Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 6, No.2, June : 2025

GENERATIVE AI WITH BLOCKCHAIN TECHNOLOGY

Shivam K. Yadao, Students, Department of Computer Science Engineering ,Rajiv Gandhi College of Engineering Research and Technology, Chandrapur, Maharashtra, India
Dr.Manisha Pise, Guide, Department of Computer Science Engineering
Tanay R. Tiwari, Students, Department of Computer Science Engineering, Rajiv Gandhi College of Engineering Research and Technology, Chandrapur, Maharashtra, India
Pranjal Zode, Students, Department of Computer Science Engineering, Rajiv Gandhi College of Engineering Research and Technology, Chandrapur, Maharashtra, India
Shubham B. Vaidya, Students, Department of Computer Science Engineering, Rajiv Gandhi College of Engineering Research and Technology, Chandrapur, Maharashtra, India
Sayyad Anasali, Students, Department of Computer Science Engineering, Rajiv Gandhi College of Engineering Research and Technology, Chandrapur, Maharashtra, India

ABSTRACT :
Generative AI has revolutionized content creation across various media formats, including text, images, speech, presentations, and videos. However, challenges such as content ownership, security, and authenticity persist. This paper presents a blockchain-integrated AI system capable of generating multimodal content through five modules: Text-to-Image, Text-to-Speech, Text-to-PPT, Text-to-Video, and Text-to-Text. The system employs deep learning models such as Stable Diffusion, Tacotron2, GPT, and Wikipedia API-based knowledge retrieval to generate relevant content from text prompts. Blockchain technology is integrated to secure ownership, prevent unauthorized modifications, and ensure transparency. This study aims to develop an efficient and scalable AI-powered tool for content generation while preserving authenticity and security. Multimodal content generation has become an essential aspect of digital media, with applications spanning education, entertainment, and business presentations. Traditional content creation processes require significant time and expertise. This paper explores a novel AI-driven system that automates content generation through five core modules: Text-to-Image, Text-to-Speech, Text-to-PPT, Text-to-Video, and Text-to-Text. By leveraging deep learning techniques, the system can generate images, speech, videos, presentations, and text-based content from textual descriptions. To ensure authenticity and security, we integrate “Blockchain technology” for ownership verification and content protection. This approach enhances scalability, efficiency, and security, making AI-generated content more reliable for real-world applications.
The proposed system features a user interface where a prompt is used to generate a selected media type. Upon successful content generation, a new block is created and added to the chain. This block securely records critical metadata, including the input prompt, a timestamp, and a hash of the generated output file, thus ensuring a permanent and tamper-proof record of the asset's origin. The framework also includes functionality to view the entire blockchain and export its data, demonstrating a practical application for tracking and verifying AI-generated assets.
By combining the creative capabilities of generative AI with the security and immutability of blockchain, this research presents a foundational model for enhancing accountability and trust in AI systems. This approach provides a robust mechanism for managing the provenance of digital assets, which can be extended to address challenges in intellectual property rights, content verification, and the responsible deployment of generative technologies.
Keywords: Generative AI, Google Gemini, Stable Diffusion, ZeroScope, Text-to-Speech (TTS), Wikipedia API, PowerPoint Automation, Multi-modal AI, Natural Language Processing (NLP), Text

Generation, Image Synthesis, Video Generation, AI Presentation Builder, Diffusers, gTTS, Python-pptx, Interactive Content Generation, AI with Blockchain, Deepfake, Content Provenance, IPFS, Smart Contracts.

INTRODUCTION :

In recent years, Generative Artificial Intelligence (AI) has revolutionized the way digital content is created, enabling machines to produce human-like text, realistic images, synthesized speech, and even short video clips. This project harnesses the power of state-of-the-art generative models to build an interactive platform where users can input any topic or prompt and receive dynamically generated content in multiple formats—including text, images, audio, video, and presentation slides. The primary objective of this system is to simplify content creation and knowledge dissemination by automating the generation of multimedia outputs. It uses Google Gemini for natural language understanding and text generation, Stable Diffusion for AI-driven image synthesis, ZeroScope for text-to-video generation, and advanced TTS models for audio narration. Wikipedia's API is also integrated to extract reliable information for generating structured PowerPoint presentations.

