues. On the basis of the most recent guidelines, this article reviews and summarises fundamental ethical concerns in clinical research and publication.

(Taquette & Borges da Matta Souza, 2022) [4] There is an obligation to adhere to ethical principles throughout the entire scientific research process, which encompasses the dissemination phase. This integrative review's goal was to synthesise and examine ethical conundrums that come up when looking into qualitative data and approaches that have been proposed to deal with them. Potential harms, misunderstandings about the roles of the researcher, therapist, and friend, confidentiality violations, and deadlocks in the study ethics committees were the main disputes. Numerous conflicts may arise during research. Continuous consent, ethical mindfulness, reflexivity, and self-awareness are the foundations of the proposed solutions.

(Piasecki et al., 2021) [20] Implemented the constant comparative method to identify common ethical motifs. The study exposes the scope, profundity, and complexity of ethical issues that are linked to the use of EHRs in research. Most of the ethical problems with EHR-based research derive from rapid cultural change. Individual and public dimensions of beneficence, as well as the formulation of concepts of privacy, are evolving. At the present time, we are in the midst of this transitional phase. Laws, brain processes, and human emotions are not operating up to speed with technology improvements. The individual patient's health has always been the primary concern in the medical tradition. Our comprehension of research ethics, public health ethics, and health care ethics seems to be impacted by the digitisation and change of healthcare.

(Hosseini & Gordijn, 2020) [21] The findings of a literature study on the moral dilemmas surrounding scientific authorship are presented in the paper at hand. When it comes to reporting, authorship, and publishing research findings, these challenges are seen as queries and/or worries over duties, principles, or virtues. Ten ethical themes—some of which include several ethical issues—have been discovered as a result of the papers' analysis. The following topics are listed in order of frequency of occurrence: 1) attribution; 2) breaches of authorship norms; 3) bias; 4) responsibility and accountability; 5) authorship order; 6) citations and referencing; 7) authorship definition; 8) publishing strategy; 9) originality; and 10) punishments. The present paper maps these themes, examines significant ethical issues, and offers a critical assessment of the application of standards of conduct, different cultural interpretations, and contributing reasons to unethical behaviour.

(Colnerud, 2015) [14] The objective is to furnish empirically derived information regarding the ethical scrutiny laws and procedure. This information will pertain to the experiences of Swedish researchers with ethical issues, conflicts, and challenges in their research. Three phenomena are revealed by the analysis of the researchers' responses, which is partially consistent with similar studies conducted in other nations: (i) the law of ethical vetting restricts the pertinent research ethical questions; (ii) it is

impossible to foresee every research ethical question; and (iii) there are repercussions for the ethical vetting boards' disregard for issues that do not fall under the law.

3 Conclusion

In conclusion, this systematic review underscores the centrality of ethical considerations in research involving human participants, particularly within qualitative and medical studies. Researchers must uphold fundamental principles such as informed consent, anonymity, voluntary participation, privacy, and confidentiality to ensure both the validity of findings and the dignity of participants. The study reveals systemic ethical dilemmas arising from institutional pressures like "publish or perish" and highlights the need to shift toward a culture of quality over quantity in research output. It also stresses the importance of addressing systemic injustices that disproportionately affect underrepresented groups in academia. Moreover, the digitalization of healthcare and the rise of remote data collection through social media have introduced new ethical complexities, necessitating ongoing reflection and adaptation. To navigate these evolving challenges, researchers must practice ethical mindfulness, self-awareness, and continuous consent, while fostering open dialogue with peers. Ethical education and institutional support are essential for maintaining scientific integrity, protecting participant well-being, and fostering an inclusive and respectful research environment. Ultimately, a commitment to ethics is not only a regulatory requirement but a cornerstone of responsible and impactful scientific inquiry.

References

[1] M. I. T. D. Correia, “Ethics in research,” Clin. Nutr. Open Sci., vol. 47, pp. 121–130, 2023, doi: 10.1016/j.nutos.2022.12.010.

[2] M. Kerres and S. Bedenlier, Systematic Reviews in Educational Research. 2020. doi: 10.1007/978-3-658-27602-7.

[3] F. Li, N. Ruijs, and Y. Lu, “Ethics & AI: A Systematic Review on Ethical Concerns and Related Strategies for Designing with AI in Healthcare,” AI, vol. 4, no. 1, pp. 28–53, 2023, doi: 10.3390/ai4010003.

[4] S. R. Taquette and L. M. Borges da Matta Souza, “Ethical Dilemmas in Qualitative Research: A Critical Literature Review,” Int. J. Qual. Methods, vol. 21, pp. 1–15, 2022, doi: 10.1177/16094069221078731.

[5] S. A. Z. Zaidi, E. Ahmad, and N. Shukla, “Ethical Considerations in the Use of Artificial Intelligence (AI) for Education and Research : A Review,” Int. J. Innov. Sci. Eng. Manag., pp. 156–167, 2024, doi: 10.69968/ijisem.2024v3si2156-167.

[6] C. Calia et al., “Developing an Evidence-base to Guide Ethical Action in Global Challenges Research in Complex and Fragile Contexts: A Scoping Review of the Literature,” Ethics Soc. Welf., vol. 16, no. 1, pp. 54–72, 2022, doi: 10.1080/17496535.2021.1916830.

[7] J. I. Steinert, D. Atika Nyarige, M. Jacobi, J. Kuhnt, and L. Kaplan, “A systematic review on ethical challenges of a € field’ research in low-income and middle-income countries: Respect,

justice and beneficence for research staff?" BMJ Glob. Heal, vol. 6, no. 7, pp. 1–20, 2021, doi: 10.1136/bmjgh-2021-005380.

[8] D. Ginting, "Ethical Research Dilemmas and Their Implications In English Language Teaching Studies," Acitya J. Teach. Educ., vol. 4, no. 1, pp. 110–123, 2022, doi: 10.30650/ajte.v4i1.3200.

[9] C. Scheytt and J. Pflüger, "Ethical challenges in qualitative sociology: a systematic literature review," Front. Sociol., vol. 9, no. September, pp. 1–17, 2024, doi: 10.3389/fsoc.2024.1458423.

[10] P. Saumya and V. Shyam, "The menace of dermatophytosis in India: The evidence that we need," Indian J. Dermatol. Venereol. Leprol., vol. 84, no. 1, pp. 6–15, 2018, doi: 10.4103/ijdvl.IJDVL.

[11] G. Schofield, M. Dittborn, L. E. Selman, and R. Huxtable, "Defining ethical challenge(s) in healthcare research: a rapid review," BMC Med. Ethics, vol. 22, no. 1, pp. 1–17, 2021, doi: 10.1186/s12910-021-00700-9.

[12] H. Nobile, N. L. R. Moshtaghin, Z. Lüddecke, A. Schnarr, and M. Mertz, "What can the citations of systematic reviews of ethical literature tell us about their use?—an explorative empirical analysis of 31 reviews," Syst. Rev., vol. 12, no. 1, pp. 1–19, 2023, doi: 10.1186/s13643-023-02341-y.

[13] L. L. Dhirani, N. Mukhtiar, B. S. Chowdhry, and T. Newe, "Ethical Dilemmas and Privacy Issues in Emerging Technologies: A Review," Sensors, vol. 23, no. 3, 2023, doi: 10.3390/s23031151.

[14] G. Colnerud, "Ethical dilemmas in research in relation to ethical review: An empirical study," Res. Ethics, vol. 10, no. 4, pp. 238–253, 2015, doi: 10.1177/1747016114552339.

[15] P. Sharma, R. Wagani, and M. A. Varghese, "Bibliometric analysis of research in ethical concerns and dilemmas of digital mental health care in the last two decades," Front. Hum. Dyn., vol. 6, no. January, 2024, doi: 10.3389/fhumd.2024.1502432.

[16] M. J. Drolet, E. Rose-Derouin, J. C. Leblanc, M. Ruest, and B. Williams-Jones, "Ethical Issues in Research: Perceptions of Researchers, Research Ethics Board Members and Research Ethics Experts," J. Acad. Ethics, vol. 21, no. 2, pp. 269–292, 2023, doi: 10.1007/s10805-022-09455-3.

[17] H. S. Lathan et al., "Ethical considerations and methodological uses of Facebook data in public health research: A systematic review," Soc. Sci. Med., vol. 322, no. October 2022, p. 115807, 2023, doi: 10.1016/j.socscimed.2023.115807.

[18] M. Nii Laryeafio and O. C. Ogbewe, "Ethical consideration dilemma: systematic review of ethics in qualitative data collection through interviews," J. Ethics Entrep. Technol., vol. 3, no. 2, pp. 94–110, 2023, doi: 10.1108/jeet-09-2022-0014.

[19] H. S. Lee, "Ethical issues in clinical research and publication," Kosin Med. J., vol. 37, no. 4, pp. 278–282, 2022, doi: 10.7180/kmj.22.132.

[20] J. Piasecki, E. Walkiewicz-Żarek, J. Figas-Skrzypulec, A. Kordecka, and V. Dranseika, "Ethical issues in biomedical research using electronic health records: a systematic review," Med. Heal. Care Philos., vol. 24, no. 4, pp. 633–658, 2021, doi: 10.1007/s11019-021-10031-6.

[21] M. Hosseini and B. Gordijn, "A review of the literature on ethical issues related to scientific authorship," Account. Res., vol. 27, no. 5, pp. 284–324, 2020, doi: 10.1080/08989621.2020.1750957.

AG Publishing House

Advanced Research Techniques: Theories, Methods, And Practices (Volume-2)

Volume 2
Year: 2025

AG PH Books

A Review of Statistical Method for Hypothesis Testing

Dr. Sunita Sharma [1*]

[1]*Faculty in Delhi public world school.*

Abstract

In the context of research initiatives, hypothesis testing is an essential component of statistical analyses. It entails estimating the probability of an event occurring by coincidence. In this article review the various literature's study on statistical method for hypothesis testing. In conclusion, this review highlights the enduring popularity of Null Hypothesis Significance Testing (NHST), despite growing awareness of its limitations. While disciplines like psychology and statistics show stable use, NHST remains widely adopted. Nonparametric methods offer flexible alternatives when assumptions of parametric tests are unmet. The importance of multiple testing correction, careful model selection, and transparent reporting is emphasized to ensure validity and reproducibility. Researchers should move beyond sole reliance on p-values, incorporating confidence intervals and contextual relevance. Advancing statistical literacy and methodological rigor is essential for enhancing the quality and impact of hypothesis-driven research across diverse fields.

Keywords: Null Hypothesis Significance Testing (NHST), Statistical Method, Hypothesis Testing, Alternative Hypothesis, ANOVA, Logit/Probit models, Decision Making, etc.

1 Introduction

A statistical sample is assessed by an analyst during hypothesis testing to establish the plausibility of the null hypothesis. A statistical analyst measures and examines a random sample of the population under investigation. Every analyst employs a random population sample to test two distinct hypotheses: the null hypothesis and the alternative hypothesis [1]. Generally speaking, the null hypothesis states that

* ISBN No. - 978-93-49028-02-9

every population parameter is the same. It may claim, for instance, that the population mean return is zero. In essence, the alternative hypothesis is the opposite of the null hypothesis. Therefore, only one of them can be true, and they are mutually exclusive [2]. One of the two theories, nevertheless, will always be correct. The satirical writer John Arbuthnot, who examined male and female births in England in 1710 after noting that male births outnumbered female births by a little margin in almost every year, is credited by some statisticians with creating the first hypothesis tests. Arbuthnot determined that because there was little prospect of this occurring by accident, divine providence was responsible [3], [4].

A. Hypothesis Testing

One statistical technique for determining whether or not experiment findings are significant is hypothesis testing. It comprises the establishment of a null hypothesis as well as an alternative hypothesis. These two theories will never be compatible. This implies that the alternative hypothesis is false if the null hypothesis is true, and vice versa. A typical instance of hypothesis testing is putting up a test to see if a new drug cures a disease more successfully [5], [6].

Null Hypothesis: A concise mathematical statement known as the null hypothesis is employed to suggest that there is no distinction between two possibilities. Other words, there is no distinction between specific data characteristics. An experiment's results are presumed to be solely determined by coincidence in this hypothesis. It is designated as H0. The null hypothesis is evaluated through hypothesis testing to determine whether it can be refuted. Let us assume that an experiment is conducted to determine whether females are shorter than boys at the age of five. The null hypothesis will be that they are of the same height [7].

Alternative Hypothesis: In contrast to the null hypothesis, the alternative hypothesis is a different proposal. It serves as an indicator that the results of an investigation are attributable to a genuine effect. It signifies the existence of a statistically significant difference between two potential outcomes and is represented as H1 or Ha. The alternative hypothesis for the aforementioned example may be that females are shorter than boys at the age of five.

B. Statistical method

To turn raw data into useful insights in a variety of disciplines, statistical techniques are essential. Using these mathematical tools, researchers, analysts, and decision-makers may efficiently gather, arrange, analyse, interpret, and present data. These techniques make it easier to comprehend large, complicated data sets, find trends, and make wise choices in the social sciences, business, healthcare, and engineering domains. Statistical approaches provide a methodical approach to data analysis, ranging from using descriptive statistics to summarise data to using inferential techniques to test hypotheses and make predictions [8].

C. Descriptive and Inferential Statistics

The two primary subfields of statistics are known as descriptive statistics, which describes the properties of sample and population data, and inferential statistics, which uses these properties to test hypotheses

and draw conclusions. Variance, skewness, kurtosis, and mean (average) are examples of descriptive statistics. Analysis of variance (ANOVA), logit/Probit models, linear regression analysis, and null hypothesis testing are examples of inferential statistics.

1. Descriptive Statistics

In general, descriptive statistics focus on sample data's central tendency, variability, and distribution. The term central tendency refers to the evaluation of a characteristic or typical feature of a population or sample. Descriptive statistics like the mean, median, and mode are included. A group of statistics known as variability shows how much a population or sample's constituent parts vary from one another along the qualities under evaluation. Included are metrics like range, variance, and standard deviation [9]. The term "distribution" refers to the general "shape" of the data, which may be shown on a chart like a dot plot or histogram and include characteristics like kurtosis, skewness, and the probability distribution function. Additionally, disparities between observed attributes of a data set's components may be described using descriptive statistics. In addition to aiding in the understanding of the collective qualities of the components of a data sample, they may be employed to test hypotheses and make predictions using inferential statistics [10].

2. Inferential Statistics

Through the application of inferential statistics, statisticians may make inferences about the features of a population based on observations of a sample. Additionally, it is utilised to assess the degree of certainty they can have on the validity of those judgements. Based on the sample size and distribution, statisticians can calculate the likelihood that "statistics, which quantify the central tendency, variability, distribution, and relationships" among characteristics within a data sample, will accurately represent the corresponding parameters of the entire population from which the sample is drawn [11]. Generalisations about large groups are made using inferential statistics, such as trying to forecast future occurrences or predicting typical demand for a product by surveying the purchasing patterns of a sample of users. A securities or asset class's future return may be projected using returns from a sample period [12].

Regression analysis is a frequently employed mathematical method for statistical inference. The correlation between a dependent variable and one or more explanatory (independent) variables is assessed by examining the strength and nature of the relationship. A regression model's output is frequently subjected to statistical significance analysis, which implies that the results of testing or experimentation are unlikely to have occurred by coincidence or at random. Or, in other words, statistical significance implies that the results are due to a particular cause that is revealed by the data [13].

D. Importance of Hypothesis Testing in Research

Several critical purposes are served by hypothesis testing, which is a critical component of the scientific research process:

Guiding Research: Through the formulation of precise predictions regarding the relationship between variables, hypothesis testing offers a distinct direction for research. In this way, researchers can concentrate their resources and efforts on the investigation of particular hypotheses.

Testing Predictions: By doing so, researchers can evaluate the accuracy of their predictions or assumptions regarding the parameters of the population. Through the comparison of sample data to the hypothesised population parameters, researchers can ascertain whether their predictions are corroborated by the evidence.

Decision Making: Hypothesis testing offers a methodical framework for data-driven decision-making. The acceptance or rejection of a null hypothesis can be determined by researchers through the use of statistical tests, which in turn informs conclusions regarding the research query.

Generalizability: The generalisability of researchers' findings to the broader population can be evaluated through hypothesis testing. Based on the sample data, researchers can draw conclusions about the population by investigating hypotheses.

E. Application of hypothesis testing

Medicine and Healthcare: In clinical trials and medical research, hypothesis testing is essential. To ascertain if novel medications or therapies are successful, researchers use hypothesis testing. Data is gathered from a sample population, and they develop a null hypothesis (the medicine has no impact) and an alternative hypothesis (the drug has a substantial effect). They may determine if the evidence supports rejecting or failing to reject the null hypothesis by examining the data and doing statistical tests.

Quality Control and Manufacturing: To guarantee product quality and process efficiency in manufacturing sectors, hypothesis testing is used. For instance, a vehicle manufacturer may test the hypothesis that a new model's mean fuel economy is higher than that of the old model. Through hypothesis testing and the collection of fuel efficiency data from a sample of automobiles, they may ascertain if the claim is sufficiently supported.

Economics and Finance: In order to assess hypotheses and get well-informed conclusions, economic and financial research use hypothesis testing. An economist may, for example, examine a theory on how a certain policy affects economic growth. They may evaluate the importance of the connection between the policy and economic results by looking at relevant data and doing statistical tests.

Social Sciences: Political science, sociology, psychology, and other social sciences all make extensive use of hypothesis testing. In order to test their theories, researchers in these domains gather data on social events, human behaviour, and the impacts of policies. After then, the data is analysed using statistical tests to see whether the evidence backs up the hypotheses put forward.

Market Research: Market research uses hypothesis testing to verify presumptions and create data-driven choices. Businesses may test theories on customer preferences, market trends, or the efficacy of marketing initiatives via surveys or experiments. They may make wise business judgements and learn more about customer behaviour by using hypothesis testing to the data analysis.

2 Literature Review

(Henríquez-roldán et al., 2025) [14] In order to elucidate the importance of the scientific method in statistics, this research is the first to propose an experiential learning technique in statistical education. It does this by using an actual experimental design. Several critical phases comprise the experiment, including the design of the study, the formulation and testing of hypotheses, the participation of experimental units, the analysis of the collected data, and the interpretation of the results to derive statistical inferences. Along with making statistical ideas easier to understand practically, this approach also highlights the importance of "experimental design, hypothesis testing, data analysis, and probability interpretation in real-world situations". By participating in this immersive experience, students are anticipated to develop a more profound understanding of statistics as a scientific discipline and its practical applications in daily life.

(Emmert-Streib, 2024) [15] Examine the community's response to these discussions. To be more precise, we perform a scientometric analysis of bibliographic records to examine the publication behaviour regarding the utilisation of Null hypothesis significance testing (NHST). A trend analysis is performed for the general community, specific subject areas, and individual journals. Additionally, we perform a change-point analysis to determine whether there are genuine changes or repeated movements. Therefore, we have discovered that NHST is more popular than ever among the general public. However, there is a distinct heterogeneity in the publication behaviour of specific subject areas and journals, and no uniformity is apparent.

(Jangir et al., 2024) [16] A fundamental procedure in statistical analyses for research initiatives is hypothesis testing. It entails assessment of the probability of an event transpiring by coincidence. The subjects are divided into two categories in experimental studies: a control group and a treatment group. The results of these groups are contrasted to determine whether any observed differences are indicative of the treatment effect or are merely the result of random sampling error. Conventional testing is conducted to evaluate the null hypothesis, which asserts that there is no distinction between the categories. In a study that investigates the impact of treatment on mortality rates, the null hypothesis would assert that the treatment has no effect. This methodology is the foundation of the majority of statistical analyses conducted in individual research initiatives.

(Rogers, 2022) [17] Aims to establish a paradigm for comprehending the issue by examining the history of NHST and exploring alternative methodologies. In order to contextualise the issues, a concise overview of the history, dissemination, and criticisms of NHST is presented. The spectrum of available options to NHST is illustrated through the description of supplementary and alternative methodologies. Simulated data sets are employed to illustrate a variety of viable alternative methodologies and supplements to NHST. The Monte Carlo method simulations are employed to compare the methods that provide decision rules with NHST.

