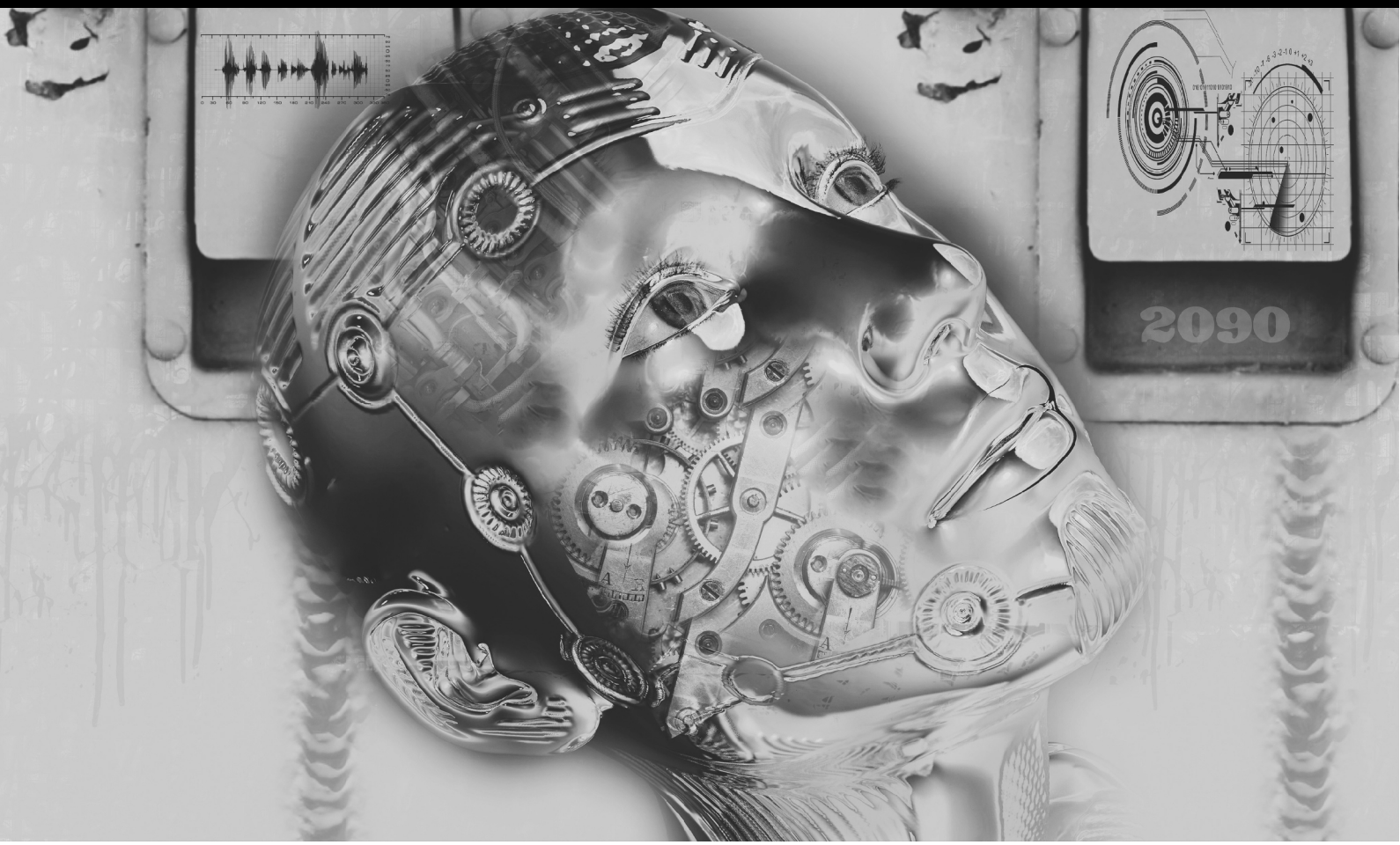
WERKSMANS
ATTORNEYS



DIGITAL MEDIA & ELECTRONIC COMMUNICATIONS

Generative AI and Copyright: Emerging Legal and Policy Questions for Africa’s Creative Industries

25 March 2026 *7 min read*

by Tebogo Sibidla, Director

Generative artificial intelligence (AI) is rapidly becoming one of the most contested issues in copyright law worldwide. At the centre of this transformation is a fundamental question: can copyrighted works be used to train AI systems without permission? Recent developments, including litigation in the US and the UK and recent UK policy reports, highlight the growing tension between AI innovation and the protection of creative works. As Africa's creative industries continue to experience significant global growth, these developments raise important legal and policy questions regarding the use of African content in AI training datasets.

The Global Debate on AI Training Data

Generative AI models are developed using text and data mining techniques, which involve copying and analysing large volumes of digital content to identify patterns. Training datasets often include copyrighted materials such as books, music, news articles and journalism archives, images and films.

The central legal question is whether this use of copyrighted material constitutes copyright infringement, or whether it falls within permissible exceptions such as fair use.

Globally, several broad approaches are emerging:

Approach	Description	Typically Supported By
Licensing Model	AI developers must obtain licences for copyrighted works used in training datasets	Creative industries, rights holders, publishers, collecting societies
Opt-Out Model	AI developers may use publicly available works unless rights holders opt out	Some regulators (historically UK/EU proposals), large AI developers
Broad Exceptions	AI training treated as fair use or equivalent exception	Technology companies, some countries (e.g. the US, to an extent)
Hybrid	Combination of licensing model and transparency mechanisms	Some countries such as the UK and EU

Global Litigation Shaping the Legal Landscape

Courts are beginning to define how copyright law applies to AI training.

