AIIA

In support of UNESCO

unesco

Democratic Infrastructure for Creative Futures: Building the AI, IP & Culture Repository

Written by Valentine Goddard and Dr. Leslie Salgado Arzuaga
with expert editorial collaboration from Farah Fancy, Karina Kesserwan and Chloe van Bussel

Context and Goals

Introduction

Context

The rapid expansion of generative artificial intelligence is profoundly reshaping how cultural and knowledge resources are created, shared, and governed, exposing significant gaps in existing frameworks for understanding, protection, and oversight. While intellectual property regimes (IPRs) remain one of the primary mechanisms available to artists, creators, cultural workers, and Indigenous knowledge holders to protect their work, safeguard cultural heritage, and derive fair value from their contributions, they are increasingly strained by the scale, speed and opacity of AI systems, which often rely on vast amounts of data drawn from public, proprietary and traditional knowledge sources. At the same time.

At the same time, these same frameworks can enable the privatization and appropriation of public domain knowledge and cultural commons with proprietary AI systems, creating tensions between artists' economic rights, cultural sovereignty, and broader economic development rights, particularly for communities in the Global Majority. These impacts are gendered and intersectional, disproportionately affecting women and communities whose knowledge, labour and decision making authority have been historically undervalued or excluded, contributing to labour precarity, cultural erasure, and unequal access to decision making. Furthermore, there is also a clear lack of accessible, independent, and balanced information to support civil society, cultural actors, and policymakers in navigating these complex dynamics. In this context, the creation of a dedicated, civil society-led and collaboratively designed Repository emerged as a necessary response to facilitate knowledge sharing, surface diverse perspectives, share best practices, protect digital cultural sovereignty, and support more equitable, informed, rights-based, and culturally sensitive approaches to AI governance in the cultural sphere.

What is the Repository?

A civil society-led Repository on AI, intellectual property regimes (IPRs) and culture (hereafter the "Repository") would be a centralized, organized, co-designed online space where information, resources, and knowledge are collected, stored, and shared. It would help navigate a growing tension" while IP frameworks remain essential tools for artists, creators, cultural workers and Indigenous Peoples and knowledge holders to protect their work, safeguard cultural heritage, and secure fair remuneration, they can also enable the privatization and appropriation of public domain, traditional, and Indigenous knowledge within proprietary AI systems, affecting digital cultural sovereignty, freedom of expression, and broader economic development rights. A dedicated Repository is needed precisely to inhabit this

tension: to document and analyze how AI and IP are reshaping cultural and knowledge commons; to make visible the gendered and intersectional impacts of these transformations; include creative advocacy materials, policy commentary, case studies, legal and legislative trackers, and community led responses; and to hold a structured, ongoing debate, grounded in artists, cultural and civil society organizations (CSO) realities, on both the protective and constraining roles of IP in the digital era. Such a space should emerge from co-design with those most affected, not be imposed as a ready-made technical solution, ensuring more equitable, rights-based, and culturally sensitive access to independent and balanced information.

UNESCO's Subgroup on Intellectual Property (IP) and Culture

UNESCO's Subgroup on Intellectual Property (IP) and Culture is dedicated to developing gender-responsive, rights based policy recommendations on the impact of generative AI on creative ecosystems, art production value-chains and the preservation of digital and cultural commons. In a context where digital cultural sovereignty is increasingly urgent, the subgroup highlights the importance of civil society and artist participation in ensuring the ethical governance of AI.

The Open Call & Co-design workshop

To ground the Subgroup's work in diverse, real-world perspectives, an international Open Call for contributions was launched, resulting in 43 submissions, of which 23 were selected[1]. Building on these inputs, a multi-stakeholder co-design workshop was held on April 8th and 13th, 2026, convening 42 participants. The findings presented in this report draw from the analysis of these contributions, the discussions held during the workshop, and the broader research conducted throughout the project.

This initiative was launched with the objective of co-designing a Repository to facilitate independent and balanced knowledge sharing on the impact of AI on creative ecosystems, art production and related value-chains, with a focus on gender-responsive, rights-based policy implications. The Repository aims to:

- Facilitate knowledge sharing and informed, ongoing dialogue on the impacts of AI on culture, intellectual property, and knowledge systems;
- Support informed and sovereign decisions-making by communities, creators and civil society as a whole;

[1] A comprehensive document compiling all accepted submissions is viewable at this link: https://docs.google.com/document/d/1V29VrXWwC463xdteCHr0z0A8_tDcd-hpRwOKETs7hYA/edit?usp=sharing

- Protect digital cultural sovereignty and preserve cultural and knowledge commons.
- Prevent the non-consensual commodification and militarization of knowledge, cultural erasure, and the concentration of power in knowledge creation and dissemination including Indigenous Peoples' knowledge systems and ways of knowing, while safeguarding the freedom of expression, addressing artists' labour precarity and economic security.
- Highlight community-driven responses such as licensing frameworks, data trusts, and collaborative governance models, while supporting policy reform aligned with data, privacy, human rights and broader sovereignty protections.