However, as AI-generated content becomes indistinguishable from reality, it poses significant societal challenges. The proliferation of "deepfakes"—hyper-realistic synthetic media—threatens to erode public trust, spread misinformation, and enable malicious fraud. This has created an "arms race" between generative models and the technologies designed to detect them. To enhance security, content authenticity, and data integrity, the system is being extended with blockchain-based functionality. Blockchain technology offers a robust solution for content authentication and provenance by creating an immutable, timestamped record of an asset's origin and lifecycle. Although this blockchain module is still under development, the goal is to implement decentralized verification and tamper-proof records for all AI-generated outputs, addressing concerns about trust, plagiarism, and content misuse. This platform not only demonstrates the potential of multi-modal generative AI in education, research, and media but also paves the way for future applications where security and trust in AI-generated content are paramount.

While both Generative AI and blockchain represent distinct technological paradigms, their convergence offers a compelling avenue for addressing critical challenges and unlocking new potentials that neither technology can achieve alone. This paper explores the synergistic relationship between Generative AI and blockchain, investigating how integrating these two fields can enhance the capabilities of generative models, particularly in areas concerning data provenance, intellectual property, security, and verifiable content creation. We delve into the foundational concepts of Generative AI and blockchain, analyze existing intersection points, and propose novel frameworks and applications that leverage the strengths of both.

SYSTEM ARCHITECTURE :

The core components of the system are:

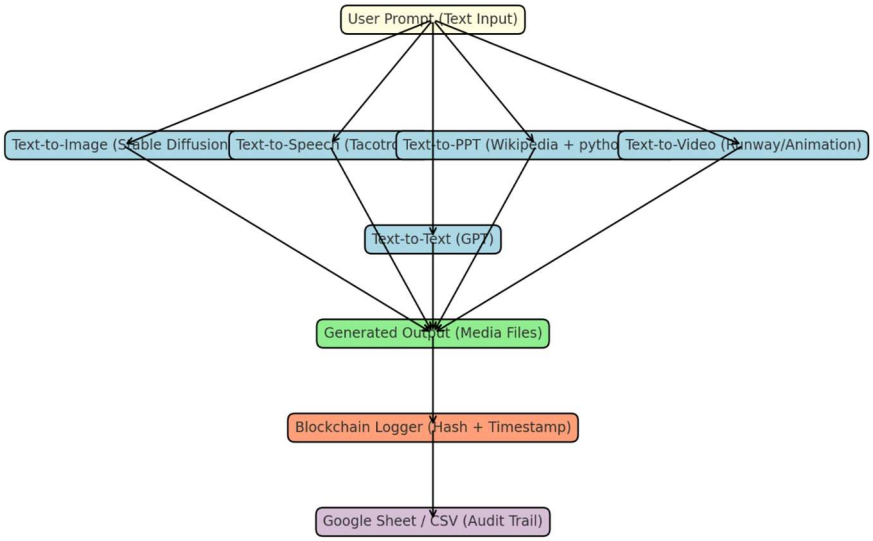
User Interface (UI): The system features an interactive interface built with ipywidgets. This UI enables users to input a prompt or topic, select the desired type of content to generate (e.g., text, image, speech, video, or PowerPoint presentations), initiate the content generation process, and view the historical record of the blockchain.

Generative AI Module: This module is responsible for the actual content creation. It likely integrates with or utilizes various generative models, as indicated by the presence of functions for generating different media types such as generate_text, generate_image, generate_speech, generate_video, and generate_slide_content.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 6, No.2, June : 2025

