Technology and Regulation

# Mitigating Generative AI's negative impact on Indigenous knowledge from international and Vietnamese laws perspectives


| Author(s) | |
|---|---|
| Contact | duongpham0827@gmail.com<br>E.P.Joubert@tilburguniversity.edu |
| Affiliation(s) | Duong Thuy Pham is a researcher at Viettel High Technology Industries Corporation - Branch of Viettel Group, Vietnam.<br>Tronel Joubert is a lecturer at the Private, business and labour law department of Tilburg Law School, Tilburg, the Netherlands. |



| Keywords | ethnic minority, generative AI, Indigenous knowledge, Indigenous peoples, Vietnam. |
|---|---|
| Citation | Duong Thuy Pham and Tronel Joubert, Mitigating Generative AI's negative impact on Indigenous knowledge from international and Vietnamese laws perspectives, Technology and Regulation, 2025, 194-213 • 10.71265/gk1gwc22 • ISSN: 2666-139X |



## Abstract

**Indigenous knowledge, which has been developed over generations and possesses a unique understanding of local environments, offers precious responses to sustainable development challenges, for instance, climate change, biodiversity loss and pollution. Despite its pivotal role, Indigenous knowledge of various ethnic minorities and Indigenous peoples is in danger of disappearing due to centuries of history of colonization, discrimination and racism. The emergence of GenAI will complicate the knowledge preservation effort as GenAI models constitute a threat, via content created by those models, to perpetuating and even amplifying inaccurate information related to Indigenous knowledge. This paper aims to discuss solutions to alleviate GenAI's adverse impact on Indigenous knowledge from international and Vietnamese laws perspectives, with the ultimate goal being to propose a feasible answer to protect Indigenous knowledge of 53 ethnic minorities in Vietnam from GenAI's threats. To arrive at the eventual outcome, this paper identifies the significance of Indigenous knowledge to sustainable development, and the vulnerability of Indigenous knowledge under GenAI's drawbacks. The paper, additionally, applies experience in addressing this issue from an international perspective to the context of Vietnam. By doing so, the paper furthermore raises the need for research to provide solutions to preserve and promote Indigenous knowledge that suits the socio-economic condition of each country, as there is no one-size-fits-all answer for ethnic minorities and Indigenous peoples on a global scale.**

## 1. Introduction

"*Is there only one kind of science we can use to truly understand the world?*"[1]

The answer is, as a legal scholar would say, it depends! At present, cultural diversity, with the wealth of knowledge transmitted through it,[2] enables each individual to understand the world from different perspectives.[3] As more knowledge vanishes due to the loss of languages and cultures worldwide,[4] individuals start to understand less about the world from viewpoints that are different from their own. In the worst scenario, next generations will only have one kind of science to explain how the world works.

Ethnic minorities and Indigenous peoples, whose knowledge has been misappropriated, together with their languages being most likely to disappear by 2115, are struggling with the challenge of knowledge preservation. Since 2023, with the breakout of GenAI models, for example, ChatGPT by OpenAI, Gemini by Google, Claude by Anthropic and Midjourney by Midjourney Inc., with their "smarter" update recently, showing GenAI models' ability to produce content based on the data they are trained on, has taken this challenge to extremes. In other words, when GenAI models learn from datasets that lack reliable information on the knowledge of ethnic minorities and Indigenous peoples (collectively called "**Indigenous knowledge**" or "**IK**"), the inaccuracy would be reflected in the content created by those models. As 90 percent of the content on the internet could be AI-generated by 2026, inaccurate content related to IK would spread over the internet. Gen AI models, consequently, would perpetuate and even amplify stereotypes related to ethnic minorities and Indigenous peoples.

This paper discusses on how GenAI models would adversely affect IK and which solutions are feasible for Vietnam, home of numerous ethnic minorities, with each minority possessing unique Indigenous knowledge that requires preservation and promotion.[5] With an aim to provide the audience with an overview of GenAI's negative impact on IK and how regulations govern such issues at present, the paper examines related concerns in a general view instead of detailed analysis. In this regard, the paper sets off by defining who are ethnic minorities and Indigenous peoples, as well as identifying the significance of Indigenous knowledge. Afterward, the paper examines GenAI models' negative impact on IK arising from data to train GenAI models, those models' outputs and power of Big Tech. The paper, in addition, discusses international perspectives on mitigating GenAI models' negative impact on IK, analyzes current Vietnamese laws regarding such issues, as well as how the international approaches would be applied to protect IK of Vietnam's ethnic minorities from GenAI's threats. By demonstrating how the international inspiration is tailored to suit the context of Vietnam, the paper furthermore emphasizes the necessity of researching socio-economic conditions and domestic regulations of the country where ethnic minorities or Indigenous peoples inhabit to propose the most appropriate solution to mitigate Gen AI models' impact on IK, as there is no one-size-fits-all answer for preservation and promotion of IK worldwide.

It should be noted that GenAI is a double-edged sword with both opportunities and challenges for preserving and promoting IK. Nevertheless, this paper only concentrates on discussing negative impact of GenAI on IK and feasible solutions to mitigate that impact, aiming to emphasize the essential of regulating GenAI models' development and operation properly to ensure that emerging technology does not threaten IK, which is both invaluable and vulnerable.

---

1. Chi Luu, "What We Lose When We Lose Indigenous Knowledge" [2023] JSTOR Daily https://daily.jstor.org/what-we-lose-when-we-lose-indigenous-knowledge.
2. *"Ostering Cultural Diversity, Intercultural Dialogue and a Culture of Peace in Tanzania" (UNESCO, September 14, 2023) https://www.unesco.org/en/fieldoffice/daressalaam/expertise/fosteringculturaldiversity accessed December 29, 2023.*
3. "Why Is Cultural Diversity Important?" (*Partnership International*) https://www.partnershipinternational.ie/why-is-cultural-diversity-important/#:~:text=It%20helps%20dispel%20negative%20stereotypes,respect%2C%20and%20understanding%20across%20cultures accessed December 29, 2023.
4. Syvlia Schmelkes, "Epistemic Justice and the Knowledge Commons for Lifelong and Lifewide Learning" (*UNESCO*, January 9, 2023) https://www.unesco.org/en/articles/epistemic-justice-and-knowledge-commons-lifelong-and-lifewide-learning accessed December 29, 2023.
5. "Ethnic Minority Empowerment | Oxfam in Vietnam" (*Oxfam in Vietnam*) https://vietnam.oxfam.org/ethnic-minority-empowerment accessed December 31, 2023.

## 2. Ethnic minorities, Indigenous peoples and the significance of Indigenous knowledge

"*Just as the world needs genetic diversity of species, it needs diversity of knowledge systems.*"[6]

At present, no international definition has been agreed on who are minorities.[7] However, it is widely accepted that the existence of a minority must include both objective factors (e.g. the existence of a shared ethnicity, language or religion) and subjective factors (e.g. individuals must identify themselves as members of a minority).[8] Furthermore, it is vital that a minority is in a non-dominant position compared with the dominant ethnic population of a country. Accordingly, in general, ethnic minority refers to an ethnic group in a country in which (i) they are fewer in number than the rest of the population, (ii) they share characteristics (such as culture, language, religion or race) that differentiate them from the majority and (iii) their members are determined to preserve those identities.[9] It is noteworthy that while some ethnic minorities have inhabited in their land for centuries, thus having deep understanding about the surrounding nature; other groups are descendants of migrants,[10] which means they may not possess their own knowledge about the place they are residing. Even though the origin of an ethnic minority varies, this paper only discusses ethnic minorities living on their land with their own knowledge about the surrounding nature.

Similarly, no universal definition of Indigenous peoples has been accepted.[11] Generally, Indigenous peoples inhabited land in a territory before it was invaded or colonized by currently dominant cultures.[12] Additionally, they possess "distinct social, economic and political systems, languages, cultures and beliefs" and seek to maintain and develop their distinct identity. [13]

The unique characteristics that distinguish an ethnic minority or Indigenous peoples from the majority of the society reflect knowledge (i.e. understandings, skills and philosophies) gained over centuries[14] through interaction with natural surroundings[15] embedded in their culture and life.[16] All that knowledge is referred to as traditional knowledge or Indigenous knowledge ("**IK**"),[17] which is defined as "knowledge, innovations and practices of Indigenous and local communities embodying traditional lifestyles relevant for the conservation and sustainable use of biological diversity".[18] IK can be expressed in various forms, including songs, folklore, proverbs, cultural values, beliefs, rituals, community laws, local language, agricultural practices and other means.[19]

6. Patrick Ngulube, "Managing and Preserving Indigenous Knowledge in the Knowledge Management Era: Challenges and Opportunities for Information Professionals" (2002) 18 Information Development 95, 96 https://doi.org/10.1177/026666602400842486.
7. https://emergency.unhcr.org/protection/persons-risk/minorities-and-indigenous-peoples 14 February 2025.
8. OHCHR, *Minority Rights: International Standards and Guidance for Implementation* (United Nations 2010) 2 https://www.ohchr.org/en/publications/special-issue-publications/minority-rights-international-standards-and-guidance.
9. *"Indigenous Peoples and Ethnic Minorities: Marginalization Is the Norm,", Report on the world social situation (2018) 97–98 https://doi.org/10.18356/14642ccc-en.*
10. Ibid.
11. https://emergency.unhcr.org/protection/persons-risk/minorities-and-indigenous-peoples 14 February 2025.
12. *Unhcr Emergency Handbook, "Minorities and Indigenous Peoples" UNHCR (January 30, 2024) https://emergency.unhcr.org/protection/persons-risk/national-ethnic-religious-and-linguistic-minorities-and-indigenous-peoples.*
13. OHCHR (n 8).
14. "Traditional Knowledge" (*UNESCO UIS*, September 12, 2023) https://uis.unesco.org/en/glossary-term/traditional-knowledge accessed April 27, 2024.
15. "Local and Indigenous Knowledge Systems (LINKS)" (UNESCO, January 6, 2022) https://en.unesco.org/links accessed December 31, 2023.
16. Ian Falefuafua Tapu and Terina Kamailelauli'I Fa'Agau, "A New Age Indigenous Instrument: Artificial Intelligence & Its Potential for (De)Colonialized Data" (2022) 57 Harvard Civil Rights-Civil Liberties Law Review 715, 719 https://journals.law.harvard.edu/crcl/wp-content/uploads/sites/80/2023/01/ANewAgeIndigenousInstrument.pdf.
17. For the avoidance of doubt, the term "Indigenous knowledge" in this paper means both "Indigenous knowledge" and "traditional knowledge", which are referrred to and used interchangeably in UNCTAD's book "*Protecting and promoting traditional knowledge: Systems, national experiences and international dimensions*".
18. UN Convention on Biological Diversity (1992). See more at: https://legal.un.org/avl/ha/cpbcbd/cpbcbd.html#3
19. "Traditional Knowledge" (n 14).