Throughout this report, we systematically draw from the Open Call's goals, the IP and Culture Subgroup's concept note priorities, pre-workshop research, and—most critically—the insights from the 23 accepted proposals. Our approach identifies key needs emerging from these sources, then reframes them as actionable design priorities for the Repository, ensuring community-driven solutions remain grounded in participant voices, lived realities, and shared commitment to equitable, rights-based, and culturally sensitive AI governance.

Conceptual Framework

Before we get started, allow us to clarify what we mean by art and cultural data. The Open Call defined art as “all kinds of art forms, from word to weaving, from sound to film, from poetry to pottery”. Submissions stressed that artisan work is art, that translation is culture, that metadata about libraries and archives, and even financial data is culture. Financial data includes patterns of saving, spending, reciprocity, and surviving, and can reflect how communities organize care, navigate constraint, sustain identity and make meaning from expressions of community life. Therefore, for the purpose of this report, art, culture and cultural data are understood broadly and grounded in what communities identify as culture, rather than what institutions define for them. Furthermore, this links to concepts of Cultural Intellectual Property defined by the Canada Council for the Arts *as Indigenous Peoples’* right to own and control their cultural heritage. Indigenous heritage is used to refer to these rights to bring the terminology more in line with language being used in international forums. Indigenous heritage comprises all objects, sites and knowledge transmitted from generation to generation. Indigenous peoples heritage is a living heritage.

Interconnected Clusters

From the selected proposals emerge resources that contribute to the conversation about the impact of AI on creative ecosystems, art production and related value-chains. These resources include existing community initiatives, research, and technical solutions, three interconnected clusters where researchers and practitioners focus on. This body of work can be grouped in three interconnected clusters: 1) projects that seek to **examine** the impacts and implications of Generative AI in culture, 2) projects seeking to **transform** the landscape through advocacy, literacy and legal support, and 3) and projects that **operationalize** solutions for the challenges AI brings to cultural rights. The next sections unpack how selected proposals contribute to each cluster. However, it is essential to note that many proposals contribute to more than one area, because of their work spans from critical examination to governance design and operational implementation.

1) Projects examining the scene of challenges in AI, cultural rights and IP

A group of proposals focuses on critically mapping the challenges at the intersection of AI, cultural rights, and IP. Contributions such as *The Translator's Right to Exist: Protecting Human Agency in AI Governance*,[2] *Algorithmic Colonial Gaze and Gender Bias in Generative AI*[3]*, Artificial Intelligence and Gender Inequality*[4] and *On hegemonic AI*[5] *discourse* expose how AI systems reproduce structural inequalities, including cultural erasure, gender and racial bias, and the devaluation of human creative labor. These works underline that AI often relies on unconsented cultural extraction while reinforcing dominant epistemologies that exclude, undervalue, and override Global South perspectives, Indigenous Peoples, and other ancestral knowledge holders, while also threatening freedom of expression and cultural sovereignty.

Other proposals deepen this idea by examining legal and economic tensions, including *EU Data Governance and Access Regimes: Implications for Cultural Commons, AI Training, and Intellectual Property Governance*, *On financial data as cultural data*[6]*, On AI in the heritage sector*[7]*, Research at the Intersection of AI, Labor, and Cultural Economy*[8] and *Artificial Intelligence and the Cultural Sector in Senegal: Challenges, Opportunities, and Perspectives*[9]. They highlight how current IP and data governance frameworks fail to address collective knowledge production, cultural intellectual property, while enabling enclosure of cultural commons, and exacerbating inequalities in access,

[2] Andreea Irina Acasandrei
[3] Hannah Maruci Aflalo
[4] Barbara Pozzo
[5] Dr. Andreas Giorgallis, Dr. Aline Iramina and Dr. Charles Mak
[6] Barbara Furiati
[7] Paula Westenberger
[8] Nicoletta Kolpakov on behalf of The Cirrus Institute Future of Work Initiative
[9] Kofi Sika Latzoo

representation, and benefit sharing. These contributions are critical in showing how existing AI and IP systems can reproduce inequity rather than protect cultural rights.

