SAJSET
ISSN: 2992-3581

Savannah Journal of Science and Engineering Technology
Volume: 03; Special Issue: 06; Pages: 479-485; 2025.
www.sajsetjournal.com.ng

DOI: xx.xxxx.x/sajset-03-2025-0062

SAJSET
ISSN: 2992-3581

Cultivating a data-driven culture with human-led AI: Balancing algorithmic power and human intuition in the age of big data

[1]Yusuf, B. J., [2]Adikpe, A. O., [1]Ajayi, J. and [3]Tsenyil, F. Y.
[1]*Department of Computer Engineering Federal Polytechnic, Nyak, Shendam, Plateau State, Nigeria*
[2]*Department of Electronic and Telecommunications Engineering, Ahmadu Bello University Zaria, Kaduna State, Nigeria*
[3]*Department of Civil Engineering, Federal Polytechnic, Nyak, Shendam, Plateau State, Nigeria*
agburuadikpe@gmail.com; Johnsonajayi300@gmail.com; tsenyilfestus@gmail.com

Paper History
Received: 01st October, 2025
Accepted: 15th October, 2025
Published: October, 2025

Corresponding author
Yusuf, B. J.
bassijeremiahyusuf@fedpolynyakshendam.edu.ng

Abstract:
The geometric increase in the volume of data creation has taken data-driven decision making from a competitive advantage to an operation essential. Artificial-Intelligence (AI) systems, especially, those, based on Machine-Learning (ML) and Natural-Language-Processing (NLP) methods, provide organizations with an opportunity to analyze and process petabyte-scale data streams in nearly real time. However, companies that rely solely on algorithmic outputs expose themselves to ethical violations, undisclosed discrimination bias and fragile black box decision pathways. The discussion in this paper presents a Human-Led AI (HLAI) model which considers the need for human intervention, ethical decision-making and situational awareness that augments the computational advantages of AI. On this premise, recent studies, industry surveys and regulatory advice that provide an organizational blueprint are discussed to develop a culture of data-driven organization based on HLAI principles. Examples covering healthcare applications and modelling risks in financial services are delineated. Furthermore, avenues that are crucial to foster a symbiosis of human and machine decision system are discussed.

Keywords: *Algorithm, Big data, Ethics, Human AI, Machine-Learning*

1. Introduction

In recent years, there has been exponential increase in the data volume demanded and used by end- users and the global projections delineated in [1] states this is expected to exceed 180 zettabytes by 2026, which represents about 180% increase from the data volume expended in 2020. Additionally, it is forecasted that the Compound Annual Growth Rate (CAGR) of the low-code development platforms market will amount to around 30% in 2030. In spite of the increased adoption in Machine Learning (ML) libraries, Business Intelligence systems, workflow orchestrators, and cloud-native analytics engines, it is observed that 65% of enterprises rather justify and carry out their actions on selective data already carried out than use the data to carry influence the strategy to be implemented [2]. With the adoption of decision intelligence, analysts and organization get to overcome the gap between data availability and data value. Therefore, it is not just powerful analytics that is needed; also strongly embedded within the culture of the organization should be the fact that evidence beats anecdote. At the same time, the level of the adoption of Artificial Intelligence (AI) is gaining momentum. Deep neural networks, Large Language Models (LLMs), and edge Machine Learning (ML) pipelines automate the collection, pattern discovery, and prediction of data [3]. Nevertheless, in strict algorithmic control, opacity is magnified [4, 5], tracking back to the source is hard [6], and this has the potential to encode history biases [4, 6]. On this premise, the following is the critical research question of this paper: What can enable organizations to harness the power that AI brings to analytics, and at the same time retain human capabilities such as intuition, creativity and ethical considerations needed to make trustworthy judgements?

The section delineates literature, policy systems, and business examples that support the dynamic process of incorporating AI into decision setting. The Human-Led AI (HLAI) approach that is discussed in this paper is based on three overlapping fields: data-driven culture, human and AI collaboration, trustworthy/responsible AI.