(Suvorov et al., 2022) [18] One of the vital procedures in contemporary medical research is statistical hypothesis testing. Scientific researchers initially develop a research hypothesis, which is subsequently

used to develop and evaluate the statistical hypothesis. This review delineates the general algorithm for evaluating the null and alternative hypotheses using a t-test, as well as the examples that are compiled for various research questions. In addition, the authors discuss type I errors, which are essential for interpreting p-values estimated from statistical tests, and type II errors, which are employed to evaluate the power of a study. The article concentrates on the distinction between statistically significant and clinically significant effects, as well as the calculation methods and measure of effect size. The connection between type II error, effect magnitude, and sample size is also examined.

(Van Witteloostuijn & van Hugten, 2022) [19] Over the past decade, there has been an additional increase in this critique, which is linked to recommendations for conducting quantitative empirical research in an alternative manner. In the present research note, we conduct a comprehensive examination of 148 articles that present the results of quantitative empirical studies published in recent issues of 18 prestigious journals in six prominent social science disciplines: business, economics, political science, psychology, public administration, and sociology. Although numerous economics studies emphasise "estimation precision" rather than "statistical significance," we conclude that NHST remains excessively prevalent in all six disciplines. In addition, we identify a few instances of innovative practices that may serve as a catalyst for change.

(Mata & Milner, 2021) [20] With a focus on their relevance for current pathology research and treatment, the statistical tests of the publications were summarised. Ninety-three percent of the 195 identified publications mentioned using one or more statistical tests. Several significant discoveries were identified through a retrospective statistical analysis of the articles. Initial reports of tests for normality were rare, and "parametric hypothesis tests" were overused. Second, research that used multisample hypothesis tests (such as analysis of variance) to look at differences between study groups seldom used post hoc testing. Third, the survival analysis, regression, and correlation methods were not used to their full potential. A primer on pertinent statistical concepts and tests is presented in light of these findings. "Regression, correlation, survival, and genetic data analysis" are among the subjects covered in this primer, along with descriptive and comparative statistics and the best research design.

(Dyckman & Zeff, 2019) [21] Suggests enhancements to the method and quality of this type of research. We address the following limitations in current research that we believe are being disregarded or not utilised appropriately: (1) situational effects that remain unaddressed due to model limitations and what has been previously termed "data carpentry," (2) the information loss that results from undervaluing what can and cannot be learnt from replications, (3) the necessary enhancements to relying on a study's computed "p-values" instead of the economic or behavioural significance of the results, and (4) the limitations and alternatives to winsorizing.

(Emmert-Streib & Dehmer, 2019) [22] From the perspective of data science, statistical hypothesis testing is one of the most misconstrued quantitative analysis methodologies. The procedural components of the system are intricately interconnected, despite their apparent simplistic nature. In this paper, we explore the formal meaning of the components and their connections, as well as the underlying logic of statistical

hypothesis testing. As a generic backbone, our presentation is applicable to all statistical hypothesis tests, rendering it useful in all application domains of artificial intelligence and data science.

(Kruschke & Liddell, 2018) [23] There is a conceptual distinction between hypothesis testing and estimation with quantified uncertainty in the context of data analysis. The New Statistics is a trend among frequentists in psychology that emphasises estimation over hypothesis testing. Bayesian methods and frequentist methods are also conceptually distinct. Our primary objective in this article is to elucidate the superiority of Bayesian methods over frequentist methods in respect to the objectives of the New Statistics. The paper examines the frequentist and Bayesian methodologies for hypothesis testing and estimation with credible or confidence intervals. Bayesian methodologies for meta-analysis, randomised controlled trials, and power analysis are additionally elaborated upon in the article.

3 Conclusion

This review highlights the enduring prominence of Null Hypothesis Significance Testing (NHST) within the research community, despite growing awareness of its limitations. While its popularity continues to rise overall, subject-specific trends—particularly in psychology, biology, and statistics—reveal a more stable or heterogeneous usage. Nonparametric methods offer flexible alternatives to traditional parametric tests, especially when assumptions of normality and homoscedasticity are violated. The widespread use of hypothesis testing in diverse domains such as medicine, economics, and data science underscores the need for careful model selection, transparency, and reproducibility. As multiple testing becomes increasingly common in fields like genomics and finance, appropriate correction techniques must be applied to mitigate errors. Researchers should move beyond an overreliance on p-values and incorporate confidence intervals, effect sizes, and the practical implications of findings to enhance interpretability and impact. Ethical and methodological rigor, combined with improved statistical literacy and communication, will be essential for advancing hypothesis testing methodologies and fostering meaningful scientific progress. The path forward involves both educating the community and adopting more nuanced, context-sensitive approaches to data analysis.

References

[1] D. P. Turner, H. Deng, and T. T. Houle, “Statistical Hypothesis Testing: Overview and Application,” Headache, vol. 60, no. 2, pp. 302–308, 2019, doi: 10.1111/head.13706.

[2] P. J. Veazie, “Understanding statistical testing,” SAGE Open, vol. 5, no. 1, 2015, doi: 10.1177/2158244014567685.

[3] S. Pednekar, R. D. Kulkarni, and P. Mahanwar, “STATISTICAL METHODS AND TESTING OF HYPOTHESIS,” 2022.

[4] Beeranna and N. G. Basappa, “The Role of Education In Shaping Awareness and Attitudes Towards Politics In Special Reference To Yadagiri District, Karnataka,” Int. J. Innov. Sci. Eng. Manag., pp. 55–62, 2024, doi: 10.69968//ijisem.2024v3i455-62.

[5] R. McNulty, "A logical analysis of null hypothesis significance testing using popular terminology," BMC Med. Res. Methodol., vol. 22, no. 1, pp. 1–9, 2022, doi: 10.1186/s12874-022-01696-5.

[6] N. Ben and S. K. Yadav, "Women Empowerment - A Study of SHGS in Satna & Maihar District of Madhya Pradesh," Int. J. Innov. Sci. Eng. Manag., pp. 324–332, 2025, doi: 10.69968/ijisem.2025v4i1324-332.

[7] D. Szucs and J. P. A. Ioannidis, "When null hypothesis significance testing is unsuitable for research: A reassessment," Front. Hum. Neurosci., vol. 11, no. August, 2017, doi: 10.3389/fnhum.2017.00390.

[8] A. J. Harrison et al., "Recommendations for statistical analysis involving null hypothesis significance testing," Sport. Biomech., vol. 19, no. 5, pp. 561–568, 2020, doi: 10.1080/14763141.2020.1782555.

[9] S. L. Pugh and A. Molinaro, "The nuts and bolts of hypothesis testing," Neuro-Oncology Pract., vol. 3, no. 3, pp. 139–144, 2015, doi: 10.1093/nop/npv052.

[10] P. Ranganathan and C. S. Pramesh, "An introduction to statistics: Understanding hypothesis testing and statistical errors," Indian J. Crit. Care Med., vol. 23, pp. S230–S231, 2019, doi: 10.5005/jp-journals-10071-23259.

[11] R. A. Cooper, "Making Decisions with Data: Understanding Hypothesis Testing & Statistical Significance," Am. Biol. Teach., vol. 81, no. 8, pp. 535–542, 2019, doi: 10.1525/abt.2019.81.8.535.

[12] S. Bahariniya and F. Madadizadeh, "Review of the statistical methods used in original articles published in iranian journal of public health from 2015-2019: A review article," Iran. J. Public Health, vol. 50, no. 8, pp. 1577–1585, 2021, doi: 10.18502/ijph.v50i8.6803.

[13] T. Hamasaki, F. Bretz, L. M. LaVange, P. Müller, G. Pennello, and J. C. Pinheiro, "Editorial: Roles of Hypothesis Testing, p-Values and Decision Making in Biopharmaceutical Research," Stat. Biopharm. Res., vol. 13, no. 1, pp. 1–5, 2021, doi: 10.1080/19466315.2021.1874803.

[14] C. F. Henríquez-roldán, S. I. Bangdiwala, R. I. Barrera-, Á. S. Bustos-rubilar, S. I. Bangdiwala, and R. I. Barrera-, "From hypothesis to conclusion: the essential role of statistics in science," Res. Stat., vol. 3, no. 1, p., 2025, doi: 10.1080/27684520.2025.2450532.

[15] F. Emmert-Streib, "Trends in null hypothesis significance testing: Still going strong," Heliyon, vol. 10, no. 21, p. e40133, 2024, doi: 10.1016/j.heliyon.2024.e40133.

[16] M. Jangir, D. Singh Chahar, R. Lal, and S. Sharma, "an Overview of Statistics: Comprehending Hypothesis Testing and Statistical Imperfections (Errors and Bias)," World J. Pharm. Med. Res. www.wjpmr.com ⎮, vol. 10, no. 4, p. 90, 2024, [Online]. Available: www.wjpmr.com

[17] B. D. Rogers, "Null Hypothesis Statistical Testing: A Survey of the History, Critiques, and Alternative Methodologies," (Doctoral Diss., 2022.

[18] A. Y. Suvorov et al., "Statistical hypothesis testing: general approach in medical research," Sechenov Med. J., vol. 13, no. 1, pp. 4–13, 2022, doi: 10.47093/2218-7332.2022.426.08.

[19] A. van Witteloostuijn and J. van Hugten, “The state of the art of hypothesis testing in the social sciences,” Soc. Sci. Humanit. Open, vol. 6, no. 1, p. 100314, 2022, doi: 10.1016/j.ssaho.2022.100314.

[20] D. A. Mata and D. A. Milner, “Statistical Methods in Experimental Pathology: A Review and Primer,” Am. J. Pathol., vol. 191, no. 5, pp. 784–794, 2021, doi: 10.1016/j.ajpath.2021.02.009.

[21] T. R. Dyckman and S. A. Zeff, “Important issues in statistical testing and recommended improvements in accounting research,” Econometrics, vol. 7, no. 2, 2019, doi: 10.3390/econometrics7020018.

[22] F. Emmert-Streib and M. Dehmer, “Understanding Statistical Hypothesis Testing: The Logic of Statistical Inference,” Mach. Learn. Knowl. Extr., vol. 1, no. 3, pp. 945–962, 2019, doi: 10.3390/make1030054.

[23] J. K. Kruschke and T. M. Liddell, “The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective,” Psychon. Bull. Rev., vol. 25, no. 1, pp. 178–206, 2018, doi: 10.3758/s13423-016-1221-4.

AG Publishing House

Advanced Research Techniques: Theories, Methods, And Practices (Volume-2)

Volume 2
Year: 2025

AG PH Books

The Role of Artificial Intelligence in Research Methodology

Dr. Anjela Swarndipika Kujur [1*]

[1]*Assistant Professor, Sai Nath University, Ranchi, Jharkhand*

Abstract

A tedious and intricate procedure, thematic analysis is a part of qualitative research. The potential for improving and partially automating theme analysis has been raised by the development of "generative artificial intelligence (A.I.)", particularly big language models. Review the many studies on the use of artificial intelligence in research technique in this article. This review emphasizes the transformative role of Artificial Intelligence, particularly tools like Chat GPT, in enhancing research methodology. AI streamlines data analysis, fosters deeper insight generation, and supports personalized workflows, especially in qualitative research. However, its effectiveness hinges on user understanding, ethical awareness, and critical engagement. While AI can supplement human expertise, it should not replace it. A synergistic relationship between AI and human judgment is essential for valid outcomes. Future efforts must focus on developing standardized guidelines and fostering transparency to ensure responsible, meaningful, and contextually appropriate integration of AI in global research practices.

Keywords: Qualitative research, ChatGPT, Artificial intelligence (AI), Large language models (LLM), Data analysis, Academic research, etc.

1 Introduction

The paradigm change in research from conventional to artificial intelligence (AI)-driven research has been brought about by recent technological innovations and advancements in machine learning and data mining. The major technology that will shape the future of human civilisation is artificial intelligence (AI), which has the ability to enhance human skills without causing undue expense [1]. In contrast to the rigorous, time-consuming, arduous, and highly technical nature of conventional qualitative and quantitative research, AI-driven research methods are "relaxed, quick, convenient, time-efficient, and sometimes contactless". However, this disparity is unavoidable and has even sparked discussions about its ethical implications [2]. The potential for "Chat GPT, Solab, Ernie Bot, Gemini, Claude, Grok, Llama,

and Copilot", to someday depose human intellect, enable plagiarism, and feed or misinterpret information has frightened many. However, the method academics collect and evaluate data online has been completely transformed by the internet of things [3]. Actually, the vast majority of journals have already begun to accept AI-driven research submissions for potential publication. This indicates that Chat GPT and other forms of artificial intelligence have influenced the field of study. Therefore, given its practical acceptance, mainstreaming AI in research is not far from reality [4].

Data analysis is one area where artificial intelligence is really helpful. It may be quite time-consuming to handle the majority of non-numerical data obtained in qualitative research via focus groups, interviews, stakeholder engagement, Delphi approach, field reconnaissance, field notes, and observations [5]. The ability to interpret and analyse data effectively will be necessary to participate in a meaningful play of codes, themes, and words. One researcher can analyse qualitative data alone, but in certain cases, two or more researchers working together can provide superior themes. On the other side, thematic analysis is required in quantitative research in order to generate themes for exploratory factor theory (EFA) [6]. Even if the statistical software groups the data, the researcher's meaningful analysis is the only factor that will determine the names of the themes or dimensions. This is the point at which artificial intelligence functions effectively and powerfully. By employing the software to statistically organise the thoughts, the AI may be asked about the appropriate codes or topics [7].

A. Artificial intelligence

Artificial intelligence (AI) technologies enable computers to do a broad variety of complex activities, including data analysis, suggestion-making, and the ability to observe, understand, and translate spoken and written language. AI serves as the foundation for innovation in contemporary computing, enabling both consumers and organisations to reap benefits [8]. For instance, "optical character recognition (OCR)" collects text and data from documents and images, converts unstructured content into organised data appropriate for commercial use, and yields useful information [9].

The core of the concept is data, even if the specifics vary throughout AI techniques. Large data sets allow AI systems to learn and develop by identifying connections and patterns that humans might miss. In this learning process, algorithms—which are sets of rules or instructions—are used to direct the AI's decision-making and analysis [10]. In the field of machine learning, which is a popular aspect of artificial intelligence, algorithms are trained on labelled or unlabelled data to generate predictions or classify information. A further specialisation, deep learning employs "multi-layered artificial neural networks" that replicate the organisation and operations of the human brain to process information. As a result of continuous learning and adaptation, AI systems improve their ability to complete specific tasks, including language translation, picture recognition, and beyond [11].

B. Use of AI in research

Acquire research knowledge and literature review: Research tools with AI capabilities for reading, taking notes, and annotating may significantly speed up the learning process. These tools may assist users choose if an item is worth reading by showing them snippets from the literary source with the most

relevant information highlighted. This may assist the user in taking notes on the topic, identifying which paragraphs to read in-depth, and rapidly finding pertinent information in research papers [12]. Users should study the original material rather than only depending on AI-generated summaries and critically evaluate the result without taking it as "the truth" in order to get the most of such an AI-powered tool for research.

Research planning and study design: AI-driven experimental design tools improve parameters using machine learning approaches. Automating experimental design processes may help researchers save more time and effort, freeing up more time for data interpretation and analysis. These AI methods may reduce human error and R&D costs. Researchers must build models that include a large number of variables and factors in order to employ AI technologies for experimental design models in an efficient manner [13]. Researchers may create ideal designs that optimise the efficacy of their studies by feeding certain criteria into these models.

Data analysis: AI-powered data analysis tools have revolutionised the field, as traditional methods of data analysis were reliant on manual processes and had limited computational capabilities. Machine learning algorithms are employed by these tools to interpret, extract, and reveal patterns in extensive datasets. This can increase the efficacy of research output production and reduce the time and cost associated with its production. In order to optimise the utilisation of AI tools for data analysis, researchers must establish explicit project objectives and pinpoint the precise insights and outcomes they intend to achieve through the analysis. Furthermore, it is imperative that they gather germane data and ensure that it is clean, well-organised, and suitable for analysis. Lastly, it is imperative that the researchers identify and determine the AI tools and algorithms that are most appropriate for their objective analysis [14].

Peer review assistance: Submission volumes for peer evaluation are growing steadily. By decreasing the amount of time spent on screening and revising, it is possible to significantly increase academic productivity and save millions of working hours. AI-powered peer review tools have the potential to realise the potential of semi-automated peer review systems. This system could identify potentially low-quality or controversial studies and match reviewers with manuscripts that align with their subject-matter expertise. Although AI is not yet capable of conducting peer review, AI tools can be effectively employed in the peer review process to recommend appropriate journals for an article, conduct initial quality control on submitted manuscripts, and identify reviewers [15].

2 Literature Review

(Bennis & Mouwafaq, 2025) [16] Thematic analysis is a technical and time-consuming component of qualitative research. There is hope that theme analysis will become better and largely automated with the advent of "generative artificial intelligence (A.I.)", especially huge language models. Particularly in complex psychosocial analysis, this study shows how artificial intelligence (AI) may be integrated with qualitative research techniques. As a result, the A.I. deep learning models demonstrated exceptional accuracy and efficiency. The results of this study indicate that "qualitative research methodology" should

prioritise the preservation of analytical rigour by utilising technology that integrates human skills and artificial intelligence capabilities, in accordance with a standardised future checklist for reporting full process transparency.

(Zhang et al., 2025) [17] According to our research, users' skills in interacting with ChatGPT are greatly improved by increasing transparency, offering rapid advice, and enhancing their comprehension of LLMs' capabilities. After conducting a comparison of the attitudes of researchers towards LLM-supported qualitative analysis before and after the co-design process, we have determined that the primary factors contributing to the shift from a negative to a positive perception are a more comprehensive understanding of the LLM's capabilities and the implementation of prompt engineering techniques that enhance response transparency and, as a result, strengthen trust. This study provides insights for qualitative researchers into how AI is perceived in LLM applications, in addition to emphasising the value of well-crafted prompts. Lastly, we highlight the possible ethical hazards and the influence that researchers—especially novices—have on future advancements in AI research and ethical standards.

(Ofosu-Ampong, 2024) [18] Seeks to find research gaps that might direct further studies. The results demonstrate that the existing literature is biassed towards the frequency of technical problems and draw attention to the comparatively insufficient attention paid to other topics, including conceptualisation, application domains, and contextual knowledge co-creation concerns. In addition, the review discovered that modern AI, which is constantly pushing the limits of computational power to address ever-more complex decision-making problems, differs from previous iterations in two key ways that have a big impact on organisational learning when addressing AI's potential: autonomy and learnability. By facilitating the identification of research spaces for subsequent studies, this study advances the field of AI research by providing insights into "current challenges, research methodologies, degree of analysis and conceptual approaches, and AI framework".

(EnP, 2024) [4] To determine the capabilities and applications of Artificial Intelligence in the development of themes in both quantitative and qualitative analysis. In order to reduce data and extract algorithms, the researcher implemented multistage data mining as an AI data collection tool. This methodology was capable of producing the final themes that pertain to the capabilities and applications of AI models. These topics include insight production, objectivity and bias reduction, effective data processing, sentiment analysis, and sentiment recognition analysis. In the fields of "data mining, data processing, and data analysis", this utilisation of AI models and its capabilities can be employed to underscore the ethical and advantageous implementations of AI in research.

(Christou, 2023) [19] Motivated by conceptualisation and critical methodological dynamics, the goal of this work is to explore the function of artificial intelligence in the process of developing theories. As such, it offers a conceptual map of the relationship between AI and theory formation, critically assesses the potential and constraints of AI in theory construction, and outlines important factors to take into account when using AI to establish new theories or advance current ones. It is debatable if AI tools are

necessary for developing theories since researchers' cognitive and evaluative abilities are seen to be crucial in this process, however the importance of AI in developing theory should not be understated.

(Fridgeirsson et al., 2023) [20] Examine the primary ways that cost, risk, and schedule will be impacted by artificial intelligence (AI) in project management. A set of Master of Project Management degree graduates were evaluated using an online survey that reflected the PMBOK's best practices and the future impact of artificial intelligence on the project management field. Resource cost estimate is thought to have the most impact on project cost management. The use of AI in project risk management will most significantly affect the probability and effect forms.

