Enhancing Research with Chatgpt: Seven Strategic Applications

Mohd Nasir Saari[*], **Mohd Adza Arshad**

Centre of Excellence for Engineering and Technology (CREaTE), Public Works Department Malaysia (PWD), Malaysia

ABSTRACT

The rise of artificial intelligence in academic research has introduced transformative opportunities to improve the efficiency, accessibility, and analytical depth of scholarly work. This study investigates the strategic integration of ChatGPT, an advanced generative AI model, into seven critical stages of the research data analysis lifecycle: data cleaning and structuring, exploratory data analysis, visualisation, statistical modelling, interactive tabulation, literature synthesis, and report generation. Drawing on empirical use cases and theoretical perspectives, we argue that ChatGPT functions not only as a computational tool but also as a cognitive collaborator capable of enhancing methodological rigor and interpretive clarity. This study demonstrates how the ChatGPT reduces technical barriers, fosters inclusive participation, and accelerates knowledge production across disciplinary boundaries. While the model offers significant pedagogical and epistemological benefits, we also identified ethical challenges related to AI bias, accuracy, authorship, and research integrity. The paper concludes by positioning ChatGPT as a partner in reimagining human–machine collaboration in the research enterprise and calls for critical frameworks to guide its responsible adoption in academic ecosystems. This study contributes to the ongoing debates on AI-assisted scholarship and provides a roadmap for integrating large language models into high-impact research.

Keywords: academic writing, artificial intelligence, ChatGPT, data analysis, methodological innovation, research workflows, scholarly communication

INTRODUCTION

The contemporary research landscape is undergoing a profound transformation, driven by the exponential growth of data and the emergence of intelligent computational tools. Researchers are becoming increasingly required to navigate complex datasets, interpret multifaceted results, and deliver actionable insights with speed and precision. Amid these challenges, artificial intelligence (AI) tools, such as ChatGPT, have emerged as powerful enablers of research efficiency and analytical capability.

ChatGPT, a generative language model powered by advanced machine learning, offers a versatile set of functions that span data cleaning, exploration, statistical analysis, visualisation, reporting, and knowledge integration. This enables researchers to engage with their data through natural language interactions, effectively bridging the gap

between technical complexity and user accessibility. By automating routine processes and enhancing cognitive tasks, the ChatGPT redefines the workflow of data-driven research across disciplines.

This article explores seven strategic applications of ChatGPT that are instrumental in supporting high-quality research. These functions are discussed in relation to real-world academic scenarios, demonstrating how AI can be effectively deployed to streamline analysis, strengthen interpretive depth, and foster methodological innovation. As the demand for rapid, reproducible, and rigorous research continues to increase, integrating ChatGPT into research workflows is not merely an advantage; it is also becoming an essential component of modern scholarly practices.
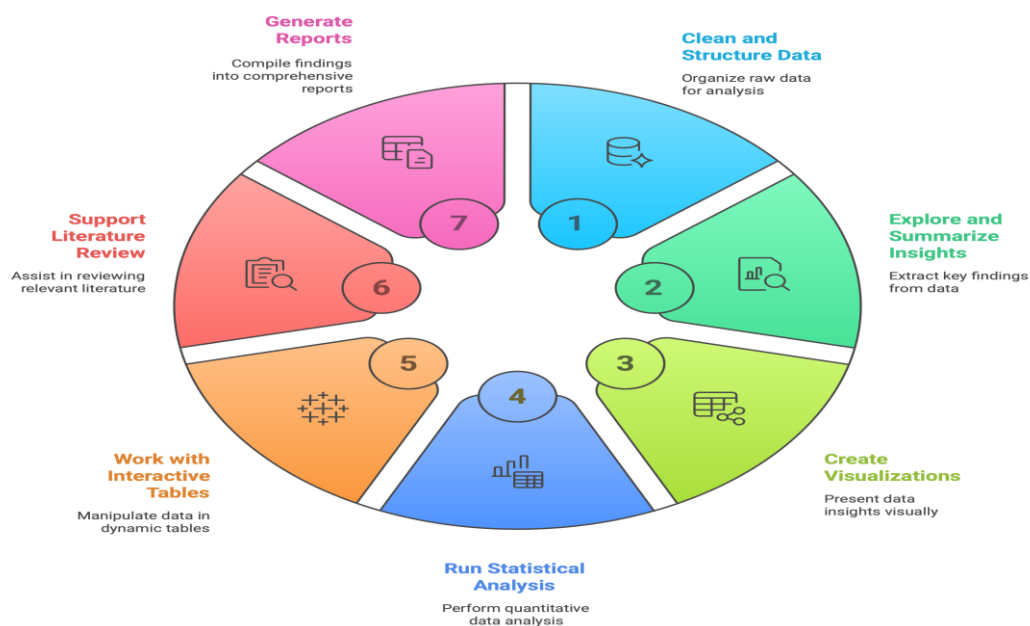

