EDITORIAL MAR CARIBE

SCIENTIFIC RESEARCH METHODOLOGY APPLIED TO ARTIFICIAL INTELLIGENCE AND DATA SCIENCE:

GENERAL APPROACH

JOSEFINA ARIMATEA GARCÍA CRUZ
MARIA MAURA CAMAC TIZA
CARLOS ALBERTO ASIÁN QUIÑONES
JOSÉ CÉSAR PIEDRA ISUSQUI
LUIS ALBERTO JUSCAMAYTA RAMÍREZ
JACKELINE ROXANA HUAMAN FERNANDEZ
JUDITH DEL ROSARIO ROSAS CAJAHUANCA

ISBN 978-9915-698-43-4

Scientific research methodology applied to artificial intelligence and data science: General approach

García Cruz, Josefina Arimatea; Camac Tiza, Maria Maura; Asián Quiñones, Carlos Alberto; Piedra Isusqui, José César; Juscamayta Ramírez, Luis Alberto; Huaman Fernandez, Jackeline Roxana; Rosas Cajahuanca, Judith del Rosario

First edition (1st ed.): October, 2025

Edited by:

Editorial Mar Caribe ®
www.editorialmarcaribe.es
Av. Gral. Flores 547, 70000 Col. del Sacramento, Department of Colonia, Uruguay.

Cover design and illustrations: *Isbelia Salazar Morote*

E-book available at:
https://editorialmarcaribe.es/ark:/10951/isbn.9789915698434

Format: Electronic

ISBN: 978-9915-698-43-4

ARK: ark:/10951/isbn.9789915698434

Editorial Mar Caribe (OASPA): As a member of the Open Access Scholarly Publishing Association, we support open access in accordance with OASPA's code of conduct, transparency, and best practices for the publication of scholarly and research books. We are committed to the highest editorial standards in ethics and deontology, under the premise of "Open Science in Latin America and the Caribbean"

Editorial Mar Caribe, signatory No. 795 of 12.08.2024 of the Declaration of Berlin
"... We feel compelled to address the challenges of the Internet as an emerging functional medium for the distribution of knowledge. Obviously, these advances can significantly modify the nature of scientific publishing, as well as the current quality assurance system...." (Max Planck Society, ed. 2003., pp. 152-153).

Editorial Mar Caribe adheres to the *UNESCO "Recommendation concerning the Preservation of and Access to Documentary Heritage, including Digital Heritage"* and the International Reference Standard for an Open Archival Information System (OAIS-ISO 14721). This book is digitally preserved by ARAMEO.NET

Editorial Mar Caribe

Scientific research methodology applied to artificial intelligence and data science: General approach

Colonia, Uruguay

2025

Scientific research methodology applied to artificial intelligence and data science: General approach

Index

Index of Tables

Dedication

“… To the researchers of tomorrow, who will inherit these tools to decipher the mysteries that we can barely glimpse today. May the Fifth Paradigm be the compass that guides your curiosity towards some knowledge without borders. And, for those who look for patterns in chaos and beauty in data..."

Introduction

The history of science has gone through various paradigms: from the empirical observation of natural phenomena and the theoretical formulation of laws to the computational simulation of complex systems. Today, we are immersed in what Jim Gray called the "fourth paradigm": *data-intensive scientific discovery*. In this new scenario, Artificial Intelligence (AI) and Data Science are not mere technical engineering tools; they have become the fundamental lens through which we interrogate reality.

However, the dizzying advance of these technologies has brought with it an epistemological challenge: the gap between predictive capacity and scientific validity. In the race to optimize hyperparameters and reduce mean square error, it is often forgotten that an AI model, in a research context, is not just a software product, but a mathematically formalized hypothesis that must be tested, disproved, and explained.

There is a latent tension in current practice. On the one hand, software engineering seeks to make the system "work"; on the other, the methodology of research requires understanding "why it works" and under what conditions it is reproducible.

This book, *Scientific research methodology applied to artificial intelligence and data science: General approach*, was born from the imperative need to systematize research in this field. It seeks to answer critical questions that many students and professionals face: How do you formulate a valid research question in a Machine Learning project?, What differentiates a technology implementation project from a scientific research thesis?, How do we address the reproducibility crisis in

"Black Box" models?, What are the ethical and bias standards that should govern data collection?.

The central objective of this text is to serve as a bridge that connects the rigor of the traditional scientific method—with its emphasis on hypothesis, controlled experimentation, and causal inference—with the flexibility and power of modern Data Science workflows (CRISP-DM, KDD, etc.).

It is not a programming manual in Python or R, nor a compendium of neural network architectures. It is a guide to scientific thinking applied to data. Here, the reader will learn how to structure big data chaos within a methodological framework that ensures robust, generalizable, and ethically responsible findings.

Through the four chapters, we will break down the AI research lifecycle from a methodological perspective:

- *Epistemological Foundations:* We will explore the nature of knowledge generated by inductive and deductive algorithms.
- *Research Design:* Definition of scope, selection of variables and the dichotomy between explanatory and predictive studies.
- *Data Management as Evidence:* Processing, cleansing, and the importance of data quality beyond volume.
- *Validation and Metrics:* Going beyond accuracy as the only metric; sensitivity analysis, robustness, and rigorous cross-validation.
- *Ethics and Communication:* The Researcher's Responsibility in the Face of Algorithmic Bias and How to Write Technical Findings for a Scientific Audience.

Ultimately, this book proposes that the best Artificial Intelligence is not the one that simply processes data faster, but the one that helps us better understand the world with greater rigor and truth. Welcome to the study of scientific methodology in the age of data.

Chapter 1.

The Fifth Paradigm: Redefining Scientific Inquiry through Artificial Intelligence and Big Data Methodologies

1. The Epistemological Shift in Research Methodology

The history of scientific discovery is often categorized into distinct paradigms, each defined by the primary instrument of inquiry. The first paradigm was empirical, driven by direct observation of natural phenomena. The second was theoretical, characterized by the formulation of laws and generalizations, such as Maxwell's equations or Newton's laws of motion. The third, emerging in the mid-20th century, was computational, utilizing simulations to model complex systems that were analytically intractable. The fourth paradigm, articulated by Jim Gray in the early 2000s, was data-intensive, predicated on the analysis of massive datasets generated by instruments and sensors. Today, we stand at the precipice of a "Fifth Paradigm".[1] This new era is not merely an extension of the data-intensive fourth paradigm; it represents a fundamental qualitative shift driven by the integration of Artificial Intelligence (AI) and Big Data into the very fabric of the scientific method.

In this Fifth Paradigm, the relationship between the researcher and the object of study is mediated—and in some cases, fully managed—by intelligent algorithms. We are witnessing the transition from "human-generated hypotheses tested on data" to "AI-generated hypotheses validation by humans".[3] The epistemological implications are profound. Traditional research methodologies, grounded in the scarcity of data and the necessity of sampling, are being upended by an abundance of data and the capability for "total population analysis".[4] The linear progression of the scientific method—observation, hypothesis, experimentation, analysis, conclusion—is being compressed into iterative, high-velocity loops executed by autonomous agents.[6]

This report provides an exhaustive analysis of these emerging methodologies. It explores how Large Language Models (LLMs) are revolutionizing literature synthesis, transforming it from a manual bottleneck into an automated, semantic reasoning process. It examines the "Causal Revolution," which seeks to move AI beyond mere correlation to the understanding of cause-and-effect relationships essential for policy and scientific intervention.[8] It investigates the rise of "Generative Social Science," where synthetic populations mimic human behavior, allowing for in silico sociological experiments.[10] Finally, it addresses the critical challenges of this new era: the reproducibility crisis exacerbated by data leakage, the ethical imperatives of algorithmic fairness, and the changing role of the human scientist in a loop increasingly dominated by silicon intelligence.

2. The Revolution in Literature Review and Knowledge Synthesis

The initial phase of any research project—the literature review—has traditionally been a labor-intensive process of manual discovery, reading, and synthesis. The exponential growth of scientific publishing, with millions of papers published annually, has made comprehensive manual review nearly impossible. The integration of AI, specifically through semantic search and Retrieval-Augmented Generation (RAG), has fundamentally altered this landscape, transitioning the researcher's role from a consumer of text to an architect of inquiry.

2.1 From Lexical Search to Semantic Reasoning

For decades, information retrieval in academia relied on lexical or keyword-based search. This methodology is inherently brittle; a search for "cardiovascular disease" might miss relevant papers that strictly use the term "heart failure" unless complex Boolean strings are constructed. The Fifth Paradigm introduces **Semantic Search**, powered by vector embeddings. In this model, text is converted into high-dimensional vectors, and retrieval is based on the proximity of these vectors in semantic space rather than exact word matching.[11]

Semantic Scholar, a pioneer in this field, utilizes machine learning to understand the *context* of citations. It distinguishes between a citation that merely mentions a prior work and one that is foundational to the methodology, classifying them as "highly influential citations".[12] This moves the metric of scientific impact from raw counts to semantic relevance. Furthermore, tools like Elicit and Consensus leverage LLMs to perform "automated extraction." Upon receiving a natural language query (e.g., "What is the impact of continuous glucose monitors on long-term diabetes complications?"), these systems do not just return a list of links. They scan the full text of relevant papers, extract the specific findings, sample sizes, and methodologies, and present them in a structured matrix.[13]

This capability represents a methodological leap: the instantaneous generation of a "Matrix of Evidence." A task that previously required weeks of manual coding—extracting P, I, C, O (Population, Intervention, Comparison, Outcome) elements from dozens of papers—can now be performed in minutes (see Table 1).

Table 1: AI-Enhanced Research Platforms

Feature	**Traditional Bibliographic Databases (e.g., Web of Science, PubMed)**	**AI-Enhanced Research Platforms (e.g., Elicit, Semantic Scholar, Scite)**
Search Mechanism	Boolean Logic / Keyword Matching	Vector-Space Semantic Retrieval
Unit of Retrieval	Document (Title/Abstract)	Specific Answer / Claim / Data Point
Citation Context	Count only	Classification (Supporting, Contrasting, Mentioning)

Synthesis	Manual by Researcher	Automated "TL;DR", Summaries, & Comparison Tables
Handling Synonyms	Requires manual "OR" strings	Native understanding of conceptual similarity

2.2 Automated Systematic Reviews and Screening Efficiency

The systematic review is the gold standard of evidence synthesis, particularly in medicine and the social sciences. However, it is resource-intensive, often requiring teams of reviewers to screen thousands of titles and abstracts to identify the few dozen that meet inclusion criteria. New methodologies employ LLMs to automate this screening phase, acting as a "second screener" or even a primary filter.

Recent studies comparing LLM performance to human reviewers have demonstrated transformative results. In one comprehensive evaluation, an LLM-assisted workflow reduced the total screening time from 54.7 hours (using traditional semi-automated software like Rayyan) to just 25.5 hours—a time saving of over 50%.[15] More strikingly, when the LLM was used for the initial title/abstract screening, it achieved a Negative Predictive Value (NPV) of 100%, meaning it successfully excluded irrelevant studies without discarding a single relevant one.[15] Other research indicates that LLM-assisted screening can reduce manual effort by up to 95% in specific contexts.[16]

The methodology for systematic reviews is thus evolving from a "dual-human" process to a "Human-AI" hybrid model. In this workflow, the AI serves as a high-recall filter, flagging potential studies and providing rationales for exclusion, while the human expert focuses on the final inclusion decisions and the subtle interpretation of complex findings. This shift allows researchers to conduct "living systematic reviews"—reviews that are continuously updated as new literature is published—rather than static snapshots that become outdated within months.

2.3 Knowledge Graphs and GraphRAG: Solving the Hallucination Problem

A critical limitation of pure LLMs in research is "hallucination"—the generation of plausible but fictitious text, citations, or data points. In scientific writing, a hallucination rate of even 1% is unacceptable. To mitigate this, advanced methodologies are moving beyond simple "chat" interfaces to Retrieval-Augmented Generation (RAG) systems grounded in Knowledge Graphs (KGs).[17]

A Knowledge Graph is a structured representation of facts, where entities (e.g., "Protein A," "Algorithm B," "Author C") are nodes and their relationships (e.g., "Upregulates," "Developed by," "Cites") are edges. The integration of KGs with LLMs, known as GraphRAG, allows for multi-hop reasoning.

Consider a research question: "What is the connection between Gene X and Disease Z?" A standard LLM might hallucinate a connection based on statistical probability in its training data. A GraphRAG system, however, will traverse the knowledge graph:

1. Paper A links Gene X to Enzyme Y.
2. Paper B links Enzyme Y to Disease Z.
3. The system infers a potential pathway: Gene X -> Enzyme Y -> Disease Z.

The LLM then generates a response citing Paper A and Paper B as the evidence for this inferred link.[19] This methodology transforms the literature review from a retrieval task to a logic discovery task, allowing researchers to uncover latent connections in the scientific corpus that no single paper has explicitly articulated.[20]

2.4 The Risk of Bibliographic Simulacra and Algorithmic Literacy

Despite these advances, the "hallucination rate" (HR) remains a critical methodological variable. Research evaluating tools like GPT-4 and Claude 2 in scientific contexts mandates the reporting of HR and "Prompt Sensitivity" (PS).[21] While models have improved, they still occasionally fabricate citations or misattribute findings, particularly in niche fields where training data is sparse.[22]

Furthermore, the ease of generating summaries poses a risk of "bibliographic simulacra"—a state where researchers interact primarily with AI-generated syntheses rather than the primary texts. This can lead to the propagation of subtle misinterpretations or the loss of nuance. The Fifth Paradigm demands that researchers develop algorithmic literacy: the ability to audit AI outputs, verify citations against primary sources, and understand the provenance of the information presented.[24] We are moving toward a methodology of "Trust but Verify," where AI acceleration is balanced by rigorous human validation.

3. Big Data Methodologies: Beyond Sampling and Significance

The traditional research paradigm was constrained by the high cost and logistical difficulty of data collection. This constraint necessitated the development of sampling theory—the mathematical framework for inferring the properties of a whole population from a small, representative part. Big Data, characterized by the "Three Vs" (Volume, Velocity, Variety) and now expanded to dimensions including Veracity, Value, and Variability [25], fundamentally challenges the necessity of sampling.

3.1 "N=All": Total Population Analysis

In the digital age, researchers often have access to the entire population of interest—every transaction in a financial market, every click on a website, or every genomic sequence in a biobank. This shift to "N=All" renders traditional notions of statistical significance (p-values) less relevant.[4]

When analyzing a total population, the concept of sampling error—the error introduced by observing only a subset—vanishes. Differences observed in the data are, by definition, real differences in that population. The methodological challenge shifts from calculating sampling error to managing measurement error and model error. As noted by Mayer-Schönberger and Cukier, "N=All" allows researchers to embrace the messiness of real-world data rather than curating pristine, small samples.[26]

However, "N=All" is not a panacea. The phenomenon of "Big Data Hubris"—the assumption that massive datasets automatically yield truth—remains a risk. The failure of Google Flu Trends, which attempted to predict flu outbreaks based on search queries (N=All searches), demonstrated that a massive dataset can still be biased if the underlying data generation process (user search behavior) drifts over time or is influenced by media coverage.[5] Thus, the new methodology emphasizes representativeness over volume. A dataset of 100 million tweets is an "N=All" of Twitter users, but it is not an "N=All" of the global population. Researchers must now account for algorithmic selection bias inherent in the platforms generating the data.[27]

3.2 Real-Time Streaming Analytics: Lambda and Kappa Architectures

Traditional research is retrospective: data is collected over a period, cleaned, and then analyzed in batches. The "Velocity" of Big Data necessitates **streaming analytics**, where data is processed in motion. This is particularly transformative in fields like public health (epidemic tracking), sociology (sentiment analysis), and finance.[28]

Methodologies utilizing Apache Kafka, Flink, or Spark Streaming allow researchers to analyze data streams in real-time.[29] This requires a shift in data architecture.

- **Lambda Architecture:** A hybrid approach that uses a "speed layer" for real-time views and a "batch layer" for comprehensive, high-latency analysis. This allows researchers to see immediate trends while correcting for errors later.[28]
- **Kappa Architecture:** A simplified model that treats everything as a stream, allowing for a single processing framework.

In this context, the research "instrument" is no longer a static survey but a dynamic algorithm. Anomaly detection algorithms run continuously on these streams, identifying outliers (e.g., a sudden spike in emergency room visits) that would be smoothed out in aggregate batch analysis.[30] This requires defining "windows" of analysis—sliding windows or tumbling windows—rather than static cross-sections.

3.3 The Integration of Unstructured Data via Multimodal Analysis

The "Variety" of Big Data refers to the influx of unstructured data—images, video, free text, and sensor logs. Traditional methodology required this data to be manually coded before analysis. AI, particularly Deep Learning (CNNs and Transformers), allows for the direct analysis of unstructured data.[31]

For example, in migration studies, researchers traditionally relied on lagged government statistics. New methodologies integrate Call Detail Records (CDRs) from mobile phones, satellite imagery indicating settlement growth, and social media activity to map migration flows in real-time.[33] This triangulation of disparate data sources—multi-modal analysis—provides a holistic view but introduces complex ontological challenges. How does a researcher reconcile the "truth" of a satellite pixel with the "truth" of a tweet? The methodology requires a robust framework for data fusion, assigning confidence weights to different modalities based on their known reliability and bias profiles.

4. Generative Social Science and Synthetic Populations

One of the most avant-garde developments in research methodology is the use of Generative AI to simulate human subjects. This field, termed **Generative Social Science**, leverages the capabilities of LLMs to create "silicon subjects" or "generative agents" that mimic human attitudes, behaviors, and social interactions.[10]

4.1 Generative Agents: The Silicon Subject

Traditional social science relies on surveys and human-subject experiments, which are expensive, slow, and often difficult to replicate (WEIRD—Western, Educated, Industrialized, Rich, Democratic—bias is a known issue). Generative Agents offer a radical alternative. These are computational agents instantiated with specific personas (e.g., "a 35-year-old nurse with two children and conservative political views") and placed in a simulated environment.[35]

Research by Stanford and others has shown that these agents can replicate the responses of real human populations on surveys like the General Social Survey (GSS) with high fidelity, achieving up to 85% correlation with human responses.[36] This allows researchers to perform in silico experiments: testing the impact of a policy intervention, a marketing campaign, or a social rumor on a synthetic population before deploying it in the real world.

The methodology involves "Inverse Generative Social Science" (iGSS). Instead of designing agents with simple, hand-crafted rules (as in traditional Agent-Based Modeling), researchers use evolutionary computing or LLMs to *evolve* agents that can reproduce a known macro-phenomenon.[34] This shifts the focus from "what happens if rules X apply?" to "what micro-rules must exist to explain macro-phenomenon Y?"

4.2 Synthetic Data in Medicine: Balancing Privacy and Utility

Beyond simulation, synthetic data is revolutionizing privacy-sensitive research, particularly in healthcare. Access to patient records is often restricted by regulations like HIPAA or GDPR. Synthetic data generation involves training an AI model (often a Generative Adversarial Network or GAN) on real patient data to learn the statistical distributions and correlations. The model then generates a new, artificial dataset that preserves these statistical properties without containing any real individuals.[38]

This approach allows for Total Population Analysis without privacy compromise. Researchers can share synthetic datasets openly, enabling reproducibility and collaboration that was previously impossible. However, the methodology requires rigorous validation. Researchers must perform:

- **Utility Evaluations:** Does the synthetic data yield the same regression coefficients and predictive power as the real data?
- **Privacy Evaluations:** Is it possible to re-identify real individuals from the synthetic set via linkage attacks?.[40]

Recent studies propose metrics to quantify "privacy leakage" in synthetic datasets, ensuring that the trade-off between data utility and privacy is transparent. This "tiered access" model—open synthetic data for exploration, restricted real data for final validation—is becoming a standard workflow in medical research.[41]

4.3 The "Echo Chamber" Risk and the Limits of Simulation

While promising, Generative Social Science faces the **"Echo Chamber" problem**. LLMs are trained on internet text, which reflects specific biases, cultural norms, and linguistic patterns. A society of generative agents might simply reproduce the stereotypes present in the training data rather than the nuanced reality of human behavior.[27]

There is a risk that researchers might mistake the simulation for the territory. The methodology requires a continuous loop of validation: real-world data calibrates the simulation, the simulation generates hypotheses, and the hypotheses are tested back in the real world. Critics argue that relying too heavily on generative agents could lead to a "positivist" drift in qualitative research, imposing rigid categories on the fluid nature of human experience.[42]

5. The Causal Revolution: Beyond Correlation

For decades, the mantra of statistics has been "correlation does not imply causation."

Traditional machine learning excels at pattern recognition—identifying correlations in high-dimensional space—but often fails to understand *why* those patterns exist. This limitation is critical in fields like medicine and policy, where the goal is intervention, not just prediction. The **Causal AI** movement, spearheaded by Judea Pearl and others, is introducing methodologies that allow researchers to infer causality from observational data.[8]

5.1 Structural Causal Models (SCMs) and Directed Acyclic Graphs (DAGs)

The core of this methodology is the **Structural Causal Model (SCM)**. Unlike a neural network, which is a "black box" of weights, an SCM is explicitly represented by a **Directed Acyclic** Graph (DAG)—a visual map where nodes are variables and arrows represent causal influence.[9]

Researchers use tools like DoWhy (Microsoft), CausalNex, and Pyro to build these models. The methodology involves a distinct two-step process:

1. **Causal Discovery:** Algorithms analyze the data to suggest potential causal structures (e.g., "does X cause Y, or does confounder Z cause both?").
2. **Causal Inference:** The researcher encodes their domain knowledge into the DAG and uses it to estimate the effect of an intervention.[44]

This approach enables counterfactual reasoning—asking "What would have happened if?" questions. For example, in a clinical trial, Causal AI can estimate the outcome for a specific patient *had they received the alternative treatment*, a calculation impossible with standard statistics that only observe the realized outcome.[8]

5.2 Prescriptive vs. Predictive AI

The shift from Predictive AI (what will happen?) to Prescriptive AI (how can we make it happen?) is the hallmark of causal methodologies. In manufacturing, Causal AI is used for Root Cause Analysis, distinguishing between a symptom (correlation) and the actual defect source (causation).[45]

This methodology is crucial for mitigating spurious correlations. In Big Data, the probability of finding a statistically significant but meaningless correlation increases with the size of the dataset (the "Look-elsewhere effect"). Causal AI acts as a filter, rejecting correlations that do not fit the causal structure of the world, thus enhancing the robustness and generalizability of research findings.[46] For instance, an AI might find a correlation between ice cream sales and drowning deaths; a Causal AI model would identify "Temperature" as the confounder causing both, preventing a policy recommendation to ban ice cream to save swimmers.

5.3 Tools for Causal Inquiry

The ecosystem of tools supporting this methodology is maturing rapidly:

- **DoWhy:** A Python library that unifies causal inference under a single API (Model, Identify, Estimate, Refute). It is particularly strong on the "Refute" step, allowing researchers to stress-test their causal assumptions.[44]
- **CausalNex:** Allows researchers to use Bayesian Networks to infer causality, offering a "Human-in-the-loop" interface where experts can manually correct the edges of the learned graph.[47]
- **TensorFlow Causal:** Brings causal inference to the scale of deep learning, allowing for causal discovery on massive datasets.[47]

6. Autonomous Experimentation: The Rise of the AI Scientist

Perhaps the most futuristic development in research methodology is the automation of the scientific process itself. Self-Driving Laboratories (SDLs) and AI Scientists are systems capable of planning, executing, and analyzing experiments with minimal or no human intervention.

6.1 Materials Acceleration Platforms (MAPs)

In materials science and chemistry, the search space for new molecules is practically infinite (estimated at 10^{60} small molecules).[48] Traditional trial-and-error is too slow. Materials Acceleration Platforms (MAPs) combine robotics, AI, and high-throughput screening to close the loop of experimentation.[49]

An SDL operates in cycles, often referred to as "Level 3" autonomy:

1. **Design:** An AI (often using Bayesian Optimization or Active Learning) selects the next best experiment to run based on previous results, balancing exploration (trying new things) and exploitation (refining known hits).
2. **Synthesize:** Robotic arms mix reagents and manage the reaction.
3. **Characterize:** Sensors measure the properties of the new material (e.g., conductivity, absorbance).
4. **Learn:** The results update the AI's model, and the cycle repeats.[49]

These systems can run 24/7, accelerating discovery by orders of magnitude. The methodology shifts the researcher's role from "bench scientist" to "system architect"—designing the search space and the optimization parameters rather than pipetting liquids.[51] Notable examples include setups at NC State and BU, where "dynamic flow experiments" redefine data utilization in fluidic labs.[52]

6.2 "The AI Scientist": Automated Paper Generation

Beyond physical experiments, AI agents are now capable of conducting end-to-end computational research. Sakana AI recently introduced "The AI Scientist," a comprehensive system that can generate novel research ideas, write the necessary code, execute the experiments, visualize the results, and write a full scientific paper—all autonomously.[6]

The workflow of such an agentic system is:

1. **Ideation:** The LLM brainstorms research directions and checks them against existing

literature (via Semantic Scholar) to ensure novelty.

2. **Coding:** An agent (like Aider) writes the experiment code (e.g., Python/PyTorch).
3. **Execution:** The code is run, and logs/results are captured.
4. **Drafting:** The LLM interprets the results and drafts a paper in LaTeX, generating figures and citations.
5. **Reviewing:** An "Automated Reviewer" module critiques the paper, prompting the scientist agent to revise the text or run additional experiments.[53]

While currently limited to computational domains, this challenges the definition of authorship. It raises the prospect of "Recursive Scientific Improvement," where AI systems design better AI systems. However, critics note that these systems currently struggle with deep methodological innovation, often producing derivative or incrementally novel work, and can suffer from "hallucinated" correctness where the code runs but the methodology is flawed.[54]

7. Methodological Integrity and Ethics in the AI Era

The integration of AI and Big Data into research methodology is not a panacea; it introduces systemic risks that, if unaddressed, threaten the integrity of science.

7.1 The Reproducibility Crisis and Data Leakage

Machine Learning-based science is facing a reproducibility crisis. A major driver is Data Leakage—the accidental inclusion of information from the test set into the training process. This leads to overly optimistic performance estimates that fail to generalize to new data.[55]

Methodologists have identified profound taxonomies of leakage, including:

- **Temporal Leakage:** Using future data to predict the past (e.g., using a patient's discharge diagnosis to predict their admission risk).
- **Feature Leakage:** Including variables that are proxies for the target label.

- **Distributional Leakage:** When the test set is not drawn from the same distribution as the deployment environment.

To combat this, new methodologies mandate the use of Model Info Sheets and rigorous Train-Test-Validation splits. The concept of reproducibility is expanding to include "Algorithmic Reproducibility"—can another researcher, using the same code and data, generate the exact same model? This is often difficult due to the stochastic nature of GPU training and random seeds.[55]

7.2 Algorithmic Bias and Ethics

Data is never neutral; it is a social artifact. Algorithmic Bias occurs when AI models amplify existing societal prejudices present in the training data. In medical research, algorithms trained on predominantly white populations have been shown to misdiagnose skin conditions in patients with darker skin tones.[27]

Methodologies for Fair AI are emerging to address this. These include:

- **Pre-processing:** Re-weighting the data to ensure balanced representation.
- **In-processing:** Adding fairness constraints to the loss function of the model.
- **Post-processing:** Adjusting the model's outputs to equalize error rates across demographic groups.[43]

Ethical research in the Fifth Paradigm requires a "Data Nutrition Label" approach—transparently documenting the provenance, composition, and limitations of the datasets used.[57]

8. Human-AI Collaboration Models

As AI systems become more autonomous, the nature of human oversight is evolving. The methodology now explicitly defines the human's role in the loop.

8.1 Human-in-the-Loop (HITL)

In HITL models, the human actively makes decisions at key points. The AI provides a recommendation, but the human must "sign off" for the process to proceed. This is the standard for high-stakes research, such as clinical trials or weapon systems research.[58] For example, in automated systematic reviews, the human is the final arbiter of inclusion.

8.2 Human-on-the-Loop (HOTL)

In HOTL models, the human plays a supervisory role. The system operates autonomously, but the human monitors the performance metrics and intervenes only when anomalies occur or confidence scores drop below a threshold. This is the typical model for Self-Driving Laboratories and high-frequency trading algorithms.[60]

8.3 Human-in-Command (HIC)

HIC emphasizes that while AI may execute tasks, the strategic direction and ethical responsibility remain solely with the human. This model is gaining traction in defense and policy research, ensuring that "agentic" workflows do not drift from their intended human-aligned goals.[58]

The consensus in methodological ethics is that for scientific discovery, HOTL is the minimum standard. The "black box" problem means that even if an AI discovers a truth, human validation is required to integrate that truth into the broader scientific canon.

9. Future Frontiers: Quantum Computing and the 2030 Outlook

Looking ahead to 2025 and beyond, the intersection of Quantum Computing and Big Data promises to shatter current computational limits.

9.1 Quantum Machine Learning (QML)

Quantum computers, utilizing qubits and superposition, can solve optimization problems (like protein folding or supply chain logistics) exponentially faster than classical computers.[62] Quantum Machine Learning (QML) will allow researchers to analyze datasets of a complexity that is currently intractable. For instance, simulating the precise quantum mechanics of a chemical reaction is impossible for classical supercomputers but native to quantum devices. This will likely lead to a "quantum leap" in pharmacology and materials science.[63] By 2025, we expect to see the first "Quantum Advantage" in specific research niches, such as optimizing catalysts for carbon capture.

9.2 The "Fifth Paradigm" by 2030

By 2030, predictions suggest that AI Agents will be proactive research partners. They will not just answer questions but independently monitor the literature, identify gaps, and propose hypotheses.[65] The global market for these AI-driven research tools is projected to reach hundreds of billions of dollars.[65] The definition of a "scientist" will expand to include those who experts in are orchestrating these silicon intelligences—"Prompt Engineers" of the physical world.

The integration of Artificial Intelligence and Big Data into research methodology is not merely an upgrade in tools; it is a fundamental restructuring of how we ask and answer questions. We are moving from a scarcity-based epistemology—defined by sampling, linear regression, and manual synthesis—to an abundance-based epistemology—defined by total population analysis, deep learning, and autonomous generation.

This transition empowers researchers to tackle "wicked problems" of organized complexity—climate change, pandemics, metabolic networks—that were previously beyond reach. The cost of drug discovery could plummet, the speed of materials innovation could skyrocket, and our understanding of social dynamics could become predictive rather than merely descriptive.

However, this potential comes with the responsibility of rigour. The researcher of the future must be a hybrid scholar: part domain expert, part data scientist, and part ethicist. We must guard against the seduction of "easy answers" provided by black-box algorithms. Our methodologies must remain anchored in causal reasoning, robust validation, and an unwavering commitment to human interpretability. As we deploy agents that can read, reason, and experiment, the ultimate goal of research remains unchanged: not just to process data, but to generate meaning (see Table 2).

