http://www.moderntechno.de/index.php/meit/article/view/meit42-03-055
DOI: 10.30890/2567-5273.2025-42-03-055

UDC 004.8:004.62

THE EVOLUTION OF DATA ANALYSIS TOOLS: FROM SPREADSHEETS TO ARTIFICIAL INTELLIGENCE AGENTS

Tumanov O.O.
PhD
ORCID:0000-0003-0674-0037
V.N.Karazina Kharkiv National University,
Kharkiv, Maydan Svobody, 4, 61022

Abstract. *The evolution of data analysis tools signals a deeper transformation in the foundations of scientific research. This article systematically analyzes the transition between three analytical paradigms - spreadsheets, statistical calculations, and agents based on artificial intelligence - emphasizing their different logics, methodological advantages, and inherent limitations. The central problem under consideration is defined by the critical divergence between traditional analytical tools and the scale, complexity, and heterogeneity of modern data. Systematizing the historical and conceptual trajectory of these tools, the study explains why each previous paradigm has become insufficient and determines how artificial intelligence agents are not just a technological innovation, but a new way of computational thinking. Particular attention is paid to the profound transformation in the researcher's role, whose responsibilities evolve from manual data manipulation (Paradigm I) to algorithmic construction and modeling (Paradigm II), and culminate in the era of Artificial Intelligence (Paradigm III) in the functions of critical supervision and ethical oversight. The article presents a comparative paradigmatic analysis, identifies the risks associated with generative analytics, and proposes a hybrid methodological approach that integrates classical statistical accuracy with automation based on artificial intelligence. The results contribute to the creation of a more coordinated system for the responsible and transparent use of artificial intelligence systems in scientific research, which is relevant for data science, health care, public policy, and other information-intensive industries.*

Keywords: *data analysis evolution, analytical paradigms, electronic tables, statistical calculations; agents of artificial intelligence, methodological accuracy.*

Introduction.

The history of scientific progress is inseparable from the evolution of the tools with which humans quantify, model, and interpret reality. From the first measuring instruments to modern computing infrastructures, each scientific era has been defined by its dominant analytical tools. As Shapin (2018) points out, scientific objectivity is not an inherent property of research, but is shaped by the methodological and technological systems that mediate observation and inference. Over the past four decades, these systems have undergone an unprecedented transformation. The growth of big data, characterized by increasing volume, velocity, and variety (Laney, 2001; Gandomi & Haider, 2015), has exposed the limitations of traditional analytical workflows and accelerated the transition to more automated, scalable solutions.

Spreadsheet-based analysis, long valued for its accessibility and flexibility, faces challenges in scalability, reproducibility, and error control (Panko, 2005). Statistical computing environments such as R, SAS, and Python have introduced algorithmic rigor, modular pipelines, and greater methodological transparency, but they also require significant expertise and manual coding. At the same time, the growth of unstructured, multidimensional data - from biomedical records to digital communications - has pushed even advanced statistical systems to their conceptual limits (Kitchin, 2014).

To address this gap the analytical tools developed to the next step - the large language models (LLMs) and autonomous AI agents. These systems allow for the creation of natural language workflows for data cleaning, modeling, interpretation, and reporting, marking a conceptual departure from manually constructed analytical pipelines (Tumanov, 2025a). Although the technical literature emphasizes their productivity and versatility (Brown et al., 2020; Bommasani et al., 2021), methodological issues such as hallucinations, opacity, and statistical bias remain under-theorized (Ji et al., 2020; Tumanov, 2025b).

This paper proposes a structured three-paradigm model of data analysis that encompasses Paradigm I: Spreadsheets, Paradigm II: Statistical Computing, and Paradigm III: LLM-Driven AI Agents. By systematizing the conceptual evolution of these paradigms, the article seeks to clarify why previous tools have become ineffective in the era of big data and to lay the foundation for understanding the epistemological and methodological challenges that arise in AI-based analysis.