1.1 Data-Driven Culture

A data-driven culture is an organizational culture and design where most decisions are made on a basis of empirical evidence, as opposed to intuition, custom or authority. The most common features of these cultures are the easy availability of clean, well-documented data; high level of analytical literacy throughout the organization; leadership that is dedicated to providing examples of data-informed behavior; and the creation of feedback loops that integrates learning into the business processes. Among the most crucial developments identified in this field by Gartner

2024 Data & Analytics Summit is the changing focus to data-driven analytics towards decision-centric analytics. The new paradigm puts more focus on the context-sensitive workflows where the data will not only facilitate the measurement but also judgment, adaption, and deployment of strategies. In other words, a data-driven organization can look at the performance metrics through dashboards, whereas a decision-centric organization associates the metrics with the changing objectives, customer requirements, and external threats. It was also highlighted in the summit that most firms are unable to convert sophisticated data abilities into actions that can make a difference, frequently because of culture rootedness or autonomous authority, or absence of belief in AI suggestions [2].

1.1.1 Human-AI Collaboration

Human AI collaboration is one of the forms of hybrid decision-making where humans and AI systems are merged with the computational expertise of the former and the contextual, ethical and strategic capabilities of the latter. Recent empirical studies have revealed how such hybrid teams have constantly performed better than AI-only and human-only strategies in most of the fields. This not only encompasses areas under the scrutiny of data heavy tasks, such as fraud detection, logistics, but also more subtle tasks such as creative writing, legal reasoning and strategic planning [2, 4, 7]. As an example, in one experiment, authors working collaboratively with the large language model generated more engaging and grammatically correct materials, but only when they were tasked to curate, edit, and revise the results of the model [7]. These results are duplicated in the financial and medical-diagnostic spheres, in which human control can be used to detect context-dependent failures or raise ethical warning signs otherwise ignored by AI. This literature supports the importance of the concept of human-in-the-loop (HITL) design, when humans are not just stack-end decision-makers-rum, but also design/co-designers that can be used to validate AI processes in real-time.

1.1.2 Trustworthy and responsible AI

With increased AI system relative influence in private and public decision-making, the desire to use ethical and accountable AI as a framework of the world has taken a regulatory and ethical forcefulness. The AI Act created by the European Union has based its regulatory framework on three tiers of risks, where high-risk applications, including employment screening, credit scoring, and diagnosis in healthcare, are required to be transparent, have an audit guideline, and contain human oversight of their use [3]. Coupled with this, EU research ethics guidance on generative AI is focused on accountability, data provenance, and fairness as concerns machine learning system development and deployment [8]. These principles are especially important now when the great foundation models and generative AI tools have started to find their way into research, policymaking, and media creation.

On the industry front, there have been similar concerns by the financial services and auditing firms. Deloitte and Ey thought leadership requests implicit ethical governance systems, explainability of models, and traceability and human responsibility [5, 6, 9, 10]. This can be implemented, in practice, by multiple AI ethics reviewing boards, documentation standards around models (like model cards and datasets datasheets), and embedded audit histories of all automated decisions. The measures do not just serve as a means of compliance, but also a much-needed trust-construction mechanism in AI adoption.

1.2 Conceptual framework of human-led artificial intelligence

Human-Led AI (HLAI) is a way of thinking about artificial intelligence as an enhancement to, but not replacement of, human opinion. It sees a combination of effort where people have a strategic perspective, with AI having analytical power. Such model is based on the idea that making complicated decisions both needs computational accuracy and contextual knowledge, the latter of which happens to be one of the most powerful features of a human being. Fig. 1 shows a closed- loop system of HLAI, where human states all steps of the AI lifecycle: defining the problem, AI deployment and retraining. All the stages are focused on transparency, accountability, and adaptability.