2) Projects seeking to transform the scene through cultural, technological, and governance practices

A second cluster of proposals seeks to transform the current AI landscape by advancing alternative governance models, ethical frameworks, and community-driven infrastructures. Contributions such as *Territorial AI and Cultural Sovereignty: GriôTech as a Public-Interest Infrastructure from the Global South*[10], *Cultural Data Trusts & Participatory AI Governance for the Digital Cultural Commons*[11], *Operationalising Indigenous Authority in AI and Cultural Data Governance*[12] and *Ethical Guidelines Working Group on AI and Indigenous Data Management*[13] propose participatory and culturally grounded approaches to AI governance. These initiatives emphasize Indigenous Peoples' authority over cultural heritage, Indigenous data sovereignty, community decision-making power, and the need to embed cultural values and rights directly into technological systems.

Additional transformative perspectives are offered by *Beyond Private Ownership: Safeguarding Indigenous Knowledge Systems from AI Commodification*[14], *A Human Rights–Based Approach to IP Law as a Counterbalance to the Systemic Risk of AI-Generated Cultural Monism*[15], *The Importance of Worker-led AI governance*[16] and *The Enduring Wisdom of Oral Traditions in the Age of AI: A Kurdish Perspective on Cultural Sovereignty and Intellectual Property*[17]. These works advocate for pluralistic, rights-based, and culturally sensitive approaches that reconfigure how knowledge, heritage, and creativity are governed in the current scenario where AI systems permeate our life. They are critical because they provide tools to rethink AI, IP and cultural rights in ways that protect freedom of expression, address labour precarity, and prevent the concentration of power in knowledge creation and dissemination. These proposals aim to redefine this relationship rather than simply adapting existing systems.

3) Operational proposals offering practical solutions

A third group of proposals focuses on operational and implementable solutions to address AI-related challenges to cultural rights. For example, *Provenance Infrastructure as a Safeguard for Cultural Commons in the Age of Generative AI*[18] introduces concrete technical systems for tracking data

[10] Marcelle Chagas
[11] Mika (Jaeyun) Noh
[12] Lisa Pizzoni
[13] Isabel Beirigo
[14] Sophia Couto Bittencourt on behalf of LACONTRA (The Laboratory for Intellectual Property Rights and Traditional Knowledge)
[15] Francesca Rotolo
[16] Rafael Grohmann
[17] Chiman Salih
[18] Katherine Elkins and Jon Chun on behalf of Human-Centered AI Lab

provenance and enforcing rights-aware AI training practices. Similarly, *Dataset Descriptor Standard for Training AI on African Music Traditions*[19] and *Metadata as Memory: Heritage Professionals' Strategies for Safeguarding Cultural Documentation Against the Risks of Generative AI*[20] provide practical tools and standards for managing cultural data responsibly, and strengthening community authority over how knowledge is accessed and used.

Other applied contributions include *CISAC and Professional Media Production (PMP) Study: Evidence-Based Insights on AI, Copyright and Creator Remuneration*[21], *Operationalising Indigenous Authority in AI and Cultural Data Governance, Wikispeech – LLM for open source speech synthesis*[22] and *Empowering Khayamiya Artisans in Historic Cairo: A Model for Protecting Cultural Sovereignty through the 'Lifelong Learner' Initiative*[23]. These proposals translate policy concerns into actionable mechanisms, ranging from licensing frameworks, fair remuneration models, and open-source tools for speech synthesis, to governance models, cultural IP protections, and capacity-building initiatives. As a common denominator across them is the idea of ensuring that cultural actors (whether they are individual artists or organizations) and institutions can actively respond to the challenges posed by AI.

Key Design Priorities

The following themes were guided by the AI, IP and Culture subgroup to frame the Repository's scope and strategic priorities. Please note that this list does not exclude other design priorities, nor does it mean other aspects of the Open Call are not important, it simply identifies who we could reach and existing initiatives. Some questions in the Open Call are particularly difficult to fund and therefore there are no existing initiatives. This does not mean they are not critical, and we will continue to seek out and amplify these emerging voices. For example, the lack of funding for research, or solution-oriented initiatives, on the gendered impact on cultural rights was noted. Therefore, it is important to keep this in mind as we move forward in our evolutive identification of design and content priorities for the Repository.

Further, please note that the summary of the themes we received is aimed at guiding the design, that is, the building of the Repository which will evolve into a living online space to facilitate dialogue, independent and quality, lobby-free knowledge-sharing, and support sovereign decisions about cultural

---

Human-Centered AI Lab. PIs representing the Modern Language Association, AI Consortium, NIST PIs, "Archival Intelligence," Schmidt Sciences HAVI.

