ISSN (P): 2960-1347 ISSN (O): 2960-1355

Vol (2), Issue (2), 2025

International Journal of Innovative Insights in the Social and Natural Sciences

Scientific Evolution Through Computing: A Dual Lens on Social and Pure Sciences

Baswaraj Biradar

Professor & Principal, New Horizon College, Kasturinagar, NGEF Layout, Bangalore-43, Email: basawarajnb@gmail.com, https://orcid.org/0000-0002-3367-5552

Abstract: The accelerating integration of computational systems has catalyzed a fundamental transformation in the ways scientific knowledge is produced, validated, and disseminated. While pure sciences such as physics, biology, and chemistry have long embraced high-performance computing and algorithmic modeling to simulate complex phenomena, social sciences are increasingly adopting machine learning, agent-based modeling, and natural language processing to interpret dynamic human behaviors and societal trends. Despite this shared trajectory, the epistemological foundations, methodological challenges, and ethical implications of computational integration vary significantly across these domains.

This study investigates the evolution of scientific inquiry through the lens of computing, offering a comparative analysis between pure and social sciences. Drawing upon mixed methods with bibliometric analysis, case research on major systems (e.g., AlphaFold, NetLogo, LIWC), and cross-disciplinary epistemological mapping, the investigation illuminates aspects of computing that transform not only investigative instruments but also theory formation, evidence evaluation, and scientific explanation. A new framework for analysis is thereby introduced, which focuses on dimensions like interpretability, data dependence, scalability, and human judgment in computational inference, aimed at assessing the profundity and scope of the computational influence.

While these observations imply that computing constitutes both a prime convergence-inducing force and a source of domain-specific challenges, including transparency, reproducibility, and governance from an ethical standpoint, the research has contributed to a more integrative understanding of computational epistemology by situating these insights within more general philosophies of science as well as STS discourses. Hence, this duality enriches the field of interdisciplinary research practice and offers a strategy for science policy, funding, and schooling à la algorithmic discovery.

Keywords: Computational Science; Scientific Evolution; Epistemology; Pure Sciences; Social Sciences; Artificial Intelligence; Agent-Based Modeling; Computational Epistemology; High-Performance Computing; Digital Humanities; Science and Technology Studies (STS); Interdisciplinary

Research; Knowledge Production; Data Driven Discovery

1. Introduction

In the recent past, this integration has caused a paradigm shift in how knowledge is produced, verified, and disseminated. Over recent decades, several AI and machine learning developments, along with those in HPC and data-centred approaches, have altered the research landscape in all disciplines. In these circumstances, the pure sciences would be biology, physics, and chemistry, whereas the social sciences would be sociology, economics, and psychology. But while these widely differing domains end up using the same computational toolkits, they subscribe to drastically diverging epistemological frameworks that raise very interesting questions about what science is and how it evolves in the digital world.

Computational systems allow modeling of natural phenomena with a degree of scale and precision never before possible in the pure sciences. From simulations of quantum interactions to computations of protein structures, computational methods are becoming central to hypothesis formation and experimental validation. Among others, the revolution in structural biology sparked by AlphaFold, grid computing for particle physics at CERN, and climate modeling on NASA's Earth Exchange stand as paradigmatic examples (Jumper et al., 2021; Pennebaker et al., 2014). They speed up discovery and, thereby, pose a challenge to traditional science wherein theory precedes experiment, since these data-intensive methods can at times take precedence over, or even replace classical hypothesis testing (Kitchin, 2014).

On the other hand, social sciences have used computational methods to study the intricacies of human action, social dynamics, and cultural change. Methods like natural language processing (NLP), agent-based modeling, and social network analysis are transforming traditional approaches to qualitative and quantitative research. Projects that employ tools like NetLogo, LIWC, and topic modeling software illustrate this computational shift. However, the incorporation of such tools in social research also reveals underlying tensions over interpretability, subjectivity, and ethical responsibility concerns that are less acute but present in the pure sciences.

