Journal of Intellectual Property and Information Technology Law (JIPIT)
ISSN: 2788-6727 – Volume 4, Issue 1 (2024) p.17-67
https:doi.org/ 10.52907/jipit.v4i1.502

The Cannibalization of Culture: Generative AI and the Appropriation of Indigenous African Musical Works

Michael Dugeri[*]

ABSTRACT

Generative Artificial Intelligence (AI) advancements amplify concerns about the potential to appropriate Indigenous African cultural expressions such as songs, dances, and other forms of art. Generative AI systems autonomously generate diverse content, including music and art, but the supply chain of this new technology presents a complex challenge that may exacerbate cultural appropriation practices. Scholarship on the intersection of technology and Africa's art and culture is animated by the theme of cultural appropriation and the need for protection against commercial exploitation. Likewise, there is a need for more research on how the unique nature of Indigenous African musical works increases their vulnerability to appropriation in the face of entrenched content digitalization practices and the cannibalization of these works as inputs to, and outputs from, generative AI systems. Therefore, this paper attempts to fill this literature gap by exploring the interplay of generative AI training datasets, Indigenous creative works, and the risk of cultural appropriation, with a particular focus on African music. The author argues that if unaddressed, generative AI systems have the potential to significantly erode the data and proprietary rights of various Indigenous communities in Africa, thereby undermining their ability to derive value from the protection of their intellectual property and sustainability of their cultural identity. Through a doctrinal analysis of extant and emerging policy, legal, and regulatory frameworks, this paper establishes the proprietary nature of Indigenous African music and its vulnerabilities in generative AI's supply chain. The author makes recommendations that serve as a vital bridge between technology and cultural integrity, offering a

* The author holds a graduate degree in Law and Technology from the University of Ottawa. His research examines the impacts of emerging technologies, such as AI, on creative endeavours and entertainment in the Global South. Email: mduge039@uottawa.ca

pathway for responsible engagement with Indigenous cultural expressions and respectful utilization of Indigenous African musical works for generative AI systems to safeguard against misappropriation.

Keywords: Generative AI, Training Datasets, Cultural Appropriation, Indigenous African Songs, Intellectual Property

TABLE OF CONTENTS

I. INTRODUCTION

In the year 1939, in the bustling city of Johannesburg, within the heart of South Africa, a Zulu musician named Solomon Linda stood before a microphone, their voice carrying a melody that would transcend time and distance, cutting across continents. They sang *Mbube*, a resonant Zulu song inspired by their childhood encounters with lions while herding cattle (Netflix, 2019; Philips, 2023). Little did they know that this humble recording, rooted deeply in African culture and history, would become emblematic of the struggles and triumphs of African creativity in the face of exploitation.

Fast forward to the mid-20th century, continents away in the United States, a folk group called The Weavers transformed *Mbube* into '*Wimoweh*,' capturing the essence of Linda's song but failing to honor their authorship. Subsequently, another adaptation emerged, '*The Lion Sleeps Tonight*,' shooting to the pinnacle of American music charts. This rendition, crafted by the doo-wop group, The Tokens, led to an avalanche of royalties and commercial success. Yet, the original creator, Solomon Linda, received a meager sum of two United States Dollars for the rights to their masterpiece (Netflix, 2019).

The story of Linda Solomon's song, *Mbube*, is a poignant illustration of how an original creative work can be diluted, losing its traditional and spiritual significance, only to be exploited commercially by powerful entities, all without consideration for the creators or their rightful compensation (Malan, 2020). This narrative reflects a broader cultural misappropriation and exploitation issue, which the data-intensive and proprietary nature of generative AI systems is poised to exacerbate significantly. Despite Linda's undeniable authorship of *Mbube*, various adaptations were made without due credit or compensation (Malan, 2020). The economic windfall primarily benefitted artists, publishers, and record labels in developed countries, while Linda and their family faced poverty and neglect. Furthermore, these

adaptations, despite offering minor modifications to the original song, were used to establish new copyright claims, effectively shutting out Linda's descendants from reaping the rewards of their creativity (Malan, 2020).

Cultural appropriation and the unauthorized dissemination of Indigenous culture have a significant historical presence (Milpurrurru v. Indofurn Pty Ltd, [1994]; Lalani, 2017). A notable instance is Laura Boulton, who, between the 1920s and late 1960s, collected Indigenous ceremonial songs from various global communities (Reed, 2019, p. 125-127). Their activities included recording Indigenous songs in Africa in 1929, which they later presented during public lectures in the United States. In 1933, they recorded Indigenous participants in 'The Indian Villages' exhibit in Chicago. They continued their recordings in the Southwestern United States and other global locations, amassing around thirty thousand song recordings. Boulton licensed these recordings to commercial labels, selling them and making significant gains for themselves up until 1986. It has been noted that while Boulton displayed a deep understanding of copyright law and negotiation, they and their affiliated labels produced numerous copies of Indigenous music without obtaining necessary copyright assignments or licenses from the Indigenous contributors (Reed, 2019).

The cases of Laura Boulton's collection of Indigenous songs and the Linda Solomon's *Mbube* song share a common thread in which well-resourced entities appropriate valuable cultural elements and treat them as raw materials for commercial exploitation, often without acknowledging or compensating the Indigenous communities from which these works originate. Although not premised on the generative AI phenomenon, both cases exemplify the power dynamics and ethical issues surrounding cultural appropriation and the exploitation of Indigenous creative works. As is often the case, well-resourced entities exploit valuable cultural elements commercially while relying on intellectual property law doctrines that may not adequately protect the rights and interests of Indigenous communities.

A recent report by the Australasian Performing Right Association and the Australasian Mechanical Copyright Owners Society (APRA AMCOS) on AI and music examines the risk of AI systems repurposing Indigenous musical works, raising concerns about cultural appropriation. According to the report, AI models trained on culturally significant music may detach these works from their cultural context, using them without proper acknowledgment or compensation to Indigenous communities (APRA AMCOS, 2024).

The report includes perspectives from Aboriginal, Torres Strait Islander, and Māori communities on AI's impact on cultural rights. Indigenous respondents described how AI-driven content creation commodifies cultural elements like music without consent, eroding cultural integrity and misrepresenting traditions (APRA AMCOS, 2024). For Aboriginal and Torres Strait Islander communities, AI's capacity to bypass cultural protocols endangers Indigenous Cultural and Intellectual Property (ICIP) protections, risking economic and cultural losses. Likewise, Māori respondents expressed concerns about the misuse of *taonga puoro* (traditional instruments) and *waiata* (songs) by AI, advocating for Māori-led AI governance to preserve cultural heritage and ensure AI systems respect *tikanga* (cultural protocols) (APRA AMCOS, 2024). These findings echo historical exploitation concerns (Cuthbert, 1998; Inawat, 2015) and highlight the need for suitable protections and Indigenous oversight to maintain cultural integrity and prevent unauthorized use.

Indigenous musical works have been referred to as 'myriad forms, manifestations, or expressions of Indigenous cultural heritage and identity, ranging from stories, folklore, and songs, to dances, rituals, symbols, protocols, and practices' (Nketia, 1974; Nzewi, 1991; Onyeji, 2019). In the African context, these creative works refer to 'the specific musical arts creations of Indigenous African societies with which they celebrate and conduct their social and cultural events and in which various aspects of their cultural lives are woven, documented, and exhibited when needed (Onyeji, 2019, p. 8).

Given the vast diversity of the African continent, Indigenous music encompasses a wide range of expressions and genres. For instance, the polyrhythmic drumming and call-and-response patterns in West African countries like Ghana and Nigeria, the pentatonic scales and unique vocal styles in Ethiopian music, the *mbira* (thumb piano) music of the Shona people in Zimbabwe, and the complex polyphonic singing traditions of the Pygmy tribes in Central Africa (Nketia, 1974; Nzewi, 1991; Onyeji, 2019). Each of these musical forms carries distinct characteristics and cultural significance that contribute to the rich style of African music.

This proprietary material encompasses artistic expressions, innovations, and cultural manifestations originating from Indigenous communities' distinct heritage, traditions, and practices. They collectively link Indigenous peoples with their land, ecological orientation and worldviews, and their past, present and future (Oguamanam, 2017, p. 4). As a result, African musical works often hold deep-rooted spiritual, historical, or communal significance, reflecting the unique identity and worldview of the respective Indigenous group (Nketia, 1974; Nzewi, 1991; Onyeji, 2019; Oguamanam, 2017). Being a cultural expression, it is rightly noted that 'every society has an inalienable right to decide for itself what constitutes music and to what ends such musical constructions are put' (Onyeji, 2019, p. 11).

Cultural appropriation may be defined as the act by a member of a relatively dominant culture of taking a traditional cultural expression and repurposing it in a different context without authorization, acknowledgment, and or compensation, in a way that causes harm to the traditional cultural expression holder(s) (Vézina, 2019). Furthermore, 'appropriation here is understood as taking something that can be claimed by, or proved to belong to, someone else' (Boateng, 2011, p. 188). Some scholars have described cultural appropriation as 'the taking—from a culture that is not one's own—of intellectual property, cultural expressions or artifacts, history and ways of knowledge and profiting at the expense of the people of that culture' (Reed, 2019, p. 111).

Cultural appropriation, especially through generative AI, is harmful because it extracts traditional cultural expressions (TCEs) from their cultural contexts, transforming them into commodities without acknowledgment or compensation to the original communities. This decontextualization erodes the meaning and significance of these works, reducing them to commercial products and ignoring the cultural and intellectual contributions of their creators. Such appropriation continues a pattern of exploitation, where dominant cultures profit from marginalized ones without restitution. While intellectual property and moral rights laws offer some protection, they fall short of addressing the specific risks generative AI poses to Indigenous African music, underscoring the need for enhanced legal safeguards.

While cultural appropriation is primarily a social and cultural issue (Young, 2008; Lockhart, 2021; Arya, 2021), its intersection with emerging technologies, such as generative AI systems introduce pressing legal challenges (Jones, 2013; Matias, 2024). This paper explores these challenges within the framework of intellectual property (IP) law, focusing on the unauthorized use of Indigenous African music. While cultural appropriation itself may not be explicitly prohibited (Scherzinger, 1999; Matias, 2024), generative AI's replication of culturally significant works may infringe copyright protections where Intellectual property (IP) rights apply (APRA AMCOS, 2024). The analysis adopts a socio-legal perspective to assess how the law interacts with cultural appropriation in the digital age, particularly for African musical works. This paper posits that unregulated AI training on cultural data risks perpetuating appropriation, diminishing Indigenous music's cultural and artistic value, and exacerbating economic inequities. While existing frameworks, including copyright and moral rights, provide some protections, they fall short in addressing the scale and complexity at which AI can reproduce and alter Indigenous works. The paper reaffirms the necessity for stronger frameworks that mandate consent, credit, and compensation for Indigenous creators, emphasizing the need for legal safeguards in generative AI development to prevent cultural appropriation.

It is beyond the scope of this paper to undertake a thorough analysis of the thorny issue of copyright in machine-generated outputs to which humans contribute. The paper's primary focus is on understanding how generative AI systems, using diverse training datasets that may include Indigenous musical works, interact with and potentially appropriate elements from such intellectual property and the implications of this interaction regarding cultural appropriation.