(Hamilton et al., 2023) [21] Examines how human-centered activities like qualitative study analysis may be enhanced by the AI chatbot ChatGPT. The findings show both parallels and discrepancies between analyses produced by AI and humans, with human coders identifying some patterns that ChatGPT missed and vice versa. According to the study's findings, artificial intelligence (AI) tools like ChatGPT provide a potent way to support intricate human-centered work. It also forecasts that these technologies will eventually be used as a supplementary tool to help with research projects. To help find omissions, different frames, and personal biases, future study should investigate putting raw interview transcripts into ChatGPT and introducing AI-generated topics into triangulation conversations.

(Morgan, 2023) [22] The potential application of artificial intelligence programs, such as a ChatGPT, to analyse qualitative data poses a variety of concerns, chiefly the feasibility of achieving comparable results without the time-consuming and laborious process of manual classification. The findings indicate that ChatGPT performed satisfactorily; however, it was less effective in identifying nuanced, interpretive themes and more successful in reproducing concrete, descriptive themes in both instances. This resulted in a program that was relatively straightforward to operate and necessitated minimal effort in comparison to methods that required manual coding. It is crucial to acknowledge, however, that both artificial intelligence-based processes and coding are merely instruments that must be implemented as part of a more comprehensive analytical process. The results of this investigation indicate that artificial intelligence may have the potential to challenge the dominant paradigm of qualitative data analysis, which is the categorisation of data segments.

(Collins et al., 2021) [23] In recent years, the information systems (IS) research community has directed more of its focus on AI. However, a rising number of people are worried that AI research may not be able to gather up as much information as IS research has in the past. 98 of the 1877 papers that were found using the search approach were deemed main studies, and a summary of the major topics that are relevant to this investigation is provided. Determining the current stated economic worth and contributions of AI, research and practical implications on the use of AI, and the possibilities for future AI research in the form of a research agenda are all significant contributions made by this study.

(Serey et al., 2021) [24] With an emphasis on the most significant and relevant works to date, this study explores the main problems, trends, technological developments, and artificial intelligence strategies developed by emerging machine learning researchers and specialists. In the study of data management

using AI techniques, the four machine learning categories that show symmetry are "unsupervised learning, semi-supervised learning, supervised learning, and reinforced learning". Additionally, K-means, Bayesian approaches, support vector machines, and artificial neural networks are the AI techniques with greater symmetry across all groups. Lastly, five research directions to enhance machine learning prediction are offered.

3 Conclusion

All things considered, this paper highlights how artificial intelligence, and in particular "large language models (LLMs) like Chat GPT", may revolutionise research technique. AI tools streamline data processing, support theme development, and facilitate deeper qualitative analysis, allowing researchers to focus more on interpretation and insight generation. However, for AI to be effectively integrated, it is essential to enhance user transparency, improve understanding of AI capabilities, and provide guidance on prompt design. While AI offers powerful assistance, it should be viewed as a supplement—not a substitute—for human expertise. The synergy between AI capabilities and human judgment is critical for achieving valid, meaningful research outcomes. Tools like Chat GPT can aid triangulation, uncover biases, and offer alternative perspectives, enriching the overall research process. Nevertheless, ethical considerations, transparency in usage, and standardized guidelines are necessary to ensure responsible and effective AI application. The future of AI in research methodology depends on developing frameworks that support critical engagement with AI outputs, fostering both innovation and integrity. As AI continues to evolve, its role in research is likely to grow—comparable to how citation tools or data analysis software have become standard—ushering in a new era of intelligent, efficient, and reflective research practice.

References

[1] İ. Yazici, I. Shayea, and J. Din, “A survey of applications of artificial intelligence and machine learning in future mobile networks-enabled systems,” Eng. Sci. Technol. an Int. J., vol. 44, 2023, doi: 10.1016/j.jestch.2023.101455.

[2] C. Malin, C. Kupfer, J. Fleiß, B. Kubicek, and S. Thalmann, “In the AI of the Beholder—A Qualitative Study of HR Professionals’ Beliefs about AI-Based Chatbots and Decision Support in Candidate Pre-Selection,” Adm. Sci., vol. 13, no. 11, 2023, doi: 10.3390/admsci13110231.

[3] A. M. Saghiri, S. M. Vahidipour, M. R. Jabbarpour, M. Sookhak, and A. Forestiero, “A Survey of Artificial Intelligence Challenges: Analyzing the Definitions, Relationships, and Evolutions,” Appl. Sci., vol. 12, no. 8, 2022, doi: 10.3390/app12084054.

[4] D. M. C. T. J. EnP, “Capabilities and Application of Artificial Intelligence (AI) Models in Qualitative and Quantitative Data Mining, Data Processing and Data Analysis,” Arch. des Sci., vol. 74, no. 3, pp. 198–200, 2024, doi: 10.62227/as/74331.

[5] C. A. Fazakarley, M. Breen, P. Leeson, B. Thompson, and V. Williamson, “Experiences of using artificial intelligence in healthcare: A qualitative study of UK clinician and key stakeholder perspectives,” BMJ Open, vol. 13, no. 12, pp. 1–8, 2023, doi: 10.1136/bmjopen-2023-076950.

[6] S. F. Ahmad et al., "Impact of artificial intelligence on human loss in decision making, laziness and safety in education," Humanit. Soc. Sci. Commun., vol. 10, no. 1, pp. 1–14, 2023, doi: 10.1057/s41599-023-01787-8.

[7] A. M. Vieriu and G. Petrea, "The Impact of Artificial Intelligence (AI) on Students' Academic Development," MDPI, vol. 2, pp. 153–158, 2025, doi: 10.30983/bicc.v1i1.121.

[8] N. H. Patil, S. H. Patel, and S. D. Lawand, "Research Paper On Artificial Intelligence And It's Applications," J. Adv. Zool., vol. 44, pp. 229–238, 2023, doi: 10.53555/jaz.v44is8.3544.

[9] S. A. Z. Zaidi, E. Ahmad, and N. Shukla, "Ethical Considerations in the Use of Artificial Intelligence (AI) for Education and Research : A Review," Int. J. Innov. Sci. Eng. Manag., pp. 156–167, 2024, doi: 10.69968/ijisem.2024v3si2156-167.

[10] Z. Slimi, "The Impact of Artificial Intelligence on Higher Education: An Empirical Study," Eur. J. Educ. Sci., vol. 10, no. 1, pp. 17–33, 2023, doi: 10.19044/ejes.v10no1a17.

[11] M. Chen et al., "Acceptance of clinical artificial intelligence among physicians and medical students: A systematic review with cross-sectional survey," Front. Med., vol. 9, no. August, 2022, doi: 10.3389/fmed.2022.990604.

[12] L. Yu and Z. Yu, "Qualitative and quantitative analyses of artificial intelligence ethics in education using VOSviewer and CitNetExplorer," Front. Psychol., vol. 14, no. March, 2023, doi: 10.3389/fpsyg.2023.1061778.

[13] S. Castagno and M. Khalifa, "Perceptions of Artificial Intelligence among Healthcare Staff: A Qualitative Survey Study," Front. Artif. Intell., vol. 3, no. October, pp. 1–7, 2020, doi: 10.3389/frai.2020.578983.

[14] M. Govindaraj, C. Gnanasekaran, S. Thandayuthapani, S. V. Gnanamanickam, and P. Khan, "Role of Artificial Intelligence across Various Media Platforms: a Quantitative Investigation of Media Expert'S Opinion," J. Law Sustain. Dev., vol. 11, no. 5, pp. 1–14, 2023, doi: 10.55908/SDGS.V11I5.1175.

[15] S. Secinaro, D. Calandra, A. Secinaro, V. Muthurangu, and P. Biancone, "The role of artificial intelligence in healthcare: a structured literature review," BMC Med. Inform. Decis. Mak., vol. 21, no. 1, pp. 1–23, 2021, doi: 10.1186/s12911-021-01488-9.

[16] I. Bennis and S. Mouwafaq, "Advancing AI-driven thematic analysis in qualitative research: a comparative study of nine generative models on Cutaneous Leishmaniasis data," BMC Med. Inform. Decis. Mak., vol. 25, no. 1, 2025, doi: 10.1186/s12911-025-02961-5.

[17] H. E. Zhang, C. Wu, J. Xie, Y. Lyu, J. Cai, and J. M. Carroll, "Harnessing the Power of AI in Qualitative Research: Exploring, Using and Redesigning ChatGPT," Comput. Hum. Behav. Artif. Humans, vol. 1, no. 1, p. 29, 2025, doi: 10.1016/j.chbah.2025.100144.

[18] K. Ofosu-Ampong, "Artificial intelligence research: A review on dominant themes, methods, frameworks and future research directions," Telemat. Informatics Reports, vol. 14, no. March, p. 100127, 2024, doi: 10.1016/j.teler.2024.100127.

[19] P. A. Christou, "The Use of Artificial Intelligence (AI) in Qualitative Research for Theory Development," Qual. Rep., no. September, 2023, doi: 10.46743/2160-3715/2023.6536.

[20] T. V. Fridgeirsson, H. T. Ingason, H. I. Jonasson, and H. Gunnarsdottir, “A Qualitative Study on Artificial Intelligence and Its Impact on the Project Schedule, Cost and Risk Management Knowledge Areas as Presented in PMBOK®,” Appl. Sci., vol. 13, no. 19, 2023, doi: 10.3390/app131911081.

[21] L. Hamilton, D. Elliott, A. Quick, S. Smith, and V. Choplin, “Exploring the Use of AI in Qualitative Analysis: A Comparative Study of Guaranteed Income Data,” Int. J. Qual. Methods, vol. 22, pp. 1–13, 2023, doi: 10.1177/16094069231201504.

[22] D. L. Morgan, “Exploring the Use of Artificial Intelligence for Qualitative Data Analysis: The Case of ChatGPT,” Int. J. Qual. Methods, vol. 22, pp. 1–10, 2023, doi: 10.1177/16094069231211248.

[23] C. Collins, D. Dennehy, K. Conboy, and P. Mikalef, “Artificial intelligence in information systems research: A systematic literature review and research agenda,” Int. J. Inf. Manage., vol. 60, no. November 2020, p. 102383, 2021, doi: 10.1016/j.ijinfomgt.2021.102383.

[24] J. Serey et al., “Artificial intelligence methodologies for data management,” Symmetry (Basel), vol. 13, no. 11, 2021, doi: 10.3390/sym13112040.

AG Publishing House

Advanced Research Techniques: Theories, Methods, And Practices (Volume-2)

Volume 2
Year: 2025

AG PH Books

The Role of Artificial Intelligence in Scientific Research

Dr. Pullaiah.Cheepi [1*]

[1]*Assistant Professor*

Abstract

In the rapidly developing area of artificial intelligence (AI), intelligence that can comprehend natural language, identify patterns, and make data-driven decisions—tasks that normally require human intellect—is developed. The many studies on artificial intelligence in scientific inquiry are reviewed in this article. It concluded that artificial Intelligence is reshaping scientific research by enhancing data analysis, accelerating results, and enabling the handling of large, complex datasets. Tools like Chat GPT streamline qualitative research by reducing manual effort and uncovering hidden patterns, making analysis more efficient and scalable. However, ethical, privacy, and governance concerns must be addressed, and AI should be viewed as a complement—not a replacement—for human cognitive input, particularly in theory development. With proper guidance, transparency, and understanding of AI capabilities, researchers can effectively integrate AI tools to advance scientific inquiry while maintaining the integrity and depth of traditional research methodologies.

Keywords: Artificial intelligence (AI), Quantitative Research, Research methodologies, Human intelligence, Scientific research, Qualitative Research, etc.

1 Introduction

While incorporating AI into research procedures must carefully address ethical, privacy, and governance concerns, it also presents previously unheard-of potential to improve data processing, boost productivity, and promote creative discoveries. With the use of artificial intelligence (AI) technology, especially generative models, researchers may test hypotheses and comprehend system dynamics without doing

[*] ISBN No. - 978-93-49028-02-9

physical experiments by simulating complicated study contexts and situations [1]. AI's ability to process and analyse large datasets allows researchers to find patterns and insights at a scale that was previously impossible, which leads to more accurate scientific enquiries and better informed decision-making. It also facilitates cross-disciplinary research by allowing the integration and analysis of data from various fields, which leads to holistic approaches to solving complex global challenges [2]. Finally, it speeds up the research lifecycle by rapidly prototyping experiments and analysing their results, which significantly accelerates the pace of innovation. Regardless of whether the tested hypothesis was validated or not by the quantitative analysis in the basic research, this work introduces innovative methodologies that will ultimately result in the creation of working models. These models will be created by starting with research hypotheses and using AI tools specific to each research hypothesis to achieve the specific objective [3].

A. Artificial Intelligence

A variety of sophisticated functions, such as the capacity to see, comprehend, and translate spoken and written language, analyse data, and provide recommendations, are made possible by artificial intelligence (AI) technologies [4]. Artificial intelligence is a scientific discipline that focusses on the development of computers and machines that are capable of reasoning, learning, and acting in a manner that would typically necessitate human intelligence or which entails data that is too large for humans to analyse [5], [6]. Artificial intelligence is a multifaceted field that includes a variety of disciplines, such as computer science, data analytics and statistics, hardware and software engineering, linguistics, neurobiology, and even philosophy and psychology [7]. A set of technologies that are predominantly based on machine learning and deep learning, AI is a set of technologies that are used for data analytics, predictions and forecasting, object categorisation, natural language processing, recommendations, intelligent data retrieval, and more on an operational level for business use [8].

B. Scientific Research

Many methods exist for defining research. In essence, research is the process of collecting, organising, and utilising information to elucidate, substantiate, or prove a theory or area of study. Research is essential in both scientific and non-scientific fields; however, scientific research is one of the most comprehensive and pertinent categories of research [9]. In the past few centuries, the field of scientific research has made remarkable progress and achieved remarkable feats. This trend is expected to persist in the near future. In order to elucidate the reasons behind specific phenomena that manifest in the real world, scientific research employs a diverse array of scientific models, theories, and data collections [10]. A variety of methods can be employed to conduct scientific research, including experiments, case studies, and focus groups. Enlarging human knowledge is the ultimate objective of scientific research. Our comprehension of the workings of the world is significantly enhanced by scientific research [11], [12]. It also contributes to the advancement of a variety of disciplines, including biology, chemistry, psychology, medicine, and other disciplines. Furthermore, scientific research addresses both existing issues and potential future ones [13].

C. Types of Scientific Research

Scientific study comes in a wide variety of forms, ranging from mixed to quantitative and qualitative. Since there are many different classifications for it, the categorisation mostly relies on the study topic, data gathering methods, and research methodology. Some of the most fundamental forms of scientific investigation are as follows:

Quantitative Research: As the name suggests, investigations based on numbers are included in quantitative research. In quantitative approaches, some kind of data is either measured or counted. The goal of this kind of scientific study is to assess the what, when, and where of the research subject or query since it is numerical in nature. Graphs, statistics, and numbers are the primary ways it is represented. Methods for gathering quantitative data include experiments, surveys, and observations. To find patterns, it essentially measures numerical data.

Qualitative Research: The quality of the data is the primary focus of qualitative research, rather than its quantity. This is a form of descriptive and exploratory research, which is why qualitative methods such as literature reviews, interviews, focus groups, and case studies are employed. Qualitative research, in contrast to quantitative research, is designed to enquire about the rationale and methodology behind the decision-making process. Through the organisation, interpretation, and summarisation of data, this is accomplished. To put it differently, qualitative research is employed when the objective is to comprehend textual data, as opposed to quantitative research, which attempts to verify a hypothesis through numerical data. This type of investigation is designed to enhance comprehension of the subject matter or the prospective issues it may raise. In the end, this results in a solution that may or may not be grounded in empirical data.

Mixed Research: Combining quantitative and qualitative methodologies is known as mixed research. The combined approach is beneficial because it balances the shortcomings of both qualitative and quantitative research while enhancing their strengths.

D. Example of AI in Scientific Research

AI-Driven Robotic Scientists: AI algorithms are currently being implemented in robotic systems, which are frequently referred to as "robot scientists." These algorithms enable them to independently analyse data, refine hypotheses, and conduct experiments. Additionally, these robotic scientists are capable of conducting genetic screenings and synthesising compounds.

Drug Discovery and Development: Advanced By utilising AI in scientific research, scientists can expedite the discovery of new medications by generating novel drug molecules with distinctive structures and characteristics. AI is capable of predicting the potential interactions between various medications. Making pharmaceuticals safer for patients can be achieved through the analysis of extensive data.

Bioinformatics: Since AI-based protein structure predictions can forecast the behaviour of specific components, they are a key tool for comprehending protein function and medication development. On

the basis of genomic data analysis, machine learning can identify genetic variations among species and, as a result, enable the development of personalised therapies. Various biological functions, such as protein-protein interactions, are employed in scientific research by AI to investigate potential drug targets.

Real-World Example: Material Science Research: Another area that benefits from AI-driven automation is material science, which develops novel chemicals and materials for uses such electronics or renewable energy. The discovery of materials with desired qualities like "strength, conductivity, or heat resistance" may be accelerated by AI-powered robotic systems, which can investigate and test material combinations significantly more quickly than human researchers.

Reducing Human Risk in Hazardous Research: AI-driven robotic systems may do investigations in areas involving radioactive materials, toxic chemicals, or other risky substances without endangering human researchers. In chemistry, where studies sometimes call for poisonous or reactive compounds, this skill is very helpful.

E. Sector-Specific Applications- specially focus on social science research

Political Science: Because AI provides new methods for data analysis, predictive modelling, and public opinion comprehension, it is being employed more and more in "political science for social science research". AI can improve research and decision-making by analysing massive databases, finding trends, and simulating the possible outcomes of political actions.

Education: AI is being used more and more in education social science research for tasks including data analysis, individualised instruction, and improved administrative effectiveness. AI can automate instructor duties, personalise learning experiences, and analyse massive datasets to determine student learning patterns.

Public Policy: AI has a number of uses in public policy, particularly for social science research. These include better policy decision-making, predictive modelling, and increased data analysis. Additionally, it may be used to improve public involvement, track social trends, and model social interactions.

Economics: AI provides strong tools for economics social science research, allowing for the study of big datasets, pattern recognition, and trend forecasting. These applications have the potential to promote multidisciplinary cooperation, improve policymaking, and deepen our knowledge of human behaviour.

2 Literature Review

(Zhang et al., 2025) [14] To create a framework for creating prompts especially made to assist younger researchers and stakeholders interested in using AI for qualitative research, a co-design session with 13 qualitative researchers and semi-structured interviews with 17 participants were undertaken. This study provides insights for qualitative researchers into how AI is perceived in LLM applications, in addition to emphasising the value of well-crafted prompts. Lastly, we highlight the possible ethical hazards and

the influence that researchers—especially novices—have on future advancements in AI research and ethical standards.

(Hitch, 2024) [15] NLP-based AI will be defined and explained in this paper, along with its advantages and disadvantages for reflective theme analysis in health research. ChatGPT is the most popular and most accessible of the several platforms. To demonstrate possible use in reality, a working example that augments reflective thematic analysis with ChatGPT is shown. The purpose of this paper is to stimulate further discussion on the use of AI in qualitative research and to provide useful advice for investigators who want to use this technology.

(Mădălina, 2024) [3] Examines the ways in which incorporating Artificial Intelligence (AI) into research methods is altering the paradigm of research by enhancing data analysis, productivity, and creativity across industries. In addition, this proposed study will take into account the ELM framework and expand its dimension, offering a more modern viewpoint on persuasion and cognitive engagement by incorporating AI and big data into the ELM methodologies dimension. This will demonstrate a paradigm shift in research and serve as an example of how to transition from methodologies based on conventional qualitative and quantitative methods to emerging ones based on AI and BIG DATA.

(Waly, 2024) [16] This article investigates the correlation between artificial intelligence (AI) and scientific inquiry. Researchers can uncover patterns, correlations, and trends that may otherwise remain obscured by utilising AI to facilitate the extraction of meaningful insights from large datasets. Additionally, artificial intelligence facilitates predictive analytics, enabling scientists to anticipate outcomes and pinpoint potential areas for additional research. In addition, the utilisation of AI systems in experimental design and optimisation is on the rise, which has the effect of expediting processes and improving efficiency in laboratory environments. Although AI offers numerous advantages, the integration of AI into scientific research poses obstacles regarding ethical considerations, interpretability, and data quality.