1. Paradigm I. Spreadsheet

The first paradigm of data analysis emerged with the widespread adoption of spreadsheet systems such as VisiCalc in 1979 and Microsoft Excel in the late 1980s. These tools democratized computing by allowing non-programmers to manipulate, visualize, and summarize data in an intuitive grid-based interface (Grossman, 2007). Spreadsheets have become the dominant analytical medium in business, science, and government due to their availability, low cost, and flexibility. The overall workflow and inherent limitations of spreadsheet-based analytics are illustrated in Figure 1.

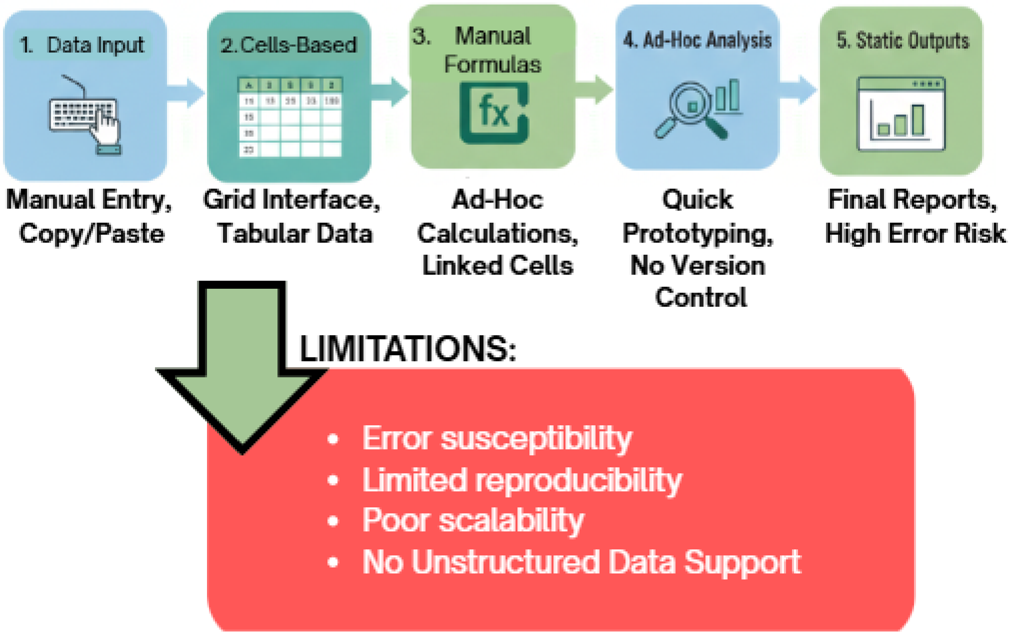

Figure 1 - Paradigm I: Spreadsheet-Based Analysis

Authoring

A schematic representation of a typical spreadsheet workflow highlighting the key steps - manual data entry, cell-based structures, formula-based calculation, ad hoc analysis, and static outputs such as error-proneness, limited reproducibility, low scalability, and lack of support.

Methodologically, spreadsheets introduced a new cognitive model of analysis: cell-level data representation, where formulas, transformations, and visualizations coexist in a single interface. This "end-user computing paradigm" enabled rapid prototyping and ad-hoc analyses without specialized statistical expertise (Nardi & Miller, 1990). However, this same flexibility created structural vulnerabilities. Empirical studies show consistently high error rates in spreadsheet models - often due to manual entry, copy-paste logic, or hidden cell dependencies - undermining reproducibility and transparency (Panko, 2018).

As data volume and complexity increased, spreadsheets faced architectural limitations. Their performance deteriorates with large datasets, they lack version control, and they do not provide systematic support for unstructured or streaming data. These limitations became particularly apparent with the expansion of large-scale data, where modern analytical tasks exceeded the scalability and methodological rigor of spreadsheets.

Thus, although Paradigm I has shaped analytical practices for over three decades and remains widely used, its inherent limitations - proneness to error, limited reproducibility, and low scalability - created methodological pressures that led to the emergence of Paradigm II: Statistical Computing.

2. Paradigm II: Statistical Computing

The limitations of spreadsheet-based analysis led to the emergence of statistical computing tools such as SAS, SPSS, R, Stata, and later Python-based ecosystems. Beginning in the late 1980s and accelerating throughout the 1990s and 2000s, these tools marked a decisive shift from cell-level manipulation to script-based, algorithmic, and reproducible computations (Chambers, 1998). Statistical computing provided a methodological foundation that allowed analysts to work with large datasets, implement formal statistical models, and document their workflows using code.

