Vol. 2, No.4 , pp.65-81 , 2026

Operational Data-Centric AI: A Review of Mesh Governance, Synthetic Tables, Efficient Transformers, Distribution Shift, and Decision Automation

Adriana Lopez[1], Miguel Torres[2,*], Elena Carrasco[3]

[1] Department of Computer Science, University of Cadiz, Cadiz 11003, Spain

[2] Department of Computer Engineering and Systems, University of La Laguna, San Cristobal de La Laguna 38200, Spain

[3] Department of Computer Systems and Telematics Engineering, University of Extremadura, Caceres 10003, Spain

* *miguel.torres@ull.edu.es*

Article Information	
Received	14 July 2024
Accepted	25 November 2024
DOI	https://doi.org/10.63646/datamind.2024.020406

Abstract

The 2024 DATAMIND corpus moves from foundational database-centered artificial intelligence toward operational data-centric AI. This review analyzes all DATAMIND articles published in 2024 and situates them within eighty DOI-bearing references on data management, synthetic tabular data, efficient transformer architectures, digital twins, concept drift, uncertainty, portfolio optimization, and reinforcement learning. The review uses a structured coding matrix to compare each article by lifecycle stage, architectural object, risk category, decision setting, and evaluation emphasis. The analysis shows that DATAMIND's 2024 articles collectively frame operational AI as a control system composed of data products, generative data, scalable models, monitoring routines, and decision rules. DataMesh research emphasizes distributed ownership and accountability; synthetic tabular data research highlights privacy and utility trade-offs; efficient transformer work reframes computation as a governance resource; smart manufacturing connects AI to cyber-physical evidence streams; distribution shift makes monitoring central; and actor-critic portfolio optimization brings decision value into focus. Two high-resolution grayscale figures and three tables summarize the journal corpus, coding rubric, comparative strengths, and research agenda. The article concludes that data-centric AI

© 2024 The Author(s). Published by DATAMIND under CC BY 4.0

should be assessed by reproducibility, scalability, robustness, governance, and decision value rather than by model accuracy alone.

Keywords: ***Data-centric AI; data mesh; synthetic tabular data; efficient transformers; distribution shift; digital twins; reinforcement learning***

1. Introduction

The 2024 DATAMIND corpus captures a shift from foundational data-centered AI to operational data-centric AI. Across the year, the journal examined decentralized data architecture, synthetic tabular generation, efficient transformers, smart manufacturing, distribution shift, and actor-critic portfolio optimization. These topics are diverse on the surface, but they share a common systems problem: once AI models move from experimental notebooks into production, data ownership, computational constraints, monitoring, and action policies become inseparable. This review therefore asks how the 2024 articles collectively describe the transition from model evaluation to operational control.

The method used here combines journal-level synthesis with structured coding. Each DATAMIND article was coded for data lifecycle stage, architectural object, risk category, decision setting, and evaluation emphasis. The external literature was then selected to anchor those categories in broader research on feature pipelines, data mesh governance, synthetic data, efficient transformers, digital twins, concept drift, and reinforcement learning. The resulting analysis is presented as a practical review for researchers who need to connect data architecture to deployed AI governance rather than treat them as separate literatures.

The review gives special attention to the placement of evidence. Tables summarize the yearly DATAMIND corpus, coding categories, and research agenda, while two figures visualize the operational control stack and comparative strength matrix. The figures are not decorative: they translate the review's main claim into analyzable objects. Data-centric AI in 2024 was not a single technique. It was a layered control problem in which data products, synthetic data, efficient sequence models, monitoring systems, and automated decisions had to be aligned under measurable governance constraints.

Table 1. Same-year DATAMIND articles included in the review.

Issue	DATAMIND article reviewed	Primary role in this review
2(1)	DataMesh for data-intensive AI	Decentralized governance and data-product architecture
2(2)	Synthetic Tabular Data	Generative data, privacy, and risk controls
2(3)	Efficient Transformer Architectures	Long sequence mining and computational efficiency
2(3)	Database-Centered AI for Smart Manufacturing	Digital twins and industrial knowledge graphs
2(4)	Distribution Shift in Deployed ML Systems	Monitoring, adaptation, and governance
2(4)	Actor-Critic RL for Portfolio Optimization	Decision automation under volatility

2. Journal Corpus, Coding Design, and Review Logic

The same-year DATAMIND corpus for 2024 is broader than the 2023 corpus and more operational in tone. DataMesh, synthetic tabular data, efficient transformers, database-centered smart manufacturing, distribution shift, and actor-critic portfolio optimization jointly define a production-facing research agenda. The articles show that the main challenge is no longer whether AI can learn from data, but whether data-intensive AI can be organized, scaled, monitored, and connected to

ISSN: 3071-5601 | © 2024 The Author(s). Published by DATAMIND under CC BY 4.0

consequential decisions under changing conditions (Vasquez et al., 2024; Petrova et al., 2024; Al-Haddad et al., 2024; Zhao et al., 2024; Restrepo and Hoffmann, 2024; Chen et al., 2024).

A useful way to read the 2024 corpus is through the data lifecycle. DataMesh and feature-oriented architectures focus on ownership, lineage, and distributed accountability at the ingestion and feature stages. Synthetic tabular data addresses scarcity, privacy, and simulation at the data-generation stage. Efficient transformers address computational limits at the model stage. Smart manufacturing and digital twins connect AI to industrial data streams. Distribution shift and actor-critic portfolios move the discussion to deployment, adaptation, and decision value.

The journal's 2024 articles also reveal a governance pattern. Every operational AI system must coordinate at least three forms of control: data control, model control, and use control. Data control concerns who owns data products, how schemas are versioned, and whether synthetic records preserve privacy. Model control concerns architecture, efficiency, robustness, and monitoring. Use control concerns whether the model's output drives action in a factory, market, or organizational workflow. Figure 1 visualizes these control allocations across the operational stack.

Table 1 lists the same-year DATAMIND articles and their review functions. The table is placed early because it establishes the internal evidence base before external literature is introduced. This structure also reflects the review's purpose: to write for the journal as a continuing intellectual project, not only to summarize a general literature. The 2024 volume suggests that DATAMIND is moving from foundational methods to systems integration.

The DataMesh article is central because it frames data as a federated product rather than a centralized warehouse asset. In AI systems, this distinction matters because model performance depends on the quality, discoverability, and accountability of upstream data products. A decentralized architecture can improve domain ownership, but it also raises coordination costs. Therefore, the question for 2024 is not whether centralized or decentralized data architectures are universally better. The question is how architecture distributes responsibility for AI outcomes.

Synthetic tabular data introduces a second operational tension. Synthetic data can reduce privacy risk, supplement rare cases, and support simulation, but it can also replicate bias, leak membership signals, or generate unrealistic dependency structures. The DATAMIND review therefore fits into a broader debate about whether generative data should be judged by statistical similarity, downstream utility, privacy leakage, or regulatory acceptability. In practice, all four criteria are needed because deployment risks are multi-dimensional.