This paper acknowledges the multifaceted nature of the topic's cultural, legal, and technological aspects and is cognizant of certain limitations. The author primarily concentrates on Indigenous African musical works within the broader realm of African creative works, acknowledging that various other Indigenous art forms face similar challenges. Furthermore, while the discussion is on the training datasets for generative AI systems and their implications for Indigenous African musical works, broader socio-economic and political contexts influencing these dynamics are acknowledged but not exhaustively explored due to the paper's specific focus.

It is essential to highlight that the term 'Indigenous' is employed in this paper in its colloquial and literal sense, devoid of entanglement in the technical intricacies surrounding terms like 'Indigenous peoples' and 'local communities' about Africa. This obviates the potential of distracting from the paper's focused and straightforward analysis. Furthermore, the term 'Africa' in this paper does not aim to represent the entire continent or encapsulate the diverse perspectives of all African populations. Instead, it seeks to examine specific musical works from various peoples and nations within Africa and to assess their relationship with emerging technologies. This includes creative works such as the *mbira* music of Zimbabwe, the traditional *maskandi* music of the Zulu nation, and the *Apala*, *Akuko n'egwu*, and *Kwagh-hir* of Nigeria. It also considers hybrid genres that combine local and Western musical cultures, such as *juju* and *fuji*.

This research paper is structured into seven parts. Part I introduces the background, defining key concepts and outlining the scope and significance of the research, particularly focusing on the challenges posed by generative AI systems to Indigenous African musical works. It highlights the risks of misrepresentation and appropriation while setting the foundation for the subsequent analysis. Part II examines the context of generative AI, including its technological underpinnings, data mining practices, and the complexities of its supply chain, with an emphasis on their intersection with Indigenous African music. Parts III and IV focus on the specific challenges, discussing the risks of cultural appropriation, the commodification of traditional works, and the limitations of current intellectual property frameworks to adequately protect the communal and oral nature of Indigenous cultural expressions. Parts V, VI, and VII explore responses from other jurisdictions, strategies for safeguarding Indigenous cultural heritage, and actionable recommendations. These sections address the need for tailored content labeling frameworks, ethical licensing systems, and enhanced transparency to ensure proper attribution and equitable compensation for Indigenous creators. The paper concludes by synthesizing its findings and emphasizing the urgency of regulatory reforms to protect the cultural, spiritual, and economic value of Indigenous African music in the age of generative AI.

II. THE CONTEXT

A. Generative AI

Generative AI is a subset of AI and refers to systems and algorithms that can autonomously create or generate new content, data, or outputs similar to those found in a given dataset or domain (Lorenz *et al.,* 2023). This kind of AI, 'enabled by the contributions of millions of humans whose work is included in the training set, often without their knowledge or permission',

(Hayes, 2023, p. 2) can generate original examples or creations that are not mere copies but possess characteristics, patterns, and features consistent with the patterns it has learned during its data training set (Hayes, 2023). Behind this sophisticated technology lies multiple layers of a complicated supply chain architecture.

Generative AI systems are built on what is known as 'Foundation Models' (Bommasani *et al.,* 2021). Foundation models are machine learning models trained on broad data (typically scraped from the Internet), generally using self-supervision at scale (Bommasani *et al.,* 2021; Tran, 2022). It has been noted that 'foundation models have fueled the recent wave of generative AI technologies: these models can be used to generate fluent text, useful code, photorealistic images, and compelling audio' (Bommasani *et al.,* 2023, p. 9). These are extensive machine-learning models designed to serve as versatile building blocks for a wide range of AI applications, including the capability to autonomously create new content, data, or outputs, such as music (Deahl, 2019). Common examples of generative AI for music include *Amper Music*, *AIVA*, *Soundraw*, *Amadeus Code*, and OpenAI's *Jukebox*.

However, achieving this level of sophistication necessitates substantial investment in data mining and curation by AI developers through a process likened to vacuuming vast amounts of data into 'state-of-the-art digital blenders' (King, 2023). As rightly observed,

> 'data to build foundation models is often sourced from the Internet, but this data can only come to be on the Internet as a result of a human data-generating process (e.g. publishing news articles, authoring personal blogs, uploading videos to YouTube, creating music)' (Bommasani *et al*, 2023, p. 9; News Media Alliance, 2023).

The 'rights laundering' potential of these models is well-documented (Wakunuma & Eke, 2024; APRA AMCOS, 2024). There is ample research that shows that:

> 'training data of many foundations and large language models include copyrighted material and material published under licenses with varying permissiveness. It is entirely possible, and many examples have been reported, where copyrighted or licensed material is retrieved or reproduced as the output of a Large Language Model (LLM)' (O'Neill & Connor, 2023, p. 8).

To cement these claims, in a recent parliamentary appearance, OpenAI, the world's most renowned generative AI company, overtly admitted that 'It would be impossible to train today's leading AI models without using copyrighted materials'. It argued that 'copyright today covers virtually every sort of human expression' and cannot be avoided (Edwards, 2024).

To train a machine learning model to aid musicians in creating lyrics, developers of AI music generators typically scrape proprietary lyrics of songs from music websites or other open Internet sources. For instance, Google's *MusicLM* model reportedly underwent training on a vast dataset comprising five million audio clips, equating to two hundred and eighty thousand hours of music (Agostinelli *et al.,* 2023). Although Google has not publicly disclosed the ownership of rights in these datasets, there is a plausible likelihood that a significant portion of it belongs to third-party intellectual property. This inference arises from Google's acknowledgment of concerns about 'cultural appropriation […and] potential misappropriation of creative content associated with its use-case' (Agostinelli *et al.,* 2023, p. 7), highlighting some inherent risks in the model's development (Weidinger *et al.,* 2022; Wakunuma & Eke, 2024; APRA AMCOS, 2024).

Cultural appropriation is a complex phenomenon stemming from various actions, with data collection 'for scientific advancement and innovation' traditionally serving as a common pretext (Oguamanam, 2020; Komminoth, 2023). Extensive research has shed light on the prevalent practice of appropriating Indigenous cultural products under the guise of 'research', subsequently utilizing them as raw materials (Boateng, 2011; Oguamanam, 2020). This allows for their appropriation within the existing framework of intellectual property laws and their transforma-

tion into 'exotic' products for commercial exploitation (Metz, 2023a; Metz, 2023b; Chanda, 2023).

Given the current convoluted nature of generative AI's supply chain and the opaque way it is often developed and deployed, the details of licenses for the underlying training data can create challenges for all parties in the model pipeline. It may also not always be clear who is liable for an infringement (Sunray, 2021, p. 21; Edwards, 2024). This is notwithstanding that since most online content has copyright protections attached at creation, using them for specific purposes could be considered infringement (Henderson *et al.,* 2023, p. 29; Guadamuz, 2023).

Scholarship on the interplay of generative AI, Indigenous creative works, cultural appropriation, and generative AI training datasets has produced many schools of thought and legal arguments (Craig, 2022a), which are beyond the remit of this paper to exhaust. Taken together, however, these legal arguments reflect a complex legal situation that tries to balance innovation, ethical considerations, creator rights, and economic interests (Cuthbert, 1998; Scherzinger, 1999; Lockhart, 2021; Matias, 2024).

However, it is imperative to note that conventional legal intellectual property arguments and theories offer little comfort in the context of legal protection for Indigenous creative works. This is because they often do not reflect the complicated relationship between Indigenous creative works and conventional copyright frameworks, which has repeatedly been shown to be an 'uneasy unfit' as a measure of protection against cultural appropriation. It has been noted that copyright is 'an imperfect fit for combatting cultural appropriation' as it remains 'a porous form of protection' (Reed, 2019, p. 116), allowing unauthorized uses of Indigenous creative works under certain exceptions, notably copyright's fair use doctrine.

To add to this complexity, technology and copyright law have always had a fraught relationship (Craig, 2021). For its part, the law trails behind the advancement of technology almost without

fail and often ends up trying to 'fit round technological pegs into square legal holes' (Sholder, 2023, p. 1). If left unaddressed, generative AI technologies would potentially erode the data sovereignty rights of various creative industries, undermining Indigenous communities' ability to derive value from the protection of their intellectual property (APRA AMCOS, 2024).

It has been astutely observed that the ability to utilize anything within reach, encountered by chance effortlessly, mirrors the ethos of *terra nullius* from the colonial era (Huron, 2017; Zuboff, 2019; McElroy, 2014). This approach aligns with the mindset of conquerors. Still, it is highly questionable in the context of Indigenous artists who consciously depict the narratives, settings, spiritual beings, fauna, flora, and communities they intimately understand and value (Huron, 2017; Wairegi & Melissa, 2023). These artists immerse themselves in and mirror a culture and spiritual essence that has endured, possibly because it predates the Doctrine of Discovery (Huron, 2017).

B. Text and data mining for generative AI systems

AI has long been a part of the music industry, but advances in generative AI systems have brought about new intricacies. Notably, AI is increasingly trained on extensive proprietary raw audio collections (Sunray, 2021). This enables these technologies to capture subtle details in recorded music, such as timbre and dynamics, which allows these technologies to outperform previous systems (Sunray, 2021). For instance, OpenAI's *Jukebox* app is engineered to create music in various genres and artist styles, including rudimentary singing (Heewoo Jun *et al.,* 2020). However, a quick review of *Jukebox's* sample library reveals numerous recognizable adaptations (Sunray, 2021, p. 210). Recognizability is a crucial goal for AI music emulators, as they must ensure that the audience can identify the source of their creations (Sunray, 2021, p. 211).

It is worth reiterating that in most jurisdictions, the copyright in a musical work protects both a song and its essence (Copyright Act 2023 Nigeria, s. 2(1); Copyright Act, 1978 South Africa, s. 2(1); Copyright Act, 2001 Kenya; Copyright *Act of 2005* Ghana s. 1)). As a result, 'when a generative model is engaged to make a new work, the audio that manifests as output is merely a tapestry of up-sampled sound recording fragments manipulated to resemble something ostensibly novel. Thus, characterizing the AI generator's output as 'original' is misleading because doing so disregards the role that reproduction of copyrighted works plays in generating the sample' (Sunray, 2021, p. 193). In other words, the mere 'reassembling of vast swathes of copyright material' should not be enough to strip the original rightsholders of their right to consent, credit, and compensation for their intellectual property usage (Gibson, 2023, p. 275).

C. The generative AI supply chain

It is important to note that 'Generative AI' constitutes a broad spectrum of interconnected technologies rather than a singular product from a specific company. This expansive ecosystem encompasses various technologies, such as music composition and video creation systems (Chen *et al,* 2023). Each generative AI model possesses distinct technical architectures, receives training from diverse data sources, and utilizes different algorithms. Understanding this supply chain is crucial for comprehending how generative AI interacts with proprietary or culturally significant material, especially in the context of this paper's focus on Indigenous African music and cultural preservation.

The term 'generative-AI supply chain' has been used to facilitate a structured understanding of the development of these systems (Lee *et al.,* 2023, p. 23). While the interests of the various groups in the supply chain are divergent, it has been noted that 'in generative AI value chains, control and ownership of data is an issue of particular importance to legal actions that aim

to redistribute financial and data resources between value chain actors such as generative AI developers, AI users, and artists and other creative workers' (Attard-Frost & Widder, 2023).