1.2.1 Defining the problem

The development of the AI starts not with data, but with a problem framing, facilitated by humans. The business objectives are expressed by domain experts, policymakers or operational leaders who define the measures of success, lay down constraints and boundary conditions that the AI system must be within. This stage involves considerable contextual and domain understanding to make certain that the correct problem is included in the resolution game plan, and that it is associated with organizational or social agenda. Mid-way problems improperly framed may result in technically correct but ethically or strategically poor results further down the line [11]. Notably, these ethical priorities and degrees of tolerance to risks are also coded into the problem definition stage by humans, for example, what type of error would be more expensive (e.g. false positives as compared to false negatives in health care diagnostic processes).

1.2.2 AI-Augmented data engineering and model training

After defining the problem, the AI systems facilitate scaling of data ingestion process, data cleaning, feature extraction, and training of the models. The level makes use of automated data pipelines, machine learning (ML) algorithms, and the platforms linking to the cloud to manage large and multifaceted data sets [12]. The activities that are important are:

a. Preprocessing of data: Standardizing and Cleaning of heterogeneous data sets.
b. Feature engineering: Discovery of variables that can have predictive value.
c. Model selection: Model training, tuning, and testing using an historical data.

The process is computationally intensive and highly automated but human beings have a supervisory role especially in setting model parameters, validating training set to ensure it is representative and data lineage and integrity. In absence of human curation, models may learn on biased, non-complete or noisy data.

1.2.3 Explainable output (XAI)

Once the model is trained, it starts to give predictions. But in the human centred architecture, the results have to be explicable and interpretable by human decision-makers. Methods developed in Explainable AI (XAI), like SHAP values, LIME and counterfactual reasoning are used to reveal which variables led to which prediction and why. Such transparency is necessary in areas as finance, law and healthcare which are high-stakes domains, where decisions which cannot be seen inside a black-box are unacceptable legally and ethically. In addition, interpretability forms a trusting eye and allows downstream validation, particularly in cases where the recommendations made by the model fail to comply with the expectations of experts [13].

1.2.4 Human validation

During this step human experts look through the model, verify or overwrite its result using their tacit knowledge, situational understanding and ethical judgments. Important Human inputs are:

a. Plausibility tests: Is the result realistic, with regard to real world restraints?
b. Contextual interpretation: Do the exogenous factors (e.g. market conditions, social dynamics) not captured by the model?
c. Ethical review: What is the impact of the outcome on the imbalance between harms or risk to some groups?

Such level of checks and balances are critical specifically in situations where AI systems are working in gray areas of rules or where human wellness is at risk. Human validation means that AI suggestions do not get out of line in relation to institutional objectives and social values [14].

1.2.5 Continuous feedback adaptive learning

Contrary to the perspective in which deployment is seen as a finish line, the HLAI model presents human feedback as a constant learning point. Either by correcting its outputs, annotating or over riding its outputs, the experts are feeding information back into the dataset or model enhancement program, completing the loop. The goal of the human-in-the-loop retraining is to enhance robustness, fairness, and generalizability of the model with time. Supporting concept drift detection is also made possible through adaptive feedback loops, for instance, in the case of the evolving conditions (e.g., consumer behavior, regulations, or market shifts), the model has to be updated. This means that HLAI system is developed alongside with the environment in which it is operating and is man-guided [15].

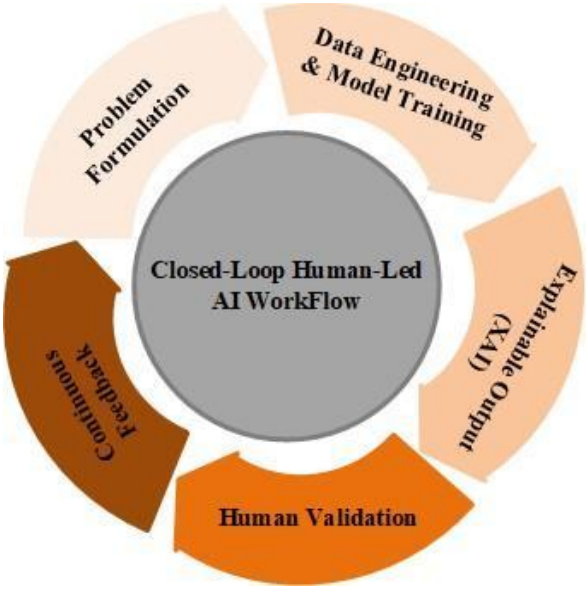

Figure 1: Human-Led AI Closed-Loop Workflow

1.3 Data-Driven culture as cultivated within HLAI

A culture conducive to Human-Led AI (HLAI) must be built in both technical and cultural ways. This section provides some major strategies that help organizations exploit the power of data without losing vital human judgment and accountability.