IK is the essence of the self-identity of ethnic minorities and Indigenous peoples. Particularly, IK connects individuals to the past, the history and the land they have lived on. By informing how their ancestors lived and interacted with other communities, other species and the cosmos, IK assists ethnic minorities and Indigenous peoples in understanding who they are at present. IK also contributes to a community's future well-being, as it is the foundation to keep languages alive, enrich cultural expressions and share cultural values with the wider world.[20]

IK benefits not only ethnic minorities and Indigenous peoples, but also the society as a whole.[21] These communities have lived in harmony with the Earth for millennia. In fact, 80 percent of global biological diversity is found in the 22 percent of land area where people have subsisted, consumed and cared for nature by using their IK. Learning that invaluable knowledge is learning how to restore environmental equilibrium.[22] Furthermore, "although IK is derived from careful observation of the environment in a particular context, it can be widely applied in many scenarios".[23] Accordingly, utilizing IK is helpful for not only the locality in which it evolves, but also for scientists and planners in designing development programs.[24]

Despite its significance for both the communities who possess the knowledge and the society, IK is at risk of becoming extinct because of various reasons, *inter alia*, language loss, and misappropriation and misuse of IK. Particularly, it is estimated that, by 2115, between 50 percent and 90 percent of the world's 7,000 mostly Indigenous languages will be extinct. [25]

As IK has generally been passed on in generations by word of mouth,[26] languages provide ethnic minorities and Indigenous peoples with unique identities, enable their cohesion and give them the means to practice and disseminate their cultures. Hence, the loss of languages contributes to the erosion of their identity and the disappearance of their culture, traditions and knowledge. Additionally, lacking protection leaves IK vulnerable to exploitation and profit-making without acknowledging or compensating the owners and producers of that knowledge.[27] Simultaneously, Indigenous knowledge-inspired products (e.g. Indigenous patterns or artefacts) are regularly appropriated without any consideration of the meaning of the knowledge behind the products. In some cases, the appropriation even disrespects or undermines the meaning of IK and its importance for the community who possess that knowledge.[28]

## 3. Negative impact of GenAI on Indigenous knowledge

"*From violations of privacy and personal integrity to the creation of fraud and misinformation, GenAI models introduce vast risks and challenges, while turbocharging others.*"[29]

Generative artificial intelligence ("**GenAI**") is a type of artificial intelligence ("**AI**") that is able to learn from data and create new data instances that mimic the properties of the input data. The general workflow to put a GenAI model to work includes: (i) collecting datasets containing examples of the type of content to be generated (e.g. text, audio, images, etc); (ii) training the GenAI model on the collected datasets to learn patterns and structures in the data; and (iii) generating new content that is a synthesis of what the

20. Eric Simons and others, "Traditional Knowledge - Fact Sheet" (2016) https://www.sfu.ca/ipinch/sites/default/files/resources/fact_sheets/ipinch_tk_factsheet_march2016_final_revised.pdf accessed April 27, 2024.
21. Peter Grant, *State of the World's Minorities and Indigenous Peoples 2016: Focus on Culture and Heritage* (Minority Rights Group 2016) 10.
22. Ibid.
23. Ngulube (n 6).
24. Ibid.
25. Grant (n 21) 49.
26. Ngulube (n 6).
27. "How to Protect Indigenous Knowledge and Creative IP from Exploitation" (*The University of Melbourne*, October 2, 2023) https://study.unimelb.edu.au/study-with-us/professional-development/blog/how-to-protect-indigenous-knowledge-and-creative-ip-from-exploitation accessed April 29, 2024.
28. Grant (n 21) 21.
29. Norwegian Consumer Council, "Ghost in the Machine – Addressing the Consumer Harms of Generative AI" (2023) 15 https://www.forbrukerradet.no/side/new-report-generative-ai-threatens-consumer-rights/ accessed December 15, 2023.

GenAI model has learned from the training data.[30] The breakout of GenAI models in 2023,[31] for example, ChatGPT by OpenAI, Gemini by Google and Midjourney by Midjourney Inc., presents potential challenges regarding how IK is owned, collected, interpreted and used.[32] This section analyses how GenAI models would negatively affect IK due to three reasons, namely data to train GenAI models, the content generated by those models and the power of Big Tech.

### 3.1 Negative impact arising from data to train GenAI models

The quality and quantity of data to train GenAI models are the bedrock of those models' effectiveness. Particularly, high-quality training data enables GenAI models to generate accurate and relevant outputs. Simultaneously, data diversity ensures that the generated outputs are accurate, inclusive and representative of varied perspectives.[33] Hence, any factors that affect the quality and quantity of datasets will also affect the operation of GenAI models trained on that data.

As GenAI models are trained on large datasets of existing content, any bias in datasets would be embedded in the GenAI models from the training stage. In practice, ethnic minorities and Indigenous peoples are often "discriminated against" and "marginalized socially, economically, politically, and culturally";[34] therefore, using datasets consisting of historical data (i.e. data collected about past events and circumstances)[35] would perpetuate existing bias towards those communities.

Furthermore, since data cannot "by nature fully reflect reality",[36] events that are not (or cannot be) recorded and quantified as data cannot be recognized by the GenAI models.[37] Considering the complexity of expressions of IK,[38] data regarding IK would only reflect pieces of knowledge "without understanding and respecting the wider context that created them".[39] Accordingly, the way GenAI models trained on such datasets learn about IK is similar to learning the words of a song but not the music.[40]

In addition, ethnic minorities and Indigenous peoples are likely to be underrepresented in the training data as a result of being among "the most marginalized communities in many societies".[41] Training datasets with only certain population groups being overrepresented would contribute to a feedback loop that continuously lessens impact of data from underrepresented populations. For example, the corpus of text of English, a widely used language, in the training data will be larger compared to unpopular Indigenous languages, which leads to providing more accurate information and better moderation.[42]

30. See more at "Generative AI: A Powerful New Technology | SAP" (SAP) https://www.sap.com/sea/products/artificial-intelligence/what-is-generative-ai.html accessed 26 Dec. 2023.
31. "The State of AI in 2023: Generative AI's Breakout Year" (*McKinsey & Company*, August 1, 2023) https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year accessed December 28, 2023.
32. "Leveraging UNESCO Normative Instruments for an Ethical Generative AI Use of Indigenous Data" (*UNESCO*, November 8, 2023) https://www.unesco.org/en/articles/leveraging-unesco-normative-instruments-ethical-generative-ai-use-indigenous-data#:~:text=Many%20indigenous%20communities%20are%20striving,knowledge%2C%20especially%20among%20younger%20generations accessed May 4, 2024.
33. Beyond Team, "Is Your Data Ready for Generative AI? A Comprehensive Guide" (*Beyond Team*, April 5, 2024) https://www.puttingdatatowork.com/post/is-your-data-ready-for-generative-ai-a-comprehensive-guide#:~:text=The%20bedrock%20of%20GenAI's%20effectiveness,both%20vast%20and%20meticulously%20curated accessed May 5, 2024.
34. *Unhcr Emergency Handbook (n 12)*.
35. Katie Terrell Hanna and Ivy Wigmore, "Definition: Historical Data" (*Tech Target*, November 2022) https://www.techtarget.com/whatis/definition/historical-data accessed May 5, 2024.
36. Animikii Inc., *DataBack: Asserting & Supporting Indigenous Data Sovereignty* (2023) 17 https://databack.animikii.com/.
37. Norwegian Consumer Council (n 29) 29.
38. Sandra Littletree, Miranda Belarde-Lewis and Marisa Duarte, "Centering Relationality: A Conceptual Model to Advance Indigenous Knowledge Organization Practices" (2020) 47 Knowledge Organization (KO) 410, 20 https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/46601/ko_47_2020_5_e.pdf?sequence=1&isAllowed=y.
39. Grant (n 21) 50.
40. Ibid.
41. *Unhcr Emergency Handbook (n 12)*.
42. Norwegian Consumer Council (n 29) 29-31.