The initial phase of this supply chain involves creative works, encompassing songs, artwork, software, and other human creative products, which generative AI aims to understand and replicate (Lee *et al.,* 2023, pp. 3-4). Subsequently, these works and associated information, are converted into digital data, represented as digitally encoded files in recognized formats. Individual data units are insufficient for AI training; hence, aggregation is needed into comprehensive training of datasets and meticulously structured compilations of interconnected data. This consolidation process requires a combination of extensive automation and deliberate human decision-making to generate effective training material for AI models (Lee *et al.,* 2023, pp. 3-4).

AI supply chains introduce two key challenges. Firstly, they exacerbate existing issues within AI systems; problems present within individual AI systems are not only perpetuated but often magnified within an AI supply chain. Secondly, AI supply chains complicate extant AI regulatory efforts (Cobbe *et al.,* 2023). For instance, establishing liability for AI-induced infringement or appropriation becomes even more challenging against complex AI supply chains. In the context of this paper, deploying AI through a complex, interdependent, and expanding network of AI systems complicates accountability mechanisms. It heightens concerns related to infringements and cultural appropriation for underrepresented and often less-resourced entities, such as Indigenous African communities (Chen *et al,* 2023).

Generative AI systems' development typically hinges on using extensive datasets for training and other purposes. Unfortunately, developers of such technologies have gained notoriety for their lack of transparency, particularly regarding data collection and usage practices. Recent studies have increasingly focused on the opacity of the generative AI supply chain ecosystems, with a recent study concluding that:

> 'foundation models have rapidly permeated society, catalyzing a wave of generative AI applications spanning enterprise and consumer-facing contexts. While the societal impact of foundation models is growing, transparency is on the decline, mirroring the opacity that has plagued past digital technologies (e.g. social media)' (Bommasani *et al.,* 2021, p. 9; Bommasani *et al.,* 2024).

Top of form reversing this trend was deemed essential, as transparency is a fundamental prerequisite for public accountability, scientific innovation, and effective governance (Bommasani *et al.,* 2021).

The AI value chain plays a pivotal role in determining the outcomes generated by AI systems, consequently influencing the associated risk factors, such as cultural appropriation. As a result, a proposed strategy from a public policy standpoint involves adopting a 'value chain governance' framework to address the potential risks brought about by AI systems (Attard-Frost & Hayes, 2023). Value chain governance describes an approach to accounting for and intervening in the activities through which 'resource inputs are provided to, and resource outputs are received from AI systems' (Attard-Frost & Widder, 2023, p. 6). Therefore, a comprehensive approach to AI value chain governance must intervene in various impacts, such as exploitative and opaque practices observed in data preparation, model development, and labor outsourcing (Attard-Frost & Hayes, 2023).

III. SITUATING THE PROBLEM

Advances in generative AI technology have resulted in unprecedented levels of abundance, access, and duplication of digital content, resulting in an increasing risk of cultural appropriation and declining regard for proprietary materials (Sunray, 2021). This has resulted in content becoming commoditized, raising doubts about its status as property with commercial worth (Sunray, 2021). Consequently, AI developers, operating as commercial entities, perceive that content is so prevalent and ingrained in cultural and digital realms that a significant portion

can no longer be regarded as proprietary (Poritz, 2024). Given this relentless digitization of content and its significant exposure to appropriation, Indigenous communities are slowly recognizing the humbling reality of content digitalization as a pathway to the appropriation of their cherished intellectual property because of its unique nature (Oguamanam, 2017).

As noted earlier, cultural appropriation refers to the unauthorized or exploitative adoption, borrowing, or imitation of elements from another culture, often by a dominant culture. It involves taking aspects of a marginalized or minority culture without understanding or respecting its original context, significance, or history (Boateng, 2011; Reed, 2019). This act can perpetuate stereotypes, dilute cultural meanings, and lead to the commercial exploitation of the appropriated elements (Boateng, 2011).

AI technologies have the potential to assist in preserving and disseminating African indigenous musical works. For example, AI can be used to digitally archive and analyze the drumming patterns of the Yoruba people in Nigeria, helping to preserve these rhythms for future generations. Similarly, AI-driven tools can assist in transcribing and teaching the intricate mbira music of Zimbabwe or the traditional *maskandi* music of the Zulu nation, ensuring that this unique musical tradition is not lost. By recognizing and addressing the diverse forms of indigenous music across Africa, AI applications can be tailored to respect and enhance each unique musical heritage.

The digitization of Indigenous musical works for preservation might inadvertently heighten the risks of appropriation as these digital formats become more accessible and open to misuse or misrepresentation. It takes incredible data to train AI systems to perform specific tasks accurately and reliably. Some of the training data is drawn from material in the public domain, but there are also 'the contributions of millions of humans whose work is included in the training set, often without their knowledge or permission' (Hayes, 2023).

Generative AI systems powered by diverse training datasets might misrepresent or oversimplify the sophisticated cultural nuances present in Indigenous African music. The risk of misinterpretation or misattribution is significant due to the lack of complete contextual understanding. This highlights the urgency and necessity of Indigenous African communities to protect their creative works and, by extension, their culture from unauthorized exploitation.

In January 2023, Google released a research paper (Agostinelli *et al.,* 2023) detailing its *MusicLM*, an AI-driven software designed to compose high-quality music based on text descriptions across various genres (Agostinelli *et al.,* 2023, p. 1). Instructively, the research paper's authors emphasized that the development of such an AI generator carries inherent risks, notably concerns related to 'cultural appropriation […and] potential misappropriation of creative content associated with its use-case' (Agostinelli *et al.,* 2023, p. 7; Fiolet, 2023). It becomes necessary, therefore, to provide some measure of protection for Indigenous content creators against the activities of generative AI systems (APRA AMCOS, 2024).

Using creativity and various means of preserving memories can establish a sense of authority within Indigenous communities. This authority is crucial as it contributes to creating and preserving the resources necessary for Indigenous people to maintain their autonomy and self-determination (Reed, 2019, p. 116). Consequently, when non-Indigenous individuals or entities tell Indigenous stories, sing their songs, or publish their oral histories without proper consent, it erodes Indigenous sovereignty, like the historical injustices of land dispossession and the assimilation of Indigenous people into settler state (Reed, 2019, p. 117; Gausen, 2023).

Generative AI systems hold immense promise in various artistic realms, with their ability to learn from extensive datasets and produce content that mirror human creations. However, as with any dual-use technology, the technology also carries pro-

found challenges regarding the ethical and legal implications of utilizing proprietary or culturally significant materials in training, refinement, and exploitation (Attard-Frost, 2023).

In Africa, the control and provision of digital infrastructure are increasingly driven by foreign entities (Dinika, 2022; Komminoth, 2023; Hlomani, 2023). Despite media portrayals that may suggest a philanthropic interest, the investments made by Big Tech companies in the continent are better characterized as a targeted pursuit of a vast user base, akin to a 'user gold rush'. These companies are driven not merely by benevolent motives but also by the strategic objective of expanding their user bases (Chapdelaine & Rogers, 2021) and sustaining the collection of the diverse and dynamic forms of digital information used by generative AI systems. This includes audio recordings of indigenous music, which are sometimes transmitted orally and evolve through community performances and teachings. It also includes transcriptions, metadata, multimedia content, and ethnographic documentation (Ben-Tal *et al.,* 2019; Kehagia & Moriaty, 2023).

It is argued, therefore, that the attainment of data sovereignty by Indigenous communities in Africa over their creative works is difficult due to this 'digital hegemony,' stemming from heavy reliance on a limited number of foreign-owned tech companies to construct internet infrastructures in Africa (Kukutai & Taylor, 2016; Komminoth, 2023). As rightly observed, 'this resultant 'digital inequality paradox' means that as more people are connected to – or become data subjects of – advanced technologies, the more inequality increases' (Lockhart, 2021; Gillwald, 2021, p. 48; Jennafer & Montoya, 2023; Birhane, 2023).

In analyzing Africa's perception of technology's promise to promote the development and economic prosperity of the continent, scholars have drawn comparisons between Africa's colonial past and current technological trends (Zuboff, 2019; de Souza *et al.,* 2024). Abeba Birhane (2023) asserts that traditional colonial powers aimed for unilateral control and dominance over colonized populations, leveraging social, economic, and political

realms to their benefit (Kwet, 2018; Kwet, 2019; Couldry & Mejias, 2019a; Coleman, 2019; Couldry & Mejias, 2019b; Couldry & Mejias, 2019c; Menon, 2023; de Souza *et al.,* 2024). At present, this control is not exercised through overt physical force, but through more subtle means, such as manipulating digital ecosystems and infrastructure.

This phenomenon can be understood as a form of 'digital neo-colonialism.' The term refers to the exploitation and imposition of digital technologies developed by powerful nations or corporations onto less developed regions, often without equitable benefits for the local populations (Mouton & Burns, 2021; Couldry & Mejias, 2021; Heeks, 2022). This concept draws parallels to historical colonialism, where powerful nations exploited and dominated weaker territories for economic and political gain. In the context of digital neo-colonialism, advanced AI technologies are introduced and used in Africa under the guise of innovation and development while primarily serving the interests of external entities (Heeks, 2022).

For example, the extraction of data from users in the Global South by corporations in the Global North can be seen as a form of digital resource exploitation (Mouton & Burns, 2021; Heeks, 2022). This dynamic creates a dependency on foreign technology and platforms, perpetuating economic imbalances and limiting the digital sovereignty of less powerful nations. Scholars argue that this new, indirect, yet pervasive form of influence and exploitation maintains and exacerbates existing inequalities while stifling local innovation (Kwet, 2018; Kwet, 2019; Gillwald, 2021, p. 48; Jennafer & Montoya, 2023; Birhane, 2023; de Souza et al., 2024). According to Michael Kwet (2018), 'digital colonialism is a structural form of domination exercised through the centralized ownership and control of the three core pillars of the digital ecosystem: software, hardware, and network connectivity' (Kwet, 2018, p. 2).

This growing scholarship on the complex power dynamics inherent in the development of AI systems in Africa highlights the

potential for technological advancement to perpetuate historical patterns of exploitation and control, albeit in subtler forms (Birhane, 2023, p. 391). Moreover, it emphasizes the need for critical scrutiny and regulatory oversight to guard against the risk of practices that further disadvantage marginalized communities (Birhane, 2023, p. 398).

The intersection of generative AI and Indigenous African musical works amplifies these challenges. Indigenous African music, a rich mix of historical, spiritual, communal, and emotional expressions deeply rooted in diverse cultures and traditions, is at risk of cultural misappropriation in the face of increasing content digital and data-hungry AI systems. The risk arises from the fundamental nature of generative AI processing and generating data based on diverse datasets, but often without a nuanced understanding of cultural, spiritual, and communal contexts. Research shows that 'generative AI systems have tendencies toward bias, stereotypes, and reductionism when it comes to national identities, too' (Turk, 2023).

