Analyzing AI Hallucinations In Air Canada's Chatbot: An Actor Network Theory Analysis Of AI Driven Misinformation

STS Research Paper
Presented to the Faculty of
School of Engineering and Applied Science
University of Virginia

By

Vishal Kamalakrishnan

February 25th 2025

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISOR

Benjamin J. Laugelli, Assistant Professor, Department of Engineering and Society

Introduction

On February 20, 2024, Air Canada's Artificial Intelligence (AI) Chatbot incorrectly informed customer Jake Moffatt about the airline's bereavement fare policy and stated he could book a full-price flight and apply for a partial refund within 90 days. In reality, Air Canada's bereavement policy requires that bereavement discounts be applied for before booking a flight. This misinformation led Moffatt to book a $1200 ticket following the death of his grandmother, and when he later sought the discount promised by the chatbot, Air Canada refused to honor it (Hawley, 2024). A civil court ruled the airline liable, dismissing its claim that the chatbot was a "separate legal entity," marking a precedent for corporate accountability in AI-mediated misinformation.

Scholars attribute these "AI hallucinations" to technical limitations in large language models (LLMs) such as inadequate training data, poor algorithmic design, and flawed prompt engineering (Fui-Hoon Nah et al., 2023). These studies treat chatbots as isolated tools, focusing on algorithmic behavior rather than AI integration among existing systems. This view overlooks the sociotechnical lens: companies are often pressured to adopt AI rapidly due to industry pressures, leading to security flaws (Constantin, 2024), while government agencies struggle to regulate AI products. By reducing hallucinations to code-level flaws, existing frameworks neglect how institutional priorities shape AI products.

I argue that Air Canada's failed chatbot was a result of the chatbot's actor-network rather than an isolated technical error. I will use Actor-Network Theory (ANT) - a framework that examines the interplay between human and non-human actors and their influence on sociotechnical systems - (Cressman, 2009), I will analyze how Air Canada's actor-network prioritized rapid AI integration at the expense of accuracy, security concerns, and systemic

oversight. This argument offers a deeper understanding of how underlying actor-networks such as corporate strategies, market pressures, and other institutional factors drive AI failures. To support these conclusions, I will draw on legal records, Air Canada's internal corporate policies, and research on LLMs and trends within the market at the time.

Literature Review

The majority of research on AI hallucinations in large language models (LLMs) attributes errors to technical limitations while neglecting the institutional ecosystems that influence systemic failures. The current discourse is that the leading cause of "AI hallucinations" stems from inadequate training data, poor algorithmic design, and flawed prompt engineering (Fui-Hoon Nah et al., 2023; Xu et al., 2025) and several ideas have been proposed to mitigate this.

D. S. Susanto et al. (2023) propose an AI Adaptive Learning System that utilizes predictive modeling and an adaptive learning algorithm which allows the learning model to adapt until it can generate a learning model that meets the user's specifications. D. S. Susanto et al. also propose the idea of limiting AI chatbots to static databases or imposing strict conditions on the information that AI has access to, limiting the possibility of hallucinations but hampering the flexibility of AI products. Ouyang et al. (2024) propose the STRUCTCHEM paper which utilizes a structured reasoning framework that allows LLMs to generate additional context and validate their responses through structured prompting, improving their accuracy by 30% in Chemistry problems. Both these approaches contribute to the scholarly discourse by attempting to solve the "AI hallucination problem" from a technical lens and show marked improvements in accuracy.

However, guaranteeing AI safety utilizing an entirely code-driven lens is fundamentally challenged by Xu et al. (2025).

Xu et al. (2025) mathematically prove hallucinations constitute an innate limitation of LLMs, with error rates scaling inversely to training data specificity. While increasing training data specificity can help mitigate hallucinations, their proof demonstrates that even exhaustive training datasets cannot eliminate hallucinations as uncertainty grows exponentially with the complexity of the prompt. Organizations deploy LLMs despite knowing that it is inevitable that they will fail (from a mathematical perspective at least) and often rely upon built-in guardrails to manage risks. However, these guardrails are frequently not tested extensively due to corporate and social pressures. Rushed deployments - driven by the overuse of AI - are the leading cause of system failures (Constantin, 2024) and corporate innovation imperatives often override accountability standards. These pressures materialized in Air Canada's chatbot incident, where the airline prioritized rapid AI adoption over accuracy and safe testing, later deflecting liability by framing the chatbot as a "separate legal entity" (Garcia, 2024).