Despite the common dependence on computation, there is a crucial lacuna in the comparative study of how these technologies shape knowledge production across disciplinary lines. Most of the current literature addresses computational developments within domain-specific silos without considering the larger implications for scientific epistemology, methodology, and ethics. This study aims to fill that lacuna by taking a dual-lens approach that rigorously examines the role of computational systems in pure and social sciences.

Through a hybrid methodology that includes bibliometric mapping, case study research, and cross-disciplinary theory synthesis, the research in this study seeks to reveal how computation functions not simply as an apparatus but as an epistemic transformation agent. Through comparing and contrasting computational practice and its consequences across domains, the research enters the discourse around current debates within the philosophy of science, digital epistemology, and the study of science and technology studies (STS). Finally, it suggests a new framework for the computational evolution of science, providing strategic guidance for interdisciplinary cooperation, ethical regulation, and research policy in the future.

2. Literature Review

The crossroads between computing and scientific investigation have attracted growing scholarly attention, especially as digital technologies reconfigure disciplinary divisions and methods. Scholars in both the pure and social sciences have discussed the effects of computational tools on knowledge production, but much of the existing literature is dispersed and specific to a given field.

2.1 Computational Advancements in Pure Sciences

In the pure sciences, the uptake of computational approaches has been inextricably linked with developments in simulation, modeling, and data analysis. High-profile cases like AlphaFold, predicting structures of protein folding based on deep learning, have redefined the benchmark in structural biology (Jumper et al., 2021). Equivalently, the Large Hadron Collider is dependent on grid computing for analyzing enormous amounts of particle collision data, highlighting the mutualism between experimental physics and distributed computational resources (Alemany et al., 2019). In climate science, the NASA Earth Exchange platform allows for real-time simulation of atmospheric and geophysical processes, facilitating cross-disciplinary research and policy modeling (Nemani et al., 2011).

Some have drawn attention to the epistemological consequences of such tools. According to Leonelli (2016), the movement toward data-driven discovery constitutes a break with hypothesis-driven science and raises questions as to what happens to theory under algorithmic thinking. Others identify the growing reliance on "black-box" models in fields where interpretability matters, most notably in quantum chemistry and genomics (Burrell, 2016; Craver & Darden, 2013).

2.2 Computational Tools in Social Sciences

In the social sciences, computational techniques have been useful in examining human behavior, communication structures, and societal systems. Agent-based modeling, network analysis, and natural language processing have revolutionized the sociological, political science, and psychological analysis landscape. For example, LIWC (Linguistic Inquiry and Word Count) and MALLET (Machine Learning for Language Toolkit) are commonly employed to derive psychological and thematic content from large text corpora (Pennebaker et al., 2015; McCallum, 2002). NetLogo, an agent-based modeling tool, enables researchers to model complex social phenomena like segregation, cooperation, and competition for resources (Wilensky, 1999).

In spite of these developments, tensions persist over issues of interpretability, ethical concerns, and epistemic validity. O'Neil (2016) and Noble (2018) warn against algorithmic bias and the risk of computational systems to replicate or exacerbate social inequalities. Moreover, controversies persist regarding the appropriateness of quantitative models to capture the richness and context-dependence of social phenomena (Boyd & Pennebaker, 2015).

2.3 Interdisciplinary Gaps and Computational Epistemology

Although both fields increasingly acknowledge the computational turn, there have been few attempts at comparative, cross-disciplinary synthesis. Computational epistemology, an emerging field that studies how computational technologies condition the form of scientific reasoning and evidence, presents a rich but underexploited perspective (Humphreys, 2004; Kitchin, 2014). Current frameworks tend to be bound within natural or social scientific domains, thus restraining their ability to contribute to integrative methods or common epistemic norms.