[19] John Oriwo

[20] Giulia Taurino

[21] Anna Neale on behalf of the International Confederation of Societies of Authors and Composers (CISAC)

[22] Signe Krantz

[23] Miral Victor Youssef on behalf of the General Administration for Cairo Heritage Preservation – Office of the Governor of Cairo

data rights protection. The Repository we envision is not a static archive of reports and initiatives, but as a living governance infrastructure, an assemblage of people and practices, a space where law, culture, and AI co-evolve under the direction of artists and cultural organizations.

The Co-design process established a network of projects, practitioners and researchers giving participants the opportunity to meet and connect the project leaders during the workshop and for the planned follow-up activities including a collectively authored publication. More importantly for the task at hand, designing an online resource cluster, it brought together a network of experts interested in supporting its creation: by contributing expertise, networks and resources.

This outcome is foundational for the future of this project because it suggests the need to adopt opportunities for connection and exchange as core to the Repository. The expert network, drawn from participants in our AI, IP, and cultural knowledge Repository Co-design workshop—with many pledging ongoing support—demonstrates exceptional interdisciplinary expertise, uniting over 25 organizations like Mozilla Foundation, KU Leuven, Translation Commons, and Indigenous Data Authority from more than 25 countries, including robust Global South representation (Benin, Brazil, Egypt, Ethiopia, Fiji, India, Iraq, Kenya, Kurdistan, Senegal, Togo). From Australia to Togo, plus worldwide entities like ICOM-ICME (active across 70+ countries with strong African/Latin American presence), it integrates diverse practitioners into collaborative projects via targeted stakeholder mapping. It is noteworthy that a majority of participants were women and there was a strong representation of Indigenous-led researchers and solution-providers also took part in the workshop and expressed the desire to support the evolutive development of this co-designed Repository.

Building on this expert network from the AI, IP, and cultural knowledge Repository Co-design workshop, our network-building efforts mark a significant stride in consolidating and expanding the IP and Culture Subgroup. This subgroup operates as a dedicated working group within UNESCO's Civil Society Organization (CSO) Network on the Ethics of Artificial Intelligence, itself nested under the broader Business and Human Rights Hub. Launched to amplify diverse voices in AI governance, the CSO Network fosters multistakeholder collaboration among over 100 organizations worldwide, prioritizing equitable participation from civil society, academia, and underrepresented regions to shape ethical AI frameworks that safeguard cultural rights, digital cultural sovereignty, and human-centred innovation.

These efforts strategically integrate prior groundbreaking work from the AI Impact Alliance, including the 2019 inaugural Art Impact AI workshops held across Canada, which directly informed UNDESA policy underscoring art's pivotal role in digital governance. Building on that momentum, the seed-

funded Art-Laws Program rallied artists and cultural organizations to lead the development of a vital resource cluster addressing generative AI's economic and legal implications. Designed to remain independent from Big Tech lobbying and the risk of creative regulatory capture, where well-resourced corporations infringe on countries' sovereign cultural policy decision-making, this initiative empowers grassroots practitioners to retain agency and steer AI's cultural impacts toward equity and sustainability.

The Role, Limits and Alternatives to Intellectual Property Regimes

Not surprisingly, since the Open Call was designed to interrogate legal and economic protections of artists and cultural workers, IPRs were central to the discussion. Current IP laws are increasingly failing to protect art and cultural data from commodification and militarization, while leaving communities and collective ownership unprotected.

Several contributors expose that the public domain, in theory, a commons belonging to everyone, is in practice becoming a primary feeding ground for proprietary AI systems. Proposals discussing this topic underline how open access, designed to democratize knowledge, can instead accelerate its enclosure. This discussion underscores the limit of IPRs to equal benefit-sharing, particularly when communities whose knowledge generates value are excluded from authority over how that value is used. In response to this, participants highlight the need to integrate into the Repository a dynamic civic watch, a civil society-led monitoring and sense-making process/space, a participatory observatory on emerging legal pathways, actively engaging community, artists and cultural rights holders in monitoring, analyzing, and responding to AI's IP and cultural impacts.

Indeed, several contributors identify legal innovations from jurisdictions typically absent from AI governance debates. These include Djibouti's December 2025 legislative framework "explicitly addressing AI in relation to cultural heritage[24]." The Brazilian Indigenous mapping describes the ongoing "Marco Legal de Proteção dos Conhecimentos Tradicionais e Culturas Populares" and the Brazilian General Data Protection Law (LGPD), which classifies "ethnic and racial origins as sensitive personal data." These proposals illustrate approaches that recognize culture, community authority, and sovereignty as central rather than peripheral to legal design.