Figure 1. Overview of the Seven Strategic Applications of ChatGPT in Research Workflow.

This figure provides a visual summary of ChatGPT's roles across the research lifecycle, serving as a conceptual framework for the structured applications discussed in this study.

Clean and Structure Data

Robust data analysis is predicated on the integrity and consistency of the underlying dataset. Raw data frequently present challenges such as typographical errors, missing values, inconsistent units, and formatting irregularities. These issues, if not addressed systematically, can undermine the validity of the research outcomes. ChatGPT provides researchers with an intelligent dialogic interface to perform data-cleaning tasks using natural language prompts. The model can detect anomalies and standardised formats and offers context-specific recommendations for restructuring datasets (Lund et al., 2023; OpenAI, 2023). This functionality is particularly crucial in large-scale, longitudinal, and multi-source datasets, which are often employed in public health, environmental science, and economic research. Moreover, ChatGPT facilitates batch processing and reproducible workflow by guiding users through stepwise transformations. By automating these preparatory steps, researchers can obtain more time and accuracy for higher-order analytical tasks, thereby ensuring the reliability of the subsequent inferences.

In practice, data preprocessing often consumes 60–80% of a research project's analytical cycle (Zhou et al., 2023). This is because of the heterogeneous nature of the data collection methods, especially in multisite or collaborative studies. ChatGPT can alleviate this burden by automating repetitive tasks and flagging inconsistencies across the datasets. For example, they can detect and harmonise inconsistent labelling schemes, standardise numerical entries, and reconcile data formats from various international systems (Dwivedi et al. 2023). This intelligent preprocessing allows researchers to maintain methodological rigor without dedicating

disproportionate time to mechanical adjustments. Furthermore, ChatGPT can document every transformation performed, supporting transparency, auditability, and compliance with data governance frameworks.

In multilanguage datasets, such as those common in comparative education, global health, or migration studies, ChatGPT's multilingual capabilities are particularly useful. It can standardise terminology across diverse linguistic sources, translating column headers or textual responses while preserving semantic consistency. For projects involving semi-structured data, such as open-ended survey responses or field notes, ChatGPT can also perform early categorisation, keyword tagging, and sentiment scoring, making unstructured data more amenable to quantitative or thematic analyses. These functions dramatically extend the reach of traditional preprocessing tools and reduce the dependence on expensive software or dedicated data engineering teams.

Explore and Summarise Insights

Exploratory Data Analysis (EDA) serves as the foundation for hypothesis generation and pattern recognition in empirical studies. ChatGPT enhances EDA by allowing researchers to pose qualitative queries—such as "What are the dominant patterns in this dataset?"and received quantitative and interpretative summaries. These insights encompass measures of central tendency, variability, frequency distribution, and anomaly detection (Zhou et al., 2023). In contrast to traditional EDA methods that often require coding proficiency, ChatGPT democratises access to data exploration, thereby supporting inclusive research participation across disciplines (Lund et al. 2023). Importantly, the model can contextualise trends by integrating domain-specific knowledge, offering early theoretical framing of the findings. This capability strengthens construct validity and assists in aligning data interpretation with established empirical literature.