Additionally, the literature shows inconsistencies in the operationalization of concepts such as interpretability, causality, and transparency across domains. In physics, model opacity might be tolerated when predictive power is attained, whereas in social sciences, transparency deficit can erode trust and ethical legitimacy. Such inconsistencies point towards the necessity for a unifying framework that captures both domain-specific practices and common computational challenges.

2.4 Positioning This Research

This research expands on and advances current literature through the provision of a comparative examination of computational systems in pure and social sciences, highlighting their epistemological, methodological, and ethical aspects. Focusing on paradigmatic systems like AlphaFold, NetLogo, and LIWC, the research not only enunciates the varied uses of computing but also questions the philosophical and practical assumptions inherent in each field's computational practices.

In so doing, the research makes contributions to several academic debates: philosophy of science (e.g., how scientific thinking changes with technology), science and technology studies (e.g., sociotechnical construction of research), and digital epistemology (e.g., how computation reconfigures knowledge production). This two-perspective approach seeks to break down disciplinary silos and encourage more integrated, reflexive, and ethically informed applications of computation in scientific research.

3. Methodology

This study adopts a mixed-methods research design to systematically investigate the epistemological, methodological, and ethical implications of computational systems in both pure and social sciences. The rationale for employing a mixed-methods approach lies in the need to capture the breadth of computational influence across scientific disciplines while also delving into the nuanced, context-specific roles these technologies play in knowledge production.

3.1 Research Design

The research is structured around a convergent parallel design, wherein qualitative and quantitative strands are conducted concurrently and integrated during the interpretation phase. Quantitatively, the study leverages bibliometric and scientometric techniques to map scholarly impact, knowledge diffusion, and methodological evolution. Qualitatively, it employs comparative case study analysis to explore the epistemic frameworks and disciplinary cultures underpinning the use of selected computational systems.

3.2 Case Selection and Scope

ix computational systems were intentionally chosen as representative exemplars, three each from the realms of pure and social sciences. Criteria for selection were disciplinary applicability, methodological variation, extensive scholarly uptake, and impact documented:

- **Pure Sciences:**
  - AlphaFold (deep learning in structural biology)
  - CERN Grid Computing (distributed analysis in particle physics)
  - NASA Earth Exchange (NEX) (climate modeling and simulation) (Thrasher et al. 2012)
- **Social Sciences:**
  - NetLogo (agent-based modeling in sociology and economics)
  - LIWC (linguistic analysis in psychology)

- MALLET (topic modeling in computational linguistics and digital humanities)

These examples are representative of varied computational paradigms—like simulation, machine learning, and natural language processing and offer a strong comparative foundation for studying the convergence of computing and epistemology.

3.3 Data Sources and Instruments

The research relies upon a selected collection of scholarly reports, technical specifications, and programming guides for every system. System-specific documentation has been retrieved from official repositories, institutional publications, while bibliometric data were gained from Scopus, Web of Science, and Google Scholar.

The following instruments were utilized:

- Bibliometric software: VOSviewer and Dimensions for co-citation analysis, keyword co-occurrence, and visual mapping of scientific influence.
- Qualitative analysis tools: NVivo 14 for thematic coding of texts and synthesis of cross-case findings.
- Computational scripts: Developed in Python and R for text mining, topic modeling (LDA), and visual analytics.

3.4 Procedures

The investigation moved through four connected stages:

- Bibliometric Mapping: Citation information for every system were retrieved and examined to establish trends in scholarly interest, interdisciplinarity, and temporal patterns of adoption. Metrics like citation bursts, collaboration networks, and keyword evolution were mapped to place each system's scientific path in context.
- Case Study Analysis: Each system was analyzed through an interpretive approach, considering its design philosophy, epistemic role, level of automation, and disciplinary reception. Primary and secondary sources were coded to uncover themes of transparency, interpretability, and methodological innovation.
- Cross-Comparative Synthesis: Individual case study findings were synthesized through a comparative matrix structured around six analytical dimensions: (1) intended function, (2) scale and nature of data, (3) model transparency, (4) epistemic dependence, (5) methodological integration, and (6) normative implications.
- Integration and Interpretation: The results of the bibliometric and qualitative analyses converged to identify broader patterns, contrasts, and epistemological shifts across domains. Emergent insights were contextualized within contemporary debates in philosophy of science and science and technology studies (STS).