Moreover, the commercial nature of AI development can prioritize profit over cultural understanding, leading to the exploitation of African-style music without adequate acknowledgment or compensation to the Indigenous communities or artists that own proprietary rights in these works. The misuse or misrepresentation of traditional African sounds and instruments also perpetuates stereotypes, distorts cultural meanings, and disrespects the spiritual aspects associated with them (Turk, 2023).

For companies that develop and deploy AI, the collection of more data to create profitable AI systems rather than the welfare of individual people or communities is often at the top of the agenda (Birhane, 2023). The discourse around 'data mining', 'abundance of data', and 'data-rich continent' shows how much individuals behind each data point are disregarded (Birhane, 2023). This muting of the Indigenous communities behind the data points is indicative of how little attention is given to matters such as people's well-being and consent, which should be

the primary concern if the goal indeed is to 'help' those in need (Birhane, 2023). Furthermore, this discourse of 'mining' people for data is reminiscent of the colonizer attitude that declares humans as raw material free for the taking (Birhane, 2023).

The lack of transparency in developing generative AI and other AI technologies adds another layer of complication to data sovereignty and accountability challenges for Indigenous African communities (Kukutai & Taylor, 2016; Bommasani *et al,* 2021; Cobbe *et al.,* 2023; Bommasani *et al.,* 2024). This opacity hampers the ability of Indigenous communities, many of whom lack the necessary exposure, technical expertise, and resources, to hold AI developers accountable for potential infringements on their intellectual property and data rights. However, AI development should not be a black hole for human agency and responsibility (Chen *et al,* 2023).

IV. INDIGENOUS CULTURAL WORKS AND THE LEGAL PROTECTION CHALLENGES

One of the primary challenges in the intersection of Indigenous cultural works, such as African music and generative AI, lies in the applicability of copyright law. Copyright laws, which often include provisions for fair use or fair dealing, permit limited use of protected material without explicit permission for specific purposes such as research, education, or commentary (Copyright Act, 2023, Nigeria, s. 20; Copyright Act, 1978, South Africa, s. 12; Copyright Act, 2001, Kenya, s. 26; Copyright Act, 2005, Ghana, s. 19). While these provisions may provide a backdoor for the use of Indigenous works in AI development, they also introduce a complex ethical dilemma. Fair use, as Gibson (2023) notes, acts as a 'gatekeeping mechanism' for the unauthorized appropriation of cultural materials (p. 107), especially when applied to generative AI's vast consumption of data.

The issue intensifies when considering the communal and oral nature of Indigenous African creative works. The reliance

on public domain principles or the advancement of technology as justifications for exploiting these works may ignore the communal ownership of such cultural expressions and the need for informed consent (Poritz, 2024; Wakunuma & Eke, 2024). While the digitization of content may facilitate innovation, it risks furthering cultural appropriation, especially when the original creators or communities are not consulted or compensated (Henderson et al., 2023).

The principle that copyright protection begins the moment a creative work is fixed in a tangible form is a cornerstone of intellectual property law globally, and this automatic attachment of rights provides immediate legal protection to creators (Henderson et al., 2023, p. 29; Copyright Act, 2023, Nigeria, s. 2(1); Copyright Act, 1978, South Africa, s. 2(1); Copyright Act, 2001, Kenya, s. 22; Copyright Act, 2005, Ghana, s. 1). This legal framework is designed to encourage the creation of original works by granting creators the exclusive rights to reproduce, distribute, and publicly perform their creations. In essence, it reflects a system built for individual ownership and control. However, this framework becomes problematic when applied to Indigenous creative works, which often do not fit neatly into the individualistic mold of copyright law.

Indigenous cultural expressions are typically communal and passed down orally through generations, evolving in meaning and form. Such works, by their nature, are collectively owned by a community and cannot be traced back to a single creator or fixed in a specific moment. The fluidity and shared ownership of these works are at odds with the rigid structure of copyright law, which is built on the assumption that a creative work has a singular author and a fixed moment of creation (Boggs, 2022; Oguamanam, 2017). This misalignment creates a gap in legal protection that makes Indigenous works vulnerable to misappropriation, particularly in the digital age, where content is easily accessible and used without acknowledgment of its origins.

Furthermore, while copyright law allows creators to control the use of their works, it does not adequately account for the cultural and spiritual significance of Indigenous expressions. These works are often imbued with meanings that extend beyond the economic considerations of copyright law, encompassing the identity, heritage, and collective memory of a community. Therefore, the automatic attachment of copyright protection, while beneficial for individual creators in traditional contexts, fails to capture the essence of Indigenous creativity, which is intertwined with communal rights and responsibilities (Oguamanam, 2017). This inadequacy becomes more pronounced with the rise of technologies like generative AI, which can exploit Indigenous works by ingesting them into training datasets without regard for the cultural contexts from which they originate.

In this sense, copyright's individual-centric framework may inadvertently facilitate the appropriation of Indigenous creative works (Lawal-Arowolo, 2015; Reed, 2019). The automatic protection it provides does not extend to the collective ownership of these works, leaving them vulnerable to being classified as part of the public domain or misused without proper attribution or compensation (Henderson et al., 2023, p. 2). As a result, the misapplication of copyright laws can perpetuate historical injustices, where Indigenous works are systematically extracted and commodified, further eroding the cultural and intellectual sovereignty of Indigenous communities.

While acknowledging that cultural appropriation may not be inherently illegal, the challenges posed by generative AI's replication and potential commercialization of Indigenous African musical works reveal critical gaps in current legal frameworks (Boateng, 2011; Oguamanam, 2017). By framing the problem as a socio-legal issue, it is argued, therefore, that IP law, moral rights, and potential sui generis protections offer avenues for addressing these issues. However, any legal response must balance the protection of cultural heritage with the realities of technological innovation. Generative AI's ability to synthesize and replicate

traditional African musical works without acknowledgment or compensation poses not only cultural but also legal challenges. While cultural appropriation is largely a social issue, existing legal frameworks do address its most exploitative aspects, especially when traditional cultural expressions are commercialized without permission. For example, the Swakopmund Protocol (ARIPO, 2010) and Kenya's Protection of Traditional Knowledge and Cultural Expressions Act (2016) explicitly affirm communities' rights over their cultural expressions, providing legal mechanisms against unauthorized appropriation. Although these frameworks do not universally criminalize appropriation, they recognize its harmful impacts and offer limited protections in cases of misuse.

Despite the protections these frameworks provide, they face significant limitations. Their application is geographically restricted to ARIPO member states and specific national jurisdictions, limiting their reach over cultural appropriation by international AI developers. Furthermore, the reliance on formal documentation of traditional cultural expressions complicates enforcement, as many African cultural works are based on oral and communal traditions. This disconnect complicates legal claims, particularly as generative AI can rapidly reproduce and distribute these works across borders. Nevertheless, these frameworks represent progress in safeguarding traditional cultural expressions, even as further reforms are necessary to address the challenges posed by AI and other advancing technologies.

V. LESSONS FROM OTHER JURISDICTIONS

Policymakers globally have been apprised by content creators and performers whose works are being employed to train AI without their consent, fair compensation, or acknowledgment, often camouflaged as 'research' (CISAC, 2023; Reisner, 2023). This comes from mounting shared concern regarding the impacts of generative AI systems on human creativity and, by extension, on artists and content creators (Attard-Frost, 2023).

In striving to 'future-proof' the law, policymakers are increasingly called upon to establish a stable set of rules and principles that remain adequate against new technologies and changing circumstances (Craig, 2022; Acemoglu & Lensman, 2023; Goujard & Volpicelli, 2023). However, achieving this regulatory goal is challenging, especially considering technology's growing capabilities, which have continually tested intellectual property law (Craig, 2022a), and the inherent difficulties of enforcing legal provisions when the infringing entities do not have a physical presence within a government's jurisdiction (Gillwald, 2021, p. 48). AI, in particular, is compelling the law to adapt and undergo complete re-evaluation (**Geist, 2021;** Hlomani & Ncube, 2023).

Indigenous creative works are unique in their holistic nature, deeply rooted in daily life and cultural identity. However, existing intellectual property frameworks, especially copyright laws, are inadequate in safeguarding this priceless intellectual property, which is continually at risk of appropriation in the face of technological advancements and content digitization. Furthermore, IP rights designed for individualistic ownership conflict with the communal nature of Ingenuous creative works, making it difficult for copyright to fully protect these forms of art (Oguamanam, 2017).

In response to the growing calls for regulation, pushbacks from creative industries, and unrelenting lawsuits—Awad and Tremblay v. OpenAI Inc., 2023; Silverman v. OpenAI Inc., 2023; Alter & Harris, 2023—in various jurisdictions, primarily the US, alleging improper use of substantial amounts of proprietary data and copyrighted materials for training these AI systems, AI developers have consistently argued that generative AI fits within already settled case law on fair use like previous generations of copy-reliant technology, including software reverse engineering, and automated plagiarism detection systems (Reed, 2019; Brittain, 2023; Sag, 2023). AI developers and some scholars have also highlighted concerns regarding the potential negative im-

pact on innovation if strict intellectual property laws are consistently applied to regulate their practices (Poritz, 2024; Wakunuma & Eke, 2024). In a specific case, Google asserted that such lawsuits harm Google's services and undermine the very concept of generative AI (Sag, 2023, p. 106). This challenge presents a significant judicial, policy, and regulatory dilemma, compelling nations worldwide to grapple with effective responses (Craig, 2022; Goujard & Volpicelli, 2023).

The resolution of these complex legal and policy issues is uncertain and would likely remain so in the coming years. The outcomes of the ongoing cases will be closely monitored for future guidance. However, it is concerning that, in the interim, many AI providers are proceeding aggressively, collecting as much data as possible and planning to address potential infringement issues when they arise (Poritz, 2024). This approach exacerbates shared concerns about AI models, particularly their negative impacts on marginalized communities.

It is even more striking that the nature and patterns of the various lawsuits against AI developers in terms of where the lawsuits are being instituted and by whom, have introduced a critical perspective on the intellectual property law challenges posed by generative AI and the differential capacity of well-resourced rightsholders, like big music publishers in western countries, to defend their rights (MUSIC - Z SONGS; and ABKCO MUSIC, INC., v. ANTHROPIC PBC., 2023). It also highlights the need for addressing the disproportionate impact on marginalized communities, including Indigenous groups, who face similar infringements but lack the resources and legal support to protect their creative works from unauthorized use by generative AI systems.

Some AI companies have begun entering into licensing deals with major publishers to use their content for AI model development and to attribute summaries to these publishers (Reuters, 2024; Fairly Trained, 2024). While this gradual shift towards licensing is encouraging, AI companies must license all their data,

not just content from prominent publishers likely to pursue legal action. Furthermore, there needs to be a consistent and organized approach to licensing for all creators, especially from underrepresented communities in developing economies. It is also worth questioning whether these deals are for fine-tuning generative AI already trained on unlicensed data, as starting with exclusively licensed material would be preferable.

