Human Oversight in Automation-Driven Financial Services

Patricia Marcella Evite

Doctoral Fellow, Marie Skłodowska-Curie Actions (MSCA) DIGITAL
(IRP 9: Audience-Dependent XAI)

Department of Political Science
Università degli Studi di Napoli Federico II (UNINA)

In partial fulfillment of the requirements for

Cluj-Napoca on Foundations of Data Science and Project Management
Assignment from Banca Transilvania on
Paradox of Data Quality: When Accuracy Clashes with Practicality

May 2025

1 Introduction

The history of automation in financial services is a story of continuous innovation, driven by the need for efficiency, accuracy, and customer satisfaction. From the Automated Teller Machines (ATMs) in the 1960s, we now have Artificial Intelligence (AI)-driven systems, such as chatbots and robo-advisors, became increasingly popular, offering customers 24/7 support and tailored financial solutions (Adeyeri, 2024; Kamuangu, 2024). The integration of AI with other technologies, such as blockchain and the Internet of Things (IoT), is expected to further revolutionize the financial services landscape in the coming years (Ahmed et al., 2024; Dhawas et al., 2025).

The technological advancement, however, has also introduced a paradox: while automation enhances efficiency, it raises concerns about a host of problems such as "loss of autonomy and freedom for individuals" (Gehl Sampath, 2021), bias (Mutasa et al., 2024), among others. In 2019, an AI-driven credit limit algorithm from Goldman Sach's was accused of gender bias, granting lower limits to women than men with similar credit profiles (BBC News). Investigations revealed flawed training data and lack of transparency. Another case was when Zest Finance's used machine learning to assess creditworthiness but faced regulatory scrutiny because its models could not fully explain denials (Ennis and Cook). While AI automation enhances efficiency, human oversight is critical to correct biases, ensure regulatory compliance, prevent systemic risk, and maintain ethical standards.

2 The Case for Human Oversight against AI Blind Spots in Developing Countries

The previously mentioned issues are exacerbated in the Global South because of historical and socio-economic vulnerabilities, which include a lack of governance capacity and dependence on technology imported from developed countries.

2.1 Low AI-Readiness and Fraud Vulnerability

Paradoxically, AI has led to increased financial *ex*clusion as those unable to access or understand them remain invisible or worse, get taken advantage of because of lack of digital/AI literacy. For instance, in South Africa, fraudsters have turned to deepfaking their scam operations, using technology to impersonate bank staff with unprecedented similarity, turning uneducated customers into victims and losing as much as EUR 280,000 for one client (Panchia, 2025). In Nigeria, "Yahoo Boys" or romance scammers build trust before stealing

money (Knutsson). If fraud monitoring is solely up to AI, it lacks cultural context and thus miss such schemes that involve social engineering.

2.2 AI Misuse and Debt Traps

AI-driven lending expands access but often prioritizes profit over responsible borrowing. This is possible through the use of alternative data such as social media and phone usage to approve loan within minutes. However, the risk arises when aggressive lending leads to harrassment by recovery or debt collection agents and suicides among over-indebted borrowers. Such is the case in KreditBee in India, where a 22-year old student committed suicide over unpaid loans (Tiwari). Humans are needed as loan officers could access borrower's *real* payment capacity and not just digital behavior. This is much prevalent in Global South because of low-income users. Human judgment is needed to prevent such exploitation.

2.3 Governance as Oversight on a Large-Scale basis

We can also view human oversight in a much larger context, not only aiding or complementing AI decisions, but also how we govern the processes that affect the outcomes of our finances. In its current state where AI systems develop faster than regulations, low-income countries suffer from lack of protection locally and at the global digital arena.

2.3.1 Proneness to Algorithmic Colonialism

According to Mohamed et al. (2020), AI is also subject to colonialism. They present three (3) applications of 'algorithmic coloniality' (Mohamed et al., 2020) which is "buil(t) upon data colonialism in the context of the interactions of algorithms across societies, which impact the allocation of resources, human socio-cultural and political behavior and extant discriminatory systems". First is *algorithmic oppression* where AI systems that automate and amplify systemic biases, reinforcing racial, gender, or class hierarchies. Second is *algorithmic exploitation* or the extraction of cheap labor and data from vulnerable groups to fuel AI development. Kenyan and Filipino 'ghost workers' label AI training dataset for 2$/hour while companies profit in tech hubs like Silicon Valley. Last is *algorithmic dispossession*, as in Western-dominated AI governance and infrastructure strip agency from the Global South. As a response, India and South Africa refused to sign OECD AI Principles and G20 Osaka Track due to unfair data flow terms.

