Chapter 9
Generative AI for Fake News Detection in Data Architectures:
From Literature to Challenges and Future Directions

Bahaa Eddine Elbaghazaoui
https://orcid.org/0000-0002-3206-5162
National School of Applied Sciences, Sultan Moulay Slimane University, Beni Mellal, Morocco

Tarik el Moudden
https://orcid.org/0000-0002-6963-6686
Computer Science Research Laboratory, Faculty of Science, Ibn Tofail University, Kenitra, Morocco

Khalid Benabbes
https://orcid.org/0000-0001-9793-2581
ENS, Moulay Ismail University, Meknès, Morocco.

ABSTRACT

The proliferation of fake news threatens data integrity and public trust, challenging the foundations of modern data-driven systems. This chapter explores how Generative AI, particularly large language models (LLMs), can enhance fake news detection and support trustworthy data architectures. We review key techniques in misinformation detection and analyze the emerging role of GenAI in enabling automated content validation, explanation generation, and credibility assessment. The integration of these models into data pipelines introduces new opportunities but also raises concerns such as hallucinations, bias propagation, and ethical risks. We outline

DOI: 10.4018/979-8-3373-5616-7.ch009

the main challenges and propose future research directions to design autonomous, explainable, and resilient systems capable of mitigating misinformation in real time while upholding transparency and data governance principles.

1. INTRODUCTION

The rapid growth of digital content has significantly amplified the dissemination of information across a wide array of platforms, including social media networks, online news portals, blogs, and various communication channels. This digital acceleration has undeniably enhanced accessibility to information and contributed to the democratization of knowledge by enabling users from diverse backgrounds and locations to access content in real time. However, this same phenomenon has also given rise to an unprecedented surge in the spread of fake news. Fake news defined as false or misleading information that is deliberately presented as factual or credible poses a substantial challenge to societies worldwide. It undermines public trust in institutions, media, and authoritative sources, disrupts the public's ability to make informed decisions, and ultimately threatens the stability of democratic and social systems. As highlighted by Ifenthaler et al. (2021), the problem of misinformation is not merely a side effect of digital communication but a pressing concern that requires urgent attention from policymakers, researchers, and technology practitioners. Accordingly, detecting, classifying, and managing misinformation has become a critical issue, particularly for data-driven systems where the quality of input directly impacts the reliability of analytical outputs and policy recommendations.

In the context of modern data architectures and analytical ecosystems, fake news transcends being a simple media concern; it represents a deeper, more systemic issue related to the integrity and quality of data (Aïmeur et al., 2023). Today's organizations and platforms rely heavily on automated data ingestion pipelines, real-time stream processing, and machine learning algorithms to derive actionable insights from massive datasets. If the information fed into these systems is tainted by misinformation or lacks credibility, the entire chain of data-driven operations from analysis to visualization to decision-making can be severely compromised. Ensuring that only reliable and verified data enters these pipelines is therefore critical to maintaining the validity and trustworthiness of the resulting insights. The growing complexity and scale of data systems make manual validation infeasible, creating a pressing need for intelligent and scalable solutions to ensure data fidelity.

In this context, Generative Artificial Intelligence (Generative AI), and particularly the advancements in Large Language Models (LLMs), have introduced transformative possibilities (AlAli et al., 2024). These AI models are designed to not only process and generate natural language but also to understand semantic nuances, detect incon-