Some regions are introducing AI-specific laws, regulations, or guidelines, usually in the form of better privacy and data protection laws, accountability, transparency, and algorithmic decision-making measures (Congressional Research Service, 2023; Office of the Privacy Commissioner of Canada, 2023). However, there is a global near-absence of AI-specific regulations that directly address the use of proprietary works in AI training datasets. While regulatory efforts to address these issues are underway in various parts of the world, such as the European Union (EU), the US, and Canada, Africa has been slower to embark on such legal and policy reforms as it continues to rely on general privacy and data protection laws, consumer protection regulations, and human rights laws, which are inadequate for addressing the unique regulatory challenges posed by AI (Yilma, 2022; CIPIT, 2023; Eke *et al.,* 2023; Ndemo *et al.,* 2023). This makes Indigenous African creative works prone to appropriation and commercial exploitation by data-hungry generative AI systems, whose reach for training datasets is boundless.

For instance, when the European Commission initially proposed the Artificial Intelligence Act (2021), generative AI was not a primary concern for regulators. However, this changed with the recent surge in generative AI systems and their far-reaching impacts on several industries, including entertainment. In response to these developments, the European Parliament made significant amendments to the European Commission's initial proposal. Notably, it introduced specific rules targeting generative AI systems, placing obligations on their providers to use copyrighted training data and to observe enhanced transparency

requirements (Amendment 399, Proposal for a regulation, aa. 28 a & b).

The European Union Artificial Intelligence Act (EU AIA) mandates that providers of general-purpose AI models adhere to transparency obligations concerning the data utilized for training their models. Generative AI is not explicitly defined in the EU AIA, but models capable of generating content such as text and images are categorized as general-purpose AI models (GPAI). According to the EU AIA, GPAI is defined as:

> 'an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market' (Article 3(63)).

Article 52(1)(c) of the EU AIA requires these providers to 'put in place a policy to respect Union copyright law, in particular, to identify and respect, including through state-of-the-art technologies, the reservations of rights expressed under Article 4(3) of Directive (EU)2019/790'. Furthermore, the language in Recital 160 mandates that any provider of a general-purpose AI model must comply with EU copyright law 'regardless of the jurisdiction in which the copyright-relevant acts underpinning the training of those general-purpose AI models take place', potentially extending the reach of EU copyright law.

Similarly, in September 2023, France proposed amendments to the French Intellectual Property Code to address the use of copyright-protected content by generative AI technology (PROPOSITION DE LOI, visant à encadrer l'intelligence artificielle par le droit d'auteur). The proposal introduces four essential obligations, including obtaining authorization from the author or IP rights (IPR) holder before integrating copyright-protected material into an AI system (PROPOSITION DE LOI, visant à en-

cadrer l'intelligence artificielle par le droit d'auteur, a1, 2). As a result, AI providers must verify whether the content they use to train their AI systems is protected under copyright and obtain the prerequisite authorization from the IPR holder in advance (Guadamuz, 2023a). Akin to the EU and France, other countries like the US and Canada have also shown readiness to confront the challenges posed by 'high impact' generative AI systems (Bill C27, 2022, cl 39; ISED, 2023).

To effectively address the risks of cultural appropriation associated with generative AI, African countries must enact and enforce more application-specific and sector-specific AI policies and legislations rather than relying solely on extant general privacy and data protection laws. They must also strengthen the national data governance frameworks necessary for safeguarding data sovereignty of the Indigenous communities whose proprietary data remain prone to third-party commercial exploitation and appropriation (Kukutai & Taylor, 2016).

Africa's stance on AI regulation is nuanced, recognizing both the opportunities and obstacles it presents (AUDA-NEPAD, 2023; Abungu & Muhindi, 2023; Oxford Insights, 2023). Policymakers are responsible for creating effective regulations that minimize the adverse effects of AI implementation while leveraging its capabilities to tackle key developmental issues on the continent. A review of the continent's approach to AI regulation reveals a deliberate effort to balance these aims (Gwagwa, 2020; Hlomani, 2023). As a result, the emerging regulation that applies to AI is primarily in the form of data protection laws, which may not meet the standard of a value chain governance framework .

The status of data governance frameworks in Africa is improving steadily and shows the promise of exploiting the potential of data for the development and empowerment of the continent, but challenges remain (Effoduh *et al.,* 2023). The African Union's (AU) endorsement of the AU Data Policy Framework (2022) and the Malabo Convention (2014) is indicative of AI pol-

icy direction in Africa and shows some political will to invest in data governance frameworks (Asiegbu & Okolo, 2024)

For instance, the AU Data Policy Framework emphasizes the importance of empowering individuals, firms, and governments to control data while highlighting the need for policies that clearly define the obligations and responsibilities of all parties to ensure a balanced approach to data governance (King'ori *et al.,* 2023). The Framework recommends that Member State policies should, at a minimum, establish data subject rights to enable personal data control. It also points to emerging ownership models, such as data trusts and stewardships, as alternatives to the traditional individual-rights-focused model. On a national level, the Framework recognizes data sovereignty and localization as methods through which states currently exercise control over data but warns against implementing both without carefully tailored justifications (King'ori *et al.,* 2023).

At a regional level, various sub-regional data frameworks exist, such as the Supplementary Act on Personal Data Protection (2010) within the Economic Community of West African States (ECOWAS) for West Africa, the East African Community Legal Framework for Cyberlaws (2008) for the East, and the Southern African Development Community (SADC) Model Law (2013) on data protection for the southern part of the continent. Moreover, several African states have established data protection laws domestically, although some are still in the drafting stage or have no legislation yet (Effoduh *et al.,* 2023; Bommasani *et al,* 2021).

VI. PROTECTION OF INDIGENOUS AFRICAN ARTS IN THE ERA OF GENERATIVE AI

Addressing cultural appropriation by generative AI systems necessitates the identification of Indigenous African musical works within AI training datasets. This is essential for assessing the risk of infringement and promoting rules that ensure consent, attribution, and compensation for Indigenous communities. The

oral transmission of many African musical traditions complicates the process, as content labeling frameworks traditionally rely on written records and metadata (Nketia, 1974; Nzewi, 1991; Onyeji, 2019). Effective labeling systems must adapt to this reality, potentially through community-driven documentation efforts that preserve the evolving nature of these works.

Digitalization, while offering increased access to Indigenous works, also heightens the risk of cultural exploitation. AI systems depend on vast datasets, often compiled through web crawlers that collect publicly available content without consent (Hays & Barr, 2023). This practice obscures the origin of appropriated material, creating barriers to accountability. This situation can heighten the challenges in holding infringing entities responsible, as it obscures the origin of cannibalized content, granting plausible deniability to beneficiaries of theft with little incentive to stop such actions.

Sometimes, people and businesses are unaware they are infringing or do not care to avoid infringing (Edwards, 2024). For example, in response to a lawsuit alleging that Google improperly used substantial amounts of personal data and copyrighted materials for training its generative AI, *Bard*, Google argued that 'utilizing publicly available information for learning purposes does not constitute theft, invasion of privacy, conversion, negligence, unfair competition, or copyright infringement' (J.L. v. Alphabet Inc, 2023). Content labeling frameworks that enable clear tracking of ownership can help mitigate these risks by facilitating proactive licensing and reducing unauthorized use (Ferrandis & Hughes, 2023).

Legislative initiatives are underway in certain jurisdictions to foster the traceability of the source materials utilized in training AI systems and potentially preserve the essence and characteristics of these materials within the outputs generated by these AI systems. Similar to the case of France's proposed statutory framework (PROPOSITION DE LOI, visant à encadrer l'intelligence artificielle par le droit d'auteur), earlier alluded to.

The US' Congress is currently considering several bills aimed at addressing concerns related to data scraping and the use of copyrighted works in AI training. Among these proposed measures are the AI Labeling Act (2023) and the AI Consent Act (2024). Additionally, the recently introduced Generative AI Copyright Disclosure Act (2024) seeks to require a notice be submitted to the Register of Copyrights, detailing the use of copyrighted works in the development of generative AI systems and addressing related issues. These bills are still under discussion and have not yet been enacted into law.

If enacted, the AI Labeling Act, on the one hand, will mandate AI-generated systems, producing audio, images, videos, or other multimedia content to display distinct labels or disclosures visibly. For instance, an AI-generated image must exhibit an explicit notification identifying it as AI-generated content, disclosure of the AI tool's identity, and the date and time the content was created.

The AI Consent Act, on the other hand, would require online platforms to obtain consent before using personal data to train AI models. It directs the Federal Trade Commission (FTC) to implement regulations to improve transparency by requiring companies to disclose when an individual's data will be used to train AI and receive consumer opt-in to this use. The Act also provides strong guidelines for these regulations, such as disclosure standards and what constitutes consumer consent.

The Generative AI Copyright Disclosure Act appears to mirror the provisions of Article 52(c) of the EU AI Act. It applies to a person who creates a training dataset or alters a training dataset (including by making an update to, refining, or retraining the dataset) in a significant manner that is used in building a generative AI system. It defines 'Generative AI system' as a software product or service that (a) substantially incorporates one or more generative AI models; and (b) is designed for use by consumers.

These efforts, however, are largely absent in African legal systems, leaving Indigenous workers particularly vulnerable to

exploitation (AUDA-NEPAD, 2023; Okolo, 2023). Instead, many countries have had to rely on existing legal structures related to privacy, data protection, consumer rights, and human rights, which were established before these technologies emerged and are, therefore, not well-equipped to regulate them adequately (Okolo, 2023).

Indigenous communities across Africa have implemented various strategies to protect their intellectual property and cultural expressions from exploitation and appropriation. The San Code of Research Ethics, established in 2017, serves as a pioneering example of how African communities are developing ethical guidelines to protect their cultural heritage, including musical works, from exploitation and appropriation. This code requires prior informed consent, benefit-sharing agreements, and respect for San culture when accessing or using their traditional knowledge (Callaway, 2017). Such ethical frameworks provide a foundation for safeguarding indigenous cultural expressions in the face of emerging technologies like generative AI.

Similarly, the *Maasai Intellectual Property Initiative* (MIPI) demonstrates how African communities are asserting control over the commercial use of their cultural imagery and designs. The MIPI has established a legal entity to represent Maasai interests across Kenya and Tanzania, enabling them to license their intellectual property and challenge unauthorized use by international companies (Brindle & Florman, 2021). This approach not only protects cultural heritage but also creates potential revenue streams for community development.

Other African communities are employing various strategies to protect their cultural expressions. In Ghana, the Adinkra symbols are now protected under the country's Copyright Act as 'works of folklore', restricting unauthorized commercial exploitation (Boateng, 2011; Torkornoo, 2012; OseiTutu, 2017). These legal protections, combined with community-led initiatives, form a multi-faceted approach to safeguarding African cultural heritage. They also provide a framework for ensuring that AI de-

velopment respects cultural ownership and promotes fair use of African musical traditions.

However, enforcing these protections remains challenging, particularly across international borders and digital platforms. As generative AI systems evolve, it is crucial to develop collaborative frameworks that respect indigenous data sovereignty while promoting responsible innovation (Kukutai & Taylor, 2016). To prevent exploitation, content labeling frameworks must integrate ethical principles that prioritize transparency and cultural sensitivity. This involves the development of metadata systems that capture not only the provenance of Indigenous works but also the cultural and spiritual significance that these works hold for their communities.