2.3.2 Lack of Legal Frameworks

In terms of legal recourse and mandate for human oversight, the EU AI Act (enforcement starting 2026) establishes a legally binding, risk-based framework with strict human oversight requirements for high-risk AI systems (e.g., healthcare, law enforcement). In contrast, most Global South countries are still in early stages, with only Brazil approaching the EU's regulatory rigor, while others rely on voluntary guidelines or draft laws with weak enforcement (Table 1). As seen in the Table 2, the EU explicitly defines high-risk AI and mandates technical safeguards, while Global South regulations are vague or nonexistent. Without mandated human oversight, AI systems in most Global South countries operate with minimal safeguards against algorithmic bias or misuse. Brazil's risk-based approach contrasts with India's sectoral proposals and Indonesia's *laissez-faire* strategy, revealing divergent policy philosophies. Lastly, the EU's rules may become a *de facto* global standard, but in any case, Global South countries need tailored, enforceable frameworks to avoid regulatory dependency or *algorithmic dispossession*.

Table 1: Comparison of AI Regulations: EU AI Act vs. Global South Countries

Region/Country	Binding?	Risk Classes	Transparency	Accountability	Human Oversight	Enforcement
EU AI Act (2026)	Mandatory	Annex III list	Required (high-risk)	Strict liability	• Override • Stop button • Dual verification	€35M/7% revenue
Brazil (Draft)	Mandatory (high-risk)	High/limited risk	Public sector	Legal liability	Required (high-risk)	2% revenue
India (Proposed)	Recommended	Proposed tiers	Suggested	Proposed	Recommended	None
South Africa	Encouraged	No	Encouraged	Partial	Encouraged	None
Argentina	Proposed	In discussion	Proposed	Included	Proposed	—
Nigeria	Suggested	No	Recommended	Mentioned	Suggested	—
Indonesia	Voluntary	No	Encouraged	Weak	Voluntary	—

Note: EU data reflects enacted legislation; Global South data as of mid-2024. "—" = no mechanism.

3 Learning from Tradition

As a recommendation, I present two cases from the Global South where we can learn how to promote a hybrid approach to the automation vs human oversight debate. Instead of looking at the Global South as mere passive recipients of innovation, the following narratives and real-life application of cultural wisdom amidst new technology is insightful.

Table 2: Comparison of Human Oversight Scope in AI Regulations: EU vs. Global South

Region/Country	High-Risk AI Coverage	Specific Oversight Mechanisms	Technical Requirements
EU AI Act (2026)	• Biometric identification • Critical infrastructure • Law enforcement • Education/employment	• Mandatory override • Dual human verification • Fundamental rights assessments	• "Stop button" • Logging of interventions • Real-time monitoring
Brazil (Draft)	• Public sector AI • Healthcare • Criminal risk	• Human review • Right to explanation	• Audit trails • No details yet
India (Proposed)	• Healthcare • Financial services	• Optional review • Sectoral guidelines	• None
South Africa	• None	• Ethical boards	• None
Argentina/Nigeria/Indonesia	• No classification	• Voluntary	• —

Note: EU requirements binding from 2026; Global South data reflects current drafts/strategies. "—" indicates no standards.

3.1 Local Mediation in Digital Community Currencies — Grassroots Economics in Kenya

Grassroots Economics is a Kenyan non-governmental organization (NGO) focused on community empowerment through alternative economic systems, particularly by promoting Community Asset Vouchers (CAVs) and localized trade networks (Sarafu). My previous ethnographic research (Evite, 2024) documents how a commons approach can successfully mediate between blockchain efficiency and community governance needs. As a background, Sarafu's blockchain infrastructure automates critical functions: 1) smart contracts execute demurrage (holding tax) calculations; 2) liquidity pools enable autonomous CAV exchanges; and 3) immutable ledgers record all transactions. Yet as the research revealed, these automated systems created new opacity, being an inherently complex system, as much as (if not more than) AI. Community members, particularly elders and those with limited digital literacy, struggled to understand the mechanisms behind their own issued voucher, that ultimately represented their stock of goods and services.

Grassroots Economics developed, through their two decades of experience, an oversight model combining community 'champions', adaptive governance protocols, and commons-based infrastructure. The champions are local leaders or staff trained to physically verify the transactions during weekly *chama* or local lending group meetings, translate automated

processes into oral or written accounting systems for the group's own verification, and act as technical support. In case of disputed transactions, the system allowed consensus-based overrides along with regular review sessions where rules are explained, debated, and ratified according to their community rules.

3.2 Human-administered XAI — *Griots* in modern-day Senegal and Tanzania

In their fieldwork, Effoduh (2024) explores how culturally embedded intermediaries can enhance the explainability and trustworthiness of AI systems in the Global South. In Senegal, *griots* have served for centuries as oral historians, storytellers, and mediators, translating complex cultural knowledge into accessible narratives for their communities. At present, those who are adept in digital tools, still perform the tradition of local mediation to democratize AI understanding, albeit only limited to explaining how to use digital tools and not explaining decisions behind them as with the technical topic of XAI. Effoduh (2024) also observed local NGO staff in Tanzania acting like "AI explainers", bridging gaps between an AI-powered prenatal ultrasound tool and expectant mothers. Trained midwives and health workers translated technical outputs into culturally resonant explanations (e.g., using Swahili, analogies, and patient-centered dialogue). The NGO's explainers addressed ambiguities in AI assessments (e.g., conflicting fetal measurements) by clarifying limitations and next steps. In this way, the *griots* perform the 'post-hoc' interpretation which makes oversight reactive, explaining after AI acts, rather than preventive, but at the same time, culturally embedded rather than technically focused.