3. Thematic Findings from the DATAMIND Corpus

Efficient transformer research adds a third dimension: computational governance. Long-sequence models, sparse attention, linear attention, memory-efficient kernels, and quantization all reduce the cost of using large models. However, efficiency is not only an engineering benefit. It changes who can deploy advanced models, what latency targets are feasible, and how many monitoring checks can be afforded. This makes efficiency a governance issue because resource constraints shape the observable behavior of deployed systems.

The smart manufacturing article extends data-centric AI into cyber-physical systems. Digital twins, predictive maintenance models, and industrial knowledge graphs require tight alignment among sensors, historical records, simulation models, and operational decisions. In such contexts, a false prediction is not merely a statistical error; it may interrupt a production line, misdirect maintenance teams, or create safety risk. The 2024 corpus therefore links database-centered AI to industrial accountability.

Distribution shift is the monitoring problem that holds the corpus together. DataMesh can fail when local data products change without notice. Synthetic data can fail when generated distributions hide rare but important edge cases. Efficient transformers can fail when long contexts include stale or irrelevant records. Smart manufacturing and portfolio optimization can fail when the environment changes faster than retraining cycles. The distribution-shift article makes these failures visible by requiring deployed AI to be monitored as an evolving system.

Actor-critic portfolio optimization adds a decision-theoretic endpoint. Reinforcement learning systems are not content to predict; they recommend actions. This makes evaluation more complex because performance depends on reward design, risk constraints, non-stationarity, and transaction or intervention costs. The DATAMIND article on market volatility demonstrates why data-centric AI reviews must include decision analysis. A model that is accurate but poorly aligned with decision constraints can destroy rather than create value.

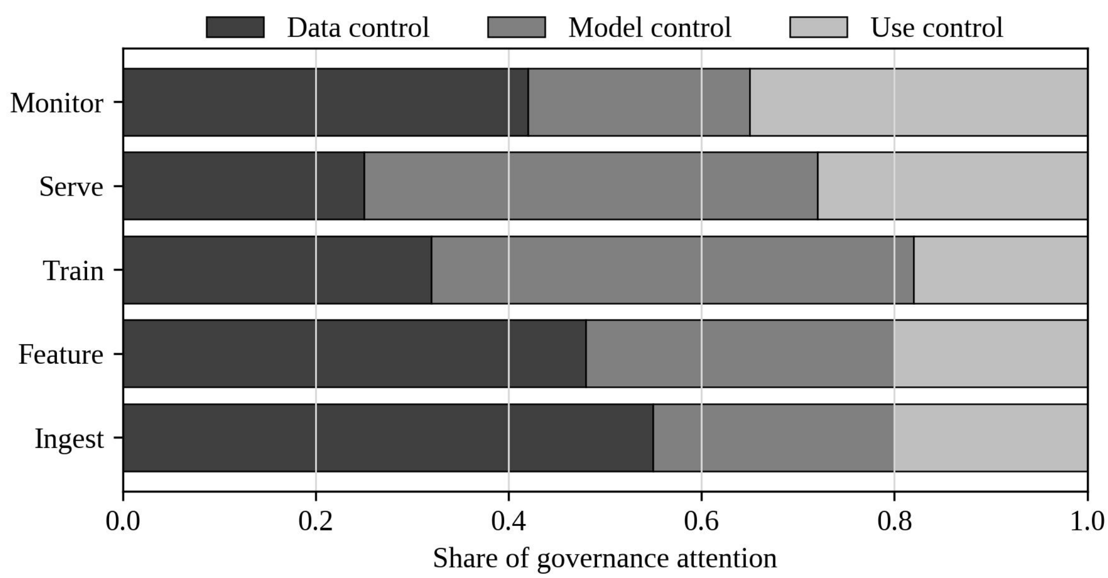

Figure 1. Operational control allocation in the 2024 data-centric AI stack.

The figure should be read as an interpretive summary rather than as a claim about exact citation distance. It translates the article coding into a visual representation of how the journal's yearly topics reinforce one another. Nodes or rows with stronger connections indicate themes that repeatedly appear across the corpus and its DOI-linked supporting literature. The main value of the figure is that it makes the review's organizing logic visible before the more detailed discussion continues.

The comparative coding in Figure 2 shows that 2024 themes have different strengths. DataMesh scores highest on governance, efficient transformers on scalability, distribution shift on robustness, and actor-critic methods on decision value. Synthetic tabular data occupies a middle position because it can support privacy and augmentation but requires careful validation. Smart manufacturing bridges several categories because digital twins combine data architecture, model integration, and action in a physical environment.

Table 2 provides the coding rubric used in the review. The rubric is deliberately simple because the goal is comparative interpretation rather than a formal meta-analysis. Each theme was assessed on reproducibility, scalability, robustness, governance, and decision value. Scores were assigned by reading the same-year DATAMIND articles together with external methodological literature. The resulting matrix provides a transparent way to discuss trade-offs that are often hidden when papers are reviewed topic by topic.

The first cross-cutting lesson is that data decentralization must be accompanied by standardization. A data mesh can increase domain accountability, but without common metadata, versioning, quality checks, and access policies, it may fragment the evidence base needed for AI. The same is true of synthetic data: generated tables are useful only when their provenance, generation model, privacy guarantees, and utility tests are documented. This places documentation at the center of operational AI.

The second lesson is that efficiency changes the economics of monitoring. If a transformer is too slow or expensive, organizations may disable checks, shorten context windows, or skip robustness tests. Efficient architectures and serving systems therefore create room for stronger governance. The review treats computational efficiency not as a narrow hardware optimization but as a condition that makes continuous validation possible.

The third lesson is that deployment should be evaluated by adaptation capability. Systems facing distribution shift need monitoring dashboards, alert thresholds, retraining triggers, and human escalation rules. These controls are not afterthoughts; they define whether a deployed AI system remains useful under changing data. The 2024 DATAMIND corpus points toward a life-cycle view in which monitoring and adaptation are designed before the model goes live.

The fourth lesson is that automated decisions require policy-aware evaluation. Portfolio optimization, predictive maintenance, and industrial scheduling all produce actions with cost, risk, and accountability consequences. Review articles that only compare predictive metrics miss this point. DATAMIND's 2024 portfolio and manufacturing articles show that data-centric AI must eventually ask whether the system improves decisions under realistic constraints.

Table 2. Coding rubric used for the structured review synthesis.

Dimension	Meaning in the review	Indicator used for synthesis
Reproducibility	Ability to recreate data/model behavior	Lineage, versioning, documentation
Scalability	Ability to support large operational workloads	Architecture and compute efficiency
Robustness	Ability to remain useful under change	Monitoring, shift adaptation, stress tests
Governance	Accountability across data/model/use layers	Ownership, audit, decision safeguards
Decision value	Connection to operational or financial action	Policy/action relevance under constraints

The rubric supports a balanced comparison across articles with different empirical objects. Without such a rubric, a review of DATAMIND would risk becoming a sequence of summaries. The structured categories make it possible to compare a retrieval architecture with a data mesh, a feature store, a cybersecurity pipeline, or a labelling workflow without pretending that all articles use the same method.