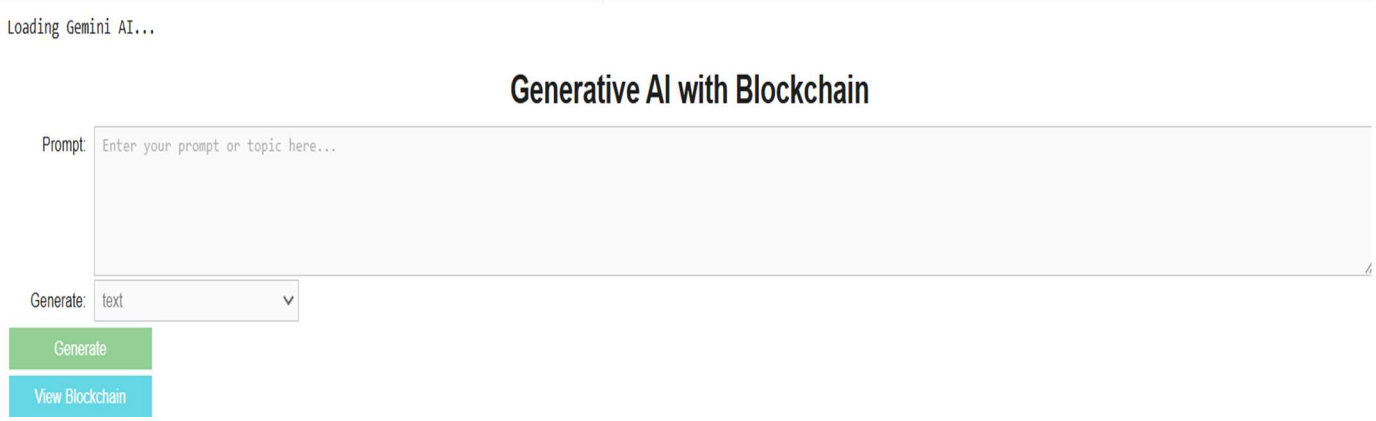
Blockchain Module: A custom blockchain implementation manages the recording of content generation activities. Each successful generation event triggers the addition of a new block to this blockchain, encapsulating metadata like the original prompt and the type of content generated. This ensures a transparent and immutable ledger of all generated assets.
Content Output and Storage: Once content is generated, it is made available to the user. For instance, audio content is saved as a .wav file, and PowerPoint presentations are created as .pptx files, with download links provided in the UI.

The proposed system consists of five AI-powered modules that convert text into various content types. Blockchain is used to log input prompts, timestamps, and output hashes to ensure secure auditing.
The following diagram shows the architecture:

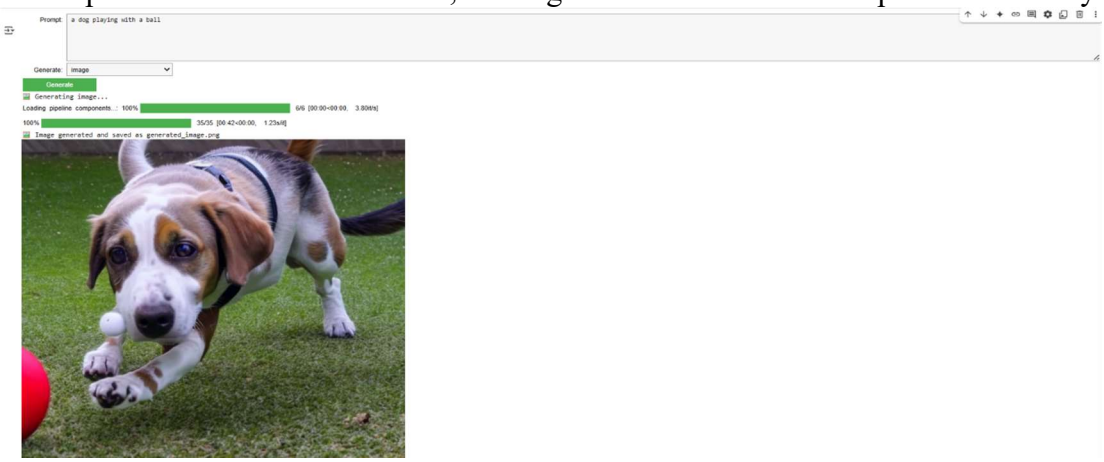


Input and Output
Input: The primary input to the system is provided by the user through an interactive interface. This includes:
Prompt/Topic: A textual description or subject that guides the generative AI in creating content.
Generation Type Selection: The user chooses the desired output format, such as text, image, speech, video, or PowerPoint presentation.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 6, No.2, June : 2025