As per the goals of the IP and Culture group and the ones set out in the Open Call, any adaptations to IPRs should be grounded in a gender-responsive and fundamental human rights-based approach; and

[24] In their proposal On Hegemonic AI Discourse, Dr. Andreas Giorgallis, Dr. Aline Iramina and Dr. Charles Mak discuss underrepresented case studies, including Djibouti's legislative framework and Ubuntu-informed ethics in Zimbabwe's AI Strategy (2026-2030).

proactive efforts to ensure they are not directed by industry or other power appropriating lobbying efforts to capture the value of cultural data. Legal reform must also protect freedom of expression, equitable benefit-sharing, and the rights of creators, Indigenous Peoples, and other ancestral knowledge holders to define the terms under which their cultural knowledge is accessed, used, and governed.

Artists and Cultural Workers' Rights in the AI Production Pipeline in Creative Economies

Not all cultural workers around the world have the same degree of worker power, whether institutional, discursive, or economic, not all artists have the same right or ability to associate, organize or participate in decisions that shape the future of their work and livelihoods. Many have asked not only for greater access to information, but for spaces to exchange with other artists globally, to make sense of these changes and make informed decisions that will impact the future of their art and livelihoods. Participants' work addresses a structural asymmetry where those who generate cultural value and data are often not the ones who capture its value or hold authority over how it is used. A connected theme that emerged in support of a need for a Repository rests on the authority to make decisions, and the consistent finding that this authority has been displaced or denied to workers, artists, and communities in the context of AI. These discussions point to the need for worker-led governance and challenge the assumption that governance power is equally shared across stakeholders. The Open Call responses emphasized the need for creators to have clearer pathways to understand, assert, and protect their rights across the AI production pipeline. The Co-design workshop began this work by creating space for participants to shape the Repository's design, while future funding would allow broader participation from artists, cultural workers, and cultural organizations who may not otherwise have the resources to contribute.

The Repository responds to a clear structural problem: AI systems too often produce cultural flattening, erasure, and homogenization, defaulting to Global North standards even when operating across languages and contexts. This can strip heritage materials of provenance and relational meaning, turning documentation into raw data, while also reinforcing gendered, racial, colonial, regional, economic, legal, and safety harms for artists, cultural workers, and knowledge holders, including labour precarity, unequal access to decision-making, and the weakening of freedom of expression. By creating a shared space for emerging research, initiatives, and practice-based resources, the Repository is designed to make visible the evidence needed to better understand these harms and to support work that actively counters them.

At the same time, the Repository is intended as a constructive intervention: a place to gather and organize knowledge that helps improve AI systems and the political, legal, and economic frameworks in

which they operate. Its goal is not only to document risks, but to help ensure that data science and machine learning do not deepen socioeconomic inequalities, and that the benefits they generate are more fairly distributed. It should support interpretation and choice, not prescribe a single path forward, allowing artists, cultural workers, and communities to make decisions based on their own contexts and authority. In that sense, the Repository becomes both a resource and a corrective, supporting artists and cultural and knowledge workers in shaping more accountable, context-sensitive, and equitable AI futures.

Cultural Authority Frameworks: Indigenous & Shared Stewardship

The proposals shared pathways to advance the governance of data and cultural knowledge in the context of AI, anchored in community and Indigenous-led perspectives. Among these approaches are included *Free, Prior, and Informed Consent (FPIC)* as the cornerstone of Indigenous cultural protection[25]; *community-governed digital cultural commons* using participatory datathons as a model for island/regional heritage governance[26]; and *Sovereign Digital Documentation,* which proposes documenting traditional craft patterns as Community-Based Cultural Data to distinguish them from AI-generated content[27]. These proposals underline the need for governance practices where communities, Indigenous Peoples, Nations, and ancestral knowledge holders whose knowledge, artistic or artisan creation are affected, are not only consulted, but hold authority to participate in decisions, to refuse participation, and to define the terms of access, use, attribution and benefit sharing.

The Āku / Ōku Framework, foundational concepts in te reo Māori (the Māori language), presented by Lisa Pizzoni proposes a distinction between data as an asset (transferable) versus data as relational, inherited, and bound through whakapapa (not freely transferable). Āku designates "that which can be held, used, or transacted," while Ōku refers to "that which is relational, inherited, and bound through whakapapa." "Current governance systems are designed to manage Āku, data as an asset , and are not equipped to recognize or enforce Ōku, data as relationship, authority, and responsibility." This

[25] In the proposal Beyond Private Ownership: Safeguarding Indigenous Knowledge Systems from AI Commodification, Sophia Couto Bittencourt discusses the Marco Legal de Proteção dos Conhecimentos Tradicionais e Culturas Populares (Legal Framework for the Protection of Traditional Knowledge and Cultural Expressions) in Brazil which intends to protect traditional expressions and knowledge against access, use, and undue exploitation by third parties without the Free, Prior, and Informed Consent (FPIC) of the communities.