1.3.1 Investment on data literacy

One of the first steps to a data-driven culture is to make sure that data and AI output are more than capable of being handled by all workers, not just by data scientists. Data literacy initiatives at organization level will debunk the black box nature of AI, make people less technophobic and able to judge the accuracy and reliability of model outcomes instead of blindly believing them [9]. By training the employees on the main principles of statistics, detection of possible biases, and interpretability of models, they are more likely to detect instances when the outputs misunderstand the logic or ethics of the domain. The upskilling opportunities must be role related so that it is relevant within the business functions such as marketing to compliance.

1.3.2 Promote cross- functional teaming

Data science is not an island; it flourishes when a part of interdisciplinary teams. The successful analytics projects unite data scientists, domain experts, operational leaders and policy stakeholders. This partnership narrows the divide between technical competence and real world applicability making them relevant and accurate. The cross-functional teams facilitate fast feedback and expose not-data capturing tacit knowledge and make the AI outcome together. Making such collaborations institutional (using co-located teams, joint project ownership, and shared KPIs) serves to avoid the danger of AI solutions being out of touch with the organization.

1.3.3 Put a Strong Data Governance in Place

Reliable AI is constructed by reliable data. Such stringent data governance structures form the grounds of data integrity, data security and adherence to regulatory laws. This consists of establishing the ownership of the data, imposing retention and retention rules, making certain anonymization is suitable, and making sure that practices concur with legislations, like the EU General Data Protection Regulation (GDPR). Metadata should also be governed, auditable and traceable to enable explainability and accountability within AI workflows. Transparent structures of stewardship avoid the problem of data silo and reduce the possibility of creating a data- byproduct culture and become a strategic asset-driven culture [8].

1.3.4 Implement explainable AI (XAI)

Since AI is increasingly becoming a part of high-stakes decision-making, there is no longer an option of explaining. It is a necessity. Human stakeholders are able to comprehend the factors affecting the decisions of a model and why with the help of XAI methods like SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-agnostic Explanations) and counterfactual analysis, where the relevant negative and positive factors of the model are presented in relation to various human stakeholders [16]. XAI helps with user trust and drives transparency needs and compliance audits in regulated industries such as finance, health care and law to name a few. To the extent that model outputs can be understood, human decision-makers may find it easier to confirm, challenge, or supersede them, and therefore, the human-driven factor of AI.

1.3.5 Embed an ethical oversight

Ethical governance is necessary in the proactive implementation of responsible AI. The formation of periodic AI ethics decisions and various oversight committees will help prevent and raise the presence of risks linked to bias and fairness and social harm [16-18]. Red-teaming: using antagonistic teams to identify blind spots in AI systems, or attempt to provoke them into demonstrating discriminatory behaviors, can be done prior to live AI systems being deployed. Having ethicists, social scientists, and marginalized communities in the development cycle enhances diverse stakeholders in the AI solutions. Ethical oversight should not be considered a one-time audit but should be built into the R&D as well as post-deployment monitoring.

2. Methodology

The proposed study will use a conceptual research design, as it aims at building and explaining the Human-Led AI (HLAI) framework. Instead of verifying a hypothesis using primary data, the study incorporates the experience of various scholarly, industrial, and regulatory literatures to create a model that can inform organizations in the process of balancing automation of AI with human decision making. The methodology was based on three broad steps as follows.