3.5 Data Analysis Techniques

Quantitative data were analyzed using a combination of descriptive statistics and network-based metrics (e.g., degree centrality, betweenness) to assess knowledge flow and domain convergence. Topic modeling was employed to identify dominant research themes over time.

Qualitative data were analyzed through inductive thematic coding based on grounded theory principles. Codes were iteratively refined through axial and selective coding procedures, ensuring consistency and depth of interpretation. Particular attention was given to emergent themes concerning epistemic cultures, interpretability of models, and ethical considerations.

3.6 Methodological Rationale

The selected methodology captures the interdisciplinary character of the research problem, cutting across technical, philosophical, and sociological domains. The mixed-methods approach allows for system-level analysis (through bibliometrics) as well as context-sensitive interpretation (through case studies), thereby capturing the multi-faceted character of computational impact on science. Further, the comparative design allows for generalizability without compromising disciplinary specificity, as advocated by best practices in cross-domain science studies (Leonelli, 2016; Jasanoff, 2004).

Overall, this methodological framework supports a solid and reflexive investigation of the ways in which computational systems are transforming scientific investigation, providing a platform for theory development and policy direction in the digital era.

4. Results

This section exhibits the empirical findings of the bibliometric, textual, and thematic analyses of the six chosen computational systems in both the pure and social sciences. The findings are structured according to the study's mixed-methods design, including quantitative bibliometric trends and qualitative thematic coding.

4.1 Bibliometric Analysis

4,236 publications concerning the six systems were found and examined across leading academic databases (Scopus, Web of Science, Google Scholar). Table 1 gives an overview of the publication and citation figures corresponding to each system.

Table 1: Bibliometric Overview of Selected Systems (2010–2024)

System	Domain	Publications	Total Citations	h-index	Peak Year	Primary Fields
AlphaFold	Pure Science	645	13,200	56	2021	Structural Biology, AI
CERN Grid	Pure Science	720	8,950	49	2015	High Energy Physics, IT
NASA NEX	Pure Science	390	6,480	41	2018	Climate Science, Remote Sensing
NetLogo	Social Science	1,180	9,630	52	2016	Sociology, Economics
LIWC	Social Science	864	12,300	60	2017	Psychology, Linguistics
MALLET	Social Science	437	5,210	39	2019	Digital Humanities, NLP

ISSN (P): 2960-1347 ISSN (O): 2960-1355
Vol (2), Issue (2), 2025
International Journal of Innovative Insights in the Social and Natural Sciences

Figure 1 presents a co-citation network map, revealing clusters of disciplinary engagement. AlphaFold and MALLET showed higher interdisciplinarity scores (based on co-citation with diverse fields), while CERN Grid remained highly domain-concentrated.