Collectively, existing literature shows that the "AI hallucination problem" cannot be solved through technical means alone and gives some reasons why AI products fail in real-world scenarios. However, current scholarships neglect to adequately explain how human and nonhuman actors co-produce AI-related failures because they do not delve deeply into the sociotechnical issues behind these failures. I aim to use Actor-Network Theory (ANT) to trace how actor-networks create AI systems from concept to implementation and the key gaps where this can fail. In my analysis, I will advance current understanding in the scholarly discourse by examining how corporate policies, regulatory bodies, and other institutional networks encode themselves into the performance AI systems. By applying ANT, this analysis advances the

understanding of the Air Canada AI Chatbot's failure as emerging from sociotechnical networks rather than from technical flaws.

Conceptual Framework

To frame my analysis of the failure of the Air Canada chatbot, I will draw on the science, technology, and society (STS) concept of actor-network theory (ANT) which describes how human and non-human actors participate in networks that shape sociotechnical systems. ANT rejects the idea that technologies exist in isolation through the argument of symmetry - both social and technical elements interact as equal "actors" in a shared network. These actors can be human, such as corporate entities and executives, or non-human such as legal policies, user trust, and market trends. ANT describes these systems and the actors within them by studying the "associations between heterogeneous actors" to describe how networks gain and lose power and shape technologies. These "associations" are shaped by a network builder - an entity that aligns these human and non-human actors together to form a cohesive network towards a shared goal (Cressman, 2009).

In the Air Canada case, the network builder is the airline's executive leadership which sought to rapidly integrate AI into their customer service. Their goal was to streamline the airline's operating efficiency, reduce costs, and maintain market competitiveness. I will use ANT to analyse how Air Canada's chatbot became a product of this network by applying ANT's translation framework. Translation occurs in four phases: problematization, where actors define a shared objective, interessement, where actors convince other actors to accept roles, enrollment, where alliances between actor-networks are formalized; and mobilization, where the network remains stable despite internal conflict.

To reconstruct this network, I will rely on court documents, Air Canada's AI adoption plans, news media articles about Air Canada's chatbot, and external market forces. I will trace the translation process that led Air Canada's executive leadership to define the alliances and roles of each actor within the network. I will then identify critical conflicts within the network to show how competing priorities shaped the chatbot's failure. Finally, I will examine how Air Canada redistributed accountability after the failure of the chatbot, using legal and technical actors to deflect responsibility. By applying ANT, I aim to show that the Air Canada chatbot's failure was not a result of mere coding errors but due to Air Canada's executive leadership creating a network where fast deployment of AI products was prioritized over proper safeguards.

Analysis

My analysis of Air Canada using Actor-Network Theory (ANT) examines how the chatbot's failure emerged from the sociotechnical dynamics engineered by the network builder — the airline's executive leadership. The problematization, interessment, enrollment, and mobilization phases of ANT's translation process offer a framework for tracking how the network builder stabilized the actor-network and how conflicting goals finally caused it to destabilize (Cressman, 2009). I argue that Air Canada's chatbot failure was not an isolated technical issue but the result of Air Canada executives prioritizing rapid AI adoption over accuracy and accountability. I begin my analysis by exploring Air Canada's corporate investment in AI as part of its modernization strategy and the importance Air Canada placed on this strategy. I then look at industry-wide trends and pressures for rapid AI adoption and examine Nvidia's role in enabling cost-efficient AI adoption. I will then look at the legal dimensions and examine

Canadian law, highlighting how the Air Canada case set a precedent for addressing chatbot misinformation in court.

Air Canada's Investment In AI: Prioritizing Innovation And Cost Reduction

Air Canada's chatbot emerged from an actor-network that framed AI adoption as a strategic imperative for modernizing customer service and enhancing operational efficiency. This framing was rooted in Air Canada's broader corporate strategy to leverage technological innovation to gain a competitive advantage. In 2019, CEO Calin Rovinescu articulated this vision by stating, "Big data and AI are now a big part of our business" (Garcia, 2024). The airline established Artificial Intelligence Labs to integrate AI into various aspects of its business operations, aiming to position itself as an industry leader in technology. The importance of this investment is reflected in Air Canada's own 2023 Annual Report which lists Air Canada's key initiatives and the consequences if those initiatives fail.

> A delay or failure to identify and devise, invest in and implement certain important initiatives could have a material impact on Air Canada, its business, results from operations and financial condition (Air Canada, 2023).