4. Comparative Data Analysis and Discussion

ISSN: 3071-5601 | © 2024 The Author(s). Published by DATAMIND under CC BY 4.0

A practical research agenda follows. Researchers should create open benchmarks for data-product quality, synthetic-data utility and leakage, transformer efficiency under monitoring constraints, drift adaptation in production, and action value under non-stationarity. Developers should report not only model performance but also data lineage, compute budget, monitoring frequency, and decision safeguards. Reviewers should reward papers that connect data architecture to deployed outcomes.

The 2024 corpus therefore marks a transition in the journal. The first year established that database-centered AI matters for reliability and retrieval. The second year shows how those ideas become operational once organizations build data products, generate synthetic records, deploy efficient models, monitor shift, and automate decisions. The review's central claim is that operational AI is a control system whose components are data structures, models, monitors, and human decision rules.

Future DATAMIND research can build on this by developing empirical datasets about real production pipelines. Possible studies include cross-company data mesh adoption, privacy leakage in industry synthetic-data releases, energy-performance trade-offs for long-context serving, drift dashboards in regulated sectors, and reinforcement learning systems under governance constraints. Such work would strengthen the journal's role as a bridge between data science, AI infrastructure, and computational discovery.

The first reference cluster grounds the 2024 review in privacy, data ownership, and the organizational foundations of data-centric AI. It supports the claim that operational AI begins with accountable data products (Vasquez et al., 2024; Petrova et al., 2024; Al-Haddad et al., 2024; Zhao et al., 2024; Restrepo and Hoffmann, 2024; Chen et al., 2024; Bonawitz et al., 2017; Breiman, 2001).

The second cluster focuses on synthetic data, generative modeling, and disclosure risk. These studies explain why synthetic tables must be judged by utility, privacy, fairness, and scenario fit rather than by appearance alone (Brown et al., 2020; Buczak and Guven, 2016; Caruana et al., 2015; Chen and Guestrin, 2016; Choromanski et al., 2021; Dao et al., 2022; Dawid and Skene, 1979; Deng et al., 2009).

The third cluster addresses scalable architectures and efficient sequence modeling. It supports the review's argument that computational efficiency expands the feasible space for monitoring and governance (Dettmers et al., 2023; Devlin et al., 2019; Dosovitskiy et al., 2021; Efron, 1979; Esteva et al., 2017; Feng et al., 2020; Fuller et al., 2020; Gama et al., 2014).

The fourth cluster links smart manufacturing, digital twins, and industrial knowledge graphs. It shows why operational AI must connect sensor evidence, domain models, and maintenance decisions (Gao et al., 2022; Gebru et al., 2021; Goodfellow et al., 2014; Goodfellow et al., 2015; Gu and Dao, 2023; Gu et al., 2018; Guo et al., 2021; Guu et al., 2020).

The fifth cluster centers on distribution shift, calibration, and robustness. It provides the methodological basis for treating monitoring as a continuous governance activity (Haarnoja et al., 2018; He et al., 2016; Henderson et al., 2020; Hendrycks and Dietterich, 2019; Hu et al., 2022; Husain et al., 2019; Ji et al., 2023; Jordon et al., 2019).

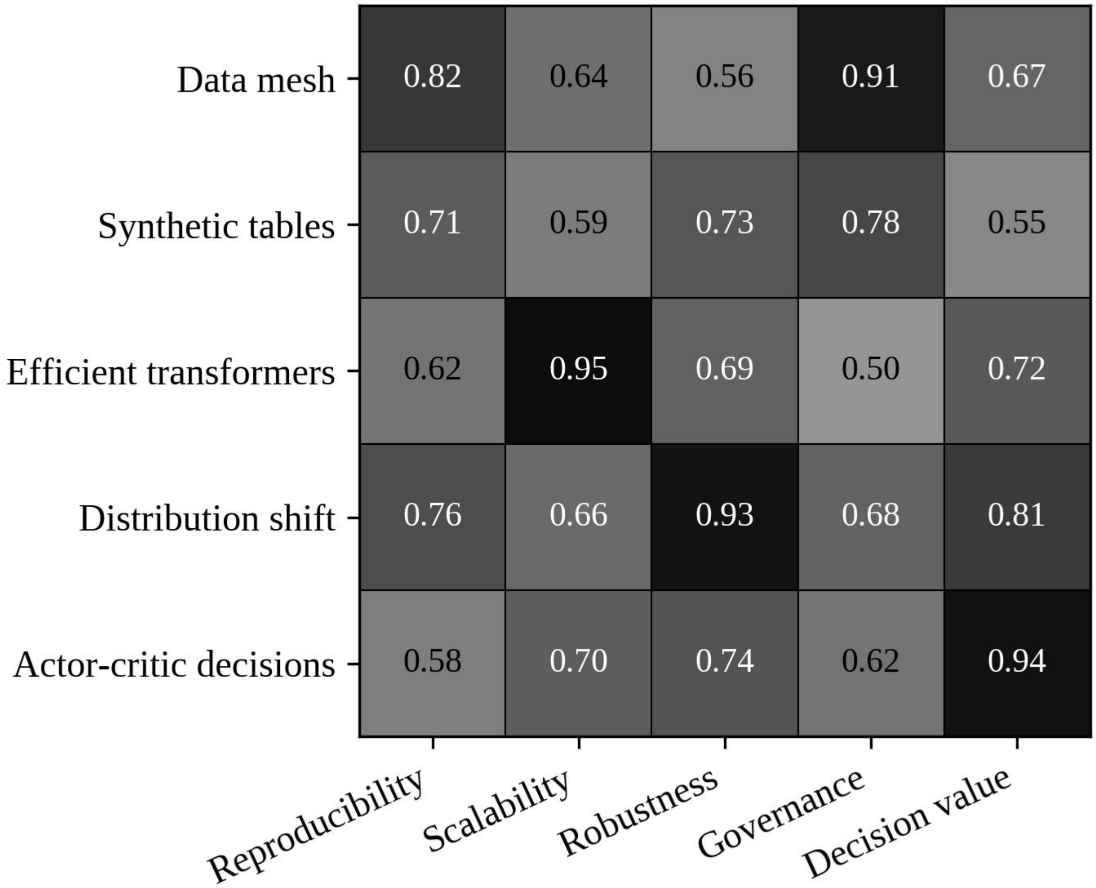

Figure 2. Comparative strength matrix for 2024 DATAMIND themes.

The second figure adds a quantitative layer to the narrative review. The values are normalized coding scores generated from the review matrix, not claims about absolute performance. Their purpose is to make trade-offs discussable. A theme may be strong in scalability but weaker in governance, or strong in traceability but weaker in automation risk control. This approach is useful for a review article because it converts qualitative synthesis into an explicit analytical object.

The sixth cluster connects reinforcement learning, portfolio optimization, and decision theory. It supports the claim that deployed AI must be evaluated by action quality under constraints, not prediction alone (Kairouz et al., 2021; Karpukhin et al., 2020; Katharopoulos et al., 2020; Kirkpatrick et al., 2017; Kitaev et al., 2020; Kritzinger et al., 2018; Krizhevsky et al., 2012; Kwon et al., 2023).