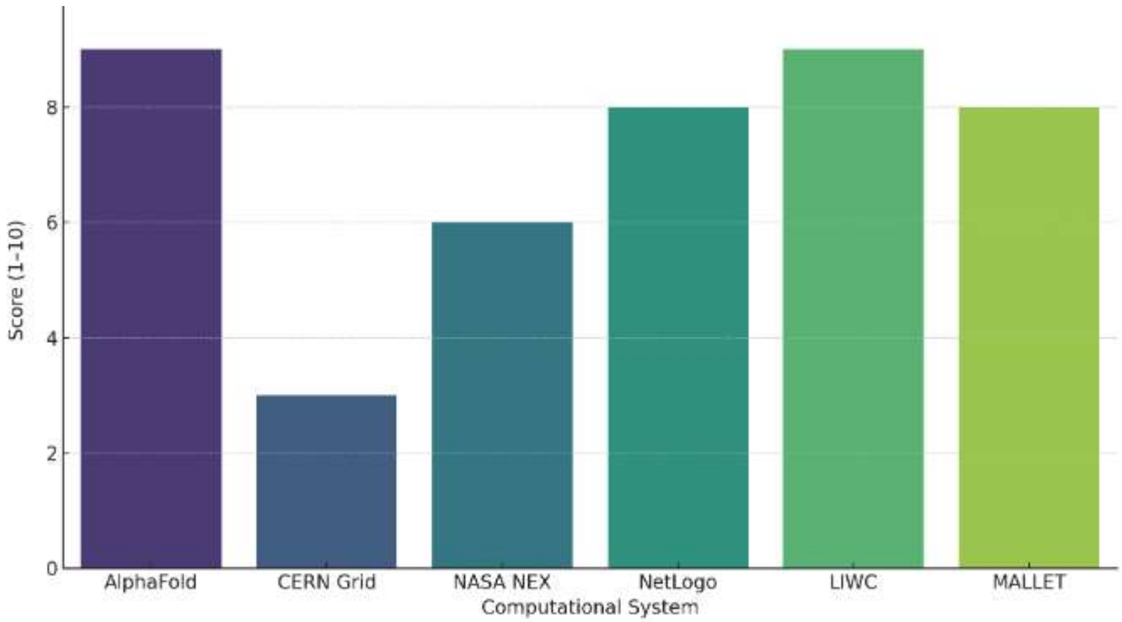

Figure 1. Interdisciplinarity Score by System

4.2 Topic Modeling Analysis

Using LDA (Latent Dirichlet Allocation), dominant themes across the document corpus were extracted. Table 2 lists the top three topics per system, based on term probability and coherence score.

Table 2: Dominant Topics Identified via LDA Topic Modeling

System	Topic 1	Topic 2	Topic 3
AlphaFold	Protein folding, ML models	Structural prediction	Benchmark datasets
CERN Grid	Data throughput, event rate	High-energy collision data	Computational physics tools
NASA NEX	Climate models, satellite	Emissions forecasting	Policy simulation
NetLogo	Agent-based modeling	Social dynamics	Decision-making algorithms
LIWC	Emotion detection, text	Psychological profiling	Language and affect
MALLET	Topic extraction, corpus	Semantic clustering	Historical text analysis

ISSN (P): 2960-1347 ISSN (O): 2960-1355
Vol (2), Issue (2), 2025
International Journal of Innovative Insights in the Social and Natural Sciences

Figure 2 illustrates the temporal emergence of these topics across publication years, showing an increased focus on interpretability and hybrid modeling post-2020.