Among these initiatives was the implementation of new technologies such as AI and automating business processes. Other key initiatives listed in the report included enhancing revenues and reducing costs highlighting that Air Canada's reasons for implementing its chatbot were not only for an improved customer experience. While these goals are standard among large businesses, 2023 represented a vital year for Air Canada where these goals were ever more important due to the COVID-19 pandemic.

The urgency to adopt AI was further amplified by the financial pressures imposed by the COVID-19 pandemic during the problematization phase of the network. According to

Gabriele-Rivet et al. (2024), at the onset of the pandemic, travel volume decreased by 90% and only slowly recovered once travel restrictions were lifted. This is exacerbated by rising fuel costs due to the pandemic, which Air Canada noted as “one of Air Canada’s largest operating cost items” (Air Canada). Air Canada's 2023 annual report also stressed the importance of keeping customer prices to a minimum due to the competitive nature of the airline industry. These pressures forced Air Canada’s leadership to prioritize cost-cutting measures while maintaining competitive pricing for customers and automation through AI was viewed as a solution to this challenge.

This context shaped the interessement phase of Actor-Network Theory (ANT), where Air Canada’s leadership defined AI adoption as essential for maintaining competitiveness during a difficult period. By emphasizing innovation and efficiency over accuracy and accountability, executives minimized concerns about known risks with large language models (LLMs) which can hallucinate responses between 3% and 27% of the time (Hawley, 2024). As Cressman (2009) notes, problematization often involves defining goals in ways that obscure conflicts among actors. In this case, financial concerns for both the business operating side and the consumer side overshadowed the need to deploy a product that was vigorously tested.

The mobilization phase of the network showcases Air Canada’s reluctance to raise prices for the consumer and seek cost-cutting measures through other means. This was due in part to Air Canada’s commitment towards their customers, and prioritizing AI adoption to reduce operational costs could be seen as a rational decision aimed at maximizing overall welfare during a period of economic instability such as the COVID-19 pandemic. However, the network-builder’s omission of safeguards in chatbot accuracy caused the network to destabilize as it led to a discrepancy between Air Canada's commitment towards its customers and its

internal financial struggles. This is reflected in the legal case where Tribunal member Christopher Rivers criticized Air Canada for failing to ensure its chatbot was accurate, stating that the airline “did not take reasonable care” in its deployment (Garcia, 2024). Ultimately, Air Canada's decision to prioritize cost and innovation exemplifies the vulnerabilities introduced into the actor-network during its enrollment stage, when alliances between actor-networks were formalized and led to the eventual collapse of the network in the mobilization stage.

Broader Industry Context: Pressures For Rapid AI Integration

The broader industry context provides critical insight into why Air Canada's actor-network prioritized rapid AI deployment despite known risks. As of 2023, 79% of organizations had integrated AI into their customer experience toolsets (Hawley, 2024). This widespread adoption created competitive pressures for companies like Air Canada to follow suit or risk falling behind. The increasing prevalence of AI framed AI as a vital technology, and many companies often rushed to adopt it amidst security concerns from experts (Constantin 2024). Air Canada’s rival WestJet implemented its chatbot in 2018 (which also had concerns about reliability) (Garcia, 2024) and AI announcements by competitors created a sense of urgency in capitalizing AI, leading Air Canada to accelerate its development of the chatbot.

During the problematization phase of Actor-Network Theory (ANT), Air Canada’s executive leadership defined AI adoption as essential for maintaining competitiveness and meeting customer expectations. Industry trends acted as influential non-human actors that shaped this objective. However, this framing obscured the risks and failure rate of deploying AI systems. Rschmelzer (2023) notes that 70-80% of AI products fail and that organizations often fail to align AI projects with tangible business goals or allocate sufficient resources for testing and maintenance. Rschmelzer also notes that organizations fail to supply AI products with quality

data that represents real-world scenarios leading to the classic case of “Garbage In Garbage Out.” During the interessment phase, Air Canada’s executive leadership utilized the prevalence of AI within the airline industry, citing competitor’s AI products as well as financial promise, to steer the company into investing within AI. Air Canada’s executives justified the large initial investment into AI and rationalized the risks of AI products.

> (Air Canada) believes investing in automation and machine learning technology will lower its expenses" and "fundamentally" create "a better customer experience" (Belanger, 2024).

As shown above, Air Canada’s executives framed AI adoption as highly imperative during problematization and gathered actors towards that goal during enrollment citing competitor actions and market trends. Some may argue that the reason for the failure is inherent within LLMs which can hallucinate responses between 3% and 27% of the time (Hawley, 2024). They may take Air Canada’s argument that since LLMs are non-human and can give false information, that the chatbot’s failure is a result of overreliance of the user and that customers such as Jake Moffart are responsible for validating the chatbot’s answers.