Table 2: Comparative Analysis of Research Methodologies

Dimension	**Traditional Methodology**	**AI & Big Data Methodology (Fifth Paradigm)**
Data Source	Surveys, Controlled Experiments, Archives	Streaming Sensors, Web Scrapes, Synthetic Data, "N=All"
Analysis Type	Hypothesis Testing (Deductive)	Pattern Recognition & Generative (Inductive/Abductive)
Causality	Randomized Controlled Trials (RCTs)	Structural Causal Models (SCMs), Counterfactuals
Literature Review	Manual search & synthesis	Semantic Search, Automated Extraction, Knowledge Graphs
Sampling	Representative Random Sampling	Total Population Analysis, Bias Mitigation in Data
Role of Technology	Tool for calculation (Passive)	Partner in hypothesis & design (Active/Agentic)
Key Risk	Sampling Error, p-hacking	Algorithmic Bias, Hallucination, Data Leakage

Chapter bibliographies

1. Abisoye, A. (2024). A Conceptual Framework for Integrating Artificial Intelligence into STEM Research Methodologies for Enhanced Innovation. *International Journal of Future Engineering Innovations*, *1*(1). https://doi.org/10.54660/ijfei.2024.1.1.56-66
2. Abu Sharshouh, A. (2025). The Use of Artificial İntelligence in Accounting and Auditing TT - Muhasebe ve Denetim Alanında Yapay Zekâ Kullanımı. *Karadeniz Ekonomi Araştırmaları Dergisi*, *6*(1).
3. Arslantaş, S. (2024). Artificial intelligence and big data from digital health applications: publication trends and analysis. *Journal of Health Organization and Management*. https://doi.org/10.1108/JHOM-06-2024-0241
4. Axiotis, E., Kontogiannis, A., Kalpoutzakis, E., & Giannakopoulos, G. (2021). A personalized machine-learning-enabled method for efficient research in ethnopharmacology. The case of the southern balkans and the coastal zone of asia minor. *Applied Sciences (Switzerland)*, *11*(13). https://doi.org/10.3390/app11135826
5. Benedeti Rosa, M., & Kubota, L. C. (2025). Artificial intelligence: where does Brazil stand in global scientific production and what are the main technical determinants of adoption by Brazilian companies. *EconomiA*. https://doi.org/10.1108/ECON-01-2025-0010
6. Bhardwaj, M., Kumar, P., & Singh, A. (2025). Bibliometric review of digital transformation in agriculture: innovations, trends and sustainable futures. *Journal of Agribusiness in Developing and Emerging Economies*. https://doi.org/10.1108/JADEE-02-2025-0083
7. Budiman, A. S., Putri, D. N. N., Candra, H., Djuana, E., Sari, T. K., Aji, D. P. B., Putri, L. R., Sitepu, E., Speaks, D., & Pasang, T. (2024). Crack Catcher AI – Enabling smart fracture mechanics approaches for damage control of thin silicon cells or wafers. *Solar Energy Materials and Solar Cells*, *273*. https://doi.org/10.1016/j.solmat.2024.112927

8. Cheng, J. C. P., Chen, W., Chen, K., Wang, Q., Sadri, H., Karki, B. R., Basnet, S., Xiang, J., Montoya, J., Porras, J., Ucar, A., Karakose, M., Kırımça, N., Abdelalim, A. M., Essawy, A., Alnaser, A. A., Shibeika, A., Sherif, A., van Roosmale, S., … Liu, G. (2024). applied sciences Artificial Intelligence for Predictive Maintenance Applications : *Journal of Building Engineering*, *12*(1).
9. Colina Vargas, A. M., Espinoza Mina, M. A., Catalán, L. L., & Catalán, B. L. (2024). Computational intelligence, educational robotics, and artificial intelligence in the educational field. A bibliometric study and thematic modelling. *International Journal of Educational Research and Innovation*, *2024*(22). https://doi.org/10.46661/ijeri.10369
10. Cricchio, J. (2025). Balancing openness and ownership: open innovation strategies for AI development. *European Journal of Innovation Management*. https://doi.org/10.1108/EJIM-04-2024-0470
11. Demidenko, S. Y., & Romanovskiy, N. v. (2023). THEORETICAL UNDERSTANDING OF SOCIAL REALITY: THE 25th ANNIVERSARY OF KHARCHEV READINGS. *Sotsiologicheskie Issledovaniya*, *2023*(9). https://doi.org/10.31857/S013216250027774-4
12. Educause. (2018). Horizon report: 2018 higher education edition. Retrieved from EDUCAUSE Learning Initiative and The New Media Consortium. *Liila*.
13. Evolution Analysis and Prediction of Scientific Topic Popularity in the Field of LIS. (2021). *Documentation, Information and Knowledge*, *38*(2). https://doi.org/10.13366/j.dik.2021.02.035
14. Gong, J., Zhao, Z., Niu, X., Ji, Y., Sun, H., Shen, Y., Chen, B., & Wu, B. (2025). AI reshaping life sciences: intelligent transformation, application challenges, and future convergence in neuroscience, biology, and medicine. In *Frontiers in Digital Health* (Vol. 7). https://doi.org/10.3389/fdgth.2025.1666415
15. Govindan, K., Cheng, T. C. E., Mishra, N., Shukla, N., Pei-Chen, L., Arya, V., Sharma, P., Singh, A., de Silva, P. T. M., Chatfield, A. T., Reddick, C. G., Karlsson, P.-M., Lal, S., Rexhausen, D., Ajak, A. D., Lilford, E., Topal, E., Zhang, Y. Y., Fu, Z., … Perez-Franco, R. (2018). Industry Surveys IT Consulting & Other Services. *Journal of Cleaner*

Production, *25*(1).

16. Guo, Z. (2024). Innovative Functionality of the Scientific and Technical Information System of China. *Visnyk of Kharkiv State Academy of Culture*, *65*. https://doi.org/10.31516/2410-5333.065.04
17. Harahulia, S., & Symonenko, T. (2025). RESEARCH SUPPORT SERVICES IN DIGITAL HUMANITIES:APPROACHES, TECHNOLOGIES, TOOLS. *Manuscript and Book Heritage of Ukraine*, *2025*(2). https://doi.org/10.15407/rksu.37.189
18. Irtyshcheva, I., Kim, V., & Kletsov, Y. (2022). MANAGING BUSINESS PROCESSES IN THE CONDITIONS OF THE DEVELOPMENT OF THE DIGITAL ECONOMY: GLOBAL AND NATIONAL EXPERIENCES. *Baltic Journal of Economic Studies*, *8*(5). https://doi.org/10.30525/2256-0742/2022-8-5-101-107
19. Iurasova, O., Ivashko, L., Maksymova, J., & Maksymov, O. (2025). INNOVATION TRENDS: IMPACT ON DEVELOPMENT OF FINANCIAL SECTOR. *Смарт-Економіка, Підприємництво Та Безпека*, *3*(1). https://doi.org/10.60022/sis.3.(01).1
20. Jeon, Y. (2021). A Study on the Ethics of the Concept and Purpose of Culture Technology in the Post-COVID Era. *Society for International Cultural Institute*, *14*(1). https://doi.org/10.34223/jic.2021.14.1.1
21. Jirandehi, A. B., & Sajadi, S. M. (2024). Co-Citation Analysis and Co-Occurrence Analysis of Resilience Keywords in StartUps. *Scientometrics Research Journal*, *10*(2). https://doi.org/10.22070/rsci.2024.17648.1667
22. Khabsa, M., Elmagarmid, A., Ilyas, I., Hammady, H., Ouzzani, M., Kouznetsov, A., Japkowicz, N., Visser, E., Felizardo, K. K. R., Souza, S. R. S., Maldonado, J. C. J. C. J. C. J. C., Salleh, N., Martins, R. M., Mendes, E. E., MacDonell, S. G., Maldonado, J. C. J. C. J. C. J. C. J. C. J. C., Li, D., Wang, Z. Z., Wang, L., … Demner-Fushman, D. (2020). A New Method to Enhance. *Systematic Reviews*, *9*(1).
23. Khmelkov, A. (2023). INSTITUTIONAL DESIGN OF ANALYTICAL TOOLS OF THE FINANCIAL CONTROL INSTITUTION. *Social Economics*, *66*. https://doi.org/10.26565/2524-2547-2023-66-08
24. Kim, T. Y., Gang, J. Y., Oh, H. J., Ahmad, K., JianMing, Z., Rafi, M., Allen, G. G.,

Travis, T. A., Ramirez, C., ع. و. ،. م. غ. س, کوچکی., Kyrillidou, M., Manifesto, L. S., Johnson, L., Smith, R., Willis, H., Levine, A., Haywood, K., Vice-chancellor, S. D., Vice-chancellor, D. V.-S. D., … Briggs, T. (2019). Using Data Visualization and Dashboards to Support the Decision Making of Public Library Branch Managers. *Library Hi Tech*, *38*(1).

25. Lausiri, N., & Nahda, K. (2022). Pengaruh Rasio Likuiditas, Leverage Dan Profitabilitas Terhadap Kondisi Financial Distress Dengan Kepemilikan Manajerial Sebagai Variabel Moderasi. *Selekta Manajemen: Jurnal Mahasiswa & Manajemen*, *1*(5).
26. Litynska, V. (2025). MODERN METHODS OF CONDUCTING MARKETING RESEARCH. *Social Economics*, *70*. https://doi.org/10.26565/2524-2547-2025-70-06
27. Luo, J., Zhuo, W., Liu, S., & Xu, B. (2024). The Optimization of Carbon Emission Prediction in Low Carbon Energy Economy under Big Data. *IEEE Access*, *12*. https://doi.org/10.1109/ACCESS.2024.3351468
28. Mekhovych, K. (2025). TRANSFORMATION OF THE THEORY OF INNOVATIVE MARKETING EFFICIENCY IN THE CONDITIONS OF AN UNDETERMINED FUTURE. *Energy Saving. Power Engineering. Energy Audit.*, *1 (204)*. https://doi.org/10.20998/2313-8890.2025.01.07
29. Mi, S., Dong, C. Q., & Liu, Y. (2025). Data insights and talent portraits of science and technology talents: Principles and applications. *Studies in Science of Science*, *43*(6).
30. Morozova, I. A., Smetanin, A. S., & Smetanina, A. I. (2024). Management of Business Competitiveness Based on Artificial Intelligence and Big Data for Its Sustainable Development. *Journal of Modern Competition*, *18*(1). https://doi.org/10.37791/2687-0657-2024-18-1-29-40
31. Mosleh, G., Hemmati, S., Mohagheghzadeh, A., & Zarshenas, M. M. (2023). Integrating Systems Biology and Artificial Intelligence in Traditional Pharmacy Research: Advancements, Challenges, and Opportunities. *Trends in Pharmaceutical Sciences*, *9*(3).
32. Mossalanejad, A. (2025). Artificial Intelligence Policymaking and the Power Equation Priorities. *Geopolitics Quarterly*, *21*(3).

https://doi.org/10.22034/igq.2025.511096.2023

33. Nazari, M., Ansari, M., & Pour, M. J. A. (2025). Bibliographic Analysis and Future Research Directions in the Field of Sustainable and Innovative Business Models. *Scientometrics Research Journal*, *11*(2). https://doi.org/10.22070/rsci.2024.19650.1759
34. Nedoshytko, I., & Patriak, O. (2022). Electronic Document Management and Its Value for Business. *Digital Platform: Information Technologies in Sociocultural Sphere*, *5*(2). https://doi.org/10.31866/2617-796x.5.2.2022.270142
35. Nyomba, Muh. A. (2021). Faktor Yang Berhubungan Dengan Kejadian Infeksi Saluran Pernapasan Akut (ISPA) Pada Balita di Sekitar Wilayah Tempat Pembuangan Akhir (TPA) Sampah Antang Kota Makassar Tahun 2021. *Factors Associated of The Incidence of Acute Respiratory Infection (ARI) in. Universitas Hasanuddin.*, *9*.
36. Pan, C., Bowen, W., Peng, Z., & Shao, L. (2024). Biological network-based exploration of the principles of traditional Chinese medicine. *Kexue Tongbao/Chinese Science Bulletin*, *69*(1). https://doi.org/10.1360/TB-2023-0730
37. Perano, M., Cammarano, A., Varriale, V., del Regno, C., Michelino, F., & Caputo, M. (2023). Embracing supply chain digitalization and unphysicalization to enhance supply chain performance: a conceptual framework. *International Journal of Physical Distribution and Logistics Management*, *53*(5–6). https://doi.org/10.1108/IJPDLM-06-2022-0201
38. Pysmennyi, V., & Rudan, V. (2024). Methods of financial management in local communities under the conditions of the russian-Ukrainian war. *Economic Analysis*, *34(3)*. https://doi.org/10.35774/econa2024.03.351
39. RAHMAN, G. O. H. S. E. Y. T. A. (2020). DETERMINANTS OF FOOD DELIVERY PLATFORMS THROUGH CUSTOMER LOYALTY TOWARD SUSTAINABILITY OF BUSINESS. *Human Relations*, *3*(1).
40. Rejter, A. (2024). STUDIES IN ONOMASTICS AND THE NEW HUMANITIES. *Onomastica*, *68*. https://doi.org/10.17651/ONOMAST.68.2
41. Shakil, M., Ali, M., Illahi, T., & Ahmed, F. (2025). FinTech 5.0 and the Future of Global

Finance: Harnessing Artificial Intelligence, Blockchain, and Big Data to Build Inclusive, Sustainable, and Resilient Financial Ecosystems. *Inverge Journal of Social Sciences*, *4*(4). https://doi.org/10.63544/ijss.v4i4.179

42. Sipa, M. (2023). Trends and challenges in talent management – mind map. *Scientific Papers of Silesian University of Technology Organization and Management Series*, *2023*(164). https://doi.org/10.29119/1641-3466.2022.164.34
43. Skoczkowski, T., Bielecki, S., Wołowicz, M., & Węglarz, A. (2025). Redefining Energy Management for Carbon-Neutral Supply Chains in Energy-Intensive Industries: An EU Perspective. *Energies*, *18*(15). https://doi.org/10.3390/en18153932
44. Skrypnyk, S. v. (2025). Key Concepts and Models of Strategic Innovation Management in Ukraine: Modern Approaches. *The Problems of Economy*, *2*(64). https://doi.org/10.32983/2222-0712-2025-2-77-85
45. Soni, A., Jain, A., & Kaur, B. (2022). A bibliometric analysis of effect of artificial intelligence on customer satisfaction in banking sector. *International Research Journal of Modernization in Engineering*, *4*(3).
46. Sydorovych, O., Bukovskyi, M., & Yaroshovych, A. (2024). DIGITAL TOOLS OF TAX ADMINISTRATION: CURRENT TRENDS AND PERSPECTIVES. *INNOVATIVE ECONOMY*. https://doi.org/10.37332/2309-1533.2024.4.26
47. Torres-Sainz, R., de-Zayas-Pérez, M. R., Trinchet-Varela, C. A., Pérez-Vallejo, L. M., & Rodríguez, R. P. (2025). INTELLIGENT PREDICTIVE MAINTENANCE: A BIBLIOMETRIC APPROACH THROUGH THE SCOPUS DATABASE. *Bibliotecas, Anales de Investigacion*, *21*(1).
48. Vollmar, H. C., Kramer, U., Müller, H., Griemmert, M., Noelle, G., & Schrappe, M. (2017). Digitale Gesundheitsanwendungen - Rahmenbedingungen zur Nutzung in Versorgung, Strukturentwicklung und Wissenschaft - Positionspapier der AG Digital Health des DNVF. *Gesundheitswesen*, *79*(12). https://doi.org/10.1055/s-0043-122233
49. VOLOSHYN, R. (2024). COMMODITY EXPERTISE IN E-COMMERCE. *"ECONOMY. FINANCES. MANAGEMENT: Topical Issues of Science and Practical*

Activity," 4 (70). https://doi.org/10.37128/2411-4413-2024-4-4

50. Wang, Y., Kang, M., Liu, Y., Li, J., Xue, K., Wang, X., Du, J., Tian, Y., Ni, Q., & Wang, F. Y. (2023). Can Digital Intelligence and Cyber-Physical-Social Systems Achieve Global Food Security and Sustainability? *IEEE/CAA Journal of Automatica Sinica*, *10*(11). https://doi.org/10.1109/JAS.2023.123951
51. Wijanarti, H. L., & Anisyah, T. D. A. (2022). Hubungan Antara Kualitas Tidur, Beban Kerja Fisik Terhadap Perasaan Kelelahan Kerja Pada Perawat Rawat Inap Kelas 3 Di Rs Pku Muhammadiyah Gamping. *Jurnal Lentera Kesehatan Masyarakat*, *1*(1).
52. Yubo, S., Ramayah, T., & Hongmei, L. (2022). Research on the Impact of Intelligent Manufacturing Technology on Business Model Based on CNKI's Evidence. *Global Business and Management Research, Suppl. Special Issue*, *14*(3s).
53. Zavalii, T., & Lehenchuk, S. (2025). DEVELOPMENT OF RESEARCH METHODS IN MARKETING IN THE CONDITIONS OF FORMING THE DIGITAL MARKETING PARADIGM. *Economic Scope*, *204*. https://doi.org/10.30838/ep.204.101-109
54. Zhao, R., Yang, X., Zhang, D., Li, J., Huang, Y., Xian, G., Kou, Y., & Sun, T. (2025). Artificial Intelligence for Agricultural Science (AI4AS): Key Elements, Challenges and Pathways. *Smart Agriculture*, *7*(3). https://doi.org/10.12133/j.smartag.SA202502019
55. Zhao, X., Wang, X., Research, R. C.-A. in R., & 2025, undefined. (2025). Research on optimizing water resource management and improving crop yield using big data and artificial intelligence. *Jstage.Jst.Go.Jp*, *5*(2).
56. Zharov, A. M. (2024). THE DYNAMICS OF KNOWLEDGE:AN EVOLUTIONARY ANALYSIS OF SCIENTIFIC DEVELOPMENT THEORIES. *Vestnik Tomskogo Gosudarstvennogo Universiteta. Filosofiya, Sotsiologiya, Politologiya, 80*. https://doi.org/10.17223/1998863x/80/3
57. Zheldibayeva, R. (2025). THE IMPACT OF AI AND PEER FEEDBACK ON RESEARCH WRITING SKILLS: A STUDY USING THE CGSCHOLAR PLATFORM AMONG KAZAKHSTANI SCHOLARS. *Scientific Journal of Astana IT University*. https://doi.org/10.37943/21tbax7694

58. Zozulov, O., & Boboshko, Y. (2025). Classification of product development and market entry models. *Marketing and Digital Technologies*, *9*(2). https://doi.org/10.15276/mdt.9.2.2025.4

59. Zuo, H. (2025). Human-Environment Relationship Research from a Multidisciplinary Integration Perspective in Geography. *Frontiers in Sustainable Development*, *5*(8). https://doi.org/10.54691/64t83m16

60. Игнатьев, М. Е., & Степанов, В. В. (2025). Методология диссертационных исследований в юридических науках: отдельные аспекты и проблемы. *СОВРЕМЕННОЕ ПРАВО*, *7*. https://doi.org/10.25799/ni.2025.32.84.001

61. Ксения Юрьевна, Б., & Елена Иосифовна, Л. (2024). Identification and regulation of the macro-regional ecosystem' human resources potential development trends: global and national experience. *STATE AND MUNICIPAL MANAGEMENT SCHOLAR NOTES*, *1*(4). https://doi.org/10.22394/2079-1690-2024-1-4-28-33

62. Muhtar, E. (2022). PENGARUH PENDEKATAN KETRAMPILAN PROSES DALAM MENINGKATKAN HASIL BELAJAR FISIKA PADA KONSEP GETARAN DAN GELOMBANG PADA PESERTA DIDIK KELAS VIII SMP NEGERI 17 HALMAHERA SELATAN. *Jurnal Pembelajaran Dan Sina Fisika*, *7*(1).

63. Permenkes. (2022). Peraturan Menteri Kesehatan Nomor 24 Tahun 2022 Tentang Rekam Medis. Jakarta: Menteri Kesehatan Republik Indonesia. *Science*, *7*(1).

64. Putri, S. H. (2022). Pengaruh Brand Image terhadap keputusan pembelian produk skincare. *Science*, *7*(1).

65. Candaika, P. (2022). Gambaran Kesiapan Rumah Sakit Umum Daerah (RSUD) Menghadapi Kebijakan Kelas Rawat Inap Standar (KRIS) Tahun 2022. *Science*, *7*(1).

Chapter 2.

The Epistemology of Algorithms: A Comprehensive Framework for Scientific Methodology in Artificial Intelligence and Data Science

1. The Fourth Paradigm and the Crisis of Reproducibility

The emergence of Artificial Intelligence (AI) and Data Science (DS) has catalyzed a fundamental transformation in the structure of scientific inquiry, shifting the locus of discovery from the traditional hypothetico-deductive model toward a data-intensive, computational paradigm often referred to as the "Fourth Paradigm" of science.[1] This shift is not merely instrumental—replacing analog tools with digital ones—but epistemological, altering the very nature of how knowledge is generated, validated, and interpreted. Where the industrial revolution mechanized physical labor through heat engines, the information revolution is mechanizing cognitive labor through "data engines" capable of generating actionable knowledge from vast, unstructured datasets.[1] However, this rapid mechanization has outpaced the development of rigorous methodological standards, leading to a landscape where innovation frequently precedes validation.

The pathway from algorithmic innovation to reliable scientific deployment is non-linear and fraught with complexity. Unlike classical software engineering, where logic is deterministic and explicitly coded, modern AI systems—particularly those based on deep learning—operate as probabilistic "black boxes" where internal decision-making processes are opaque and emergent.[2] This opacity presents a profound challenge to the scientific method, which relies on transparency, reproducibility, and the falsifiability of hypotheses.

The integration of AI into high-stakes domains such as healthcare, climate science, and criminal justice demands a transition from ad-hoc experimentation to a rigorous "Diffusion Engine" of methodologies.[1] This engine must bridge the gap between abstract computer science and domain-specific application, ensuring that AI models are not just predictive, but robust, fair, and scientifically valid.

Current literature suggests that the "reproducibility crisis" in AI is largely a symptom of methodological immaturity. Issues such as "p-hacking," weak baselines, and data leakage—long recognized in statistics—have resurfaced in machine learning research, exacerbated by the stochastic nature of training algorithms and the proprietary nature of large datasets.[4] Furthermore, the lack of standardized reporting protocols has made it difficult to distinguish between genuine algorithmic advances and performance gains achieved through hyperparameter overfitting or random chance.

To address these challenges, this report synthesizes a unified methodological framework for AI and Data Science. It draws upon diverse streams of research—from software engineering lifecycles like CRISP-DM and TDSP to statistical rigor in hypothesis testing and the emerging discipline of Scientific Machine Learning (SciML). By integrating principles of data provenance, rigorous experimental design (including ablation and sensitivity analysis), and ethical stewardship, this framework aims to establish a standard for expert-level research that is exhaustive, reproducible, and deeply integrated with the scientific method.

2. Structural Lifecycles: From Linear Process to Circular Inquiry

The management of scientific research in AI requires a structured lifecycle that accommodates the unique characteristics of data-driven projects: high uncertainty, iterative experimentation, and the need for continuous validation. While software development has settled on Agile methodologies, data science requires a hybrid approach that fuses engineering discipline with scientific rigor.

2.1 The Evolution of Process Models: CRISP-DM and Beyond

The Cross-Industry Standard Process for Data Mining (CRISP-DM) remains the foundational reference model for the field, widely adopted for its clarity and industry-agnostic applicability.[6] It delineates the research process into six phases: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment.

However, the modern application of CRISP-DM acknowledges its limitations. The original model's implicit linearity is often insufficient for the complexities of modern AI, where "Data Preparation" and "Modeling" are deeply intertwined through feature engineering and representation learning. Consequently, the methodology has evolved into a circular approach.[7] For instance, insights gained during the "Model Evaluation" phase often force a return to "Problem Definition," creating a feedback loop that refines the scientific question itself.[8] This circularity is essential; a linear progression assumes that the initial hypothesis and data quality are sufficient, which is rarely the case in high-dimensional scientific problems.

2.2 The Team Data Science Process (TDSP) and Agile Integration

To address the collaborative and operational needs of modern research teams, Microsoft introduced the Team Data Science Process (TDSP). TDSP modernizes CRISP-DM by integrating it with Agile practices (Scrum, Kanban) and emphasizing the lifecycle of the *team* rather than just the *project*.[9]

The TDSP framework is distinct in its focus on:

1. **Standardized Project Structure:** Enforcing a consistent directory structure and document template system to reduce "technical debt" and facilitate knowledge transfer between researchers.[10]

2. **Version Control and Collaboration:** Unlike solitary academic research of the past, TDSP mandates the use of Git-based version control for both code and documentation, treating the experimental process as a collaborative software engineering effort.[9]
3. **Iterative Sprints:** Research is broken down into fixed-length iterations (Sprints), allowing for rapid feedback and the flexibility to pivot based on early experimental results. This "fail fast" mentality is crucial in AI, where training large models can be resource-intensive (see Table 3).[11]

Table 3: Comparative Analysis of Research Lifecycles

Feature	CRISP-DM	TDSP (Team Data Science Process)	Agile/Kanban for Data Science
Philosophical Basis	Industrial Data Mining	Collaborative Software Engineering	Lean Manufacturing / Flow
Structure	6 Phases (Cyclical)	Role-Based, Iterative Lifecycle	Continuous flow, Minimize WIP
Key Artifacts	Phase Reports, Model	Git Repos, Charters, Exit Reports	Backlogs, Kanban Boards
Strengths	Comprehensive, standard terminology, focus on business goals [6]	Integration with MLOps, strong team focus, reproducibility support [9]	High flexibility, maximizes throughput, adapts to uncertainty [11]
Limitations	Can become a "waterfall" trap; lacks specific team roles [11]	Higher overhead for small, academic teams; documentation heavy	Can lack strategic coherence without a broader framework

2.3 The Scientific AI Lifecycle

In a purely research context—distinct from industrial deployment—the lifecycle focuses heavily on *validity* and *discovery*. The "Unified View" of AI research methodology incorporates a long-term dimension based on the scientific method, identifying strategies such as hypothetical/deductive and hermeneutical/inductive analysis.[12]

The critical distinction in the scientific lifecycle is the Problem Formulation phase. Here, abstract scientific inquiries must be translated into concrete data science problems.[7] This involves defining the problem in terms of symmetry groups (e.g., supervised vs. unsupervised learning) and identifying the mathematical metrics that will serve as proxies for scientific truth.[13] For example, in clinical AI, the problem formulation must explicitly define the "beneficiary" and the "end-user" to preemptively address ethical downstream effects.[14] This phase sets the trajectory for the entire project; a poorly framed problem will lead to technically accurate but scientifically irrelevant models.

3. The Science of Data: Curation, Provenance, and Ethics

In the classical scientific method, "data" was often the result of controlled experiments designed to minimize noise. In the AI paradigm, data is often "found" (observational), unstructured, and noisy. Therefore, the methodology of *data curation* has become a scientific discipline in its own right, critical for ensuring the validity of downstream inference.

3.1 Data Provenance and Lineage

Data provenance—the documentation of the origin, history, and transformations of data—is the bedrock of reproducibility in AI.[15] Without a verifiable lineage, an AI model is scientifically essentially worthless, as its predictions cannot be audited for artifacts or errors introduced during preprocessing.

Best practices for provenance in research include:

- **Metadata Schemas:** Utilizing standardized schemas to record provenance details (creator, date, source, license) consistently across all datasets.[16]
- **Immutable Logging:** Implementing systems that create a permanent, unalterable record of every transformation applied to the raw data. This allows researchers to "replay" the data processing pipeline to verify results.[16]
- **Hash Verification:** Using cryptographic hashes to verify the integrity of data files, ensuring that the dataset analyzed today is bit-for-bit identical to the one analyzed yesterday.[16]

The necessity of this rigor was highlighted by a recent audit of AI training datasets, which found a systemic failure in provenance tracking: over 70% of datasets examined lacked license information, and 50% had miscategorized licenses.[17] Such negligence not only creates legal liability but undermines scientific trust. Tools like the Data Provenance Explorer have emerged to help researchers trace the lineage of fine-tuning datasets, ensuring that the "supply chain" of ideas remains transparent and robust.[17]

3.2 Datasheets for Datasets

To standardize the communication of data characteristics, the AI community is increasingly adopting the "Datasheets for Datasets" framework.[18] Modeled after the datasheets used in the electronics industry (which describe the operating limits of a resistor or capacitor), this document provides a structured summary of a dataset's motivation, composition, collection process, and recommended uses.

The scientific value of a Datasheet lies in its ability to force the researcher to articulate *context*.

- **Motivation:** Why was this data collected? (e.g., specific gap in literature vs. general purpose).
- **Composition:** Does the dataset contain subpopulations? Are there missing data patterns?

- **Collection:** How was the data acquired? (e.g., API, scraping, sensors).
- **Uses:** What are the tasks for which the dataset is valid? Crucially, what are the tasks for which it should *not* be used?.[18]

In the Earth Sciences, this methodology has been adapted to document geospatial biases and technical limitations, proving that the framework is transferable across scientific domains.[20] Empirical studies show that the act of completing a Datasheet increases the "ethical sensitivity" of machine learning engineers, helping them recognize potential biases that might otherwise go unnoticed.[21]

3.3 Algorithmic Fairness and Bias Metrics

Rigorous methodology demands that "performance" be defined beyond simple accuracy. In many scientific and social applications, a model that achieves high accuracy by exploiting biases in the training data is considered a failure. Therefore, bias assessment is a mandatory component of the evaluation phase.

Research identifies three primary metrics for assessing fairness, each representing a different philosophical definition of equity:

1. **Demographic Parity (Statistical Parity):** This metric requires that the probability of a positive outcome be independent of the protected attribute (e.g., gender or race).[22] $P(\hat{Y}=1 \mid A=0) = P(\hat{Y}=1 \mid A=1)$. While intuitively appealing, this metric can be problematic if the base rates of the target variable differ genuinely between groups.
2. **Equality of Opportunity:** This focuses on the True Positive Rate (TPR). It requires that qualified individuals in both groups have an equal chance of being selected.[22] $P(\hat{Y}=1 \mid Y=1, A=0) = P(\hat{Y}=1 \mid Y=1, A=1)$. This is often the preferred metric in medical diagnostics (e.g., ensuring a cancer detection model works equally well for all races).
3. **Disparate Impact:** A ratio-based metric mandated by US employment law. If the ratio of the selection rate of the protected group to the unprotected group is less than 0.8 (the

"four-fifths rule"), the model is considered biased.[22]

The "Objective Fairness Index" (OFI) has recently been proposed to provide a legally grounded and context-aware perspective on these metrics, addressing gaps where traditional disparate impact calculations fail.[24] Integrating these metrics into the standard evaluation loop ensures that the "scientific discovery" is not merely an artifact of systemic inequality.

4. Rigorous Experimental Design in Machine Learning

An ML experiment is a controlled procedure designed to falsify a hypothesis regarding the relationship between data features and target variables.[25] However, the stochastic nature of training (random initialization, data shuffling) and the complexity of hyperparameters make isolation of variables difficult. A robust experimental design is the only defense against "alchemy"—the trial-and-error tweaking of models until they appear to work.

4.1 The Importance of Strong Baselines

A pervasive issue in AI literature is the use of "weak baselines" to inflate the perceived novelty of a proposed method. A rigorous study must compare the new model not just against other state-of-the-art complex models, but against well-tuned simple models.[26]

Guidelines for Baseline Construction:

- **Classical Baselines:** Always test simple, interpretable models (Linear/Logistic Regression, Random Forests) first. These establish the "floor" of performance. If a deep neural network barely outperforms a logistic regression, the complexity is likely unjustified.[26]
- **Hyperparameter Tuning:** Baselines must be tuned with the same rigor as the experimental model. A comparison between a hyper-tuned neural net and a default-

parameter Random Forest is scientifically invalid.

- **Preprocessing consistency:** Both the baseline and the new model must consume the exact same data splits and preprocessing steps to prevent data leakage.[26]

4.2 Cross-Validation Methodologies

Estimating the generalization error—how the model performs on unseen data—is the central task of evaluation. The standard "train/test split" is often insufficient due to the high variance in result estimates.

- **k-Fold Cross-Validation:** The dataset is partitioned into k disjoint subsets. The process involves k iterations; in each, a different subset is held out for testing while the remaining k-1 are used for training. The final performance metric is the average of the k scores.[27] This reduces variance and ensures every data point is used for testing exactly once.
- **Stratified k-Fold:** For classification problems with imbalanced classes, standard k-fold can result in folds with no positive examples. Stratification ensures that the class distribution in every fold preserves the distribution of the whole dataset.[28]
- **Time Series Split (Rolling Window):** In temporal data (e.g., stock prices, climate data), standard cross-validation introduces "look-ahead bias" (training on future data to predict the past). The correct methodology is a rolling window where the training set consists of indices [0, t] and the test set consists of [t+1, t+k].[27]

4.3 Hyperparameter Optimization and Ablation

The search for optimal hyperparameters (learning rate, layer depth) is part of the experimental design. However, distinguishing between performance gains from architecture vs. hyperparameters is critical.

Ablation Studies:

An ablation study is the systematic removal of components of a machine learning system to measure their marginal contribution to performance.29 It is the AI equivalent of a "gene knockout" experiment.

- **Methodology:** Deconstruct the model into its additive components (e.g., "Architecture A + Feature B + Regularization C"). Train and evaluate variants where one component is removed or replaced with a counterfactual (zeroed out or randomized).[30]
- **Interpretation:** If removing a complex attention mechanism result in a <1% drop in accuracy, the mechanism is likely redundant or the model is learning via a "Compensatory Masquerade" (where other parts of the network compensate for the missing signal).[31]
- **Efficiency:** While full grid search for ablations is ideal, it is computationally expensive. Researchers effectively use "Leave-One-Component-Out" strategies or automated tools like MAGGY to parallelize ablation trials.[29]

5. Statistical Significance and Model Comparison

In many published papers, a model is declared "superior" if its accuracy is 0.1% higher than the baseline. Without statistical testing, such claims are scientifically vacuous. The field has moved toward specific non-parametric tests to validate these comparisons.

5.1 The Failure of the Paired t-test

Traditionally, the paired Student’s t-test was used to compare the means of two models' performance. However, this test assumes that the differences in performance are normally distributed and, crucially, that the samples are independent. In k-fold cross-validation, the training sets overlap significantly (sharing k-2 folds worth of data), violating the independence assumption. This leads to an elevated Type I error rate (false positives).[5]

5.2 Recommended Statistical Tests

To address the shortcomings of the t-test, the AI community has converged on more robust alternatives:

1. McNemar’s Test:

This is the standard for comparing two classifiers on a single dataset. It operates on the contingency table of predictions:

- N_{01}: Number of examples misclassified by Model A but correctly classified by Model B.
- N_{10}: Number of examples correctly classified by Model A but misclassified by Model B.

