Chapter 14
Exploring Cloud-Based AI Music Tools and Their Implications for IT and Multimedia Professionals

Gerli Ryza Delos Santos Reyes
http://orcid.org/0009-0002-9404-1286
Pangasinan State University, Philippines

ABSTRACT

Cloud-based artificial intelligence (AI) music tools have rapidly evolved into scalable, production-ready systems embedded within cloud-native infrastructures, fundamentally reshaping multimedia workflows and digital content ecosystems. Leveraging transformer-based architectures, diffusion models, and distributed GPU-enabled environments, these platforms generate full-length musical compositions from minimal prompts and parameter controls. Despite growing commercial adoption, limited empirical research has examined their professional and technological implications for Information Technology (IT) and multimedia practitioners. This study investigates the architectural foundations, workflow transformation dynamics, ethical risk perceptions, and competency requirements associated with cloud-based AI music systems. Using a mixed-method descriptive-correlational design, data were collected from 184 IT and multimedia professionals across academic and industry sectors to examine AI literacy, adoption intention, ethical concern, and workflow integration patterns.

DOI: 10.4018/979-8-3373-6279-3.ch014

INTRODUCTION

Artificial Intelligence (AI) has undergone a profound transformation over the past decade, evolving from rule-based computational systems and narrow machine learning applications into large-scale generative architectures capable of producing outputs that approximate human creativity. Early AI systems were primarily confined to deterministic problem-solving, pattern recognition, and predictive analytics within clearly defined parameters. In contrast, contemporary generative AI systems leverage deep neural networks, large-scale representation learning, and transformer-based architectures to synthesize text, images, code, audio, and music with increasing contextual coherence and stylistic fluency (Goodfellow et al., 2016; Vaswani et al., 2017). This shift reflects not merely incremental algorithmic refinement but a structural transformation in computational capability, marking a transition from analysis-focused systems to synthesis-oriented intelligence. Foundation models trained on massive multimodal datasets now demonstrate the capacity to generate complex creative artifacts that were previously assumed to require exclusively human cognition (Bommasani et al., 2021). As a result, generative AI has begun to redefine production ecosystems across creative, industrial, and academic domains.

Parallel to advancements in AI architectures, cloud computing has fundamentally reshaped the deployment and accessibility of advanced computational systems. The emergence of elastic cloud infrastructures, GPU-accelerated computing clusters, containerized microservices, and orchestration platforms such as Kubernetes has enabled scalable, distributed machine learning systems to operate reliably in production environments (Fox & Patterson, 2020; Burns et al., 2016). Cloud-native architectures facilitate horizontal scaling, automated resource provisioning, real-time inference, and continuous integration/continuous deployment (CI/CD) pipelines. These infrastructure developments eliminate the need for high-cost local hardware while enabling professionals to access powerful generative models via web-based interfaces and application programming interfaces (APIs). Consequently, generative AI systems are increasingly delivered through cloud-based service models, commonly conceptualized as AI-as-a-Service (AIaaS), thereby democratizing access to computationally intensive creative tools (Armbrust et al., 2010). The convergence of AI innovation and cloud infrastructure has created a technological ecosystem in which generative systems function as distributed, service-oriented platforms rather than standalone software applications.

Within this broader generative landscape, music synthesis represents one of the most technically demanding and commercially disruptive applications. Unlike text generation, which primarily operates on symbolic token sequences, music modeling requires capturing complex temporal dependencies, harmonic progressions, rhythmic patterns, melodic contour, timbral texture, and dynamic variation across