(Christou, 2023) [17] Motivated by conceptualisation and critical methodological dynamics, the goal of this work is to explore the function of artificial intelligence in the process of developing theories. As such, it offers a conceptual map of the relationship between AI and theory formation, critically assesses the potential and constraints of AI in theory construction, and outlines important factors to take into account when using AI to establish new theories or advance current ones. It is debatable if AI tools are necessary for developing theories since researchers' cognitive and evaluative abilities are seen to be crucial in this process, however the importance of AI in developing theory should not be understated.

(Hamilton et al., 2023) [18] Investigates how the artificial intelligence chatbot ChatGPT may support human-centered activities like qualitative study analysis. In this research, the themes that emerged from qualitative analyses of interviews with guaranteed income pilot participants conducted by humans and AI are compared. The findings show that human and AI-generated analyses vary and are comparable in that human coders identified some motifs that Chat GPT did not, and vice versa. According to the study's findings, artificial intelligence (AI) technologies like Chat GPT provide a potent way to support intricate

human-centered work. It also forecasts that these tools won't be the only ones used to make research jobs easier. To find mistakes, different perspectives, and biases, future studies should investigate feeding Chat GPT raw interview transcripts and adding AI-generated topics to triangulation conversations.

(Morgan, 2023) [19] Particularly, it investigates the degree to which the themes that were initially selected to summarise the two previous analyses can be replicated in the responses from Chat GPT. The findings indicate that Chat GPT performed satisfactorily; however, it was less effective in identifying nuanced, interpretive themes and more successful in reproducing concrete, descriptive themes in both instances. This resulted in a program that was relatively straightforward to operate and necessitated minimal effort in comparison to methods that required manual coding. It is crucial to acknowledge, however, that both artificial intelligence-based processes and coding are merely instruments that must be implemented as part of a more comprehensive analytical process. The results of this investigation indicate that artificial intelligence may have the potential to challenge the dominant paradigm of qualitative data analysis, which is the categorisation of data segments.

(Mungara, 2022) [20] Examine the studies conducted on artificial intelligence methods in big data analytics. The Systematic Literature Review (SLR) approach is used by the writers to choose relevant research articles. Four groups—machine learning, knowledge-based and reasoning techniques, decision-making algorithms, search methods, and optimisation theory—are taken into consideration to study these processes. Within each category, many articles are examined. Additionally, this assessment compares the scalability, efficiency, accuracy, and privacy of the chosen AI-driven big data analytics methodologies, highlighting their advantages and disadvantages. Additionally, some key topics are offered to improve big data analytics methods in the future.

3 Conclusion

The integration of Artificial Intelligence (AI) and Big Data tools into scientific research is revolutionizing research paradigms by accelerating data analysis, enhancing accuracy, and enabling the processing of large and complex datasets. AI tools like ChatGPT offer promising advantages in qualitative research by reducing labor-intensive tasks and uncovering hidden patterns through advanced modeling techniques. However, their use also raises ethical, privacy, and governance concerns that necessitate careful adaptation of research methodologies. While AI significantly improves efficiency and scalability in qualitative analysis, the cognitive and evaluative roles of researchers remain vital, especially in tasks such as theory development. Our findings suggest that improved transparency, guidance on AI prompts, and a deeper understanding of AI capabilities can enhance user interaction and foster positive attitudes toward AI use in research. AI should be seen as a complementary tool that supports, rather than replaces, human insight. Despite the challenges in implementation, the potential of AI to transform qualitative and quantitative research is substantial. For meaningful results, researchers must use these tools judiciously, ensuring responsible and ethical integration. As scientific inquiry evolves, AI will continue to shape the future of research by expanding the boundaries of knowledge generation and analysis.

References

[1] M. Elbadawi, H. Li, A. W. Basit, and S. Gaisford, "The role of artificial intelligence in generating original scientific research," Int. J. Pharm., vol. 652, no. January, p. 123741, 2024, doi: 10.1016/j.ijpharm.2023.123741.

[2] M. Khalifa and M. Albadawy, "Using artificial intelligence in academic writing and research: An essential productivity tool," Comput. Methods Programs Biomed. Update, vol. 5, no. March, p. 100145, 2024, doi: 10.1016/j.cmpbup.2024.100145.

[3] C. Mădălina, "NEW APPROACHES IN RESEARCH METHODOLOGY BASED ON ARTIFICIAL INTELLIGENCE AND BIG DATA," no. 2, pp. 65–81, 2024.

[4] J. J. Prieto-Gutierrez, F. Segado-Boj, and F. D. S. França, "Artificial Intelligence in Social Science: a Study Based on Bibliometrics Analysis," Hum. Technol., vol. 19, no. 2, pp. 149–162, 2023, doi: 10.14254/1795-6889.2023.19-2.1.

[5] J. Gao and D. Wang, "Quantifying the Benefit of Artificial Intelligence for Scientific Research," 2023, [Online]. Available: http://arxiv.org/abs/2304.10578

[6] S. A. Z. Zaidi, E. Ahmad, and N. Shukla, "Ethical Considerations in the Use of Artificial Intelligence (AI) for Education and Research : A Review," Int. J. Innov. Sci. Eng. Manag., pp. 156–167, 2024, doi: 10.69968/ijisem.2024v3si2156-167.

[7] T. Magash, "Data Science and Artificial Intelligence in Scientific Research," Res. gate, no. April 2024, pp. 332–360, 2024, doi: 10.58532/v3bcag7p1ch20.

[8] P. K. Verma and L. Kaur, "Research Paper on Role of Data Features and Data Collection Tools in Artificial Intelligence," Int. J. Res. Appl. Sci. Eng. Technol., vol. 12, no. 3, pp. 111–116, 2024, doi: 10.22214/ijraset.2024.58693.

[9] S. Nagar, "Artificial Intelligence in Scientific Research: Lessons for SPIs," no. May, 2024.

[10] M. A. Kacena, L. I. Plotkin, and J. C. Fehrenbacher, "The Use of Artificial Intelligence in Writing Scientific Review Articles," Curr. Osteoporos. Rep., vol. 22, no. 1, pp. 115–121, 2024, doi: 10.1007/s11914-023-00852-0.

[11] S. Erduran and O. Levrini, "The impact of artificial intelligence on scientific practices: an emergent area of research for science education," Int. J. Sci. Educ., 2024, doi: 10.1080/09500693.2024.2306604.

[12] S. Dehariya and D. Sharma, "Study On Highway Roadside Safety Prabhat Chauraha to Bajrang Chauraha Raisen Road Bhopal," pp. 44–50, 2024.

[13] P. A. Christou, "Thematic Analysis through Artificial Intelligence (AI)," Qual. Rep., vol. 29, no. 2, pp. 560–576, 2024, doi: 10.46743/2160-3715/2024.7046.

[14] H. E. Zhang et al., "Harnessing the Power of AI in Qualitative Research: Exploring, Using and Redesigning ChatGPT," Comput. Hum. Behav. Artif. Humans, vol. 1, no. 1, p. 29, 2025, doi: 10.1016/j.chbah.2025.100144.

[15] D. Hitch, "Artificial Intelligence Augmented Qualitative Analysis: The Way of the Future?" Qual. Health Res., vol. 34, no. 7, pp. 595–606, 2024, doi: 10.1177/10497323231217392.

[16] M. F. Waly, “Artificial Intelligence and Scientific Research,” Sustain. Educ. Globe, 2024, doi: 10.4293/JSLS.2024.00006.

[17] P. A. Christou, “The Use of Artificial Intelligence (AI) in Qualitative Research for Theory Development,” Qual. Rep., no. September, 2023, doi: 10.46743/2160-3715/2023.6536.

[18] L. Hamilton, D. Elliott, A. Quick, S. Smith, and V. Choplin, “Exploring the Use of AI in Qualitative Analysis: A Comparative Study of Guaranteed Income Data,” Int. J. Qual. Methods, vol. 22, pp. 1–13, 2023, doi: 10.1177/16094069231201504.

[19] D. L. Morgan, “Exploring the Use of Artificial Intelligence for Qualitative Data Analysis: The Case of ChatGPT,” Int. J. Qual. Methods, vol. 22, pp. 1–10, 2023, doi: 10.1177/16094069231211248.

[20] P. Mungara, “The Role of Artificial Intelligence in Data Analytics,” Int. J. Sci. Res., vol. 3, no. 3, pp. 1722–1729, 2022, doi: 10.1017/dap.2021.20.

AG Publishing House

Advanced Research Techniques: Theories, Methods, And Practices (Volume-2)

Volume 2
Year: 2025

AG PH Books

Machine Learning for Data Analysis in Research: A Review

N V S K Vijayalakshmi K [1*], Dr. J Sasikala [2]

[1]*Assistant professor, Sir C R reddy College of Engineering, Eluru, AP, India.*
[2]*Professor,Dept of IT,Annamalai University,Chidambaram,Tamilnadu.*

Abstract

The implementation of machine learning (ML) in data analytics has drawn significant attention in recent years due to its potential to revolutionise a variety of domains. Reviewing the varied literature on the use of machine learning in research data analysis is the aim of this essay. This review underscores the growing necessity of integrating machine learning (ML) techniques with big data (BD) analytics to address challenges in scalability, speed, and complexity. ML offers promising tools for generating valuable insights and driving smarter decision-making in various domains. However, to fully harness its potential, future research must focus on developing adaptable, cost-effective, and robust ML models capable of handling noisy, incomplete, and real-time data. As data volumes grow, the synergy between ML and BD will remain vital in advancing analytical capabilities, particularly for Internet of Things (IoT) applications and business intelligence.

Keywords: Machine Learning, Data Analysis, Deep Learning, Internet of Things (IoT), Data Processing, Decision-Making, etc.

1 Introduction

The commercial value of Big Data (BD) is enormous in a number of industries, such as "healthcare, transportation, e-business, power monitoring, and economic services". However, when confronted with this vast quantity of data, the conventional method is unable to effectively conduct data analysis. “ABI (Advance Business Intelligence)” conducted research that suggests that there will be more than 30 billion interconnected devices available to satisfy informational requirements [1]. Performing data

* ISBN No. - 978-93-49028-02-9

administration, processing, and analysis is a complex and challenging task due to the immense volume of data that these actual systems can generate from a variety of resources. It is a challenging issue for numerous industries and organizations to integrate the IT departments, healthcare companies, government agencies, and research institutions of today [2]. To solve this problem, BD science was given its own domain, and in order to support effective and quick growth, new trends in research and teaching are needed. The effectiveness and efficiency of Machine Learning (ML) in BD analysis are contingent upon the algorithms and the dataset's configuration, which necessitates numerous time-consuming operations. In practice, not every system can guarantee peak performance without changing the module [3]. With the use of new scientific discoveries that can be combined with machine learning (ML) systems to make decisions, BD solutions may provide high-performance outcomes quickly. ML is an influential instrument for managing BD, according to a variety of studies. As seen in, this is similar to the connection between BD and the ML correlation between the sources and individual learning. Based on this viewpoint, individuals are capable of acquiring knowledge from the sources in order to address innovative issues. Similarly, they are capable of resolving novel issues by acquiring knowledge from BD [4].

A. Machine learning

Machine learning (ML), a branch of artificial intelligence, is the study and development of "statistical algorithms" that may gain insight from data and generalise to new data, performing tasks without explicit instructions. Neural networks are a class of statistical algorithms that have outperformed many earlier machine learning techniques in a subdiscipline of machine learning thanks to developments in the area of deep learning [5]. Machine learning is used in many fields, including "computer vision, speech recognition, email filtering, natural language processing, agriculture, and health". Predictive analytics is the use of machine learning to solve business issues. Techniques for mathematical optimisation, or mathematical programming, and statistical analysis form the foundation of machine learning. The study of exploratory data analysis (EDA) using unsupervised learning is the focus of the related discipline of data mining [6].

B. Importance of machine learning

The amount of data being generated is increasing daily. Every day, the globe generates more data than it has ever done before. Without machine learning, it would be very difficult to evaluate and use all that data. Machine learning is therefore creating a whole new range of possibilities for human interaction with computers and other devices [7]. Machine learning has many applications for businesses, such as data analysis, translation and transcription, chatbots for automated customer support, fraud detection, security threat identification, personalisation, and recommendations. Additionally, machine learning is propelling the exciting innovations of the future, like robots, augmented and virtual reality, drones, and autonomous cars [8].

C. Data Analytics

Data analytics is the act of examining, transforming, and evaluating raw data to identify important "trends, patterns, and insights that might guide decisions". This involves using statistical techniques, tools, and algorithms to transform massive volumes of data into insights that may be put to use. Data analytics is employed extensively in a variety of sectors to enhance corporate performance, optimise operations, anticipate trends, and resolve complex problems [9]. The following lists the essential elements of data analytics:

Data Collection: First, information is gathered from a variety of sources, including databases, websites, sensors, and surveys. Data may be either unstructured (text, graphics, etc.) or structured (rows and columns).

Data Cleaning: Prior to analysis, the data is cleaned to remove errors, duplicates, missing values, and inconsistencies. This guarantees the data's accuracy and dependability.

Data Processing: The data is cleaned and then converted into an analysis-ready format. This might include aggregating data, classifying variables, or normalizing values.

Data Analysis: To find patterns, correlations, and trends in the data, analytical methods including machine learning algorithms, regression analysis, and descriptive statistics are used.

Data Visualization: Understanding the analysis's findings and sharing insights with stakeholders is facilitated by the visual representation of data provided by charts, graphs, and dashboards.

Interpretation and Decision-Making: The analysis's conclusions are interpreted to help guide choices. Businesses might use these insights to improve productivity, address particular issues, or optimize their plans.

D. Roles of Machine Learning in Data Analytics

Automating Data Processing and Analysis: Data processing and analysis automation is one of machine learning's main functions in data analytics. Cleaning, organizing, and processing data using traditional data analysis techniques involves a lot of human labor and often entails tedious job completion. By automating these processes, machine learning saves time and effort needed to prepare data for analysis, allowing data analysts and scientists to focus on more complex and strategic tasks [10].

Identifying Hidden Patterns and Insights: One of the most significant benefits of machine learning in data analytics is its ability to uncover "hidden patterns, trends, and correlations in data" that may be too complex for traditional methods to detect. Large datasets may be rapidly analyzed by ML algorithms, which can provide insights that would otherwise go unnoticed. Machine learning improves businesses' capacity to make data-driven choices based on more profound and nuanced insights by spotting these patterns [11].

Enhancing Predictive Analytics: Predictive analytics is one of machine learning's most well-known applications in data analytics. Machine learning greatly increases the precision and effectiveness of

predictive analytics' forecasts, which are based on previous data and forecast future patterns or occurrences. By applying machine learning to improve prediction accuracy, organisations may make better decisions by anticipating changes in "the market, customer behaviour, and operational risks" [12].

Enabling Real-Time Analytics and Decision-Making: In today's fast-paced business environment, organisations need to make decisions promptly. Traditional data analysis methods often cannot keep up with the volume and speed of incoming data. As a result of machine learning's capacity to assess vast quantities of data in real time, organisations can make data-driven decisions more promptly. Enabling real-time analytics, machine learning provides companies with a competitive edge by allowing them to respond promptly to shifts in "the market, consumer behaviour, or operational circumstances" [13].

Enhancing Data Visualization and Interpretability: The output of "machine learning algorithms" must be actionable and understandable to humans, regardless of their capabilities.. Machine learning has made a substantial contribution to enhancing data visualisation and interpretability, which helps decision-makers better understand complex data and insights. By making data simpler to comprehend and explain, machine learning helps bridge the gap among "technical data analysis and practical business insights" [13].

Driving Prescriptive Analytics: In addition to predictive analytics, machine learning is also playing a critical role in prescriptive analytics, which involves the recommendation of the most effective course of action to achieve desired results. Prescriptive analytics, which is facilitated by machine learning, offers organizations actionable insights that assist them in optimizing their operations and achieving superior results.

2 Literature Review

(Banu et al., 2024) [14] The integration of machine learning algorithms into data-driven projects is examined in this paper, with a focus on a methodical approach to project conception, data collection, preprocessing, model building, assessment, deployment, and monitoring. The research moves forward by choosing and training suitable machine learning algorithms, using techniques like hyperparameter tweaking and cross-validation to maximize model accuracy and generalizability. Along with techniques for real-time prediction and analysis, the shift from model building to deployment in a production setting is covered. In order to ease stakeholder comprehension, the study's conclusion entails producing actionable findings and creating clear visualizations. For openness and future reference, thorough documenting of methods and model setups is recommended. By using machine learning (ML) in data analytics, this all-encompassing strategy seeks to promote operational efficiency and well-informed decision-making. (Chakraborty et al., 2024) [15] This article reviews some of the important data-driven features of "artificial intelligence" in the medical domain. The article showcased the latest developments in "data-driven medical science" through the use of machine learning (ML) and deep learning (DL) in two different areas: first, the use of chabot technologies in medicine and healthcare, specifically ChatGPT, and second, the latest developments in data science in healthcare through the use of ML and DL. DL-enabled ChatGPT technology was recently developed and illustrated in the article. In

conclusion, we provide a succinct summary of the wide-ranging applications of machine learning and deep learning in medicine, as well as the significant challenges that need to be addressed to incorporate the latest "ML to DL technologies into healthcare". Researchers would greatly benefit from the article's thorough summary of "the data-driven paradigm" change in medicine, which is presented using ML to DL technologies.

(Gupta & Gupta, 2023) [16] In this paper, we examine the state-of-the-art in big data analytics and the learning model for healthcare machine learning decision support systems, and we talk about how they could affect patient outcomes and healthcare delivery. To define the big data field, a preliminary survey was conducted. Many hospitals have embraced Hadoop technologies for data analysis because they provide an effective distributed data storage solution that lowers the need for expensive storage devices like disks and others. For the intelligent behavior of the proposed model, this method suggests supervised learning using a hidden Markov model.

(Towler et al., 2023) [17] Evaluated MATA's ability to assist healthcare treatments by contrasting it with "human-only" thematic analysis methods on the same dataset (1,472 user answers from a behavioral intervention for COVID 2019). Findings that varied somewhat were the result of human judgments or subtleties in the study. Conversation: MATA's results were comparable to those of theme analysis done by humans, although it took a lot less time. MATA can help qualitative researchers swiftly comprehend and analyze huge datasets for basic studies that don't need a deep or nuanced comprehension of the data. During public health crises, for example, this method may provide quick optimization and aid in the creation and execution of interventions.

(Alsharif et al., 2020) [18] Network systems are equipped to analyze the data, gain knowledge, and make judgments based on the available dataset. They can also access a variety of experimental symmetric data from a wide range of network devices. This research only looks at supervised and unsupervised machine learning (ML) approaches, which are thought to be the foundation of smart data analysis for the Internet of Things. In addition to discussing the benefits and drawbacks of each algorithm, this paper covers and analyzes important topics pertaining to supervised and unsupervised machine learning approaches. It also looks at research trends and suggests areas for further investigation.

(Kolisetty & Rajput, 2020) [4] Understanding the significance of machine learning in large-scale data analysis is the goal of this research. It helps to comprehend the ramifications and difficulties associated with data heterogeneity, classification error, and computational complexity in big data. The capacity to use data transformation and knowledge extraction to extract value from massive amounts of data for predictive analysis and decision-making is covered. In addition to discussing the degree to which machine learning may be used to big data analysis, it will provide suggestions on how big data may affect real-time data analysis. Additionally, it will highlight the significance and potential from the perspective of promoting feature research advancement in the big data ML sector. (Qiu et al., 2016) [19] summarise the most recent discoveries in the field of machine learning that pertain to the processing of substantial quantities of data. First, examine the various machine learning approaches. Subsequently, identify some of the most intriguing approaches from recent research, such as deep learning,

representation learning, distributed and parallel learning, transfer learning, active learning, and kernel-based learning". The focus should then shift to the analysis and discussion of the challenges and possible solutions related to machine learning for huge data. Then, for large data processing, look at how machine learning and signal processing methods are closely related. Give a summary of a few unresolved problems and current research directions.