Moreover, in the event that GenAI models are trained on material collected from the internet, the training data usually contains personal data, for example, photos of real people or conversations between individuals. People whose personal data are used to train a GenAI model would never be informed this would happen, never consent to such use of their data and likely not be aware of violation of their privacy and personal data.[43] Whisper API, the speech-to-text model launched by OpenAI,[44] which offers audio transcription and translation into Engish for numerous languages, including Māori, would serve as an example for this issue. Particularly, as Whisper API was trained on 1,381 hours of the Māori language from websites, without the community's consent, Indigenous technology and culture experts have raised concerns that harvesting their data without consent would abuse and distort Indigenous culture.[45]

### 3.2 Negative impact arising from content generated by GenAI models

By 2026, about 90 percent of online content could be synthetic (i.e. AI-generated).[46] As the volume of synthetic content grows, "people will not be able to know whether an image, text, sound, or video is real or synthetic".[47] Accordingly, it is easier to intentionally create and spread misleading content (disinformation) regarding IK across the internet. In the worst-case scenario, the proliferation of synthetic content even results in the deniability of authentic content. [48]

Even if the content is not intentionally generated to spread disinformation, such content would unintentionally cause negative impact on IK as it reflects existing issues regarding the quality and quantity of training data. Particularly, discriminatory factors against IK in the trained datasets would be reinforced through being recreated in GenAI models' generated content. Additionally, as the models do not "understand context and the content it produces",[49] content regarding IK created by GenAI models would look convincing and correct but is actually incorrect. Moreover, GenAI models' ability to generate content about an individual, such as portrait images or claims about a person, means that those models would create "new" personal data in a way that the individual has no control over. Due to ethnic minorities and Indigenous peoples' distinctive understanding of what information is inappropriate to be shared, and how cultural heritage should be treated, the aforementioned issues would worsen their concern regarding sharing information into datasets that are accessible to the public.[50]

It is also worth noting that GenAI models have the tendency to "make-up" non-existing sources themselves, by presenting either sources that do not exist or sources that do not contain relevant content to support the generated content.[51] As GenAI models' generated content has been increasingly persuasive and engaging, individuals, especially those who are not familiar with relevant subjects, would not be able to notice or reveal these mistakes.[52]

### 3.3 Negative impact arising from power of Big Tech

Nowadays, the market for GenAI is seized in the hands of Big Tech, such as Google, Meta and Microsoft. The ability to control the technology, provides Big Tech with the power to patch or otherwise modify GenAI models, add or remove functionality, as well as ban, filter or otherwise restrict content.[53] With this power, Big Tech can decide what generated content is harmful and what is permitted based on their point of view.

43. Ibid.
44. OpenAI, "Introducing APIs for GPT-3.5 Turbo and Whisper" (*OpenAI*, April 24, 2024) https://openai.com/index/introducing-chatgpt-and-whisper-apis/ accessed May 5, 2024.
45. Rina Chandran, "Indigenous Groups in NZ, US Fear Colonisation as AI Learns Their Languages" (April 3, 2023) https://www.context.news/ai/nz-us-indigenous-fear-colonisation-as-bots-learn-their-languages accessed May 5, 2024.
46. Norwegian Consumer Council (n 29).
47. Ibid 26.
48. Ibid 26.
49. Ibid 22..
50. Rebecca A Tsosie, "Tribal Data Governance and Informational Privacy: Constructing 'Indigenous Data Sovereignty'" (2019) 80 Montana Law Review 229, 236 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3454632.
51. Chris Moran, "ChatGPT Is Making up Fake Guardian Articles. Here's How We're Responding" *The Guardian* (April 6, 2023) https://www.theguardian.com/commentisfree/2023/apr/06/ai-chatgpt-guardian-technology-risks-fake-article.
52. Norwegian Consumer Council (n 29) 22.
53. Ibid 17.

Content moderation[54] of GenAI models, to some extent, helps remove discrimination or disinformation content regarding ethnic minorities and Indigenous peoples; however, attempts at removing unacceptable content also entail the risk of removing intrinsic content related to IK. Utilizing content moderation to set limits on what content a GenAI model is allowed to generate could also reinforce underrepresentation.[55]

Furthermore, Big Tech often "mines" data solely for their corporate interests while infringing benefits of various groups, including ethnic minorities and Indigenous peoples.[56] Particularly, as GenAi models need data as much as possible to train on, Big Tech tends to make use of data on the Internet, to power their systems. In addition to internet content, Big Tech would also collect data from users of their products, by stating in privacy policies that "data may be used to improve their existing products or develop new ones",[57] including GenAI models. Big Tech, however, are not willing to reveal their data sources and often simply state that the data used to train GenAI models are publicly available.[58] As it is not easy to verify whether training data, which contains various types of information, including IK, are collected and used with permission/consent, some actors have been raising questions regarding whether Big Tech should be allowed to use the collective knowledge of humanity in whole, or IK in part, to turn a profit. Additionally, ethical concerns regarding how values generated based on the knowledge should be distributed are under discussion. These concerns extend to data governance issues regarding who should control how data is used in some specific circumstances, such as a GenAI model trained on IK is commercialized.[59]

Meta, with an aim to get rid of experts' biases and perspective in content moderation, announced on 7 January 2025 that the company will end its fact-checking program on Facebook, Instagram and Threads' posts.[60] Accordingly, instead of hiring independent experts to manage content on its platforms, Meta will move to a Community Notes model, which relies on users to decide "when posts are potentially misleading and need more context" and "what sort of context is helpful for other users to see".[61] In this regard, Meta expects to ensure that beliefs and experiences are shared on their platforms with less restrictions. However, Mark Zuckerberg acknowledged that this change is a "tradeoff",[62] as this moderation change would enable more harmful content to appear on Meta's platforms. At present, it is uncertain whether Meta's change in content moderation policy would lead to positive or negative impact for its platforms' users. In the long run, in case information affected by Meta's new policy becomes "public available sources" and thus used to train GenAi models, the consequence is also unpredictable.

## 4. Solutions to mitigate GenAI's negative impact on Indigenous knowledge from international perspectives

"*As AI continues to reshape society, it is crucial to ensure that technological advancements respect the unique characteristics of IK, particularly, how owned, collected, interpreted and used in AI systems.*"[63]

54. Content moderation is the process of reviewing and monitoring user-generated content to ensure that it meets certain standards and guidelines. This includes removing inappropriate or offensive content and enforcing community guidelines and terms of service. See more at: Jonas Strandell and Jonas Strandell, "What Is Content Moderation? (Plus Best Practices)" (*Besedo*, May 24, 2024) https://besedo.com/blog/what-is-content-moderation accessed May 26, 2024.
55. Norwegian Consumer Council (n 12) 30-31.
56. Jason Edward Lewis and others, "Indigenous Protocol and Artificial Intelligence Position Paper" [2020] 61 https://spectrum.library.concordia.ca/986506/.
57. https://www.vox.com/technology/2023/7/27/23808499/ai-openai-google-meta-data-privacy-nope 16 February 2025
58. Ibid.
59. Norwegian Consumer Council (n 29) 17.
60. Joel Kaplan, "More Speech and Fewer Mistakes" *Meta* (January 7, 2025) https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/.
61. Ibid.
62. Ibid.
63. UNESCO "Leveraging UNESCO Normative Instruments for an Ethical Generative AI Use of Indigenous Data" (*UNESCO*, November 8, 2023) https://www.unesco.org/en/articles/leveraging-unesco-normative-instruments-ethical-generative-ai-use-indigenous-data#:~:text=Many%20indigenous%20communities%20are%20striving,knowledge%2C%20especially%20among%20younger%20generations accessed May 4, 2024.

The breakout of GenAI models has undeniably resulted in negative impact on IK. Considering the essence of IK for both the society and communities possessing it, plus the vulnerable nature of IK, it is fundamental to determine how to duly address those issues to protect the knowledge. This section scrutinizes solutions proposed by UNESCO, the EU and Australia to tackle challenges related to GenAI's negative impact on IK.

### 4.1 UNESCO's Recommendation on the Ethics of Artificial Intelligence

The United Nations Educational, Scientific and Cultural Organization ("**UNESCO**") is a specialized agency of the United Nations, with its mission being the promotion of "knowledge sharing and the free flow of ideas to accelerate mutual understanding and a more perfect knowledge of each other's lives". [64] Considering the development of GenAI would present challenges regarding using knowledge without cultural sensitivity, UNESCO provides instruments to support Member States to protect IK and cultural expressions.[65] The Recommendation on the Ethics of Artificial Intelligence, which is UNESCO's first-ever global standard on AI ethics adopted by all Member States in 2021,[66] is one of those provided instruments, with it seeking to guide AI development in a responsible direction. [67] As GenAI is a type of AI, recommendations regarding AI life cycle[68] in general would be applicable to GenAI should be taken into account when addressing issues related to GenAI models.

In relation to IK, the Recommendation on the Ethics of Artificial Intelligence emphasizes that attention in AI ethics debates must be paid to "neglecting local knowledge" and "cultural pluralism".[69] The Recommendation, in addition, recognizes that data in which AI models have been trained would impact social and cultural processes, with biases embedded in those models potentially resulting in "discrimination, inequality, digital divides, exclusion and a threat to cultural, social and biological diversity and social or economic divides".[70] Moreover, it is acknowledged that the development of technologies should parallel a commensurate increase in data, and access to trusted sources of data would contribute to mitigating the risk of disinformation.[71] The Recommendation also notes that the protection against AI's negative implications should adapt to different circumstances of different countries so that no country and no one is left behind.[72]

Particularly, Member States, on a voluntary basis, are encouraged to apply guidance as follows:

i) *Regarding fairness and non-discrimination*, the Recommendation suggests that the specific needs of each group (including, *inter alia*, different language groups, cultural systems, and marginalized and vulnerable people) should be considered when sharing AI benefits. Additionally, Member States should ensure that local communities are able to access and engage in the life cycle of AI models "with locally relevant content and services, and with respect for multilingualism and cultural diversity". Simultaneously, reinforcing or perpetuating discriminatory and biased outcomes needs to be minimized and avoided.[73]
ii) *Regarding multi-stakeholder and adaptive governance and collaboration*, the Recommendation proposes that the participation of different stakeholders, including but not limited to governments, intergovernmental communities and private sector companies is necessary throughout the AI life cycle. Furthermore, the Recommendation emphasizes the significant participation of marginalized groups in AI development.[74]
iii) *Regarding ethical impact assessment*, the Recommendation suggests that frameworks for impact assessment, as well as due diligence and oversight mechanisms to identify impact, predict consequences, mitigate risks and avoid harmful consequences of AI models, need to be developed. In addition, the