[26] The proposal *Insular Data Trusts: AI and the Cycladic Cultural Commons* by Jacob Moe and Marina Markellou on behalf of Archipelago Network proposes "combining open infrastructures, data governance expertise and participatory "datathons" that invite local stakeholders to actively shape how their cultural data is described, licensed, and reused."

[27] In *Empowering Khayamiya Artisans in Historic Cairo: A Model for Protecting Cultural Sovereignty through the "Lifelong Learner" Initiative,* Miral Victor Youssef shares an initiative where traditional artisanal Khayamiya patterns are dcoumented as "Community-Based Cultural Data" to ensure that they are distinguished from AI-generated content.

distinction raises an adaptive challenge for the Repository: many existing governance systems are designed to manage ownership and access, but not relationship, responsibility, refusal, or inherited authority.

Ethical questions around what happens when cultural data is detached from its context and fed into computational processes also emerge in the proposals. These proposals signal the relevance of keeping all knowledge included in the Repository contextualized and advance provenance mechanisms as part of AI governance. They also raise technical and adaptive questions about how authority, consent, validation, access, and withdrawal can be operationalized, especially where Nations, communities, or knowledge holders face geographic, technological, linguistic, legal, or institutional barriers.

- *Provenance Infrastructure as a Safeguard for Cultural Commons* (Katherine Elkins & Jon Chun, Human-Centered AI Lab): "Without provenance infrastructure, an AI-restored image of a 1919 Creole newspaper becomes indistinguishable from AI-generated content. Its public domain status can be obscured, and its cultural context can be severed. It can also enter proprietary training datasets without the knowledge or consent of descendant communities."

- *Dataset Descriptor Standard for African Music Traditions* (John Oriwo): "African music traditions embody complex systems of rhythm, language, performances, spirituality and social meaning... My proposed African Music Dataset Descriptor Standard is designed to guide the ethical documentation, governance, and use of African music datasets intended for AI models training. The standard aims to ensure that datasets reflect the cultural origins, contextual meanings, the rights of creators and knowledge holders while enabling responsible technological innovation."

- *Metadata as Memory* (Giulia Taurino, Getty Research Institute): "Generative AI models are trained on vast, un-curated corpora, including digitized cultural heritage materials: museum catalogue records, archival finding aids, oral history transcripts, and image collections. Yet the provenance, context, and relational meaning embedded in these materials are lost during ingestion, processing and output generation. When a large language model generates a photograph's metadata without retaining its attribution, or when a text-to-image system reproduces stylistic elements of a cultural tradition without crediting its origin, the result is not merely an IP concern, but rather an act of cultural flattening. Documentation turns into raw data, and the communities who created the original records become invisible."

These proposals translate into the need for collective governance, where co-design principles, not extractivism guide the collection of experiences, practices and ideas. Opposite to the way LLM systems simplify, strip context from, or make knowledge appear generic, the Repository should seek complexity, diversity in perspectives, visible gaps, contextual integrity, and clear mechanisms to recognize authorship, respect authority, support refusal, protect provenance, and ground knowledge in context.

Data Trusts as Collective Cultural Governance Infrastructure

The same way UNESCO fights to preserve tangible and intangible cultural heritage from the impacts of war, it emerges from this workshop that as cultural data is militarized, we, as humanity and united nations, must protect cultural data from geoeconomic warfare strategies. As one of the goals of the IP and Culture group is to address the commodification and militarization of cultural data, our Open Call was designed to bring together a body of works that explores tools to prevent enclosure, establish collective rights and benefits distribution.

The examples below are not an exhaustive list of all cultural data trusts worldwide. On the contrary there are noteworthy initiatives such as Cultupedia in Quebec as well as a collective data licensing initiative (Serpentine) which were introduced during AI Impact Alliance's Art-Laws program and surely others we are not aware of. However, we are very pleased to have learned about new ones, adding to a particularly important aspect of the Repository design. Indeed, building the Repository would likely require a consent-based and context-sensitive access to the information it aims to hold. Thus a purely open source system would not be possible. Not all knowledge should be universally accessible, and openness without consent can lead to misuse or unfair advantage instead of supporting equity. The Repository itself could be seen as a data trust for the evolving legal and economic implications of AI for art and culture. This is why this last "theme" is introduced before the second part, which is an estimate of costs depending on technical design choices, such as limited access and multilingual capabilities. But first, allow us to provide a brief definition of a data trust and list the initiatives presented at the Co-design workshop.