Moreover, the ChatGPT supports the development of iterative insights. Researchers can drill down into subsets of data or compare variables across temporal and demographic dimensions (Vinuesa et al. 2020). This interactivity facilitates hypothesis refinement and improves internal validity. ChatGPT's interpretive support is not limited to numerical summary. It can also highlight counterintuitive trends, suggest potential covariates, and flag variables with high leverage or multicollinearity risks. These functions are particularly valuable in exploratory research, where emergent patterns guide the selection of future research designs, sample segmentation, and experimental modelling.

Advanced applications include the use of ChatGPT for preliminary clustering and segmentation analyses. Through unsupervised learning support or scripting assistance, ChatGPT can help identify natural groupings in data such as behavioural archetypes or market segments. It can also generate narratives describing the characteristics of each group in an accessible language. This contributes to both the methodological depth and communicative clarity of early research outputs. Furthermore, ChatGPT's capacity to summarise complex trends for different audiences, such as policymakers, clinicians, and community stakeholders, enhances the societal relevance and translational potential of the research findings.

Create Visualisations

Visualisation is indispensable for data interpretation and dissemination. Effective visuals enable the detection of structural relationships, identification of outliers, and presentation of findings to nontechnical audiences. ChatGPT supports the generation of diverse graphical outputs including bar charts, histograms, pie charts, boxplots, and scatterplots, which can be iteratively refined based on user feedback (Heaven, 2023). The conversational interface allows researchers to specify parameters, group variables, and label axes without recourse to specialised software or programming. In high-impact research, where clear communication of data trends is vital for peer review, policymaking, and interdisciplinary dialogue, ChatGPT's visualisation support fosters accessibility, transparency, and interpretive accuracy.

A key strength of ChatGPT's visualisation capacity is its adaptability. It can generate visual comparisons between experimental groups, illustrate regression lines with confidence bands, and display time-series fluctuations to highlight seasonal or cyclical patterns (OpenAI, 2023). The ability to create customised dashboards for interactive presentations also positions ChatGPT as an ally for real-time stakeholder consultation.

This visual responsiveness is especially valuable in applied research fields, such as public health, education policy, and disaster risk management.

Beyond static graphs, ChatGPT can generate HTML-compatible or script-ready visual codes, allowing researchers to deploy interactive plots on web-based dashboards and mobile applications. This capability is becoming increasingly valuable in participatory research, where real-time feedback loops and data visualisation for lay audiences are required. In disciplines such as urban planning and environmental monitoring, ChatGPT can be used to convert geospatial data into heat maps or choropleth charts, facilitating spatiotemporal analyses that would otherwise require specialised GIS tools. Its ability to tailor visual output to various dissemination formats, from academic articles to policy briefs and conference presentations, renders it a powerful cross-platform visual assistant.

Run Statistical Analysis

Inferential statistics are essential for validating the hypotheses and generalising the findings. ChatGPT assists in conducting a wide range of statistical procedures, from descriptive analytics to advanced inferential models, while simultaneously offering methodological scaffolding (Lund et al. 2023). This tool provides real-time guidance for test selection (e.g. ANOVA, regression, chi-square), assumption diagnostics, and interpretation of outputs such as p-values, confidence intervals, and effect sizes (OpenAI, 2023). This dual functionality enhances analytical rigor and reduces the common misinterpretations of statistical results.

ChatGPT also provides transparency in analytical decision-making by explaining the rationale behind test selection and model interpretation in plain languages (Dwivedi et al., 2023). This is particularly beneficial in interdisciplinary projects involving collaborators from nonstatistical backgrounds. Additionally, it supports simulation-based inference methods and can assist with bootstrapping techniques, Bayesian analysis, and longitudinal modelling approaches. These advanced functionalities enable researchers to manage complex data structures such as nested data or repeated measures, thereby increasing the depth and validity of empirical enquiries.