The test statistic is approximated by $\chi^2 = \frac{(|N_{01} - N_{10}| - 1)^2}{N_{01} + N_{10}}$. It tests if the models make errors on the same examples. It is computationally efficient as it requires running the models only once.5

2. 5x2 Cross-Validation Paired t-test:
Proposed by Dietterich, this method performs 5 iterations of 2-fold cross-validation. It is designed to balance the trade-off between power and Type I error, specifically accounting for the variation arising from the choice of training sets. It is considered slightly more powerful than McNemar's test but requires 10 training runs.33
3. Wilcoxon Signed-Rank Test:
This is the recommended non-parametric test for comparing two classifiers across multiple datasets (e.g., a benchmark study on 20 different domains). It ranks the differences in performance and checks if the distribution of differences is symmetric around zero. It is robust to outliers and does not assume normality (see Table 4).[27]

Table 4: Selection Matrix for Statistical Tests in AI

Scenario	Recommended Test	Rationale
Two models, Single Dataset	McNemar's Test	Low computational cost, acceptable Type I error.[33]
Two models, Single Dataset (High Rigor)	5x2cv Paired t-test	Better estimation of training set variance.[33]

Two models, Multiple Datasets	Wilcoxon Signed-Rank	Robust to non-normal distributions across domains.[34]
Multiple models, Multiple Datasets	Friedman Test + Nemenyi Post-hoc	Corrects for multiple comparisons problem.[34]

6. Sensitivity Analysis and Interpretability

As models grow in complexity, "interpretability" becomes a requirement for scientific validity. Sensitivity Analysis (SA) provides the methodological toolkit to peer inside the black box by quantifying how changes in inputs affect outputs.

6.1 Input Perturbation and Feature Importance

The simplest form of SA involves perturbing input variables (e.g., adding Gaussian noise, masking pixels, or shifting values) and observing the degradation in model output.

- **Local Analysis:** Explains a specific prediction. For example, in an image classification model, masking a specific region of the image to see if the classification changes identifies that region as the "cause" of the prediction.[35]
- **Global Analysis:** Averages the sensitivity across the entire dataset to rank features by overall importance. This helps in feature selection and model simplification.[36]

6.2 Advanced Jacobian-Based Analysis

For deep neural networks, sensitivity is mathematically formalized using the Jacobian matrix of the function. The norm of the input-output Jacobian ||\nabla_x f(x)||_F measures the local sensitivity of the network.

- **Generalization Insight:** Research has shown a strong correlation between the Jacobian norm and generalization error. Models that are robust to small input perturbations (low sensitivity) tend to generalize better to unseen data.[2]
- **Robustness Training:** This insight has led to training methodologies where the

Jacobian norm is penalized (regularized) during training, forcing the model to learn smoother decision boundaries that are less brittle.[2]

7. Scientific Machine Learning (SciML): Bridging Data and Physics

Scientific Machine Learning (SciML) represents the frontier where data-driven methodology meets mechanistic modeling. Unlike pure AI, which learns patterns solely from data, SciML incorporates domain knowledge (physics, biology, chemistry) directly into the learning process. This creates a "gray box" model that combines the flexibility of neural networks with the interpretability of differential equations.[37]

7.1 Physics-Informed Neural Networks (PINNs)

The defining methodology of SciML is the Physics-Informed Neural Network (PINN). In a standard neural network, the optimization objective is to minimize the difference between predictions and data labels (Loss_{data}). In a PINN, the loss function is augmented with a "residual" term derived from the governing physical laws (e.g., Navier-Stokes equations for fluid dynamics).

Loss_{Total} = Loss_{data} + \lambda \cdot Loss_{Physics}

By minimizing this composite loss, the network is constrained to find solutions that not only fit the observed data points but also satisfy the underlying differential equations in the spaces *between* the data points.[37] This regularization allows PINNs to generalize well even in "small data" regimes where pure deep learning would fail due to overfitting.[38]

7.2 Operator Learning and Model Discovery

SciML also expands the scope of learning from functions to *operators*.

- **Neural Operators (e.g., DeepONet, FNO):** These architectures learn the mapping between infinite-dimensional function spaces. For example, instead of solving a heat equation for one specific initial condition, a Neural Operator learns the *solution operator* that maps *any* initial temperature profile to its future state. This allows for real-time simulation of complex physical systems at a fraction of the cost of traditional numerical solvers.[37]

- **Automated Model Discovery (SINDy):** This methodology uses sparse regression to discover the governing equations from data. It takes time-series data and finds the sparsest combination of mathematical terms (e.g., x, x^2, $\sin(x)$) that describe the dynamics. The output is not a black-box model, but a symbolic equation (e.g., $\frac{dx}{dt} = \sigma(y-x)$) that scientists can interpret and verify.[39]

8. Reproducibility, MLOps, and the Crisis of Trust

The ultimate test of a scientific methodology is reproducibility. In AI, this is challenging due to the complex stack of software and hardware dependencies. The field has moved toward "MLOps" (Machine Learning Operations) to provide the infrastructure for reproducible science.

8.1 The Hierarchy of Reproducibility

Methodologists distinguish between three levels of validity [40]:

1. **Repeatability:** The same team, using the same code and hardware, obtains the same result. This is a check of the *stability* of the code.
2. **Reproducibility (Dependent):** A different team, using the *original* artifacts (code, data), obtains the same result. This verifies that the result is not an artifact of a specific local environment.
3. **Replicability (Independent):** A different team reimplements the algorithm from the paper's description (without seeing the original code) and obtains a consistent result. This is the highest standard, verifying the *scientific truth* of the method rather than just the code correctness.

8.2 The Technological Stack for Reproducibility

Achieving reproducibility requires specific tools to control entropy in the computational environment:

- **Containerization (Docker):** Docker packages the operating system, libraries, and dependencies into an immutable image. This ensures that a model trained on a Linux server in 2024 can be re-run on a Windows laptop in 2026 with the exact same environment.[42]
- **Data Versioning (DVC):** Standard Git is poor at handling large binary files. Tools like DVC (Data Version Control) allow researchers to version control datasets alongside code. A specific Git commit can be linked to a specific "snapshot" of the data, ensuring that if the data changes, the provenance of the model is not lost.[44]
- **Determinism and Seeds:** Deep learning libraries (PyTorch, TensorFlow) often default to non-deterministic algorithms for speed (especially on GPUs). Rigorous methodology requires setting global random seeds and forcing deterministic algorithms, even at the cost of performance, to ensure that results are bit-wise identical across runs.[42]

8.3 The NeurIPS Checklist as a Standard

To enforce these standards, the NeurIPS conference (a premier venue for AI research) has introduced a mandatory reproducibility checklist. Authors must explicitly declare whether they have provided code, data, error bars, and details on computing infrastructure.[47]

Recent experiments using Large Language Models (LLMs) to audit these checklists have shown that AI tools themselves can help enforce methodological rigor. An "Author Checklist Assistant" powered by GPT-4 was used to vet papers, providing feedback to authors on whether their claims of reproducibility were substantiated by the provided artifacts.[49]

9. Future Horizons: AI-Driven Hypothesis Generation

As the methodology matures, AI is transitioning from a tool for *testing* hypotheses to a system for *generating* them. This closes the loop of the scientific method, automating the discovery process itself.

9.1 The HypoGeniC Framework

The "HypoGeniC" framework demonstrates this new capability. It uses an iterative "agentic" workflow:

1. **Initialization:** An LLM reviews vast literature and initial data to propose a set of scientific hypotheses.
2. **Refinement:** A separate AI agent tests these hypotheses against challenging data examples ("counter-examples").
3. Update: Hypotheses that fail are discarded; those that survive are refined and made more specific. This process mimics the peer-review and revision cycle of human science but operates at a speed and scale impossible for human researchers.[51]

9.2 Case Study: AlphaFold and the Protein Universe

DeepMind's AlphaFold serves as the exemplar of this new scientific methodology. Its architecture is built on an iterative "recycling" mechanism where the model generates a structural hypothesis, assesses it, and feeds it back as input for refinement.[52] This loop effectively performs in silico experiments.

The impact is a paradigm shift in biology: researchers can now bypass the years-long process of X-ray crystallography for determining protein structures. Instead, they use the AI prediction as a high-confidence hypothesis to guide downstream experiments in drug discovery and molecular dynamics.[53] In this regime, the AI methodology becomes the scientific method.

The application of scientific methodology to AI and data science is no longer optional; it is the prerequisite for progress. The field has graduated from the "wild west" of ad-hoc scripts and leaderboard-chasing to a mature discipline governed by rigorous frameworks.

The Unified Framework proposed in this report relies on three pillars:

1. **Structural Rigor:** Adopting circular lifecycles (TDSP, Scientific Lifecycle) that emphasize feedback loops and team collaboration.
2. **Experimental Rigor:** Utilizing solid baselines, ablation studies, and appropriate statistical tests (McNemar/Wilcoxon) to validate claims.
3. **Ethical and Data Rigor:** Treating data curation as a first-class scientific activity, documenting provenance, and actively measuring bias.

For the modern researcher, adhering to these methodologies is not merely about getting a paper accepted; it is about ensuring that the "Intelligence" we are building is robust, transparent, and ultimately beneficial to the human enterprise. As we stand on the brink of automated discovery, the rigor of our methods will determine the reliability of our future knowledge.

Chapter bibliography

1. T. P., & . V. G. (2025). INVESTIGATION OF DEEP LEARNING FRAMEWORK TO ACCELERATE AI FOR IMAGE PROCESSING APPLICATIONS. *International Journal For Multidisciplinary Research*, *7*(3). https://doi.org/10.36948/ijfmr.2025.v07i03.45548
2. Abdul-Rahman, G., Zwitter, A., & Haleem, N. (2025). A systematic literature review on the role of artificial intelligence in citizen science. *Discover Artificial Intelligence*, *5*(1). https://doi.org/10.1007/s44163-025-00437-z
3. Acomi, N., Surugiu, G., Raicu, G., & Stanca, C. (2025). A Thematic Review of Port Services and Emission Reduction Strategies Using ATLAS.ti. *TransNav*, *19*(2). https://doi.org/10.12716/1001.19.02.14
4. Aggarwal, K., Mijwil, M. M., Sonia, Al-Mistarehi, A. H., Alomari, S., Gök, M., Zein Alaabdin, A. M., & Abdulrhman, S. H. (2022). Has the Future Started? The Current Growth of Artificial Intelligence, Machine Learning, and Deep Learning. *Iraqi Journal for Computer Science and Mathematics*, *3*(1). https://doi.org/10.52866/ijcsm.2022.01.01.013
5. Ahmadpour, M. (2024). Emerging Approaches in Enhancing Mental and Physical Health of the Elderly: Integration of Strength Training, Mindfulness, and Artificial Intelligence. *Journal of Psychological Dynamics in Mood Disorders*, *3*(5). https://doi.org/10.61838/kman.pdmd.3.5.19
6. Alshadidi, A. A. F., Alshahrani, A. A., Aldosari, L. I. N., Chaturvedi, S., Saini, R. S., Hassan, S. A. bin, Cicciù, M., & Minervini, G. (2023). Investigation on the Application of Artificial Intelligence in Prosthodontics. In *Applied Sciences (Switzerland)* (Vol. 13, Issue 8). https://doi.org/10.3390/app13085004
7. Asaduzzaman, M., & Flowra, M. T. (2018). Resurgence of infectious diseases due to forced migration: is planetary health and One Health action synergistic? *The Lancet Planetary Health*, *2*(10).

8. Bock, T., & von der Oelsnitz, D. (2025). Leadership-competences in the era of artificial intelligence – a structured review. In *Strategy and Leadership* (Vol. 53, Issue 3). https://doi.org/10.1108/SL-09-2024-0100
9. Boukrouh, I., & Azmani, A. (2024). ARTIFICIAL INTELLIGENCE APPLICATIONS IN E-COMMERCE: A BIBLIOMETRIC STUDY FROM 1995 TO 2023 USING MERGED DATA SOURCES. *International Journal of Professional Business Review*, *9*(4). https://doi.org/10.26668/businessreview/2024.v9i4.4537
10. Buchkremer, R., Demund, A., Ebener, S., Gampfer, F., Jagering, D., Jurgens, A., Klenke, S., Krimpmann, D., Schmank, J., Spiekermann, M., Wahlers, M., & Wiepke, M. (2019). The Application of Artificial Intelligence Technologies as a Substitute for Reading and to Support and Enhance the Authoring of Scientific Review Articles. In *IEEE Access* (Vol. 7). https://doi.org/10.1109/ACCESS.2019.2917719
11. Cafuta, D., Dodig, I., Cesar, I., & Kramberger, T. (2021). Developing a modern greenhouse scientific research facility — A case study. *Sensors*, *21*(8). https://doi.org/10.3390/s21082575
12. Chango, X., Flor-Unda, O., Gil-Jiménez, P., & Gómez-Moreno, H. (2024). Technology in Forensic Sciences: Innovation and Precision. In *Technologies* (Vol. 12, Issue 8). https://doi.org/10.3390/technologies12080120
13. Chen, S., & Garcia, A. (2017). Embodiment of Cognition: Exploring the Interplay between Body and Mind. *Studies in Science of Science | ISSN: 1003-2053*, *35*(9).
14. Espinel, R., Herrera-Franco, G., Rivadeneira García, J. L., & Escandón-Panchana, P. (2024). Artificial Intelligence in Agricultural Mapping: A Review. In *Agriculture (Switzerland)* (Vol. 14, Issue 7). https://doi.org/10.3390/agriculture14071071
15. Fil, N., Slisarenko, R., Deineko, Z., & Morozova, L. (2025). Trends in Artificial Intelligence Research on Education: Topic Modeling Using Latent Dirichlet Allocation. *Bulletin of Kharkov National Automobile and Highway University*, *108*. https://doi.org/10.30977/bul.2219-5548.2025.108.0.17
16. Georgiev, G. (2021). THE BREAKDOWN OF THE PUBLIC-PRIVATE DIVIDE IN SECURITIES LAW: CAUSES, CONSEQUENCES, AND REFORMS GEORGE.

Frontiers in Neuroscience, *14*(1).

17. Guizzardi, G., & Guarino, N. (2024). Explanation, semantics, and ontology. *Data and Knowledge Engineering*, *153*. https://doi.org/10.1016/j.datak.2024.102325
18. Gupta, S., Mahmood, A., Shetty, P., Adeboye, A., & Ramprasad, R. (2024). Data extraction from polymer literature using large language models. *Communications Materials*, *5*(1). https://doi.org/10.1038/s43246-024-00708-9
19. Hou, J., Zheng, B., Li, H., & Li, W. (2025). Evolution and impact of the science of science: from theoretical analysis to digital-AI driven research. In *Humanities and Social Sciences Communications* (Vol. 12, Issue 1). https://doi.org/10.1057/s41599-025-04617-1
20. Jungmeier, G. (2017). The Biorefinery Fact Sheet. *The International Journal of Life Cycle Assessment*, *23*(1).
21. Júnior, G. H. K., & Gottschalg-Duque, C. (2023). Information architecture applied on natural language processing: a proposal Information Science contributions on data preprocessing for training and learning of artificial neural networks. *Revista Digital de Biblioteconomia e Ciencia Da Informacao*, *21*. https://doi.org/10.20396/rdbci.v21i00.8671396/30919
22. Kharimah, W. (2021). DAMPAK PENGUNDURAN DIRI AMERIKA SERIKAT DARI PERJANJIAN JCPOA (JOINT COMPREHENSIVE PLAN OF ACTION). *Frontiers in Neuroscience*, *14*(1).
23. Kirkpatrick, A. W., McKee, J. L., McBeth, P. B., Ball, C. G., Laporta, A. T., Broderick, T., Leslie, T., King, D., Wright Beatty, H. E., Keillor, J., Tien, H., de'Angelis, N., Khan, J., Marchegiani, F., Bianchi, G., Aisoni, F., Alberti, D., Ansaloni, L., Biffl, W. L., … MOORI, S. H. (2020). Translating Scientific Evidence into Practice using Digital Medicine and Electronic Patient Reported Outcomes. *Telemedicine and E-Health*, *17*(1).
24. Koçak, M., & Akçalı, Z. (2024). Development trends and knowledge framework of artificial intelligence (AI) applications in oncology by years: a bibliometric analysis from 1992 to 2022. *Discover Oncology*, *15*(1). https://doi.org/10.1007/s12672-024-01415-

0

25. Leong, W. D., Teng, S. Y., How, B. S., Ngan, S. L., Rahman, A. A., Tan, C. P., Ponnambalam, S. G., Lam, H. L., Yadav, G., Luthra, S., Jakhar, S. K., Mangla, S. K., Rai, D. P., Tortorella, G. L., Narayanamurthy, G., Thurer, M., Dombrowski, U., Wullbrandt, J., Fochler, S., … Gierej, S. (2019). Elimination of waste: creation of a successful Lean colonoscopy program at an academic medical center. *Surgical Endoscopy*.
26. Luo, X., Tham, Y. C., Daher, M., Bian, Z., Chen, Y., Estill, J., & group, the G. working. (2025). Protocol for developing the reporting guideline for the use of chatbots and other Generative Artificial intelligence tools in MEdical Research (GAMER). *medRxiv*.
27. Masso, A., Gerassimenko, J., Kasapoglu, T., & Beilmann, M. (2025). Research ethics committees as knowledge gatekeepers: The impact of emerging technologies on social science research. *Journal of Responsible Technology*, *21*. https://doi.org/10.1016/j.jrt.2025.100112
28. Morante-Carballo, F., Arcentales-Rosado, M., Caicedo-Potosí, J., & Carrión-Mero, P. (2025). Artificial intelligence applications in hydrological studies and ecological restoration of watersheds: A systematic review. In *Watershed Ecology and the Environment* (Vol. 7). https://doi.org/10.1016/j.wsee.2025.05.004
29. Mortaheb, R., & Jankowski, P. (2023). Smart city re-imagined: City planning and GeoAI in the age of big data. *Journal of Urban Management*, *12*(1). https://doi.org/10.1016/j.jum.2022.08.001
30. Munen, J. (2021). SURVEY OF CONTEMPORARY CHALLENGES FACED IN MARKETING OF LIFE INSURANCE POLICIES IN EMBU COUNTY. *Frontiers in Neuroscience*, *14*(1).
31. Murikah, W., Nthenge, J. K., & Musyoka, F. M. (2024). Bias and ethics of AI systems applied in auditing - A systematic review. *Scientific African*, *25*. https://doi.org/10.1016/j.sciaf.2024.e02281
32. Palazzo, F., Zambetta, G., & Palazzo, S. (2025). Artificial Intelligence at the

Crossroads of Engineering and Innovation. *Computing&AI Connect*, *2*(1). https://doi.org/10.69709/caic.2025.141254

33. Palazzo, S., Palazzo, F., & Zambetta, G. (2025). The Role of Artificial Intelligence in Advanced Engineering: Current Trends and Future Prospects. *Journal of Intelligent Communication*, *4*(1). https://doi.org/10.54963/jic.v4i1.959
34. Papaneophytou, C., & Nicolaou, S. A. (2025). Promoting Critical Thinking in Biological Sciences in the Era of Artificial Intelligence: The Role of Higher Education. In *Trends in Higher Education* (Vol. 4, Issue 2). https://doi.org/10.3390/higheredu4020024
35. Pepic, I., Feldt, R., Ljungström, L., Torkar, R., Dalevi, D., Maurin Söderholm, H., Andersson, L. M., Axelson-Fisk, M., Bohm, K., Sjöqvist, B. A., & Candefjord, S. (2021). Early detection of sepsis using artificial intelligence: a scoping review protocol. *Systematic Reviews*, *10*(1). https://doi.org/10.1186/s13643-020-01561-w
36. Perdana, R., Yulianti, D., & Bertiliya, W. A. (2025). Enhancing scientific literacy through AI-integrated Inquiry Social Complexity module in primary education. *Journal of Educational Management and Instruction (JEMIN)*, *5*(2). https://doi.org/10.22515/jemin.v5i2.10891
37. Posunko, A., Chavarha, M., Bashkirova, L., Vasylyuk-Zaitseva, S., & Abramchuk, O. (2025). Machine Learning in Medicine: A New Paradigm for Diagnosis and Treatment. *Futurity Medicine*, *4*(3). https://doi.org/10.57125/fem.2025.09.30.01
38. Rattan, P., Penrice, D. D., & Simonetto, D. A. (2022). Artificial Intelligence and Machine Learning: What You Always Wanted to Know but Were Afraid to Ask. In *Gastro Hep Advances* (Vol. 1, Issue 1). https://doi.org/10.1016/j.gastha.2021.11.001
39. Rincon, I. B., Soledispa-Cañarte, B. J., Sumba-Bustamante, R. V., del Carmen Burbano-Gómez, Z., & Jiménez-Granizo, F. P. (2023). Methodological Approaches and Paradigms in Scientific and Technological Research: a Bibliometric Review | Abordajes Metodológicos Y Paradigmas En La Investigación Científica Y Tecnológica: Una Revisión Bibliométrica. *Bibliotecas, Anales de Investigacion*, *19*(1).
40. Rincon Soto, I. B., Soledispa-Cañarte, B. J., Sumba-Bustamante, R. V., del Carmen

Burbano-Gómez, Z., & Jiménez-Granizo, F. P. (2023). METHODOLOGICAL APPROACHES AND PARADIGMS IN SCIENTIFIC AND TECHNOLOGICAL RESEARCH: A BIBLIOMETRIC REVIEW. *Bibliotecas, Anales de Investigacion*, *19*(1).

41. Rodríguez Flores, E. A., Garcés Giraldo, L. F., Valencia, J., & Valencia-Arias, A. (2025). Research trends in the use of artificial intelligence techniques in scientific research. *Revista Venezolana de Gerencia*, *30*(109). https://doi.org/10.52080/rvgluz.30.109.23
42. Saurbayev, R. Zh., Zhetpisbay, A. K., & Yerekhanova, F. T. (2024). Philological Studies: Computer Aspects of Stylometry Automation. *Iasaýı Ýnıversıtetiniń Habarshysy*, *133*(3). https://doi.org/10.47526/2024-3/2664-0686.72
43. Schneider, R., Angwin, C., Jenkinson, C., Jones, A. S. K., Jennison, C., Henley, W., Farmer, A., Sattar, N., Holman, R. R., Pearson, E., Shields, B., al., et, Blackwell, K., Donskih, R., Jones, C. M., Nixon, A., Vidal, M. J., Nakov, R., Singh, P., … Ewel, C. (2021). Effectiveness of olive oil for the prevention of pressure ulcers caused in immobilized patients within the scope of primary health care: study protocol for a randomized controlled trial. *Trials*, *21*(1).
44. Spivakovsky, O., Omelchuk, S., Kobets, V., Poltoratskyi, M., Malchykova, D., & Lemeshchuk, O. (2025). INSTITUTIONAL REGULATION OF USING AI-BASED SERVICES THROUGH THE KANO MODEL: THE CASE OF KHERSON STATE UNIVERSITY. *Information Technologies and Learning Tools*, *108*(4). https://doi.org/10.33407/itlt.v108i4.6269
45. Taha Kandil, T. (2025). Artificial intelligence (AI) and alleviating supply chain bullwhip effects: social network analysis-based review. *Journal of Global Operations and Strategic Sourcing*, *18*(1). https://doi.org/10.1108/JGOSS-04-2023-0038
46. Trejos-Gil, C. A., & Gómez-Monsalve, W. D. (2024). Artificial Intelligence in Media and Journalism. Systematic Review on Spain and Latin America in Scopus and Web of Science Databases (2018-2022). In *Palabra Clave* (Vol. 27, Issue 4). https://doi.org/10.5294/pacla.2024.27.4.1
47. Wang, C., Nan, J., Kang, Y., Xu, X., Li, T., & Zhu, Z. (2025). Design of a

comprehensive teaching experiment for mechanoluminescent human biomechanical detection. *Experimental Technology and Management*, *42*(9). https://doi.org/10.16791/j.cnki.sjg.2025.09.028

48. Williamson, B. (2020). New digital laboratories of experimental knowledge production: Artificial intelligence and education research. *London Review of Education*, *18*(2). https://doi.org/10.14324/LRE.18.2.05
49. Zhang, B., Liu, Q., Li, X., Liu, L., Yang, B., Letu, H., Gao, L., Zhang, W., Zhang, H., Bian, Z., Qi, M., Chen, C., & Shang, H. (2025). The core concepts and fundamental issues of remote sensing science. *National Remote Sensing Bulletin*, *29*(1). https://doi.org/10.11834/jrs.20244503
50. Zhao, T., Zhou, Y., Shi, R., Cao, Z., & Ren, Z. (2024). GWAI: Artificial intelligence platform for enhanced gravitational wave data analysis. *SoftwareX*, *28*. https://doi.org/10.1016/j.softx.2024.101930
51. Zheng, Z., He, Z., Khattab, O., Rampal, N., Zaharia, M. A., Borgs, C., Chayes, J. T., & Yaghi, O. M. (2023). Image and Data Mining in Reticular Chemistry Using GPT-4V. *ArXiv*.
52. Casanova Pistón, A., & Martínez Domínguez, M. (2024). Producción científica sobre Inteligencia Artificial y educación: un análisis cienciométrico. *Hachetetepé. Revista Científica de Educación y Comunicación*, *28*. https://doi.org/10.25267/hachetetepe.2024.i28.1102
53. Guliev, I. A., Mammadov, A., & Ibrahimli, K. (2024). The Use of Artificial Intelligence Technologies in Energy and Climate Security. *Review of Business and Economics Studies*, *12*(4). https://doi.org/10.26794/2308-944X-2024-12-4-58-71

Chapter 3.

Beyond the Binary: The Convergence of Qualitative Inquiry and Artificial Intelligence

The landscape of social scientific inquiry is currently undergoing a seismic transformation, characterized by the collapsing of the traditional dichotomy between qualitative and quantitative research. Historically, these two domains have been viewed as distinct, often opposing, epistemological camps: qualitative research focused on "thick description," interpretivism, and the subjective nuance of the human experience, while quantitative research prioritized generalizability, statistical significance, and the objective measurement of variables. However, the rapid ascent of Artificial Intelligence (AI)—specifically Large Language Models (LLMs), Natural Language Processing (NLP), and machine learning (ML)—has birthed a hybrid discipline: Qualitative Data Science.[1]

This emerging field does not merely represent the application of computational tools to textual data; rather, it fundamentally reconfigures the epistemology of social inquiry. It promises to resolve the "scale vs. depth" trade-off that has long plagued researchers, offering the analytical breadth of "big data" while attempting to preserve the interpretive depth of hermeneutics.[3] As unstructured data—from social media feeds to digitized historical archives—proliferates, the ability to analyze millions of documents with qualitative sensitivity is no longer a luxury but a necessity for addressing what scholars term the "futures knowledge deficit".[4]

However, this convergence is fraught with profound methodological and ethical tensions. The integration of algorithmic systems into qualitative workflows raises critical questions about "positivist drift"—the risk that the statistical logic of AI will flatten the rich, subjective variance of human meaning into standardized, computable categories.[5]

Furthermore, the deployment of AI in social analysis introduces the threat of "epistemological violence," where Western-centric algorithms silence or misinterpret Indigenous and marginalized knowledge systems, perpetuating a form of "algorithmic colonialism".[7]

This report provides an exhaustive, expert-level analysis of this convergence. It dissects the theoretical frameworks enabling this shift, such as Computational Grounded Theory and Blended Reading; it scrutinizes the specific technological architectures facilitating it, from BERT embeddings to GPT-4's chain-of-thought reasoning; and it evaluates the practical applications of these tools in diverse sectors, from clinical health research to corporate ethnography. Ultimately, it argues that the future of qualitative inquiry lies not in the automation of the researcher, but in the cultivation of "Safe Qualitative AI"—systems designed to function as dialogic partners that enhance, rather than replace, human meaning-making.[3]

1. The Epistemological Crisis and the Rise of Hybrid Methodologies

The integration of artificial intelligence into the qualitative domain is not simply a technical upgrade; it is an epistemological event. It forces a confrontation between two distinct theories of knowledge: the Positivist assumption that truth is objective, observable, and scalable, and the Interpretivist/Constructivist assumption that truth is subjective, socially constructed, and context-dependent.

1.1 The Schism: "Thick Description" vs. "Big Data"

For decades, qualitative research has prided itself on its "human-centered" approach, relying on the researcher's subjectivity as a primary instrument of analysis. Methods like ethnography and phenomenology require deep immersion in the field, producing "thick descriptions" that capture the cultural webs of meaning in which social actions are suspended. In contrast, the "Big Data" revolution of the early 21st century was largely a quantitative phenomenon, driven by the belief that with enough data, theory becomes unnecessary—a concept famously critiqued as the "end of theory."

The rise of Qualitative Data Science challenges this schism. It posits that computational methods can be used to perform "distant reading" of massive corpora to identify structural patterns, which can then be interrogated through "close reading" to derive meaning.[10] This suggests a move towards Methodological Hybridity, where the boundaries between the "qualitative" and "quantitative" are porous. As noted in recent scholarship, AI-driven methodologies are showing improvements in consistency and reproducibility compared to standard qualitative methods, yet they are met with resistance from scholars who fear the loss of the "human element" essential for true interpretivism.[12]

1.2 Theoretical Frameworks for the Digital Age

To navigate this new landscape, scholars have developed rigorous frameworks that justify the use of computation within interpretivist paradigms. These frameworks move beyond viewing software as a passive container for codes (as in early Computer-Assisted Qualitative Data Analysis Software, or CAQDAS) to viewing algorithms as active, generative agents in the analytical process.

1.2.1 Computational Grounded Theory (CGT)

One of the most robust and systematically developed frameworks is Computational Grounded Theory (CGT), proposed by Laura Nelson and others to bridge the gap between expert human knowledge and the pattern recognition capabilities of computers.[13] Traditional Grounded Theory, developed by Glaser and Strauss, relies on inductive reasoning—building theory from the data up—but has historically been limited by the human capacity to process large volumes of text (the "small N" problem).

CGT addresses this scalability issue without sacrificing the recursive, iterative nature of the method. The framework operates through a rigorous three-step process, transforming the "black box" of machine learning into a transparent partner in theory generation [14]:

- Step 1: Pattern Detection (Computational/Inductive) This initial phase involves the inductive computational exploration of text. The researcher employs unsupervised machine learning techniques, such as topic modeling, clustering, or word scores, to "read" the entire corpus. This step is purely inductive; the algorithm identifies latent structures, lexical co-occurrences, and novel patterns that might escape human notice due to cognitive bias or sheer data volume. It serves as a "lens" or a "map" to view the data's topography, highlighting outliers and clusters that warrant attention.13
- Step 2: Pattern Refinement (Qualitative/Interpretive) This is the critical "human-in-the-loop" phase. The researcher returns to the data with a hermeneutic approach, engaging in "deep reading" of the computationally identified clusters. Here, the raw outputs of the machine—which are merely statistical associations—are interrogated, contextualized, and refined into meaningful sociological or psychological concepts. The machine suggests a pattern (e.g., words "home" and "trap" appearing together); the human interprets the meaning (e.g., "domestic confinement during pandemic lockdowns").14 This step honors the interpretivist commitment to context and meaning.
- Step 3: Pattern Confirmation (Computational/Deductive) In the final phase, the researcher uses further computational techniques (such as supervised learning, specific NLP queries, or network analysis) to test the validity of the refined patterns across the entire corpus. This step ensures that the insights derived from deep reading are not idiosyncratic to a few documents but are representative of the broader dataset. It provides a measure of rigor and reproducibility often lacking in purely manual qualitative analysis.14

This framework allows for "resampling"—an iterative process of accessing the field, creating a model, and computationally searching for relevant cases—thereby integrating the principle of constant comparison at a scale previously impossible.[17]

1.2.2 Blended Reading and the Digital Humanities

Parallel to developments in sociology, the Digital Humanities have pioneered the concept of Blended Reading. This methodology reconciles "close reading" (the careful, nuanced analysis of individual texts) with "distant reading" (the quantitative analysis of massive corpora, as popularized by Franco Moretti).[10]

Blended Reading acknowledges that in the digital age, the "classic duality of interpreter and text has changed" due to the immense volume of digitally available data.[10] It proposes a modular analysis process where computational tools are used not to replace the scholar, but to augment their capacity to navigate "archives of abundance (see Table 5)."

