SHIFT

MetamorphoSis of cultural Heritage
Into augmented hypermedia assets
For enhanced accessibiliTy
and inclusion

Funded by
the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

Document info

Document ID:	D4.2 - Distribution SHIFT Curation Repository for Cultural Assets with DRM capabilities
Version date:	31.03.2025
Total number of pages:	33
Abstract:	The aim of the deliverable is to report on the software design, architecture and implementation for tracking and tracing the provenance of cultural assets. The term provenance is used to refer to the identification of content ownership belonging to the cultural heritage institutions and to identify the derivative copyrights that are generated through digital content transformation using AI technologies. The information related to the ownership and other metadata is then encoded using international standards and the results are stored in an immutable datastore. A series of technologies have been identified that once integrated offers the flexibility and transparency to track the relevant information through the digital archives. In this deliverable, the list of such solutions that have been reported on include (i) minio (the physical file repository system); (ii) ISO/IEC 230001-23: Smart Contract ontology to encode the cultural asset ownership; (iii) Ethereum private network (to track the provenance and digital transformation of cultural assets. All of these components are integrated using a Python Django framework that implements the necessary application programming interface (APIs) to support the data storage, retrieval and identification of derivative copyrights for new content generated using GenAI technologies.
Keywords	Content ownership, copyright, digital transformation, smart contract, blockchain, GenAI

Authors

Name	**Organisation**	**Role**
Krishna Chandramouli	QMUL	Researcher
Ebroul Izquierdo	QMUL	Researcher
Tomas Piatrik	QMUL	Researcher

REVIEWERS

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

Name	Organisation	Role
Razvan Purcarea	SIMAVI	Reviewer
Ioana Crihana	ANBPR	Reviewer

VERSION HISTORY

Version	Description	Date
0.1	Table of contents distributed among partners	09/12/2024
0.2	First round of contribution	20/01/2025
0.3	Second round of input contribution	19/02/2025
0.4	Final revision and submission to the review	31/03/2025
0.5	Final review and submission	31/03/2025

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

EXECUTIVE SUMMARY

This project establishes an integrated digital pipeline designed to support the secure, transparent, and standards-compliant management of cultural assets, particularly original artworks and their derivatives produced through generative AI. The system brings together key technologies—namely generative AI, object storage, and private blockchain—to enable robust tracking of provenance, ownership rights, and transformation history. The process begins with the digital submission of artwork, where each image is hashed to create a unique identifier. These assets can then be transformed into short video sequences using generative AI models, enhancing public engagement while preserving a clear link to the original source.

All assets—including the original artworks, AI-generated videos, and related documents—are securely stored using MinIO, a private object storage platform compatible with S3 protocols. For every asset submitted, a machine-readable contract is automatically generated in accordance with ISO/IEC 21000-23 (Contract Expression Language). These contracts formally capture the parties involved, usage rights, licensing terms, and validity periods, providing a structured and interoperable way to manage intellectual property and derivative rights. Each contract is then stored alongside the associated asset in MinIO.

To ensure integrity and long-term traceability, metadata for each asset—including the contract's hash, creation details, and transformation lineage—is registered on a private Ethereum-based blockchain. This step guarantees that all asset interactions are auditable and tamper-resistant, fostering trust in the asset's digital provenance.

The need for software integration across these systems is paramount. It eliminates manual overhead in rights management, reduces risks associated with asset duplication or tampering, and ensures that the entire asset lifecycle—from creation to AI-based enhancement—is securely tracked and verifiable. By harmonizing digital asset processing with international standards such as ISO/IEC 21000-23, the system offers cultural institutions a future-proof solution for managing digital heritage, derivative rights, and AI-enhanced interpretation within a secure, transparent, and legally grounded framework.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

Content

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

Abbreviations and Acronyms

Abbreviation / Acronym	Description
AI	Artificial Intelligence
CEL	Contract Expression Language
CIDOC	International Committee for Documentation
CRM	Conceptual Reference Model
DID	Digital Item Declaration
DII	Digital Item Identification
DRM	Digital Rights Management
GenAI	Generative AI
geth	Go-Ethereum
IPFS	Interplenary File System
ISO	International Standards Organisation
JSON	Javascript Object Notation
MPEG	Moving Pictures Expert Group
MVCO	Media Value Chain Ontology
ODRL	Open Digital Rights Language
PROV-O	Provenance Ontology
W3C	World Wide Web Consortium
XML	eXtensible Markup Language

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

1. INTRODUCTION

The digitization of cultural heritage assets has become an essential strategy for preservation, accessibility, and scholarly engagement in the 21st century. However, as cultural institutions increasingly embrace digital workflows—including the use of generative artificial intelligence (GenAI) to reinterpret or animate historical content—new challenges have emerged regarding authenticity, provenance, and rights management. Traditional systems for recording the history and legal status of artworks are often siloed, manually intensive, and unsuited to the complexities introduced by derivative digital works. There is a growing need for a digitally integrated infrastructure that can securely manage original artworks and their transformations, while ensuring compliance with international standards and maintaining the integrity of historical provenance.