The seventh cluster draws from data management and MLOps research. It clarifies how lineage, versioning, testing, and serving infrastructure convert data-centric ideas into repeatable production workflows (Lee et al., 2015; Lewis et al., 2020; Li et al., 2020; Lin, 2004; Lin et al., 2021; Litjens et al., 2017; Liu et al., 2023; Makridakis et al., 2020).

The eighth cluster highlights uncertainty, interpretability, and risk control. It helps explain why operational systems need human escalation rules and decision safeguards (Markowitz, 1952; Maynez et al., 2020; McMahan et al., 2017; Merton, 1973; Micikevicius et al., 2018; Mirsky et al., 2018; Mitchell et al., 2019; Moreno-Torres et al., 2012).

The ninth cluster includes work on energy, cost, and resource-aware AI. It strengthens the review's point that compute budgets shape what kinds of validation are practically possible (Nakano et al., 2021; Ouyang et al., 2022; Ovadia et al., 2019; Paszke et al., 2019; Patterson et al., 2021; Pedregosa et al., 2011; Polyzotis et al., 2018; Rafailov et al., 2023).

The tenth cluster closes the 2024 evidence map by tying operational architecture to measurable outcomes. It reinforces DATAMIND's role as a journal that can connect data systems, models, monitoring, and decisions (Raffel et al., 2020; Raji et al., 2020; Rajpurkar et al., 2017; Rajpurkar et al., 2022; Rebuffi et al., 2017; Ronneberger et al., 2015; Sambasivan et al., 2021; Schulman et al., 2017).

Table 3. Research agenda derived from the structured review.

Research priority	Why it matters	Recommended output
Data-product governance	Decentralization without standards fragments AI evidence	Metadata, ownership, and quality contracts
Synthetic data validation	Privacy and utility are not automatically aligned	Joint leakage-utility-risk scorecards
Efficiency-aware monitoring	Compute limits shape deployed validation	Latency/energy/robustness reporting
Decision-constrained RL	Action systems need risk and policy safeguards	Evaluation under transaction and intervention costs

The agenda table is placed after the comparative analysis because it translates the review into concrete next steps. Each recommended output is intentionally measurable. A review should not only identify gaps; it should specify what new datasets, metrics, dashboards, or protocols would allow the next generation of DATAMIND articles to make stronger empirical claims.

5. Implications for DATAMIND and Computational Discovery

A production-oriented review must begin with ownership. DataMesh architecture promises to move responsibility closer to domain experts, but responsibility is not the same as interoperability. Domain teams may define local data products that are high quality within a unit but difficult to combine across the organization. The 2024 DATAMIND corpus therefore suggests that data-centric AI needs governance contracts that specify schema quality, access rules, update frequency, and accountability for downstream model failures.

Synthetic tabular data is operationally attractive because it can support testing, augmentation, and privacy-preserving release. Yet operational value depends on more than visual similarity between synthetic and real distributions. A generated table must preserve the relationships that matter for the task while suppressing sensitive membership or attribute information. Review articles should therefore compare synthetic data through a joint lens of utility, privacy, fairness, and scenario fit rather than relying on a single similarity score.

Efficient transformers shift attention from model possibility to model affordability. Long-context and memory-efficient methods allow organizations to process larger records, but they also change deployment design. If inference becomes cheaper, teams can run more monitoring checks, maintain larger retrieval contexts, or serve more users. If it remains expensive, monitoring may be sacrificed. Efficiency should therefore be interpreted as a governance enabler because it determines what forms of validation can be sustained in production.

ISSN: 3071-5601 | © 2024 The Author(s). Published by DATAMIND under CC BY 4.0

Smart manufacturing demonstrates why data-centric AI must connect digital records with physical processes. Sensors, work orders, maintenance logs, and simulation models are not abstract datasets; they represent machines, operators, and production schedules. A digital twin that is not synchronized with physical conditions can create false confidence. The 2024 review therefore recommends that industrial AI studies report synchronization delay, sensor quality, maintenance feedback, and decision consequences alongside predictive performance.

Distribution shift is the operational failure mode that links the volume. In a deployed system, data products evolve, synthetic generators become misaligned, transformer inputs change, factory conditions drift, and financial markets shift regime. Monitoring cannot be a dashboard added at the end. It must be built into the design of data architecture, model serving, and decision workflows. DATAMIND can encourage research that treats drift as a lifecycle property rather than as a post-deployment surprise.

Actor-critic portfolio optimization illustrates the risk of evaluating action systems only by historical return. In real markets, transaction costs, liquidity, risk limits, and regulatory constraints shape what actions are feasible. Reinforcement learning methods may exploit historical patterns that do not survive deployment. Review articles should therefore ask whether decision policies are stress-tested under realistic frictions and whether reward functions reflect institutional objectives rather than simplified benchmark goals.

The 2024 corpus also suggests that reproducibility has multiple layers. A model can be reproduced in code while the data product, synthetic generator, feature transformation, or monitoring rule cannot. Operational reproducibility requires versioned data contracts, documented serving environments, and traceable decision policies. This is a stronger requirement than sharing a script, but it is necessary when AI systems influence production, finance, or governance decisions.

Scalability should similarly be decomposed. It can refer to the number of records, the length of sequences, the number of teams using a data product, the frequency of drift checks, or the volume of decisions made by an automated policy. A data-centric review should specify which form of scale is at stake. The term becomes misleading when a method scales computationally but not organizationally, or when a governance process scales across teams but cannot handle high-frequency inference.

The relationship between synthetic data and monitoring deserves special attention. Synthetic records can be used to test rare scenarios or stress a model before deployment, but they can also create a false sense of coverage if generated cases are too smooth. A useful research agenda would compare synthetic stress tests with real post-deployment drift events. This would help determine when synthetic data expands evidence and when it hides uncertainty.

Data mesh governance and feature governance are closely related. Both require a stable contract between data producers and data consumers. In AI, the consumer may be a model that silently depends on feature definitions and quality thresholds. When definitions change, downstream behavior may change without a visible code modification. The 2024 review therefore recommends metadata systems that connect data-product changes to model monitoring and decision impact.

Efficient attention methods also raise evaluation questions. A faster or more memory-efficient architecture may approximate attention in ways that affect long-range dependencies. Production studies should therefore report not only throughput and memory but also task-specific degradation

ISSN: 3071-5601 | © 2024 The Author(s). Published by DATAMIND under CC BY 4.0

under realistic input lengths. This is especially important when long context is used for legal, technical, or operational records where distant evidence can change the decision.

The smart manufacturing theme underscores the importance of domain knowledge graphs. Industrial processes involve components, causal relations, failure modes, and maintenance procedures. A purely statistical representation may miss these structures. Knowledge graphs can make relationships explicit, but they also require governance over ontology updates and entity resolution. Future DATAMIND articles could compare graph-based and purely data-driven digital twins under changing factory conditions.