64. "UNESCO in Brief" (*UNESCO*) https://www.unesco.org/en/brief accessed May 8, 2024.
65. UNESCO (n 63.
66. UNESCO, *Recommendation on the Ethics of Artificial Intelligence* (UNESCO 2022) https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence.
67. Ibid.
68. The AI life cycle range from research, design and development to deployment and use, including maintenance, operation, trade, financing, monitoring and evaluation, validation, end-of-use, disassembly and termination. See more at UNESCO (n 66).
69. UNESCO (n 66).
70. Ibid.
71. UNESCO (n 66) 5-6.
72. Ibid 4.
73. Ibid 20.
74. Ibid 23.

assessment should take into consideration cultural and social diversity, including local customs and religious traditions.[75]

iv) *Regarding culture*, the Recommendation encourages Member States to examine and tackle AI's cultural impact on nuances of human language and expression, focusing on negative implications such as reduction of use, which would result in "the disappearance of endangered languages, local dialects, and tonal and cultural variations associated with human language and expression". Moreover, visibility and discoverability of local content should be ensured, and a diverse supply of and plural access to cultural expressions should be promoted.[76]

As knowledge possessed by ethnic minorities and Indigenous peoples is unique and precious, considering those communities' specific needs, supporting their participation throughout the AI life cycle, as well as identifying and tackling technology's consequences on IK would facilitate the mitigation of GenAI's negative impact on IK.

It is also noteworthy that, in addition to potential risks, AI technologies carry positive impact on IK, such as preservation, enrichment, understanding, promotion, management and accessibility of IK.[77] Accordingly, addressing ethical concerns should contribute to providing opportunities that anchor innovation in "values and principles, and moral and ethical reflection" instead of hampering technology development.[78]

## 4.2 EU's regulations related to mitigating GenAI's negative impact on Indigenous knowledge

The European Union ("**EU**") has a unique "unilateral ability"[79] to regulate the global marketplace with its regulations. The EU's General Data Protection Regulation ("**GDPR**") has set a standard for jurisdictions worldwide regarding privacy and data protection,[80] and now continues to affect the operation of numerous GenAI models as parts of their data contain personal data. In addition, the EU's Artificial Intelligence Act ("**AI Act**"), the world's first comprehensive AI law,[81] is expected to shape "how multinational companies manage data and navigate AI".[82] Hence, discussing how the GDPR and the AI Act address issues related to GenAI models' impact on IK is vital to anticipate how other countries would follow the EU's footsteps in tackling those issues in the future.

### *4.2.1 Artificial Intelligence Act*

The AI Act, which was adopted on 13 March 2024, came into effect as of 2 August 2024,[83] with an aim to guarantee that the use of AI in the EU is "safe, transparent, traceable, non-discriminatory and environmentally friendly".[84] The AI Act primarily applies to (i) all AI systems[85] placed on the EU market or their output is used in the EU and (ii) providers and deployers of AI systems having their place of establishment or located in the EU.[86] Governance of AI systems is accessed on a risk-based approach, with four levels of risks being

75. Ibid 26-28.
76. Ibid 32-33.
77. Ibid 31.
78. Ibid 5.
79. Anu Bradford, "The Brussels Effect," *Oxford University Press eBooks* (2019) 1 https://doi.org/10.1093/oso/9780190088583.003.0003.
80. Jennifer Bryant, "3 Years in, GDPR Highlights Privacy in Global Landscape" (*IAAP*, May 25, 2021) https://iapp.org/news/a/three-years-in-gdpr-highlights-privacy-in-global-landscape accessed May 18, 2024.
81. "EU AI Act: First Regulation on Artificial Intelligence" (*European Parliament*, August 6, 2023) https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence accessed May 12, 2024.
82. Isabel Gottlieb, "EU Poised to Enact Sweeping AI Rules With US, Global Impact (1)" *Bloomberg Law* (March 6, 2024) https://news.bloomberglaw.com/artificial-intelligence/eu-poised-to-enact-sweeping-ai-rules-with-us-global-impact.
83. "The Artificial Intelligence Act" (*European Council*) https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401689 accessed 3 September 2024.
84. "EU AI Act: First Regulation on Artificial Intelligence" (*European Parliament*, August 6, 2023) https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence accessed May 12, 2024.
85. The AI Act, Article 3 (1) defines an AI system as a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.
86. The AI Act, Article 2.

minimal risk, transparency risk, high risk and unacceptable risk.[87] In general, the AI Act focuses on regulated high-risk AI systems while minimal-risk systems are unregulated, transparency-risk systems are subject to light transparency obligations and unacceptable-risk systems are prohibited.[88]

The AI Act states that high-quality data and access to high-quality data are crucial to ensure AI systems perform safely and do not become a source of discrimination. In particular, any potential biases in the datasets, for example, those generated when using historical data or implementing systems in real-world settings, would perpetuate and amplify existing discrimination.[89] To mitigate this risk, the training datasets are required to follow data governance and management practices; be "relevant, sufficiently representative, and to the best extent possible free of errors and complete"; and consider "features, characteristics or elements that are particular to the specific geographical, contextual, behavioral or functional setting".[90] Moreover, the AI Act prescribes that bias detection and correction are allowed to be processed only when conditions regarding safeguards of personal data are satisfied.[91] As mentioned in Section 3.1, datasets to train GenAI models are likely to include biases against IK influenced by historical data, data taken out of context, as well as ethnic minorities and Indigenous peoples being discriminated and underrepresented. Hence, applying requirements to achieve high-quality data, and duly detect and tackle biases in the training data would help mitigate adverse impact of operating GenAI models on those communities.

The AI Act, in addition, recognizes that the creation of large quantities of synthetic content by AI systems would result in misinformation and manipulation on a large scale.[92] Accordingly, it is mandatory for AI systems generating synthetic content to mark outputs in a machine-readable format, detect them as "artificially generated" or "manipulated", as well as take into account "specificities and limitations of various types of content".[93] Applying those requirements would be useful to identify whether data regarding IK is real or generated by GenAI models. Thus, in case synthetic content is related to IK, regardless of whether the expressed content is convincing or unconvincing, a person, even who is not familiar with the subject, would recognize that the content is artificially generated, which means that the content potentially contains discrimination, disinformation or is based on non-existing sources.

Together with governing risks arising from using AI systems, the AI Act also attempts to support innovation, and ensure free movement, cross-border, of AI-based goods and services by preventing EU Member States from imposing any restrictions on AI systems that are not explicitly authorized by the AI Act.[94]

#### *4.2.2 General Data Protection Regulation*

The GDPR, which was adopted in 2016 and has been effective since 2018,[95] is crafted for the purpose of "modernizing EU data protection legislation, making it fit for protecting fundamental rights in the context of the digital age's economic and social challenges".[96] The GDPR applies to the processing of personal data (i) by companies established in the EU and (ii) of data subjects[97] who are in the EU.[98] As GenAI models analyze massive amounts of data (i) collected from the internet, with parts of it being undeniably personal

87. "Briefing EU Legislation in Progress - Artificial Intelligence Act" (*European Parliment*, March 2024) https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf accessed May 12, 2024.
88. "High-Level Summary of the AI Act | EU Artificial Intelligence Act" (*EU Artificial Intelligence Act*, February 27, 2024) https://artificialintelligenceact.eu/high-level-summary/ accessed May 12, 2024.
89. The AI Act, Recital 67.
90. The AI Act, Article 10 (2), (3) and (4).
91. The AI Act, Article 10 (5).
92. The AI Act, Recital 133.
93. The AI Act, Article 50.
94. The AI Act, Recital 1.
95. "The General Data Protection Regulation" (*European Council*) (n 83).
96. European Union Agency for Fundamental Rights and others, Handbook on European data protection law (2018 edn, Imprimerie Centrale, 2018) 30.
97. The GDPR, Article 4 (1) defines a "data subject" as an identified or identifiable natural person who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.
98. The GDPR, Articles 2 (1) and 3 (1).

data and (ii) associating with individuals through personal accounts to use those models, the GDPR is therefore applicable to some input and output of GenAi models.[99] Nevertheless, considering the fact that the GDPR has been enacted for a few years before the bloom of GenAI models, how the GDPR is going to govern issues related to personal data processed by those models remains uncertain, with some noteworthy issues being a lawful basis to process personal data, as well as the implementation of the rights to rectification and erasure.