Unlike spreadsheets, statistical environments introduced modularity, version control, and automation. Scripts allowed the same analysis to be repeated on different datasets and over time, providing methodological transparency and reducing human error. This paradigm also expanded the analytical repertoire: regression models, generalized linear models, survival analysis, multivariate methods, resampling methods, and Bayesian inference became widely available (Venables & Ripley, 2002). The training requirements were higher than in Paradigm I, but the resulting analytical rigor was significantly greater. Statistical computing also corresponded to the early era of big data. Figure 2 illustrates the workflow within this second paradigm, from large, diverse datasets to algorithmic models (R and Python) and scalable extensions. Despite its advantages, this approach is limited by bottlenecks associated with multidimensional data, complex multiplatform workflows, and the need for specialized programming skills.

Tools such as R and Python were extended by open source ecosystems (NumPy, SciPy, pandas, tidyverse), allowing for the processing of millions of records and the integration of machine learning methods (McKinney, 2010). However, even this paradigm faced structural bottlenecks as the volume of data grew exponentially.

Multidimensional text, image, genomic, and streaming data required architectural capabilities beyond the limits of classical statistics.

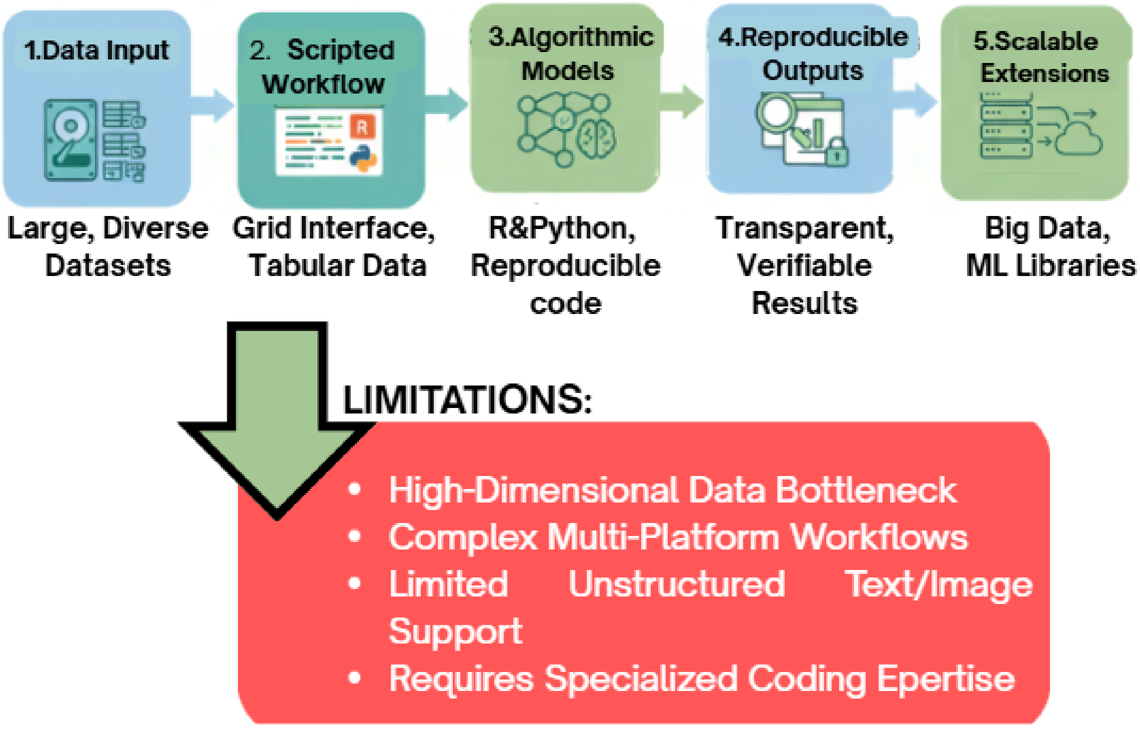

Figure 2 - Paradigm II: Statistical Computing

Authoring

Moreover, modern workflows became fragmented across multiple languages, environments, and pipelines, increasing the complexity and cognitive burden on researchers (Blei & Smyth, 2017).