For AI developers, the ethical responsibility extends beyond legal compliance. Transparent content identification systems and culturally sensitive metadata can ensure that Indigenous works are used with consent, proper attribution, and compensation. Although some AI developers have introduced opt-out mechanisms (Hays, 2023; OpenAI, 2024), these systems are often cumbersome. A better approach would involve opt-in systems and automatic attribution mechanisms that safeguard Indigenous intellectual property and promote equitable engagement with AI technologies (Henderson et al., 2023).

African nations require broad policy and legal reforms at both national and continental levels to safeguard the interests of Indigenous African content creators. This involves revising and harmonizing laws to protect intellectual property rights and ensure fair compensation for using Indigenous creative works, including Indigenous African songs, as training datasets for generative AI systems. To achieve this, African countries must develop more application-specific and sector-specific AI policies and make more significant efforts to promote public participation and collaboration on AI governance across various segments of society, including Indigenous communities.

At the very least, foundation model developers should demonstrate adherence to national copyright law regimes and, where required, allow rightsholders to opt out seamlessly and transparently if they choose to do so. Policy and governance frameworks should adapt and respond to changing local, national, and global digital ecosystems. Regulation should be introduced early in the development of new technologies. Early intervention is crucial if policymakers are to avert the harm that emerging technologies can do (Mazibuko-Makena & Kraemer-Mbula, 2021).

As mentioned earlier, African Indigenous communities have taken proactive measures to safeguard their intellectual property. For example, the San people of southern Africa developed the *San Code of Research Ethics*, while the Maasai of Kenya and Tanzania established the *Maasai Intellectual Property Initiative* (MIPI). In Ghana, the Adinkra symbols are legally protected under the Copyright Act. These community-led initiatives highlight the proactive measures Indigenous groups are taking to safeguard their cultural heritage in the digital era. While these examples primarily focus on visual and traditional knowledge aspects of culture, the principles and strategies employed can be equally applied to protect indigenous African musical works. The legal frameworks, community-led initiatives, and ethical guidelines established in these cases provide valuable models for safeguarding traditional music against unauthorized use or appropriation, including in the context of generative AI systems (Brindle & Florman, 2021).

Extant copyright frameworks need enhancement and adaptation to protect the cultural heritage of Indigenous African music. Specific provisions recognizing and safeguarding traditional cultural expressions may require the establishment of a *sui generis* system for the protection of Indigenous intellectual property like the case in jurisdictions such as Australia and the US Enforcing legal provisions for moral rights is also crucial to ensure proper attribution to Indigenous communities and individual creators, even within AI-generated content (Copyright

Act, 2023 Nigeria, s. 14; Copyright Act, 1978 South Africa, s. 20). It has been suggested that 'Instance attribution can also address the credit assignment problem by providing a clear attribution page that lists all works which contributed to the output, along with licensing information' (Henderson *et al.,* 2923). Providing customizable attribution models through content identification tools aligns with the emphasis of copyright law on recognizing creators' rights and preferences in the presentation of their work.

Earlier, this paper mentions that several Indigenous African musical works are passed down through oral tradition rather than written notation (Nketia, 1974; Nzewi, 1991; Onyeji, 2019). This means that these works are often dynamic and constantly evolving, shaped by community performances and teachings. As a result, traditional content identification and labeling methods, which rely on written records and digital metadata, may not be well-suited for these works. To address this, any effective framework must consider the oral nature of these works and incorporate community-based documentation efforts or oral history projects that honor and preserve the long-standing cultural traditions of these communities.

It is essential to develop specialized licensing frameworks for the fair utilization of Indigenous musical works in AI training datasets. These frameworks should allow permissible use while respecting the rights and interests of Indigenous communities. Adequate compensation and acknowledgment are fundamental components within these frameworks to ensure fair use while protecting Indigenous content creators' interests. To achieve this, it will be necessary to legislate for increased transparency in AI system development, mandating the disclosure of dataset sources to facilitate scrutiny and ensure the ethical use and representation of Indigenous cultural content.

VII. CONCLUSION

This paper critically examines the complex relationship between generative AI and Indigenous African musical works, emphasizing the risk of cultural appropriation through the use of Indigenous songs in AI training. The core argument is that generative AI training datasets, comprising diverse inputs, including proprietary and culturally significant materials, directly influence the AI models' outputs. The unchecked inclusion of Indigenous African works in these datasets not only risks appropriating cultural expressions but also undermines the collective ownership of these works, diminishing their cultural, spiritual, and economic value.

The author avers and demonstrates that current copyright frameworks are inadequate for protecting the dynamic, evolving nature of Indigenous African musical works, which are often orally transmitted. This gap leaves these communities vulnerable to exploitation by AI developers and other commercial entities. Additionally, the lack of comprehensive regulatory frameworks in several African countries exacerbates this vulnerability, as it leaves gaps that could be exploited by AI developers.

To address these challenges, this paper calls for more suitable content labeling frameworks and ethical guidelines prioritizing informed consent, transparency, and equitable compensation for Indigenous creators. Such frameworks should integrate culturally sensitive metadata and precise mechanisms for attribution and consent.

REFERENCES

Acemoglu, D., & Lensman, T. (2023). Regulating Transformative Technologies. *NBER Working Paper (No. w31461).* https://ssrn.com/abstract=4512495

African Commission on *Human and Peoples' Rights.* (2021) Resolution 473 on the need to Undertake a Study on the Right to Cultural Heritage in Africa.

African Regional Intellectual Property Organization (ARIPO). (2010). *Swakopmund Protocol on the Protection of Traditional Knowledge and Expressions of Folklore*. ARIPO.

African Union (AU). (2020). *Digital Transformation Strategy for Africa (2020-2030).* https://au.int/en/documents/20200518/digital-transformation-strategy-africa-2020-2030

African Union Development Agency–NEPAD (AUDA-NEPAD). (2023). *Regulation and Responsible Adoption of AI in Africa Towards Achievement of AU Agenda 2063.* White Paper.

African Union. (2022). *AU Data Policy Framework.*

Agostinelli A., Denk, T. I., Borsos, Z., Engel J., Verzetti, M. et al. (2023). MusicLM: Generating Music From Text. *Google Research*. https://arxiv.org/pdf/2301.11325.pdf

AI Labeling Act of 2023, S. 2691, 118th Congress (2023).

Alter, A., Harris, E. A. (2023). Franzen, Grisham and Other Prominent Authors Sue OpenAI. *The New York Times*. https://www.nytimes.com/2023/09/20/books/authors-openai-lawsuit-chatgpt-copyright.html

APRA AMCOS. (2024). *AI and Music.* APRA AMCOS. https://www.apraamcos.com.au/about/supporting-the-industry/research-papers/aiandmusic

Arya, R. (2021). Cultural appropriation: What it is and why it matters? *Sociology Compass*, 15(6), e12923.. https://doi.org/10.1111/soc4.12923

Attard-Frost, B. (2023). Generative AI Systems: Impacts on Artists & Creators and Related Gaps in the Artificial Intelligence and Data Act. *SSRN* https://ssrn.com/abstract=4468637

Attard-Frost, B., Hayes, H. A. (2023). Valuing Value Chains: On Canadian AI Regulation, Co-Governance, and the Scope of AI Value Chains. In Helen A. Hayes & Nicole Goodman (eds.) *Regulating Digital.* (Forthcoming) University of Toronto Press.

Attard-Frost, B., Widder, D. G. (2023). The Ethics of AI Value Chains. *arXiv*. https://arxiv.org/abs/2307.16787v2

Awad and Tremblay v. OpenAI Inc., No. 3:23-cv-03223 (2023).

Ayça, A. (2022). The Potential Economic Empowering Role of Cross-border Data Flows for Data Protection in Africa. pp. 14 – 21 In P. Gehl Sampath & F. Tregenna. (Eds.) *Digital Sovereignty: African Perspectives*. DSI/NRF South African Research Chair in Industrial Development.

Baio, A. (2020). With Questionable Copyright Claim, Jay-Z Orders Deepfake Audio Parodies Off YouTube. *WAXY*. https://waxy.org/2020/04/jay-z-orders-deepfake-audio-parodies-off-youtube/

Ben-Tal, O., et al. (2019). Artificial Intelligence and Music: Open Questions of Copyright Law and Engineering Praxis. Proceedings of the 2019 International Conference on Music Information.

Bengio, Y. (2012). Representation Learning: A Review and New Perspectives. *arXiv* https://arxiv.org/pdf/1206.5538.pdf

Bill C-27, *An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts,* 44th Parliament, 1st Session, cl. 39 (2022).

Birkstedt, T., Minkkinen, A., Tandon, A., Mantymaki, M. (2023). AI governance: themes, knowledge gaps and future agendas. https://www.emerald.com/insight/publication/issn/1066-2243

Boateng, B. (2011). *The Copyright Thing Doesn't Work Here*. University of Minnesota Press.

Boggs, J. (2022). Protecting Indigenous Artists Against Infringement and Appropriation. *Copyright Alliance*. https://copyrightalliance.org/protecting-indigenous-artists-infringement-appropriation/

Bogle, A. (2023). New York Times, CNN and Australia's ABC block OpenAI's GPT Bot web crawler from accessing content. *The Guardian*. https://www.theguardian.com/technology/2023/aug/25/new-york-times-cnn-and-abc-block-openais-gptbot-web-crawler-from-scraping-content

Bommasani R., Klyman K., Longpre S., Kapoor S., Maslej N., Xiong B., Zhang D., Liang P. (2023). The Foundation Model Transparency Index. https://crfm.stanford.edu/fmti/

Bommasani, R., Hudson, D. A., Adeli, E. et al., (2021). On the Opportunities and Risks of Foundation Models. *arXiv*. https://arxiv.org/abs/2108.07258

Brindle, M., & Florman, N. (2021). The Maasai Intellectual Property Initiative: A 20th-century model for turning assets into income. *Journal of Fair Trade*. 5-12.

Brittain, B. (2023). Google says data-scraping lawsuit would take 'sledgehammer' to generative AI. *Reuters*. https://www.reuters.com/legal/litigation/google-says-data-scraping-lawsuit-would-take-sledgehammer-generative-ai-2023-10-17/

Brittain, B. (2023b). Music publishers sue AI company Anthropic over song lyrics. *Reuters*. https://www.reuters.com/legal/music-publishers-sue-ai-company-anthropic-over-song-lyrics-2023-10-18/

Brown, I. (2023). Expert explainer: Allocating accountability in AI supply chains. *Ada Lovelace Institute*. https://www.adalovelaceinstitute.org/resource/ai-supply-chains/

Callaway, E. (2017). South Africa's San people issue ethics code to scientists. *Nature* 543(7646), 475-476.

Canadian Media Fund. (2016). Discoverability: Toward a Common Frame of Reference – Part 1.

Canadian Media Fund. (2016). Discoverability: Toward a Common Frame of Reference – Part 2. https://telefilm.ca/en/etudes/discoverability-toward-common-frame-reference-part-2-audience-journey

Caramiaux, B., Alaoui, S. F. (2022). Explorers of Unknown Planets: Practices and Politics of Artificial Intelligence in Visual Arts. *Proceedings of the ACM on Human-Computer Interaction*.

CDCE. (2020). The challenge of discoverability. The urgency to act to defend the diversity of cultural expressions. https://cdec-cdce.org/en/publications/the-challenge-of-discoverability-the-urgency-to-act-to-defend-the-diversity-of-cultural-expressions/

Chakrabarti, G. (2014). Vulnerable Position of Traditional Knowledge Under IPR: Concern for Sustainable Development. *OIDA International Journal of Sustainable Development*, 67 – 94.