*Regarding the lawful basis*, the GDPR, in general, requires personal data to be legitimately processed if it is based on data subjects' consent for all purposes carried out.[100] This requirement, however, is likely to clash with the operation of GenAI models. Particularly, considering the vast number of data subjects whose data have been used to train those models, plus each model's ability to add an infinite number and range of applications, obtaining individual consent would be a complex exercise.[101] In addition to data subjects' consent, other legal bases for processing personal data under the GDPR include contract, legal obligation, vital interests, public task and legitimate interests.[102] Since GenAI models' development stages and usage vary, it would be difficult to determine which basis will be applied. For instance, under Article 6 (1) (b) of the GDPR, a contract may form a lawful basis for processing personal data, provided that "processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract". A contract as a lawful basis is, however, unsuitable in the case of GenAI models as typically no contract exists between the data subjects whose data is used to train such models and the companies being responsible for training those models.[103]

*Regarding the right to rectification*, the GDPR states that data subjects have the right to "obtain ... without undue delay the rectification of inaccurate personal data" and "have incomplete personal data completed".[104] In fact, input to train GenAI models is frequently not accurate[105] and would be out-of-date[106] while content created by the models is sometimes affected by bias and discrimination, or "plausible-sounding but incorrect or nonsensical".[107] In case those inaccuracy input and output contain personal data, it is, however, unclear how GenAI models would correct data that is embedded in the models, as modifying the data is both costly and time-consuming.[108]

*Regarding the right to be forgotten,* the GDPR allows data subjects to request companies to erase their personal data without undue delay under some circumstances.[109] Removing data from GenAI models' datasets, however, is a complex task, with the same reason as implementing the right to rectification, i.e. the task consumes substantial costs and time.[110]

As discussed in Section III, existing challenges related to IK include ethnic minorities and Indigenous peoples' personal data collected and used to train GenAI models with neither notification nor asking for

99. Norwegian Consumer Council (n 29) 45.
100. The GDPR, Recitals 32 and 40.
101. "Walking the Tightrope: As Generative AI Meets EU Regulation, Pragmatism Is Likely" (*Deloitte Insights*, January 18, 2024) https://www2.deloitte.com/xe/en/insights/industry/technology/technology-media-and-telecom-predictions/2024/tmt-predictions-eu-generative-ai-regulation.html accessed May 14, 2024.
102. The GDPR, Article 6 (1).
103. Confederation of European Data Protection Organisations, *Generative AI: The Data Protection Implications* (2023) 8-9 https://cedpo.eu/generative-ai-the-data-protection-implications/.
104. The GDPR, Article 16.
105. In March 2023, the Italian Data Protection Authority highlighted that "As confirmed by the tests carried out so far, the information made available by ChatGPT does not always match factual circumstances, so that inaccurate personal data are processed". See more at: https://www.gpdp.it/web/guest/home/docweb/-/docweb-display/docweb/9870847#english accessed May 17, 2024.
106. GenAI models, such as ChatGPT, Gemini or Claude, rely on data scraping activities up to a certain point in time, which means the data may not be up-to-date and users may obtain personal data that is no longer relevant, lack context or outrightly inaccurate. See more at: Confederation of European Data Protection Organisations, *Generative AI: The Data Protection Implications* (2023) 5 https://cedpo.eu/generative-ai-the-data-protection-implications/.
107. OpenAI, "Introducing ChatGPT" (*OpenAI*, November 30, 2022) https://openai.com/index/chatgpt/ accessed May 17, 2024.
108. Confederation of European Data Protection Organisations (n 103).
109. The GDPR, Article 17.
110. Confederation of European Data Protection Organisations (n 103).

consent; data related to IK being inaccurate, containing bias and discrimination; new personal data about those individuals generated beyond their control; and Big Tech mining data from those communities without ethical concerns. Considering the GDPR's regulations as aforementioned would, to some extent, be the solution for the challenges, it is, therefore, necessary to provide guidance on GenAI models' compliance with the GDPR as soon as possible.

### 4.3 Australia's approach to a stand-alone legislation to protect Indigenous knowledge

Aboriginal and Torres Strait Islander peoples, who are the first peoples of Australia or "Indigenous Australian", are distinct groups with different cultures, languages, beliefs and practices.[111] With an aim to better support Indigenous Australian, the IP Australia[112] has been discussing what can be improved in protecting their IK. During their work, the IP Australia recognized that existing Australian Intellectual Property ("**IP**") laws, which focus on individual rights with a limited term of protection, are "ill suited" to protecting IK, which is communal and passed down over generations.[113] Simultaneously, the IP Australia confirmed six primary concerns of Indigenous Australian regarding the IK, namely (i) misappropriation of Indigenous arts and crafts, (ii) misuse of Indigenous languages and clan names, (iii) ownership of recordings and digitized records of IK, (iv) commercial exploitation of IK without benefits flowing to Indigenous communities, (v) unauthorized use of IK relating to genetic resource and (vi) misuse of sensitive secret sacred knowledge.[114]

Considering these issues, the IP Australia has been working on a stand-alone legislation to support Indigenous Australian to protect and commercialise IK. Four noteworthy recommendations of this approach are:

- A definition of IK should cover a range of knowledge held and developed by Indigenous Australian,[115] with the definition recognizing that IK is "not static but living and evolving and capable of adapting over time to contemporary and commercial realities".[116]
- IP rights concentrating on ownership and protection of IK need to be created. Particularly, protection of IK should be similar to copyright, i.e. being dependent on either registration or inclusion in a database. In addition, any third parties (e.g. collaborators, businesses and researchers) who would like to use IK are responsible for obtaining free, prior and informed consent of the Indigenous community owning such knowledge, as well as entering into a license agreement to share financial and non-financial benefits.[117]
- Establishment of a national IK authority is necessary to enforce action against unauthorized use and misappropriation of IK, help third parties identify and secure each community's consent to use their IK, as well as educate and advise Indigenous Australian on how to protect their IK.[118]
- A partnership between the Australian Government and Indigenous communities is essential during design and development process of the stand-alone legislation, with such legislation being guided by views, priorities and knowledge of the First Peoples of Australia.[119]

111. Australian Institute of Aboriginal and Torres Strait Islander Studies, "Australia's First Peoples" (*AIATSIS*) https://aiatsis.gov.au/explore/australias-first-peoples#toc-what-term-is-best-to-use- accessed May 19, 2024.
112. IP Australia is the Australian Government agency that administers IP rights and legislation for patents, trademarks, design rights and plant breeder's rights (PBR). See more at: "About Us" (*IP Australia*) https://www.ipaustralia.gov.au/about-us accessed May 19, 2024.
113. "Scoping Study on Standalone Legislation to Protect and Commercialise Indigenous Knowledge - Final Report" (2023) 6 https://www.ipaustralia.gov.au/tools-and-research/professional-resources/data-research-and-reports/publications-and-reports/scoping-study-on-standalone-ik-legislation.
114. "Interim Report: Scoping Study on stand-alone legislation to protect and commercialise Indigenous Knowledge" (2023) 4 https://www.ipaustralia.gov.au/tools-and-research/professional-resources/data-research-and-reports/publications-and-reports/scoping-study-on-standalone-ik-legislation.
115. A range of IK encompasses "the intangible and tangible aspects of the cultural practices, cultural expressions, resources and knowledge systems that have been and continue to be developed, nurtured and refined as part of expressing cultural identity and that includes, without limitation: story, songlines, ceremonies, languages and traditional land management practices". See more at: "Scoping Study on Standalone Legislation to Protect and Commercialise Indigenous Knowledge - Final Report" (n 113).
116. "Scoping Study on Standalone Legislation to Protect and Commercialise Indigenous Knowledge - Final Report" (n 113).
117. "Interim Report: Scoping Study on stand-alone legislation to protect and commercialise Indigenous Knowledge" (n 114).
118. Idem
119. "Scoping Study on Standalone Legislation to Protect and Commercialise Indigenous Knowledge - Final Report" (n 113).

The approach of the Australian Government, when being enforced, would be a promising answer for better protection of IK against issues related to negative impact of GenAI models, namely Big Tech collecting and using IK without consent of ethnic minority and Indigenous peoples, benefits generated by exercising IK not distributed to communities who own the knowledge, and expressing IK without understanding its context and meaning.

## 5. Mitigating GenAI's negative impact on Indigenous knowledge from Vietnamese laws perspective – Issues and Proposals

"*Culture is the soul of the nation, reflecting its identity. The nation remains as long as culture remains.*"[120]

Vietnam, the home of 54 ethinic groups[121] and 93 living languages[122], is an ethnically diverse society. The Kinh ethnic, which accounts for 87 percent of the population, is the majority of the country. Despite the difference in the total number of people, the Kinh ethnic and other Vietnam's ethnic groups have established a fundamental solidarity from the first century of the history for economic purposes. This solidarity has been reinforced among 54 ethnic groups for centuries as a result of unceasingly defending Vietnam's independence.[123]

Throughout history, Vietnam's ethnic minorities[124] have contributed to not only the country's defence against invasion, but also its cultural diversity. At present, even though ethnic minorities only account for a small percentage of the population, these communities possess a vibrant diversity of IK that has been created over thousands of years.[125] Various research has been conducted to emphasize the significance of IK to not only ethnic minorities themselves but also the society in numerous sectors, such as agriculture activities,[126] forest protection[127] and adaptation to salinity.[128] Despite being beneficial, IK of Vietnam's ethnic minorities is fading away, with languages coming close to extinction and local cultural values being weakened.[129]

The global breakout of GenAI models, as discussed in Section 3, entails potential negative impact on IK regardless of the knowledge's origin. IK possessed by ethnic minorities of Vietnam, accordingly, cannot escape the influence of this universal phenomenon. Particularly, as many GenAI models have been mainly

120. Phu Trong Nguyen, "Strive to build, preserve and promote the unique values of Vietnamese culture, which is advanced and rich in national identity" (*Communist Review*, November 24, 2021) https://www.tapchicongsan.org.vn/media-story/-/asset_publisher/V8hhp4dK31Gf/content/tiep-tuc-xay-dung-giu-gin-chan-hung-va-phat-trien-nen-van-hoa-cua-dan-toc accessed May 22, 2024.

121. "Ethnic Groups in Vietnam" (*Vietnam Government Portal*) https://vietnam.gov.vn/ethnic-groups-in-viet-nam accessed May 21, 2024.

122. "Socialist Republic of Vietnam" (*Ethnologue*) https://www.ethnologue.com/country/VN/ accessed May 21, 2024.

123. "Ethnic Groups in Vietnam" (*Vietnam Government Portal*) https://vietnam.gov.vn/ethnic-groups-in-viet-nam accessed February 10, 2025.

124. At present, the official number of ethnic minorities in Vietnam is 53. However, since not all 53 ethnic groups possess their own IK as defined in Chapter 1, the paper only discusses about Vietnam's ethnic minorities with IK.

125. Ha Thi Song Nguyen, "Policy to preserve and promote cultural values of ethnic minorities in Vietnam" (*Communist Review*, August 12, 2023) https://www.tapchicongsan.org.vn/web/guest/van_hoa_xa_hoi/-/2018/828114/chinh-sach-bao-ton-va-phat-huy-gia-tri-van-hoa-cac-dan-toc-thieu-so-o-viet-nam.aspx accessed May 20, 2024.