Importantly, the ChatGPT can serve as an adaptive learning environment. Walking researchers through hypothesis testing, degrees of freedom, and statistical power analysis functions as tutors for graduate students and novice analysts. This pedagogical role elevates the ChatGPT beyond being a mere computational tool, positioning it as a research mentor capable of fostering statistical literacy and methodological independence.

Work with Interactive Tables

Static data presentation limits researchers' analytical flexibility, particularly when the datasets are large or multidimensional. ChatGPT enhances table-based data interaction by enabling real-time querying, filtering, and restructuring of the tabular information. Users can request, for example, "Sort by mean test score, descending", or "Group by department and calculate the average publication output", allowing for hands-free manipulation of data tables that mirror advanced spreadsheet or database functionalities (Zhou et al., 2023).

This capability is particularly valuable in environments where datasets evolve over time or require tailored views for diverse audiences. ChatGPT can assist in generating comparison tables, pivot summaries, and crosstabs, which are useful for program evaluation, monitoring indicators, and equity analysis. Additionally, its ability to convert data into HTML or LaTeX table formats supports high-quality formatting for journal submissions and web-based publications.

For researchers managing meta-analyses or systematic reviews, ChatGPT can structure the extracted study data into evidence tables, summarising effect sizes, sample characteristics, and quality assessments. It can also automate table updates when new studies are added, thereby ensuring the integrity of data during the review process. This integration of interactivity, automation, and presentation utility aligns well with the dynamic needs of contemporary research workflows and publication standards.

SUPPORT LITERATURE REVIEW

One of the most time-intensive stages of research is literature review, which requires identifying relevant works, extracting core arguments, synthesising diverse perspectives, and locating gaps in knowledge. ChatGPT significantly reduces this burden by leveraging natural language processing to efficiently scan, summarise, and compare vast bodies of the literature (Lund et al., 2023; Dwivedi et al., 2023). It can generate annotated bibliographies, perform thematic analysis, and outline contrasting positions within scholarly debate.

Moreover, the ChatGPT supports theory development by identifying key constructs, mediators, and moderators across studies. For example, when researching motivation in online learning, the tool might surface prevalent theories, such as Self-Determination Theory or Expectancy-Value Theory, and indicate how these are operationalised in the literature. It also assists in tracing the historical evolution of concepts and demonstrates how definitions and models have shifted over time, which is crucial for producing nuanced conceptual frameworks.

In systematic and scoping reviews, ChatGPT can draft inclusion/exclusion criteria, organise findings using PRISMA-aligned formats, and support the critical appraisal of evidence using tools such as CASP or GRADE. Its multilingual capabilities allow cross-language literature exploration, which is valuable for global and comparative research. By accelerating and enhancing literature integration, ChatGPT allows researchers to engage more critically and creatively with existing scholarships.

Generate Reports

The ability to communicate research findings effectively is as important as the analysis itself. ChatGPT meaningfully contributes to this process by generating draft manuscripts, reports, executive summaries, and policy briefs tailored to diverse audiences and publication formats (Heaven, 2023; OpenAI, 2023). Researchers can prompt the model to produce components of academic papers, such as abstracts, introductions, methods, results, and discussions, based on raw data, notes, or outlines.

What sets ChatGPT apart is its stylistic adaptability. For example, it can produce a formal tone appropriate for peer-reviewed journals, or a more accessible tone suitable for outreach reports. It aligns content with disciplinary conventions such as APA or IEEE formatting and supports multilingual translations for international dissemination. These capacities are especially useful in collaborative, multi-author projects where consistency in language and argumentation is vital.

Moreover, ChatGPT allows recursive drafting, whereby sections are revised based on feedback loops from coauthors, reviewers, and editors. It facilitates alignment with journal scopes, addresses reviewer comments, and ensures coherence across sections. This iterative role underscores ChatGPT's value, not just as a drafting assistant but also as a key partner in the scholarly communication process.