3 Conclusion

In conclusion, this review underscores the growing significance of machine learning (ML) in addressing the challenges of big data (BD) analysis across diverse research and industry domains. As data volumes expand and diversify, traditional ML techniques often fall short due to limitations in scalability, speed, and adaptability. Integrating ML with BD analytics enables more accurate predictions, real-time processing, and efficient decision-making, fostering improved business outcomes and research insights. The fusion of ML with moderate-cost systems offers practical alternatives to high-end computational setups, making advanced analytics more accessible. However, challenges such as data noise, incompleteness, and algorithm sensitivity persist, especially in smart data and IoT contexts. This highlights the need for developing more robust, scalable, and noise-resilient ML models. Future research should focus on creating adaptive frameworks, exploring semi-supervised learning, and enhancing ML's ability to process heterogeneous and high-velocity data. Ultimately, the synergy between ML and BD analytics holds transformative potential, but achieving its full benefits requires continuous innovation, ethical considerations, and contextual understanding of the data environments. By evolving to meet the demands of big data, ML can serve as a powerful enabler of operational excellence, informed decision-making, and scientific discovery in the digital age.

References

[1] A. E. Ezugwu, J. Greeff, and Y. S. Ho, "A comprehensive study of groundbreaking machine learning research: Analyzing highly cited and impactful publications across six decades," J. Eng. Res., vol. 13, no. 1, pp. 371–383, 2025, doi: 10.1016/j.jer.2023.10.027.

[2] R. Pugliese, S. Regondi, and R. Marini, "Machine learning-based approach: Global trends, research directions, and regulatory standpoints," Data Sci. Manag., vol. 4, no. December, pp. 19–29, 2021, doi: 10.1016/j.dsm.2021.12.002.

[3] H. S. Sanil et al., "Role of machine learning in changing social and business eco-system – a qualitative study to explore the factors contributing to competitive advantage during COVID pandemic," World J. Eng., vol. 19, no. 2, pp. 238–243, 2021, doi: 10.1108/WJE-06-2021-0357.

[4] V. V. Kolisetty and D. S. Rajput, "A review on the significance of machine learning for data analysis in big data," Jordanian J. Comput. Inf. Technol., vol. 6, no. 1, pp. 41–57, 2020, doi: 10.5455/jjcit.71-1564729835.

[5] J. Kumari, E. Kumar, and D. Kumar, "A Structured Analysis to study the Role of Machine Learning and Deep Learning in The Healthcare Sector with Big Data Analytics," Arch. Comput. Methods Eng., vol. 30, no. 6, pp. 3673–3701, 2023, doi: 10.1007/s11831-023-09915-y.

[6] G. Gouthami, N. Siddartha, and D. P. Borugadda, "Data Science: the Impact of Machine Learning," Futur. Trends Artif. Intell. Vol. 3 B. 8, no. May, pp. 153–170, 2024, doi: 10.58532/v3bgai8p2ch7.

[7] S. Tufail, H. Riggs, M. Tariq, and A. I. Sarwat, "Advancements and Challenges in Machine Learning: A Comprehensive Review of Models, Libraries, Applications, and Algorithms, Electron" vol. 12, no. 8, 2023, doi: 10.3390/electronics12081789.

[8] S. Uddin, S. Yan, and H. Lu, "Machine learning and deep learning in project analytics: methods, applications and research trends," Prod. Plan. Control, vol. 0, no. 0, pp. 1–20, 2024, doi: 10.1080/09537287.2024.2320790.

[9] B. Khosravi et al., "Demystifying Statistics and Machine Learning in Analysis of Structured Tabular Data," J. Arthroplasty, vol. 38, no. 10, pp. 1943–1947, 2023, doi: 10.1016/j.arth.2023.08.045.

[10] A. Brnabic and L. M. Hess, "Systematic literature review of machine learning methods used in the analysis of real-world data for patient-provider decision making," BMC Med. Inform. Decis. Mak., vol. 21, no. 1, 2021, doi: 10.1186/s12911-021-01403-2.

[11] S. Uddin, S. Ong, and H. Lu, "Machine learning in project analytics: a data-driven framework and case study," Sci. Rep., vol. 12, no. 1, pp. 1–13, 2022, doi: 10.1038/s41598-022-19728-x.

[12] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," SN Comput. Sci., vol. 2, no. 3, pp. 1–21, 2021, doi: 10.1007/s42979-021-00592-x.

[13] N. Ben and S. K. Yadav, "Women Empowerment - A Study of SHGS in Satna & Maihar District of Madhya Pradesh," Int. J. Innov. Sci. Eng. Manag., pp. 324–332, 2025, doi: 10.69968/ijisem.2025v4i1324-332.

[14] A. Banu, J. Sravanthi, S. Bolugoddu, L. Panjala, and S. Hasanoddin, "Data-Driven Insights : Applying Machine Learning in Data Analytics," Intell. Syst. Appl. Eng., 2024.

[15] C. Chakraborty, M. Bhattacharya, S. Pal, and S. S. Lee, "From machine learning to deep learning: Advances of the recent data-driven paradigm shift in medicine and healthcare," Curr. Res. Biotechnol., vol. 7, no. August 2023, p. 100164, 2024, doi: 10.1016/j.crbiot.2023.100164.

[16] D. Gupta and S. Gupta, "Big Data Analysis using Machine Learning for Health Systems," Int. J. Comput. Appl., vol. 185, no. 9, pp. 8–13, 2023, doi: 10.5120/ijca2023922745.

[17] L. Towler et al., "Applying machine-learning to rapidly analyze large qualitative text datasets to inform the COVID-19 pandemic response: comparing human and machine-assisted topic analysis techniques," Front. Public Heal. vol. 11, no. October, pp. 1–11, 2023, doi: 10.3389/fpubh.2023.1268223.

[18] M. H. Alsharif, A. H. Kelechi, K. Yahya, and S. A. Chaudhry, "Machine learning algorithms for smart data analysis in internet of things environment: Taxonomies and research trends," Symmetry (Basel), vol. 12, no. 1, 2020, doi: 10.3390/SYM12010088.

[19] J. Qiu, Q. Wu, G. Ding, Y. Xu, and S. Feng, "A survey of machine learning for big data processing," EURASIP J. Adv. Signal Process., vol. 2016, no. 1, 2016, doi: 10.1186/s13634-016-0355-x.

AG Publishing House

Advanced Research Techniques: Theories, Methods, And Practices (Volume-2)

Volume 2
Year: 2025

AG PH Books

Advancements in Data Collection Methods: Online Surveys, Interviews, and Beyond

Dr. Mintu Debnath [1*]

[1]*Assistant Professor, Chakdaha College, Chakdaha, Nadia, West Bengal.*

Abstract

After a well-designed study, data is the primary source of information required for studying a research topic. Data collection is essential because without the precise information obtained, research cannot be carried out. In this article review the various literature's study on advancements in data collection methods. It concluded that advancements in online survey and interview methods have revolutionized data collection, especially in the wake of the COVID-19 pandemic. The rise of digital tools and self-recorded video statements has enabled multimodal data capture, enhancing both qualitative and quantitative research. Online platforms offer cost-effective, wide-reaching, and convenient solutions for collecting accurate data. While some areas of survey research adopt innovations cautiously, others embrace rapid transformation driven by digital engagement and user preferences. As internet usage grows, online data collection continues to evolve, supporting dynamic research strategies and fostering a more agile, visually oriented, and globally connected research environment.

Keywords: Data collection method, Online survey, Qualitative methods, Interview, Quantitative methods, Focus groups, Observation, etc.

1 Introduction

The task of gathering data starts when the study topic has been determined and the research strategy or plan has been laid out. There are two primary methods for collecting data about circumstances, particular problems, or any other occurrence. Sometimes all that is needed is to retrieve the information that is already there [1]. The researcher must next decide which of the two data collection procedures will be

[*] ISBN No. - 978-93-49028-02-9

employed to determine the kind of data that will be used for the study. Data collection is, by definition, the process of acquiring, evaluating, and analysing precise understandings for study using accepted, established techniques. The information acquired may be used by a researcher to assess the hypothesis [2]. The most critical phase of the procedure is almost always the collection of data, irrespective of the studied subject. The methods of data collection for different fields of research vary depending on the critical information [3]. The primary objective of data collection is to safeguard "the information-rich and unquestioning data" that is collected for statistical analysis and to facilitate the development of data-driven research judgements [4].

A. Data collection

Data collection is the process of collecting and analysing data or information from a variety of sources to resolve "research issues, provide answers to queries, assess results, and predict trends and probability". This critical phase is essential for research, analysis, and decision-making in a variety of disciplines, such as "the social sciences, business, and healthcare". Researchers are required to ascertain the types of data, its sources, and the methodologies employed during the data collection process. The collection of data is essential in "the government, business, and research sectors" [5].

1. Quantitative Methods

Quantitative methods frequently employ statistical instruments for market research and demand forecasting. These methods anticipate demand by using previous data. These fundamental methods of gathering data are often used to create long-term estimates. Techniques for statistical analysis are particularly reliable since they are very objective [6].

Time Series Analysis: When a variable's values are arranged sequentially at equal intervals of time, it's called a time series. Using patterns, a business may predict how much demand there will be for its products and services over a certain period of time.

Smoothing Techniques: When there are no discernible patterns in the time series, smoothing methods may be used. In order to determine patterns and demand levels for estimating future demand, they remove random fluctuation from the previous demand. For demand forecasting smoothing, the most popular methods are "the weighted moving average and simple moving average approaches".

Barometric Method: Based on current developments, researchers use this method—also known as the leading indicators approach—to forecast future trends. Events from the past that are utilised to predict future events are known as leading indicators.

2. Qualitative Methods

Qualitative data is information that provides a description and explanation of an object. It can be visible, observed, and documented. This data type is not numerical in character. Among other techniques, focus groups, observations, and one-on-one interviews are used to collect this kind of data. Qualitative data in statistics is also referred to as categorical data [7]. This type of data is categorised according to the attributes and properties of an object or a phenomenon. Qualitative data collection is a potent instrument

for revealing the complex tapestry of human experiences. Through techniques like focus groups, interviews, and observations, researchers may chronicle the why behind behaviours and reveal motives, emotions, and narratives that data alone cannot convey. This method is indispensable for comprehending the tangible consequences of policies, programs, and products [8].

Interviews: One-on-one interviews are a commonly used data collecting strategy in qualitative research because they allow for the direct collection of highly personalised information from the subject. Interviews are significantly more advantageous for gathering data on sensitive subjects, as respondents are more likely to discuss "their beliefs, motivations, opinions, and experiences in a one-on-one setting than in a group setting" [9].

Focus Groups: While focus groups and unstructured interviews have many characteristics, the main difference is that the goal is to collect data from many participants at once. One of the most frequently used data collection instruments in qualitative research is focus groups, which are effective in gathering information based on collective perspectives. When doing a series of individual interviews is too lengthy or difficult to arrange, they are very helpful [10]. In the collection of data from a specific group of individuals, such as donors or clients from a particular demographic, focus groups are most beneficial. In an effort to ascertain the perspectives of the participants and the rationale behind them, the researcher should meticulously moderate and guide the discussion to a specific topic. Because participants are more likely to divulge when others are doing the same, group feedback often produces more thorough data than one-on-one interviews. Furthermore, the sharing of information by one participant may prompt an insight from another that would not have been apparent otherwise [11].

Observation: Qualitative research employs observation as one of its most effective data collection instruments for the acquisition of information through subjective methods. In addition to psychologists, sociologists, behaviour experts, and product developers, modern marketers also often use qualitative observation as a method. Information that is not readily quantifiable or measurable is the primary objective. The participants themselves provide virtually no cognitive input [12]. Researchers typically observe subjects and their responses as they conduct their daily activities and record comprehensive field notes from which to extract data. The extent of participant interaction during observational techniques varies [13]. Certain qualitative observations necessitate the researcher's complete immersion over an extended period. For example, participating in the same volunteer organisations, society meetings, church, or clinic as the participants. Researchers will almost certainly observe the most authentic responses in this scenario, as opposed to relying on behaviours that are elicited in a simulated environment. Based on the nature of the investigation and the intended objective, they may or may not elect to identify themselves as researchers during the process [14].

Open-Ended Surveys and Questionnaires: In the absence of an in-person meeting, organisations can gather the views and opinions of respondents through open-ended surveys and questionnaires. They can be transmitted electronically and are regarded as one of the most cost-effective qualitative data collection instruments [15]. Open-ended questions, in contrast to closed-question surveys and questionnaires, enable participants to provide extended and comprehensive responses that can be used to extrapolate

substantial quantities of data. There are no consistent responses, which can make it difficult to analyse the results of open-ended surveys and questionnaires. It is a common practice to categorise sentiments as positive, negative, or neutral and then analyse the data in greater detail [16].

2 Literature Review

(Lau & Bratby, 2024) [17] The video statements presented in this research are an alternative video method. Participants self-record their experiences while adhering to a guideline to capture multimodal (visual, auditory, and textual) data, which is both cost-effective and time-efficient. Consequently, it is feasible to accumulate video statements remotely. During our sample research, the methodology produces authentic impressions, offering a glimpse into a new organisational phenomenon. Participants employ the output as a source of data and as a foundation for subsequent discussions in order to enhance the significance of the video representations. In general, video statements provide a unique method of data acquisition that assists researchers in the provision of more comprehensive knowledge for business organisations and management.

(Wallwey & Kajfez, 2023) [6] Research designs that creatively use mixed methodologies may accept and support complex theoretical frameworks, like those often used in psychology research. This manuscript investigates a mixed-methods engineering education study that utilised artifact-based interviewing to enhance data collection. Visual representations of quantitative artefacts from the mixed methods research phase were implemented during subsequent interviews. Using quantitative artefacts as visual aids during interviews improved the quality of the data collected because they gave participants the language and visuals they needed to create data on the fly. This integration ensured that the theoretical underpinnings of the study were in alignment with the interview discussions, while simultaneously fostering an engaging interview environment.

(Jain, 2021) [18] Comparing surveys with in-person interviews as methods for gathering data in qualitative exploratory research is the aim of this work. Analysis was based on memos that were recorded throughout the data gathering process. A methodical three-step coding approach was used to examine the memoranda in order to determine the advantages and disadvantages of using each of the two data gathering instruments. For the chosen case study, the author contrasts the advantages and disadvantages of each approach using content analysis of the field notes and memoranda that were recorded throughout the research. For exploratory research, interviews are a helpful substitute for surveys when conducted methodically. Other research approaches and more data gathering instruments might be compared in this study's expansion.

(Braun et al., 2020) [19] Qualitative researchers may benefit greatly from "surveys that prioritise qualitative research values" and use the vast potential of qualitative data, especially with the advent of online delivery methods. Nevertheless, the method is not widely employed, and there is still a lack of methodological discourse regarding qualitative surveys. The underutilisation and little methodological debate may be explained by the prevalence of interviews in "qualitative research" and (erroneous) presumptions about the depth of qualitative survey data. Through an analysis of our interaction with the

creation of online surveys as a resource for qualitative research, we aim to challenge preconceived notions about qualitative surveys and to demonstrate that they are a flexible, exciting method with a broad range of applications and benefits for participants as well as researchers.

(Brühlmann et al., 2020) [20] However, crowdsourced samples are becoming more prevalent in a variety of academic disciplines, despite recent apprehensions regarding data quality. The objective of this study was to employ a variety of detection methods and measures to comprehensively evaluate negligence in a crowdsourced sample (N = 394). An analysis of the latent profile indicated that 45.9% of the participants exhibited some form of negligent behaviour. In an experiment that was included in the survey, the effect size was enhanced by excluding these participants. In accordance with our discoveries, we have provided a number of straightforward recommendations for evaluating data quality.

(Utibe Monday, 2020) [21] Investigated, in light of the researcher's fieldwork experience, the consequences of using interviews as a data gathering technique in the social sciences. The study came to the conclusion that although interviews are a useful tool for learning about interviewees' perspectives, they may also complement other techniques that provide deeper insights into participants' core values and beliefs. For example, by supplementing interviews with personal observation, researchers might examine participants' internal beliefs and exterior behaviours. Thus, the report emphasised that, although it depended on the research topics, adopting several data collecting instruments would assist get richer data and validate the research results.

(Mergel et al., 2019) [22] In order to improve public service delivery, boost efficiency and effectiveness in their designs, and achieve objectives such as "greater transparency, interoperability, or citizen satisfaction", governments are modifying their working methods in response to the influence of supranational agreements and the evolution of expectations. In contrast to the accessibility of consulting reports, there is a dearth of systematic comprehension regarding the definition of digital transformation in the daily operations of public administrators, their approach to digital transformation projects, and the anticipated outcomes. Describe digital transformation using an empirically supported definition derived from expert interviews. Describe the rationale, steps, and expected outcomes of the public sector's digital transformation in a conceptual framework.

(Couper, 2017) [23] Explores the most recent technological and methodological advancements in the acquisition of survey data. For the purpose of expanding and improving the survey instrument, consider how the survey profession has responded to these developments and obstacles. Look at the shift from "telephone surveys to self-administered (mail and/or Web) modalities", the rise in address-based sampling that followed, and the fall in random digit dialling over time. Analyse the increase in nonprobability sampling methods, particularly those that are linked to online data collection. Additionally, I evaluate surveys that are frequently referred to as "big data alternatives." At last, examine a variety of contemporary technological and methodological developments that have been implemented to transform the survey methodology. Conclude that surveys continue to be a reliable and adaptable approach to the collection of data on, and the formulation of inferences about, populations, despite the numerous significant obstacles they encounter.

3 Conclusion

In conclusion, the advancement of online survey and interview methods has significantly transformed data collection in the digital era. The COVID-19 pandemic accelerated this transformation, pushing researchers to adopt innovative, contactless approaches. One such evolution is the use of video statements—self-recorded, multimodal data capturing tools—which cater to a visually engaged generation and offer rich, nuanced insights beyond traditional text-based responses. While some areas of survey research remain cautious in adopting change, competitive sectors like political polling and market research have embraced digital tools more readily. Online surveys, enabled through SMS, email, websites, and social media platforms, offer vast reach, cost-efficiency, and convenience, making them ideal for both qualitative and quantitative research. These methods not only improve response rates but also provide real-time data for faster decision-making. As digital transformation continues to influence public and private sectors, researchers must align with emerging technologies and preferences to remain effective. Overall, the integration of video-based self-reporting and online data collection has reshaped the survey landscape, making it more adaptable, inclusive, and efficient for modern research needs.

References

[1]. M. P. Couper and K. A. Mcgonagle, “Recent Developments in Web-Based Data Collection for Longitudinal Studies,” no. Ses 1623864, 2019, [Online]. Available: https://micda.psc.isr.umich.edu/

[2]. K. Xu et al., “Advanced Data Collection and Analysis in Data-Driven Manufacturing Process,” Chinese J. Mech. Eng. (English Ed., vol. 33, no. 1, 2020, doi: 10.1186/s10033-020-00459-x.

[3]. S. Harrison, F. Alderdice, and M. A. Quigley, “Impact of sampling and data collection methods on maternity survey response: a randomised controlled trial of paper and push-to-web surveys and a concurrent social media survey,” BMC Med. Res. Methodol., vol. 23, no. 1, pp. 1–12, 2023, doi: 10.1186/s12874-023-01833-8.

[4]. Syeda Ayeman Mazhar, R. Anjum, A. I. Anwar, and A. A. Khan, “Methods of Data Collection: A Fundamental Tool of Research,” J. Integr. Community Heal, vol. 10, no. 01, pp. 6–10, 2021, doi: 10.24321/2319.9113.202101.

[5]. H. Taherdoost, “Data Collection Methods and Tools for Research; A Step-by-Step Guide to Choose Data Collection Technique for Academic and Business Research Projects,” Int. J. Acad. Res. Manag., no. 1, pp. 10–38, 2021, [Online]. Available: https://hal.science/hal-03741847

[6]. C. Wallwey and R. L. Kajfez, “Quantitative research artifacts as qualitative data collection techniques in a mixed methods research study,” Methods Psychol., vol. 8, no. August 2022, p. 100115, 2023, doi: 10.1016/j.metip.2023.100115.