However, this view overlooks the role that Air Canada's executive leadership placed in normalizing the trust that customers have with AI chatbots. The enrollment phase illustrates how these industry trends shaped user expectations and interactions with Air Canada's chatbot. Passengers became enrolled actors by relying on the chatbot for information without questioning its accuracy—a behavior normalized by the abundance of AI products across industries (Hawley, 2024). The airline’s decision to link the chatbot to its official website created a facade of legitimacy and discouraged user skepticism while also conditioning the user to perceive the chatbot as authoritative. And in this case, Jake Moffart did not validate the chatbot’s response

even though it presented a link to the correct bereavement policy on Air Canada's website because he had "no reason" to believe that the page provided was "inherently more trustworthy than its chatbot"(Belanger, 2024). The chatbot's failure highlights how gaps between the user's trust in the AI product and the actual response of the chatbot materialized into consequences for users during the mobilization stage. This dilemma between the prevalence of AI products across industries and the trust that customers place in them underscores the importance of validating AI responses.

Nvidia's Role In Enabling Easy And Efficient AI Deployment

Nvidia played a pivotal role in Air Canada's actor-network by providing the technological infrastructure that made AI adoption more accessible and cost-efficient. As Hwang (2023) explains, Nvidia's advancements in Graphical Processing Unit (GPU) technology significantly reduced computational costs associated with deploying LLMs. Between 2021 and 2023, Nvidia introduced DGX Cloud services that allowed companies to access supercomputing capabilities on a subscription basis rather than investing in costly physical infrastructure. These innovations enabled organizations like Air Canada to integrate AI systems into their workflow without significant upfront investments.

Nvidia's GPUs acted as non-human actors that stabilized the network by aligning corporate goals of AI adoption with better technology and reduced prices. The availability of affordable computing power made AI adoption desirable, reinforcing Air Canada's executive leadership to push towards modernization through AI during the problematization and enrollment stages of the network. This is a classic case of technological determinism causing Air Canada's executive leadership to rush toward adopting AI without being prepared to manage the

complexities of generative AI systems (as noted by Rschmelzer on the high failure rate of AI products).

By reducing computational costs and increasing access to generative AI tools, Nvidia enabled organizations across industries to adopt generative AI technologies that improved efficiency and accessibility for millions of users (Hwang, 2023). For instance, DGX Cloud eliminated depreciation costs associated with physical infrastructure while offering on-demand scalability—features that benefited companies seeking cost-effective solutions, especially during a period of economic uncertainty such as the COVID-19 pandemic. However, Nvidia was not truly altruistic in its advancement of AI and had a significant monetary gain from the rollout of new GPUs, experiencing an 88% increase in share price throughout 2023 (Hwang, 2023). While Nvidia enabled rapid deployment through scalable solutions like DGX Cloud, access to this additional technology did not help in improving the core concerns of the accuracy of LLMs. This gap between the accessibility of AI technology advanced by Nvidia and the dismissal of accuracy concerns of LLMs enabled Air Canada's executive leaderships' investment into the chatbot and contributed to the network's destabilization during the mobilization stage.

Legal Accountability And Setting Precedents For AI Misinformation

This legal ambiguity acted as a non-human actor that shaped how Air Canada's executive leadership defined their objectives during the problematization stage of the network. The absence of clear precedents regarding chatbot liability created a perception of low legal risk, encouraging the airline to prioritize fast AI adoption over validating its responses to avoid spreading misinformation. The representation of Canadian law at the time as it pertains to this case could be seen through the concept of agency. According to the Canada Revenue Agency (2003), an agency relationship exists when one party (the agent) is authorized to act on behalf of another

(the principal) in dealings with third parties. The disconnect in the network was due to this concept of agency being traditionally applied to human representatives. It was unknown at the time how this concept could apply to a non-human actor and this ambiguity allowed Air Canada's executive leadership to formalize alliances between technical and legal actors without clearly defining accountability mechanisms for errors from the chatbot. This ambiguity is also the baseline behind Air Canada's argument that the chatbot represented a "separate legal entity" due to the non-human nature of the chatbot.

The Air Canada tribunal ruling marked a significant departure from earlier cases by holding the company accountable under the principles of "negligent misrepresentation" during the mobilization stage(Garcia, 2024). Tribunal member Christopher Rivers emphasized that companies are responsible for all information provided on their platforms regardless of whether it originates from static pages or automated tools like chatbots (Garcia, 2024). The Tribunal decided that Air Canada's chatbot had met all three criteria of agency: it was deployed with the airline's consent, it provided policy information affecting customers, and was under the control of Air Canada's engineers.