A cultural data trust is a stewardship model where stakeholders collectively pool, govern, and share cultural data (e.g., heritage archives, creative works) through a trusted entity, ensuring ethical use, provenance tracking, and benefit-sharing while preventing exploitation by AI platforms or corporations. In simple terms, a data trust helps answer: who can use cultural data, under what conditions, with whose permission, and who benefits from that use. In Quebec's civil code context, exemplified by Culturepédia (Culture pour tous), it takes form as a *fiducie d'utilité sociale*, a purpose-bound, non-profit patrimony prioritizing public interest over commercial gain, but globally adapts to other legal regimes, from common law trusts in the UK to Indigenous governance models and the initiatives shared during

the co-design workshop, centering consent to data sharing, collective governance, and equitable benefit sharing.

- Cultural Data Trusts & Participatory AI Governance (Mika Noh, AI Culture Lab): "Current AI systems extract value from cultural datasets drawn from public domains, proprietary platforms, and Indigenous and community-based knowledge systems — often without equitable attribution, compensation, or consent. I propose developing and contributing a working model for Cultural Data Trusts that: Prevent enclosure of digital cultural commons; Establish collective licensing and benefit-sharing frameworks; Embed moral rights and provenance tracking into AI training pipelines; Align copyright, database rights, privacy law, and human rights frameworks."

- Insular Data Trusts: AI and the Cycladic Cultural Commons (Jacob Moe & Marina Markellou, Archipelago Network): "Over the next year, we will explicitly align these practices with data governance and privacy standards via data trust mechanisms [...] Create our own Archipelago-inclusive governance framework to assert sovereignty, provenance, and to ensure recognition of the collective dimension of our cultural data, to require fair access and equitable remuneration for the communities whose cultural data significantly contribute to new AI models and products."

- On protecting cultural expressions in the age of AI (Luke McDonagh, LSE): "A distinctive contribution of this proposal is to connect these legal debates with promising community-driven responses to AI's impact on the cultural commons. It will draw on emerging practices in cultural data commons, data trusts and collaboratives, and community-governed licensing schemes that embed social license, benefit-sharing and cultural sovereignty into the terms on which cultural data can be used to train AI systems. Particular attention will be given to models where institutions act as stewards, such as galleries, libraries, archives and museums that host shared datasets, negotiate with AI developers, and enforce community-defined conditions on access, attribution and reuse."

- Provenance Infrastructure as a Safeguard for Cultural Commons (Katherine Elkins & Jon Chun, Human-Centered AI Lab): "Rights-aware AI training data loader: A technical enforcement layer that automatically excludes copyrighted, culturally restricted, and takedown-flagged items from any downstream AI training. This operationalizes at the infrastructure level what data trust models aspire to at the governance level."

- Metadata as Memory (Giulia Taurino, Getty Research Institute) : "Libraries and community archives are piloting data trust and data collaborative models that place governance of digitized cultural materials in the hands of source communities. By establishing clear protocols about which materials may be used for AI training and under what conditions, heritage professionals translate abstract sovereignty claims into actionable stewardship strategies. This work is particularly important for marginalized and diaspora communities whose cultural records, often held in institutional archives far from their communities of origin, are vulnerable to appropriation even before AI ingestion."

- Empowering Khayamiya Artisans in Historic Cairo (Miral Victor Youssef, Cairo Governorate : "Use Case: A proposed 'Cultural Data Trust' model to protect Khayamiya's motifs from unauthorized AI training."

These proposals suggest that community grounded cultural data trusts could have a critical role in addressing the challenges that AI poses to cultural rights in creative (digital) economies. They move beyond access alone and focus on authority, consent, attribution, fair remuneration, and the right to refuse participation. For the design of the Repository, this means thinking beyond open access toward stewardship models that protect digital cultural sovereignty, support collective governance, and recognize that not all knowledge should be equally or universally available. It also requires understanding how cultural data trusts may be adapted to and adopted across different sociolegal environments, while remaining grounded in local authority, Indigenous governance models, and community-defined conditions of use.

Building a Repository: A Preliminary Cost Estimate

The objective of this estimate[28] is to support the design and phased deployment of a digital Repository on AI, IP and culture, co-designed with UNESCO stakeholders and international partners. The platform is intended for a consent-based community of experts, practitioners, artists, cultural workers, civil society organizations (CSOs), Indigenous knowledge holders, and decision-makers who need to contribute, consult, and manage structured content; explore that content through a conversational AI assistant; exchange in a secure and fully traceable environment; and perform efficient search across all platform materials. The project is structured in two phases: a short-term pilot, followed by broader deployment depending on funding and partnerships.

[28] The preliminary framing for a cost estimate was prepared by Étienne Pilon, CEO, Mhosaic.