126. Ho Ngoc Son, Dong Thi Linh Chi and Aaron Kingsbury, "Indigenous Knowledge and Climate Change Adaptation of Ethnic Minorities in the Mountainous Regions of Vietnam: A Case Study of the Yao People in Bac Kan Province" (2019) 176 Agricultural Systems 102683 https://doi.org/10.1016/j.agsy.2019.102683.

127. Han Mai Tuyet, "Indigenous knowledge in plant management and use of Non-Timber Forest products Of the Van Kieu community in La To Village, Huc Nghi Commune, Da Krong District, Quang Tri Province" (*VNU - Central Institute for Natural Resources and Environmental Studies*, March 29, 2012) https://cres.edu.vn/kin-thc-bn-a-trong-qun-ly-va-s-dng-thc-vt-lam-sn-ngoai-g-ca-cng-ng-ngi-van-kiu-thon-la-to-xa-huc-nghi-huyn-a-krong-tnh-qung-tr/ accessed May 21, 2024.

128. Dai-Long Ngo-Hoang and Nguyen Thi Thuy, "Local Knowledge Ben Tre's Coastal on Resilience to Climate Change Background Study" [2019] Social Science Research Network https://doi.org/10.2139/ssrn.3367754.

129. Vu Hoi, "Preserve and Promote the Culture of Ethnic Minorities" *Nhan Dan (People) Online Newspaper* (July 29, 2022) https://nhandan.vn/bao-ton-phat-huy-van-hoa-cac-dan-toc-thieu-so-post707842.html.

trained in English,[130] with "relatively little data from Vietnam to access in that language",[131] issues related to data diversity, bias, and being underrepresented in the training data are inevitable. Hence, the accuracy of generated content regarding Vietnamese culture and history is likely to be lessened,[132] which would result in the spread of misleading content or the creation of non-existing information. Together with GenAI models developed by international Big Tech, companies in Vietnam have joined the race to develop GenAI models that are customized for the domestic market, with VinGroup's ViGPT being the first homegrown GenAI model of the country, aiming to embed local languages and cultures in GenAI technology.[133] Whether GenAI models are developed with a focus on the Vietnamese market or to apply on an international scale, establishing a regulatory framework to govern the development and operation of those models would be significant to ensure that GenAI-related risks for IK are identified and addressed properly. Taking this demand into consideration, this section analyses issues in the current regulatory framework of Vietnam related to AI and IK. The section then discusses the authors' proposals regarding how the laws should be amended to mitigate potential GenAI models' negative impact on the knowledge of Vietnamese ethnic minorities, based on exisitng issues in Vietnam's legal framework, as well as lessons learned from international perspectives.

## 5.1 Issues in current regulatory framework regarding AI and Indigenous knowledge

The Vietnamese laws define "*ethnic minority*" as an ethnic group with a population smaller than that of the ethnic majority group within the territory of Vietnam.[134] Acknowledging the significance of ethnic minorities' IK for the country's socio-economic development, the Government of Vietnam sets out the framework for protection of these values. Particularly, the Law on Cultural Heritage 2024 specified a framework to protect and promote values of intangible cultural heritage, which includes knowledge, skills, customs, cultural expressions and related objects, artifacts, and spaces practiced and transmitted by communities, groups, and individuals through many generations and forming their cultural identity, constantly transmitted, protected, recreated, created, and having historical, cultural, and scientific value.[135] With a focus on ethnic minorities, the Decree No. 05/2011/ND-CP specified five types of policy regarding culture conservation and development for ethnic minorities in Vietnam, including policies on traditional cultural values, languages, historical and cultural relics, enjoyment of cultural benefits, as well as traditional festivals.[136] These regulations, however, only state general policies regarding intangible cultural heritage in whole, or cultural of Vietnam's ethnic minorities in part, plus lacking concentration on issues caused by AI.

Regarding AI, at present, the Vietnamese legislation is silent on specific regulations on this issue. However, the Government of Vietnam, in accordance with Decision No. 127/QD-TTg,[137] has taken a proactive approach to the establishment of an AI-specific regulatory framework. Particularly, Decision No. 127/QD-TTg acknowledges the intrinsic need for AI-related regulations, with the law-making process concentrating on, *inter alia*, "developing and applying AI with people ... at the center", "avoiding technology abuse and infringement upon legitimate rights and interests of ... individuals" and "avoiding disclosure and leakage of private data of individuals".[138] As Decision No. 127/QD-TTg outlines only key targets and directions for developing and applying AI in Vietnam, further guidance need to be provided in the future to create a clear path forward.[139]

130. Madeleine North, "Generative AI Is Trained on Just a Few of the World's 7,000 Languages. Here's Why That's a Problem – and What's Being Done about It" (*World Economic Forum*, May 17, 2024) https://www.weforum.org/agenda/2024/05/generative-ai-languages-llm/ accessed May 21, 2024.
131. Yuji Nitta, Shunsuke Tabeta and Kotaro Hosokawa, "Vietnam Jumps into Made-in-Asia Generative AI Race" *Nikkei Asia* (January 24, 2024) https://asia.nikkei.com/Business/Technology/Vietnam-jumps-into-made-in-Asia-generative-AI-race.
132. Ibid.
133. Ibid.
134. Decree No. 05/2011/ND-CP, Article 4.2.
135. Law on Cultural Heritage 2024, Article 3.1.
136. Decree No. 05/2011/ND-CP, Article 13.
137. Decision No. 127/QD-TTg of the Prime Minister dated 26 January 2021 on Issuing the National Strategy on Research, Development and Application of Artificial Intelligence until the Year 2030. The official English translation of the Decision No. 127/QD-TTg can be found at: https://en.baochinhphu.vn/national-strategy-on-rd-and-application-of-artificial-intelligence-11140663.htm accessed 24 May 2024.
138. Decision No. 127/QD-TTg, Section III.
139. Mai Chi Tran, "AI in Vietnam: A Quick Guide" (*Vietnam Briefing*, August 2, 2023) https://www.vietnam-briefing.com/news/ai-in-vietnam-opportunities-and-challenges.html/ accessed May 24, 2024.

Regulations on personal data protection, which are specified under Decree No. 13/2023/ND-CP,[140] is also noteworthy. As the name suggests, this Decree provides protection for personal data of individuals who are identified by such data (i.e. the data subject), with its scope being the territory of Vietnam.[141] Decree No. 13/2023/ND-CP, however, lacks guidance about how to protect personal data processed[142] by GenAI models. Particularly, apart from five exceptions as provided under Article 17,[143] all activities related to personal data processing require data subjects' consent.[144] Hence, it is uncertain about the possibility of obtaining consent from all individuals whose data have been processed to train GenAI models regarding the countless numbers of purposes for which their data has been used. Furthermore, how to implement data subjects' rights over their personal data processed by GenAI models, such as the right to edit or request to edit personal data[145] and the right to delete or request for deletion of personal data,[146] remains unclear, considering the fact that modifying or removing data embedded in the models would be expensive and time-consuming.

Considering the importance of Vietnamese ethnic minorities' IK, the urgent need for preserving such knowledge, as well as potential negative impact on IK resulting from the performance of GenAI models, the lack of a regulatory framework on AI and existing issues related to personal data, protecting ethnic minorities' IK would be more challenging in the era of AI. Hence, it is necessary to establish a regulatory framework specifically governing and providing clear guidelines for other uncertain legal issues to ensure that GenAI's negative impact on IK is identified and mitigated properly under the laws of Vietnam.

### 5.2 Proposals for the laws of Vietnam to mitigate GenAI's negative impact on Indigenous knowledge of ethnic minorities

As the Government of Vietnam highlights international cooperation in improving institutions and policies related to AI,[147] as well as in law-making and practice of personal data protection,[148] taking inspiration from approaches from UNESCO and other countries (i.e the EU and Australia) would be essential to ensure how the laws of Vietnam mitigate negative impact of GenAI on IK follow international standards.

Particularly, the establishment of a Vietnamese AI-specific regulatory framework, in comparison with UNESCO's Recommendation on the Ethics of Artificial Intelligence and the EU's AI Act, should include the following regulations:

*Regarding the general development and operation of AI models*, AI-related regulations should (i) require that operation of AI models respects cultural diversity, with marginalized groups (including ethnic minorities) participating in AI development; (ii) ensure AI's cultural impact are examined to identify impact, predict consequences, mitigate risks and avoid harmful consequences of AI models on local customs and traditions, as well as to promote the visibility and discoverability of local content; and (iii) emphasize that specific needs of all groups/communities being considered when sharing AI benefits.

140. Decree No. 13/2023/ND-CP of the Government dated 17 April 2023 in Personal Data Protection. An unofficial English translation of the Decree No. 13/2023/ND-CP can be found at: https://eurochamvn.org/wp-content/uploads/2023/02/Decree-13-2023-PDPD_EN_clean.pdf accessed 25 May 2024.
141. Decree No. 13/2023/ND-CP, Articles 1 and 2(5).
142. Decree No. 13/2023/ND-CP, Article 2 (7) defines the processing of personal data as "one or more operations that affect personal data, such as: obtaining, recording, analysis, confirmation, storage, alteration, publicity, combination, access, retrieval, recovery, encryption, decryption, duplication, sharing, transmission, provision, transfer, deletion, destruction of personal data or other relevant operations".
143. Decree No. 13/2023/ND-CP, Article 17 specifies five exceptions where personal data is allowed to process without a data subject's consent, namely emergency cases to protect the life and health of the data subject or others, disclosure of personal data in accordance with the law, a state competent authority processing personal data in a state of emergency, fulfilment of the data subject's contractual obligations and serving the operations of state agencies as prescribed by specialized laws.
144. Decree No. 13/2023/ND-CP, Article 11(1).
145. Decree No. 13/2023/ND-CP, Article 9(3).
146. Decree No. 13/2023/ND-CP, Article 9(5).
147. Decision No. 127/QD-TTg, Section III(5).
148. Decree No. 13/2023/ND-CP, Article 7.