2.1 Systematic review

Systematic review was conducted to determine the principles, challenges, and practices associated with human-AI collaboration. The review used peer-reviewed journal articles, industry reports, and documents on policies published between 2018 and 2025. Databases including IEEE Xplore and Google Scholar were used to find the sources with such keywords as human-in-the-loop AI, trustworthy AI, data-driven culture, AI governance, and responsible AI frameworks.

Inclusion criteria were focused on literature that:

a. Discuss AI decision-making implementation in organizations,
b. Talk about human control or the morality of AI, or,
c. Current case studies in high-reward areas (e.g. healthcare, finance).

The step formed the basis of grasping the prospects as well as the threats of implementing AI and its integration with human decision-making.

2.2 Synthesis and framework development

In selecting the needed literature, some recurrent themes like transparency, accountability, adaptability, and ethical oversight were identified and overlaid on the AI life cycle. It was pointed out that while the synthesis focused on a closed-loop process where humans come in during the final part of the stage of decision validation, they also come in during the problem definition, the model training, and the continuous feedback cycles. This step changed the abstract principles into the Human-Led AI (HLAI) model, where the computational efficiency is systematically balanced with the human contextual awareness, moral reasoning and strategic judgment.

2.3 Case study analysis

To be able to base the framework on real-life practice, two high-impact areas were chosen to analyze the cases: financial credit risk modeling and healthcare triage. These industries were selected since they (a) deal with stakes decision-making directly related to social outcomes, (b) are under intense regulatory scrutiny and (c) already feature active debates on the topic of algorithmic bias and explainability.

The healthcare scenario discusses the application of LLMs to emergency department triage, comparing the results of the models with that of clinical experts. The finance case takes into account the internal audits of AI-based credit risk models, which brings out the issue of fairness and compliance. Both cases were examined in a qualitative manner with the emphasis on the way in which HLAI principles, including explainability, human override, and feedback loops, are reflected in practice.

The combination of the methods of gathering systematic evidence (literature) and empirical illustrations (case studies) helps to establish a strong basis of the suggested HLAI framework. The method proves that the HLAI has not only theoretical foundations but it can be applied in the areas where trust, fairness, and accountability are paramount.

3. Case studies

3.1 Healthcare

A recent cross-sectional article in Scientific Reports compared the two versions of ChatGPT to determine the effectiveness of ChatGPT-4 performing triage tasks in a high-capacity emergency department, relative to a validated instrument, the Emergency Severity Index (ESI) [19]. The sample was representative of 745 consecutive real-world patient through 24-hour observation of 18 and older adult patients at tertiary academic hospital in Ankara, Turkey. The classifications of the triage of each patient were determined both separately and in parallel by a jointly unified professional group of physicians of five people of the expert commission (EC) and ChatGPT. Gold standard used was the median ESI of the EC assignment. Surprisingly, compared with ChatGPT and EC decisions, an overall agreement of Cohen Kappa of 0.659 suggests that ChatGPT is accurate, even at the highest ESI level, with accuracy of 76.6%. Interestingly, ChatGPT model had the highest accuracy in predicting the high-acuity patients (ESI-1 and ESI-2) with a Cohen Kappa of 0.828, specificity of 95.63 and negative predictive value of 98.17 in such critical cases. Although these findings are promising, 12.6% of ChatGPT triage decisions were under-triaged (lower than the expert assessment), which is a matter of safety concern on rare or ambiguous conditions. On the other hand, ChatGPT had 10.7 percent occurrences of over-triaging, which may arise inefficient utilisation of resources when applying it in practice. Nevertheless, such bi-directional tendency is not unusual in the early-phase clinical AI systems.

This work presents the usefulness of the closed-loop, human-in-the-loop triage workflow in the context of HLAI: Expert clinicians on the use of ESI led to Problem Formulation.

a. The model output (triage prediction offered by ChatGPT) was explainable and auditable.
b. Human-based validation entailed a professional assessment by a panel of clinicians.
c. Feedback Loops are insinuated by mentioning the limitations of the model as well as re training requirements.