Table 5: Distant, close and blended reading

Component	Methodological Focus	Role of AI/Computation	Analytical Outcome
Distant Reading	Macro-analysis, structural trends, metadata	Identifies outliers, aberrations, inconsistencies, and large-scale temporal shifts across millions of words.[11]	A "topological map" of the discourse; identification of "hot spots" for further analysis.
Close Reading	Micro-analysis, nuance, sentiment, cultural context	None; relies on human hermeneutics. Provides the "ground truth" and interpretive depth.[18]	"Thick description" of specific texts; validation of computational findings.
Blended Reading	Integration of macro and micro	Creates a feedback loop: distant reading guides text selection for close reading; close reading refines the algorithms for distant	A multi-scalar understanding that links individual narrative to structural phenomenon.

		reading.[19]	

This approach is particularly vital for mitigating the "black box" problem. By forcing a dialogue between the algorithmic output and the raw text, Blended Reading ensures that the researcher remains anchored in the data's reality while leveraging the computer's ability to see the "longue durée" or structural topology of the discourse.[11]

1.2.3 Abductive Analysis and Machine Learning

While induction builds theory from data and deduction tests theory against data, Abduction is the logic of discovery—the leap to the best explanation for a surprising observation. Machine learning is increasingly viewed by methodologists as an engine for abductive analysis.[20]

In this context, algorithms act as generators of "surprising observations." When an unsupervised model clusters data in a way that defies existing theoretical expectations—for example, clustering "economic anxiety" with "health optimization" in a dataset of political discourse—it creates a "breakdown" in understanding. This breakdown invites abductive reasoning, forcing the researcher to generate a new hypothesis to explain this juxtaposition.[21]

Scholars argue that ML systems are inherently abductive because they rely on the contingent biases of their training data to make predictions, effectively "guessing" the nature of new data based on a learned model of the world.[22] This makes them powerful tools for "conjectural narrativization," helping qualitative researchers identify non-obvious relationships that require theoretical elaboration.[22]

1.3 The "Quantitized Qualitative" Paradigm

A growing trend within this convergence is the "Quantitized Qualitative" paradigm, which involves the systematic transformation of qualitative observations into quantitative data without losing the descriptive validity of the original observation.[23]

- **Methodology:** Qualitative themes derived from interviews or open-ended survey responses are converted into binary (0/1) or ordinal variables. For instance, a theme of "distrust in medical authority" identified in a patient interview is coded as a variable.
- **Application:** These variables are then used in advanced statistical models, such as regression models or Structural Equation Modeling (SEM), to test causal relationships between qualitative themes and quantitative outcomes (e.g., health status).[25]
- **Collinearity Risks:** A major critique of this approach is the risk of "collinearity," where response categories are linked because of the coding strategy rather than reality. Mixed methods researchers must rigorously validate that their "quantitized" data retains its semantic integrity and is not an artifact of the coding frame.[23]

Structural Topic Modeling (STM) represents the state-of-the-art in this paradigm. Unlike standard Latent Dirichlet Allocation (LDA), which treats documents as bags of words, STM allows researchers to incorporate metadata (covariates) into the model. This means the model can estimate how the prevalence of a qualitative topic varies according to attributes like the author's gender, date of publication, or political affiliation.[26] STM essentially automates the "quantitizing" process, preserving the semantic richness of the text while allowing for rigorous statistical testing of how themes relate to external variables.[28]

2. The Technics of Qualitative AI: Tools, Architectures, and Performance

The theoretical frameworks described above are implemented through a rapidly evolving stack of technologies. The shift from simple keyword searching to "semantic understanding" has been driven by the evolution of Natural Language Processing (NLP) architectures, specifically Transformers.

2.1 Large Language Models (LLMs) in Coding and Analysis

The release of generative models like **ChatGPT** (OpenAI), **Claude** (Anthropic), and **Gemini** (Google) has "upended scientific and educational paradigms," specifically in the domain of qualitative coding.[29] These models possess an unprecedented ability to parse syntax, semantics, and pragmatics, allowing them to perform tasks that previously required human intuition.

2.1.1 Performance and Reliability: Human vs. Machine

Research comparing human coders to LLMs reveals a complex landscape where performance is highly dependent on the model's size and the sophistication of the prompting strategy.

- **Coding Fidelity:** Studies utilizing GPT-4 have demonstrated "human-equivalent interpretations" for certain socio-historical codes, achieving high inter-coder reliability (Cohen's Kappa \ge 0.79).[30] In direct comparisons, GPT-4 significantly outperformed earlier models like GPT-3.5, which struggled with nuance and achieved much lower reliability scores (Mean Kappa = 0.34).[30] This highlights that the *capacity for nuance* is a function of model scale and architectural sophistication.
- **Chain-of-Thought (CoT) Prompting:** The "push-button" efficacy of these tools is largely illusory without sophisticated interaction. Research indicates that **Chain-of-Thought (CoT) prompting**—where the model is instructed to explain its reasoning *before* assigning a code—considerably improves coding fidelity.[30] This mirrors the human qualitative process of writing memos or justifications for coding decisions, suggesting that AI performs best when forced to simulate the reflexive steps of a human researcher.
- **Cost and Efficiency:** Automated analysis using LLMs is significantly more cost-effective than human coding, often reducing the time and financial investment by orders of magnitude. However, this comes at the cost of specificity and depth.[31]

2.1.2 The Problem of Consistency and Hallucination

Despite their promise, LLMs struggle with consistency, a trait often referred to as being a "stochastic parrot."

- **Inconsistent Output:** "LLMq" (Large Language Model qualitative) values can stabilize over iterations, but the actual analytical output may remain inconsistent across runs.[32] A code identified in one pass may be missed in the next, or a quote attributed to a specific theme may be hallucinated. This unpredictability conflicts with the rigorous audit trails required in qualitative inquiry.[5]
- **Prompt Sensitivity:** LLMs can be "distracted" by the vagaries of natural language interfaces. Slight variations in prompts—even those that appear semantically identical to a human—can lead to divergent analytical outcomes.[33] This fragility necessitates a new form of methodological rigor: "Prompt Engineering as Research Method".[34]

2.2 BERT and Embedding Models: Mapping the Semantic Space

While LLMs generate text, BERT (Bidirectional Encoder Representations from Transformers) and similar embedding models are used to map the *semantic space* of a dataset. These models convert text into high-dimensional vectors, allowing researchers to measure the mathematical distance between concepts.

- **Thematic Clustering:** In thematic analysis, BERT has been employed to cluster interview transcripts and open-ended survey responses with high precision.[35] The **BERT+UMAP+HDBSCAN** pipeline—which first generates embeddings (BERT), reduces their dimensionality (UMAP), and then clusters them (HDBSCAN)—has been identified as particularly effective for semi-structured interviews. This approach yields topics that are both diverse and interpretable, often outperforming traditional LDA models in coherence.[36]
- **Predictive Validity:** Studies show that BERT-based topic modeling (like BERTopic) can outperform traditional human coding in predicting specific variables. For instance, in an analysis of 552 psychotherapy transcripts, BERT-derived topics related to "negative experiences" successfully predicted symptom severity.[37]

- **The "Black Box" Limitation:** However, the use of these models introduces a persistent "black box" element. While a human coder can explain *why* they grouped two statements (e.g., "both reflect underlying anxiety"), an embedding model groups them based on vector proximity. This vector proximity is a mathematical abstraction of semantic usage, which may or may not correspond to a conceptual or theoretical link meaningful to a sociologist.[38]

2.3 The Evolution of CAQDAS (Computer-Assisted Qualitative Data Analysis Software)

The market for Qualitative Data Analysis software is shifting rapidly as traditional players integrate AI and new "AI-native" tools emerge (see Table 6).

Table 6: Tool category of key platforms

Tool Category	Key Platforms	AI Integration Features	Strengths & Weaknesses
Traditional CAQDAS	**ATLAS.ti, NVivo, MAXQDA**	**ATLAS.ti:** "AI Coding" (powered by OpenAI) automates initial coding passes, reducing time from days to hours.[39] **NVivo:** AI for sentiment analysis and autocoding by theme; features framed as assistive.[40]	**Strength:** deeply integrated into established workflows; robust data management. **Weakness:** AI features can feel "bolted on"; legacy interface complexity.
AI-Native Tools	**Delve, AILYZE, Looppanel, Dovetail**	**Delve:** "AI-assisted" workflows; positions AI as a "peer debriefer" or "junior researcher" to suggest themes.[41] **Looppanel:**	**Strength:** intuitive, built for speed and collaboration. **Weakness:** can over-simplify analysis; risk of generating surface-level "topic

		Specialized for UX/Focus groups; auto-tags sentiment and behavior patterns.[42]	summaries" rather than deep themes.[44]
Specialized Research AI	**Infranodus**	Uses network analysis and BERT to visualize discourse structure and identify "structural gaps" in narratives.[35]	**Strength:** visualizes the topology of ideas. **Weakness:** steeper learning curve for non-technical researchers.

Critically, while tools like ATLAS.ti promise to "remove the headache" of coding, methodological experts warn that they can only generate descriptive themes that "barely scratch the surface" of the data's true depth.[44] The consensus is that these tools are best for **deductive** coding (finding known patterns) rather than **inductive** theory building (discovering new meaning).[45]

3. Ethics, Bias, and the Politics of Algorithms: "Epistemological Violence"

The use of AI in qualitative research introduces profound ethical questions that go far beyond standard concerns of privacy and data security. There are deeper issues regarding the *politics of knowledge* produced by algorithmic systems—issues that threaten to undermine the very purpose of qualitative inquiry.

3.1 Epistemological Violence and Algorithmic Colonialism

A critical area of concern, drawing on postcolonial theory, is **Epistemological Violence**. This concept, articulated by scholars like Gayatri Chakravorty Spivak and Miranda Fricker (via "testimonial injustice"), refers to the harm inflicted when a dominant knowledge system silences, invalidates, or marginalizes the knowledge systems of others.[7]

In the context of AI, this violence is structural. AI models are trained on vast corpora of text scraped from the internet (e.g., Common Crawl), which is disproportionately English, Western, and hegemonic. When these models are used to analyze data from non-Western cultures or marginalized communities, they inevitably impose Western frameworks of understanding on that data.[47] This is not merely "bias" in the statistical sense; it is a form of Algorithmic Colonialism that erases local nuance and enforces a "monoculture of the mind".[7]

- **Invalidation of Knowledge:** AI systems may flag Indigenous concepts, non-linear narrative structures, or dialect-specific epistemologies as "incoherent," "irrelevant," or "errors" because they do not fit the statistical patterns of the dominant training data.[49]
- **The "Othering" of Data:** Research on migration in the Balkans has shown how "epistemic violence" is central to the EU's border regime. Similarly, AI systems used in asylum processing can perpetuate "testimonial injustice" by failing to recognize the credibility of asylum seekers' narratives due to linguistic or cultural mismatches.[46]
- **Indigenous Data Sovereignty:** In response, groups like the *Indigenous Protocol and Artificial Intelligence Working Group* argue for "decolonial AI." This involves developing systems trained on and governed by Indigenous data sovereignty principles, ensuring that the technology respects distinct epistemologies rather than flattening them.[51]

3.2 Automated Bias: Intrinsic and Extrinsic

Qualitative researchers have long acknowledged their own subjectivity (reflexivity). However, AI systems are often falsely perceived by stakeholders as neutral or objective. In reality, AI harbors deep-seated biases:

- **Intrinsic Bias:** This stems from the training data itself. Generative AI tends to reproduce dominant discourses and stereotypes found in its corpus. In qualitative coding, an AI might consistently code descriptions of poverty as "personal failure" rather than "structural inequality" if its training data reflects neoliberal ideologies that prioritize individual responsibility.[5]

- **Extrinsic Bias:** This emerges from the deployment context. A model that works "fairly" in a Western clinical setting may fail catastrophically when applied to a Global South context due to different cultural expressions of symptoms or distress.[47]
- **The "Formalism Trap":** Organizations often fall into the trap of thinking that using AI makes a process "fair" or "objective" purely because it is mathematical. This ignores the fact that fairness is a constructed, contextual social concept, not a mathematical constant. Qualitative researchers have empirically illustrated how organizations fall into this trap, failing to account for the "full meaning of social concepts" when they delegate decision-making to algorithms.[52]

3.3 Qualitative Auditing of Algorithms

Interestingly, qualitative methods are becoming the primary tool for *fixing* these quantitative problems. Qualitative Auditing involves using ethnography, interviews, and document analysis to inspect how algorithms are actually functioning in society.[53]

Case Study: Hiring Algorithms – An Ethnography of Fairness

In a landmark ethnographic study of a multinational company ("MultiCo") implementing AI for hiring, researchers van den Broek, Sergeeva, and Huysman revealed the "formalism trap" in practice.

- **Pre-AI:** The HR team viewed themselves as the "guardians of fairness," defined as suppressing bias through human judgment.
- **Post-AI:** The implementation of AI shifted the definition of fairness. Fairness became synonymous with "consistency" (everyone gets the same algorithm) and "accuracy" (predictive validity).
- **The Shift:** The researchers observed that the HR team began to "enroll" other stakeholders by emphasizing these new, narrower definitions of fairness. The qualitative inquiry uniquely uncovered that the *use* of AI did not just automate the process; it fundamentally altered the *values* of the organization, shifting focus from "candidate experience" to "operational efficiency".[52] This demonstrates the power of

qualitative auditing to reveal the socio-technical reality of AI systems.

4. Applied Case Studies and Sector Analysis

The theoretical and ethical debates are grounded in a growing body of applied research. Across health, education, and corporate sectors, the "Quantitized Qualitative" approach is yielding mixed but potent results.

4.1 Health and Clinical Research: HIV and Psychotherapy

In health research, qualitative data (patient narratives) is critical for understanding the "lived experience" of illness, but the volume of data often limits sample sizes. AI is being used to scale this analysis.

- **HIV Research Ethics:** A commentary from Johns Hopkins researchers emphasizes that while AI can efficiently code HIV-related qualitative data (e.g., identifying risk behaviors in text), it must be aligned with the "underlying epistemology" of the study. Using AI to categorize "risk behaviors" is methodologically feasible but using it to interpret the "lived experience of stigma" risks pathologizing patients if the AI lacks the necessary cultural nuance.[29] The researchers propose a framework where the *goal* of the research determines the appropriateness of AI: strictly descriptive tasks are AI-suitable, while phenomenological interpretation remains human-bound.
- **Psychotherapy and Topic Modeling:** A mixed-methods study used BERTopic to analyze 552 psychotherapy transcripts. The goal was to predict symptom severity and "therapeutic alliance" (the bond between therapist and patient).
  - **Findings:** The model successfully identified topics (e.g., "negative experiences," "health") that strongly correlated with symptom severity ($r=0.45$). This demonstrates that "quantitized" qualitative themes can be robust predictors in clinical models.
  - **Nuance:** However, the "therapeutic alliance" was better predicted by the *therapist's* speech patterns than the patient's, a subtle dynamic that required qualitative

domain expertise to interpret. The study concludes that AI allows for "treatment-relevant metrics" to be predicted with reasonable accuracy, but only when "explainable AI" (XAI) techniques are used to validate the topics.[37]

4.2 Corporate Ethnography: The "Future of Work" at Anthropic

One of the most revealing studies of AI's impact comes from within the industry itself. Anthropic conducted an internal qualitative study, involving 53 in-depth interviews with their own engineers, to understand how using their model (Claude) was changing their work practices.[57]

- **Loss of Craft:** Engineers reported a complex emotional response. While productivity soared, many expressed a "sense of loss" regarding the act of coding itself. One engineer noted, "I thought that I really enjoyed writing code, and I think instead I actually just enjoy what I get out of writing code."
- **Shift in Social Dynamics:** The study found that AI was altering workplace socialization. Claude became the "first stop" for questions that previously would have been directed to senior colleagues. This led to a decrease in mentorship opportunities and human collaboration ("I like working with people and it's sad that I 'need' them less now").
- **Existential Uncertainty:** Employees expressed "genuine uncertainty" about the future of their profession, with some fearing they were "automating themselves out of a job."
- **Methodological Significance:** This study underscores that even in a hyper-quantitative environment like an AI lab, **qualitative methods** (interviews, phenomenology) were deemed essential to understand the *human* impact of the technology. Quantitative metrics could measure code output, but only qualitative inquiry could reveal the shifting professional identity of the engineers.

4.3 Education and Focus Groups

Focus groups present a unique challenge for AI due to the multi-speaker dynamics and overlapping speech.

- **Focus Group Analysis:** Tools like **Looppanel** and **Dovetail** are increasingly used to transcribe and analyze focus group data. AI excels at tracking "who said what" and identifying high-level consensus. However, it often fails to capture the "group effect"—the specific dynamic of *how* participants influence each other's opinions, which is the methodological core of focus group research.[58]
- **Student Evaluations:** Structural Topic Modeling (STM) has been applied to analyze nearly 300,000 open-ended student evaluations of teaching. The study found that topic correlations were consistent across instructor genders, challenging the persistent narrative that male and female instructors are evaluated on fundamentally different criteria.[28] This finding, which has significant policy implications for higher education, was only possible due to the *scale* afforded by AI analysis of qualitative comments.

5. The "Safe Qualitative AI" Manifesto

In response to the dangers of "positivist drift" and "epistemic violence," a movement for "Safe Qualitative AI" has emerged. This framework argues that researchers should not reject AI, but rather build "dedicated qualitative AI systems" from the ground up, designed specifically for interpretive goals rather than borrowing general-purpose tools optimized for commercial tasks.[9]

5.1 Principles of Safe Qualitative AI

The "Safe Qualitative AI" framework outlines several core design principles intended to preserve the integrity of qualitative inquiry:

1. **Explanatory over Action-Oriented:** AI systems should be designed to build explanatory models (why did this happen? what does it mean?) rather than just predictive models (what will happen next?).[60] This aligns with the "Scientist AI" concept, where the goal is understanding rather than optimization.
2. **Explicit Uncertainty Quantification:** The AI should articulate the limits of its

interpretation. Instead of presenting a code as an authoritative fact, the system should express the *certainty* of its coding and provide alternative interpretations.[3]

3. **Transparency and Reproducibility:** The "black box" must be opened. Researchers need to see *how* the AI arrived at a theme. This necessitates tools that provide the "chain of thought" or the specific vector path used for classification, allowing for a rigorous audit of the machine's "reasoning".[9]
4. **Privacy-First Architecture:** Given the sensitivity of qualitative data (often involving vulnerable populations), Safe Qualitative AI must operate on local or private architectures. Data should never be fed back into public base models for training, ensuring strict adherence to confidentiality protocols.[9]

5.2 Human-in-the-Loop (HITL) Workflows

The consensus across the literature is that AI cannot replace the qualitative researcher; it must be a Human-in-the-Loop (HITL) system.

- **Augmentation, not Automation:** AI is best used for "low-stakes," "boring," or "verifiable" tasks (e.g., initial open coding, transcription cleaning), allowing the researcher to focus on high-level synthesis, theory building, and "taste-based" judgments.[57]
- **The Dialogic Partner:** The AI should be treated as a "dialogic partner"—an entity to argue with, to test hypotheses against, and to use for "peer debriefing." The goal is to use the AI to *challenge* the researcher's bias, not just to confirm it.[40]
- **Final Interpretive Authority:** The human must always retain the final say on the validity of a code or theme. AI suggestions are treated as *heuristic starting points*, not endpoints.[61]

6. Future Horizons: The Quantitized Qualitative Researcher

As we look forward, the boundaries between the qualitative and quantitative are likely to blur further, giving rise to a new professional identity: the Quantitized Qualitative Researcher.

6.1 The Rise of "Scientist AI"

The future trajectory involves AI moving from a passive tool to an active methodologist. The concept of Scientist AI envisions systems that can:

- **Propose Hypotheses:** Scan vast literatures to identify theoretical gaps and propose new hypotheses.
- **Design Sampling Strategies:** Analyze current data saturation in real-time and suggest which stakeholders to interview next to achieve maximum variation.[62]
- **Synthetic Piloting:** Conduct "synthetic interviews" with persona-based AI agents to pilot interview guides before entering the field, allowing researchers to refine their questions.[62]

6.2 Professional Implications

This shift will require a re-skilling of the social science workforce. Researchers will need to be fluent in "Prompt Engineering as Methodology" and understand the basics of Vector Semantics to effectively audit their tools. The danger is a bifurcation of the discipline into "purists" who reject AI and "computationalists" who embrace it. However, the most impactful research will likely come from those who can inhabit the middle ground—using "Safe Qualitative AI" to scale their inquiry while maintaining the "thick description" that makes qualitative research irreplaceable.

7. Conclusions and Strategic Outlook

The integration of AI into qualitative research represents a second Copernican Revolution in the social sciences—a shift from the human as the sole center of meaning-making to a system where meaning is co-constructed by humans and algorithms.[63]

Key Takeaways and Recommendations:

1. **Hybridity is Inevitable:** The volume of digital data makes purely manual qualitative analysis increasingly untenable for large-scale societal questions. Methodologies like **Computational Grounded Theory** and **Blended Reading** are the necessary adaptations to this new reality. They offer a rigorous path to integrating "big data" with "thick description."
2. **Epistemology First, Technology Second:** The successful use of AI depends not on the power of the model, but on the clarity of the epistemological framework. Researchers must explicitly define *why* they are using AI (e.g., for pattern detection, not truth verification) to avoid "positivist drift."
3. **Ethics as Method:** Auditing for "epistemological violence" is no longer an optional ethical step; it is a core methodological requirement. Researchers must actively interrogate their tools for Western-centric bias and algorithmic colonialism, prioritizing tools that allow for local / Indigenous data sovereignty.
4. **The Human Remains Essential:** Far from becoming obsolete, the qualitative researcher's role is elevating. The value shifts from the *labor* of coding to the *insight* of synthesis. The researcher becomes the guarantor of validity, the auditor of the algorithm, and the bridge between mathematical pattern and human meaning.

In conclusion, AI offers qualitative research a telescope to see the universe of data, but the researcher must remain the astronomer who interprets the stars. The goal is not to automate the understanding of the human condition, but to deepen it through a reflexive, critical, and "safe" collaboration with the machine (see Table 7).

Table 7: Key Methodological Comparisons

Feature	Traditional Qualitative	Computational Grounded Theory	Fully Automated (AI) Analysis
Primary Processor	Human Brain	Human + Algorithm	Algorithm

		(Iterative)	(LLM/BERT)
Scale	Small (N < 100)	Large (N > 1,000s)	Massive (N > 1,000,000s)
Logic of Inquiry	Inductive / Abductive	Hybrid (Inductive detection, Qualitative refinement)	Predominantly Deductive / Pattern Matching
Role of Context	Deep, "Thick Description"	Contextualized Patterns	Often Decontextualized / "Flattened"
Key Risk	Researcher Bias / Burnout	"Black Box" opacity	Hallucination / Epistemological Violence
Outcome	Rich Narrative Theory	Reproducible, Scalable Theory	Surface-level Topics / Quantitized Data

Chapter bibliographies

1. Abdul Rahman, H., Masri, N. A., Husaini, A., Shaharuddin, M. Y., & Zaim, S. N. N. (2025). Integration of Machine Learning and Artificial Intelligence Techniques for Qualitative Research: The Rise of New Research Paradigms. *International Journal of Qualitative Methods* , *24*. https://doi.org/10.1177/16094069251347545
2. Access, O., Adams, W. C., Alfonita, F., Alfred, M., Briggs, S., Aloisi, A., Amorim, H., Moda, F., Anderson, D. N., Anwar, M. A., Graham, M., Apouey, B., Roulet, A., Solal, I., Stabile, M., Background, P., Bajwa, U., Gastaldo, D., di Ruggiero, E., … ZACCARIA, M. L. (2019). No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析Title. *Economic and Labour Relations Review*, *14*(1).
3. Anis, S., & French, J. A. (2023). Efficient, Explicatory, and Equitable: Why Qualitative Researchers Should Embrace AI, but Cautiously. In *Business and Society* (Vol. 62, Issue 6). https://doi.org/10.1177/00076503231163286
4. Asiva Noor Rachmayani, Inaam Akhtar, M., Njie, B., Asimiran, S., Alase, A., Etikan, I., Laughlin, R., Francis, J. J., Johnston, M., Robertson, C., Glidewell, L., Entwistle, V., Eccles, M. P., Grimshaw, J. M., Alhojailan, M. I., Ibrahim, M., Braun, V., Clarke, V., Taherdoost, H., … Zairul, M. (2019). No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析Title. *ISRA International Journal of Islamic Finance*, *8*(1).
5. Avula, S., Chadwick, G., Arguello, J., Capra, R., Lippold, D., Bana e Costa, C. A., de Corte, J.-M., Vansnick, J.-C., Linstone, H. A., Carver, C. S., Scheier, M. F., Segerstrom, S. C., Berger, G., Armand, L., Bresard, S., Grimanelli, P., Landucci, A., Schwartz, B., Bolger, F., … Birbilis, G. (2015). Editorial Board. *Technological Forecasting and Social Change*, *77*(1).
6. Bhatia, A., Chandani, A., Atiq, R., Mehta, M., & Divekar, R. (2021). Artificial intelligence in financial services: a qualitative research to discover robo-advisory services. *Qualitative Research in Financial Markets*, *13*(5).

https://doi.org/10.1108/QRFM-10-2020-0199

7. Brailas, A. (2025). Artificial Intelligence in Qualitative Research: Beyond Outsourcing Data Analysis to the Machine. *Psychology International*, *7*(3). https://doi.org/10.3390/psycholint7030078
8. Brondani, M., Alves, C., Ribeiro, C., Braga, M. M., Garcia, R. C. M., Ardenghi, T., & Pattanaporn, K. (2024). Artificial intelligence, ChatGPT, and dental education: Implications for reflective assignments and qualitative research. *Journal of Dental Education*, *88*(12). https://doi.org/10.1002/jdd.13663
9. Cheekaty, S., & Muneeswari, G. (2024). Exploring Sparse Gaussian Processes for Bayesian Optimization in Convolutional Neural Networks for Autism Classification. *IEEE Access*, *12*. https://doi.org/10.1109/ACCESS.2024.3351168
10. Christou, P. (2023). How to Use Artificial Intelligence (AI) as a Resource, ow Methodological and Analysis Tool in Qualitative Research? *Qualitative Report*, *28*(7). https://doi.org/10.46743/2160-3715/2023.6406
11. Christou, P. A. (2023a). A Critical Perspective Over Whether and How to Acknowledge the Use of Artificial Intelligence (AI) in Qualitative Studies. *Qualitative Report*, *28*(7). https://doi.org/10.46743/2160-3715/2023.6407
12. Christou, P. A. (2023b). The Use of Artificial Intelligence (AI) in Qualitative Research for Theory Development. *Qualitative Report*, *28*(9). https://doi.org/10.46743/2160-3715/2023.6536
13. de Seta, G., Pohjonen, M., & Knuutila, A. (2024). Synthetic ethnography: Field devices for the qualitative study of generative models. *Big Data and Society*, *11*(4). https://doi.org/10.1177/20539517241303126
14. Durrani, B. A., Lecturer, S., Ivanauskien, N., Volung, J., Rahmani-Nejad, L., Firoozbakht, Z., Taghipoor, A., Arokiasamy, A. R. A., Pourdehghan, A., Tjan, S., Sukamto, R., Lumintan, D. B., Dhurup, M., Mafini, C., Dumasi, T., Kheng, L. L., Mahamad, O., Ramayah, T., Mosahab, R., … Wirawan, V. (2014). Implementation of E-Government in Welcoming the Contemporary Industrial Revolution 4.0 Era in Indonesia. *International Journal of Marketing Studies*, *1*(1).

15. Fahyuni, E. F., Wasis, Bandono, A., & Arifin, M. B. U. B. (2020). Integrating islamic values and science for millennial students' learning on using seamless mobile media. *Jurnal Pendidikan IPA Indonesia*, *9*(2). https://doi.org/10.15294/jpii.v9i2.23209

16. Fieldhouse, J. K., Ge, J., Randhawa, N., Wolking, D., Genovese, B. N., Mazet, J. A. K., & Desai, A. (2025). The intersection of artificial intelligence with qualitative or mixed methods for communicable disease research: a scoping review. In *Public Health* (Vol. 248). https://doi.org/10.1016/j.puhe.2025.105961

17. Flores, L. G. R., & Mejía Rivera, K. A. (2024). ARTIFICIAL INTELLIGENCE IN QUALITATIVE RESEARCH: BIBLIOMETRIC ANALYSIS OF SCIENTIFIC PRODUCTION INDEXED IN SCOPUS. *New Trends in Qualitative Research*, *20*(4). https://doi.org/10.36367/ntqr.20.4.2024.e1116

18. Friedman, C., Owen, A., & VanPuymbrouck, L. (2025). Should ChatGPT help with my research? A caution against artificial intelligence in qualitative analysis. *Qualitative Research*, *25*(5). https://doi.org/10.1177/14687941241297375

19. Göçer, Ö., Hua, Y., Göçer, K., Gunhan, S., Moaveni, S., Chou, K. C., Charlson, J., Chinyio, E., Zhu, Y. Q., Zuo, J. J., Korpela, J., Kerosuo, H., Abuelmaatti, A., Ahmed, V., Jing, X., Wang, Q. K. Q., Li, Y. Y. Y., Zhao, D., McCoy, A. P., … Wu, W. W. Y. (2014). A Ship in a Box. *Applied Mechanics and Materials*, *29*(2).

20. Goyanes, M., Lopezosa, C., & Jordá, B. (2025). Thematic analysis of interview data with ChatGPT: designing and testing a reliable research protocol for qualitative research. *Quality and Quantity*, *59*(6). https://doi.org/10.1007/s11135-025-02199-3

21. Gunewardena, D., Seck, A., Warnecke, T., Green, E. P., Blattman, C., Jamison, J., Annan, J., Dvouletý, O., Pepelasis, I. S., Protogerou, A., Krasniqi, B. A., Williams, N., Osirim, M. J., Tobias, J. M., Boudreaux, K. C., Botterill, L. C., Free, M., Demirgüc-Kunt, A., Klapper, L. F., … Naszályi, P. (2018). Abstracts. *Journal of Planning Literature*, *29*(2).

22. Güzer, B., Caner, H., Fathoni, A. F., van der Mescht, H., Whittaker, J. K., Staker, H., Horn, M. B., Crawford, R., Richey, R. C., Christensen, C. M., Horn, M. B., Staker, H., Thoms, C. L. v, Chen, J. J., Petty, J., Albiladi, W. S., Alshareef, K. K., Zhang, W., Zhu,

C., … Province, Q. (2019). No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析Title. *Computers and Education*, *7*(1).