*Regarding datasets to train AI models*, the high quality of those datasets should be ensured by implementing requirements such as data being sufficiently representative or considering features, characteristics or elements that are particular to specific circumstances.

R*egarding content generated by AI models*, all outputs should be required to be marked as "artificially generated" to avoid biases, misinformation and manipulation by using those outputs.

In terms of regulations on personal data protection, it is fundamental to provide guidelines on implementing current laws in the context of GenAI models' development and operation, with some worth-noting issues being asking for data subjects' consent in case their personal data is used to train GenAI models, and the enforcement of personal data-related rights, such as the right to edit or the right to delete.

Furthermore, to ensure that IK of ethnic minorities is protected under IP laws, the approach of Australia (i.e. establishment of specific IP rights concentrating on ownership and protection of IK, in combination with setting up a national authority to enforce action against unauthorized use and misappropriation of IK) would be considered.

Regardless of how the laws of Vietnam would be inspired by international approaches, regulations on mitigating GenAI's negative impact on IK of Vietnamese ethnic minorities must follow the Government of Vietnam's direction of fostering a Vietnamese culture that is rich in national identity, unified in the diversity of the ethnic communities, with national, humanistic, democratic, and scientific characteristics. In other words, as Vietnam is "a unified yet diverse culture", [149] the protection of ethnic minorities' IK diversity needs to be consistent with the maintenance of Vietnam's cultural identity unification. Moreover, as culture "being the driving force behind socio-economic development",[150] the laws must strike a balance between mitigating GenAI's negative impact on IK and nurturing AI technology's growth in Vietnam.

## 6. Conclusion

"*As the types of IK, and related concerns and objectives, are unique to each country and community, solutions must also be tailored to local circumstances*."[151]

IK is the essence of both the self-identity of ethnic minorities and Indigenous peoples who possess the knowledge, as well as the society as a whole. Despite its significance, IK is at risk of disappearing and being vulnerable to exploitation and profit-making, with the breakout of GenAI models worsening challenges of protecting IK. Particularly, negative impact of GenAI on IK result from three main reasons, namely data to train GenAI models, content generated by those models and power in the hands of Big Tech. Acknowledging the importance of IK, UNESCO has provided its Recommendation on the Ethics of Artificial Intelligence, seeking to guide AI development in a responsible direction. In addition, the EU, to ensure that the use of AI is safe and non-discriminatory, adopted the world's first comprehensive AI law. The AI Act, together with the GDPR, is expected to set a standard for other countries regarding the governance of AI and data protection. Australia has also taken its approach to better protection for IK by working on new IP rights concentrating on ownership and protection of IK. All those perspectives can serve as inspiration for Vietnam, a country with a vibrant diversity of IK, to establish its regulatory framework on mitigating GenAI's negative impact on ethnic minorities' knowledge. Regardless of how the laws of Vietnam would be inspired by international approaches, all regulations must be tailored to suit the Government of Vietnam's direction of ensuring the protection of IK is consistent with the maintenance of national cultural identity unification.

149. Giang Dinh, "Cultural value with establishment of Vietnamese culture, which is advanced and rich in national identity, creating motivation to realize the country's development aspirations" (*Communist Review*, August 29, 2022) https://www.tapchicongsan.org.vn/web/guest/van_hoa_xa_hoi/-/2018/825804/gia-tri-van-hoa-voi-xay-dung-nen-van-hoa-viet-nam-tien-tien%2C-dam-da-ban-sac-dan-toc%2C-tao-dong-luc-thuc-hien-khat-vong-phat-trien-dat-nuoc.aspx accessed May 15, 2024.

150. Ibid.

151. Sophia Twarog and Promila Kapoor, *Protecting and Promoting Traditional Knowledge: Systems, National Experiences and International Dimensions* (United Nations Publication 2004) iii.

Furthermore, regulations need to mitigate GenAI's negative impact on IK while nurturing AI technology's growth in Vietnam. Taking into consideration how each country pursues its own perspective in governing AI and IK-related issues, the research of GenAI's negative impact needs to be performed on a case-by-case basis, considering socio-economic conditions, laws and policies of the country where ethnic minorities or Indigenous people are residing, to propose the most feasible solution for preserving and promoting their IK. Considering the fact that GenAI models are still in the early stages of development, plus those models' impact on IK requires further analysis in detail, this paper aims to be a preliminary work on the topic to uncover issues related to GenAI's negative impact on IK, with it serving as a suggestion for future research on related topics.

# 7. REFERENCES

## 7.1 Primary sources

1. UN Convention on Biological Diversity (1992)
2. European Union's Artificial Intelligence Act
3. European Union's General Data Protection Regulation
4. Decree No. 13/2023/ND-CP of the Government dated 17 April 2023 in Personal Data Protection. An unofficial English translation of the Decree No. 13/2023/ND-CP
5. Decision No. 127/QD-TTg of the Prime Minister dated 26 January 2021 on Issuing the National Strategy on Research, Development and Application of Artificial Intelligence until the Year 2030.

## 7.2 Secondary sources

### *7.2.1 Articles*

6. Dinh G, "Cultural value with establishment of Vietnamese culture, which is advanced and rich in national identity, creating motivation to realize the country's development aspirations" (*Communist Review*, August 29, 2022) https://www.tapchicongsan.org.vn/web/guest/van_hoa_xa_hoi/-/2018/825804/gia-tri-van-hoa-voi-xay-dung-nen-van-hoa-viet-nam-tien-tien%2C-dam-da-ban-sac-dan-toc%2C-tao-dong-luc-thuc-hien-khat-vong-phat-trien-dat-nuoc.aspx accessed May 15, 2024
7. Gottlieb I, "EU Poised to Enact Sweeping AI Rules With US, Global Impact (1)" *Bloomberg Law* (March 6, 2024) https://news.bloomberglaw.com/artificial-intelligence/eu-poised-to-enact-sweeping-ai-rules-with-us-global-impact
8. Han MT, "Indigenous knowledge in plant management and use of Non-Timber Forest products Of the Van Kieu community in La To Village, Huc Nghi Commune, Da Krong District, Quang Tri Province" (*VNU - Central Institute for Natural Resources and Environmental Studies*, March 29, 2012) https://cres.edu.vn/kin-thc-bn-a-trong-qun-ly-va-s-dng-thc-vt-lam-sn-ngoai-g-ca-cng-ng-ngi-van-kiu-thon-la-to-xa-huc-nghi-huyn-a-krong-tnh-qung-tr/ accessed May 21, 2024
9. Littletree S, Belarde-Lewis M and Duarte M, "Centering Relationality: A Conceptual Model to Advance Indigenous Knowledge Organization Practices" (2020) 47 Knowledge Organization (KO) 410 https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/46601/ko_47_2020_5_e.pdf?sequence=1&isAllowed=y
10. Lewis JE and others, "Indigenous Protocol and Artificial Intelligence Position Paper" [2020] https://spectrum.library.concordia.ca/986506/
11. Moran C, "ChatGPT Is Making up Fake Guardian Articles. Here's How We're Responding" *The Guardian* (April 6, 2023) https://www.theguardian.com/commentisfree/2023/apr/06/ai-chatgpt-guardian-technology-risks-fake-article
12. Ngo-Hoang D-L and Thuy NT, "Local Knowledge Ben Tre's Coastal on Resilience to Climate Change Background Study" [2019] Social Science Research Network https://doi.org/10.2139/ssrn.3367754
13. Ngulube P, "Managing and Preserving Indigenous Knowledge in the Knowledge Management Era: Challenges and Opportunities for Information Professionals" (2002) 18 Information Development 95 https://doi.org/10.1177/026666602400842486
14. Nguyen HTS, "Policy to preserve and promote cultural values of ethnic minorities in Vietnam" (*Communist Review*, August 12, 2023) https://www.tapchicongsan.org.vn/web/guest/van_hoa_xa_hoi/-

/2018/828114/chinh-sach-bao-ton-va-phat-huy-gia-tri-van-hoa-cac-dan-toc-thieu-so-o-viet-nam.aspx accessed May 20, 2024
15. Nitta Y, Tabeta S and Hosokawa K, "Vietnam Jumps into Made-in-Asia Generative AI Race" *Nikkei Asia* (January 24, 2024)https://asia.nikkei.com/Business/Technology/Vietnam-jumps-into-made-in-Asia-generative-AI-race
16. Son HN, Chi DTL and Kingsbury A, "Indigenous Knowledge and Climate Change Adaptation of Ethnic Minorities in the Mountainous Regions of Vietnam: A Case Study of the Yao People in Bac Kan Province" (2019) 176 Agricultural Systems 102683 https://doi.org/10.1016/j.agsy.2019.102683
17. Tapu IF and Fa'Agau TK, "A New Age Indigenous Instrument: Artificial Intelligence & Its Potential for (De)Colonialized Data" (2022) 57 Harvard Civil Rights-Civil Liberties Law Review 715 https://journals.law.harvard.edu/crcl/wp-content/uploads/sites/80/2023/01/ANewAgeIndigenousInstrument.pdf
18. Tsosie RA, "Tribal Data Governance and Informational Privacy: Constructing 'Indigenous Data Sovereignty'" (2019) 80 Montana Law Review 229 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3454632
19. Vu H, "Preserve and Promote the Culture of Ethnic Minorities" *Nhan Dan (People) Online Newspaper* (July 29, 2022) https://nhandan.vn/bao-ton-phat-huy-van-hoa-cac-dan-toc-thieu-so-post707842.html