[7]. J. Zahle, “Reactivity and good data in qualitative data collection,” Eur. J. Philos. Sci., vol. 13, no. 1, pp. 1–18, 2023, doi: 10.1007/s13194-023-00514-z.

[8]. N. Saxena, “A systematic Survey of Natural Language Processing (NLP) Models & its Application,” pp. 25–31, 2022.

[9]. E. Gilbert, “New technologies and innovative methods in data collection: scoping review,” 2020, [Online]. Available: www.cls.ucl.ac.uk.

[10]. J. Sutton and Z. Austin, “Qualitative reserch: data collection,analysis,and managment,” Can. J. Hosp. Pharm., vol. 68, no. 3, pp. 226–231, 2015.

[11]. J. Zhang, D. Wolfram, and F. Ma, “The impact of big data on research methods in information science,” Data Inf. Manag., vol. 7, no. 2, p. 100038, 2023, doi: 10.1016/j.dim.2023.100038.

[12]. N. Van Quaquebeke, M. Salem, M. van Dijke, and R. Wenzel, Conducting organizational survey and experimental research online: From convenient to ambitious in study designs, recruiting, and data quality, vol. 12, no. 3. 2022. doi: 10.1177/20413866221097571.

[13]. J. Sanchez, A. Trofholz, and J. M. Berge, “Best Practices and Recommendations for Research Using Virtual Real-Time Data Collection: Protocol for Virtual Data Collection Studies,” JMIR Res. Protoc., vol. 13, pp. 1–8, 2024, doi: 10.2196/53790.

[14]. A. B. Hamilton and E. P. Finley, “Qualitative Methods in Implementation Research: An Introduction,” Physiol. Behav., vol. 176, no. 3, pp. 139–148, 2019, doi: 10.1002/hep.30150.Ductular.

[15]. T. H. Khan and E. MacEachen, “An Alternative Method of Interviewing: Critical Reflections on Videoconference Interviews for Qualitative Data Collection,” Int. J. Qual. Methods, vol. 21, pp. 1–12, 2022, doi: 10.1177/16094069221090063.

[16]. S. L. Thomas, H. Pitt, S. McCarthy, G. Arnot, and M. Hennessy, “Methodological and practical guidance for designing and conducting online qualitative surveys in public health,” Health Promot. Int., vol. 39, no. 3, 2024, doi: 10.1093/heapro/daae061.

[17]. A. Lau and M. Bratby, “Collecting qualitative data via video statements in the digital era,” Labour Ind., vol. 34, no. 2, pp. 101–113, 2024, doi: 10.1080/10301763.2023.2209923.

[18]. N. Jain, “Survey versus interviews: Comparing data collection tools for exploratory research,” Qual. Rep., vol. 26, no. 2, pp. 541–554, 2021, doi: 10.46743/2160-3715/2021.4492.

[19]. V. Braun, V. Clarke, E. Boulton, L. Davey, and C. McEvoy, “The online survey as a qualitative research tool,” Int. J. Soc. Res. Methodol., vol. 24, no. 6, pp. 641–654, 2020, doi: 10.1080/13645579.2020.1805550.

[20]. F. Brühlmann, S. Petralito, L. F. Aeschbach, and K. Opwis, “The quality of data collected online: An investigation of careless responding in a crowdsourced sample,” Methods Psychol., vol. 2, no. March, p. 100022, 2020, doi: 10.1016/j.metip.2020.100022.

[21]. T. Utibe Monday, “Impacts of Interview as Research Instrument of Data Collection in Social Sciences,” J. Digit. Art Humanit., vol. 1, no. 1, pp. 15–24, 2020, doi: 10.33847/2712-8148.1.1_2.

[22]. I. Mergel, N. Edelmann, and N. Haug, “Defining digital transformation: Results from expert interviews,” Gov. Inf. Q., vol. 36, no. 4, p. 101385, 2019, doi: 10.1016/j.giq.2019.06.002.

[23]. M. P. Couper, “New developments in survey data collection,” Annu. Rev. Sociol., vol. 43, pp. 121–145, 2017, doi: 10.1146/annurev-soc-060116-053613.

AG Publishing House

Advanced Research Techniques: Theories, Methods, And Practices (Volume-2)

Volume 2
Year: 2025

AG PH Books

Ensuring Research Integrity: Plagiarism, Fabrication, and Falsification

Dr. Jassim Rahiman K [1*]

[1]*B.A.M.S, M.D(Ay.), Assistant Professor, Dept. of Kayachikitsa, Apex Institute of Ayurvedic Medicine and Hospital, Samaspur, Chunar, Mirzapur, U.P.*

Abstract

In the corporate world, investors and founders place a high value on "financial stability, and business intelligence (BI)" is a key instrument for greatly boosting this stability. The main goal of financial management is to raise money for the business at a minimal cost and use that money to generate enormous profits. In this article review the various literature's study on business intelligence tools in corporate financial management. It concluded that the evolution of business intelligence (BI) tools has revolutionized corporate financial management by enhancing decision-making, optimizing operations, and mitigating risks. These tools empower organizations with advanced data analytics, enabling strategic financial planning and resource efficiency. By identifying and forecasting financial risks, BI ensures stability and sustainability in business practices. Finance departments leverage BI to correlate financial outcomes with business activities, aligning strategies with corporate objectives. Technologies like OLAP and data mining facilitate intelligent decision-making in complex financial environments. As businesses become increasingly data-driven, BI tools serve as essential assets, fostering growth, competitiveness, and long-term success in an ever-evolving financial landscape.

Keywords: Research (scientific) integrity, Plagiarism, Fabrication, Falsification, Research (scientific) misconduct, P-hacking etc.

[*] ISBN No. - 978-93-49028-02-9

Dr. Jassim Rahiman K

1 Introduction

One cannot undervalue the significance of Research Integrity (RI). RI is thought of as a mechanism that maintains public trust in scientists and research while simultaneously protecting the professional careers and reputations of researchers. RI is important for promoting social and economic progress [1]. Researchers, people, research funders, and government officials in charge of scientific policy all gain from it and trust is fostered. At the core of the research process is RI. It provides the basis for the public's confidence in research findings, evidence, and scientists as well as enabling scientists to trust one another and the data. The public and patients are shielded by RI from the detrimental effects of inaccurate and misleading data [2].

Finding scientific truth requires a lot of creativity, critical thinking, honesty, integrity, candour, perseverance, and persistence. These are the cornerstones that need to be used in every scientific undertaking in order to preserve and uphold the integrity of scientific literature for the advancement of humanity. Several months of meticulous project preparation and execution culminate in the publishing of a scholarly article. In the interest of science, the task must be conducted impartially, honestly, and objectively, and the findings must be communicated honestly. However, there are situations when deliberate deceptions or ignorance cause departures from the ideal. Scientific misconduct is the term used to describe these intentional or unintentional departures from the ideal [3].

Research misconduct has existed since the early days of art and science's development, but it has recently increased sharply due to academic competition, the desire for quick success through shortcuts, and the desire for publications to boost one's reputation. Perhaps the United States Public Health Service has the most thorough and legally sound description of research misconduct: plagiarism, fabrication, or falsification in the proposed, carried out, or evaluated study, or in the reporting of research findings. Disagreements or honest mistakes are not considered research misconduct. Science advances because of scientific literature. On the other hand, research misconduct undermines it and has a negative impact on the reliability, validity, and relevance of scientific results [4].

A. Research integrity

As a dynamic and wide concept, research integrity may be defined as the ethical, robust, and transparent conduct of the research process when proposing, carrying out, assessing, and reporting research results. It entails adherence to generally recognised professional standards and conventions as well as laws, rules, and guidelines. Robustness, honesty, openness, respect, and accountability are the fundamental tenets of research integrity [5]. Both individual researchers and the whole research community—including academic institutions, funding organisations, regulatory agencies, and scientific journals—are impacted by research integrity. Research integrity, which is part of the larger idea of responsible conduct of research (RCR), is governed by values and principles. RCR acts as a useful framework that converts the broad ethical and research integrity principles into doable rules, encouraging moral conduct and judgement in researchers' daily work [6].

B. Research misconduct

Research misconduct is defined as behaviours or dubious research methods that do not meet the norms of scholarship, research, and ethics necessary to maintain the integrity of study. It wastes money, compromises the scientific record, harms people and the environment, and erodes the credibility of research. The term "falsification, fabrication, and plagiarism" is often used to describe it [7]. Examples include fabricating data or findings, giving false credit to authors, manipulating research tools, materials, or procedures, or altering or deleting data, graphs, pictures, or results. Poor research design, misdemeanours, and other harmful activities are all part of the much larger category of questionable research techniques. Often referred to as "sloppy science," some of these may be used without considering the possible repercussions for the integrity of the study rather than in an effort to deceive. The main issues in avoiding research misconduct are falsification, fabrication, and plagiarism (FFP), which are referred to as the three "cardinal sins" of research conduct. The integrity of research for that person, lab, university/company, and the field at large is compromised by any departure from these standards [8].

C. Elements of research misconduct

Fabrications: Fabrication is the practice of creating data, such as reporting on patients or studies that never happened. Falsification may be mistaken for fabrication in many situations. False data or outcomes are recorded or reported, which is known as fabrication. The act of producing data without doing necessary investigation is known as fabrication. If research results have not been conducted honestly, they should not be discussed, shared, or published [9], [10].

Falsification: Falsification may be described as the alteration or omission of data or findings, or the manipulation of research materials, tools, or procedures such that the study is not correctly documented in the research record. Falsification is the act of altering an experiment's design or outcomes in a way that is not supported by science. In most cases, to improve or eliminate findings that don't fit the hypothesis. Falsification is regarded as serious research misconduct, much like fabrication and plagiarism. It is defined as altering, deleting, or concealing data or findings without cause, as well as modifying research materials, tools, or procedures, according to the European Code of Conduct [9], [11].

Plagiarism: The act of presenting someone else's words, ideas, or work as one's own without proper attribution or permission is known as plagiarism. Academia and other disciplines that depend on the integrity of intellectual property regard it as a severe ethical violation [9].

D. Scientific misconduct's impacts

Numerous stakeholders and the scientific community at large are greatly impacted by scientific misconduct [1]. The following are some major effects of scientific misconduct:

Damage to scientific integrity: Scientific research is compromised by scientific misconduct, which undermines its integrity and credibility. It weakens the scientific method and the foundation of knowledge, which are underpinned by the precise reporting of data and findings. Instances of misconduct

result in a decline in confidence in scientific discoveries, as it undermines the reliability of research findings [12].

Harm to public trust: The public's confidence in science and scientists may be damaged by scientific misconduct. The public's view, policy choices, and financing for scientific endeavours may all suffer when wrongdoing is revealed or suspected since it can breed scepticism towards scientific research. For scientific information to be accepted and used to solve social issues, public trust must be maintained [13].

Misdirection of resources: Resources may be misallocated as a consequence of misconduct in scientific research. Other researchers may devote time, energy, and resources to following lines of inquiry based on inaccurate or manufactured data when study results that are deceptive or untrue are publicised. This may impede scientific advancement and squander money that might have been used for more productive research directions [14].

2 Literature Review

(Armond et al., 2024) [15] Any violations of research integrity in the biomedical sciences have the potential to have a cascading effect on patient care, medical treatments, and the wider application of healthcare legislation. In order to address these breaches, steps including strict research procedures, open reporting, and a shift in the research culture are needed. Fostering a culture of research integrity requires institutional support in the form of mentoring, thorough training, and explicit rules. However, research behaviour is significantly influenced by institutional and structural elements, such as recognition programs and research incentives. In order to preserve public confidence in the scientific community and guarantee the validity of knowledge, preserving research integrity requires a team effort from all parties involved. Here, we go over a few definitions and guiding concepts, their significance for the biomedical sciences, and practical measures that may be taken to promote research integrity.

(Alam, 2024) [16] The academic publisher's role in screening, protecting, and examining information before and after publication has been solidifying into a fundamental publisher skill set in an age of growing concerns about research integrity that are changing in scope and form. A crucial step in resolving the issues is investing in tools and training, but improved cooperation between players from different industries is also required. In order to fully understand post-publication notices (such as retractions) and make informed decisions about which published content should be used as a foundation for future research as well as the advancement of guidelines and policies, "consumers" of published scholarly content must also be made aware of the trends in research integrity and publishing ethics misconduct.

(Chen et al., 2024) [17] emphasises the necessity for the academic community to enhance researcher qualifications, strengthen ethical norms, and establish rigorous review mechanisms in order to resolve these challenges. Mandatory AI ethics and integrity training for researchers is being implemented with the objective of promoting ethical research practices and nurturing a comprehensive understanding of potential AI misuses. Development of unified ethical standards for AI in research and the exchange of

best practices through the establishment of international collaboration frameworks. These recommendations are urgently required for the scientific community to consider and take action, as protecting research integrity is essential for preserving public trust in science.

(Francesca et al., 2024) [18] a study to evaluate the efficacy of a training course on RI by mapping the attitudes of early-career researchers on this topic through a questionnaire that was constructed using the revised version of the Scientific Misconduct Questionnaire and administered to all participants at the commencement and conclusion of the course. There is consensus among early-career researchers regarding the significance of sharing any ethical concerns that may arise in research with their colleagues and superiors, as well as establishing a work environment that promotes RI awareness. Overall, the findings indicate that the course is effective. Early-career researchers should be provided with research methodology, integrity, and ethical consultation services, in addition to RI training, by their institutions. In order to cultivate effective practices that are consistent with the principles of RI, senior scientists should encourage engagement in peer-to-peer dialogue and incorporate RI into their research practices.

(Khot et al., 2024) [19] According to a study of scientists who get funding from the National Institutes of Health (NIH), there are at least 2325 instances of scientific misconduct annually. Therefore, this study was conducted to ascertain the knowledge, attitude, and practices (KAPs) of scientific misconduct and research integrity among postgraduate students and teachers employed by medical institutions in Central India (CI) and North Karnataka (NK). This study emphasises how academic culture shapes ethical research practices and the need for better education and policy implementation to maintain research integrity in medical institutions.

(Nguyen & Tuamsuk, 2024) [20] Investigate the factors that influence the scientific integrity of scholarly publishing among researchers in Thailand. The results confirm that there are five factors that influence scientific integrity in scholarly publishing: university/faculty requirements/policies, university support, publishers' aspects, aspects of publishing, and researcher-related factors. It is clearly evident that the highest factor loading values are achieved by establishing a process and criteria for authorship agreement and verification, as well as dedicating offices or divisions within the university to address research integrity concerns and provide support to researchers. The findings guarantee that researchers affiliated with prestigious research universities in Thailand prioritise scientific integrity. It is advised that universities prioritise initiatives that are designed to cultivate scientific integrity and increase the ethical awareness of researchers in order to address this challenge.

(Rao et al., 2024) [21] p-hacking, which involves the misuse of another individual's research results or ideas without appropriate attribution, is a potential activity that researchers may engage in. In the context of research, conflict of interest (COI) arises when an individual's personal, financial, or professional interests have the potential to influence their judgement or actions. The propagation of deceptive or misleading information can result from hypotheses that are made after the results are evident. Researchers should be forthcoming with their methodologies and their discoveries should be reported with precision and honesty. Clear and rigorous policies regarding scientific misconduct ought to be implemented by research institutions. This knowledge must be disseminated to ensure that researchers

and readers comprehend which statistical analysis and reporting methods constitute scientific misconduct.

(Skarbek, 2024) [22] In the context of the role that legislators can play in combating this practice, the article addresses the issue of fabrication and falsification of scientific research. The article presents the findings of a survey that polled 70 Polish academicians regarding their perspectives on the role of legislators in preventing the fabrication and falsification of research. For instance, the results indicate that the majority of respondents (81%) advocate for disciplinary accountability for malfeasance of this nature. Additionally, they believe that legislators should combat this practice through "soft" measures, such as the promotion of research ethics codes (70%). The criminalisation of such conduct was desired by 50% of the respondents. The Polish legislator is not effective in combating such malfeasance, according to nearly half of the respondents.

(Elali & Rachid, 2023) [23] One's reputation is damaged and honest writers are undermined when research is falsified within the scientific community. Show that utilising a chatbot with an AI-based language model to fabricate research is feasible. We will evaluate the accuracy of detecting manufactured works using AI and human detection. There will be an emphasis on the dangers of using research produced by artificial intelligence as well as the reasons why research may be fabricated.

(Zhaksylyk et al., 2023) [24] Adherence to the most stringent ethical standards is mandatory for researchers. Research institutions are responsible for fostering an environment that promotes integrity principles and offers researchers valuable guidance, instruction, and support. The dissemination of research results through publishing is facilitated by editors and reviewers, who serve as protectors by maintaining ethical standards and standards of quality. The ongoing and multifaceted battle against scientific malfeasance is a continuous activity. It necessitates a collaborative endeavour and upholding the principles of rigorous science, transparency, and honesty. The scientific community may protect its fundamental principles and continue to make a meaningful contribution to the welfare of society by fostering a culture of RI. It serves as a foundation for future scientific advancements and augments current research.

3 Conclusion

In conclusion, ensuring research integrity in the digital age requires a proactive, multifaceted approach to address growing concerns such as plagiarism, fabrication, and falsification—now further amplified by AI technologies. The ease of producing fabricated content, lack of robust AI-detection tools, and academic hyper-competitiveness create opportunities for research misconduct, potentially influencing critical domains like healthcare policy and scientific standards. Motivations range from career advancement to financial gain, making it imperative for all stakeholders—researchers, institutions, funders, and publishers—to uphold ethical responsibilities. Misconduct, including data falsification, p-hacking, and failure to disclose conflicts of interest, can lead to false scientific conclusions, damaged reputations, and resource waste. Strengthening ethical standards through transparent research practices, continuous education, CME workshops, and responsible conduct training is essential. Institutions must

enforce strict policies, peer-review protocols, and foster collaboration among stakeholders to build a culture of trust and accountability. As AI becomes more integrated into research workflows, enhancing technical capabilities and investigative skills, while promoting open dialogue and resource sharing, is vital to safeguarding the credibility and reliability of scholarly communication. Only through collective vigilance and commitment can we ensure that published research remains a trusted source of knowledge and progress.

References

[1]. K. K. Lal and B. Sharma, "Research Integrity & Ethics: Scientific Misconduct," Academic, no. September, 2023.

[2]. K. R, "Research and publication-integrity and ethics," J. Dent. Res. Rev., vol. 212, no. 2, pp. 383–385, 2014, doi: 10.1016/j.atherosclerosis.2010.01.050.

[3]. M. A. Deshmukh, A. S. Dodamani, M. R. Khairnar, and R. G. Naik, "Research misconduct: A neglected plague," Indian J. Public Health, vol. 61, no. 1, pp. 33–36, 2017, doi: 10.4103/0019-557X.200255.

[4]. M. Khadem-Rezaiyan and M. Dadgarmoghaddam, "Research misconduct: A report from a developing country," Iran. J. Public Health, vol. 46, no. 10, pp. 1374–1378, 2017.

[5]. S. A. Z. Zaidi, E. Ahmad, and N. Shukla, "Ethical Considerations in the Use of Artificial Intelligence (AI) for Education and Research : A Review," Int. J. Innov. Sci. Eng. Manag., pp. 156–167, 2024, doi: 10.69968/ijisem.2024v3si2156-167.

[6]. A. Mousa et al., "Research Integrity in Guidelines and evIDence synthesis (RIGID): a framework for assessing research integrity in guideline development and evidence synthesis," eClinicalMedicine, vol. 74, p. 102717, 2024, doi: 10.1016/j.eclinm.2024.102717.

[7]. R. A. Feenstra, E. Delgado López-Cózar, and D. Pallarés-Domínguez, "Research Misconduct in the Fields of Ethics and Philosophy: Researchers' Perceptions in Spain," Sci. Eng. Ethics, vol. 27, no. 1, pp. 1–21, 2021, doi: 10.1007/s11948-021-00278-w.

[8]. A. Muthanna, Y. Chaaban, and S. Qadhi, "A model of the interrelationship between research ethics and research integrity," Int. J. Qual. Stud. Health Well-being, vol. 19, no. 1, 2024, doi: 10.1080/17482631.2023.2295151.