23. Hogg, H. D. J., Al-Zubaidy, M., Keane, P. A., Hughes, G., Beyer, F. R., & Maniatopoulos, G. (2023). Evaluating the translation of implementation science to clinical artificial intelligence: a bibliometric study of qualitative research. In *Frontiers in Health Services* (Vol. 3). https://doi.org/10.3389/frhs.2023.1161822
24. Hogg, H. D. J., Sendak, M. P., Denniston, A. K., Keane, P. A., & Maniatopoulos, G. (2023). Unlocking the potential of qualitative research for the implementation of artificial intelligence-enabled healthcare. In *Journal of Medical Artificial Intelligence* (Vol. 6). https://doi.org/10.21037/jmai-23-28
25. Inaam Akhtar, M., Njie, B., Asimiran, S., Alase, A., Etikan, I., Laughlin, R., Francis, J. J., Johnston, M., Robertson, C., Glidewell, L., Entwistle, V., Eccles, M. P., Grimshaw, J. M., Alhojailan, M. I., Ibrahim, M., Braun, V., Clarke, V., Taherdoost, H., SWELHEIM, T., … Zairul, M. (2019). No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析Title. *ISRA International Journal of Islamic Finance*, *8*(1).
26. Jain, H., & KUMAR, Prof. D. (2025). Artificial Intelligence as a Catalyst for Supply Chain Evolution: A Multi-Case Thematic Study. *INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, *09*(09). https://doi.org/10.55041/ijsrem52737
27. Jiang, Y., Ko-Wong, L., & Valdovinos Gutierrez, I. (2025). The Feasibility and Comparability of Using Artificial Intelligence for Qualitative Data Analysis in Equity-Focused Research. In *Educational Researcher* (Vol. 54, Issue 3). https://doi.org/10.3102/0013189X251314821
28. John, A. O., & Chukwumerije, T. (2014). Perceptions of Accountants on Factors Affecting Auditor's Independence in Nigeria. *International Journal of Research in Commerce, IT & Management*, *2*(11).
29. Kang, R., Xuan, Z., Tong, L., Wang, Y., Jin, S., & Xiao, Q. (2025). Nurse Researchers'

Experiences and Perceptions of Generative AI: Qualitative Semistructured Interview Study. *Journal of Medical Internet Research*, *27*. https://doi.org/10.2196/65523

30. Kozinets, R. v., & Gretzel, U. (2024). Netnography evolved: New contexts, scope, procedures and sensibilities. *Annals of Tourism Research*, *104*. https://doi.org/10.1016/j.annals.2023.103693
31. Lei, P., Duo, E. N., Gao, Y., & Zhu, X. (2025). Research on the experience of utilizing artificial intelligence in nursing learning from the perspective of nursing students – a qualitative study. *BMC Nursing*, *24*(1). https://doi.org/10.1186/s12912-025-03704-7
32. Li, Z. Q., Shroff, P. K., Venkataraman, R., Zhang, I. X., Gu, F., Lev, B., He, G., Kang, J. K. H. K., Shivdasani, A., Yamada, T., Bhattacharyya, S., Nain, A. S., Lee, K. B. H., Homburg, C., Bucerius, M., Erel, I., Liao, R. C., WEISBACH, M. S., Andrade, G., … Lys, T. Z. (2019). CORPORATE TAKEOVERS AND PRODUCTIVITY - LICHTENBERG,FR. *JOURNAL OF BANKING & FINANCE*, *37*(2).
33. Lim, W. M. (2025). What Is Qualitative Research? An Overview and Guidelines. *Australasian Marketing Journal*, *33*(2). https://doi.org/10.1177/14413582241264619
34. Longo, L. (2020). Empowering Qualitative Research Methods in Education with Artificial Intelligence. *Advances in Intelligent Systems and Computing*, *1068*. https://doi.org/10.1007/978-3-030-31787-4_1
35. Luongo, M., Ponticorvo, M., Ligorio, M. B., Crescenzo, P., & Ritella, G. (2024). Artificial Intelligence to Enhance Qualitative Research: Methodological Reflections on a Pilot Study. *Journal of Educational, Cultural and Psychological Studies*, *2024*(30). https://doi.org/10.7358/ecps-2024-030-luon
36. Marshall, D. T., & Naff, D. B. (2024). The Ethics of Using Artificial Intelligence in Qualitative Research. *Journal of Empirical Research on Human Research Ethics*, *19*(3). https://doi.org/10.1177/15562646241262659
37. Nashwan, A. J., & Abukhadijah, H. (2023). Harnessing Artificial Intelligence for Qualitative and Mixed Methods in Nursing Research. *Cureus*. https://doi.org/10.7759/cureus.48570
38. Nicmanis, M., & Spurrier, H. (2025). Getting Started with Artificial Intelligence

Assisted Qualitative Analysis: An Introductory Guide to Qualitative Research Approaches with Exploratory Examples from Reflexive Content Analysis. *International Journal of Qualitative Methods* , *24*. https://doi.org/10.1177/16094069251354863

39. Olawade, D. B., Omeni, D., Gore, M. N., & Hadi, M. (2025). Enhancing qualitative research through virtual focus groups and artificial intelligence: A review. *International Journal of Medical Informatics*, *203*. https://doi.org/10.1016/j.ijmedinf.2025.106004
40. Owoahene Acheampong, I., & Nyaaba, M. (2024). Review of Qualitative Research in the Era of Generative Artificial Intelligence. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4686920
41. Pratt, M. G. (2025). On the Evolution of Qualitative Methods in Organizational Research. In *Annual Review of Organizational Psychology and Organizational Behavior* (Vol. 12). https://doi.org/10.1146/annurev-orgpsych-111722-032953
42. Roberts, J., Baker, M., & Andrew, J. (2024). Artificial intelligence and qualitative research: The promise and perils of large language model (LLM) 'assistance.' *Critical Perspectives on Accounting*, *99*. https://doi.org/10.1016/j.cpa.2024.102722
43. Rodríguez Flores, E. A., & Sánchez Trujillo, M. de los Á. (2025). Scientific research and artificial intelligence in graduate students. A qualitative analysis. *European Public and Social Innovation Review*, *10*. https://doi.org/10.31637/epsir-2025-1049
44. Rosenberg, J. (2024). Automated Transcription of Interviews in Qualitative Research Using Artificial Intelligence A Simple Guide. *Journal of Surgery Research and Practice*. https://doi.org/10.46889/jsrp.2024/5204
45. Saheb, T., Sidaoui, M., & Schmarzo, B. (2024). Convergence of artificial intelligence with social media: A bibliometric & qualitative analysis. In *Telematics and Informatics Reports* (Vol. 14). https://doi.org/10.1016/j.teler.2024.100146
46. Sampaio, R. C., Chagas, V., Sanchez, C. S., Gonçalves, J., Borges, T., Alison, M. B., Tigrinho, C. S., de Souza, J. R., & Paz, F. S. (2024). AN ARTIFICIAL INTELLIGENCE (AI)-ASSISTED SCOPING REVIEW OF EMERGING USES OF AI IN QUALITATIVE

RESEARCH AND ITS ETHICAL CONSIDERATIONS. *Revista Pesquisa Qualitativa*, *12*(30). https://doi.org/10.33361/RPQ.2024.v.12.n.30.729

47. Schmidt, D. A., Alboloushi, B., Thomas, A., & Magalhaes, R. (2025). Integrating artificial intelligence in higher education: perceptions, challenges, and strategies for academic innovation. *Computers and Education Open*, *9*. https://doi.org/10.1016/j.caeo.2025.100274
48. Sposato, M. (2025). Introducing AI to qualitative research: Theoretical and practical implications. In *Exploring Qualitative Research in Business: Approaches, Strategies, And Methods*. https://doi.org/10.4018/979-8-3693-7698-0.ch010
49. Tierney, A., Peasey, P., & Gould, J. (2025). Student perceptions on the impact of AI on their teaching and learning experiences in higher education. *Research and Practice in Technology Enhanced Learning*, *20*. https://doi.org/10.58459/rptel.2025.20005
50. Tong, L., Niu, Y., Zhou, L., Jin, S., Wang, Y., & Xiao, Q. (2025). Perceptions and experiences of generative artificial intelligence training to support research for Chinese nurses: A qualitative focus group study. *International Journal of Nursing Sciences*, *12*(3). https://doi.org/10.1016/j.ijnss.2025.04.010
51. Yang, Y., & Ma, L. (2025). Artificial intelligence in qualitative analysis: a practical guide and reflections based on results from using GPT to analyze interview data in a substance use program. *Quality and Quantity*, *59*(3). https://doi.org/10.1007/s11135-025-02066-1
52. Yang, Y., & Siau, K. (2018). A Qualitative Research on Marketing and Sales in the Artificial Intelligence Age. *Midwest Association for Information Systems ConferenceAssociation for Information Systems Conference*, *5*(17-18 May 2018).
53. Yıldırım, H. M., Güvenç, A., & Can, B. (2025). Reporting Quality in Qualitative Research Using Artificial Intelligence and the COREQ Checklist: A Systematic Review of Tourism Literature. *International Journal of Qualitative Methods* , *24*. https://doi.org/10.1177/16094069251351487
54. Choudrie, J., Manandhar, N., Castro, C., & Obuekwe, C. (2023). Hey Siri, Google! Can you help me? A qualitative case study of smartphones AI functions in SMEs.

Technological Forecasting and Social Change, *189*. https://doi.org/10.1016/j.techfore.2023.122375

55. Dai, Y., Lai, S., Lim, C. P., & Liu, A. (2023). ChatGPT and its impact on research supervision: Insights from Australian postgraduate research students. *Australasian Journal of Educational Technology*, *39*(4). https://doi.org/10.14742/ajet.8843
56. Djakfar Musthafa, F. A. (2024). Penggunaan Artificial Intelligence (AI) dalam Pembelajaran: Fenomena Transformasi Otoritas Pengetahuan di Kalangan Mahasiswa. *Journal of Contemporary Islamic Education*, *4*(1). https://doi.org/10.25217/jcie.v4i1.4386
57. Hogg, H. D. J., Al-Zubaidy, M., Talks, J., Denniston, A. K., Kelly, C. J., Malawana, J., Papoutsi, C., Teare, M. D., Keane, P. A., Beyer, F. R., & Maniatopoulos, G. (2023). Stakeholder Perspectives of Clinical Artificial Intelligence Implementation: Systematic Review of Qualitative Evidence. In *Journal of Medical Internet Research* (Vol. 25). https://doi.org/10.2196/39742
58. Kantor, J. (2024). Best practices for implementing ChatGPT, large language models, and artificial intelligence in qualitative and survey-based research. In *JAAD International* (Vol. 14). https://doi.org/10.1016/j.jdin.2023.10.001
59. Naeem, M., Smith, T., & Thomas, L. (2025). Thematic Analysis and Artificial Intelligence: A Step-by-Step Process for Using ChatGPT in Thematic Analysis. *International Journal of Qualitative Methods* , *24*. https://doi.org/10.1177/16094069251333886
60. Orea-Giner, A., Fusté-Forné, F., & Soliman, M. (2025). How do tourists perceive green customer-love service in restaurants? A qualitative exploration of ai and human collaboration. *International Journal of Hospitality Management*, *131*. https://doi.org/10.1016/j.ijhm.2025.104300
61. Tschisgale, P., Wulff, P., & Kubsch, M. (2023). Integrating artificial intelligence-based methods into qualitative research in physics education research: A case for computational grounded theory. *Physical Review Physics Education Research*, *19*(2). https://doi.org/10.1103/PhysRevPhysEducRes.19.020123

62. Venter, I. M., Blignaut, R. J., Cranfield, D. J., Tick, A., & Achi, S. el. (2025). AI versus tradition: shaping the future of higher education. *Journal of Applied Research in Higher Education*, *17*(7). https://doi.org/10.1108/JARHE-12-2024-0702

63. Wang, S., & Kim, R. Y. (2025). Exploring the role of generative AI in enhancing critical thinking in multicultural classrooms: a case study from korea. *Multicultural Education Review*, *17*(1). https://doi.org/10.1080/2005615X.2025.2467787

Chapter IV.

Quantitative Research Methodology in Artificial Intelligence and Data Science: A Comprehensive Framework for Empirical Analysis

1. Epistemological Foundations of Quantitative Analysis in Computational Intelligence

The evolution of Artificial Intelligence (AI) and Data Science has fundamentally shifted the discipline of computer science from a deterministic, logic-based field into an empirical science heavily reliant on quantitative research methodologies. In the early eras of symbolic AI, systems were constructed upon rigid, rule-based logic where validation was a binary state: a theorem was either proven or it was not; a logic gate returned true or false. However, the modern paradigm—dominated by stochastic machine learning (ML), deep learning (DL), and probabilistic graphical models—requires a fundamental epistemological shift. We no longer validate systems based on logical absolutes but rather evaluate them within a probabilistic state space.[1]

In this contemporary context, quantitative research is defined as the systematic investigation of phenomena through the collection of quantifiable data and the application of statistical, mathematical, and computational techniques.[2] The objective is not merely to build a system that functions but to construct a rigorous evidentiary basis that quantifies *how well* it functions, under *what conditions* it fails, and *how confident* we can be in its predictions. This transition necessitates a rigorous framework for empirical evaluation that mirrors the experimental rigor found in physics or medicine. A model achieving 95% accuracy is not "correct" in the absolute sense; it is merely statistically likely to be correct under a specific, observed distribution of data.[3] Therefore, the methodologies used to evaluate these systems must be rooted in statistical theory, experimental design, and quantitative metrics that can capture the nuance of probabilistic performance.[4]

The scope of quantitative methodology in AI extends far beyond simple performance metrics like accuracy or precision. It encompasses the rigorous design of experiments to control for confounding variables, the statistical comparison of algorithms to ensure reproducibility, the quantification of bias and fairness to align with ethical standards, and the measurement of operational efficiency metrics such as latency and carbon footprint.[5] As AI systems become integral to critical decision-making processes in healthcare, finance, and autonomous systems, the demand for "scientific rigor"—defined by reproducibility, statistical significance, and transparent methodology—has become paramount.[5] The "reproducibility crisis" currently facing the field serves as a stark reminder that without robust quantitative standards, the rapid pace of innovation risks producing fragile, unreliable, or biased systems.

1.1 The Shift from Qualitative to Quantitative Paradigms

Historically, computer science education and research utilized a mix of methodologies, often leaning heavily on theoretical proofs or qualitative demonstrations of capability. However, recent bibliometric analyses indicate a dominant, accelerating trend toward quantitative methods in AI research.[1] This shift is driven by the necessity of empirical validation for high-dimensional, non-linear models where theoretical bounds—such as the Vapnik–Chervonenkis (VC) dimension—are often too loose to be practically useful for predicting performance on real-world data. Instead, researchers rely on large-scale empirical testing on benchmark datasets to quantify progress.

This quantitative dominance brings its own set of challenges. The reliance on empirical metrics has led to a "leaderboard culture," where marginal improvements in quantitative scores are prioritized over methodological soundness or interpretability.

Furthermore, the "reproducibility crisis" in AI suggests that while quantitative metrics are widely used, the *methodologies* surrounding their reporting often lack the necessary rigor.[5] Issues such as data leakage, improper splitting strategies, and the lack of statistical significance testing contribute to inflated claims of state-of-the-art (SOTA) performance. Consequently, a nuanced understanding of quantitative methodology is not merely an academic exercise but a critical requirement for valid engineering practice in the age of AI. The following report details a comprehensive framework for this quantitative practice, spanning experimental design, metric selection, statistical validation, fairness auditing, and operational assessment.

2. Experimental Design and Sampling Strategies

The foundation of any rigorous quantitative study is the experimental design. In the context of AI and Data Science, this primarily concerns how data is collected, partitioned, and utilized to train and evaluate models. The objective of experimental design is to estimate the generalization error—the expected performance of the model on unseen data drawn from the same underlying distribution—as accurately and unbiasedly as possible.[8] Without a robust design, even the most sophisticated model architectures and metrics yield meaningless results.

2.1 The Bias-Variance Decomposition in Data Partitioning

The central challenge in estimating model performance is the bias-variance trade-off inherent in resampling methods. If we use all available data for training, we maximize the information available to the model (reducing bias), but we have no independent data left for evaluation (preventing variance estimation). Conversely, if we reserve a large portion of data for testing, the model trained on the smaller remainder may not represent the true potential of the algorithm (increasing bias), and the small test set may yield unstable performance estimates (increasing variance).

2.1.1 The Holdout Method and Its Limitations

The simplest quantitative method is the holdout method, where the dataset D is partitioned into two disjoint sets: a training set D_{train} and a test set D_{test}. A common split ratio is 70:30 or 80:20. While computationally efficient, the holdout method is highly sensitive to the specific partition of data. A "lucky" split might place all easy-to-classify instances in the test set, resulting in an optimistic bias, while an "unlucky" split does the reverse.[8] This variance is particularly problematic in small to medium-sized datasets, making the holdout method less robust for rigorous scientific comparison. In modern "Deep Research," reliance on a single holdout split is often considered insufficient evidence of superiority unless the dataset size is massive (e.g., millions of examples) such that the law of large numbers stabilizes the error estimates.

2.1.2 K-Fold Cross-Validation: The Gold Standard

To mitigate the high variance of the holdout method, K-Fold Cross-Validation (CV) has emerged as the standard practice in quantitative evaluation. The dataset is randomly partitioned into k equal-sized subsamples (folds). The process is an iterative rotation: of the k subsamples, a single subsample is retained as the validation data for testing the model, and the remaining k-1 subsamples are used as training data.[8] This process is repeated k times, with each of the k subsamples used exactly once as the validation data.

The k results are then averaged to produce a single estimation. The advantage of this method is that all observations are used for both training and validation, and each observation is used for validation exactly once. The formula for the cross-validation estimate is:

$$CV_{(k)} = \frac{1}{k} \sum_{i=1}^{k} L_i$$

where L_i is the loss or error metric on the i-th fold.

Stratified K-Fold CV: In classification tasks, particularly with imbalanced datasets (e.g., fraud detection where positive cases are <1%), random splitting can result in folds that have no positive examples, rendering training or testing impossible. Stratification addresses this by ensuring that the proportion of samples for each class in every fold matches the proportion in the complete dataset. This reduces the variance of the performance estimate and ensures that the model is evaluated on a representative distribution of the minority class.[9]

2.1.3 Repeated Cross-Validation and the 5x2cv Protocol

While 10-fold CV is standard, it still possesses variance due to the randomness of the initial partitioning. Repeated K-Fold CV involves repeating the entire K-Fold process multiple times with different random seeds for the splits. This provides a better Monte-Carlo estimate of the complete cross-validation performance.[9]

A specific and statistically critical variant is 5x2 Cross-Validation, proposed by Dietterich (1998) specifically for the statistical comparison of machine learning algorithms. It involves performing 2-fold cross-validation five times. This protocol was developed to address the high Type I error rates (false positives) often observed when comparing algorithms using standard 10-fold CV. In standard 10-fold CV, the training sets for each fold overlap by 90%, violating the independence assumption required by most statistical tests. The 5x2cv method creates training sets that are more disjoint within each replication, providing a more robust basis for hypothesis testing.[10]

2.2 Resampling Techniques: The Bootstrap

Bootstrapping is a powerful statistical method for estimating the sampling distribution of an estimator by resampling with replacement from the original data.[12] In AI evaluation, bootstrapping serves a distinct purpose from cross-validation and is particularly valuable for quantifying uncertainty.

In the bootstrap method, a dataset of size n is resampled with replacement n times to create a "bootstrap sample." Probability theory dictates that the probability of any specific instance *not* being chosen in a sample of size n is $(1 - 1/n)^n$. As $n \to \infty$, this converges to $1/e \approx 0.368$. Thus, on average, a bootstrap sample contains approximately 63.2% of the original unique instances, leaving about 36.8% of the data as "out-of-bag" (OOB) samples.[13]

The.632+ Bootstrap Method:

The OOB samples can serve as a test set. However, since the training set (the bootstrap sample) only contains $\approx 63\%$ of unique data, models might underperform compared to those trained on the full n samples (pessimistic bias). The.632 bootstrap estimator attempts to correct this bias by taking a weighted average of the training error (which is usually optimistically low, often zero for overfitted models) and the OOB error:

$$Err_{.632} = 0.368 \cdot Err_{train} + 0.632 \cdot Err_{OOB}$$

While useful for small datasets, bootstrapping is computationally expensive. Its modern application is increasingly focused not just on point estimates of accuracy, but for generating Confidence Intervals (CIs) around performance metrics. By calculating the metric on thousands of bootstrap replicates, researchers can report a 95% confidence interval (e.g., "Accuracy: 85% $\pm$ 2%"), which is a critical aspect of rigorous quantitative reporting.[14]

2.3 Dealing with Data Leakage and Temporal Dependencies

A rigorous experimental design must explicitly prevent data leakage, a phenomenon where information from the test set improperly influences the training process, leading to overly optimistic performance estimates.[16] This is one of the most common causes of reproducibility failure in AI (see Table 8).

Table 8: Common causes of reproducibility failure in AI

Type of Leakage	**Description**	**Prevention Strategy**

Train-Test Contamination	The same sample appears in both training and test sets (e.g., due to duplication).	Deduplication before splitting; strict set isolation.
Temporal Leakage	In time-series, future data is used to predict past events.	**Time-Series Split:** Train on $t_{0} \dots t_{k}$, test on t_{k+1}. No random shuffling.
Group Leakage	Samples from the same subject (e.g., patient, user) are split across train/test.	**Group K-Fold:** Ensure all samples from one subject are in the same fold.
Feature Leakage	Features (e.g., "Future Sales") that constitute the target are included in input.	Rigorous feature engineering review; removing proxies for the target.

For time-series data or spatial data, random splitting (like standard K-Fold) constitutes severe data leakage. If a model is predicting stock prices, training on data from Tuesday and Thursday to predict Wednesday is invalid because it violates the causality of the information flow. In such cases, Rolling Window Validation or Walk-Forward Validation must be used, where the training set consists only of data historically preceding the validation set.[16]

2.4 Reproducibility and Rigor in Experimental Design

The "reproducibility crisis" in AI has catalyzed a movement toward stricter quantitative standards. Evaluating an AI model is not merely about reporting a final number but documenting the entire experimental process. The NeurIPS Reproducibility Checklist has become a de facto standard for quantitative rigor.[17] Adherence to this checklist ensures that the experimental design is transparent and replicable.

Key elements of reproducible quantitative design include:

- **Data Splitting Transparency:** Explicitly stating how train/validation/test splits were generated, including the exact random seed used. This allows other researchers to reconstruct the exact data subsets.[17]

- **Hyperparameter Search Space:** Reporting the range of hyperparameters considered and the method of selection (e.g., grid search, random search, Bayesian optimization). Reporting only the best result from a massive search without adjusting for the "multiple comparisons problem" leads to overfitting the validation set, a practice sometimes called "p-hacking" in statistics or "gradient descent on the test set" in ML.[5]
- **Infrastructure Specifications:** Documenting the hardware (GPU/TPU types) and software environment (library versions), as floating-point arithmetic can vary slightly across platforms, impacting the exact reproducibility of training trajectories.[17]

3. Performance Evaluation Metrics: Quantifying Success

In quantitative research, the choice of metric defines the optimization landscape. A mismatch between the quantitative metric and the actual business or scientific objective can lead to "successful" models that fail in practice. This phenomenon, often summarized by Goodhart's Law ("When a measure becomes a target, it ceases to be a good measure"), necessitates a careful selection of diverse metrics to capture the full behavior of the system.

3.1 Metrics for Classification: Beyond Accuracy

For categorical prediction tasks, Accuracy is the most intuitive metric but is often misleading, particularly in imbalanced datasets—a situation known as the "accuracy paradox." In a fraud detection dataset where only 0.1% of transactions are fraudulent, a naive classifier that predicts "valid" for every single transaction achieves 99.9% accuracy but has zero utility. Thus, rigorous quantitative analysis requires a decomposition of errors.[3]

3.1.1 The Confusion Matrix and Derived Metrics

The foundational tool for this decomposition is the Confusion Matrix, an N \times

N table that cross-tabulates predicted classes against actual classes. For a binary problem, this breaks down predictions into True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).[20] From these primitives, we derive metrics that isolate specific types of performance:

- **Precision (Positive Predictive Value):** $\frac{TP}{TP+FP}$. This measures the trustworthiness of a positive prediction. In a spam filter, high precision is critical to avoid flagging legitimate emails as spam.
- **Recall (Sensitivity/True Positive Rate):** $\frac{TP}{TP+FN}$. This measures the ability of the model to capture all positive instances. In medical diagnosis for a lethal but curable disease, high recall is paramount, even at the cost of precision.[3]
- F1 Score: The harmonic mean of Precision and Recall. The harmonic mean is used because it punishes extreme values more than the arithmetic mean; if either Precision or Recall is low, the F1 score will be low.

  $$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

3.1.2 Probabilistic Threshold-Invariant Metrics

Classifiers typically output a probability score (e.g., 0.85 chance of churn) rather than a hard label. Converting this to a label requires a threshold (e.g., >0.5). Metrics like Accuracy and F1 are sensitive to this threshold choice. To evaluate the discriminative power of the model independent of the threshold, we use the ROC (Receiver Operating Characteristic) curve and the AUC (Area Under the Curve).

- **ROC Curve:** Plots the True Positive Rate (Recall) against the False Positive Rate (1 - Specificity) at every possible classification threshold.
- **AUC-ROC:** Represents the probability that a randomly chosen positive instance is ranked higher (has a higher predicted probability) than a randomly chosen negative instance. An AUC of 0.5 implies random guessing, while 1.0 implies perfect separability. The AUC is robust to class imbalance and provides a holistic view of the

classifier's ability to distinguish between classes.[20]

3.2 Metrics for Regression: Quantifying Error Magnitude

For tasks predicting continuous output variables (e.g., price, temperature), metrics must quantify the magnitude of the distance between predicted values ($\hat{y}$) and actual values (y).[23]

- **RMSE (Root Mean Squared Error):** $\sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$. By squaring the errors before averaging, RMSE penalizes large errors disproportionately. This is desirable when large errors are particularly costly (e.g., predicting the trajectory of a self-driving car).
- **MAE (Mean Absolute Error):** $\frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|$. MAE provides a linear score, meaning all errors are weighted equally. It is often more interpretable as the "average mistake."
- **R^2 (Coefficient of Determination):** Represents the proportion of the variance in the dependent variable that is predictable from the independent variables. While useful for explanation, R^2 can be misleading in non-linear models and does not indicate whether the predictions are biased.

3.3 Advanced Metrics for Natural Language Processing (NLP)

The quantitative evaluation of Generative AI and NLP is significantly more complex because "correctness" is subjective. Unlike a classification label, there is rarely a single correct sequence of words for a translation or summary. Evaluation thus relies on measuring similarity to human-generated reference texts.

3.3.1 N-gram Based Metrics: BLEU and ROUGE

BLEU (Bilingual Evaluation Understudy): Originally designed for machine translation, BLEU measures the precision of n-grams (sequences of n words) in the candidate text compared to the reference text(s). It is a Precision-oriented metric.[25] The BLEU score formulation includes a Brevity Penalty (BP) to prevent the model from gaming the metric by outputting very short, high-precision sentences (e.g., outputting just "The" when the reference is "The cat sat on the mat" would yield 100% unigram precision without the penalty).

BLEU = BP \cdot \exp\left(\sum_{n=1}^{N} w_n \log p_n\right)
where p_n is the modified n-gram precision and w_n are weights (typically uniform). The use of the geometric mean ensures that if precision for any n-gram order is zero, the entire score is zero (unless smoothed), enforcing quality across different granularities (1-gram for adequacy, 4-gram for fluency).[26]

ROUGE (Recall-Oriented Understudy for Gisting Evaluation): Commonly used for text summarization, ROUGE focuses on Recall—how much of the reference summary was captured by the generated text?.[28]

- **ROUGE-N:** Measures the overlap of N-grams.
- **ROUGE-L:** Based on the Longest Common Subsequence (LCS). ROUGE-L captures sentence structure by identifying the longest co-occurring sequence of words in sequence, even if they are not consecutive (allowing for interruptions). This makes it more flexible than strict n-gram matching.[28]

METEOR (Metric for Evaluation of Translation with Explicit ORdering): METEOR addresses weaknesses in BLEU by calculating similarity based on unigram matching but extends this to include stemmed words (e.g., "running" matches "run") and synonyms (using WordNet). It calculates the harmonic mean of precision and recall (weighted towards recall) and includes a penalty for poor ordering, providing a better correlation with human judgment than BLEU in many contexts.[29]

3.3.2 Embedding-Based Metrics: The Semantic Shift

Traditional n-gram metrics (BLEU, ROUGE) fail to capture semantic similarity. If a model generates "automobile" and the reference is "car," n-gram metrics score 0, despite the perfect semantic match. BERTScore represents a paradigm shift toward semantic evaluation. It uses pre-trained contextual embeddings (like BERT) to represent tokens and calculates the cosine similarity between the embeddings of the candidate and reference tokens.[30]

- **Mechanism:** BERTScore computes a pairwise cosine similarity matrix between all tokens in the candidate and reference. It then performs a "greedy matching" to find the most similar reference token for each candidate token.
- **Advantage:** It captures paraphrasing and semantic equivalence that surface-level metrics miss.
- **Robustness:** BERTScore has been shown to correlate better with human judgments of quality, particularly for complex generation tasks where diverse vocabulary is used.[30]

4. Statistical Significance and Hypothesis Testing

In rigorous quantitative AI research, reporting a higher average metric is insufficient to claim superiority. Due to the stochastic nature of initialization, data shuffling, and non-convex optimization landscapes, performance differences may be due to random chance. Statistical hypothesis testing provides the framework to distinguish signal from noise.

4.1 The Problem with Simple t-tests in Cross-Validation

A common methodological error in AI research is using a standard Student's t-test on the k scores resulting from a single run of k-fold cross-validation. This approach violates the fundamental assumption of the t-test: the independence of samples. In 10-fold CV, the training sets for any two folds share approximately 90% of the same data. Consequently, the performance estimates are highly correlated. This correlation leads to a gross underestimation of the variance and an inflated Type I error rate (detecting a significant difference where none exists).[10]

4.2 Recommended Statistical Tests for Algorithm Comparison

To address the violation of independence, specialized statistical tests have been developed for the specific structure of machine learning experiments.

4.2.1 McNemar's Test

For comparing two classifiers on a single test set, McNemar's Test is the recommended non-parametric standard. It operates not on the accuracy scores themselves, but on the contingency table of the two algorithms' predictions on individual test instances:

- n_{00}: Number of examples where both Model A and Model B were wrong.
- n_{11}: Number of examples where both were correct.
- n_{01}: Number of examples where Model A was correct, but Model B was wrong.
- n_{10}: Number of examples where Model A was wrong, but Model B was correct.

The test focuses entirely on the discordant pairs (n_{01} and n_{10}). Under the null hypothesis that the classifiers have equal error rates, n_{01} and n_{10} should be roughly equal. The test statistic is calculated as:

$$\chi^2 = \frac{(|n_{01} - n_{10}| - 1)^2}{n_{01} + n_{10}}$$

This statistic follows a Chi-Squared distribution with 1 degree of freedom. If the p-value is below the significance threshold (usually 0.05), we reject the null hypothesis and conclude that the models have different performance profiles.[10]

4.2.2 The 5x2cv Paired t-Test

For comparing algorithms using cross-validation (when a single large test set is unavailable), Dietterich (1998) analyzed various testing schemes and recommended the **5x2cv paired t-test**. This protocol involves performing five replications of 2-fold cross-validation.

- **Why 2-fold?** In 2-fold CV, the training sets for the two folds are completely disjoint (they are the inverse of each other). This maximizes the independence between the two estimates in a single run.

- Why 5 repetitions? To gather enough samples for a t-test without re-introducing excessive correlation. The resulting test statistic provides a much better balance between Type I errors (false positives) and Type II errors (low power) compared to the naive k-fold t-test.10

4.3 Bootstrap Confidence Intervals and Effect Sizes

Beyond binary hypothesis testing (significant vs. not significant), modern quantitative practice emphasizes Effect Sizes and Confidence Intervals (CIs). A result can be statistically significant but practically meaningless (e.g., an accuracy improvement of 0.001% with a massive sample size).

Effect Size (Cohen's d): This metric quantifies the magnitude of the difference between two groups in terms of standard deviations.

d = \frac{\bar{x}_1 - \bar{x}_2}{s_{pooled}}

A Cohen's d of 0.2 is considered a small effect, 0.5 a medium effect, and 0.8 a large effect. Reporting effect size helps practitioners understand whether an "improvement" justifies the computational or financial cost of deploying the new model.[33]

Bootstrap Confidence Intervals: Bootstrapping the test set predictions allows for the construction of non-parametric CIs around metrics like Accuracy, F1, or AUC.

1. Resample the test set predictions B times (e.g., 1000).
2. Recalculate the metric for each resample.
3. The distribution of these B metrics approximates the sampling distribution.
4. The 2.5th and 97.5th percentiles define the 95% CI.
   If the 95% CI of the difference between two models (Metric_A - Metric_B) does not contain zero, the result is statistically significant. This method is robust to non-normal

distributions of errors.15

5. Algorithmic Fairness: Quantitative Ethics

As AI systems increasingly impact human lives (hiring, lending, criminal justice), "Fairness" has moved from an abstract ethical concept to a quantifiable engineering constraint. Quantitative research in fairness involves defining mathematical criteria for unbiased decision-making and measuring deviations from these criteria.

5.1 The Taxonomy of Fairness Metrics

There are three primary families of fairness metrics used in quantitative auditing. Crucially, these metrics are often mathematically incompatible with one another, forcing researchers to make explicit trade-offs.

5.1.1 Disparate Impact (Demographic Parity)

This metric represents the legal concept of "disparate impact." It requires that the probability of a positive outcome (e.g., getting a loan) be equal across groups (e.g., gender, race), regardless of the ground truth or other variables.

$P(\hat{Y}=1 \mid A=a) = P(\hat{Y}=1 \mid A=b)$The **Disparate Impact Ratio** is often calculated as:$\frac{P(\hat{Y}=1 \mid A=minority)}{P(\hat{Y}=1 \mid A=majority)}$

A common heuristic is the "four-fifths rule" (80%), which suggests this ratio should be at least 0.80 to avoid prima facie evidence of discrimination.[35]

5.1.2 Equalized Odds (Separation)

Equalized Odds requires that the model performs equally well for both groups, specifically mandating equality of True Positive Rates (TPR) and False Positive Rates (FPR) across groups.

$$P(\hat{Y}=1 \mid Y=y, A=a) = P(\hat{Y}=1 \mid Y=y, A=b), \quad y \in \{0, 1\}$$

This ensures that qualified individuals in both groups have the same chance of selection, and unqualified individuals have the same chance of rejection. This metric allows for different selection rates if the base rates of qualification differ between groups.36

5.1.3 Predictive Parity (Sufficiency)

This requires that the Positive Predictive Value (PPV)—the probability that a positive prediction is actually positive—be equal across groups.

$$P(Y=1 \mid \hat{Y}=1, A=a) = P(Y=1 \mid \hat{Y}=1, A=b)$$

This is often used in risk assessment tools (like recidivism prediction) to ensure that a "High Risk" score means the same probability of re-offense regardless of the demographic of the defendant.37

5.2 The Impossibility Theorems

A critical insight from quantitative fairness research is the proof that, in the presence of unequal base rates (different prevalence of the target variable Y between groups), it is mathematically impossible to satisfy disparate impact, equalized odds, and predictive parity simultaneously.

For example, if Group A has a higher rate of loan repayment than Group B (unequal base rates), and we enforce Predictive Parity (equal PPV), mathematics dictates that we *must* have unequal False Positive Rates or False Negative Rates, thereby violating Equalized Odds.[37] This "Impossibility Theorem" (proven by Chouldechova, Kleinberg, and others) forces researchers to choose which definition of fairness is appropriate for their specific context. It shifts the burden from "optimizing for fairness" to "choosing the fairness metric that aligns with the specific ethical and legal goals of the application."

6. Operational Efficiency and Green AI Metrics

The quantitative evaluation of AI is incomplete without considering the computational cost. As models scale exponentially (e.g., Large Language Models), operational metrics become as critical as accuracy. The "Red AI" trend (buying performance with massive compute) is giving way to "Green AI," which prioritizes efficiency.