### 7.2.2 Books

20. Animikii Inc., *DataBack: Asserting & Supporting Indigenous Data Sovereignty* (2023) https://databack.animikii.com/
21. Bradford A, "The Brussels Effect," *Oxford University Press eBooks* (2019) https://doi.org/10.1093/oso/9780190088583.003.0003
22. Confederation of European Data Protection Organisations, *Generative AI: The Data Protection Implications* (2023) https://cedpo.eu/generative-ai-the-data-protection-implications/
23. Grant P, *State of the World's Minorities and Indigenous Peoples 2016: Focus on Culture and Heritage* (Minority Rights Group 2016)
24. Handbook UE, "Minorities and Indigenous Peoples" *UNHCR* (January 30, 2024) https://emergency.unhcr.org/protection/persons-risk/national-ethnic-religious-and-linguistic-minorities-and-indigenous-peoples
25. OHCHR, *Minority Rights: International Standards and Guidance for Implementation* (United Nations 2010) https://www.ohchr.org/en/publications/special-issue-publications/minority-rights-international-standards-and-guidance
26. Twarog S and Kapoor P, *Protecting and Promoting Traditional Knowledge: Systems, National Experiences and International Dimensions* (United Nations Publication 2004)
27. UNESCO, *Recommendation on the Ethics of Artificial Intelligence* (UNESCO 2022) https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence

### 7.2.3 Reports

28. "Indigenous Peoples and Ethnic Minorities: Marginalization Is the Norm", *Report on the world social situation* (2018) https://doi.org/10.18356/14642ccc-en
29. "Interim Report: Scoping Study on stand-alone legislation to protect and commercialise Indigenous Knowledge" (2023) https://www.ipaustralia.gov.au/tools-and-research/professional-resources/data-research-and-reports/publications-and-reports/scoping-study-on-standalone-ik-legislation
30. Norwegian Consumer Council, "Ghost in the Machine – Addressing the Consumer Harms of Generative AI" (2023) https://www.forbrukerradet.no/side/new-report-generative-ai-threatens-consumer-rights/ accessed December 15, 2023.
31. "Scoping Study on Standalone Legislation to Protect and Commercialise Indigenous Knowledge - Final Report" (2023) https://www.ipaustralia.gov.au/tools-and-research/professional-resources/data-research-and-reports/publications-and-reports/scoping-study-on-standalone-ik-legislation
32. Simons E and others, "Traditional Knowledge - Fact Sheet" (2016) https://www.sfu.ca/ipinch/sites/default/files/resources/fact_sheets/ipinch_tk_factsheet_march2016_final_revised.pdf accessed April 27, 2024

### 7.2.4 Websites


33. "About Us" (*IP Australia*) https://www.ipaustralia.gov.au/about-us accessed May 19, 2024
34. Australian Institute of Aboriginal and Torres Strait Islander Studies, "Australia's First Peoples" (*AIATSIS*) https://aiatsis.gov.au/explore/australias-first-peoples#toc-what-term-is-best-to-use- accessed May 19, 2024
35. Beyond Team, "Is Your Data Ready for Generative AI? A Comprehensive Guide" (*Beyond Team*, April 5, 2024) https://www.puttingdatatowork.com/post/is-your-data-ready-for-generative-ai-a-comprehensive-guide#:~:text=The%20bedrock%20of%20GenAI's%20effectiveness,both%20vast%20and%20meticulously%20curated accessed May 5, 2024
36. "Briefing EU Legislation in Progress - Artificial Intelligence Act" (*European Parliment*, March 2024) https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf accessed May 12, 2024
37. Bryant J, "3 Years in, GDPR Highlights Privacy in Global Landscape" (*IAAP*, May 25, 2021) https://iapp.org/news/a/three-years-in-gdpr-highlights-privacy-in-global-landscape accessed May 18, 2024
38. Chandran R, "Indigenous Groups in NZ, US Fear Colonisation as AI Learns Their Languages" (April 3, 2023) https://www.context.news/ai/nz-us-indigenous-fear-colonisation-as-bots-learn-their-languages accessed May 5, 2024
39. Luu C, "What We Lose When We Lose Indigenous Knowledge" [2023] JSTOR Daily https://daily.jstor.org/what-we-lose-when-we-lose-indigenous-knowledge
40. "Ethnic Groups in Vietnam" (*Vietnam Government Portal*) https://vietnam.gov.vn/ethnic-groups-in-vietnam accessed May 21, 2024
41. "Ethnic Minority Empowerment | Oxfam in Vietnam" (*Oxfam in Vietnam*) https://vietnam.oxfam.org/ethnic-minority-empowerment accessed December 31, 2023
42. "EU AI Act: First Regulation on Artificial Intelligence" (*European Parliament*, August 6, 2023) https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence accessed May 12, 2024
43. Hanna KT and Wigmore I, "Definition: Historical Data" (*Tech Target*, November 2022) https://www.techtarget.com/whatis/definition/historical-data accessed May 5, 2024
44. "High-Level Summary of the AI Act | EU Artificial Intelligence Act" (*EU Artificial Intelligence Act*, February 27, 2024) https://artificialintelligenceact.eu/high-level-summary/ accessed May 12, 202
45. "How to Protect Indigenous Knowledge and Creative IP from Exploitation" (*The University of Melbourne*, October 2, 2023) https://study.unimelb.edu.au/study-with-us/professional-development/blog/how-to-protect-indigenous-knowledge-and-creative-ip-from-exploitation accessed April 29, 2024
46. Kaplan J, "More Speech and Fewer Mistakes" *Meta* (January 7, 2025) https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/
47. "Leveraging UNESCO Normative Instruments for an Ethical Generative AI Use of Indigenous Data" (*UNESCO*, November 8, 2023) https://www.unesco.org/en/articles/leveraging-unesco-normative-instruments-ethical-generative-ai-use-indigenous-data#:~:text=Many%20indigenous%20communities%20are%20striving,knowledge%2C%20especially%20among%20younger%20generations accessed May 4, 2024
48. "Local and Indigenous Knowledge Systems (LINKS)" (UNESCO, January 6, 2022) https://en.unesco.org/links accessed December 31, 2023
49. North M, "Generative AI Is Trained on Just a Few of the World's 7,000 Languages. Here's Why That's a Problem – and What's Being Done about It" (*World Economic Forum*, May 17, 2024) https://www.weforum.org/agenda/2024/05/generative-ai-languages-llm/ accessed May 21, 2024
50. Nguyen PT, "Strive to build, preserve and promote the unique values of Vietnamese culture, which is advanced and rich in national identity" (*Communist Review*, November 24, 2021) https://www.tapchicongsan.org.vn/media-story/-/asset_publisher/V8hhp4dK31Gf/content/tiep-tuc-xay-dung-giu-gin-chan-hung-va-phat-trien-nen-van-hoa-cua-dan-toc accessed May 22, 2024
51. OpenAI, "Introducing APIs for GPT-3.5 Turbo and Whisper" (*OpenAI*, April 24, 2024) https://openai.com/index/introducing-chatgpt-and-whisper-apis/ accessed May 5, 2024
52. OpenAI, "Introducing ChatGPT" (*OpenAI*, November 30, 2022) https://openai.com/index/chatgpt/ accessed May 17, 2024
53. "Ostering Cultural Diversity, Intercultural Dialogue and a Culture of Peace in Tanzania" (*UNESCO*, September 14, 2023) https://www.unesco.org/en/fieldoffice/daressalaam/expertise/fosteringculturaldiversity accessed December 29, 2023

54. Schmelkes S, "Epistemic Justice and the Knowledge Commons for Lifelong and Lifewide Learning" (*UNESCO*, January 9, 2023) https://www.unesco.org/en/articles/epistemic-justice-and-knowledge-commons-lifelong-and-lifewide-learning accessed December 29, 2023
55. "Socialist Republic of Vietnam" (*Ethnologue*) https://www.ethnologue.com/country/VN/ accessed May 21, 2024
56. Strandell J and Strandell J, "What Is Content Moderation? (Plus Best Practices)" (*Besedo*, May 24, 2024) https://besedo.com/blog/what-is-content-moderation accessed May 26, 2024
57. "The General Data Protection Regulation" (*European Council*) https://www.consilium.europa.eu/en/policies/data-protection/data-protection-regulation/#:~:text=The%20EU%20general%20data%20protection%20regulation%20(GDPR)%20is%20the%20strongest,application%20on%2025%20May%202018. accessed May 16, 2024
58. "The State of AI in 2023: Generative AI's Breakout Year" (*McKinsey & Company*, August 1, 2023) https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year accessed December 28, 2023
59. "Traditional Knowledge" (*UNESCO UIS*, September 12, 2023) https://uis.unesco.org/en/glossary-term/traditional-knowledge accessed April 27, 2024
60. Tran MC, "AI in Vietnam: A Quick Guide" (*Vietnam Briefing*, August 2, 2023) https://www.vietnam-briefing.com/news/ai-in-vietnam-opportunities-and-challenges.html/ accessed May 24, 2024
61. "UNESCO in Brief" (*UNESCO*) https://www.unesco.org/en/brief accessed May 8, 2024
62. "Walking the Tightrope: As Generative AI Meets EU Regulation, Pragmatism Is Likely" (*Deloitte Insights*, January 18, 2024) https://www2.deloitte.com/xe/en/insights/industry/technology/technology-media-and-telecom-predictions/2024/tmt-predictions-eu-generative-ai-regulation.html accessed May 14, 2024
63. "Why Is Cultural Diversity Important?" (*Partnership International*) https://www.partnershipinternational.ie/why-is-cultural-diversity-important/#:~:text=It%20helps%20dispel%20negative%20stereotypes,respect%2C%20and%20understanding%20across%20cultures accessed December 29, 2023