Chanda, P. (2023). Is Soundraw Free to Use? A Comprehensive Analysis. https://ambcrypto.com/blog/is-soundraw-free-to-use-a-comprehensive-analysis/

Chapdelaine, P., Rogers, J. M. (2021). Contested Sovereignties: States, Media Platforms, Peoples, and the Regulation of Media Content and Big Data in the Networked Society. 10(3) *Laws* 66.

Chen, J., Yoshida, N., Takada, H. (2023). An investigation of licensing of datasets for machine learning based on the gqm model. https://arxiv.org/abs/2303.13735

Chen, S. H., Hopkins, A., Ilyas, A., Madry, A., Struckman, I., Videgaray, L. (2023). AI supply chains (and why they matter). *AI Policy Substack*. https://aipolicy.substack.com/p/supply-chains-2

Cobbe, J., Veale, M., Singh, J. (2023). Understanding accountability in algorithmic supply chains. *ACM Conference on Fairness, Accountability, and Transparency*. ACM. https://dl.acm.org/doi/10.1145/3593013.3594073

Concord Music Grp. Music, Inc., v. Anthropic PBC, PBC, No. 3:23-cv-01092 (2023).

Congressional Research Service. (2023). Generative Artificial Intelligence and Data Privacy: A Primer. https://crsreports.congress.gov/product/pdf/R/R47569

Copyright Act 2001. (Kenya).

Copyright Act 2002. (South Africa).

*Copyright Act 2005. (*Ghana).

Copyright Act 2023. (Nigeria).

Couldry, N., & Mejias, U. A. (2021). The decolonial turn in data and technology research: what is at stake and where is it heading? *Information, Communication & Society*, 26(4), 786–802. https://doi.org/10.1080/1369118X.2021.1986102

Craig, C. J. (2021). AI and Copyright. In Florian Martin-Bariteau & Teresa Scassa (eds.) *Artificial Intelligence and the Law in Canada*. LexisNexis Canada.

Craig, C. J. (2022). Legal reform to enhance global text and data mining research. *SCIENCE* 951-53.

Craig, C. J. (2022a). The AI-Copyright Challenge: Tech-Neutrality, Authorship, and the Public Interest. In Ryan Abbott (ed). *Research Handbook on Intellectual Property and Artificial Intelligence*. Edward Elgar Press.

Creamer, E. (2023). Authors file a lawsuit against OpenAI for unlawfully 'ingesting' their books. *The Guardian*. https://www.theguardian.com/books/2023/jul/05/authors-file-a-lawsuit-against-openai-for-unlawfully-ingesting-their-books

Cuthbert, D. (1998). Beg, borrow or steal: The politics of cultural appropriation. *Postcolonial Studies*, *1*(2), 257–262. https://doi.org/10.1080/13688799890174

Data Protection Act 2012. (Ghana).

Data Protection Act 2023. (Nigeria).

David, E. (2023). Musicians are eyeing a legal shortcut to fight AI voice clones. *The Verge*. https://www.theverge.com/2023/9/21/23836337/music-generative-ai-voice-likeness-regulation

de Beer, J., Elahi, S., Oguamanam, C., Rizk, N. (2013). Knowledge & Innovation in Africa: Scenarios for the Future. In Shirin Elahi et al (eds). *Knowledge & Innovation in Africa: Scenarios for the Future*. Open AIR.

Deahl, D. (2019). We've Been Warned About AI and Music for Over 50 Years, But No One's Prepared. *VERGE*. https://www.theverge.com/2019/4/17/18299563/ai-algorithm-music-law-copyright-human

Diakopoulos, N. (2016). Accountability in algorithmic decision making. *Communications of the ACM*, 56-62.

Dinika, A. (2022). Rethinking Digital Infrastructure Development in Africa. In P. Gehl Sampath & F. Tregenna (eds.), *Digital Sovereignty: African Perspectives*. DSI/NRF South African Research Chair in Industrial Development.

East African Community. (2008). *Legal framework for cyberlaws*.

Economic Community of West African States (ECOWAS). (2010). Supplementary Act A/SA.1/01/10 on personal data protection.

Effoduh, J., Akpudo, U. E., Kong, J. D. (2023). Towards an Inclusive Data Governance Policy for the Use of Artificial Intelligence in Africa. *SSRN*. https://ssrn.com/abstract=4581619

Eke, D. O., Wakunuma, K., Akintoye, S. (eds) (2023) *Responsible AI in Africa: Challenges and Opportunities*. Palgrave McMillan.

European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending certain Union legislative acts (Artificial Intelligence Act).* Official Journal of the European Union, L 168, 12 July 2024, pp. 1–144. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Ferrandis, C. M., Hughes, S. (2023). BigCode: toward open AI governance processes in LLMs for coding. https://oecd.ai/en/wonk/bigcode

Filespin. (2022). How AI/ML Can Enhance Digital Asset Management. http://filespin.io/blog/2022-03-03-how-aiml-can-enhance-digital-asset-management/

Fill-Flynn S.M. (2022). Legal reform to enhance global text and data mining research. 378 *SCIENCE* 951-53.

Fiolet, E. (2023). Google's MusicLM AI System Creates Music from Text Descriptions. https://www.ubergizmo.com/2023/01/googles-musiclm-ai-system-creates-music-from-text-descriptions/

Fuentes, J. A. (2023). Protecting the rights of indigenous cultures under the current intellectual property system: is it a good idea?. *J. Marshall Rev. Intell. Prop. L.* 88.

Gary Jr., A. (2021). Solving The Discoverability Dilemma In The Creator Economy. https://www.antoniogaryjr.com/essays/solving-the-discoverability-dilemma-in-the-creator-economy

Geist, M. (2021) AI and International Regulation. In Florian Martin-Bariteau & Teresa Scassa, (eds). *Artificial Intelligence and the Law in Canada*. LexisNexis Canada.

Gibson, J. (2023). Page against the machine: the death of the author and the rise of the producer?. *Queen Mary Journal of Intellectual Property* 275.

Gillwald, A. (2021). A new digital deal rather than a Fourth Industrial Revolution policy?. In Zanzima Mazibuko-Makena and Erika Kraemer-Mbula (eds). *Leap 4.0: African Perspectives on the Fourth Industrial Revolution*. MISTRA.

Goanta, C., Aletras, N., Chalkidis I., Ranchordas, S., Spanakis, G. (2023). Regulation and NLP (RegNLP): Taming Large Language Models. *arXiv*. https://arxiv.org/abs/2310.05553

Gokaslan, A., Cohen, V. (2019). Openwebtext corpus. https://skylion007.github.io/OpenWebTextCorpus/

Goujard, C., Volpicelli, G. (2023). ChatGPT is entering a world of regulatory pain in Europe. *Politico*. https://www.politico.eu/article/chatgpt-world-regulatory-pain-eu-privacy-data-protection-gdpr/

Guadamuz, A. (2023). A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs. *SSRN*. https://ssrn.com/abstract=4371204

Guadamuz, A. (2023). French lawmakers propose new copyright law about generative AI. https://www.technollama.co.uk/french-lawmakers-propose-new-copyright-law-about-generative-ai

Gwagwa, A., Kazim, E., Hilliard, A. (2022). The role of the African value of Ubuntu in global AI inclusion discourse: A normative ethics perspective. https://www.cell.com/patterns/pdf/S2666-3899(22)00042-3.pdf

Hard, L., Bordoli, J. (2022). How automated content tagging improves findability. *TechTarget*.

Hardie, M. (1999). The Bulun Bulun Case: John Bulun Bulun & Anor v R & T Textiles Pty Ltd (1998). *Indigenous Law Bulletin* 24.

Harrison, M. (2023). Lawsuit Claims Google Is Vacuuming Up People's Whole Lives To Train AI. https://futurism.com/the-byte/lawsuit-google-train-ai

Hayes, C. M. (2023). Generative Artificial Intelligence and Copyright: Both Sides of the Black Box. *SSRN*. https://ssrn.com/abstract=4517799

Hays, K. (2023). OpenAI offers a way for creators to opt out of AI training data. It's so onerous that one artist called it 'enraging'. *Business Insider*. https://www.businessinsider.com/openai-dalle-opt-out-process-artists-enraging-2023-9

Heeks, R. (2022). Digital inequality beyond the digital divide: conceptualizing adverse digital incorporation in the global South. *Information Technology for Development*, 28(4), 688–704. https://doi.org/10.1080/02681102.2022.2068492

Henderson, P., Li, X., Jurafsky, D., Hashimoto, T., Lamley, M. A., Liang, P. (2023). Foundation Models and Fair Use. *Journal of Machine Learning Research* 1.

Henshall, W. (2023). How the U.N. Plans to Shape the Future of AI. *Time*. https://time.com/6316503/un-ai-governance-plan-gill/

Hlomani, H., Ncube, C. B. (2023). Data Regulation in Africa: Free Flow of Data, Open Data Regimes and Cyber Security. *AERC Working Paper DG-004, African Economic Research Consortium*, Nairobi.

https://www.techtarget.com/searchcontentmanagement/tip/AI-in-content-management-supports-tagging-search

Huron, D. (2017). Canada needs a law protecting Indigenous art from appropriation. *Ricochet Public Journalism*. https://ric-

ochet.media/en/1808/canada-needs-a-law-protecting-indigenous-art-from-appropriation

Inawat, R. J. (2015). Music as Cultural Heritage: Analysis of the Means of Preventing the Exploitation of Intangible Cultural Heritage. 14 *J. Marshall Rev. Intell. Prop. L.* 228.

Indian Arts and Crafts Act, Pub. L. No. 101-644, tit. I, 104 Stat. 4662 (1990) (as amended). U.S. Innovation, Science and Economic Development Canada (ISED), *Canadian Guardrails for Generative AI – Code of Practice*, 2023.

J.L. v. Alphabet Inc, (2023) No. 3:23-cv-03440.

Jones, R. (2013). Technology and the cultural appropriation of music. In Jones, R & Moore, R

Kehagia, N., Moriaty, M. (2023). Recurring patterns: An ethnographic study on the adoption of AI music tools by practitioners of electroacoustic, contemporary and popular musics. *Journal of Pervasive Media*. 8. 51-64. 10.1386/jpm_00004_1.

Kenya. (2016). *Protection of Traditional Knowledge and Cultural Expressions Act No. 33 of 2016*. Government Printer.

King, S. (2023). My Books Were Used to Train AI. *The Atlantic*. https://www.theatlantic.com/books/archive/2023/08/stephen-king-books-ai-writing/675088/

Komminoth, L. (2023). Can Africa achieve 'digital sovereignty' in an era of Big Tech?. *African Business*. https://african.business/2023/07/technology-information/can-africa-achieve-digital-sovereignty-in-an-era-of-big-tech

Kukutai, T., & Taylor, J. (2016). *Indigenous data sovereignty: Toward an agenda*. ANU Press.