Thus, while Paradigm II laid the foundation for reproducible, statistical-based analysis, the exponential expansion of multimodal data and the need for integrated natural language interfaces revealed its limitations. These factors, in turn, laid the foundation for the development of Paradigm III: AI agents controlled by LLMs.

3. Paradigm III: LLM-Driven AI Agents

The third paradigm marks a transformational shift from manual, code-driven analytical work to natural language-mediated computations powered by large-language models (LLMs). Unlike traditional tools that require explicit formulas, scripts, or predefined statistical procedures, LLM-driven agents operate using generative reasoning: they interpret goals, infer analytical steps, generate code, perform computations, and autonomously summarize results. This capability positions AI agents not as tools in the classical sense, but as semi-autonomous workers capable of making complex decisions within analytical pipelines (Bommasani et al., 2022).

The Large-language models fundamentally change the locus of analytical practice. Instead of finishing calculations, researchers now formulate questions, check data generated, and evaluate the results based on statistical, ethical and methodological grounds. Figure 3 presents a new analytical framework (Paradigm III) that starts with a natural language query, includes code generation and execution by an AI agent, and concludes with human validation of the results and bias control. However, this approach is severely limited by risks such as hallucinations, fabricated references, methodological errors, and opacity of the decision-making process by the AI agent.

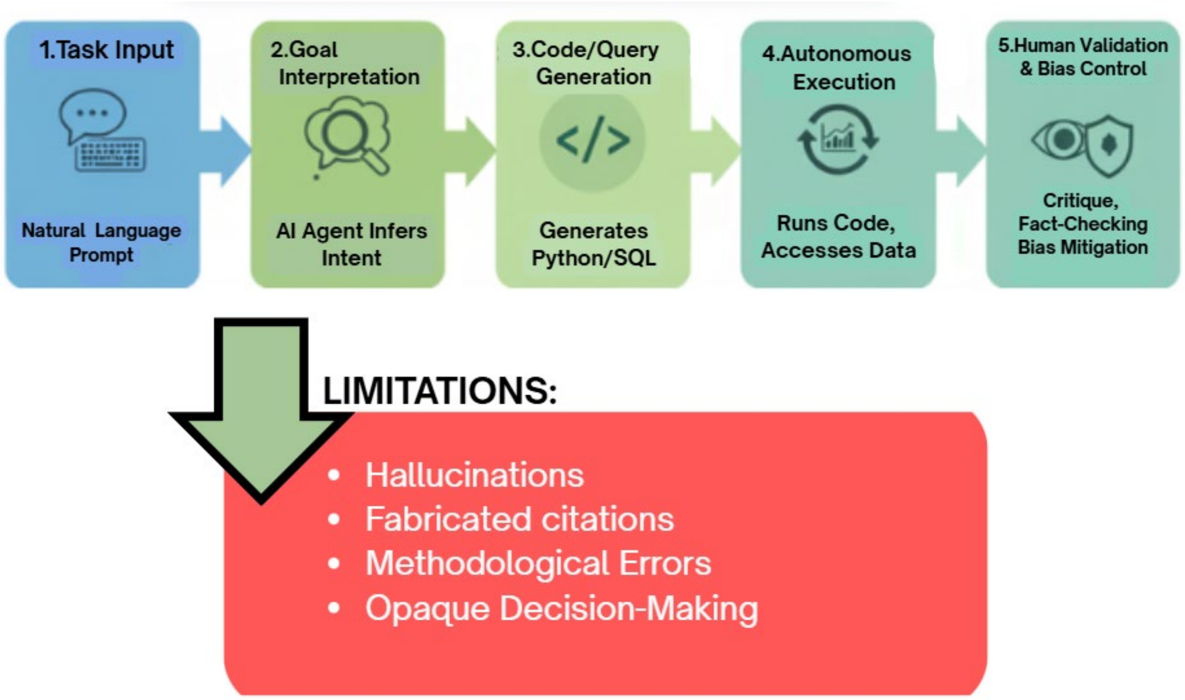

Figure 3 - Paradigm III: LLM-Driven AI Agents

Authoring

As Tumanov (2025a) asserts, the paradigm shift lies not only in automation, but also in the epistemic inheritance of the delegation of analytical stages to models, the internal mechanisms of which are non-transparent and probabilistic. The literature and statistical engineering literature also expects that the LLM may introduce new findings that arise from initial data that may expand at the time of analysis. they should not be criticized (Bender et al., 2021; Tumanov, 2025b).