[9]. A. A. Sanni and S. M. Enefu, "Research Misconduct (fabrication, falsification, plagiarism etc.)/ Latest Bibliography and Proper in Bibliography in research work Practicum or workshop if any," Res. gate, no. July 2020, 2019.

[10]. C. Chou, I. J. Lee, and J. Fudano, "The present situation of and challenges in research ethics and integrity promotion: Experiences in East Asia," Account. Res., vol. 31, no. 6, pp. 576–599, 2024, doi: 10.1080/08989621.2022.2155144.

[11]. L. M. Bouter, J. Tijdink, N. Axelsen, B. C. Martinson, and G. ter Riet, "Ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity," Res. Integr. Peer Rev., vol. 1, no. 1, pp. 1–8, 2016, doi: 10.1186/s41073-016-0024-5.

[12]. R. Roje, A. Reyes Elizondo, W. Kaltenbrunner, I. Buljan, and A. Marušić, “Factors influencing the promotion and implementation of research integrity in research performing and research funding organizations: A scoping review,” Account. Res., vol. 30, no. 8, pp. 633–671, 2023, doi: 10.1080/08989621.2022.2073819.

[13]. N. Allum et al., “Researchers on research integrity: a survey of European and American researchers,” F1000Research, vol. 12, p. 187, 2023, doi: 10.12688/f1000research.128733.1.

[14]. Y. Xie, K. Wang, and Y. Kong, “Prevalence of Research Misconduct and Questionable Research Practices: A Systematic Review and Meta-Analysis,” Sci. Eng. Ethics, vol. 27, no. 4, p. 41, 2021, doi: 10.1007/s11948-021-00314-9.

[15]. A. C. V. Armond, K. D. Cobey, and D. Moher, “Key concepts in clinical epidemiology: research integrity definitions and challenges,” J. Clin. Epidemiol, vol. 171, p. 111367, 2024, doi: 10.1016/j.jclinepi.2024.111367.

[16]. S. Alam, “Trends in research integrity concerns and the evolving role of the publisher,” Insights UKSG J., vol. 37, 2024, doi: 10.1629/uksg.663.

[17]. Z. Chen et al., “Research integrity in the era of artificial intelligence: Challenges and responses,” Med. (United States), vol. 103, no. 27, p. e38811, 2024, doi: 10.1097/MD.0000000000038811.

[18]. G. Francesca, S. Ceruti, S. Martini, M. Picozzi, M. Cosentino, and F. Marino, “Educating and Training in Research Integrity (RI): A Study on the Perceptions and Experiences of Early Career Researchers Attending an Institutional RI Course,” J. Acad. Ethics, vol. 22, no. 3, pp. 413–430, 2024, doi: 10.1007/s10805-023-09497-1.

[19]. A. Khot, C. A. Chindhalore, and A. Naikwadi, “Knowledge, Attitude, and Practices About Research Integrity and Scientific Misconduct Among the Faculty and Medical Postgraduates Working in Medical Colleges in North Karnataka and Central India: A Cross-Sectional Online Survey,” Cureus, vol. 16, no. 4, 2024, doi: 10.7759/cureus.59200.

[20]. L. T. Nguyen and K. Tuamsuk, “Factors influencing researchers’ scientific integrity in scholarly publishing: a path analysis approach,” Res. Ethics, 2024, doi: 10.1177/17470161241257445.

[21]. K. N. Rao, M. Mair, R. D. Arora, P. Dange, and N. M. Nagarkar, “Misconducts in research and methods to uphold research integrity,” Indian J. Cancer, vol. 55, no. 1, pp. 382–389, 2024, doi: 10.4103/ijc.IJC.

[22]. O. Skarbek, “Fabrication and Falsification of Scientific Research – a Challenge for Legislators or an Irrelevant Issue?” Kryt. Prawa, vol. 16, no. 1, pp. 200–212, 2024, doi: 10.7206/kp.2080-1084.666.

[23]. F. R. Elali and L. N. Rachid, “AI-generated research paper fabrication and plagiarism in the scientific community,” Patterns, vol. 4, no. 3, p. 100706, 2023, doi: 10.1016/j.patter.2023.100706.

[24]. A. Zhaksylyk, O. Zimba, M. Yessirkepov, and B. F. Kocyigit, “Research Integrity: Where We Are and Where We Are Heading,” J. Korean Med. Sci., vol. 38, no. 47, pp. 1–11, 2023, doi: 10.3346/jkms.2023.38.e405.

AG Publishing House

Advanced Research Techniques: Theories, Methods, And Practices (Volume-2)

Volume 2
Year: 2025

AG PH Books

Big Data and Its Impact on Research Methodology

Dr. Pilli Lalitha Kumari[1*], Dr. Prasanna Kumar Yekula[2]

[1]*Associate Professor, Department of Computer Science and Engineering, Visakha Institute of Engineering & Technology, 88th Division, Narava Visakhapatnam.*

[2]*Professor, School of Mining Engineering, Faculty of Engineering, PNG University of Technology Private Mail Bag Lae 411 Morobe Province, Papua New Guinea.*

Abstract

The integration of information systems research and social sciences enables a new era of innovation. In this context, the social sciences are provided with new insights through the implementation of Big Data and analytics technology. It also provides innovative, adaptable solutions to critical social issues and obstacles. This article offers a thorough analysis of the wide range of literature on big data and how it affects research technique. It concluded that big data is reshaping research methodology by enabling access to vast datasets, novel analytical techniques, and dynamic data visualization tools. It fosters deeper understanding across disciplines, offering transformative potential in areas like customer behavior and sport management. However, effective application requires context-aware methods, high-quality data, and rigorous preprocessing. Conceptual frameworks assessing organizational readiness for big data adoption further enhance decision-making capabilities. As research evolves, big data's value lies in its ability to generate responsive, socially aware insights. With appropriate methodologies and a focus on data integrity, big data continues to revolutionize research and knowledge creation across domains.

Keywords: Big data, Research methodology, Social science research, Quantitative research, Qualitative research, Social implications, etc.

1 Introduction

Despite its extraordinary digital ubiquity, the study of big data is motivated by the epistemic belief that vast data sets provide higher kinds of intelligence and erudition. The reason for this is the vast array of natural occurrences and human actions that are digitally recorded. Common features of big data include

[*] ISBN No. - 978-93-49028-02-9

the well-known "five Vs: volume, velocity, variety, veracity, and value" [1]. These include the enormous volume of data that is produced and preserved, the rapid growth of data, the variety of data kinds and formats, the quality of data collection, and the usefulness of data extraction, in that order. These data properties have created previously unheard-of possibilities and problems for information science data usage and analysis techniques. Sophisticated research methods are required for big data [2].

The following are the fundamental components of any empirical research study: “research objectives/aims, research methods, data, and their analysis”. They are inextricably linked. In order to accomplish the goals of a study, research methods are the strategies, tactics, and processes used to gather, process, analyse, produce, and draw conclusions. A dataset's availability, nature, and size may have a big impact on the choice of research methodologies and even study subjects [3]. Analysis and processing methodologies that are appropriate for a large dataset are necessary, as they are not self-evident. A dataset that is broad, diverse in type, and content-rich would broaden the scope of research topic selections. To the study themes, academics would also be able to use a variety of research approaches [4]. The growing big data trend has had a substantial impact on information science data-driven empirical research projects. A research study must be carried out in order to do an in-depth examination and discussion of the trend's impact on information science research methodologies [5]. The study's importance is multifaceted. It encourages the information science research community to have a productive conversation and reflect on the challenges and possibilities posed by the big data movement. Additionally, it supports academics in the area in creating effective research designs for studies focused on big data and helps information practitioners overcome the difficulties they face in the big data age [6].

A. Big data

The term "big data" refers mostly to data volumes that are either very vast or complex for traditional data processing technologies to handle. Data with a greater number of entries (rows) may provide greater statistical power, while data with a higher complexity (more attributes or columns) may result in a higher false discovery rate [7]. Traditionally used data processing tools, such as spreadsheets, are incapable of managing or analysing extremely large and intricate data sets, which are referred to as "big data". Big data includes unstructured data like social media postings or videos, structured data like inventory databases or financial transaction records, and hybrid data sets like those used to train sophisticated artificial intelligence language models. The sources of these data sets may range from Shakespeare's plays to the financial spreadsheets of a company over the past decade [8].

Big data has only grown as a result of recent technological developments that have significantly reduced the cost of processing and storage, making it easier and less expensive than ever to store more data. The growing amount of data allows businesses to utilise it to make more accurate and precise business decisions. Though this is a different benefit, there is more to making the most of big data than just evaluating it. "Astute analysts, business users, and executives" are necessary for the whole discovery process because they must be able to recognise patterns, pose pertinent queries, create knowledgeable hypotheses, and predict behaviour [9].

B. The Vs of big data

Big data is often defined by volume, velocity, and variety, however definitions may vary greatly. The term "3 Vs of big data" is often used to describe these features.

Volume: As the name suggests, the most common attribute linked to "big data" is its volume. This phrase describes the enormous amount of data that is constantly produced and accessible for gathering from a wide range of devices and sources.

Velocity: The rate at which data is generated is referred to as "big data velocity". Since data is now often created in real-time or very real-time, it needs to be processed, retrieved, and evaluated at the same rate to make a meaningful difference.

Variety: There are many different sources of data, and it might be semi-structured, unstructured, or structured. Thus, it is heterogeneous. Currently, traditional structured data (e.g., spreadsheets or relational databases) is supplemented by "unstructured text, images, audio, video files, or semi-structured formats, such as sensor data", that are unable to be organised in a fixed data schema.

In addition to the three basic Vs, three additional variables--value, variability, and veracity--are frequently mentioned in the context of utilising potential of big data.

Veracity: It can be difficult to manage the quality and accuracy of big data because of "its chaotic, congested, and error-prone design". Larger datasets may be vexing and burdensome, while smaller datasets may offer an incomplete perspective. The data is more reliable when its veracity is higher.

Variability: The concept of data that has been collected is subject to constant evolution, which may result in inconsistency over time. In addition to data acquisition methods that are contingent upon the information that organisations intend to capture and analyse, context and interpretation are also altered.

Value: Determining the commercial value of the data you accumulate is indispensable. Big data has to be loaded with the right information and then thoroughly examined in order to provide insights that may aid in decision-making.

C. Social implications

There are numerous implications for society inherent in big data. Companies, industries, and nations employ them to enhance their "decision-making and effectiveness". Numerous industries heavily depend on data-driven decision-making, such as "banking and other financial services, insurance, telecommunications, information technology (IT), healthcare, automobile, oil, energy, and utilities, among others" [10].

Big data has also been used to predict weather and market trends, prevent catastrophes, deter fraud, increase profits, and track the environment to find the latest market trends. Furthermore, big data contributes to national security and commercial transactions [11]. Applications that use big data include customer relationship management (CRM) and enterprise resource planning (ERP), which are used for "deal monitoring, payroll, the sales pipeline, inventories, contacts, payables, authorising requests, and

more". In addition, big data applications encompass web applications that are employed for "advertising, collaboration, e-commerce, digital marketing, recommendation, weblogs, and mobile computing". Big data is also utilised in IoT applications, which include "log files, sensors, devices, radio-frequency identification, text, images, weather, audio, video, spatial and GPS coordinates, eGov feeds, clickstreams, social sentiments, and data market feeds" [12], [13].

There is also a substantial impact of big data on management processes, including distributed process planning and supply management. In the manufacturing sector, they are also employed to enhance operational performance, manage production safety, and conduct mechanical assembly [14]. Big data are also employed to fuel design innovation that is motivated by consumer insight, as it has an impact on business-to-business processes, advertising, and marketing initiatives. Big data has facilitated the monitoring and evaluation of customer satisfaction by organisations, thereby improving customer service in the finance and banking sector [15], [16].

2 Literature Review

(Kgakatsi et al., 2024) [17] The purpose of this assessment is to examine the influence of BD on SMEs, with a particular emphasis on revenue growth, economic performance, and business enhancement. SMEs experienced substantial enhancements in operational efficiency, revenue generation, and competitiveness as a result of the adoption of BD. The research, however, demonstrates that there are ongoing obstacles, including inadequate financial resources and technical proficiency. A potential overreliance on quantitative approaches may be emphasised by this imbalance, which may restrict the profundity of insights that can be obtained. Although BD has the potential to significantly drive innovation and improve the competitiveness of SMEs, it is essential to address the current methodological biases and resource-related barriers in order to fully realise its benefits. To achieve a comprehensive comprehension of the effects of BD on SMEs, future research should emphasise the use of a variety of methodologies.

(Khan & Shao, 2024) [18] Aim to conduct a comprehensive literature review that examines the influence of big data and knowledge management on consumer interactions and consumption patterns from an applied science basis. Several theoretical and practical implications are associated with the results of this investigation. Theoretically, this review contributes to the expanding corpus of literature on the intersection of consumer behaviour, knowledge management, and big data. In order to enhance consumer interactions and consumption patterns, the findings can provide policymakers and practitioners with practical insights on how to leverage big data and knowledge management.

(J. Zhang et al., 2023) [2] Examine how big data is affecting information science research approaches. “The information science research community” is encouraged to engage in constructive dialogue and contemplation regarding pertinent research methodologies. In the age of big data, this may help information practitioners overcome the difficulties they face and help researchers in the area create effective study designs for big data-oriented research investigations. Any research study is essential and fundamentally dependent on research methods. The relationship between research methods and big data

analysis is quite natural. The paper investigates "an emerging and significant research field": big data, from a unique perspective on research methodologies.

(Mamo, 2023) [19] Focusses on the ethical and privacy implications of the use of modern, innovative techniques for social media data. We look at how social media data is accessed, the natural language processing methods that are utilised, the problems they solve, the advantages and disadvantages of each approach, and how social media data is accessed. Additionally, the commentary demonstrates the effectiveness and suitability of sentiment-analysis tools (such Syuzhet, Bing, and AFFIN) for the examination of social media data in the sports domain. Thus, the future of sport management research can be substantially influenced by the rigorous application of contemporary innovative techniques. Prior to employing sophisticated analytical methods, researchers must exercise caution regarding the data's origin and preprocessing.

(Omoyiola, 2022) [20] In the current era, the implications of big data are profound, and they have become increasingly relevant to society. The utilisation of big data technology has been widely adopted, with its applications being implemented "at the national, organisational, and industry levels". Big data's impact on industry transformation is changing how government, business, education, and the humanitarian sector operate. It has a positive impact on our world in "all sectors, resulting in legal, economic, political, social, and ethical implications". Additionally, it provides insights. Innovation, efficiency, improved decision-making, and a higher return on investments are other outcomes of this change. Big data's societal ramifications, risks, difficulties, and prospects are all examined in this study.

(Lytras et al., 2020) [21] The social sciences are provided with new insights through the implementation of Big Data and analytics technology. It also provides innovative, adaptable solutions to critical social issues and obstacles. Offer this edited collection on big data research's societal ramifications. It is among the first worldwide efforts to look at how this kind of study affects people and societal concerns. Three pillars form the framework of the pertinent debate: Big Data Research's Social Impact is covered in Section A. Section B. Methods and Techniques for Social Sciences and Social Impact Research Driven by Big Data. Section C. Strategies for Researching Big Data. Complementary social and technical variables are the effects of Big Social Networks on Business Intelligence and Sustainable Economic Development.

(Lytras & Visvizi, 2019) [22] "Big data and big data-based services" were examined in order to determine their precise impact on individuals and societies. This paper provides a detailed examination of the ways in which individuals perceive data, with a particular emphasis on their perceptions of the actual sharing of their data. The purpose of this paper is to establish a value space for the social impact of big data that is relevant "to three factors: the intention to share personal data, individual concerns, and the social impact of big data". The main contribution of this research is the understanding it offers of the still-emerging area of study that takes place at the nexus of social science and computer science. We anticipate that this field of research will become increasingly significant in the coming years.

3 Conclusion

In conclusion, big data has revolutionized research methodology by enabling access to vast datasets and offering advanced analytical tools to uncover complex patterns and trends. Its influence spans various disciplines, empowering researchers to visualize relationships both statically and dynamically over time. However, the successful application of big data methods depends on the research context, objectives, and design. The current study introduces a conceptual framework assessing organizational readiness for big data adoption, highlighting the importance of temporal dynamics and real-time analytics in decision-making processes. Despite significant progress, key gaps remain in understanding how big data and knowledge management influence consumer behavior, presenting opportunities for further research. By bridging these gaps, researchers and companies can optimize customer engagement strategies and enhance competitive positioning. Furthermore, big data offers powerful tools in domains such as sport management, particularly through social media analytics. Techniques like Syuzhet, Bing, and AFFIN show promise in sentiment analysis, provided that researchers prioritize data quality, representativeness, and preprocessing. Ultimately, the integration of innovative, context-appropriate research methodologies will determine the extent to which big data enhances knowledge creation. This review affirms that big data's role in shaping responsive, socially aware services and evidence-based decision-making is just beginning to unfold.

References

[1] U. Sivarajah, S. Kumar, V. Kumar, S. Chatterjee, and J. Li, "A study on big data analytics and innovation: From technological and business cycle perspectives," Technol. Forecast. Soc. Change, vol. 202, no. October 2023, p. 123328, 2024, doi: 10.1016/j.techfore.2024.123328.

[2] J. Zhang, D. Wolfram, and F. Ma, "The impact of big data on research methods in information science," Data Inf. Manag., vol. 7, no. 2, p. 100038, 2023, doi: 10.1016/j.dim.2023.100038.

[3] M. Mallappallil, J. Sabu, A. Gruessner, and M. Salifu, "A review of big data and medical research," SAGE Open Med., vol. 8, 2020, doi: 10.1177/2050312120934839.

[4] S. Fanelli, L. Pratici, F. P. Salvatore, C. C. Donelli, and A. Zangrandi, "Big data analysis for decision-making processes: challenges and opportunities for the management of health-care organizations," Manag. Res. Rev., vol. 46, no. 3, pp. 369–389, 2023, doi: 10.1108/MRR-09-2021-0648.

[5] Z. Sun, "A Research Methodology for Big Data Intelligence," PNG UoT BAIS, vol. 7, no. 3, pp. 1–8, 2022, doi: 10.13140/RG.2.2.12956.97926/2.

[6] D. Riswantini, E. Nugraheni, A. Arisal, P. H. Khotimah, D. Munandar, and W. Suwarningsih, "Big data research in fighting COVID-19: Contributions and techniques," Big Data Cogn. Comput., vol. 5, no. 3, 2021, doi: 10.3390/bdcc5030030.

[7] L. Japec et al., Big data in survey research: Aapor task force report, vol. 79, no. 4. 2015. doi: 10.1093/poq/nfv039.

[8] S. Suvivuo and V. K. Tuunainen, Challenges and Solutions in Qualitative Big Data Research: A Methodological Literature Review, vol. 55. 2024. doi: 10.17705/1CAIS.05502.

[9] N. U. Shahid and N. J. Sheikh, "Impact of Big Data on Innovation, Competitive Advantage, Productivity, and Decision Making: Literature Review," Open J. Bus. Manag., vol. 09, no. 02, pp. 586–617, 2021, doi: 10.4236/ojbm.2021.92032.

[10] D. P. Acharjya and K. A. P, "A Survey on Big Data Analytics: Challenges, Open Research Issues and Tools," Int. J. Res. Appl. Sci. Eng. Technol., vol. 10, no. 12, pp. 1899–1908, 2016, doi: 10.22214/ijraset.2022.48294.

[11] M. S. Kumar and J. Prabhu, "Recent development in big data analytics: Research perspective," Res. Anthol. Artif. Intell. Appl. Secur., vol. 3–4, no. January, pp. 1640–1663, 2021, doi: 10.4018/978-1-7998-7705-9.ch072.

[12] A. Zhang and N. Lv, "Research on the Impact of Big Data Capabilities on Government's Smart Service Performance: Empirical Evidence from China," IEEE Access, vol. 9, pp. 50523–50537, 2021, doi: 10.1109/ACCESS.2021.3056486.