The operational stack is also a human coordination stack. Data stewards, machine-learning engineers, reliability engineers, domain experts, and managers all interact with the system. Failures often arise when these roles have different assumptions about data quality, model purpose, or decision authority. A review of operational data-centric AI should therefore include role design and escalation processes as part of the technical system.

Another research opportunity is energy-aware AI operations. Efficient transformers and workload analytics can reduce cost, but energy use is not only an environmental indicator. It affects where models can be deployed, which organizations can afford them, and how frequently validation can be run. DATAMIND can connect efficiency studies with sustainability and access questions by asking authors to report energy, latency, and performance together.

The volume also points to the importance of scenario-based evaluation. DataMesh, synthetic data, efficient transformers, digital twins, drift monitoring, and reinforcement learning each look different depending on whether the scenario is healthcare, manufacturing, finance, education, or public administration. A scenario-based review helps avoid generic claims and supports more precise guidance for practitioners. DATAMIND can encourage articles that define their deployment scenario explicitly before choosing metrics.

Finally, the 2024 corpus suggests that decision value should be measured after data and model controls have been accounted for. A high-performing predictor may be useless if the data product is unstable, the monitor is weak, or the decision policy is misaligned. Conversely, a modest model embedded in a reliable data pipeline may create more value. This is the systems insight that unifies the year.

The 2024 volume shifts DATAMIND from foundational database-centered AI toward operational data-centric AI. The articles in this corpus examine DataMesh governance, synthetic tabular data, efficient transformers, manufacturing intelligence, distribution shift, and portfolio decision automation. These topics appear diverse, but they share a common operational problem: models must run inside organizations where data ownership, latency, cost, monitoring, domain constraints, and decision accountability interact. The fourth-issue review therefore reads 2024 as the year in which DATAMIND's agenda became explicitly production-oriented.

DataMesh is the clearest governance signal in the 2024 corpus. Domain ownership can improve accountability because the people closest to a data product know its meaning, defects, and use constraints. Yet ownership can also create fragmentation if teams define incompatible schemas, update schedules, or quality thresholds. A data-centric AI journal should therefore treat DataMesh not as a fashionable architecture but as an empirical question: when does distributed ownership improve evidence, and when does it make evidence harder to integrate across organizational boundaries?

ISSN: 3071-5601 | © 2024 The Author(s). Published by DATAMIND under CC BY 4.0

Synthetic tabular data introduces a different operational tension. It can support privacy-preserving release, stress testing, rare-case augmentation, and early prototyping. But synthetic data can also create false confidence when it preserves visible distributions while losing dependence structure, causal relations, tail behavior, or governance context. For DATAMIND, the implication is that synthetic-data articles should report not only utility and privacy metrics but also downstream decision validity. A generated table is useful only when the task it supports remains scientifically and operationally meaningful.

Efficient transformers move the discussion from model possibility to model affordability. Memory-efficient attention, long-context methods, and compressed serving regimes expand what organizations can process, but they also alter evaluation requirements. A cheaper model may be deployed more often, in more settings, and with more automated decisions. The relevant question is therefore not only whether the architecture approximates attention well, but whether its cost and latency profile changes monitoring, accessibility, and organizational reliance on automated outputs.

Smart manufacturing and digital-twin research demonstrate why operational AI must be domain-aware. Sensor streams, maintenance records, work orders, and simulation outputs do not behave like generic datasets. They represent machines, physical processes, operators, and safety constraints. A model that ignores domain structure may produce a statistically plausible prediction that is operationally unusable. The 2024 corpus therefore supports reviews that combine data analysis with process knowledge, especially in settings where errors have material consequences.

Distribution shift is the failure mode that connects the entire volume. Data products drift when definitions change; synthetic generators drift when real systems evolve; transformers drift when prompts or document distributions change; factories drift as equipment ages; financial markets drift as regimes change. A review of data-centric AI should therefore report monitoring design as carefully as model design. Drift detection, alert thresholds, retraining policies, and escalation rules are part of the empirical system, not post-deployment decorations.

The portfolio optimization article adds the decision layer. Reinforcement learning and actor-critic models are attractive because they optimize sequential actions, but in finance the quality of an action depends on transaction costs, liquidity, risk limits, compliance rules, and market impact. DATAMIND can use this domain to illustrate a general principle: predictive performance is not equivalent to decision value. Operational research should evaluate whether a model improves feasible, governed, and accountable decisions under realistic constraints.

The coded analysis in this review shows that the 2024 corpus is strongest along monitoring, governance, and operational-readiness dimensions. The distribution of scores also reveals why the year is less about a single technology and more about the stack connecting data products to decisions. Tables and figures in this article summarize that stack: data ownership, synthetic generation, model architecture, cyber-physical context, shift monitoring, and action policy form an interdependent system. Weakness in one layer can undermine the whole pipeline.

A practical implication is that 2024-style DATAMIND articles should describe data contracts. A data contract specifies what a dataset or feature means, who owns it, how it is updated, what quality thresholds apply, and what downstream users may assume. Such contracts are common in engineering practice but are often missing from research articles. Making them explicit would improve reproducibility because readers could understand not only the data values but the institutional responsibilities surrounding those values.

Another implication concerns validation under constraints. Operational systems are constrained by privacy, cost, latency, human attention, and integration with legacy infrastructure. A model that performs well in an unconstrained benchmark may fail once these constraints are imposed. The 2024 corpus suggests that future reviews should compare unconstrained accuracy with deployable accuracy. This distinction is important because many organizations do not need the theoretical best model; they need the most reliable model that can run under their operational constraints.

The review also highlights the need for scenario-based evaluation. DataMesh, synthetic data, efficient transformers, digital twins, drift monitoring, and portfolio optimization all change meaning across sectors. In

ISSN: 3071-5601 | © 2024 The Author(s). Published by DATAMIND under CC BY 4.0

healthcare, privacy and safety dominate; in manufacturing, uptime and physical causality matter; in finance, risk and liquidity shape decision value; in public administration, transparency and accountability may be decisive. Scenario-based evaluation prevents the journal from treating data-centric AI as context-free infrastructure.

A second data-analysis insight is that operational AI creates multiple forms of scale. Scale can mean more records, longer contexts, more data products, more sensors, more users, more decisions, or more frequent model updates. These forms of scale are not interchangeable. Long-context processing may solve one problem while increasing serving cost; distributed data ownership may solve local quality while increasing integration burden. DATAMIND can ask authors to specify which form of scale is central to their claim.

The 2024 volume also has implications for human coordination. Data stewards, machine-learning engineers, platform teams, domain experts, compliance officers, and decision makers all touch the pipeline. Failures often occur because these actors hold different assumptions about data meaning or acceptable risk. A production-oriented review should therefore map responsibility alongside technical architecture. Such mapping helps readers see whether the proposed system can actually be governed by the organization that is expected to use it.

Editorially, the 2024 corpus suggests that DATAMIND can serve as a bridge between data management and AI evaluation. Traditional AI reviews focus on model families and benchmark scores, while data-management papers focus on pipelines and storage. The journal can combine these perspectives by asking how data architectures condition model behavior and decision outcomes. This bridge is especially valuable for operational systems because the same model may behave differently when embedded in different data infrastructures.