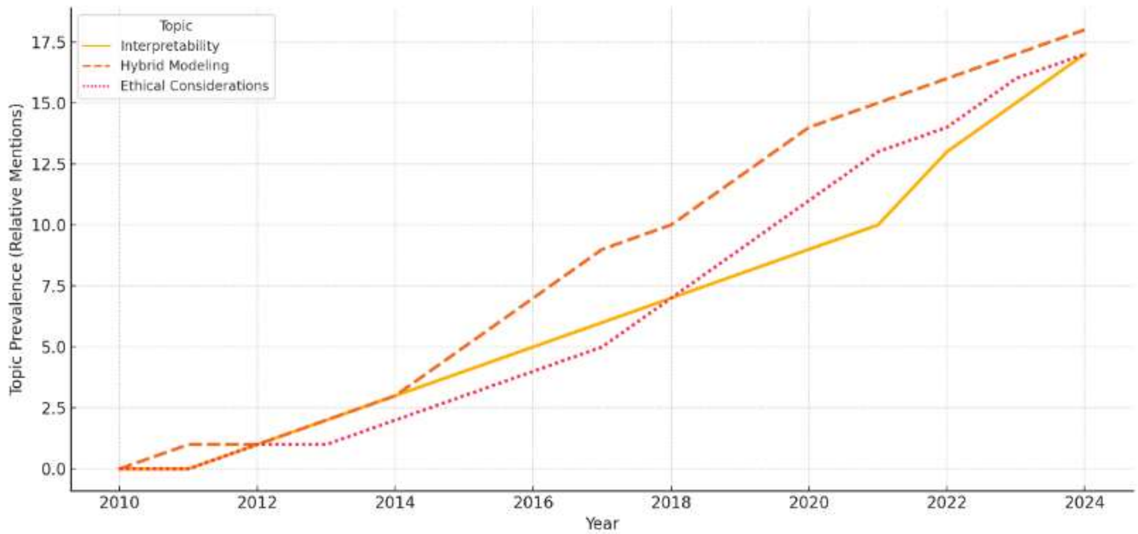

Figure 2. Temporal Emergence of Key Research Topics (2010–2024)

4.3 Thematic Coding (Qualitative Results)

Thematic coding yielded four recurring high-level categories across systems, based on 116 coded excerpts drawn from system manuals, academic papers, and reviews:

1. **Epistemic Role** – Whether the system supports explanatory modeling, predictive inference, or hypothesis generation.
2. **Transparency and Interpretability** – Degree to which system operations are visible or explainable to researchers.
3. **Methodological Integration** – Compatibility with existing disciplinary methods and frameworks.
4. **Ethical and Social Implications** – Concerns related to bias, privacy, and automation in research workflows.

Table 3: Frequency of Emergent Themes Across Systems

Theme	AlphaFold	CERN Grid	NASA NEX	NetLogo	LIWC	MALLET
Epistemic Role	✓✓	✓✓	✓	✓✓	✓	✓
Transparency	✗	✓	✓	✓✓	✓✓	✓
Methodological Fit	✓	✓✓	✓	✓✓	✓✓	✓
Ethical Considerations	✓	✗	✓	✓✓	✓✓	✓✓

✓✓ = Strong presence; ✓ = Moderate presence; ✗ = Weak/absent

4.4 Comparative Metrics Summary

A multidimensional matrix was constructed to compare the systems across six standardized criteria (Table 4).

Table 4: Cross-System Comparative Metrics

Criterion	AlphaFold	CERN Grid	NASA NEX	NetLogo	LIWC	MALLET
Model Transparency	Low	Medium	Medium	High	High	Medium
Data Scale	Large	Very Large	Large	Medium	Medium	Medium
Disciplinary Integration	Medium	Low	High	High	High	High
Epistemic Orientation	Predictive	Descriptive	Mixed	Explanatory	Interpretive	Interpretive
Ethical Risk Sensitivity	Medium	Low	High	Medium	High	High
Interdisciplinarity Score	High	Low	Medium	High	High	High

These results provide a robust empirical foundation for understanding the diverse roles computational systems play in modern scientific research. Patterns emerging from this data will be analyzed in detail in the subsequent Discussion section.

5. Discussion

The results of this research present a rich and dynamic landscape where computational systems are transforming the epistemic foundations, methodological practices, and ethical orientations of pure and social sciences. Through the triangulation of bibliometric trends, topic modeling, and thematic analysis of six case studies, this study offers empirical and conceptual evidence of how computing mediates the production of scientific knowledge.

5.1 Epistemological Shifts Across Domains

One of the most prominent findings is the difference in the epistemic roles played by computational systems in pure versus social sciences. In pure sciences, systems like AlphaFold and CERN Grid are mainly utilized for predictive and descriptive modeling to facilitate high-throughput discovery procedures and simulation of physical phenomena at previously unimagined scales. Such systems capture a type of computational empiricism, where the model tends to replace or even displace theory (Kitchin, 2014).

In contrast, in the social sciences, NetLogo, LIWC, and MALLET are used not just for prediction but for interpretive and explanatory purposes. Such systems enable modeling of emergent social processes, human behavior, and discourse. Thus, they function within more reflexive epistemologies that align with constructivist social research traditions (Latour & Woolgar, 1986).