Figure 3. Key Dimensions of ChatGPT Contribution to Research Practice.

This figure highlights the four core dimensions of ChatGPT's integration into research: strategic functionality, inclusivity, educational value, and ethical responsibility.

CONCLUSION: STRATEGIC INTEGRATION OF CHATGPT IN RESEARCH WORKFLOWS

This study critically examines the multifaceted roles that ChatGPT plays in enhancing contemporary research data analysis workflow. By outlining seven strategic applications—data cleaning and structuring, exploratory data analysis, visualisation, statistical analysis, interactive tabulation, literature synthesis, and report generation—we demonstrated how ChatGPT can serve as a comprehensive digital assistant across the research continuum. Each function contributes not only to increased efficiency but also to greater epistemic clarity, methodological rigor, and communicative effectiveness. These qualities align with the evolving demands of the Q1-level scholarly production, where innovation, transparency, and impact are paramount.

The implications of integrating ChatGPT into the research environment are far from realised. First, it empowers researchers from diverse disciplinary backgrounds to engage with complex data, without requiring advanced coding or statistical expertise. This democratisation of research analytics can help bridge the divide between data-rich disciplines and those historically underrepresented in quantitative research. Second, the ChatGPT offers significant pedagogical benefits by providing real-time, contextualised explanations that foster learning and skill development, which is particularly relevant in graduate training and interdisciplinary education.

Moreover, the capabilities of ChatGPT extend beyond technical tasks. It functions as an intellectual collaborator to enhance conceptual synthesis and encourages reflexivity in data interpretation. By facilitating more iterative, dialogic, and transparent interactions with research material, the tool promotes a more inclusive and critically engaged mode of scholarship. Its utility in multilingual contexts and ability to synthesise vast volumes of literature further posits it as a bridge between global knowledge systems, enabling richer comparative and transdisciplinary research.

However, the responsible adoption remains a critical concern. Although ChatGPT introduces new efficiency, its use must be governed by ethical protocols. Researchers should be vigilant regarding AI-generated inaccuracies, unacknowledged biases, and the need for human oversight in analytical judgments. To ensure the integrity of scholarly output, questions on authorship, intellectual contribution, and disclosure must be addressed within institutional frameworks and journal policies.

In conclusion, the ChatGPT represents a paradigm shift in research practice. It redefines the human–machine relationship in knowledge production, not as a tool of automation, but as a partner in cognitive augmentation. As AI technologies mature, their integration into academic workflow is likely to become more seamless and transformative. Future research should evaluate the longitudinal impact of such integration on research quality, participation diversity and epistemological innovation. By embedding ChatGPT thoughtfully and critically into a research enterprise, scholars can harness its potential to advance not only the pace but also the substance of scientific discovery.

ACKNOWLEDGEMENT

REFERENCES

1. Dwivedi, Y. K., Hughes, D. L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... & Williams, M. D. (2023). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 71, 102642. https://doi.org/10.1016/j.ijinfomgt.2022.102642
2. Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E. (2022). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

3. Heaven, W. D. (2023). AI-powered tools like ChatGPT are changing how we interact with data. *MIT Technology Review*. Retrieved from https://www.technologyreview.com/

4. Lund, B. D., Wang, T., Mannheimer, S., & West, J. D. (2023). ChatGPT and artificial intelligence in academia: A new era of research productivity. *College & Research Libraries News*, 84(2), 80–85. https://doi.org/10.5860/crln.84.2.80

5. OpenAI. (2023). *GPT-4 technical report*. Retrieved from https://openai.com/research/gpt-4

6. Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., ... & Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature Communications*, 11(1), 233. https://doi.org/10.1038/s41467-019-14108-y

7. Zhou, Y., Zhang, R., & Guo, Y. (2023). Using artificial intelligence to enhance research analytics: A systematic review. *Information Processing & Management*, 60(1), 102045. https://doi.org/10.1016/j.ipm.2022.102045