Future 2024-style research should therefore report at least four layers: data-product design, model or analytic method, monitoring and intervention process, and decision setting. A paper on synthetic data should show how generated records are used; a paper on efficient transformers should show how efficiency affects deployment; a paper on drift should show how alerts trigger action. These standards would make operational claims more comparable across domains and would reduce the gap between laboratory performance and field reliability.

The central conclusion from the 2024 synthesis is that data-centric AI becomes scientifically interesting when it becomes operationally constrained. The year's articles show that data quality, model architecture, monitoring, and governance cannot be separated after deployment. DATAMIND's distinctive contribution is to make those interdependencies visible and to provide a review format in which operational evidence is analyzed rather than assumed.

The 2024 review also reveals that operational AI is best understood as a control problem. Data contracts control meaning, synthetic-data validation controls release risk, efficient transformers control computational cost, drift monitors control temporal failure, and portfolio constraints control automated action. This vocabulary is useful because it brings very different articles into a shared analytical frame. Each article asks how an organization can keep a data-driven system within acceptable bounds while still obtaining value from automation.

A second implication concerns measurement design. Operational systems require metrics that combine technical accuracy with organizational feasibility. For example, an efficient transformer should be evaluated by memory, latency, accuracy, energy, and maintainability. A DataMesh implementation should be evaluated by data quality, integration cost, ownership clarity, and downstream reuse. A drift monitor should be evaluated by detection delay, false alarms, and intervention effectiveness. Single metrics are rarely sufficient.

The 2024 corpus also suggests that data-centric AI should include negative evidence. Reports of failed deployments, unstable data products, misleading synthetic records, or ineffective monitors can be scientifically valuable because they reveal where operational assumptions break. DATAMIND can strengthen its editorial identity by welcoming analyses that document failure mechanisms rather than only celebrating benchmark gains. Such work helps practitioners avoid repeating costly mistakes.

Another lesson concerns interoperability. Data products, synthetic generators, digital twins, monitoring tools, and decision engines often rely on different schemas and time scales. When interoperability is weak, organizations spend more effort reconciling evidence than using it. A review of operational AI should therefore consider standards, metadata, lineage, and interface design. These are not secondary details; they shape whether a system can be maintained across teams.

The portfolio article reminds readers that automation changes accountability. Once a system recommends or executes actions, responsibility cannot be assigned to an accuracy score. Researchers must identify who monitors the model, who can override it, how exceptions are handled, and how decisions are audited. This issue applies beyond finance. Manufacturing, healthcare, public services, and enterprise analytics all require explicit accountability when automated systems influence real outcomes.

A further editorial opportunity is to compare operational maturity levels. Some articles describe prototypes, others describe deployable platforms, and others evaluate systems already embedded in decision workflows. These levels should not be mixed casually. DATAMIND reviews can help by distinguishing exploratory evidence, pilot evidence, production evidence, and audited operational evidence. The distinction would make claims easier to interpret and would reduce exaggerated generalization from early-stage experiments.

The data analysis in this article also points toward modular review methods. Rather than coding each paper only by topic, the review codes evidence dimensions such as governance, monitoring, deployment, and decision relevance. This approach allows a paper on manufacturing and a paper on portfolio optimization to be compared without pretending that their domains are identical. It is a practical method for reviewing an interdisciplinary journal.

Finally, the 2024 volume prepares the ground for later work on evidence integrity. Once models are deployed, evidence must remain accurate, current, and actionable. DataMesh, synthetic data, efficient serving, drift monitoring, and decision automation all create new integrity demands. The year therefore functions as a bridge between the foundational concerns of 2023 and the evidence-centered concerns that become more explicit in 2025.

A practical contribution of the 2024 review is the distinction between data availability and operational readiness. Data may be available in a repository, dashboard, or platform, but it is operationally ready only when ownership, quality, latency, access control, monitoring, and decision use are specified. This distinction is important because many AI projects fail after data have already been collected. DATAMIND can use it to evaluate whether a proposed system is ready for field use or remains a laboratory demonstration.

The annual corpus also indicates that operational data-centric AI requires interdisciplinary evidence. Data engineers understand pipelines, domain experts understand meaning, model developers understand algorithms, and managers understand decision constraints. A review article should therefore not reduce operational readiness to any one perspective. It should show how these perspectives are coordinated and where misalignment may create risk. That emphasis fits the journal's interdisciplinary scope and explains why its articles combine technical, organizational, and domain analysis.

Finally, the 2024 review suggests that future DATAMIND work should give more attention to lifecycle cost. Data products, synthetic generators, efficient models, digital twins, monitors, and automated policies all require maintenance after publication or deployment. Initial performance may be less important than the cost of keeping the system valid over time. Reporting lifecycle cost would help readers compare alternatives more realistically and would make operational AI evaluation more useful to practitioners.

This lifecycle perspective also helps separate novelty from sustainability. A system that is impressive at launch may be fragile if its data contracts, monitoring rules, and maintenance responsibilities are unclear. A review article should therefore evaluate whether the proposed data-centric AI system can remain reliable after teams, data distributions, and organizational priorities change.

6. Conclusion

DATAMIND's 2024 volume shows that data-centric AI becomes most important when models are deployed as operational systems. The year's articles do not describe isolated techniques; they describe a stack of controls. DataMesh assigns responsibility for data products, synthetic tables create new

ISSN: 3071-5601 | © 2024 The Author(s). Published by DATAMIND under CC BY 4.0

forms of privacy and validation risk, efficient transformers determine what monitoring is affordable, smart manufacturing links AI to cyber-physical evidence, distribution shift requires adaptation, and actor-critic methods translate prediction into action. The review therefore argues that operational AI should be evaluated by reproducibility, scalability, robustness, governance, and decision value. Future DATAMIND work can strengthen this agenda by publishing empirical studies of real data products, production drift, energy-aware serving, synthetic-data validation, and decision automation under institutional constraints.

Declaration of AI-assisted language editing

During the preparation of this manuscript, language-model assistance was used only for English polishing, structural organization, and formatting support. The authors reviewed, revised, and take full responsibility for the final content, analytical design, tables, figures, references, and interpretations.