Lalani, A. (2017). Leslieville exhibition cancelled after Toronto artist's work called cultural appropriation. *Toronto Star*. https://www.thestar.com/news/gta/leslieville-exhibition-cancelled-after-toronto-artist-s-work-called-cultural-appropriation/article_e227d75d-a0d9-554d-a65f-8d8726679c89.html

Lawal-Arowolo, A. (2011). Copyright Law and the Recognition of 'Folkloric Creations' and 'Folk Medicine' in Africa. *Journal of Black and African Arts and Civilization*, 33.

Lawal-Arowolo, A. (2015). Losing Their Status: Traditional Peoples in Africa, Their Intellectual Properties and Laws. *SSRN*. https://ssrn.com/abstract=2694904

Lee, K., Cooper, A. F., Grimmelmann, J. (2023). Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain. *SSRN*. https://ssrn.com/abstract=4523551

Leibowicz, C. (2023). Why watermarking AI-generated content won't guarantee trust online. *MIT Technology Law Review* 1.

Lephoko, K. (2019). Language In Language Out: Natural Language Processing in the Context of Indigenous South African Languages. https://issuu.com/kgothatsolephoko1/docs/klephoko_langauge_in_01?utm_medium=referral&utm_source=cdn.embedly.com

Lixinski, L., Young, S. (2021). Creative Differences: Indigenous Artists and the Law at 20th Century Nation-Building Exhibitions. *Hastings Int'l & Comp. L. Rev.* 3.

Lockhart, A. (2021). A Stolen Culture: The Harmful Effects of Cultural Appropriation. *Honors Theses*. 1804. https://egrove.olemiss.edu/hon_thesis/1804

Lorenz, P., Perset, K., Berryhill, J. (2023). Initial Policy Considerations for Generative Artificial Intelligence. *OECD Artificial Intelligence Papers*. https://www.oecd.org/publications/initial-policy-considerations-for-generative-artificial-intelligence-fae2d1e6-en.htm

Malan, R. (2020). In the Jungle: Inside the Long, Hidden Genealogy of 'The Lion Sleeps Tonight'. *RollingStone*. https://www.rollingstone.com/feature/in-the-jungle-inside-the-long-hidden-genealogy-of-the-lion-sleeps-tonight-108274/

Matias, C. F. (2024). Generative AI, Copyright and Emancipation: The Case of Digital Art. *Law, Technology and Humans*, November. https://doi.org/10.5204/lthj.3567.

Mazibuko-Makena, Z., Kraemer-Mbula, E. (eds.) (2021) *Leap 4.0: African Perspectives on the Fourth Industrial Revolution*. MISTRA.

McKelvey, F., Hunt, R. (2019). Discoverability: Toward a Definition of Content Discovery

Metz, C. (2023). OpenAI in Talks for Deal That Would Value Company at $80 Billion. *New York Times*. https://www.nytimes.com/2023/10/20/technology/openai-artifical-intelligence-value.html?smid=nytcore-android-share

Metz, C. (2023). OpenAI to Offer New Version of ChatGPT for a $20 Monthly Fee. *New York Times*. https://www.nytimes.com/2023/02/01/technology/openai-chatgpt-plus-subscription.html

Milpurrurru v. Indofurn Pty Ltd (1994) FCA 975; 130 ALR 659; 30 IPR 209.

Modic, D., Hafner, A., Damij, N., Zajc, L. C. (2019). Innovations in intellectual property rights management: Their potential benefits and limitations. *European Journal of Management and Business Economics* 189.

Mouton, M., & Burns, R. (2021). (Digital) neo-colonialism in the smart city. *Regional Studies*, 55(12), 1890–1901. https://doi.org/10.1080/00343404.2021.1915974

Ndemo B., Ndung'u N., Odhiambo, S., Shimeles, A. (eds.) (2023). *Data Governance and Policy in Africa*. Palgrave McMillan.

Netflix (2019). ReMastered: The Lion's Share: 'The Lion Sleeps Tonight' Story. https://www.netflix.com/title/80191050

News Media Alliance. (2023). White Paper: How the pervasive copying of expressive works to train and fuel generative artificial intelligence systems is copyright infringement and not a fair use. https://www.newsmediaalliance.org/wp-content/uploads/2023/10/AI-White-Paper-with-Technical-Analysis.pdf

Nketia, K. (1974). *The Music of Africa*. Sussex: R.J Acford Ltd.

Nzewi, M. (1991). Musical Practice and Creativity: An African Traditional Perspective. Bayreuth: Iwalewa-Haus, University of Bayreuth, 1991. 164 pp., musical examples, figures, photographs, bibliography. *Yearbook for Traditional Music*. 1993; 25:154-156. 10.2307/768694

Office of the Privacy Commissioner of Canada. (2023). Joint statement on data scraping and the protection of privacy. https://www.priv.gc.ca/en/opc-news/speeches/2023/js-dc_20230824/

Oguamanam, C. (2017). Rethinking Copyright for Indigenous Creative Works. https://www.oguamanam.com/publications/2017-6-28/rethinking-copyright-for-indigenous-creative-works

Oguamanam, C. (2020). Indigenous peoples, data sovereignty, and self-determination: Current realities and imperatives. *The African Journal of Information and Communication* 1.

Oguamanam, C. (2021). From science, technology and innovation to Fourth Industrial Revolution strategies in Africa: The case for indigenous knowledge systems. In Zanzima Mazibuko-Makena and Erika Kraemer-Mbula (eds), *Leap 4.0: African Perspectives on the Fourth Industrial Revolution*. MISTRA.

Oguamanam, C. (2022). Transition to the Fourth Industrial Revolution: Africa's Science, Technology and Innovation Framework and Indigenous Knowledge Systems. *African Journal of Legal Studies* 12.

Okolo C. T. (2023). AI in the Global South: Opportunities and challenges towards more inclusive governance. The Brookings Institution. https://www.brookings.edu/articles/ai-in-the-global-south-opportunities-and-challenges-towards-more-inclusive-governance/

Onyeji, C. (2019). Composing art music from indigenous African musical paradigms. In Emily Achieng'Akuno (ed.) *Music Education in Africa: Concept, Process, and Practice*. Routledge, London.

Organisation for Economic Co-operation and Development (OECD). (2023). Regulatory sandboxes in artificial intelligence. *OECD Digital Economy Working Papers*. https://www.oecd.org/publications/regulatory-sandboxes-in-artificial-intelligence-8f80a0e6-en.htm

OseiTutu, J.J. (2017). Harmonizing Cultural IP across Borders: Fashionable Bags & Ghanaian Adinkra Symbols. *Akron Law Review*, 51(2).

Philips, L. (2023). The Lion Sleeps Tonight: one song's journey from 1930s South Africa to Disney money-spinner. *The Guardian*. https://www.theguardian.com/music/2023/may/08/mbube-the-lion-sleeps-tonight-lion-king

Protection of Personal Information Act (POPIA) 2021. South Africa.

Quach, K. (2023). How 'AI watermarking' system pushed by Microsoft and Adobe will and won't work. *The Registrar*. https://www.theregister.com/2023/10/15/microsoft_adobe_ai_watermark/

Raffel, Colin et al., "Exploring the limits of transfer learning with a unified text-to-text transformer" *arXiv* (2019), online: https://arxiv.org/abs/1910.10683

Reed, T. (2019). Fair Use as Cultural Appropriation. *California Law Review* 101.

REFERENCES

Reisner, A. (2023). Revealed: The Authors Whose Pirated Books Are Powering Generative AI. *The Atlantic*. https://www.theatlantic.com/technology/archive/2023/08/books3-ai-meta-llama-pirated-books/675063/.

Resolution adopted by the General Assembly on 13 September 2007 [without reference to a Main Committee (A/61/L.67 and Add.1)] 61/295. *United Nations Declaration on the Rights of Indigenous Peoples*.

ROPOSITION DE LOI, visant à encadrer l'intelligence artificielle par le droit d'auteur (2023). France.

Rosenblatt, E. L. (2019). Copyright's One-Way Racial Appropriation Ratchet. *UC Davis Law Review* 591.

Sag, M. (2023). Copyright Safety for Generative AI. *Houston Law Review* 101.

Sampath G. P., Tregenna F. (eds.) (2022). *Digital Sovereignty: African Perspectives*. DSI/NRF South African Research Chair in Industrial Development.

Scherzinger, M. (1999). Music, Spirit Possession and the Copyright Law: Cross-Cultural Comparisons and Strategic Speculations. *Yearbook for Traditional Music*. 31:102-125.

Sholder, S. J. (2023). Protecting Digital Content Online in the New Wild West of Generative AI and the Blockchain. *The Licensing Journal* 1.

Silvermann v. OpenAI Inc. (2023) Case 3:23-cv-03416.

Small, Z. (2023). Black Artists Say A.I. Shows Bias, With Algorithms Erasing Their History. *New York Times*. https://www.nytimes.com/2023/07/04/arts/design/black-artists-bias-ai.html

Sunray, EE (2021). Sounds of Science: Copyright Infringement in AI Music Generator Outputs. *Cath. U. J. L. & Tech* 185.

Through Platforms. *Social Media and Society*.

Torkornoo, G. (2012). Creating Capital from Culture - Re-thinking the Provisions on Expressions of Folklore in Ghana's Copyright Law. *Annual Survey of International & Comparative Law* 18. http://digitalcommons.law.ggu.edu/annlsurvey/vol18/iss1/4

Tran, T. H. (2022). Image Apps Like Lensa AI Are Sweeping the Internet, and Stealing From Artists. *The Daily Beast*. https://www.thedailybeast.com/how-lensa-ai-and-image-generators-steal-from-artists

Turk, V. (2023). How AI reduces the world to stereotypes. https://restofworld.org/2023/ai-image-stereotypes/

Vézina, B. (2019). Curbing cultural appropriation in the fashion industry with intellectual property. *WIPO Magazine*. https://www.wipo.int/wipo_magazine/en/2019/04/article_0002.html

Wairegi, A., Omino, M., Rutenberg, I. (2020). AI in Africa: Framing AI through an African Lens. *Communication, technologies et développement*. http://journals.openedition.org/ctd/4775

Wakunuma, K., Eke, D. (2024). Africa, ChatGPT, and Generative AI Systems: Ethical Benefits, Concerns, and the Need for Governance. Philosophies 9, no. 3: 80. https://doi.org/10.3390/philosophies9030080

Warrington, J. (2023). BBC blocks ChatGPT maker from using its content over AI copyright concerns. *The Telegraph*. https://www.telegraph.co.uk/business/2023/10/06/bbc-blocks-chatgpt-content-use-ai-copyright-concerns/

Weidinger, L., Uesato, J., Rauh, M., Griffin, C., Huang, P. et al. (2022). Taxonomy of Risks posed by Language Models. https://dl.acm.org/doi/10.1145/3531146.3533088

Xie, T., Poritz, I. (2023). Creator of buzzy ChatGPT is sued for vacuuming up 'vast amounts' of private data to win the 'A.I. arms race'. *Fortune*. https://fortune.com/2023/06/28/openai-chatgpt-sued-private-data/

Yilma, K. (2022). African Union's Data Policy Framework and Data Protection in Africa. *Journal of Data Protection and Privacy* 3.

Young, J. O. (2008). *Cultural Appropriation and the Arts*. Blackwell Publishing, USA.