From a methodological perspective, the third paradigm offers unprecedented advantages: speed, scalability, multimodal integration, and the ability to synthesize diverse data sources. AI agents can generate statistical models, perform diagnostics, extract context from the literature, and dynamically adjust workflows. However, these advantages coexist with new forms of epistemic risk. Law professors can distort numerical results (Ji et al., 2023), falsify citations (Borji, 2023), or incorrectly apply statistical procedures, producing results that appear plausible but lack methodological validity. Moreover, a natural language interface risks hiding the technical complexity of analytical operations. This raises concerns about reproducibility, transparency, and verification of statistical assumptions—the cornerstones of scientific integrity. As Suresh and Guttag (2021) point out, opaque computing systems require explicit constraints to prevent the reinforcement of existing social or structural biases.

Paradigm III is thus not simply an advanced toolkit; it introduces a new epistemological model in which human control becomes central. The researcher evolves from coder to supervisor, auditor, and ethical validator, responsible for ensuring that AI-generated results meet statistical rigor and scientific standards. The emergence of this paradigm emphasizes the need for hybrid methodologies that integrate generative intelligence with explicit classical analytical principles.

The evolution from spreadsheets to statistical computing and, subsequently, to artificial intelligence agents represents not a linear technological improvement, but a significant conceptual transformation. Each paradigm embodies a distinct analytical logic, epistemic stance, and risk structure. Table 1 summarizes the key differences between the three paradigms.

This comparison illustrates the variable nature of analytical responsibility. Paradigm I focused on manual logic construction in a flexible environment, but with limited methodological oversight. Paradigm II enhanced accuracy and reproducibility by formalizing analytical workflows with code, statistical theory, and modular pipelines. Paradigm III breaks with both traditions: analysis becomes a dialogue between the researcher and the model, in which computational thinking is partially externalized to an autonomous agent. This increases analytical speed and breadth, but introduces epistemic uncertainty, making observation central to the scientific process.

Table 1. Evolution of analytical paradigms by key dimensions

Dimension	Paradigm I: Spreadsheet	Paradigm II: Statistical computing	Paradigm III: AI Agents
Analytical logic	Manual, cell-based operations	Algorithmic, rule-based modeling	Generative, reasoning-like inference
Reproducibility	Low	High	Variable; depends on constraints and oversight
Cognitive demand	Low	High (programming, modeling theory)	High in oversight, low in coding
Primary risk	Manual errors, hidden dependencies	Model mis-specification, coding errors	Bias, hallucination, opacity
Researcher role	Operator	Programmer - methodologist	Supervisor - validator - ethicist

Authoring

Conclusion.

Have considered the evolution of data analytical tools across three paradigms, demonstrating that the transition from spreadsheets to statistical computing and ultimately AI agents is a profound conceptual transformation rather than a simple technological development. Each paradigm embodies a distinct analytical logic, epistemological stance, and risk structure, fundamentally shaping the role and responsibilities of the researcher.

Spreadsheets democratized data processing but imposed structural constraints on reproducibility and methodological rigor. Statistical computing introduced formal algorithms, code-based workflows, and reproducible pipelines, but at the same time required extensive computational expertise and remained constrained by the increasing complexity of modern data ecosystems. AI agents, in turn, externalize the components of analytical thinking, enabling unprecedented scalability and automation, while simultaneously creating new risks related to bias, hallucinations, opacity, and reduced traceability.

The comparative analysis highlights that the role of the researcher has shifted from operator (Paradigm I) to programmer-methodologist (Paradigm II), and now to manager, validator, and ethical controller (Paradigm III). This transformation reflects a broader epistemic reorientation: analytical powers are increasingly shared between humans and machines, requiring new forms of oversight, critical analysis, and methodological governance.