These results indicate that by reframing ChatGPT as a decision-support system, but not a decision- maker, it can supplement the actual work of triage because it completes preliminary decision- making quickly, but clinicians have time to interfere, confirm, and correct, when the case is at the edge. The research advocates a hybrid system in which speed and scale will be achieved with the help of AI, whereas safety, ethics, and accountability will be provided with human supervision. The authors correctly warn that even ESI in its turn is human-purposed and comprises subjective thresholds that AI might not be able to interpret on a regular basis. They also suggest that triage requirements should co-evolve with AI application, and future triage systems must have designs with the interpretation of AI in mind which is another critical aspect to research and develop. This case study demonstrates the exemplary way to employ AI potential in healthcare and combine its advantages with clinical safety and ethical protection.

3.2 Accounting

The AI models, especially those employed in the credit risk scoring, have become dominant in credit approval in the financial services industry underwriting. Nevertheless, such models commonly work with complicated historical data which can be biased in terms of gender, race, generational income geography, or even type of employment. Another element that can be seen as a significant challenge is that algorithmic models become translators of structural inequalities when they operate in a black box of the decision-making process. In this regard, a large multinational financial institution undertook the internal audit of its Gradient-Boosting Machine (GBM) model consumed in consumer credit approvals. It was triggered by witnessed disproportionate rates of approval among the demographic categories that are entitled to protection irrespective of the risk criteria.

Another source of information used in the audit procedure was Deloitte Argus and GRAPA tools, which are expert in the field of cognition risk examination and target benchmark comparison [20]. The review team was comprised of a member of a cross functional task force: data scientists, risk officers, compliance experts and ethics advisors. The team applied the counterfactual fairness testing, where a question such as, “What would happen to the outcomes of the model on holding all other factors unchanged, had the applicant had different protected attributes (e.g., gender or ethnicity)?” is one of the questions being asked.

As part of the Helix GLAD and Comply model proposed by EY, which emphasizes on detecting anomalies in the journal entries, this financial institution has created the internal utilities so that they identify and flag bias indicators, including female representation with consistently low approval rates in a particular income range. There was also a human override system whereby, the loan officers were to involve themselves when the outcome of the models did not appear to match the ground-level financial realities. As a result of such intervention:

a. The difference in approval rates between genders and ethnicity decreased by 12 % relative to a compound fairness index.
b. Default rate was steady, which indicated no violation of risk control as a result of improving equity.
c. Internal audit records were more explainable and traceable as they passed the regulatory standards of scrutiny.
d. The case highlights some of the principles of the Human-Led AI (HLAI) system:
e. Transparency with the help of audit tools such as Argus and GLAD.
f. Human override checks points.
g. Accountability that was achieved as a result of fairness metrics being owned by compliance teams.
h. Feedback loops of constant overrides into retraining pipelines of the models.

The further implication is that the financial institutions cannot be guided by the black-box optimization only when making the decisions. Rather, when the human governance is integrated into AI systems, more just, stable, and regulatorily sustainable results become possible. With a further increase in the use of AI in the finance sector, human-AI co-monitoring is not only a good business practice anymore, but an actual business requirement.

4. Challenges and potential research directions

The two case studies in healthcare and financial service demonstrate how HLAI can be used to further improve decision-making by integrating the precision of algorithms with human judgment. But they also show the ongoing obstacles that organizations encounter in implementing HLAI into practice. These are all cultural and technical issues, which cover the organizational inertia, and skill gaps, data architecture, model transparency, and regulation uncertainty. They need to be addressed to help HLAI shift away towards isolated successes into a sustainable paradigm.

4.1 Resistance to change

One of the most widespread obstacles is cultural inertia. A large number of employees used to making decisions based on intuition or authority might not trust or accept AI-based workflows. This resistance can destroy adoption without careful change management strategies. Data storytelling, where technical knowledge is converted into interesting stories, and pilot projects, led by internal champions, can assist organizations to get first wins and trust in HLAI systems.