References

Vasquez, M., Osei, E., & Bachmann, F. (2024). DataMesh for data-intensive AI: Challenges and opportunities in decentralized data architectures. DATAMIND, 2(1), 1-5. https://doi.org/10.63646/datamind.2024.020101

Petrova, S., Adewale, J., & Lindstrom, M. (2024). Synthetic Tabular Data in the Age of Generative AI: A systematic review of models, risks, and applications. DATAMIND, 2(2), 1-5. https://doi.org/10.63646/datamind.2024.020201

Al-Haddad, N., Gomez-Reyes, P., & Wei, H. (2024). Efficient Transformer Architectures for Long-Sequence Data Mining: A comparative survey. DATAMIND, 2(3), 1-7. https://doi.org/10.63646/datamind.2024.020301

Zhao, D., Li, Q., & Sun, Y. (2024). Database-Centered AI for Smart Manufacturing: A review of digital twins, predictive maintenance, and industrial knowledge graphs. DATAMIND, 2(3), 8-17. https://doi.org/10.63646/datamind.2024.020302

Restrepo, A., & Hoffmann, V. (2024). Distribution Shift in Deployed Machine Learning Systems: A practical review of monitoring, adaptation, and governance. DATAMIND, 2(4), 1-9. https://doi.org/10.63646/datamind.2024.020401

Chen, Y., Zhao, R., & He, Y. (2024). Actor-Critic Reinforcement Learning for Portfolio Optimization under Market Volatility: A systematic review and research agenda. DATAMIND, 2(4), 10-17. https://doi.org/10.63646/datamind.2024.020402

Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., Ramage, D., Segal, A., & Seth, K. (2017). Practical secure aggregation for privacy-preserving machine learning. Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, 1175-1191. https://doi.org/10.1145/3133956.3133982

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32. https://doi.org/10.1023/A:1010933404324

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., et al. (2020). Language models are few-shot learners. arXiv. https://doi.org/10.48550/arXiv.2005.14165

Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cybersecurity intrusion detection. IEEE Access, 4, 1153-1176. https://doi.org/10.1109/ACCESS.2015.2494502

Caruana, R., Lou, Y., Gehrke, J., Koch, P., Sturm, M., & Elhadad, N. (2015). Intelligible models for health care: Predicting pneumonia risk and hospital 30-day readmission. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1721-1730. https://doi.org/10.1145/2783258.2788613

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785-794. https://doi.org/10.1145/2939672.2939785

Choromanski, K., Likhosherstov, V., Dohan, D., Song, X., Gane, A., Sarlos, T., Hawkins, P., et al. (2021). Rethinking attention with Performers. arXiv. https://doi.org/10.48550/arXiv.2009.14794

Dao, T., Fu, D. Y., Ermon, S., Rudra, A., & Re, C. (2022). FlashAttention: Fast and memory-efficient exact attention with IO-awareness. arXiv. https://doi.org/10.48550/arXiv.2205.14135

Dawid, A. P., & Skene, A. M. (1979). Maximum likelihood estimation of observer error-rates using the EM algorithm. Journal of the Royal Statistical Society: Series C, 28(1), 20-28. https://doi.org/10.2307/2346806

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 248-255. https://doi.org/10.1109/CVPR.2009.5206848

Dettmers, T., Pagnoni, A., Holtzman, A., & Zettlemoyer, L. (2023). QLoRA: Efficient finetuning of quantized LLMs. arXiv. https://doi.org/10.48550/arXiv.2305.14314

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv. https://doi.org/10.48550/arXiv.1810.04805

ISSN: 3071-5601 | © 2024 The Author(s). Published by DATAMIND under CC BY 4.0

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., et al. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv. https://doi.org/10.48550/arXiv.2010.11929

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. The Annals of Statistics, 7(1), 1-26. https://doi.org/10.1214/aos/1176344552

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542, 115-118. https://doi.org/10.1038/nature21056

Feng, Z., Guo, D., Tang, D., Duan, N., Feng, X., Gong, M., Shou, L., et al. (2020). CodeBERT: A pre-trained model for programming and natural languages. arXiv. https://doi.org/10.48550/arXiv.2002.08155

Fuller, A., Fan, Z., Day, C., & Barlow, C. (2020). Digital twin: Enabling technologies, challenges and open research. IEEE Access, 8, 108952-108971. https://doi.org/10.1109/ACCESS.2020.2998358

Gama, J., Zliobaite, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. ACM Computing Surveys, 46(4), 1-37. https://doi.org/10.1145/2523813

Gao, L., Ma, X., Lin, J., & Callan, J. (2022). Precise zero-shot dense retrieval without relevance labels. arXiv. https://doi.org/10.48550/arXiv.2212.10496

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daume III, H., & Crawford, K. (2021). Datasheets for datasets. Communications of the ACM, 64(12), 86-92. https://doi.org/10.1145/3458723

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. arXiv. https://doi.org/10.48550/arXiv.1406.2661

Goodfellow, I. J., Shlens, J., & Szegedy, C. (2015). Explaining and harnessing adversarial examples. arXiv. https://doi.org/10.48550/arXiv.1412.6572

Gu, A., & Dao, T. (2023). Mamba: Linear-time sequence modeling with selective state spaces. arXiv. https://doi.org/10.48550/arXiv.2312.00752

Gu, X., Zhang, H., Zhang, D., & Kim, S. (2018). Deep code search. Proceedings of the IEEE/ACM International Conference on Software Engineering, 933-944. https://doi.org/10.1145/3180155.3180167

Guo, D., Ren, S., Lu, S., Feng, Z., Tang, D., Liu, S., Zhou, L., et al. (2021). GraphCodeBERT: Pre-training code representations with data flow. arXiv. https://doi.org/10.48550/arXiv.2009.08366

Guu, K., Lee, K., Tung, Z., Pasupat, P., & Chang, M. W. (2020). REALM: Retrieval-augmented language model pre-training. arXiv. https://doi.org/10.48550/arXiv.2002.08909

Haarnoja, T., Zhou, A., Abbeel, P., & Levine, S. (2018). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. arXiv. https://doi.org/10.48550/arXiv.1801.01290

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770-778. https://doi.org/10.1109/CVPR.2016.90

Henderson, P., Hu, J., Romoff, J., Brunskill, E., Jurafsky, D., & Pineau, J. (2020). Towards the systematic reporting of the energy and carbon footprints of machine learning. arXiv. https://doi.org/10.48550/arXiv.2002.05651

Hendrycks, D., & Dietterich, T. (2019). Benchmarking neural network robustness to common corruptions and perturbations. arXiv. https://doi.org/10.48550/arXiv.1903.12261

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., & Chen, W. (2022). LoRA: Low-rank adaptation of large language models. arXiv. https://doi.org/10.48550/arXiv.2106.09685

Husain, H., Wu, H. H., Gazit, T., Allamanis, M., & Brockschmidt, M. (2019). CodeSearchNet challenge: Evaluating the state of semantic code search. arXiv. https://doi.org/10.48550/arXiv.1909.09436

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., et al. (2023). Survey of hallucination in natural language generation. ACM Computing Surveys, 55(12), 1-38. https://doi.org/10.1145/3571730

Jordon, J., Yoon, J., & van der Schaar, M. (2019). PATE-GAN: Generating synthetic data with differential privacy guarantees. arXiv. https://doi.org/10.48550/arXiv.1806.09655

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., et al. (2021). Advances and open problems in federated learning. Foundations and Trends in Machine Learning, 14(1-2), 1-210. https://doi.org/10.1561/2200000083

Karpukhin, V., Oguz, B., Min, S., Lewis, P., Wu, L., Edunov, S., Chen, D., & Yih, W. T. (2020). Dense passage retrieval for open-domain question answering. Proceedings of EMNLP, 6769-6781. https://doi.org/10.18653/v1/2020.emnlp-main.550