The proposed platform combines five main functional components: structured content management for documents, images, and videos; an AI chatbot that answers only from platform content and cites sources; secure role-based member access with community-defined authority over contribution, retrieval, and stewardship of knowledge a peer exchange space with searchable, traceable discussions; and a transversal search engine with both text-based and semantic retrieval. Non-functional requirements include Canadian hosting, no U.S. data transit, encryption in transit and at rest, GDPR-aligned compliance for European partners, audit logs, scalability, accessibility, *multilingual capability, and ongoing maintenance. These requirements make the platform more sophisticated than a standard content portal and are a major reason the estimate spans a wide range. The Repository must be designed not only as a technical platform, but as relational infrastructure where access, participation, and governance reflect the authority and consent of those whose knowledge it holds.

Mhosaic recommends an agile delivery model over 12 to 24 months, with development phased from a basic functional prototype to chatbot and search integration, then peer exchange features, and finally consolidation and wider deployment. The proposed commercial model is a continuous monthly partnership, estimated at $10,000 to $20,000 per month, covering backend and frontend development, AI integration, infrastructure, project management, and support. Infrastructure costs are estimated at approximately $100 to $300 per month for the pilot phase and $300 to $900 per month for the deployment phase, excluding any premium for self-hosted AI. On that basis, the overall indicative budget is approximately $125,000 to $500,000 for the technical development components. This excludes the costs of the coordination for the governance and outreach aspects outlined above.

The main factors influencing cost are the structural decisions still to be made. The largest variable is the choice of language model: a self-hosted open model in Canada improves digital cultural sovereignty, data stewardship and alignment with UNESCO's ethical positioning, but adds infrastructure and implementation costs, including about $300 to $500 per month in GPU hosting.

Multilingualism is another major driver, because each additional language increases interface translation, testing, and chatbot validation costs. Cost also depends on whether the peer exchange space is custom-built or integrated from an existing tool; whether videos are simply stored or also streamed; whether the chatbot indexes only Repository documents or also indexes member discussions; and whether contribution workflows require review, moderation, and approval before publication. These choices affect architecture, development time, infrastructure, and long-term operating cost, which is why the estimate remains preliminary rather than fixed.

*Please note that implementing a multilingual information retrieval and chat system raises significant legal, social, cultural, and technical challenges not covered here. For instance, AI-assisted translation demands recognition of creative adaptation as authorship, plus robust consent for language as cultural data (see above discussion on ownership). These ethical and technical complexities make multilingual functionality a key design priority for future Repository phases.

Conclusion

The implication of the growing precarity of artists extends far beyond individual livelihoods, they are closely tied to broader risks of cultural erasure of communities who have been marginalized and to an increasing concentration of power over knowledge creation and dissemination. Artists and cultural workers are, in many ways, democracy's canaries in the coal mine: when their rights are weakened, cultural sovereignty is weakened for all. Protecting their rights is therefore essential to sustaining culture as a form of democratic infrastructure.

The Repository responds directly to that need. The co-design workshop enabled those who are most impacted to share their input towards developing a framework that reflects their realities, priorities and needs, ensuring that it remains grounded in lived experience. This work underscores the fundamental principle: governance must be produced with communities, not just for them. This is particularly critical in the context of generative AI, where global dependencies continue to shape the political economy of knowledge, often reinforcing existing inequities.

Culture itself must be understood as public infrastructure, encompassing art, traditional knowledge, and the systems through which information is created, translated and shared. In this context, emerging notions of informational, knowledge and cognitive sovereignty extend and reinforce the concept of cultural sovereignty. The role of a Repository as a living space that facilitates dialogue and knowledge sharing among its members, becomes central to supporting these forms of sovereignty.

In conclusion, the Repository is envisioned as a dynamic and facilitative space for knowledge sharing and sovereign decision making. This report serves as a reference point for what comes next. By offering a diverse set of evolving tools, perspectives and resources, it enables users to navigate complexity on their own terms. This supports democratic access to knowledge, accommodates differing viewpoints, and strengthens cultural agency, ensuring the platform remains an evergreen, inclusive hub for more equitable AI futures. Investing in this initiative means supporting an approach that is grounded, collaborative and a reflection of the reality being felt by the cultural sector globally and it is necessary for

it to continue to thrive, because when IP and cultural protections fail, cultural sovereignty is weakened, having consequences across democracies and economies alike.

**An Addendum and Executive Summary will be added soon.*

Please kindly contact Valentine Goddard via AI Impact Alliance's contact form or on LinkedIn if you would like to support the development of the Repository. You can also become a member of AI Impact Alliance to become part of our community.