6.1 Computational Complexity and FLOPS

Counting parameters is a poor proxy for computational cost. A sparse model with many parameters might be faster than a dense model with fewer. The standard unit for training cost is FLOPS (Floating Point Operations).

For Transformer-based Large Language Models (LLMs), the training cost is often approximated using the Kaplan/OpenAI scaling laws:

C \approx 6 N D

where C is total FLOPs, N is the number of parameters, and D is the dataset size (in tokens). This factor of 6 arises from the mechanics of the backpropagation algorithm: the forward pass requires approximately 2 FLOPs per parameter (multiply-add), and the backward pass requires approximately 4 FLOPs per parameter.[38]

Model FLOPs Utilization (MFU): This metric describes how efficiently the hardware is being used. It is calculated as the ratio of observed FLOPs per second to the theoretical peak FLOPs of the hardware. High MFU indicates that the training run is not bottlenecked by memory bandwidth or communication latency.[39]

6.2 Latency and Throughput

In production environments, the average latency is often less important than the tail distribution.

- **P99 and P95 Latency:** The time below which 99% or 95% of requests fall. This captures the "tail latency" which affects the worst-case user experience. A system might have a fast average response but occasional massive hangs; P99 exposes this.
- **Throughput:** The number of inferences processed per second.
- **Cold vs. Warm Start:** Quantitative analysis must distinguish between the latency of a system spinning up from zero (cold start) versus one already resident in memory (warm start). Cold starts are a critical metric for serverless AI deployments.[40]

6.3 Carbon Footprint and Energy Consumption

"Green AI" introduces metrics to quantify environmental impact. Tools like **CodeCarbon** and **CarbonTracker** have been developed to measure the energy consumption (E) of hardware (GPU/CPU/RAM) in real-time and multiply it by the **Carbon Intensity** (CI) of the local power grid.[42]

Carbon (gCO_2) = E_{kWh} \times CI_{gCO_2/kWh}

This formula highlights a critical operational insight: the carbon footprint depends heavily on *where* the model is trained. Training a Transformer in a region powered by coal (high carbon intensity) emits vastly more CO_2 than training the exact same model in a region powered by hydroelectricity or wind.[43] Consequently, reporting the "training carbon cost" has become a standard requirement in high-quality research papers, promoting the migration of heavy workloads to greener data centers.[6] Furthermore, researchers distinguish between Training Carbon (a one-time fixed cost) and Inference Carbon (a variable cost that scales with usage). For widely used models like ChatGPT, inference carbon quickly eclipses training carbon, making model compression and distillation critical quantitative goals.[45]

7. Production-Grade Experimentation: A/B Testing and Beyond

While offline metrics (Accuracy, F1, AUC) are useful for model development, they are merely proxies for business value. A model with higher AUC might fail to drive user engagement due to latency or unexpected behavioral dynamics. The final phase of quantitative validation occurs in production through online experimentation.

7.1 A/B Testing for ML

A/B testing (Split testing) is the gold standard for causal inference in production. It compares a Control (current model) against a Treatment (new model) on live traffic.
Unlike standard UI A/B testing, ML A/B tests face specific challenges:

- **Effect Stability:** ML models may degrade over time (concept drift) or adapt to user behavior. Short tests may miss long-term degradation.
- **OEC (Overall Evaluation Criterion):** The success metric must be a business KPI (e.g., conversion rate, watch time), not a model metric (e.g., RMSE). There is often a disconnect between the offline loss function (minimizing error) and the online OEC (maximizing revenue).[46]
- **Power Analysis:** Before starting the test, researchers must perform a power analysis to determine the minimum sample size needed to detect a statistically significant difference (Minimum Detectable Effect). This prevents "peeking" at results and stopping early, which inflates false positives.[46]

7.2 Multivariate Testing and Bandits

- **Multivariate Testing (MVT):** Allows testing multiple variables simultaneously (e.g., Model Architecture + UI Layout + Copy). While it allows for the detection of interaction effects (e.g., the model works better with Layout A than Layout B), it requires significantly larger sample sizes than A/B testing.[47]
- **Multi-Armed Bandits (MAB):** A dynamic form of A/B testing where traffic is strictly not split 50/50 for the duration. Instead, algorithms (like Thompson Sampling or Upper Confidence Bound) progressively route more traffic to the winning variation *during* the test. This balances **Exploration** (gathering data to find the winner) and **Exploitation**

(serving the best model to maximize reward). MABs are preferred when the cost of routing users to an inferior model is high (high "regret"), such as in news recommendation or ad targeting.[46]

8. Conclusion: The Holistic Quantitative Framework

Quantitative research methodology in AI is a multi-dimensional discipline that has evolved far beyond simple accuracy reporting. It is a hybrid discipline integrating statistical theory (for hypothesis testing and error estimation), software engineering (for reproducibility and testing), ethics (for fairness quantification), and environmental science (for carbon accounting).

A rigorous quantitative report in modern AI must encompass a holistic view:

1. **Robust Experimental Design:** Utilizing stratified cross-validation, preventing data leakage, and documenting hyperparameter search spaces.
2. **Comprehensive Metrics:** Reporting precision, recall, AUC, and domain-specific metrics (BLEU/BERTScore) rather than a single summary number.
3. **Statistical Validity:** Applying appropriate hypothesis tests (McNemar's, 5x2cv t-test) and reporting confidence intervals and effect sizes to distinguish signal from noise.
4. **Operational Viability:** Quantifying latency distributions (P99), FLOPs, and carbon footprint to ensure sustainability.
5. **Fairness Auditing:** Measuring disparate impact and equalized odds to ensure ethical deployment, while acknowledging the mathematical trade-offs involved.
6. **Production Validation:** Moving beyond offline metrics to rigorous online A/B testing and bandit strategies.

As the field matures, the "State of the Art" is defined not just by the novelty of the neural architecture, but by the rigor of the quantitative evidence supporting it. This empirical framework transforms AI from a speculative art form into a measurable, predictable, and reliable engineering science. The transition from "It works on my machine" to "It works statistically significantly on the population" is the hallmark of the modern quantitative AI researcher (see Table 9).

Table 9: Summary of Quantitative Metrics and Methods by Domain

Domain	Method/Metric	Application	Key Characteristic
Exp. Design	**Stratified K-Fold CV**	General ML	Preserves class distribution; reduces variance.
	5x2cv	Algo Comparison	Low Type I error; ideal for statistical testing.
	Bootstrapping	Uncertainty	Generates Confidence Intervals; robust to non-normality.
Performance	**F1 Score**	Classification	Harmonic mean of Precision/Recall; good for imbalance.
	AUC-ROC	Classification	Threshold-independent; holistic view of discrimination.
	RMSE	Regression	Penalizes large errors heavily; sensitive to outliers.
NLP	**BLEU**	Translation	Precision-based; n-gram overlap; brevity penalty.
	ROUGE-L	Summarization	Recall-based; Longest Common

			Subsequence.
	BERTScore	GenAI / Semantic	Embedding cosine similarity; captures meaning.
Statistics	**McNemar's Test**	Binary Classifiers	Tests disagreement between models; non-parametric.
	Cohen's d	Effect Size	Standardized difference; measures magnitude of effect.
Fairness	**Disparate Impact**	Bias Auditing	Ratio of positive outcomes between groups.
	Equalized Odds	Bias Auditing	Equality of TPR and FPR across groups.
Efficiency	**FLOPS**	Compute Cost	Floating Point Operations; scaling law estimation.
	P99 Latency	Production	99th percentile response time; measures tail lag.
	Carbon Intensity	Green AI	CO_2 emitted per kWh; depends on energy mix.

Chapter bibliographies

1. Abadio, A. K. R., Kioshima, E. S., Martins, N. F., Maigret, B., Felipe, M. S. S., Freitas, S. M., Maigret, B., Felipe, M. S. S., Abagyan, R. A., Kufareva, I., Totrov, M., Abbasi, M., Sadeghi-Aliabadi, H., Hassanzadeh, F., Amanlou, M., Abdelmonsef, A. H., Dulapalli, R., Dasari, T., Padmarao, L. S., … Olson, A. J. A. J. (2015). A pharmacophore-based virtual screening approach for the discovery of Trypanosoma cruzi GAPDH inhibitors (vol 5, pg 2019, 2013). *J Chem Inf Model*. https://doi.org/Book_Doi 10.1002/9783527633326
2. Ackoff, R. L., Aguilera-Caracuel, J., Hurtado-Torres, N. E., Aragón-Correa, J. A., Al-Mawali, H., Al-Shbiel, S. O., Alavi, M., Kayworth, T. R., Leidner, D. E., Albino, V., Garavelli, A. C., Gorgoglione, M., Alfaadhel, S., Allame, S. M., Nouri, B. A., Tavakoli, S. Y., Shokrani, S. A. R., Allame, S. M., Nouri, B. A., … van Tartwijk, J. (2010). Business research methods. *Journal of Knowledge Management*, *14*(1).
3. Al, P., Tentang, G., Arab, P. B., Relevansinya Di Era, D., Lailatul, G., & Maftuhah, H. (2025). Perspekif Al Ghazali Tentang Pendidikan Bahasa Arab Dan Relevansinya di Era Globalisasi. *Taqdir*, *11*(1).
4. Bris, A., Wang, T. Y. H., Zatzick, C. D., Miller, D. J. P., Fern, M. J., Cardinal, L. B., Gregoire, D. A., Shepherd, D. A., Westphal, J. D., Shani, G., Troster, C., van Quaquebeke, N., Lanaj, K., Hollenbeck, J. R., Ilgen, D. R., Barnes, C. M., Harmon, S. J., Feldman, E. R., DesJardine, M. R., … Sangiorgi, F. (2021). KNIGHTS, RAIDERS, AND TARGETS - THE IMPACT OF THE HOSTILE TAKEOVER - COFFEE,JC, LOWENSTEIN,L, ROSEACKERMAN,S. *JOURNAL OF BANKING & FINANCE*, *37*(1).
5. Cantarelli, C. C., Flybjerg, B., Molin, E. J. E., & Wee, B. van. (2010). Cost Overruns in Large-Scale Transport Infrastructure Projects. *European Journal of Transport and Infrastructure Research*, *10*(1).
6. Chodvadiya, J., Shukla, G., & Sharma, G. (2025). 16S-23S-5S rRNA Database: a comprehensive integrated database of archaeal and bacterial rRNA sequences,

alignments, intragenomic heterogeneity, and secondary structures. *Sustainability (Switzerland)*, *11*(1).

7. Committee, S. (1998). IEEE Standard for Software Verification and Validation IEEE Standard for Software Verification and Validation. In *IEEEInstitute of Electrical and Electronics Engineers* (Vol. 1998, Issue July).
8. Darrick Jerome Willis. (2025). Spiritual Disciplines and Christian Formative Education: Tools to Jump-Start Spiritual Formation and Eliminate Stagnation. *Sustainability (Switzerland)*, *11*(1).
9. Goddard, A., King, V., Garcia-Perez, A., Graham, R., Jones, C. C., Tickle, A., Wilson, L., Press, F., Sumsion, J., Preston-Shoot, M., Elmes, J., Burke, P. P.-J., Ratke, D. C., Tite, P. L., Heille, G., Elias, J. W., Jagessar, M. N., Hepokoski, C., Esterline, D., … Koswatte, I. (2005). Editorial. *Times Higher Education Supplement*, *20*(1).
10. Gottlieb, M. M. M., Hetzroni, O., Agada, H., Leikin, M., Tan, P., Alant, E., Hughes, E. M. E. M. E. M. E. M. M., Yakubova, G., Weng, P.-L. L., Bouck, E. C., Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. J.-C., Clubley, E., Bae, Y. S., Chiang, H.-M. M., Hickson, L., Titeca, D., … Kageyama, R. (2021). Preface. *Journal of Autism and Developmental Disorders*, *11*(1).
11. Hopkins, Al. (2019). Wirtschafts- und Insolvenzrecht. *SSRN Electronic Journal*, *1*(1).
12. Hopkins, K., Mutasam, U., Ibrohim, I., Susilo, H., Puttick, S., Cullinane, A., Voss, S., Kruse, J. W., Kent-Schneider, I., Wei, B., Jiang, Z., Gai, L., Torres, J., Vasconcelos, C., Cavallo, A. A., Burgers, H., Ghezzi, A., van de Vrande, V., Erduran, S., … Schmidt Kjærgaard, R. (2020). Exploring the Use of Lesson Study to Develop Elementary Preservice Teachers' Pedagogical Content Knowledge for Teaching Nature of Science. *Science and Education*, *16*(2).
13. Jung, E. E. J., Jung, E. E. J., Ye, J., Kurashima, Y., Kobayashi, T., Tsuda, H., Takahara, T., Sakurai, W., Lima, L. M. M. de, Sá, L. R. de, Macambira, A. F. U. dos S., Nogueira, J. de A., Vianna, R. P. de T., Moraes, R. M. de, Barzegar, M., Sadeghi-Niaraki, A., Shakeri, M., Choi, S.-M., Oduntan, G., … Buchanan, R. (2020). Abstracts. *Journal of Planning Literature*, *10*(1).

14. Khan, M., Hussain, M., Gunasekaran, A., Ajmal, M. M., Helo, P. T., Unger, S., Landis, A., Katircioglu, K., Gooby, R., Helander, M., Drissi, Y., Chowdhary, P., Johnson, M., Yonezawa, T., Logrieco, A. F., Miller, J. D., Eskola, M., Krska, R., Ayalew, A., … Singh, A. (2019). Impact of Occupational Exposure to Chemicals in Life Cycle Assessment: A Novel Characterization Model Based on Measured Concentrations and Labor Hours. *Journal of Cleaner Production*, *11*(1).
15. Kolzow, D. R., Smith, C. C. C., Serrat, O., Dilie, H. M., Zeeshan, S., Ng, S. I., Ho, J. A., Jantan, A. H., Massey, J., Sulak, T., Sriram, R., Dennis, R. S., Bocarnea, M., Hai, T. N., Van, Q. N., Herbert, S. L., So-Jung Kim, Kyoung-Seok Kim, Y.-G. C., Guillaume, Dr. O., Honeycutt, Dr. A., … Ingram, O. C. Jr. (2021). Unit 5 Theories of Leadership. *International Journal of Organizational Leadership*, *1*(1).
16. Ministry of Education. (2018). Education strategic plan 2018 - 2030. In *Accra, Republic of Ghana.*
17. Muhammad, N. I., Judijanto, L., Sulistiani, I., Syarifuddin, & Syahrin, M. (2023). STRATEGIC CONTENT MARKETING: A MULTI-CHANNEL ANALYSIS OF THE IMPACT ON BRAND AWARENESS, CONSUMER ENGAGEMENT, AND CONVERSION RATES IN THE CONTEXT OF EVOLVING DIGITAL LANDSCAPES. *International Journal of Society Reviews (INJOSER)*.
18. Rahmawati. (2025). EFEKTIVITAS PELAYANAN DUKCAPIL ONLINE PADA DINAS KEPENDUDUKAN DAN CATATAN SIPIL DI KOTA MAKASSAR. *SKRIPSI UNIVERSITAS NEGERI MAKASSAR*, *11*(1).
19. Ronzon, T., Gurria, P., Carus, M., Cingiz, K., El-Meligi, A., Hark, N., Iost, S., M'barek, R., Philippidis, G., van Leeuwen, M., Wesseler, J., Medina-Lozano, I., Grimplet, J., Díaz, A., Tejedor-Calvo, E., Marco, P., Fischer, M., Creydt, M., Sánchez-Hernández, E., … Miras Ávalos, J. M. (2025). Understanding Material Recovery Facility Policies in the Philippines: A Scoping Review. *Sustainability (Switzerland)*, *2*(1).
20. Sani, A., Maharani, A. R., & Nazimah. (2025). Peran Komunitas Dosen Dalam Peningkatan Kualitas Pembelajaran Di Perguruan Tinggi. *Journal Genta Mulia*, *15*(1).
21. Shimray, S. R., & Subaveerapandiyan, A. (2025). AI and workforce dynamics: a

bibliometric analysis of job creation, displacement and reskilling. *Global Knowledge, Memory and Communication*. https://doi.org/10.1108/GKMC-12-2024-0824

22. Shivarama Krishna, K., & Sathish Kumar, K. (2015). A review on hybrid renewable. *Renewable and Sustainable Energy Reviews*, *52*.
23. Skyrme, D., & Amidon, D. (2010). Business research methods. *Journal of Knowledge Management*, *14*(1).
24. Sovia Lenny. (2006). RINGKASAN Puding merah (Gruptophyllum pictum L Griff). *Jurnal Farmasi Galenika (Galenika Journal of Pharmacy) (e-Journal)*, *1*(1).
25. Tàpias, M. (2019). Clothing as Solar Radiation Protection. *Autex Research Journal*, *2*(1).
26. Tursunbayeva, A., Chalutz-Ben Gal, H., Santos Fonseca, S. S., Benito, P. A., Piña Ramírez, C., Rubeis, G., Filz, M.-A., Bosse, J. P., Herrmann, C., Cavalieri, A., Reis, J., Amorim, M., Santarsiero, F., Carlucci, D., Schiuma, G., Sukums, F., Wamala-Larsson, C. W., Kisenge, R., Ravi, S. J., … Kazancoglu, Y. (2024). A Spotlight on Environmental Sustainability in View of the European Green Deal. *Sustainability (Switzerland)*, *16*(1).
27. UU Republik Indonesia, Munawir, M., Nasional, B. S., Susanta, G., Jatimnet.com, MajaMojokerto.net, Radarmojokerto.jawapos.com, Wijayakusuma, D. M. S., Nahman, A., Godfrey, L., Jacobsen, R., Buysse, J., Gellynck, X., Bayu, D., Depkes, R., Soekidjo Notoatmodjo, Boulanger, L., Ismoyo, I. H., Tarigan, R., … Sanyal, S. (2022). PENENTUAN ALTERNATIF LOKASI TEMPAT PEMBUANGAN AKHIR (TPA) SAMPAH DI KABUPATEN SIDOARJO. *Energies*, *15*(1).
28. Verifier, S. D., & Drive, A. H. (2015). Simulink ® Verification and Validation ™ Reference. *ReVision*.
29. Wang, Z., Wang, D., Xu, Y., Zhou, L., & Zhou, Y. (2025). Intelligent Computing Social Modeling and Methodological Innovations in Political Science in the Era of Large Language Models. *Journal of Chinese Political Science*. https://doi.org/10.1007/s11366-025-09917-6
30. Warhurst, A. (2002). Sustainability indicators and sustainability performance management. *Sustainability Indicators and Sustainability Performance Management*, *43*(43).

31. Waryono, T. (2019). RINGKASAN Puding merah (Gruptophyllum pictum L Griff). *Jurnal Farmasi Galenika (Galenika Journal of Pharmacy) (e-Journal)*, *1*(1).
32. Abdellatief, M., Sultan, A. B. M., Jabar, M. A., Abdullah, R., Achsan, H. T. Y., Wibowo, W. C., Agarwal, S., Tomar, D., Ahmadpour, A., Mirdamadi, M., Hosseini, J. F., Chizari, M., Al-Dosari, H., Al-Khanjari, Z. A., Kutti, N. S., Dorvlo, A. S. S., Ramadhan, H. A., AlQahtani, S. A., Al-Saadoon, G. M. W., … Srinath. (2014). What are Learning Analytics? *Journal of Computer Science*, *3*(1).
33. Dini, M. (2022). Pengaruh academic self efficacy, komitmen beragama dan prokrastinasi terhadap ketidak jujuran akademik dimoderasi oleh social desirability pada mahasiswa diperguruan tinggi keagamaan islam swasta. *Science*, *7*(1).
34. Annuš, N., & Kmeť, T. (2024). Learn with M.E.—Let Us Boost Personalized Learning in K-12 Math Education! *Education Sciences*, *14*(7). https://doi.org/10.3390/educsci14070773
35. Fardian, D., Suryadi, D., Prabawanto, S., & Jupri, A. (2025). Integrating Chat-GPT in the Classroom: A Study on Linear Algebra Learning in Higher Education. *International Journal of Information and Education Technology*, *15*(4). https://doi.org/10.18178/ijiet.2025.15.4.2279
36. Foy, P. (2023). *Large Language Models (LLMs) in Venture Capital - This Week in VC*. MLQ AI.
37. Kulkarni, R. v., Choudhari, G., Bhavik, M., Bura, R., & Bhosale, V. (2025). GAMIFYING EARLY MATH EDUCATION: AN INNOVATIVE AI GAME FOR CHILDREN. *Journal of Information Technology Education: Innovations in Practice*, *24*. https://doi.org/10.28945/5595
38. Kumar Kanvaria, V., & Tarance Suraj, M. (2024). The Role Of AI In Mathematics Education: Assessing The Effects Of An "Auto Draw" Webtool On Middle-Level Achievement. *The Online Journal of Distance Education and E-Learning, 12*(1).
39. Maguey-González, J. A., Nava-Ramírez, M. de J., Gómez-Rosales, S., Ángeles, M. de L., Solís-Cruz, B., Hernández-Patlán, D., Merino-Guzmán, R., Hernández-Velasco, X., Figueroa-Cárdenas, J. de D., Vázquez-Durán, A., Hargis, B. M., Téllez-Isaías, G.,

Méndez-Albores, A., Angelico, R., Colombo, C., di Iorio, E., Brtnický, M., Fojt, J., Conte, P., … Chorover, J. (2018). Modeling of Soil Cation Exchange Capacity Based on. *Geoderma*, *3*(4).

40. Muñoz-Olano, J. F., Hurtado-Parrado, C., Troll, E. S., Friese, M., Loschelder, D. D., Alblwi, A., McAlaney, J., al Thani, D. A. S., Phalp, K., Ali, R., Elhai, J. D., Sapci, O., Yang, H., Amialchuk, A., Rozgonjuk, D., Montag, C., Bäulke, L., Daumiller, M., Dresel, M., … Cevallos Pozo, G. (2020). A review of web-based support systems for students in higher education. *International Journal of Environmental Research and Public Health*, *10*(1).
41. Pérez-Santín, E., De-La-fuente-valentín, L., García, M. G., Bravo, K. A. S., Hernández, F. C. L., & Sánchez, J. I. L. (2023). Applicability domains of neural networks for toxicity prediction. In *AIMS Mathematics* (Vol. 8, Issue 11). https://doi.org/10.3934/math.20231426
42. Saragih, B., & , Ika Karyati, D. S. (2010). Pengaruh Pewarna Ekstrak Cair Alami Bawang Tiwai (Eleutherine americana Merr) Terhadap Mutu Selai Kulit Pisang Kepok (Musa paradisiaca Linn). *Jurnal Tenknologi Pertanian*, *6*(2).
43. Sart, G. (2017). HOW DOES THE MAKER MOVEMENT AFFECT THE SCHOOL ADMINISTRATIONS' ROLES ON IMPROVING THE STEM EDUCATION? *INTED2017 Proceedings*, *1*. https://doi.org/10.21125/inted.2017.2452
44. Sentioningtias, I. (2025). Pemanfaatan Aplikasi Math Kids Untuk Mengembangkan Berhitung Anak Usia Dini Di TK Islam Nrhasanah Bandar Lampung. In *Sustainability (Switzerland)* (Vol. 11, Issue 1).
45. Starrett, A., Yow, J., Lotter, C., Irvin, M. J., Adams, P., Guillemot, F., Lacroix, F., Nocus, I., Assen, J. H. E., Otting, H., Jin, X., Tigelaar, D., van der Want, A., Admiraal, W., Peschl, H., Deng, C., Larson, N., Kuno, C. B., Hein, S., … León, J. (2022). Editorial Board. *Teaching and Teacher Education*, *90*(3).
46. van Merrienboer, J. J. G., Paas, F., Traxler, A., Brewe, E., Hazari, Z., Cass, C., Beattie, C., Kelly, A. M., Wilcox, B. R., Lewandowski, H. J., Bao, L., Xiao, Y., Koenig, K., Han, J., Eisinga, R., Grotenhuis, M., Pelzer, B., Cronbach, L. J., Butts, M. M., … VanLehn,

K. A. (2016). Evaluation of Colorado Learning Attitudes about Science Survey. *Physical Review Physics Education Research*, *12*(2).

47. Vicente, N. R. (2010). Sistemas agroflorestais sucessionais como estratégia de uso e conservação de recursos florestais em zonas ripárias da Microbacia Arroio Primeiro de Janeiro, Anchieta-SC. *Journal of Mental Science*, *7*(37).

Conclusion

To do research in Artificial Intelligence (AI) and Data Science, we must first understand the nature of the knowledge that these disciplines produce. Unlike traditional programming, where knowledge is explicit and encoded by humans, modern AI infers knowledge from observation. This chapter explores that paradigm shift and its implications for scientific validity.

Historically, classical computing operated under a deductive logic. The programmer acted as a legislator, writing explicit rules (if-then) that the system blindly obeyed. If we wanted to detect spam, we wrote: *"If the email contains the word 'free' and 'urgent', then it is spam."*

Artificial Intelligence, specifically Machine Learning, reverses this process towards inductive logic.

- *Traditional approach:* data + rules = answers.
- *Machine Learning Approach:* Data + Answers = Rules.

In research, this implies that we no longer design the solution directly; we design the *mechanism* that will find the solution. Epistemologically, this is risky: the model can find patterns that are statistically correct in the training data, but conceptually false in the real world (*overfitting*). AI research means auditing these learned "rules" to ensure that they represent valid knowledge and not mere numerical matches.

A common mistake in early childhood researchers is to confuse the *dataset* with reality. In the methodology of science, we must remember that data is always a reduction of reality.

- *Representation Bias:* If we train a facial recognition model only with images of light-skinned people, the model has not learned to "see humans", it has learned to "see clear pixel patterns".
- *The fallacy of objectivity:* It is often believed that "data does not lie". However, data collection, cleaning, and labeling are subjective processes fraught with human decisions.

Deep learning models often act as "black boxes". We can observe the input and output, but the internal transformations are so complex (millions of parameters) that they are unintelligible to human cognition. This raises a fundamental question for further research on the subject: Is it science if I can't explain how it works?

With *Big Data*, it is easy to find spurious correlations. If we analyze enough variables, we will find by pure chance that "cheese consumption per capita" correlates perfectly with "the number of people who died entangled in their sheets" (a classic example of spurious correlation).

Algorithms are machines for looking for correlations, not for understanding causes. The researcher must provide the theoretical framework. Scientific methodology demands that we move from the question *"What will happen?"* (prediction) to *"What if...?"* (causal inference). Without this step, our models are fragile in the face of any change in the environment.

Research in AI and Data Science is not just software engineering; it is a form of radical empiricism. As researchers, our responsibility does not end with obtaining high *accuracy*. Our job is to validate that this *accuracy* comes from real, causal and generalizable patterns, and not from hidden biases in the data or statistical coincidences. In the following chapters, we will see how to translate these philosophical principles into a concrete experimental design.

Bibliography

Abadio, A. K. R., Kioshima, E. S., Martins, N. F., Maigret, B., Felipe, M. S. S., Freitas, S. M., Maigret, B., Felipe, M. S. S., Abagyan, R. A., Kufareva, I., Totrov, M., Abbasi, M., Sadeghi-Aliabadi, H., Hassanzadeh, F., Amanlou, M., Abdelmonsef, A. H., Dulapalli, R., Dasari, T., Padmarao, L. S., … Olson, A. J. A. J. (2015). A pharmacophore-based virtual screening approach for the discovery of Trypanosoma cruzi GAPDH inhibitors (vol 5, pg 2019, 2013). *J Chem Inf Model*. https://doi.org/Book_Doi 10.1002/9783527633326

Abdellatief, M., Sultan, A. B. M., Jabar, M. A., Abdullah, R., Achsan, H. T. Y., Wibowo, W. C., Agarwal, S., Tomar, D., Ahmadpour, A., Mirdamadi, M., Hosseini, J. F., Chizari, M., Al-Dosari, H., Al-Khanjari, Z. A., Kutti, N. S., Dorvlo, A. S. S., Ramadhan, H. A., AlQahtani, S. A., Al-Saadoon, G. M. W., … Srinath. (2014). What are Learning Analytics? *Journal of Computer Science*, *3*(1).

Abdul Rahman, H., Masri, N. A., Husaini, A., Shaharuddin, M. Y., & Zaim, S. N. N. (2025). Integration of Machine Learning and Artificial Intelligence Techniques for Qualitative Research: The Rise of New Research Paradigms. *International Journal of Qualitative Methods* , *24*. https://doi.org/10.1177/16094069251347545

Abdul-Rahman, G., Zwitter, A., & Haleem, N. (2025). A systematic literature review on the role of artificial intelligence in citizen science. *Discover Artificial Intelligence*, *5*(1). https://doi.org/10.1007/s44163-025-00437-z

Abisoye, A. (2024). A Conceptual Framework for Integrating Artificial Intelligence into STEM Research Methodologies for Enhanced Innovation. *International Journal of Future Engineering Innovations*, *1*(1). https://doi.org/10.54660/ijfei.2024.1.1.56-66

Abu Sharshouh, A. (2025). The Use of Artificial İntelligence in Accounting and Auditing TT - Muhasebe ve Denetim Alanında Yapay Zekâ Kullanımı. *Karadeniz Ekonomi Araştırmaları Dergisi*, *6*(1).

Access, O., Adams, W. C., Alfonita, F., Alfred, M., Briggs, S., Aloisi, A., Amorim, H., Moda, F., Anderson, D. N., Anwar, M. A., Graham, M., Apouey, B., Roulet, A., Solal, I., Stabile, M., Background, P., Bajwa, U., Gastaldo, D., di Ruggiero, E., … ZACCARIA, M. L. (2019). No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共

分散構造分析 Title. *Economic and Labour Relations Review*, *14*(1).

Ackoff, R. L., Aguilera-Caracuel, J., Hurtado-Torres, N. E., Aragón-Correa, J. A., Al-Mawali, H., Al-Shbiel, S. O., Alavi, M., Kayworth, T. R., Leidner, D. E., Albino, V., Garavelli, A. C., Gorgoglione, M., Alfaadhel, S., Allame, S. M., Nouri, B. A., Tavakoli, S. Y., Shokrani, S. A. R., Allame, S. M., Nouri, B. A., … van Tartwijk, J. (2010). Business research methods. *Journal of Knowledge Management*, *14*(1).

Acomi, N., Surugiu, G., Raicu, G., & Stanca, C. (2025). A Thematic Review of Port Services and Emission Reduction Strategies Using ATLAS.ti. *TransNav*, *19*(2). https://doi.org/10.12716/1001.19.02.14

Aggarwal, K., Mijwil, M. M., Sonia, Al-Mistarehi, A. H., Alomari, S., Gök, M., Zein Alaabdin, A. M., & Abdulrhman, S. H. (2022). Has the Future Started? The Current Growth of Artificial Intelligence, Machine Learning, and Deep Learning. *Iraqi Journal for Computer Science and Mathematics*, *3*(1). https://doi.org/10.52866/ijcsm.2022.01.01.013

Ahmadpour, M. (2024). Emerging Approaches in Enhancing Mental and Physical Health of the Elderly: Integration of Strength Training, Mindfulness, and Artificial Intelligence. *Journal of Psychological Dynamics in Mood Disorders*, *3*(5). https://doi.org/10.61838/kman.pdmd.3.5.19

Al, P., Tentang, G., Arab, P. B., Relevansinya Di Era, D., Lailatul, G., & Maftuhah, H. (2025). Perspekif Al Ghazali Tentang Pendidikan Bahasa Arab Dan Relevansinya di Era Globalisasi. *Taqdir*, *11*(1).

Alshadidi, A. A. F., Alshahrani, A. A., Aldosari, L. I. N., Chaturvedi, S., Saini, R. S., Hassan, S. A. bin, Cicciù, M., & Minervini, G. (2023). Investigation on the Application of Artificial Intelligence in Prosthodontics. In *Applied Sciences (Switzerland)* (Vol. 13, Issue 8). https://doi.org/10.3390/app13085004

Anis, S., & French, J. A. (2023). Efficient, Explicatory, and Equitable: Why Qualitative Researchers Should Embrace AI, but Cautiously. In *Business and Society* (Vol. 62, Issue 6). https://doi.org/10.1177/00076503231163286

Annus, N., & Kmet, T. (2024). Learn with M.E.-Let Us Boost Personalized Learning in K-12 Math Education! *EDUCATION SCIENCES*, *14*(7).

Arslantaş, S. (2024). Artificial intelligence and big data from digital health applications: publication trends and analysis. *Journal of Health Organization and Management*. https://doi.org/10.1108/JHOM-06-2024-0241

Asaduzzaman, M., & Flowra, M. T. (2018). Resurgence of infectious diseases due to forced migration: is planetary health and One Health action synergistic? *The Lancet Planetary Health*, *2*(10).