The article argues that sustainable progress in data science requires the development of hybrid analytical methodologies that combine the generative power of AI agents with the fundamental rigor of classical statistical thinking. This approach meets modern demands for transparency, reproducibility, fairness, and ethical responsibility - principles that define scientific integrity in the era of AI-mediated research.

References

1. Bender, E. M., Gebru, T., McMillan-Major, A., and Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21). New York, NY, USA: Association for Computing Machinery, pp. 610–623. doi: 10.1145/3442188.3445922.

2. Blei, D. M. and Smyth, P. (2017). Science and data science. Proceedings of the National Academy of Sciences of the United States of America, 114(33), pp. 8689–8692. doi: 10.1073/pnas.1702076114.

3. Bommasani, R., Hudson, D.A., Ross, A. et al. (2022) On the Opportunities and Risks of Foundation Models. Working paper, arXiv:2108.07258. Available at: https://arxiv.org/abs/2108.07258 (Accessed: 11 December 2025).

4. Borji, A. (2023). A Categorical Archive of ChatGPT Failures. Research

Square. Available at: https://doi.org/10.21203/rs.3.rs-2895792/v1.

5. Brown, T., Mann, B., Ryder, N., Subbiah, M., Amodei, D., Brown, A., Clark, J., Askell, S., Chen, S., Child, R., Ziegler, D., Wu, J., Winter, C., Tittel, S., Poli, J., and Radford, A. (2020). Language Models are Few-Shot Learners. In: Advances in Neural Information Processing Systems 33 (NeurIPS 2020). doi: 10.48550/arXiv.2005.14165.

6. Chambers, J.M. (1998) Programming with Data: A Guide to the S Language. New York: Springer.

7. Gandomi, A. and Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management, 35(2), pp. 137–144.

8. Grossman, T. (2007) 'Spreadsheet Engineering: A Research Framework', From the Proceedings of the European Spreadsheet Risks Interest Group Symposium, Cardiff, Wales, July 2002. Available at: https://doi.org/10.48550/arXiv.0711.0538 (Accessed: 11 December 2025).

9. Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y.J., Madotto, A. and Fung, P. (2023) Survey of Hallucination in Natural Language Generation, ACM Computing Surveys, 55(12), pp. 1–38. Available at: https://doi.org/10.1145/3571730 (Accessed: 11 December 2025).

10. Kitchin, R. (2014). Big Data, New Epistemologies and Paradigm Shift. Big Data & Society, 1, pp. 1–12. doi: 10.1177/2053951714528481.

11. Laney, D. (2001). 3D Data Management: Controlling Data Volume, Velocity and Variety. Stamford, CT: META Group.

12. McKinney, W. (2010). Data Structures for Statistical Computing in Python. In: Proceedings of the 9th Python in Science Conference (SciPy 2010). Austin, Texas: SciPy, pp. 56–61. doi: 10.25080/Majora-92bf1922-00a.

13. Nardi, B. and Miller, J. (1990) 'The spreadsheet interface: A basis for end-user programming', Communications of the ACM, 33(12), pp. 72–81.

14. Panko, R. (2005). What We Know About Spreadsheet Errors. Journal of End User Computing, 10. doi: 10.4018/joeuc.1998040102.

15. Shapin, S. (2018). A Social History of Truth: Civility and Science in

Seventeenth-Century England. Chicago: University of Chicago Press.

16. Suresh, H. and Guttag, J. (2021). A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle. In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21). New York, NY, USA: Association for Computing Machinery, pp. 1–13. doi: 10.1145/3442188.3445924.

17. Tumanov, O. (2025a). Large Language Models: A New Paradigm for Data Analysis. Business Inform, 9, pp. 106–113. doi: 10.32983/2222-4459-2025-9-106-113.

18. Tumanov, O. (2025b). The Unseen Data: A Statistical and Engineering Perspective on Biases in Large Language Models. SWorld Journal, 33(1), pp. 178–183. doi: 10.30888/2663-5712.2025-33-01-078.

19. Venables, W. N. and Ripley, B. D. (2002). Modern Applied Statistics With S. New York: Springer. doi: 10.1007/b97626.

Article sent: 19.12.2025

Tumanov O.O.