4.2 Skill Gaps

A workforce that is able to interact with AI outputs in a critical way is the key to the successful adoption of HLAI. However, the overall AI literacy outside of the data science workforce is low. In order to overcome this obstacle, companies may utilize automated ML (AutoML) systems and citizen data scientist toolkits, which have reduced technical hurdles but allow significant human control. These solutions make AI use more democratic and at the same time, they provide a level of involvement in the oversight and validation by non-technical employees.

4.3 Data Silos

Both human and algorithmic decision-making is limited by siloed data architectures. This is commonly the case with traditional organizational structures that tend to isolate data between departments thereby disincentivizing collaboration and undermining model performance. The solutions that seem to hold a promising future are emerging solutions like data mesh architecture, where the data is considered a product being shared by multiple teams, and federated learning, where the model training is not centralized, and sensitive data are not concentrated in one location. Interoperability is promoted without breaching data privacy and governance in these approaches.

4.4 Model Opacity

Complex models of AI such as deep learning are usually black boxes, their output is hard to interpret. Such lack of transparency compromises trust and makes validation difficult. New approaches, like causal explainable AI (XAI) and counterfactual explanations, concept-based models, and other new capabilities, can reduce the difference between human and algorithmic reasoning through offering intuitive understanding of model behavior. The methods are especially essential when dealing with high stakes environments, where accountability and interpretability are non negotiable.

4.5 Regulatory Uncertainty

The quick development of AI regulation paves the way to uncertainty among organizations. As an example, the EU AI Act categorizes high-risk applications including credit scoring and medical diagnostics as those which must have transparency, documentation, and human supervision. Equivalent models, such as the IEEE P7000 standards are being developed around the world. Compliance strategies should be anticipatory in nature in organizations, as model cards, datasheets, and audit trails can be used to ensure organizations are flexible to new legal and ethical requirements as they emerge.

4.6 Scalable Ethical Assurance

Ethical management of AI is commonly informal and based on an audit-by-audit basis or a periodic audit. To bring AI systems to scale and make them reflect the values of the society, organizations should shift to stability and systematic guarantees. Some of the anticipated solutions are automated bias scanners, synthetic stress testing using edge case, and Ethics-as-a-Service (EaaS) platforms offering external and independent ethical risk analysis. These tools may directly incorporate ethical surveillance into the AI lifecycle, and make sure to comply with the principles of fairness and accountability in the long run.

4.7 Future Research Directions

Though there are solutions to overcome a number of pain points in HLAI adoption, additional research is necessary to improve and scale the paradigm. Two priorities stand out:

a. Interactive XAI: Building interfaces with the capability to pose interaction-what-if question, sensitivity of the model and get human friendly explanations in real time. This interactivity increases trust and gives users more power to interact with AI systems.
b. Governance Automation: Developing active compliance and governance systems, also known as compliance bots, policy engines, AI-based observability tools, etc., that respond to changes in regulations and organizational needs. Such innovations will enable organizations to go beyond rule-checking, which involves stagnant rule-checking to active, adaptive governance.

Overall, the obstacles to HLAI are genuine and not overwhelming. HLAI can be achieved on the ground as a matter of operations, rather than just being a theoretical

model, through cultural change, technical innovation and proactive governance. The conclusion that follows outlines how such a balance between human control and data power can lead to a trustworthy, data-driven future.

5. Conclusion

AI technology and data-driven culture support each other. However, there is a sustainable competitive advantage only when the algorithmic forces are balanced by human judgement. The HLAI framework proposed proves that the critical facilitators of trustworthy AI include transparency, accountability, controllability, and ethical governance. By embracing the solutions and measures outlined in this paper, organisations will be placed in the best position to turn the data fire-hose into responsible business intelligence, driving value both at the societal and commercial level and maintaining an empowerment of humanity.

References

[1]. Liutak, O. Baula, O. Voitenko, I. Protsyk, V. and Hrytsai, O., (2024). Features of the development of the modern digital technology market," in *Economic Forum*, 2024, vol. 14, no. 4: Lutsk, Ukraine: Lutsk National Technical University, pp. 32-46.

[2]. linkurious. "From data-centric to decision-centric: Insights from 2024 Gartner Data & Analytics Summit." https://linkurious.com/blog/gartner-data-and-analytics-2024/ (accessed 01/07, 2025).