Asiva Noor Rachmayani, Inaam Akhtar, M., Njie, B., Asimiran, S., Alase, A., Etikan, I., Laughlin, R., Francis, J. J., Johnston, M., Robertson, C., Glidewell, L., Entwistle, V., Eccles, M. P., Grimshaw, J. M., Alhojailan, M. I., Ibrahim, M., Braun, V., Clarke, V., Taherdoost, H., … Zairul, M. (2019). No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析 Title. *ISRA International Journal of Islamic Finance*, *8*(1).

Avula, S., Chadwick, G., Arguello, J., Capra, R., Lippold, D., Bana e Costa, C. A., de Corte, J.-M., Vansnick, J.-C., Linstone, H. A., Carver, C. S., Scheier, M. F., Segerstrom, S. C., Berger, G., Armand, L., Bresard, S., Grimanelli, P., Landucci, A., Schwartz, B., Bolger, F., … Birbilis, G. (2015). Editorial Board. *Technological Forecasting and Social Change*, *77*(1).

Axiotis, E., Kontogiannis, A., Kalpoutzakis, E., & Giannakopoulos, G. (2021). A personalized machine-learning-enabled method for efficient research in ethnopharmacology. The case of the southern balkans and the coastal zone of asia minor. *Applied Sciences (Switzerland)*, *11*(13). https://doi.org/10.3390/app11135826

Benedeti Rosa, M., & Kubota, L. C. (2025). Artificial intelligence: where does Brazil stand in global scientific production and what are the main technical determinants of adoption by Brazilian companies. *EconomiA*. https://doi.org/10.1108/ECON-01-2025-0010

Bhardwaj, M., Kumar, P., & Singh, A. (2025). Bibliometric review of digital transformation in agriculture: innovations, trends and sustainable futures. *Journal of Agribusiness in Developing and Emerging Economies*. https://doi.org/10.1108/JADEE-02-2025-0083

Bhatia, A., Chandani, A., Atiq, R., Mehta, M., & Divekar, R. (2021). Artificial intelligence in financial services: a qualitative research to discover robo-advisory services. *Qualitative Research in Financial Markets*, *13*(5). https://doi.org/10.1108/QRFM-10-2020-0199

Bock, T., & von der Oelsnitz, D. (2025). Leadership-competences in the era of artificial intelligence – a structured review. In *Strategy and Leadership* (Vol. 53, Issue 3). https://doi.org/10.1108/SL-09-2024-0100

Boukrouh, I., & Azmani, A. (2024). ARTIFICIAL INTELLIGENCE APPLICATIONS IN E-COMMERCE: A BIBLIOMETRIC STUDY FROM 1995 TO 2023 USING MERGED DATA SOURCES. *International Journal of Professional Business Review*, *9*(4). https://doi.org/10.26668/businessreview/2024.v9i4.4537

Brailas, A. (2025). Artificial Intelligence in Qualitative Research: Beyond Outsourcing Data Analysis to the Machine. *Psychology International*, *7*(3). https://doi.org/10.3390/psycholint7030078

Bris, A., Wang, T. Y. H., Zatzick, C. D., Miller, D. J. P., Fern, M. J., Cardinal, L. B., Gregoire, D. A., Shepherd, D. A., Westphal, J. D., Shani, G., Troster, C., van Quaquebeke, N., Lanaj, K., Hollenbeck, J. R., Ilgen, D. R., Barnes, C. M., Harmon, S. J., Feldman, E. R., DesJardine, M. R., … Sangiorgi, F. (2021). KNIGHTS, RAIDERS, AND TARGETS - THE IMPACT OF THE HOSTILE TAKEOVER - COFFEE,JC, LOWENSTEIN,L, ROSEACKERMAN,S. *JOURNAL OF BANKING & FINANCE*, *37*(1).

Brondani, M., Alves, C., Ribeiro, C., Braga, M. M., Garcia, R. C. M., Ardenghi, T., & Pattanaporn, K. (2024). Artificial intelligence, ChatGPT, and dental education: Implications for reflective assignments and qualitative research. *Journal of Dental Education*, *88*(12). https://doi.org/10.1002/jdd.13663

Buchkremer, R., Demund, A., Ebener, S., Gampfer, F., Jagering, D., Jurgens, A., Klenke, S., Krimpmann, D., Schmank, J., Spiekermann, M., Wahlers, M., & Wiepke, M. (2019). The Application of Artificial Intelligence Technologies as a Substitute for Reading and to Support and Enhance the Authoring of Scientific Review Articles. In *IEEE Access* (Vol. 7). https://doi.org/10.1109/ACCESS.2019.2917719

Budiman, A. S., Putri, D. N. N., Candra, H., Djuana, E., Sari, T. K., Aji, D. P. B., Putri, L. R., Sitepu, E., Speaks, D., & Pasang, T. (2024). Crack Catcher AI – Enabling smart fracture mechanics approaches for damage control of thin silicon cells or wafers. *Solar Energy Materials and Solar Cells*, *273*. https://doi.org/10.1016/j.solmat.2024.112927

Cafuta, D., Dodig, I., Cesar, I., & Kramberger, T. (2021). Developing a modern greenhouse scientific research facility — A case study. *Sensors*, *21*(8). https://doi.org/10.3390/s21082575

Candaika, P. (2022). Gambaran Kesiapan Rumah Sakit Umum Daerah (RSUD) Menghadapi Kebijakan Kelas Rawat Inap Standar (KRIS) Tahun 2022. *Science*, *7*(1).

Cantarelli, C. C., Flybjerg, B., Molin, E. J. E., & Wee, B. van. (2010). Cost Overruns in Large-Scale Transport Infrastructure Projects. *European Journal of Transport and Infrastructure Research*, *10*(1).

Casanova Pistón, A., & Martínez Domínguez, M. (2024). Producción científica sobre Inteligencia Artificial y educación: un análisis cienciométrico. *Hachetetepé. Revista Científica de Educación y Comunicación*, *28*. https://doi.org/10.25267/hachetetepe.2024.i28.1102

Chango, X., Flor-Unda, O., Gil-Jiménez, P., & Gómez-Moreno, H. (2024). Technology in Forensic Sciences: Innovation and Precision. In *Technologies* (Vol. 12, Issue 8). https://doi.org/10.3390/technologies12080120

Cheekaty, S., & Muneeswari, G. (2024). Exploring Sparse Gaussian Processes for Bayesian Optimization in Convolutional Neural Networks for Autism Classification. *IEEE Access*, *12*. https://doi.org/10.1109/ACCESS.2024.3351168

Chen, S., & Garcia, A. (2017). Embodiment of Cognition: Exploring the Interplay between Body and Mind. *Studies in Science of Science | ISSN: 1003-2053*, *35*(9).

Cheng, J. C. P., Chen, W., Chen, K., Wang, Q., Sadri, H., Karki, B. R., Basnet, S., Xiang, J., Montoya, J., Porras, J., Ucar, A., Karakose, M., Kırımça, N., Abdelalim, A. M., Essawy, A., Alnaser, A. A., Shibeika, A., Sherif, A., van Roosmale, S., … Liu, G. (2024). applied sciences Artificial Intelligence for Predictive Maintenance Applications : *Journal of Building Engineering*, *12*(1).

Chodvadiya, J., Shukla, G., & Sharma, G. (2025). 16S-23S-5S rRNA Database: a comprehensive integrated database of archaeal and bacterial rRNA sequences, alignments, intragenomic heterogeneity, and secondary structures. *Sustainability (Switzerland)*, *11*(1).

Choudrie, J., Manandhar, N., Castro, C., & Obuekwe, C. (2023). Hey Siri, Google! Can you help me? A qualitative case study of smartphones AI functions in SMEs. *Technological Forecasting and Social Change*, *189*. https://doi.org/10.1016/j.techfore.2023.122375

Christou, P. (2023). How to Use Artificial Intelligence (AI) as a Resource, ow Methodological and Analysis Tool in Qualitative Research? *Qualitative Report*, *28*(7). https://doi.org/10.46743/2160-3715/2023.6406

Christou, P. A. (2023a). A Critical Perspective Over Whether and How to Acknowledge the Use of Artificial Intelligence (AI) in Qualitative Studies. *Qualitative Report*, *28*(7). https://doi.org/10.46743/2160-3715/2023.6407

Christou, P. A. (2023b). The Use of Artificial Intelligence (AI) in Qualitative Research for Theory Development. *Qualitative Report*, *28*(9). https://doi.org/10.46743/2160-3715/2023.6536

Colina Vargas, A. M., Espinoza Mina, M. A., Catalán, L. L., & Catalán, B. L. (2024). Computational intelligence, educational robotics, and artificial intelligence in the educational field. A bibliometric study and thematic modelling. *International Journal of Educational Research and Innovation*, *2024*(22). https://doi.org/10.46661/ijeri.10369

Committee, S. (1998). IEEE Standard for Software Verification and Validation IEEE Standard for Software Verification and Validation. In *IEEEInstitute of Electrical and Electronics Engineers* (Vol. 1998, Issue July).

Cricchio, J. (2025). Balancing openness and ownership: open innovation strategies for AI development. *European Journal of Innovation Management*. https://doi.org/10.1108/EJIM-04-2024-0470

Dai, Y., Lai, S., Lim, C. P., & Liu, A. (2023). ChatGPT and its impact on research supervision: Insights from Australian postgraduate research students. *Australasian Journal of Educational Technology*, *39*(4). https://doi.org/10.14742/ajet.8843

Darrick Jerome Willis. (2025). Spiritual Disciplines and Christian Formative Education: Tools to Jump-Start Spiritual Formation and Eliminate Stagnation. *Sustainability (Switzerland)*, *11*(1).

de Seta, G., Pohjonen, M., & Knuutila, A. (2024). Synthetic ethnography: Field devices for the qualitative study of generative models. *Big Data and Society*, *11*(4).

https://doi.org/10.1177/20539517241303126

Demidenko, S. Y., & Romanovskiy, N. v. (2023). THEORETICAL UNDERSTANDING OF SOCIAL REALITY: THE 25th ANNIVERSARY OF KHARCHEV READINGS. *Sotsiologicheskie Issledovaniya*, *2023*(9). https://doi.org/10.31857/S013216250027774-4

Dini, M. (2022). Pengaruh academic self efficacy, komitmen beragama dan prokrastinasi terhadap ketidak jujuran akademik dimoderasi oleh social desirability pada mahasiswa diperguruan tinggi keagamaan islam swasta. *Science*, *7*(1).

Djakfar Musthafa, F. A. (2024). Penggunaan Artificial Intelligence (AI) dalam Pembelajaran: Fenomena Transformasi Otoritas Pengetahuan di Kalangan Mahasiswa. *Journal of Contemporary Islamic Education*, *4*(1). https://doi.org/10.25217/jcie.v4i1.4386

Durrani, B. A., Lecturer, S., Ivanauskien, N., Volung, J., Rahmani-Nejad, L., Firoozbakht, Z., Taghipoor, A., Arokiasamy, A. R. A., Pourdehghan, A., Tjan, S., Sukamto, R., Lumintan, D. B., Dhurup, M., Mafini, C., Dumasi, T., Kheng, L. L., Mahamad, O., Ramayah, T., Mosahab, R., … Wirawan, V. (2014). Implementation of E-Government in Welcoming the Contemporary Industrial Revolution 4.0 Era in Indonesia. *International Journal of Marketing Studies*, *1*(1).

Educause. (2018). Horizon report: 2018 higher education edition. Retrieved from EDUCAUSE Learning Initiative and The New Media Consortium. *Liila*.

Espinel, R., Herrera-Franco, G., Rivadeneira García, J. L., & Escandón-Panchana, P. (2024). Artificial Intelligence in Agricultural Mapping: A Review. In *Agriculture (Switzerland)* (Vol. 14, Issue 7). https://doi.org/10.3390/agriculture14071071

Evolution Analysis and Prediction of Scientific Topic Popularity in the Field of LIS. (2021). *Documentation, Information and Knowledge*, *38*(2). https://doi.org/10.13366/j.dik.2021.02.035

Fahyuni, E. F., Wasis, Bandono, A., & Arifin, M. B. U. B. (2020). Integrating islamic values and science for millennial students' learning on using seamless mobile media. *Jurnal Pendidikan IPA Indonesia*, *9*(2). https://doi.org/10.15294/jpii.v9i2.23209

Fardian, D., Suryadi, D., Prabawanto, S., & Jupri, A. (2025). Integrating Chat-GPT in the Classroom: A Study on Linear Algebra Learning in Higher Education. *International Journal of Information and Education Technology*, *15*(4). https://doi.org/10.18178/ijiet.2025.15.4.2279

Fieldhouse, J. K., Ge, J., Randhawa, N., Wolking, D., Genovese, B. N., Mazet, J. A. K., & Desai, A. (2025). The intersection of artificial intelligence with qualitative or mixed methods for communicable disease research: a scoping review. In *Public Health* (Vol. 248). https://doi.org/10.1016/j.puhe.2025.105961

Fil, N., Slisarenko, R., Deineko, Z., & Morozova, L. (2025). Trends in Artificial Intelligence Research on Education: Topic Modeling Using Latent Dirichlet Allocation. *Bulletin of Kharkov National Automobile and Highway University*, *108*. https://doi.org/10.30977/bul.2219-5548.2025.108.0.17

Flores, L. G. R., & Mejía Rivera, K. A. (2024). ARTIFICIAL INTELLIGENCE IN QUALITATIVE RESEARCH: BIBLIOMETRIC ANALYSIS OF SCIENTIFIC PRODUCTION INDEXED IN SCOPUS. *New Trends in Qualitative Research*, *20*(4). https://doi.org/10.36367/ntqr.20.4.2024.e1116

Foy, P. (2023). *Large Language Models (LLMs) in Venture Capital - This Week in VC*. MLQ AI.

Friedman, C., Owen, A., & VanPuymbrouck, L. (2025). Should ChatGPT help with my research? A caution against artificial intelligence in qualitative analysis. *Qualitative Research*, *25*(5). https://doi.org/10.1177/14687941241297375

Georgiev, G. (2021). THE BREAKDOWN OF THE PUBLIC-PRIVATE DIVIDE IN SECURITIES LAW: CAUSES, CONSEQUENCES, AND REFORMS GEORGE. *Frontiers in Neuroscience*, *14*(1).

Göçer, Ö., Hua, Y., Göçer, K., Gunhan, S., Moaveni, S., Chou, K. C., Charlson, J., Chinyio, E., Zhu, Y. Q., Zuo, J. J., Korpela, J., Kerosuo, H., Abuelmaatti, A., Ahmed, V., Jing, X., Wang, Q. K. Q., Li, Y. Y. Y., Zhao, D., McCoy, A. P., … Wu, W. W. Y. (2014). A Ship in a Box. *Applied Mechanics and Materials*, *29*(2).

Goddard, A., King, V., Garcia-Perez, A., Graham, R., Jones, C. C., Tickle, A., Wilson, L., Press, F., Sumsion, J., Preston-Shoot, M., Elmes, J., Burke, P. P.-J., Ratke, D. C., Tite, P. L., Heille, G., Elias, J. W., Jagessar, M. N., Hepokoski, C., Esterline, D., … Koswatte, I. (2005). Editorial. *Times Higher Education Supplement*, *20*(1).

Gong, J., Zhao, Z., Niu, X., Ji, Y., Sun, H., Shen, Y., Chen, B., & Wu, B. (2025). AI reshaping life sciences: intelligent transformation, application challenges, and future convergence in neuroscience, biology, and medicine. In *Frontiers in Digital Health* (Vol. 7). https://doi.org/10.3389/fdgth.2025.1666415

Gottlieb, M. M. M., Hetzroni, O., Agada, H., Leikin, M., Tan, P., Alant, E., Hughes, E. M. E. M. E. M. E. M. M., Yakubova, G., Weng, P.-L. L., Bouck, E. C., Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. J.-C., Clubley, E., Bae, Y. S., Chiang, H.-M. M., Hickson, L., Titeca, D., … Kageyama, R. (2021). Preface. *Journal of Autism and Developmental Disorders*, *11*(1).

Govindan, K., Cheng, T. C. E., Mishra, N., Shukla, N., Pei-Chen, L., Arya, V., Sharma, P., Singh, A., de Silva, P. T. M., Chatfield, A. T., Reddick, C. G., Karlsson, P.-M., Lal, S., Rexhausen, D., Ajak, A. D., Lilford, E., Topal, E., Zhang, Y. Y., Fu, Z., … Perez-Franco, R. (2018). Industry Surveys IT Consulting & Other Services. *Journal of Cleaner Production*, *25*(1).

Goyanes, M., Lopezosa, C., & Jordá, B. (2025). Thematic analysis of interview data with ChatGPT: designing and testing a reliable research protocol for qualitative research. *Quality and Quantity*, *59*(6). https://doi.org/10.1007/s11135-025-02199-3

Guizzardi, G., & Guarino, N. (2024). Explanation, semantics, and ontology. *Data and Knowledge Engineering*, *153*. https://doi.org/10.1016/j.datak.2024.102325

Guliev, I. A., Mammadov, A., & Ibrahimli, K. (2024). The Use of Artificial Intelligence Technologies in Energy and Climate Security. *Review of Business and Economics Studies*, *12*(4). https://doi.org/10.26794/2308-944X-2024-12-4-58-71

Gunewardena, D., Seck, A., Warnecke, T., Green, E. P., Blattman, C., Jamison, J., Annan, J., Dvouletý, O., Pepelasis, I. S., Protogerou, A., Krasniqi, B. A., Williams, N., Osirim, M. J., Tobias, J. M., Boudreaux, K. C., Botterill, L. C., Free, M., Demirgüc-Kunt, A., Klapper, L. F., … Naszályi, P. (2018). Abstracts. *Journal of Planning Literature*, *29*(2).

Guo, Z. (2024). Innovative Functionality of the Scientific and Technical Information System of China. *Visnyk of Kharkiv State Academy of Culture*, *65*. https://doi.org/10.31516/2410-5333.065.04

Gupta, S., Mahmood, A., Shetty, P., Adeboye, A., & Ramprasad, R. (2024). Data extraction from polymer literature using large language models. *Communications Materials*, *5*(1). https://doi.org/10.1038/s43246-024-00708-9

Güzer, B., Caner, H., Fathoni, A. F., van der Mescht, H., Whittaker, J. K., Staker, H., Horn, M. B., Crawford, R., Richey, R. C., Christensen, C. M., Horn, M. B., Staker, H., Thoms, C. L. v, Chen, J. J., Petty, J., Albiladi, W. S., Alshareef, K. K., Zhang, W., Zhu, C., … Province,

Q. (2019). No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析 Title. *Computers and Education*, *7*(1).

Harahulia, S., & Symonenko, T. (2025). RESEARCH SUPPORT SERVICES IN DIGITAL HUMANITIES:APPROACHES, TECHNOLOGIES, TOOLS. *Manuscript and Book Heritage of Ukraine*, *2025*(2). https://doi.org/10.15407/rksu.37.189

Hogg, H. D. J., Al-Zubaidy, M., Keane, P. A., Hughes, G., Beyer, F. R., & Maniatopoulos, G. (2023). Evaluating the translation of implementation science to clinical artificial intelligence: a bibliometric study of qualitative research. In *Frontiers in Health Services* (Vol. 3). https://doi.org/10.3389/frhs.2023.1161822

Hogg, H. D. J., Al-Zubaidy, M., Talks, J., Denniston, A. K., Kelly, C. J., Malawana, J., Papoutsi, C., Teare, M. D., Keane, P. A., Beyer, F. R., & Maniatopoulos, G. (2023). Stakeholder Perspectives of Clinical Artificial Intelligence Implementation: Systematic Review of Qualitative Evidence. In *Journal of Medical Internet Research* (Vol. 25). https://doi.org/10.2196/39742

Hogg, H. D. J., Sendak, M. P., Denniston, A. K., Keane, P. A., & Maniatopoulos, G. (2023). Unlocking the potential of qualitative research for the implementation of artificial intelligence-enabled healthcare. In *Journal of Medical Artificial Intelligence* (Vol. 6). https://doi.org/10.21037/jmai-23-28

Hopkins, Al. (2019). Wirtschafts- und Insolvenzrecht. *SSRN Electronic Journal*, *1*(1).

Hopkins, K., Mutasam, U., Ibrohim, I., Susilo, H., Puttick, S., Cullinane, A., Voss, S., Kruse, J. W., Kent-Schneider, I., Wei, B., Jiang, Z., Gai, L., Torres, J., Vasconcelos, C., Cavallo, A. A., Burgers, H., Ghezzi, A., van de Vrande, V., Erduran, S., … Schmidt Kjærgaard, R. (2020). Exploring the Use of Lesson Study to Develop Elementary Preservice Teachers' Pedagogical Content Knowledge for Teaching Nature of Science. *Science and Education*, *16*(2).

Hou, J., Zheng, B., Li, H., & Li, W. (2025). Evolution and impact of the science of science: from theoretical analysis to digital-AI driven research. In *Humanities and Social Sciences Communications* (Vol. 12, Issue 1). https://doi.org/10.1057/s41599-025-04617-1

Inaam Akhtar, M., Njie, B., Asimiran, S., Alase, A., Etikan, I., Laughlin, R., Francis, J. J., Johnston, M., Robertson, C., Glidewell, L., Entwistle, V., Eccles, M. P., Grimshaw, J. M., Alhojailan, M. I., Ibrahim, M., Braun, V., Clarke, V., Taherdoost, H., SWELHEIM, T., … Zairul, M. (2019). No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析 Title. *ISRA International Journal of Islamic Finance*, *8*(1).

Irtyshcheva, I., Kim, V., & Kletsov, Y. (2022). MANAGING BUSINESS PROCESSES IN THE CONDITIONS OF THE DEVELOPMENT OF THE DIGITAL ECONOMY: GLOBAL AND NATIONAL EXPERIENCES. *Baltic Journal of Economic Studies*, *8*(5). https://doi.org/10.30525/2256-0742/2022-8-5-101-107

Iurasova, O., Ivashko, L., Maksymova, J., & Maksymov, O. (2025). INNOVATION TRENDS: IMPACT ON DEVELOPMENT OF FINANCIAL SECTOR. *Смарт-Економіка, Підприємництво Та Безпека*, *3*(1). https://doi.org/10.60022/sis.3.(01).1

Jain, H., & KUMAR, Prof. D. (2025). Artificial Intelligence as a Catalyst for Supply Chain Evolution: A Multi-Case Thematic Study. *INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, *09*(09). https://doi.org/10.55041/ijsrem52737

Jeon, Y. (2021). A Study on the Ethics of the Concept and Purpose of Culture Technology in the Post-COVID Era. *Society for International Cultural Institute*, *14*(1). https://doi.org/10.34223/jic.2021.14.1.1

Jiang, Y., Ko-Wong, L., & Valdovinos Gutierrez, I. (2025). The Feasibility and Comparability of Using Artificial Intelligence for Qualitative Data Analysis in Equity-Focused Research. In *Educational Researcher* (Vol. 54, Issue 3). https://doi.org/10.3102/0013189X251314821

Jirandehi, A. B., & Sajadi, S. M. (2024). Co-Citation Analysis and Co-Occurrence Analysis of Resilience Keywords in StartUps. *Scientometrics Research Journal*, *10*(2). https://doi.org/10.22070/rsci.2024.17648.1667

John, A. O., & Chukwumerije, T. (2014). Perceptions of Accountants on Factors Affecting Auditor’s Independence in Nigeria. *International Journal of Research in Commerce, IT & Management*, *2*(11).

Jung, E. E. J., Jung, E. E. J., Ye, J., Kurashima, Y., Kobayashi, T., Tsuda, H., Takahara, T., Sakurai, W., Lima, L. M. M. de, Sá, L. R. de, Macambira, A. F. U. dos S., Nogueira, J. de A., Vianna, R. P. de T., Moraes, R. M. de, Barzegar, M., Sadeghi-Niaraki, A., Shakeri, M., Choi, S.-M., Oduntan, G., … Buchanan, R. (2020). Abstracts. *Journal of Planning Literature*, *10*(1).

Jungmeier, G. (2017). The Biorefinery Fact Sheet. *The International Journal of Life Cycle Assessment*, *23*(1).

Júnior, G. H. K., & Gottschalg-Duque, C. (2023). Information architecture applied on natural language processing: a proposal Information Science contributions on data preprocessing for training and learning of artificial neural networks. *Revista Digital de Biblioteconomia e Ciencia Da Informacao*, *21*. https://doi.org/10.20396/rdbci.v21i00.8671396/30919

Kang, R., Xuan, Z., Tong, L., Wang, Y., Jin, S., & Xiao, Q. (2025). Nurse Researchers' Experiences and Perceptions of Generative AI: Qualitative Semistructured Interview Study. *Journal of Medical Internet Research*, *27*. https://doi.org/10.2196/65523

Kantor, J. (2024). Best practices for implementing ChatGPT, large language models, and artificial intelligence in qualitative and survey-based research. In *JAAD International* (Vol. 14). https://doi.org/10.1016/j.jdin.2023.10.001

Khabsa, M., Elmagarmid, A., Ilyas, I., Hammady, H., Ouzzani, M., Kouznetsov, A., Japkowicz, N., Visser, E., Felizardo, K. K. R., Souza, S. R. S., Maldonado, J. C. J. C. J. C. J. C., Salleh, N., Martins, R. M., Mendes, E. E., MacDonell, S. G., Maldonado, J. C. J. C. J. C. J. C., Li, D., Wang, Z. Z., Wang, L., … Demner-Fushman, D. (2020). A New Method to Enhance. *Systematic Reviews*, *9*(1).

Khan, M., Hussain, M., Gunasekaran, A., Ajmal, M. M., Helo, P. T., Unger, S., Landis, A., Katircioglu, K., Gooby, R., Helander, M., Drissi, Y., Chowdhary, P., Johnson, M., Yonezawa, T., Logrieco, A. F., Miller, J. D., Eskola, M., Krska, R., Ayalew, A., … Singh, A. (2019). Impact of Occupational Exposure to Chemicals in Life Cycle Assessment: A Novel Characterization Model Based on Measured Concentrations and Labor Hours. *Journal of Cleaner Production*, *11*(1).

Kharimah, W. (2021). DAMPAK PENGUNDURAN DIRI AMERIKA SERIKAT DARI PERJANJIAN JCPOA (JOINT COMPREHENSIVE PLAN OF ACTION). *Frontiers in Neuroscience*, *14*(1).

Khmelkov, A. (2023). INSTITUTIONAL DESIGN OF ANALYTICAL TOOLS OF THE

FINANCIAL CONTROL INSTITUTION. *Social Economics*, *66*. https://doi.org/10.26565/2524-2547-2023-66-08

Kim, T. Y., Gang, J. Y., Oh, H. J., Ahmad, K., JianMing, Z., Rafi, M., Allen, G. G., Travis, T. A., Ramirez, C., کوچکی, س. غ. م. ، و. ع., Kyrillidou, M., Manifesto, L. S., Johnson, L., Smith, R., Willis, H., Levine, A., Haywood, K., Vice-chancellor, S. D., Vice-chancellor, D. V.-S. D., … Briggs, T. (2019). Using Data Visualization and Dashboards to Support the Decision Making of Public Library Branch Managers. *Library Hi Tech*, *38*(1).

Kirkpatrick, A. W., McKee, J. L., McBeth, P. B., Ball, C. G., Laporta, A. T., Broderick, T., Leslie, T., King, D., Wright Beatty, H. E., Keillor, J., Tien, H., de'Angelis, N., Khan, J., Marchegiani, F., Bianchi, G., Aisoni, F., Alberti, D., Ansaloni, L., Biffl, W. L., … MOORI, S. H. (2020). Translating Scientific Evidence into Practice using Digital Medicine and Electronic Patient Reported Outcomes. *Telemedicine and E-Health*, *17*(1).

Koçak, M., & Akçalı, Z. (2024). Development trends and knowledge framework of artificial intelligence (AI) applications in oncology by years: a bibliometric analysis from 1992 to 2022. *Discover Oncology*, *15*(1). https://doi.org/10.1007/s12672-024-01415-0

Kolzow, D. R., Smith, C. C. C., Serrat, O., Dilie, H. M., Zeeshan, S., Ng, S. I., Ho, J. A., Jantan, A. H., Massey, J., Sulak, T., Sriram, R., Dennis, R. S., Bocarnea, M., Hai, T. N., Van, Q. N., Herbert, S. L., So-Jung Kim, Kyoung-Seok Kim, Y.-G. C., Guillaume, Dr. O., Honeycutt, Dr. A., … Ingram, O. C. Jr. (2021). Unit 5 Theories of Leadership. *International Journal of Organizational Leadership*, *1*(1).

Kozinets, R. v., & Gretzel, U. (2024). Netnography evolved: New contexts, scope, procedures and sensibilities. *Annals of Tourism Research*, *104*. https://doi.org/10.1016/j.annals.2023.103693

Kulkarni, R. v., Choudhari, G., Bhavik, M., Bura, R., & Bhosale, V. (2025). GAMIFYING EARLY MATH EDUCATION: AN INNOVATIVE AI GAME FOR CHILDREN. *Journal of Information Technology Education: Innovations in Practice*, *24*. https://doi.org/10.28945/5595

Kumar Kanvaria, V., & Tarance Suraj, M. (2024). The Role Of AI In Mathematics Education: Assessing The Effects Of An "Auto Draw" Webtool On Middle-Level Achievement. *The Online Journal of Distance Education and E-Learning*, *12*(1).

Lausiri, N., & Nahda, K. (2022). Pengaruh Rasio Likuiditas, Leverage Dan Profitabilitas Terhadap Kondisi Financial Distress Dengan Kepemilikan Manajerial Sebagai Variabel Moderasi. *Selekta Manajemen: Jurnal Mahasiswa & Manajemen*, *1*(5).

Lei, P., Duo, E. N., Gao, Y., & Zhu, X. (2025). Research on the experience of utilizing artificial intelligence in nursing learning from the perspective of nursing students – a qualitative study. *BMC Nursing*, *24*(1). https://doi.org/10.1186/s12912-025-03704-7

Leong, W. D., Teng, S. Y., How, B. S., Ngan, S. L., Rahman, A. A., Tan, C. P., Ponnambalam, S. G., Lam, H. L., Yadav, G., Luthra, S., Jakhar, S. K., Mangla, S. K., Rai, D. P., Tortorella, G. L., Narayanamurthy, G., Thurer, M., Dombrowski, U., Wullbrandt, J., Fochler, S., … Gierej, S. (2019). Elimination of waste: creation of a successful Lean colonoscopy program at an academic medical center. *Surgical Endoscopy*.

Li, Z. Q., Shroff, P. K., Venkataraman, R., Zhang, I. X., Gu, F., Lev, B., He, G., Kang, J. K. H. K., Shivdasani, A., Yamada, T., Bhattacharyya, S., Nain, A. S., Lee, K. B. H., Homburg, C., Bucerius, M., Erel, I., Liao, R. C., WEISBACH, M. S., Andrade, G., … Lys, T. Z. (2019). CORPORATE TAKEOVERS AND PRODUCTIVITY - LICHTENBERG,FR. *JOURNAL OF BANKING & FINANCE*, *37*(2).

Lim, W. M. (2025). What Is Qualitative Research? An Overview and Guidelines. *Australasian Marketing Journal*, *33*(2). https://doi.org/10.1177/14413582241264619

Litynska, V. (2025). MODERN METHODS OF CONDUCTING MARKETING RESEARCH. *Social Economics*, *70*. https://doi.org/10.26565/2524-2547-2025-70-06

Longo, L. (2020). Empowering Qualitative Research Methods in Education with Artificial Intelligence. *Advances in Intelligent Systems and Computing*, *1068*. https://doi.org/10.1007/978-3-030-31787-4_1

Luo, J., Zhuo, W., Liu, S., & Xu, B. (2024). The Optimization of Carbon Emission Prediction in Low Carbon Energy Economy under Big Data. *IEEE Access*, *12*. https://doi.org/10.1109/ACCESS.2024.3351468

Luo, X., Tham, Y. C., Daher, M., Bian, Z., Chen, Y., Estill, J., & group, the G. working. (2025). Protocol for developing the reporting guideline for the use of chatbots and other Generative Artificial intelligence tools in MEdical Research (GAMER). *medRxiv*.

Luongo, M., Ponticorvo, M., Ligorio, M. B., Crescenzo, P., & Ritella, G. (2024). Artificial Intelligence to Enhance Qualitative Research: Methodological Reflections on a Pilot Study. *Journal of Educational, Cultural and Psychological Studies*, *2024*(30). https://doi.org/10.7358/ecps-2024-030-luon

Maguey-González, J. A., Nava-Ramírez, M. de J., Gómez-Rosales, S., Ángeles, M. de L., Solís-Cruz, B., Hernández-Patlán, D., Merino-Guzmán, R., Hernández-Velasco, X., Figueroa-Cárdenas, J. de D., Vázquez-Durán, A., Hargis, B. M., Téllez-Isaías, G., Méndez-Albores, A., Angelico, R., Colombo, C., di Iorio, E., Brtnický, M., Fojt, J., Conte, P., … Chorover, J. (2018). Modeling of Soil Cation Exchange Capacity Based on. *Geoderma*, *3*(4).