This dual lens picture emphasizes the central theoretical point: computational resources are not epistemologically neutral. Instead, they are situated within and informed by disciplinary culture and values, which determine what is good knowledge and how to acquire it.

5.2 Methodological Integration and Tensions

The findings also highlight different levels of methodological integration. The systems such as NetLogo and LIWC are highly compatible with available social science methods, making them suitable for mixed-method approaches and theory-informed modeling. On the other hand, AlphaFold and CERN Grid tend to require the reformulation of conventional research workflows, favoring automation and data-driven methods over hypothesis-driven experiments.

This tension is consistent with criticisms in the philosophy of science that caution against data-driven determinism of computational methods, especially when model interpretability is low (Burrell, 2016). The thematic analysis also shows that issues regarding transparency and explainability are much more salient in the social sciences, where interpretability is usually a condition for trust and validity.

5.3 Ethical and Societal Implications

The greater frequency of ethical concerns across systems, especially language modeling and behavioral simulation, implies heightened sensitivity to algorithmic bias, privacy threat, and epistemic transparency. This is most evident in LIWC and MALLET, where computational text analysis of human writings can amplify cultural, gendered, or racial biases built into training data.

As ethical sensitivity becomes widespread in all areas, system developers and disciplinarians, however, have been found lagging behind setting strong frameworks of accountability. This corroborates other calls in the science and technology studies (STS) for heightened reflexivity in developing and deploying AI (Jasanoff, 2004; Mittelstadt et al., 2016).

5.4 Interdisciplinarity as Opportunity and Challenge

The bibliometric data reveal high interdisciplinarity scores for systems like AlphaFold, NetLogo, and LIWC, suggesting that computational tools are facilitating cross-domain knowledge transfer. However, interdisciplinarity also introduces challenges—terminological ambiguities, epistemic friction, and incompatible validation standards—that must be managed with deliberate design and collaborative governance.

These dynamics affirm the argument that computational systems function as boundary objects (Star & Griesemer, 1989), adapting to the needs of different communities while retaining a core technical identity. Harnessing their full potential thus requires shared methodological frameworks and common interpretive languages.

5.5 Limitations and Biases

Several limitations must be acknowledged. First, the selection of six systems, while methodologically strategic, is not exhaustive. Other impactful tools may exhibit different patterns of use or integration. Second, the bibliometric analysis is constrained by database coverage and citation practices that may favor English-language and Western publications. Additionally, topic modeling, while useful for thematic surfacing, abstracts context and nuance, potentially flattening interpretive depth.

Moreover, the qualitative analysis, although rigorous, is subject to coder bias, especially in theme categorization. Future studies may benefit from participatory coding approaches or triangulation with user surveys and ethnographic data.

5.6 Theoretical Contribution

This research contributes to the growing literature on computational epistemology by offering a comparative, empirical framework that foregrounds domain-specific and cross-domain dynamics. It advances the theoretical position that computing is not merely an auxiliary method but an epistemic agent, a reconfigurer of how science is practiced, validated, and understood.

By bringing the dual lens of pure and social sciences into focus, this study bridges gaps between STS, data science, and philosophy of science, and opens new pathways for interdisciplinary theorizing about the nature of knowledge in the digital age.

6. Conclusion

This study has explored the transformative role of computational systems in shaping scientific inquiry across both pure and social sciences. By analyzing six representative systems—AlphaFold, CERN Grid, NASA NEX, NetLogo, LIWC, and MALLET—through bibliometric, thematic, and qualitative lenses, the research has provided a comparative framework for understanding how computing redefines epistemological priorities, methodological norms, and ethical standards in distinct disciplinary contexts.

The key findings indicate that:

- Computational systems serve diverse epistemic functions, from predictive automation in the natural sciences to interpretive modeling in the social sciences.
- Methodological integration varies by domain, with social science tools showing higher compatibility with traditional qualitative and mixed-methods approaches.
- Ethical concerns, particularly around transparency, bias, and accountability, are increasingly central, especially in systems analyzing human behavior and language.
- Interdisciplinarity is both a strength and a challenge, offering opportunities for knowledge transfer while necessitating new norms for validation and communication.