In *Getty Images v Stability AI*, Getty alleges that Stability AI used millions of its copyrighted images to train an AI model without permission. The UK High Court dismissed certain copyright claims on jurisdictional grounds, noting insufficient evidence that the relevant acts occurred in the UK. However, the case highlighted the cross-border complexity of AI training, where datasets may be compiled and processed across multiple jurisdictions, and did not finally determine the substantive copyright issues.

In *The New York Times v OpenAI and Microsoft*, the newspaper alleges that millions of its articles were used to train AI systems without authorisation. The case raises important questions about the legality of training on copyrighted journalism, potential market substitution and liability for AI-generated outputs. This case is particularly important for news publishers globally, including in Africa, where digital journalism increasingly relies on subscription and advertising revenue.

Similarly, in *Andersen v Stability AI*, visual artists allege that AI models were trained on scraped images without consent, raising questions about whether such datasets infringe on copyright and whether AI models may constitute derivative works.

These cases are likely to play a significant role in shaping how courts approach AI training globally.

Policy Direction: UK and Global Signals

Recent UK developments provide useful insight into how governments may approach these issues.

In its Report on AI, Copyright and the Creative Industries, the UK House of Lords Communications and Digital Committee adopts a strongly pro-creator stance, emphasising that generative AI systems rely extensively on human-created content, cautions against weakening copyright protections and emphasises the need to ensure that creators are compensated for the use of their works.

The UK Government's Report on Copyright and Artificial Intelligence (March 2026) confirms that there is currently no consensus on how copyright law should apply to AI training and that further evidence is required before legislative reform. It stepped back from its earlier proposal for a broad opt-out exception, following strong opposition from the creative industries. It is now exploring more balanced approaches, reflecting the evolving and contested nature of the issue.

The UK is increasingly focusing on:

- supporting licensing frameworks
- improving transparency in AI training data
- encouraging market-led solutions

This reflects a broader global trend towards balancing innovation with protection of creative industries.

Why This Issue Matters for Africa's Creative Industries

Africa's creative industries have grown rapidly over the past decade and are increasingly integrated into global digital markets.

Music: African music has achieved global prominence, with genres such as Afrobeats and Amapiano reaching international audiences. Artists such as Burna Boy, Wizkid, Tyla and Black Coffee have gained global recognition, while streaming platforms distribute African music worldwide.

Film and Television: Nigeria's Nollywood remains one of the largest film industries globally, producing thousands of films annually. South Africa's television and broadcasting sector continues to produce locally relevant content, distributed through platforms such as DStv. Global streaming service providers are also investing in African content.

Digital Media and Journalism: African publishers such as Nation Media Group (Kenya), Daily Maverick (South Africa) and Premium Times (Nigeria) have built significant digital audiences. Their business models rely heavily on content monetisation through subscriptions and advertising.

As African content gains international reach, it is increasingly likely to be incorporated into global AI training datasets.

This creates a structural risk: African creative works may be used to train AI systems without the knowledge, consent or compensation of creators.

At the same time, African creators and businesses may have limited visibility into how their content is used in AI development, particularly where training occurs across multiple jurisdictions.

Most African copyright regimes were developed prior to the emergence of generative AI and do not expressly address the use of copyrighted works in AI training datasets. Copyright laws in most African jurisdictions do not yet directly regulate text and data mining or AI training.

Strategic Opportunity: African Data as an Asset

While much of the debate focuses on risk, AI training also presents a potential commercial opportunity.

Generative AI systems require large volumes of high-quality data, and AI developers are increasingly entering into licensing agreements with content owners. African datasets that may become particularly valuable include language datasets (given the diversity of African languages), music catalogues (particularly in globally popular genres such as Afrobeats and Amapiano) and media archives (including film, television and journalism).

This suggests that African creators and institutions may be able to participate in emerging markets for licensed AI training data.

What Businesses Should Do Now

The evolving legal landscape requires action from both AI developers and creative industry participants.

AI developers should:

- Ensure that training datasets are lawfully sourced.
- Implement governance processes to track data provenance.
- Monitor regulatory developments across jurisdictions.

Creative industries should:

- Strengthen copyright notices and licensing terms to include an express provision on whether content may be used for AI training.
- Implement technical measures to restrict unauthorised scraping, including tools that allow websites to block AI crawlers or charge for access to content.
- Embed copyright metadata or invisible watermarks into their content, which can help identify unauthorised use, prove ownership and track where works appear online.

- Collaborate through industry bodies to advocate for regulatory frameworks, negotiate licensing frameworks or developing collective rights management models.
- Explore opportunities to license content for AI training.

These steps can help mitigate risk while positioning businesses to benefit from emerging opportunities.

Conclusion

The global debate on AI training data remains unresolved, but its direction is becoming clearer. Governments are increasingly focusing on licensing, transparency and market-based solutions.

For Africa, the key question is how creative industries can both protect their content and participate in emerging data markets. The outcome of this debate will play a critical role in shaping the future of the continent's creative economy.

As global approaches begin to converge around licensing and transparency, early strategic positioning will be critical for African creators and businesses.

Tebogo Sibidla

Director

View Profile