[3]. J. Tang *et al.*, "Towards General Industrial Intelligence: A Survey of Large Models as a Service in Industrial IoT," *IEEE Communications Surveys & Tutorials,* 2025.

[4]. D. Gritsenko and M. Wood, "Algorithmic governance: A modes of governance approach,"*Regulation & Governance,* vol. 16, no. 1, pp. 45-62, 2022.

[5]. C. Katzenbach and L. Ulbricht, "Algorithmic governance," *Internet Policy Review,* vol. 8, no. 4, pp. 1-18, 2019.

[6]. S. Lazar, "Governing the algorithmic city," *Philosophy & Public Affairs,* vol. 53, no. 2, pp. 102-168, 2025.

[7]. A. Simkute, L. Tankelevitch, V. Kewenig, A. E. Scott, A. Sellen, and S. Rintel, "Ironies of generative AI: understanding and mitigating productivity loss in Human-AI interaction," *International Journal of Human–Computer Interaction,* vol. 41, no. 5, pp. 2898-2919, 2025.

[8]. D. Mügge, "EU AI sovereignty: For whom, to what end, and to whose benefit?," *Journal of European Public Policy,* vol. 31, no. 8, pp. 2200-2225, 2024.

[9]. S. Alam and M. F. Khan, "Enhancing AI-Human Collaborative Decision-Making in Industry 4.0 Management Practices," *IEEE Access,* 2024.

[10]. S. Wu, Y. Liu, M. Ruan, S. Chen, and X.-Y. Xie, "Human-generative AI collaboration enhances task performance but undermines human's intrinsic motivation," *Scientific Reports,* vol. 15, no. 1, p. 15105, 2025.

[11]. K. K. R. Chinthala, M. S. Thakur, M. Shuaib, and S. Alam, "Prospects of Computational Intelligence in Society: Human-Centric Solutions, Challenges, and Research Areas," *Journal of Computational and Cognitive Engineering,* 2024.

[12]. J. K. A. Oluwaferanmi, "Automating ETL Pipelines Using Artificial Intelligence: Transforming Legacy Data Integration Systems into Intelligent Data Workflows," 2025.

[13]. P. Linardatos, V. Papastefanopoulos, and S. Kotsiantis, "Explainable ai: A review of machine learning interpretability methods," *Entropy,* vol. 23, no. 1, p. 18, 2020.

[14]. M. Balas *et al.*, "Exploring the potential utility of AI large language models for medical ethics: an expert panel evaluation of GPT-4," *Journal of medical ethics,* vol. 50, no. 2, pp. 90-96, 2024.

[15]. Z. J. Wang, D. Choi, S. Xu, and D. Yang, "Putting humans in the natural language processing loop: A survey," *arXiv preprint arXiv:2103.04044,* 2021.

[16]. L. D. Oyeniyi, C. E. Ugochukwu, and N. Z. Mhlongo, "The influence of AI on financial reporting quality: A critical review and analysis," *World Journal of Advanced Research and Reviews,* vol. 22, no. 1, pp. 679-694, 2024.

[17]. M. Constantinides *et al.*, "Implications of Regulations on the Use of AI and Generative AI for Human-Centered Responsible Artificial Intelligence," in *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 2024, pp. 1-4.

[18]. W. Murikah, J. Nthenge, and F. Musyoka, "Bias and ethics of AI systems applied in auditing-A systematic review. Scientific African, 25, e02281," ed, 2024.

[19]. C. Colakca, M. Ergın, H. S. Ozensoy, A. Sener, S. Guru, and A. Ozhasenekler, "Emergency department triaging using ChatGPT based on emergency severity index principles: a cross-sectional study," *Scientific reports,* vol. 14, no. 1, p. 22106, 2024.

[20]. A. Zemánková, "Artificial intelligence and blockchain in audit and accounting: Literature review," *wseas Transactions on Business and Economics,* vol. 16, no. 1, pp. 568-581, 2019.