Output: The system produces several types of outputs based on the user's input:
Generated Content: This is the core output, which can be Textual content, Images, Speech audio files (e.g., .wav), Video files, PowerPoint presentation files (e.g., .pptx), often with a download link provided.
Blockchain Record: Each successful content generation event is recorded as a new block on a custom blockchain. This serves as an immutable log of the generative actions. The blockchain itself can be viewed or exported to an Excel file, acting as a verifiable output of the system's operations.



LITERATURE SURVEY :

Author	Year	Technology Used	Key Finding	Advantage	Disadvantage
OpenAI (Brown et al.)	2023	GPT-4, Transformer Models	GPT-4 demonstrated the ability to generate high-quality, coherent text and code across diverse applications, including text generation for automation.	State-of-the-art language generation, adaptable to various industries	Requires vast computational resources for training and operation

Ramesh et al. (OpenAI)	2022	DALL·E 2, Diffusion Models	DALL·E 2 generated highly realistic images from text prompts, showing impressive fidelity in image-text association.	High creativity in generating visually appealing images from text	Requires fine-tuning to avoid bias and ensure quality control in edge cases
Microsoft Research	2022	DeepSpeech, Tacotron 2, WaveNet	Significant improvements in speech synthesis accuracy and naturalness using Tacotron 2 and WaveNet.	Highly realistic speech generation, capable of handling various accents	Struggles with long-form content and occasional loss of coherence in audio output
Radford et al. (OpenAI)	2021	CLIP, Contrastive Language–Image Pre-training	CLIP efficiently linked image and text representations, improving image and text retrieval tasks.	High performance in zero-shot learning, flexibility in multiple tasks	Vulnerable to adversarial attacks and may inherit biases from training data
Protocol Labs	2020	IPFS (InterPlanetary File System), Decentralized Storage	IPFS enabled secure, decentralized storage of digital content, reducing dependency on centralized data storage.	Scalability and decentralized content storage	Complex for general users and requires improved integration with mainstream blockchain systems
Karras et al.	2019	StyleGAN, Generative Adversarial Networks (GANs)	Introduced a style-based architecture for GANs that improved the control over generated images, especially in facial features.	High-quality image generation, increased control over feature generation	Susceptible to mode collapse during training, computationally expensive

METHODOLOGY :

The proposed system consists of five distinct AI-driven modules that take text prompts as input and generate content across different media types. Each module is implemented using deep learning models suited for its specific task.

Text-to-Image: Utilizes Stable Diffusion, a latent diffusion model that transforms descriptive text into high-quality images. It learns to denoise random noise into meaningful images conditioned on text embeddings.

Text-to-Speech: Employs Tacotron2 for spectrogram generation and WaveGAN for waveform synthesis to convert input text into human-like audio output.

Text-to-PPT: Uses the Wikipedia API to extract content based on a given topic, then leverages the python-pptx library to format and organize it into a well-structured PowerPoint presentation.
Text-to-Video: Integrates video frame generation techniques such as Runway's Gen-2 or MoCoGAN to synthesize motion or animation based on textual input. Recent advancements like the MagicTime model show promise in improving temporal coherence, a key challenge in video generation.
Text-to-Text: Implements transformer-based NLP models like Gemini for text summarization, rewriting, explanation, and question answering.
Blockchain Integration and Logging:
To ensure authenticity, security, and traceability of outputs, a lightweight blockchain-like system is implemented using a **Google Sheets-based ledger**.
Hashing Mechanism: Each user input and corresponding output (file or response) is hashed using the **SHA-256 algorithm**.
Logging Details: The hash, timestamp, module name, and a unique identifier are logged into a Google Sheet using the **Google Sheets API**. This provides a transparent, chronological record of AI-generated activity.
No Wallet Required: This system does not rely on Ethereum, smart contracts, or wallets. It achieves blockchain benefits (immutability, verifiability, transparency) without incurring any costs.