4 The Future is Collaboration

It is not the case that we should select with manual processes or purely AI-governed ones. Per Boston Consulting Group, companies that are able to successfully deploy AI systems in a wider scale dedicate 70% of the business's efforts to "people and processes", which includes human oversight activities (Mills et al., 2025). Research from MIT Sloan also show that although large language models (LLMs) can deliver reliable financial analysis when enhanced with specialized finance modules, human supervision remains essential to clarify complexities and foster the trust necessary for sustaining client relationships (Vereckey, 2024). The presented cases in Africa invite us to think that rather than viewing specific characteristics of Global South as "constraints" on technological adoption, I argue that they represent viable alternatives to Western-centric human-automation paradigms.

References

Toluwani Babatunde Adeyeri. Economic impacts of ai-driven automation in financial services. *International Journal of Scientific Research and Management (IJSRM)*, 12(07): 6779–6791, 2024. doi: 10.18535/ijsrm/v12i07.em07. URL https://doi.org/10.18535/ijsrm/v12i07.em07. Association of Chartered Certified Accountants, Saint Joseph's University, Philadelphia, USA.

Mohammad Nadeem Ahmed, Abhijeet Anand, Mohammad Rashid Hussain, Mohammed Mohsin Ahmed, Imran Mohd Khan, and Mohammad Ashiquee Rasool. Artificial intelligence in fintech: Emerging trends and use cases. In *2024 IEEE 7th International Conference on Advanced Technologies, Signal and Image Processing (ATSIP)*, volume 1, pages 459–464, 2024. doi: 10.1109/ATSIP62566.2024.10638924.

BBC News. Apple's 'sexist' credit card investigated by US regulator. URL https://www.bbc.com/news/business-50365609.

P. Dhawas, V. K. Katankar, J. Gorakh, R. R. Shende, M. Madankar, and T. Sahni. Ai and its transformative role in financial services. In D. Darwish and S. Kumar, editors, *Utilizing AI and Machine Learning in Financial Analysis*, pages 245–266. IGI Global Scientific Publishing, 2025. doi: 10.4018/979-8-3693-8507-4.ch013. URL https://doi.org/10.4018/979-8-3693-8507-4.ch013.

J. O. Effoduh. A global South perspective on explainable AI. *Carnegie Endowment for International Peace*, 2024.

Dan Ennis and Tim Cook. Bias from AI lending models raises questions of culpability, regulation. URL https://www.bankingdive.com/news/artificial-intelligence-lending-bias-model-regulation-liability/561085/. The number of data points used in AI lending models — and a lack of diversity among people creating them — raise red flags.

P. M. Evite. Moneywork and affordances behind community currencies: The case of Sarafu network in Kenya. *RAMICS2024 Conference Proceedings*, 2024.

P. Gehl Sampath. Governing artificial intelligence in an age of inequality. *Global Policy*, 12 (S6):21–31, 2021. doi: 10.1111/1758-5899.12940.

P. Kamuangu. Advancements of ai and machine learning in fintech industry (2016-2020). *Journal of Economics, Finance and Accounting Studies*, 2024. doi: 10.32996/jefas.2024.6.1.3.

Kurt Knutsson. How 'yahoo boys' use real-time face-swapping to carry out elaborate romance scams. *CyberGuy Report*. URL https://tech.yahoo.com/general/articles/yahoo-boys-real-time-face-140022054.html.

Steven Mills, Noah Broestl, and Anne Kleppe. Responsible ai: You won't get genai right if you get human oversight wrong. *Boston Consulting Group (BCG)*, March 2025. URL https://www.bcg.com/publications/2025/wont-get-gen-ai-right-if-human-oversight-wrong. Accessed: 2025-04-05.

Shakir Mohamed, Marie-Therese Png, and William Isaac. Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology*, 33:659–684, 2020. doi: 10.1007/s13347-020-00405-8.

E. T. Mutasa, C. Dhiwwale, and S. S. A. Gopal. Artificial intelligence in developing economies: Unpacking business innovations, prospects, and challenges. *International Journal of Academic Research in Business and Social Sciences*, 14(11), 2024. doi: 10.6007/ijarbss/v14-i11/23529.

Y. Panchia. AI-powered scams drain South Africans' bank accounts—here's how to protect your money. *Daily Maverick*, 2025.

Anil Kumar Tiwari. The dark world of illegal loan apps in india. URL https://www.aljazeera.com/economy/2023/12/25/the-dark-world-of-illegal-loan-apps-in-india. Borrowers are harassed by reps of lending apps who abuse, threaten and blackmail them by accessing their contacts, pics.

Betsy Vereckey. Can generative AI provide trusted financial advice? *IDEAS MADE TO MATTER: ARTIFICIAL INTELLIGENCE*, Apr 2024. URL https://mitsloan.mit.edu/ideas-made-to-matter/can-generative-ai-provide-trusted-financial-advice.