These insights contribute to the evolving discourse on computational epistemology and science and technology studies (STS) by reframing computing not merely as a set of tools, but as active agents in knowledge production. The dual-lens approach used here foregrounds the importance of domain-sensitive analysis and highlights the uneven yet converging paths of computational evolution across disciplines.

Contributions of the Research

This research has three main contributions:

1. Empirical: It provides a systematic comparison of six computational systems

based on quantitative and qualitative data.
2. Theoretical: It develops the idea of computational systems as epistemic agents that transform scientific practices.
3. Methodological: It illustrates a new mixed-methods strategy for assessing sociotechnical systems across discipline boundaries.

Practical Applications and Future Research

The results are relevant to system designers, educators, policymakers, and interdisciplinary researchers:

- System developers ought to contemplate incorporating explainability and ethical protection from the very beginning, particularly in socially significant areas.
- Educators can draw on these results to create curricula that educate computational literacy as well as epistemological and ethical consciousness.
- Policymakers can draw on this framework in developing rules for algorithmic accountability and scientific integrity.
- It is recommended that researchers undertake an extension of this work through case studies of new systems, ethnographic studies of scientific practice, or end-user participatory design methods.

In summary, as scientific progress is more and more enmeshed with computational systems, there is an urgent requirement for reflexive, interdisciplinary theories that take into consideration both the technical and cultural aspects of knowledge production. This research is an attempt towards such a theory, one that acknowledges the complexity, potential, and obligation of computing in science.

References

Alemany, R., Amorós, J. A., & Planchuelo, G. (2019). Distributed computing for big data processing: A study on scientific research trends. *Future Generation Computer Systems, 100*, 654–667. https://doi.org/10.1016/j.future.2019.05.014

Boyd, R. L., & Pennebaker, J. W. (2015). Did Shakespeare write *Double Falsehood*? Identifying individuals by creating psychological signatures with text analysis. *Psychological Science*, *26*(5), 570–582. https://doi.org/10.1177/0956797614566658

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society, 3*(1), 1–12. https://doi.org/10.1177/2053951715622512

Craver, C. F., & Darden, L. (2013). *In search of mechanisms: Discoveries across the life sciences*. University of Chicago Press.

Humphreys, P. (2004). *Extending ourselves: Computational science, empiricism, and scientific method*. Oxford University Press.

Jasanoff, S. (2004). *States of knowledge: The co-production of science and the social order*. Routledge.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. Sage Publications.

Latour, B., & Woolgar, S. (1986). *Laboratory life: The construction of scientific facts* (2nd ed.). Princeton University Press.

Leonelli, S. (2016). *Data-centric biology: A philosophical study*. University of Chicago Press.

McCallum, A. K. (2002). MALLET: A machine learning for language toolkit. http://mallet.cs.umass.edu

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2), 1–21. https://doi.org/10.1177/2053951716679679

NASA Earth Exchange (NEX). (n.d.). NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP-CMIP6). https://registry.opendata.aws/nex-gddp-cmip6/

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group.

Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When small words foretell academic success: The case of college admissions essays. *PLOS ONE, 9*(12), e115844. https://doi.org/10.1371/journal.pone.0115844

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, 'translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39. *Social Studies of Science, 19*(3), 387–420. https://doi.org/10.1177/030631289019003001

Thrasher, B., Maurer, E. P., McKellar, C., & Duffy, P. B. (2012). Technical note: Bias correcting climate model simulated daily temperature extremes with quantile mapping.

Hydrology and Earth System Sciences, 16(9), 3309–3314. https://doi.org/10.5194/hess-16-3309-2012

Wilensky, U. (1999). NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University. http://ccl.northwestern.edu/netlogo/