Benefits:

- Easy audit trail
- Tamper-proof record
- Cost-effective for academic and prototype use.

Challenges:
While the integration of generative AI and blockchain offers significant promise, several challenges and areas for future research remain.
Scalability and Cost: Both AI model inference and blockchain transactions can be computationally expensive. While Layer-2 solutions are making blockchains more efficient, the cost of on-chain transactions for a high-volume content generation platform remains a concern. Future work should explore state channels and other scaling solutions to minimize costs.
The Deepfake Dilemma: As generative models become more powerful, the risk of malicious use for creating convincing deepfakes increases. While the proposed blockchain system establishes provenance for *newly created* content, it does not solve the problem of verifying existing, un-logged media. Future systems could integrate advanced detection models, such as those using behavioral profiling, to flag unverified content.
Data Integrity in a Decentralized System: Combining blockchain with decentralized storage like IPFS enhances data integrity. However, ensuring data availability on IPFS requires that at least one node remains online to serve the content. Pinning services or incentive layers like Filecoin are needed to guarantee permanence, adding another layer of complexity.

FUTURE WORK:
Smart Contract Royalties: Develop and deploy smart contracts that automatically manage licensing and royalty payments for AI-generated content, creating a new marketplace for creators.
Decentralized Identity (DID): Integrate DIDs to allow creators to securely and privately own and control their digital identity and the content they generate, without relying on centralized identity providers.

Domain-Specific Models: Fine-tune the generative models for specific applications, such as medical imaging, educational content, or legal document generation, to improve accuracy and relevance.
Integration with Advanced Generative Models: The system can be enhanced by incorporating more sophisticated and state-of-the-art generative AI models. This includes leveraging larger language models for more nuanced text generation, advanced diffusion models for higher-quality image synthesis, and cutting-edge techniques for realistic video and complex multimedia content creation. Exploring new modalities beyond the current scope, such as 3D model generation or interactive experiences, also presents a significant opportunity.
Improved User Experience and Accessibility: Developing a more intuitive and feature-rich decentralized application (dApp) interface would improve user accessibility and broaden the system's adoption. This could include richer visualization of the blockchain, more granular controls for content generation, and seamless integration with existing creative workflows.

CONCLUSION:

This work presents a fully integrated Generative AI system with a blockchain-inspired logging mechanism for secure and traceable media generation. The five modules—Text-to-Image, Text-to-Speech, Text-to-PPT, Text-to-Video, and Text-to-Text—demonstrate the power of AI to automate content creation across formats. The initial prototype using Google Sheets logging provides a cost-effective method for demonstrating traceability. However, the future of trusted generative media lies in more robust, decentralized systems. By integrating technologies like IPFS for storage and smart contracts for IP management, we can build a transparent, secure, and equitable ecosystem for AI-generated content. Future enhancements should focus on scalability, advanced deepfake mitigation, and the implementation of a fully decentralized infrastructure for identity, storage, and royalty management.1
The current implementation showcases the foundational principles and the significant potential of this convergence, future work can explore several avenues. These include scaling the blockchain component for production-level demands, integrating with more advanced and specialized generative models, exploring decentralized storage solutions for the generated content, and developing more sophisticated mechanisms for intellectual property management on-chain. Ultimately, the synergy between Generative AI and blockchain paves the way for a more trustworthy, transparent, and ethically sound ecosystem for AI-driven content creation.

REFERENCES:

1. Brown, T. B., et al. "Language Models are Few-Shot Learners." *NeurIPS*, 2020.
2. Karras, T., Laine, S., and Aila, T. "A Style-Based Generator Architecture for GANs." *IEEE TPAMI*, 2019.
3. Radford, A., et al. "DALL·E: Creating Images from Text." *OpenAI*, 2021.
4. Nakamoto, S. "Bitcoin: A Peer-to-Peer Electronic Cash System." *2008*.
5. Protocol Labs. "InterPlanetary File System (IPFS)." *2020*.
6. Chen, X., et al. "MoCoGAN: Decomposing Motion and Content for Video Generation." *CVPR*, 2018.