Marshall, D. T., & Naff, D. B. (2024). The Ethics of Using Artificial Intelligence in Qualitative Research. *Journal of Empirical Research on Human Research Ethics*, *19*(3). https://doi.org/10.1177/15562646241262659

Masso, A., Gerassimenko, J., Kasapoglu, T., & Beilmann, M. (2025). Research ethics committees as knowledge gatekeepers: The impact of emerging technologies on social science research. *Journal of Responsible Technology*, *21*. https://doi.org/10.1016/j.jrt.2025.100112

Mekhovych, K. (2025). TRANSFORMATION OF THE THEORY OF INNOVATIVE MARKETING EFFICIENCY IN THE CONDITIONS OF AN UNDETERMINED FUTURE. *Energy Saving. Power Engineering. Energy Audit.*, *1 (204)*. https://doi.org/10.20998/2313-8890.2025.01.07

Mi, S., Dong, C. Q., & Liu, Y. (2025). Data insights and talent portraits of science and technology talents: Principles and applications. *Studies in Science of Science*, *43*(6).

Ministry of Education. (2018). Education strategic plan 2018 - 2030. In *Accra, Republic of Ghana*.

Morante-Carballo, F., Arcentales-Rosado, M., Caicedo-Potosí, J., & Carrión-Mero, P. (2025). Artificial intelligence applications in hydrological studies and ecological restoration of watersheds: A systematic review. In *Watershed Ecology and the Environment* (Vol. 7). https://doi.org/10.1016/j.wsee.2025.05.004

Morozova, I. A., Smetanin, A. S., & Smetanina, A. I. (2024). Management of Business Competitiveness Based on Artificial Intelligence and Big Data for Its Sustainable Development. *Journal of Modern Competition*, *18*(1). https://doi.org/10.37791/2687-0657-2024-18-1-29-40

Mortaheb, R., & Jankowski, P. (2023). Smart city re-imagined: City planning and GeoAI in the age of big data. *Journal of Urban Management*, *12*(1). https://doi.org/10.1016/j.jum.2022.08.001

Mosleh, G., Hemmati, S., Mohagheghzadeh, A., & Zarshenas, M. M. (2023). Integrating Systems Biology and Artificial Intelligence in Traditional Pharmacy Research: Advancements, Challenges, and Opportunities. *Trends in Pharmaceutical Sciences*, *9*(3).

Mossalanejad, A. (2025). Artificial Intelligence Policymaking and the Power Equation Priorities. *Geopolitics Quarterly*, *21*(3). https://doi.org/10.22034/igq.2025.511096.2023

Muhammad, N. I., Judijanto, L., Sulistiani, I., Syarifuddin, & Syahrin, M. (2023). STRATEGIC CONTENT MARKETING: A MULTI-CHANNEL ANALYSIS OF THE IMPACT ON BRAND AWARENESS, CONSUMER ENGAGEMENT, AND CONVERSION RATES IN THE CONTEXT OF EVOLVING DIGITAL LANDSCAPES. *International Journal of Society Reviews (INJOSER)*.

Muhtar, E. (2022). PENGARUH PENDEKATAN KETRAMPILAN PROSES DALAM MENINGKATKAN HASIL BELAJAR FISIKA PADA KONSEP GETARAN DAN GELOMBANG PADA PESERTA DIDIK KELAS VIII SMP NEGERI 17 HALMAHERA SELATAN. *Jurnal Pembelajaran Dan Sina Fisika*, *7*(1).

Munen, J. (2021). SURVEY OF CONTEMPORARY CHALLENGES FACED IN MARKETING OF LIFE INSURANCE POLICIES IN EMBU COUNTY. *Frontiers in Neuroscience*, *14*(1).

Muñoz-Olano, J. F., Hurtado-Parrado, C., Troll, E. S., Friese, M., Loschelder, D. D., Alblwi, A., McAlaney, J., al Thani, D. A. S., Phalp, K., Ali, R., Elhai, J. D., Sapci, O., Yang, H., Amialchuk, A., Rozgonjuk, D., Montag, C., Bäulke, L., Daumiller, M., Dresel, M., … Cevallos Pozo, G. (2020). A review of web-based support systems for students in higher education. *International Journal of Environmental Research and Public Health*, *10*(1).

Murikah, W., Nthenge, J. K., & Musyoka, F. M. (2024). Bias and ethics of AI systems applied in auditing - A systematic review. *Scientific African*, *25*. https://doi.org/10.1016/j.sciaf.2024.e02281

Naeem, M., Smith, T., & Thomas, L. (2025). Thematic Analysis and Artificial Intelligence: A Step-by-Step Process for Using ChatGPT in Thematic Analysis. *International Journal of Qualitative Methods* , *24*. https://doi.org/10.1177/16094069251333886

Nashwan, A. J., & Abukhadijah, H. (2023). Harnessing Artificial Intelligence for Qualitative and Mixed Methods in Nursing Research. *Cureus*. https://doi.org/10.7759/cureus.48570

Nazari, M., Ansari, M., & Pour, M. J. A. (2025). Bibliographic Analysis and Future Research Directions in the Field of Sustainable and Innovative Business Models. *Scientometrics Research Journal*, *11*(2). https://doi.org/10.22070/rsci.2024.19650.1759

Nedoshytko, I., & Patriak, O. (2022). Electronic Document Management and Its Value for Business. *Digital Platform: Information Technologies in Sociocultural Sphere*, *5*(2). https://doi.org/10.31866/2617-796x.5.2.2022.270142

Nicmanis, M., & Spurrier, H. (2025). Getting Started with Artificial Intelligence Assisted Qualitative Analysis: An Introductory Guide to Qualitative Research Approaches with Exploratory Examples from Reflexive Content Analysis. *International Journal of Qualitative Methods* , *24*. https://doi.org/10.1177/16094069251354863

Nyomba, Muh. A. (2021). Faktor Yang Berhubungan Dengan Kejadian Infeksi Saluran Pernapasan Akut (ISPA) Pada Balita di Sekitar Wilayah Tempat Pembuangan Akhir (TPA) Sampah Antang Kota Makassar Tahun 2021. *Factors Associated of The Incidence of Acute Respiratory Infection (ARI) in. Universitas Hasanuddin.*, *9*.

Olawade, D. B., Omeni, D., Gore, M. N., & Hadi, M. (2025). Enhancing qualitative research through virtual focus groups and artificial intelligence: A review. *International Journal of Medical Informatics*, *203*. https://doi.org/10.1016/j.ijmedinf.2025.106004

Orea-Giner, A., Fusté-Forné, F., & Soliman, M. (2025). How do tourists perceive green customer-love service in restaurants? A qualitative exploration of ai and human collaboration. *International Journal of Hospitality Management*, *131*. https://doi.org/10.1016/j.ijhm.2025.104300

Owoahene Acheampong, I., & Nyaaba, M. (2024). Review of Qualitative Research in the Era of Generative Artificial Intelligence. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4686920

Palazzo, F., Zambetta, G., & Palazzo, S. (2025). Artificial Intelligence at the Crossroads of Engineering and Innovation. *Computing&AI Connect*, *2*(1). https://doi.org/10.69709/caic.2025.141254

Palazzo, S., Palazzo, F., & Zambetta, G. (2025). The Role of Artificial Intelligence in Advanced Engineering: Current Trends and Future Prospects. *Journal of Intelligent Communication*, *4*(1). https://doi.org/10.54963/jic.v4i1.959

Pan, C., Bowen, W., Peng, Z., & Shao, L. (2024). Biological network-based exploration of the principles of traditional Chinese medicine. *Kexue Tongbao/Chinese Science Bulletin*, *69*(1). https://doi.org/10.1360/TB-2023-0730

Papaneophytou, C., & Nicolaou, S. A. (2025). Promoting Critical Thinking in Biological Sciences in the Era of Artificial Intelligence: The Role of Higher Education. In *Trends in Higher Education* (Vol. 4, Issue 2). https://doi.org/10.3390/higheredu4020024

Pepic, I., Feldt, R., Ljungström, L., Torkar, R., Dalevi, D., Maurin Söderholm, H., Andersson, L. M., Axelson-Fisk, M., Bohm, K., Sjöqvist, B. A., & Candefjord, S. (2021). Early detection of sepsis using artificial intelligence: a scoping review protocol. *Systematic Reviews*, *10*(1). https://doi.org/10.1186/s13643-020-01561-w

Perano, M., Cammarano, A., Varriale, V., del Regno, C., Michelino, F., & Caputo, M. (2023). Embracing supply chain digitalization and unphysicalization to enhance supply chain performance: a conceptual framework. *International Journal of Physical Distribution and Logistics Management*, *53*(5–6). https://doi.org/10.1108/IJPDLM-06-2022-0201

Perdana, R., Yulianti, D., & Bertiliya, W. A. (2025). Enhancing scientific literacy through AI-integrated Inquiry Social Complexity module in primary education. *Journal of Educational Management and Instruction (JEMIN)*, *5*(2). https://doi.org/10.22515/jemin.v5i2.10891

Pérez-Santín, E., De-La-fuente-valentín, L., García, M. G., Bravo, K. A. S., Hernández, F. C. L., & Sánchez, J. I. L. (2023). Applicability domains of neural networks for toxicity prediction. In *AIMS Mathematics* (Vol. 8, Issue 11). https://doi.org/10.3934/math.20231426

Permenkes. (2022). Peraturan Menteri Kesehatan Nomor 24 Tahun 2022 Tentang Rekam Medis. Jakarta: Menteri Kesehatan Republik Indonesia. *Science*, *7*(1).

Posunko, A., Chavarha, M., Bashkirova, L., Vasylyuk-Zaitseva, S., & Abramchuk, O. (2025). Machine Learning in Medicine: A New Paradigm for Diagnosis and Treatment. *Futurity Medicine*, *4*(3). https://doi.org/10.57125/fem.2025.09.30.01

Pratt, M. G. (2025). On the Evolution of Qualitative Methods in Organizational Research. In *Annual Review of Organizational Psychology and Organizational Behavior* (Vol. 12). https://doi.org/10.1146/annurev-orgpsych-111722-032953

Putri, S. H. (2022). Pengaruh Brand Image terhadap keputusan pembelian produk skincare. *Science*, *7*(1).

Pysmennyi, V., & Rudan, V. (2024). Methods of financial management in local communities under the conditions of the russian-Ukrainian war. *Economic Analysis, 34(3)*. https://doi.org/10.35774/econa2024.03.351

RAHMAN, G. O. H. S. E. Y. T. A. (2020). DETERMINANTS OF FOOD DELIVERY PLATFORMS THROUGH CUSTOMER LOYALTY TOWARD SUSTAINABILITY OF BUSINESS. *Human Relations*, *3*(1).

Rahmawati. (2025). EFEKTIVITAS PELAYANAN DUKCAPIL ONLINE PADA DINAS KEPENDUDUKAN DAN CATATAN SIPIL DI KOTA MAKASSAR. *SKRIPSI UNIVERSITAS NEGERI MAKASSAR*, *11*(1).

Rattan, P., Penrice, D. D., & Simonetto, D. A. (2022). Artificial Intelligence and Machine Learning: What You Always Wanted to Know but Were Afraid to Ask. In *Gastro Hep Advances* (Vol. 1, Issue 1). https://doi.org/10.1016/j.gastha.2021.11.001

Rejter, A. (2024). STUDIES IN ONOMASTICS AND THE NEW HUMANITIES. *Onomastica*, *68*. https://doi.org/10.17651/ONOMAST.68.2

Rincon Soto, I. B., Soledispa-Cañarte, B. J., Sumba-Bustamante, R. V., del Carmen Burbano-Gómez, Z., & Jiménez-Granizo, F. P. (2023). METHODOLOGICAL

APPROACHES AND PARADIGMS IN SCIENTIFIC AND TECHNOLOGICAL RESEARCH: A BIBLIOMETRIC REVIEW. *Bibliotecas, Anales de Investigacion*, *19*(1).

Rincon, I. B., Soledispa-Cañarte, B. J., Sumba-Bustamante, R. V., del Carmen Burbano-Gómez, Z., & Jiménez-Granizo, F. P. (2023). Methodological Approaches and Paradigms in Scientific and Technological Research: a Bibliometric Review | Abordajes Metodológicos Y Paradigmas En La Investigación Científica Y Tecnológica: Una Revisión Bibliométrica. *Bibliotecas, Anales de Investigacion*, *19*(1).

Roberts, J., Baker, M., & Andrew, J. (2024). Artificial intelligence and qualitative research: The promise and perils of large language model (LLM) 'assistance.' *Critical Perspectives on Accounting*, *99*. https://doi.org/10.1016/j.cpa.2024.102722

Rodríguez Flores, E. A., & Sánchez Trujillo, M. de los Á. (2025). Scientific research and artificial intelligence in graduate students. A qualitative analysis. *European Public and Social Innovation Review*, *10*. https://doi.org/10.31637/epsir-2025-1049

Rodríguez Flores, E. A., Garcés Giraldo, L. F., Valencia, J., & Valencia-Arias, A. (2025). Research trends in the use of artificial intelligence techniques in scientific research. *Revista Venezolana de Gerencia*, *30*(109). https://doi.org/10.52080/rvgluz.30.109.23

Ronzon, T., Gurria, P., Carus, M., Cingiz, K., El-Meligi, A., Hark, N., Iost, S., M'barek, R., Philippidis, G., van Leeuwen, M., Wesseler, J., Medina-Lozano, I., Grimplet, J., Díaz, A., Tejedor-Calvo, E., Marco, P., Fischer, M., Creydt, M., Sánchez-Hernández, E., … Miras Ávalos, J. M. (2025). Understanding Material Recovery Facility Policies in the Philippines: A Scoping Review. *Sustainability (Switzerland)*, *2*(1).

Rosenberg, J. (2024). Automated Transcription of Interviews in Qualitative Research Using Artificial Intelligence A Simple Guide. *Journal of Surgery Research and Practice*. https://doi.org/10.46889/jsrp.2024/5204

Saheb, T., Sidaoui, M., & Schmarzo, B. (2024). Convergence of artificial intelligence with social media: A bibliometric & qualitative analysis. In *Telematics and Informatics Reports* (Vol. 14). https://doi.org/10.1016/j.teler.2024.100146

Sampaio, R. C., Chagas, V., Sanchez, C. S., Gonçalves, J., Borges, T., Alison, M. B., Tigrinho, C. S., de Souza, J. R., & Paz, F. S. (2024). AN ARTIFICIAL INTELLIGENCE (AI)-ASSISTED SCOPING REVIEW OF EMERGING USES OF AI IN QUALITATIVE RESEARCH AND ITS ETHICAL CONSIDERATIONS. *Revista Pesquisa Qualitativa*, *12*(30). https://doi.org/10.33361/RPQ.2024.v.12.n.30.729

Sani, A., Maharani, A. R., & Nazimah. (2025). Peran Komunitas Dosen Dalam Peningkatan Kualitas Pembelajaran Di Perguruan Tinggi. *Journal Genta Mulia*, *15*(1).

Saragih, B., & , Ika Karyati, D. S. (2010). Pengaruh Pewarna Ekstrak Cair Alami Bawang Tiwai (Eleutherine americana Merr) Terhadap Mutu Selai Kulit Pisang Kepok (Musa paradisiaca Linn). *Jurnal Tenknologi Pertanian*, *6*(2).

Sart, G. (2017). HOW DOES THE MAKER MOVEMENT AFFECT THE SCHOOL ADMINISTRATIONS' ROLES ON IMPROVING THE STEM EDUCATION? *INTED2017 Proceedings*, *1*. https://doi.org/10.21125/inted.2017.2452

Saurbayev, R. Zh., Zhetpisbay, A. K., & Yerekhanova, F. T. (2024). Philological Studies: Computer Aspects of Stylometry Automation. *Iasaýı Ýnıversıtetiniń Habarshysy*, *133*(3). https://doi.org/10.47526/2024-3/2664-0686.72

Schmidt, D. A., Alboloushi, B., Thomas, A., & Magalhaes, R. (2025). Integrating artificial intelligence in higher education: perceptions, challenges, and strategies for academic innovation. *Computers and Education Open*, *9*. https://doi.org/10.1016/j.caeo.2025.100274

Schneider, R., Angwin, C., Jenkinson, C., Jones, A. S. K., Jennison, C., Henley, W., Farmer, A., Sattar, N., Holman, R. R., Pearson, E., Shields, B., al., et, Blackwell, K., Donskih, R., Jones, C. M., Nixon, A., Vidal, M. J., Nakov, R., Singh, P., … Ewel, C. (2021). Effectiveness of olive oil for the prevention of pressure ulcers caused in immobilized patients within the scope of primary health care: study protocol for a randomized controlled trial. *Trials*, *21*(1).

Sentioningtias, I. (2025). Pemanfaatan Aplikasi Math Kids Untuk Mengembangkan Berhitung Anak Usia Dini Di TK Islam Nrhasanah Bandar Lampung. In *Sustainability (Switzerland)* (Vol. 11, Issue 1).

Shakil, M., Ali, M., Illahi, T., & Ahmed, F. (2025). FinTech 5.0 and the Future of Global Finance: Harnessing Artificial Intelligence, Blockchain, and Big Data to Build Inclusive, Sustainable, and Resilient Financial Ecosystems. *Inverge Journal of Social Sciences*, *4*(4).

https://doi.org/10.63544/ijss.v4i4.179

Shimray, S. R., & Subaveerapandiyan, A. (2025). AI and workforce dynamics: a bibliometric analysis of job creation, displacement and reskilling. *Global Knowledge, Memory and Communication*. https://doi.org/10.1108/GKMC-12-2024-0824

Shivarama Krishna, K., & Sathish Kumar, K. (2015). A review on hybrid renewable. *Renewable and Sustainable Energy Reviews*, *52*.

Sipa, M. (2023). Trends and challenges in talent management – mind map. *Scientific Papers of Silesian University of Technology Organization and Management Series*, *2023*(164). https://doi.org/10.29119/1641-3466.2022.164.34

Skoczkowski, T., Bielecki, S., Wołowicz, M., & Węglarz, A. (2025). Redefining Energy Management for Carbon-Neutral Supply Chains in Energy-Intensive Industries: An EU Perspective. *Energies*, *18*(15). https://doi.org/10.3390/en18153932

Skrypnyk, S. v. (2025). Key Concepts and Models of Strategic Innovation Management in Ukraine: Modern Approaches. *The Problems of Economy*, *2*(64). https://doi.org/10.32983/2222-0712-2025-2-77-85

Skyrme, D., & Amidon, D. (2010). Business research methods. *Journal of Knowledge Management*, *14*(1).

Soni, A., Jain, A., & Kaur, B. (2022). A bibliometric analysis of effect of artificial intelligence on customer satisfaction in banking sector. *International Research Journal of Modernization in Engineering*, *4*(3).

Sovia Lenny. (2006). RINGKASAN Puding merah (Gruptophyllum pictum L Griff). *Jurnal Farmasi Galenika (Galenika Journal of Pharmacy) (e-Journal)*, *1*(1).

Spivakovsky, O., Omelchuk, S., Kobets, V., Poltoratskyi, M., Malchykova, D., & Lemeshchuk, O. (2025). INSTITUTIONAL REGULATION OF USING AI-BASED SERVICES THROUGH THE KANO MODEL: THE CASE OF KHERSON STATE UNIVERSITY. *Information Technologies and Learning Tools*, *108*(4). https://doi.org/10.33407/itlt.v108i4.6269

Sposato, M. (2025). Introducing AI to qualitative research: Theoretical and practical implications. In *Exploring Qualitative Research in Business: Approaches, Strategies, And Methods*. https://doi.org/10.4018/979-8-3693-7698-0.ch010

Starrett, A., Yow, J., Lotter, C., Irvin, M. J., Adams, P., Guillemot, F., Lacroix, F., Nocus, I., Assen, J. H. E., Otting, H., Jin, X., Tigelaar, D., van der Want, A., Admiraal, W., Peschl, H., Deng, C., Larson, N., Kuno, C. B., Hein, S., … León, J. (2022). Editorial Board. *Teaching and Teacher Education*, *90*(3).

Sydorovych, O., Bukovskyi, M., & Yaroshovych, A. (2024). DIGITAL TOOLS OF TAX ADMINISTRATION: CURRENT TRENDS AND PERSPECTIVES. *INNOVATIVE ECONOMY*. https://doi.org/10.37332/2309-1533.2024.4.26

T. P., & . V. G. (2025). INVESTIGATION OF DEEP LEARNING FRAMEWORK TO ACCELERATE AI FOR IMAGE PROCESSING APPLICATIONS. *International Journal For Multidisciplinary Research*, *7*(3). https://doi.org/10.36948/ijfmr.2025.v07i03.45548

Taha Kandil, T. (2025). Artificial intelligence (AI) and alleviating supply chain bullwhip effects: social network analysis-based review. *Journal of Global Operations and Strategic Sourcing*, *18*(1). https://doi.org/10.1108/JGOSS-04-2023-0038

Tàpias, M. (2019). Clothing as Solar Radiation Protection. *Autex Research Journal*, *2*(1).

Tierney, A., Peasey, P., & Gould, J. (2025). Student perceptions on the impact of AI on their teaching and learning experiences in higher education. *Research and Practice in Technology Enhanced Learning*, *20*. https://doi.org/10.58459/rptel.2025.20005

Tong, L., Niu, Y., Zhou, L., Jin, S., Wang, Y., & Xiao, Q. (2025). Perceptions and experiences of generative artificial intelligence training to support research for Chinese nurses: A qualitative focus group study. *International Journal of Nursing Sciences*, *12*(3). https://doi.org/10.1016/j.ijnss.2025.04.010

Torres-Sainz, R., de-Zayas-Pérez, M. R., Trinchet-Varela, C. A., Pérez-Vallejo, L. M., & Rodríguez, R. P. (2025). INTELLIGENT PREDICTIVE MAINTENANCE: A BIBLIOMETRIC APPROACH THROUGH THE SCOPUS DATABASE. *Bibliotecas, Anales de Investigacion*, *21*(1).

Trejos-Gil, C. A., & Gómez-Monsalve, W. D. (2024). Artificial Intelligence in Media and Journalism. Systematic Review on Spain and Latin America in Scopus and Web of Science Databases (2018-2022). In *Palabra Clave* (Vol. 27, Issue 4). https://doi.org/10.5294/pacla.2024.27.4.1

Tschisgale, P., Wulff, P., & Kubsch, M. (2023). Integrating artificial intelligence-based methods into qualitative research in physics education research: A case for computational grounded theory. *Physical Review Physics Education Research*, *19*(2). https://doi.org/10.1103/PhysRevPhysEducRes.19.020123

Tursunbayeva, A., Chalutz-Ben Gal, H., Santos Fonseca, S. S., Benito, P. A., Piña Ramírez, C., Rubeis, G., Filz, M.-A., Bosse, J. P., Herrmann, C., Cavalieri, A., Reis, J., Amorim, M., Santarsiero, F., Carlucci, D., Schiuma, G., Sukums, F., Wamala-Larsson, C. W., Kisenge, R., Ravi, S. J., … Kazancoglu, Y. (2024). A Spotlight on Environmental Sustainability in View of the European Green Deal. *Sustainability (Switzerland)*, *16*(1).

UU Republik Indonesia, Munawir, M., Nasional, B. S., Susanta, G., Jatimnet.com, MajaMojokerto.net, Radarmojokerto.jawapos.com, Wijayakusuma, D. M. S., Nahman, A., Godfrey, L., Jacobsen, R., Buysse, J., Gellynck, X., Bayu, D., Depkes, R., Soekidjo Notoatmodjo, Boulanger, L., Ismoyo, I. H., Tarigan, R., … Sanyal, S. (2022). PENENTUAN ALTERNATIF LOKASI TEMPAT PEMBUANGAN AKHIR (TPA) SAMPAH DI KABUPATEN SIDOARJO. *Energies*, *15*(1).

van Merrienboer, J. J. G., Paas, F., Traxler, A., Brewe, E., Hazari, Z., Cass, C., Beattie, C., Kelly, A. M., Wilcox, B. R., Lewandowski, H. J., Bao, L., Xiao, Y., Koenig, K., Han, J., Eisinga, R., Grotenhuis, M., Pelzer, B., Cronbach, L. J., Butts, M. M., … VanLehn, K. A. (2016). Evaluation of Colorado Learning Attitudes about Science Survey. *Physical Review Physics Education Research*, *12*(2).

Venter, I. M., Blignaut, R. J., Cranfield, D. J., Tick, A., & Achi, S. el. (2025). AI versus tradition: shaping the future of higher education. *Journal of Applied Research in Higher Education*, *17*(7). https://doi.org/10.1108/JARHE-12-2024-0702

Verifier, S. D., & Drive, A. H. (2015). Simulink ® Verification and Validation ™ Reference. *ReVision*.

Vicente, N. R. (2010). Sistemas agroflorestais sucessionais como estratégia de uso e conservação de recursos florestais em zonas ripárias da Microbacia Arroio Primeiro de Janeiro, Anchieta-SC. *Journal of Mental Science*, *7*(37).

Vollmar, H. C., Kramer, U., Müller, H., Griemmert, M., Noelle, G., & Schrappe, M. (2017).

Digitale Gesundheitsanwendungen - Rahmenbedingungen zur Nutzung in Versorgung, Strukturentwicklung und Wissenschaft - Positionspapier der AG Digital Health des DNVF. *Gesundheitswesen*, *79*(12). https://doi.org/10.1055/s-0043-122233

VOLOSHYN, R. (2024). COMMODITY EXPERTISE IN E-COMMERCE. *"ECONOMY. FINANCES. MANAGEMENT: Topical Issues of Science and Practical Activity," 4 (70)*. https://doi.org/10.37128/2411-4413-2024-4-4

Wang, C., Nan, J., Kang, Y., Xu, X., Li, T., & Zhu, Z. (2025). Design of a comprehensive teaching experiment for mechanoluminescent human biomechanical detection. *Experimental Technology and Management*, *42*(9). https://doi.org/10.16791/j.cnki.sjg.2025.09.028

Wang, S., & Kim, R. Y. (2025). Exploring the role of generative AI in enhancing critical thinking in multicultural classrooms: a case study from korea. *Multicultural Education Review*, *17*(1). https://doi.org/10.1080/2005615X.2025.2467787

Wang, Y., Kang, M., Liu, Y., Li, J., Xue, K., Wang, X., Du, J., Tian, Y., Ni, Q., & Wang, F. Y. (2023). Can Digital Intelligence and Cyber-Physical-Social Systems Achieve Global Food Security and Sustainability? *IEEE/CAA Journal of Automatica Sinica*, *10*(11). https://doi.org/10.1109/JAS.2023.123951

Wang, Z., Wang, D., Xu, Y., Zhou, L., & Zhou, Y. (2025). Intelligent Computing Social Modeling and Methodological Innovations in Political Science in the Era of Large Language Models. *Journal of Chinese Political Science*. https://doi.org/10.1007/s11366-025-09917-6

Warhurst, A. (2002). Sustainability indicators and sustainability performance management. *Sustainability Indicators and Sustainability Performance Management*, *43*(43).

Waryono, T. (2019). RINGKASAN Puding merah (Gruptophyllum pictum L Griff). *Jurnal Farmasi Galenika (Galenika Journal of Pharmacy) (e-Journal)*, *1*(1).

Wijanarti, H. L., & Anisyah, T. D. A. (2022). Hubungan Antara Kualitas Tidur, Beban Kerja Fisik Terhadap Perasaan Kelelahan Kerja Pada Perawat Rawat Inap Kelas 3 Di Rs Pku Muhammadiyah Gamping. *Jurnal Lentera Kesehatan Masyarakat*, *1*(1).

Williamson, B. (2020). New digital laboratories of experimental knowledge production: Artificial intelligence and education research. *London Review of Education*, *18*(2). https://doi.org/10.14324/LRE.18.2.05

Yang, Y., & Ma, L. (2025). Artificial intelligence in qualitative analysis: a practical guide and reflections based on results from using GPT to analyze interview data in a substance use program. *Quality and Quantity*, *59*(3). https://doi.org/10.1007/s11135-025-02066-1

Yang, Y., & Siau, K. (2018). A Qualitative Research on Marketing and Sales in the Artificial Intelligence Age. *Midwest Association for Information Systems ConferenceAssociation for Information Systems Conference*, *5*(17-18 May 2018).

Yıldırım, H. M., Güvenç, A., & Can, B. (2025). Reporting Quality in Qualitative Research Using Artificial Intelligence and the COREQ Checklist: A Systematic Review of Tourism Literature. *International Journal of Qualitative Methods* , *24*. https://doi.org/10.1177/16094069251351487

Yubo, S., Ramayah, T., & Hongmei, L. (2022). Research on the Impact of Intelligent Manufacturing Technology on Business Model Based on CNKI's Evidence. *Global Business and Management Research, Suppl. Special Issue*, *14*(3s).

Zavalii, T., & Lehenchuk, S. (2025). DEVELOPMENT OF RESEARCH METHODS IN MARKETING IN THE CONDITIONS OF FORMING THE DIGITAL MARKETING PARADIGM. *Economic Scope*, *204*. https://doi.org/10.30838/ep.204.101-109

Zhang, B., Liu, Q., Li, X., Liu, L., Yang, B., Letu, H., Gao, L., Zhang, W., Zhang, H., Bian, Z., Qi, M., Chen, C., & Shang, H. (2025). The core concepts and fundamental issues of remote sensing science. *National Remote Sensing Bulletin*, *29*(1). https://doi.org/10.11834/jrs.20244503

Zhao, R., Yang, X., Zhang, D., Li, J., Huang, Y., Xian, G., Kou, Y., & Sun, T. (2025). Artificial Intelligence for Agricultural Science (AI4AS): Key Elements, Challenges and Pathways. *Smart Agriculture*, *7*(3). https://doi.org/10.12133/j.smartag.SA202502019

Zhao, T., Zhou, Y., Shi, R., Cao, Z., & Ren, Z. (2024). GWAI: Artificial intelligence platform for enhanced gravitational wave data analysis. *SoftwareX*, *28*. https://doi.org/10.1016/j.softx.2024.101930

Zhao, X., Wang, X., Research, R. C.-A. in R. (2025). Research on optimizing water resource management and improving crop yield using big data and artificial intelligence. *Jstage.Jst.Go.Jp*, *5*(2).

Zharov, A. M. (2024). THE DYNAMICS OF KNOWLEDGE:AN EVOLUTIONARY ANALYSIS OF SCIENTIFIC DEVELOPMENT THEORIES. *Vestnik Tomskogo*

Gosudarstvennogo Universiteta. Filosofiya, Sotsiologiya, Politologiya, 80. https://doi.org/10.17223/1998863x/80/3

Zheldibayeva, R. (2025). THE IMPACT OF AI AND PEER FEEDBACK ON RESEARCH WRITING SKILLS: A STUDY USING THE CGSCHOLAR PLATFORM AMONG KAZAKHSTANI SCHOLARS. *Scientific Journal of Astana IT University*. https://doi.org/10.37943/21tbax7694

Zheng, Z., He, Z., Khattab, O., Rampal, N., Zaharia, M. A., Borgs, C., Chayes, J. T., & Yaghi, O. M. (2023). Image and Data Mining in Reticular Chemistry Using GPT-4V. *ArXiv*.

Zozulov, O., & Boboshko, Y. (2025). Classification of product development and market entry models. *Marketing and Digital Technologies*, *9*(2). https://doi.org/10.15276/mdt.9.2.2025.4

Zuo, H. (2025). Human-Environment Relationship Research from a Multidisciplinary Integration Perspective in Geography. *Frontiers in Sustainable Development*, *5*(8). https://doi.org/10.54691/64t83m16

Игнатьев, М. Е., & Степанов, В. В. (2025). Методология диссертационных исследований в юридических науках: отдельные аспекты и проблемы. *СОВРЕМЕННОЕ ПРАВО*, *7*. https://doi.org/10.25799/ni.2025.32.84.001

Ксения Юрьевна, Б., & Елена Иосифовна, Л. (2024). Identification and regulation of the macro-regional ecosystem' human resources potential development trends: global and national experience. *STATE AND MUNICIPAL MANAGEMENT SCHOLAR NOTES*, *1*(4). https://doi.org/10.22394/2079-1690-2024-1-4-28-33

This edition of "*Scientific research methodology applied to artificial intelligence and data science: General approach"*, was completed in the city of Colonia del Sacramento in the Eastern Republic of Uruguay on October 3, 2025

EST. 2021 EMC
EDITORIAL MAR CARIBE

SCIENTIFIC RESEARCH METHODOLOGY APPLIED TO ARTIFICIAL INTELLIGENCE AND DATA SCIENCE:

GENERAL APPROACH

JOSEFINA ARIMATEA GARCÍA CRUZ
MARIA MAURA CAMAC TIZA
CARLOS ALBERTO ASIÁN QUIÑONES
JOSÉ CÉSAR PIEDRA ISUSQUI
LUIS ALBERTO JUSCAMAYTA RAMÍREZ
JACKELINE ROXANA HUAMAN FERNANDEZ
JUDITH DEL ROSARIO ROSAS CAJAHUANCA

ISBN 978-9915-698-43-4