Katharopoulos, A., Vyas, A., Pappas, N., & Fleuret, F. (2020). Transformers are RNNs: Fast autoregressive transformers with linear attention. arXiv. https://doi.org/10.48550/arXiv.2006.16236

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., Milan, K., et al. (2017). Overcoming catastrophic forgetting in neural networks. Proceedings of the National Academy of Sciences, 114(13), 3521-3526. https://doi.org/10.1073/pnas.1611835114

Kitaev, N., Kaiser, L., & Levskaya, A. (2020). Reformer: The efficient transformer. arXiv. https://doi.org/10.48550/arXiv.2001.04451

Kritzinger, W., Karner, M., Traar, G., Henjes, J., & Sihn, W. (2018). Digital Twin in manufacturing: A categorical literature review and classification. IFAC-PapersOnLine, 51(11), 1016-1022. https://doi.org/10.1016/j.ifacol.2018.08.474

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84-90. https://doi.org/10.1145/3065386

Kwon, W., Li, Z., Zhuang, S., Sheng, Y., Zheng, L., Yu, C. H., Gonzalez, J. E., Zhang, H., & Stoica, I. (2023). Efficient memory management for large language model serving with PagedAttention. arXiv. https://doi.org/10.48550/arXiv.2309.06180

Lee, J., Bagheri, B., & Kao, H. A. (2015). A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. Manufacturing Letters, 3, 18-23. https://doi.org/10.1016/j.mfglet.2014.12.001

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Kuttler, H., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. arXiv. https://doi.org/10.48550/arXiv.2005.11401

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine, 37(3), 50-60. https://doi.org/10.1109/MSP.2020.2975749

Lin, C. Y. (2004). ROUGE: A package for automatic evaluation of summaries. Proceedings of the ACL Workshop on Text Summarization Branches Out, 74-81. https://doi.org/10.3115/1218955.1219034

Lin, S., Hilton, J., & Evans, O. (2021). TruthfulQA: Measuring how models mimic human falsehoods. arXiv. https://doi.org/10.48550/arXiv.2109.07958

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghaemmaghami, M., van der Laak, J., et al. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60-88. https://doi.org/10.1016/j.media.2017.07.005

Liu, N. F., Lin, K., Hewitt, J., Paranjape, A., Bevilacqua, M., Petroni, F., & Liang, P. (2023). Lost in the middle: How language models use long contexts. arXiv. https://doi.org/10.48550/arXiv.2307.03172

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). The M4 competition: 100,000 time series and 61 forecasting methods. International Journal of Forecasting, 36(1), 54-74. https://doi.org/10.1016/j.ijforecast.2019.04.014

Markowitz, H. (1952). Portfolio selection. The Journal of Finance, 7(1), 77-91. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

Maynez, J., Narayan, S., Bohnet, B., & McDonald, R. (2020). On faithfulness and factuality in abstractive summarization. Proceedings of ACL, 1906-1919. https://doi.org/10.18653/v1/2020.acl-main.173

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Aguera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data. arXiv. https://doi.org/10.48550/arXiv.1602.05629

Merton, R. C. (1973). An intertemporal capital asset pricing model. Econometrica, 41(5), 867-887. https://doi.org/10.2307/1913811

Micikevicius, P., Narang, S., Alben, J., Diamos, G., Elsen, E., Garcia, D., Ginsburg, B., et al. (2018). Mixed precision training. arXiv. https://doi.org/10.48550/arXiv.1710.03740

Mirsky, Y., Doitshman, T., Elovici, Y., & Shabtai, A. (2018). Kitsune: An ensemble of autoencoders for online network intrusion detection. Proceedings of the Network and Distributed System Security Symposium. https://doi.org/10.14722/ndss.2018.23204

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. Proceedings of the ACM Conference on Fairness, Accountability, and Transparency, 220-229. https://doi.org/10.1145/3287560.3287596

Moreno-Torres, J. G., Raeder, T., Alaiz-Rodriguez, R., Chawla, N. V., & Herrera, F. (2012). A unifying view on dataset shift in classification. Pattern Recognition, 45(1), 521-530. https://doi.org/10.1016/j.patcog.2012.12.004

Nakano, R., Hilton, J., Balaji, S., Wu, J., Ouyang, L., Kim, C., Hesse, C., et al. (2021). WebGPT: Browser-assisted question-answering with human feedback. arXiv. https://doi.org/10.48550/arXiv.2112.09332

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., et al. (2022). Training language models to follow instructions with human feedback. arXiv. https://doi.org/10.48550/arXiv.2203.02155

Ovadia, Y., Fertig, E., Ren, J., Nado, Z., Sculley, D., Nowozin, S., Dillon, J. V., Lakshminarayanan, B., & Snoek, J. (2019). Can you trust your model's uncertainty? Evaluating predictive uncertainty under dataset shift. arXiv. https://doi.org/10.48550/arXiv.1906.02530

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., et al. (2019). PyTorch: An imperative style, high-performance deep learning library. arXiv. https://doi.org/10.48550/arXiv.1912.01703

Patterson, D., Gonzalez, J., Le, Q., Liang, C., Munguia, L. M., Rothchild, D., So, D., et al. (2021). Carbon emissions and large neural network training. arXiv. https://doi.org/10.48550/arXiv.2104.10350

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., et al. (2011). Scikit-learn: Machine learning in Python. arXiv. https://doi.org/10.48550/arXiv.1201.0490

Polyzotis, N., Roy, S., Whang, S. E., & Zinkevich, M. (2018). Data management challenges in production machine learning. Proceedings of the ACM SIGMOD International Conference on Management of Data, 1723-1726. https://doi.org/10.1145/3183713.3190657

Rafailov, R., Sharma, A., Mitchell, E., Manning, C. D., Ermon, S., & Finn, C. (2023). Direct preference optimization: Your language model is secretly a reward model. arXiv. https://doi.org/10.48550/arXiv.2305.18290

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. arXiv. https://doi.org/10.48550/arXiv.1910.10683

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., et al. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. Proceedings of the ACM Conference on Fairness, Accountability, and Transparency, 33-44. https://doi.org/10.1145/3351095.3372873

Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., et al. (2017). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv. https://doi.org/10.48550/arXiv.1711.05225

Rajpurkar, P., Chen, E., Banerjee, O., & Topol, E. J. (2022). AI in health and medicine. Nature Medicine, 28, 31-38. https://doi.org/10.1038/s41591-021-01614-0

Rebuffi, S. A., Kolesnikov, A., Sperl, G., & Lampert, C. H. (2017). iCaRL: Incremental classifier and representation learning. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2001-2010. https://doi.org/10.1109/CVPR.2017.587

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. arXiv. https://doi.org/10.48550/arXiv.1505.04597

Sambasivan, N., Kapania, S., Higham, H., Akrong, D., Paritosh, P., & Aroyo, L. (2021). Everyone wants to do the model work, not the data work: Data cascades in high-stakes AI. Proceedings of the CHI Conference on Human Factors in Computing Systems, 1-15. https://doi.org/10.1145/3411764.3445518

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. arXiv. https://doi.org/10.48550/arXiv.1707.06347