Chapter 4
Auditing Generative AI

Asmita Jain
https://orcid.org/0009-0005-6790-2919
University at Buffalo, SUNY, USA

Zayan Shaheen
University at Buffalo, SUNY, USA

Dollie Juneja
https://orcid.org/0009-0008-4455-2348
University at Buffalo, SUNY, USA

Anuja Ulhe
https://orcid.org/0009-0006-7481-8675
University at Buffalo, SUNY, USA

ABSTRACT

The swift growth and implementation of Generative AI systems has created several unique governance and compliance concerns, notably in terms of ethical usage, data protection, accountability, and transparency. This study presents a complete audit approach for monitoring GenAI systems, focusing on critical aspects such as risk assessment, control deployment, and management of compliance. By utilizing known governance frameworks such as COSO & COBIT 5 to provide a systematic guideline for examining the operational integrity, security, & ethical aspects of GenAI installations. The approach for the audit framework we propose focuses on addressing GenAI specific risks such as inaccuracies in AI results, data security vulnerabilities, and a lack of clarity in AI decision-making processes. Our method handles these issues using a mix of automated control evaluations, threat evaluation, and continuous monitoring. The study presents ways for resolving governance challenges, such as changing legal environments and the complexities of keeping AI models in accordance with retention laws.

DOI: 10.4018/979-8-3373-3078-5.ch004

1. INTRODUCTION

In today's digital age Artificial Intelligence (AI), specifically Generative Artificial Intelligence (GenAI), has taken the world by storm, amalgamating a huge following and changing the way we interact with technology across many industries such as entertainment, healthcare, education, and finance. This technological advancement has made everyday life much easier. Giving people new perspectives and methods of successfully completing tasks. GenAI has transcended our interactions with technology allowing for a more efficient and seamless integration. With the help of large language models (LLM), the power behind Generative AI, we are able to produce intricate and elaborate results. GenAI has capabilities that are unachievable due to its instantaneous manner.

The capacity of GenAI to influence many businesses is becoming increasingly clear as it develops. In addition to improving current procedures, the technology is opening completely new avenues for creativity. For example, GenAI is transforming personalized medicine and drug development in the healthcare industry, and it is opening the door to more immersive and engaging entertainment experiences. Automation is significantly increasing efficiency and accuracy in the banking industry by automating processes like fraud detection and client support.

Each year, Stanford publishes an AI index report (Stanford University, 2024) covering promising trends and research capabilities in AI. The 2024 report describes an important moment when AI's influence into public consciousness. The proportion of those who think AI will majorly impact their life in the next 3-5 years have grown from 60% to 66% this year. As AI becomes ubiquitous, there has been a growing investment in the private sector such as healthcare, financial services and cybersecurity, generative AI is largely influencing our daily lives. The Stanford report states two prominent examples that provide illustration on exactly how this change is occurring. In 2023, With the release of more significant AI applications - from AlphaDev (Stanford University, 2024) which makes sorting algorithms more efficient, to GNoME, which facilitates the process of material discovery. Second, Open AI launches two new models o-1 and o-1 mini with human-like reasoning that aims to spend more time computing more complex and challenging reasoning based mathematical problems.

Notwithstanding these developments, the quick uptake of GenAI has sparked serious worries, especially about its potential abuse. Millions of dollars have been lost as a result of scammers' use of generative AI techniques to produce deepfakes that mimic the sounds and appearance of family members. Such abuse emphasizes how urgently a legal framework that guarantees the ethical and responsible use of these technologies is needed.