The Tribunal ruling against Air Canada emphasizes the importance of examining these actor-networks holistically rather than blaming failures solely on technical flaws or blaming the user. As more organizations incorporate AI systems into their business, understanding how sociotechnical networks shape AI products is essential for preventing similar incidents in the future.

Conclusion

In conclusion, my analysis of the Air Canada chatbot's failure utilizing Actor Network Theory (ANT) demonstrated that Air Canada's chatbot failure was not an isolated technical issue but a product of a sociotechnical network shaped by competing priorities. From Nvidia's enabling easy and cost-efficient access to Graphical Processing Units (GPUs), industry-wide trends towards AI adoption, internal financial pressures within Air Canada, and ambiguous legal frameworks—each of these components contributed to an environment where accuracy and accountability were sacrificed in favor of efficient AI adoption.

This analysis is significant because it underscores the importance of looking at the socio-technical systems behind why AI products fail rather than blaming purely technical errors. Since AI-related hallucinations cannot be entirely solved utilizing a technical approach alone (Xu et al., 2024), it is all the more important to minimize the external systemic factors that contribute to AI failures. By examining the interactions between human and non-human actors during the translation process, the insight gained from this study can help develop AI products that are resilient from the failures experienced in the Air Canada chatbot.

3622 words

References

Air Canada 2023 Annual Report. Air Canada - Investor Relations. (n.d.).

https://aircanada.investorroom.com/

Belanger, A. (2024, February 16). *Air Canada must honor refund policy invented by airline's*

Chatbot. Ars Technica. https://arstechnica.com/tech-policy/2024/02/air-canada-must-honor-refund-policy-invented-by-airlines-chatbot/

Canada Revenue Agency. (2003, August 28). *Agency*. Government Of Canada. https://www.canada.ca/en/revenue-agency/services/forms-publications/publications/p-182r/agency.html

Constantin, L. (2024, September 19). *Companies skip security hardening in rush to adopt AI*. CSO Online. https://www.csoonline.com/article/3529615

Cressman, D. (2009). A Brief Overview of Actor-Network Theory: Punctualization, Heterogeneous Engineering & Translation.

D. S. Susanto, E. Halim, Richard, A. Gui & Nelly. (2023). Resolving Artificial Intelligence Hallucination in Personalized Adaptive Learning System. *2023 Eighth International Conference on Informatics and Computing (ICIC)*. https://ieeexplore.ieee.org/document/10382000

Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research*, 25(3), 277–304. https://doi.org/10.1080/15228053.2023.2233814

Gabriele-Rivet, V., Rees, E., Rahman, A., & Milwid, R. M. (2024, May 1). *Impact of the COVID-19 pandemic on inbound air travel to Canada*. Government Of Canada. https://www.canada.ca/en/public-health/services/reports-publications/canada-communicable-disease-report-ccdr/monthly-issue/2024-50/issue-3-4-march-april-2024/impact-covid-19-pandemic-inbound-air-travel.html

Garcia, M. (2024, February 20). *What Air Canada lost in "remarkable" lying AI chatbot case*. Forbes. https://www.forbes.com/sites/marisagarcia/2024/02/19/what-air-canada-lost-in-remarkable-lying-ai-chatbot-case

Hawley, M. (2024, April 2). *Exploring Air Canada's AI Chatbot Dilemma*. CMSWire.com. https://www.cmswire.com/customer-experience/exploring-air-canadas-ai-chatbot-dilemma

Hwang, R. (2023, May 2). *Nvidia's role in the new wave of ai*. Michigan Journal of Economics. https://sites.lsa.umich.edu/mje/2023/05/02/nvidias-role-in-the-new-wave-of-ai/

Ouyang, S., Zhang , Z., Yan , B., Liu , X., Choi , Y., Han , J., & Qin, L. (2024). Structured Chemistry Reasoning with Large Language Models (dissertation). arXivLabs. Retrieved from https://arxiv.org/abs/2311.09656

Rschmelzer. (2023, December 26). *Top reasons why AI projects fail*. Cognilytica. https://www.cognilytica.com/top-10-reasons-why-ai-projects-fail

Xu, Z., Jain, S., & Kankanhalli, M. (2025, February 13). *Hallucination is inevitable: An innate limitation of large language models*. arXiv.org. https://arxiv.org/abs/2401.11817