[13] S. Dehariya and D. Sharma, "Study On Highway Roadside Safety Prabhat Chauraha to Bajrang Chauraha Raisen Road Bhopal," pp. 44–50, 2024.

[14] O. Müller, I. Junglas, J. Vom Brocke, and S. Debortoli, "Utilizing big data analytics for information systems research: Challenges, promises and guidelines," Eur. J. Inf. Syst., vol. 25, no. 4, pp. 289–302, 2016, doi: 10.1057/ejis.2016.2.

[15] A. Gupta, A. Deokar, L. Iyer, R. Sharda, and D. Schrader, "Big Data & Analytics for Societal Impact: Recent Research and Trends," Inf. Syst. Front., vol. 20, no. 2, pp. 185–194, 2018, doi: 10.1007/s10796-018-9846-7.

[16] S. A. Z. Zaidi, E. Ahmad, and N. Shukla, "Ethical Considerations in the Use of Artificial Intelligence (AI) for Education and Research : A Review," Int. J. Innov. Sci. Eng. Manag., pp. 156–167, 2024, doi: 10.69968/ijisem.2024v3si2156-167.

[17] M. Kgakatsi, O. P. Galeboe, K. K. Molelekwa, and B. A. Thango, "The Impact of Big Data on SME Performance: A Systematic Review," pp. 632–695, 2024, doi: 10.20944/preprints202409.0985.v1.

[18] M. N. Khan and Z. Shao, "Impact of Big Data and Knowledge Management on Customer Interactions and Consumption Patterns: Applied Science Research Perspective," Eng. Technol. Appl. Sci. Res., vol. 14, no. 3, pp. 14125–14133, 2024, doi: 10.48084/etasr.7203.

[19] Y. Z. Mamo, "Big Data and Innovative Research Methods," Int. J. Sport Commun., vol. 16, no. 3, pp. 352–360, 2023, doi: 10.1123/ijsc.2023-0109.

[20] B. O. Omoyiola, "The social implications, risks, challenges and opportunities of big data," Emerald Open Res., vol. 4, p. 23, 2022, doi: 10.35241/emeraldopenres.14646.1.

[21] M. D. Lytras, A. Visvizi, and K. T. Chui, Big Data Research for Social Sciences and Social Impact. 2020. doi: 10.3390/books978-3-03928-221-0.

[22] M. D. Lytras and A. Visvizi, “Big data and their social impact: Preliminary study, Sustain” vol. 11, no. 18, 2019, doi: 10.3390/su11185067.

AG Publishing House

Advanced Research Techniques: Theories, Methods, And Practices (Volume-2)

Volume 2
Year: 2025

AG PH Books

The Impact of Digital Tools on Modern Research Practices: The Role of Institutional Review Boards (IRBs) in Research Ethics

Dr. Prasanna Kumar Yekula [1*], Dr. Wilson Kobal [2]

[1]*Professor, School of Mining Engineering, Faculty of Engineering, PNG University of Technology Private Mail Bag Lae 411 Morobe Province, Papua New Guinea.*

[2]*Lecturer, School of Mining Engineering, Faculty of Engineering, PNG University of Technology Private Mail Bag Lae 411 Morobe Province, Papua New Guinea.*

Abstract

In the 21st century, the development of digital tool technologies has led to a notable improvement in social science research worldwide. Thanks to these technological advancements, researchers can now easily access and analyse information, manage their time, and create more engaging research. This has completely changed the way that research is conducted. Review of the many studies on the influence of digital technologies on contemporary research methods is provided in this article. It concluded that digital tools, including AI and big data analytics, are transforming modern research practices by enhancing efficiency, data quality, and collaborative capabilities across healthcare, education, and social research. Institutions play a crucial role by investing in infrastructure, training, and ethical frameworks to ensure responsible and effective technology integration. While these tools offer significant advantages, challenges such as data privacy, ethical concerns, and the digital divide must be addressed. By fostering digital literacy, expanding access, and implementing clear guidelines, institutions can harness the full potential of digital tools to advance research practices in an ethical and inclusive manner.

Keywords: Digital tool technologies, Modern research practices, Data analysis, Data collection, Social research, Data privacy, Ethical concerns, Research methods, etc.

[*] ISBN No. - 978-93-49028-02-9

1 Introduction

Recent years have seen revolutionary developments brought about by the use of contemporary technology into social research methodologies, particularly in educational settings. In particular, the methods used for data gathering, processing, and interpretation have changed as a result of digital tools and platforms. Digital sociology, for example, has developed as a discipline to comprehend these shifts, particularly in the educational sector [1]. Digital technology has an impact on social functioning and human behaviour, posing new problems for social research as well as methodological advancements. Specifically, digital technologies like social media, big data analytics, and online surveys are playing a bigger role in educational research because they provide previously inaccessible information [2]. Social media platforms have developed into abundant data sources that enable academics to examine interactions, behaviours, and trends in real time. Additionally, sophisticated analytics methods like machine learning and natural language processing improve researchers' comprehension of social data, including sentiment analysis and user behaviour [3]. Online surveys, despite its inherent biases, have also gained popularity as practical and affordable means of gathering data from a variety of demographics. Our comprehension of behaviours and consequences in social and educational settings is improved by the use of data in the form of conducting the process of merging survey and social media data, such as connecting Twitter data with mental health surveys [4].

Digital technologies have complicated ethical ramifications for social research. Concerns about privacy, particularly with regard to educational data, have drawn particular attention as academics struggle with the moral need to protect student and teacher data [5]. Similar ethical issues are brought about by techno-social developments in the healthcare industry, such as the escalation of the digital gap and worries about data justice and equality. Ethical factors including openness, permission, and fairness must be balanced with the advantages of data-driven insights in education in order to ensure ethical and equitable research processes [6].

A. The evolution of digital tools in social research

When statistical software such as SPSS was introduced in the middle of the 20th century, it completely changed how researchers handled big datasets and carried out intricate analyses. This marked the beginning of the development of digital tools in social research. As increasingly complex software, like R and Python, was developed over time, researchers' skills improved even more, enabling more in-depth data analysis and visualization [7]. These instruments signalled a dramatic change from manual data processing to more automated and effective techniques, setting the groundwork for the incorporation of digital technology into all phases of the research process [8], [9]. Applications of digital technology in social research have evolved along with it. The field of data gathering saw a significant transformation with the introduction of online survey tools such as Qualtrics, SurveyMonkey, and Google Forms in the late 20th century. These platforms greatly decreased the time and expense involved with conventional survey methodologies by enabling researchers to swiftly and effectively contact larger and more varied audiences [10].

B. Impact of digital tools on social research

In social science research, digital technologies have had a significant influence on data gathering techniques. Researchers no longer have to worry about their location or lack of resources to gather data thanks to online survey platforms like Qualtrics, Google Forms, and Survey Monkey. Particularly Qualtrics has become well-known because of its complex features, which enable more intricate and customised survey designs, such as randomisation, embedded data, and logic branching [11]. In addition to making the data collecting procedure more efficient, these technologies have raised the calibre and dependability of the data that was gathered. The development of online communities for collaborative research purposes has been made possible by new technologies, which have improved community and collaboration. This is especially true when the tools are used to share research materials and information, such as Twitter and social bookmarking [10], [12].

Some of the issues with conventional data gathering techniques have been resolved thanks to the use of mobile-based data collection tools like ODK (Open Data Kit), especially in rural and isolated locations. Researchers doing fieldwork in places with poor internet access have found ODK to be a useful tool because to its offline functionality and form design flexibility [13]. More thorough and reliable data collecting has resulted from this, improving the quality of research results across a range of disciplines. It is impossible to overestimate the importance of digital tools in data analysis and interpretation because statistical software, such as SPSS and Excel, has long been a mainstay of social science research, enabling researchers to easily conduct a variety of statistical tests and data manipulations [14]. Digital technologies have also made it easier for researchers to access hard-to-reach and marginalised people, which has led to more inclusive research procedures. To ensure that diverse perspectives are represented in the research process, online surveys, for example, may be broadly dispersed across various demographics and geographical areas. This inclusion is especially significant in social research, since creating successful policies and treatments requires an awareness of all groups' viewpoints [15], [16].

2 Literature Review

(Al-Dmour et al., 2025) [17] Examines the impact of artificial intelligence (AI) and big data analytics on healthcare outcomes in Jordanian healthcare institutionss. According to the results, AI technologies substantially enhance diagnostic accuracy and treatment planning, while big data analytics improve operational efficiency and patient care. It is essential for policymakers, healthcare administrators, and technology developers to comprehend the relationship between AI, big data analytics, and healthcare delivery in order to create effective strategies that enhance operational efficiency and patient care. This investigation proposes that healthcare personnel be provided with exhaustive training, organisational capabilities be improved, and user-friendly AI and big data analytics tools be implemented.

(Aljaunid, 2025) [18] Examines how contemporary technology has affected social research techniques in educational settings, paying particular attention to the ways in which digital technologies affect the procedures of data collection, analysis, and interpretation. The results show that regular technology usage greatly improves research speed and data correctness, giving educational researchers more options for

gathering data and more analytical precision. However, difficulties including the digital divide, ethical conundrums, and data privacy concerns have surfaced as significant barriers to the successful incorporation of technology in educational research. By providing insights into the advantages and difficulties of integrating digital tools and proposing helpful suggestions for future research in an increasingly digitalised educational environment, this study adds to the conversation on technology in educational research.

(John et al., 2024) [10] Thus, the purpose of this work is to examine the many kinds of cutting-edge digital tools used in social science research and how they contribute to higher-quality research. A qualitative analysis was performed on the gathered data. According to the results, there are around six distinct cutting-edge digital tools available for use in social science research, and the calibre of study is greatly influenced by these tools. However, because of their lack of understanding, many researchers continue to underuse these tools. Therefore, in order to provide academic researchers with a comprehensive understanding of the many kinds of digital tools that are accessible and how to use them effectively in order to conduct high-quality research, the article suggests that additional awareness-raising and groundbreaking works be done.

(Ladeira et al., 2024) [19] To further comprehend these performances, a study of the empirical findings from publications addressing big data analytics in the services sector is required. To determine the pertinent aspects of big data analytics and assess artificial intelligence as a possible moderator of its impacts on service performance, this research carried out a meta-analysis based on this justification. The findings show that the adoption of big data analytics is driven by competitive pressure, resources and capabilities, and environmental dynamism. Adoption of big data analytics is more impacted by environmental dynamic than by resources and skills. According to the results, using big data analytics with artificial intelligence improves service performance more than implementing big data analytics without AI.

(Md. Aminul Islam, 2024) [20] To address any study gaps on the influence of big data analytics on digital marketing strategies, a thorough literature review in these areas should be conducted. Since big data has a significant impact on the development of digital advertising strategies and the ways in which big data influences advertising, this research assessed the existing literature on big data applications and found that digital marketing is a broad field. During one of the most devastating pandemics in history, several organisations have found that they may overcome significant obstacles by using the best big data solutions.

(Bryda & Costa, 2023) [21] The importance of qualitative research and the qualitative computer data analysis process in leveraging digital potential and influencing collaborative work is emphasised in this article's critical reflection. This process is essentially a dialectical interaction between the traditional technique and the emerging digital technological world. The researcher's personality and analytical and research processes are shaped by the use of digital tools in qualitative research activities. Additionally, it promotes the growth of new analytical, digital, and information technology (IT) abilities while teaching cooperation and collaborative thinking. It is challenging to envision modern qualitative research and data

analysis in the social sciences and humanities. Our interpretation frameworks are shaped by our exposure to contemporary computer-based qualitative data analysis methods, which also alter the perspective and perception of research issues.

(Rosa et al., 2022) [22] Gives an update on the use of technology in clinical trials during the last five years, citing instances from a range of medical illnesses and with differing degrees of technological integration. The use of real-world data (such as electronic health records) for study recruitment and the integration of artificial intelligence into diagnostic instruments are examples of digital technology integration. With the advent of fully virtual clinical trials, face-to-face contact is no longer necessary. A significant portion of the literature shows how digital methods may enhance clinical trial design and execution. Even while there are still obstacles to overcome, the previous five years have seen positive development, and with good preparation, they may be surmounted.

(Xu et al., 2022) [23] Examine how big data analytics courses contribute to engineering students' skill and employability development via digital education. According to the empirical research, engineering students' career growth aspects, personal qualities, and human capital are all positively impacted by both hard and soft talents. The study's theoretical and practical ramifications expand our understanding of the role of digitisation in improving engineering students' employability and skill sets while also adding to the body of information on engineering education.

(Sabharwal & Miah, 2021) [24] The use of Big Data Analytics (BDA) in the sector has grown significantly in the last few years. The need of BDA capacity in organisations is acknowledged as a data-driven tool to support well-informed decision-making; yet, there aren't many research that have explained BDA capabilities in a manner that may advance our theoretical understanding of using BDA in the organisational domain. Different definitions of big data have been proposed, and this study examines the literature on the categorisation of BDA and its potential. Our results increase the efficiency and use of BDA applications across different organisations.

3 Conclusion

In conclusion, the integration of AI and big data analytics into modern research practices has significantly enhanced operational efficiency, data quality, and service performance across sectors such as healthcare and education. Institutions that prioritize investments in digital technologies can deliver more accurate diagnostics, personalized solutions, and responsive services. To fully realize these benefits, user-friendly tools must be developed, supported by robust infrastructure, comprehensive training, and strong institutional leadership. In education and social research, digital tools facilitate more efficient and collaborative work, improving the quality and scope of qualitative data analysis. However, challenges such as ethical concerns, data privacy, the digital divide, and limited researcher proficiency with digital tools must be addressed. Institutions play a vital role in bridging these gaps through targeted interventions, expanding access to reliable internet and digital resources, and fostering a culture of ethical digital research. Establishing clear guidelines and frameworks for responsible use of digital tools is essential to mitigate algorithmic bias and ensure transparency. By supporting inclusive, ethical, and well-

informed adoption of digital tools, institutions can maximize the transformative potential of technology in research, advancing knowledge and innovation in a responsible and equitable manner.

References

[1] D. Wang and X. Shao, "Research on the impact of digital transformation on the production efficiency of manufacturing enterprises: Institution-based analysis of the threshold effect," Int. Rev. Econ. Financ, vol. 91, no. January, pp. 883–897, 2024, doi: 10.1016/j.iref.2024.01.046.

[2] A. Haleem, M. Javaid, M. A. Qadri, and R. Suman, "Understanding the role of digital technologies in education: A review," Sustain. Oper. Comput, vol. 3, no. February, pp. 275–285, 2022, doi: 10.1016/j.susoc.2022.05.004.

[3] M. Fırat, H. Altınpulluk, and H. Kılınç, "Determination of digital technologies preferences of educational researchers," Asian Assoc. Open Univ. J., vol. 16, no. 1, pp. 20–40, 2021, doi: 10.1108/AAOUJ-09-2020-0064.

[4] M. BAJÚZOVÁ and R. HRMO, "Digital Tools in Education The: The Impact of Digital Tools in Education on Students' Creativity," J. Technol. Inf., vol. 15, no. 2, pp. 150–171, 2024, doi: 10.5507/jtie.2023.007.

[5] T. Gkrimpizi, V. Peristeras, and I. Magnisalis, "Classification of Barriers to Digital Transformation in Higher Education Institutions: Systematic Literature Review," Educ. Sci., vol. 13, no. 7, 2023, doi: 10.3390/educsci13070746.

[6] A. I. Stoumpos, F. Kitsios, and M. A. Talias, "Digital Transformation in Healthcare: Technology Acceptance and Its Applications," Int. J. Environ. Res. Public Health, vol. 20, no. 4, 2023, doi: 10.3390/ijerph20043407.

[7] F. Loureiro, L. Sousa, and V. Antunes, "Use of digital educational technologies among nursing students and teachers: An exploratory study," J. Pers. Med., vol. 11, no. 10, pp. 1–13, 2021, doi: 10.3390/jpm11101010.

[8] S. M. Alshehri, "Perspective Chapter: The Impact of Digital Education on Modern Educational Institutions," Intech, p. 13, 2024, [Online]. Available: http://dx.doi.org/10.1039/C7RA00172J%0Ahttps://www.intechopen.com/books/advanced-biometric-technologies/liveness-detection-in-biometrics%0Ahttp://dx.doi.org/10.1016/j.colsurfa.2011.12.014

[9] S. A. Z. Zaidi, E. Ahmad, and N. Shukla, "Ethical Considerations in the Use of Artificial Intelligence (AI) for Education and Research : A Review," Int. J. Innov. Sci. Eng. Manag., pp. 156–167, 2024, doi: 10.69968/ijisem.2024v3si2156-167.

[10] P. John, S. Shettima, and A. M. Sulaiman, "The Role of Innovative Digital Tools in Enhancing Social Science Research," no. September, pp. 0–20, 2024.

[11] K. Okoye et al., Impact of digital technologies upon teaching and learning in higher education in Latin America: an outlook on the reach, barriers, and bottlenecks, vol. 28, no. 2. Springer US, 2023. doi: 10.1007/s10639-022-11214-1.

[12] S. Rafiq, S. Iqbal, and A. Afzal, "The Impact of Digital Tools and Online Learning Platforms on Higher Education Learning Outcomes," Al-Mahdi Res. J., no. May, 2024.

[13] I. T. Awidi and M. Paynter, "An Evaluation of the Impact of Digital Technology Innovations on Students' Learning: Participatory Research Using a Student-Centred Approach," Technol. Knowl. Learn, vol. 29, no. 1, pp. 65–89, 2024, doi: 10.1007/s10758-022-09619-5.

[14] C. Wang, X. Chen, T. Yu, Y. Liu, and Y. Jing, "Education reform and change driven by digital technology: a bibliometric study from a global perspective," Humanit. Soc. Sci. Commun., vol. 11, no. 1, 2024, doi: 10.1057/s41599-024-02717-y.

[15] P. Samadi-Parviznejad, "The role of big data in digital transformation," J. Data Anal., vol. 1, no. 1, pp. 42–47, 2022, doi: 10.59615/jda.1.1.42.

[16] Dr Shaheen Parveen and Shaikh Imran Ramzan, "The Role of Digital Technologies in Education: Benefits and Challenges," Int. Res. J. Adv. Eng. Manag., vol. 2, no. 06, pp. 2029–2037, 2024, doi: 10.47392/irjaem.2024.0299.

[17] R. Al-Dmour, H. Al-Dmour, E. Basheer Amin, and A. Al-Dmour, "Impact of AI and big data analytics on healthcare outcomes: An empirical study in Jordanian healthcare institutions," Digit. Heal. vol. 11, 2025, doi: 10.1177/20552076241311051.

[18] H. Aljaunid, "The Impact of Modern Technology on Social and Education Research Methods in Educational Contexts : An Applied Study," Int. J. Educ. Teach. Zo., vol. 4, no. 1, pp. 1–2, 2025.

[19] W. J. Ladeira et al., "Big data analytics and the use of artificial intelligence in the services industry: a meta-analysis," Serv. Ind. J., 2024, doi: 10.1080/02642069.2024.2374990.

[20] Md. Aminul Islam, "Impact of Big Data Analytics on Digital Marketing: Academic Review," J. Electr. Syst., vol. 20, no. 5s, pp. 786–820, 2024, doi: 10.52783/jes.2327.

[21] G. Bryda and A. P. Costa, "Qualitative Research in Digital Era: Innovations, Methodologies and Collaborations," Soc. Sci., vol. 12, no. 10, 2023, doi: 10.3390/socsci12100570.

[22] C. Rosa, L. A. Marsch, E. L. Winstanley, M. Brunner, and A. N. C. Campbell, "Using digital technologies in clinical trials: current and future applications," HHS Public Access, 2022, doi: 10.1016/j.cct.2020.106219.Using.

[23] L. Xu et al., "Assessing the impact of digital education and the role of the big data analytics course to enhance the skills and employability of engineering students," Front. Psychol., vol. 13, no. October, pp. 1–20, 2022, doi: 10.3389/fpsyg.2022.974574.

[24] R. Sabharwal and S. J. Miah, "A new theoretical understanding of big data analytics capabilities in organizations: a thematic analysis," J. Big Data, vol. 8, no. 1, 2021, doi: 10.1186/s40537-021-00543-6.