Generative AI has introduced novel forms of engagement with cultural assets, enabling dynamic visualizations and creative reimaginings that are compelling to contemporary audiences. While this innovation expands the cultural and educational value of heritage content, it also complicates the attribution, licensing, and ownership of the derived outputs. As argued by McCormack et al. (2019) [12], the creative capacity of AI systems challenges traditional notions of authorship, thereby demanding robust frameworks for intellectual property management. In the context of museums, galleries, and archives, this means that not only must the original asset be preserved and identified, but each generative derivative must also be traceable back to its source, with clear records of the rights granted, transformations applied, and responsible parties involved.

To address these challenges, this project proposes an integrated system that combines generative AI processing, object storage, metadata management, and blockchain technology into a single provenance-aware workflow. Central to the system is the adoption of ISO/IEC 21000-23:2018, a standard developed by MPEG for the expression of digital contracts and licensing terms (known as the Contract Expression Language or CEL). CEL allows for the formalization of rights, duties, parties, and validity periods in a machine-readable format, aligning well with the complexities of GenAI workflows and derivative rights. As explored by Vetro et al. (2003) [4] and Wang & Vetro (2011) [5], MPEG-21's broader framework offers a robust foundation for digital content management in distributed systems, particularly in contexts where multiple stakeholders are involved in content creation, transformation, and distribution.

In parallel, the use of a private blockchain infrastructure (such as Ethereum or Hyperledger) ensures immutability and verifiability of asset registration and rights documentation. Blockchain-based provenance tracking has been explored in various domains, including fine arts (Zhao et al., 2019) [17], cultural informatics

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

(Liu et al., 2020) [16], and digital identity (Hardjono et al., 2018) [15]. When used in conjunction with hashed metadata and decentralized storage (e.g., MinIO[1]), blockchain provides a tamper-proof ledger that not only secures the authenticity of digital artifacts but also enables transparent audits of their use and transformation history.

The proposed system automates the generation of rights contracts in CEL format for both original and AI-derived works, storing these alongside the digital files in MinIO and anchoring the associated metadata on a local blockchain network. This ensures that each asset, regardless of its origin or transformation state, is embedded within a comprehensive metadata framework that links content, authorship, permissions, and contractual obligations. Moreover, the integration of smart contracts enables programmable enforcement of rights—allowing institutions to define, for example, time-limited or non-commercial licenses for generated content.

In summary, the integration of GenAI, blockchain, and ISO/IEC 21000-23 represents a forward-looking approach to cultural heritage management. It addresses the urgent need for provenance transparency, legal clarity, and interoperable metadata in an era where digital transformation is redefining how cultural content is created, accessed, and interpreted. By leveraging these technologies in unison, cultural institutions can preserve the integrity of their collections while embracing the creative potential of artificial intelligence.

1.1. SCOPE AND OBJECTIVES

The deliverable is built upon the research outcomes of WP2 and WP3 aimed at strengthening and improving museum and cultural experience for the wider range of audiences. In this context, the use of digital tools to transform using computer vision tools and more broadly the use and impact of using GenAI has been considered. Therefore, to protect the rights of the cultural heritage institutions in maintaining the copyright of the paintings and other cultural objects, it is vital to ensure the appropriate information about the assets are being coded using digital technologies. The proposed implementation in the deliverable will be a part of the overall architecture design and in close consultation with the end-users the acceptability of the digital rights management solution will be evaluated.

[1] https://min.io/

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

1.2 Overview of ownership, copyright, derivative copyrights

The transformation of cultural content through digital technologies has brought renewed focus on the legal and ethical dimensions of ownership, copyright, and derivative works. As cultural institutions embrace digitization and explore generative artificial intelligence (GenAI) to reimagine and reinterpret heritage assets, questions surrounding intellectual property rights become increasingly complex. This literature review examines the key theoretical and practical frameworks that govern ownership and copyright, with a focus on the emerging discourse around derivative copyrights in the context of AI-generated content.

Ownership and Copyright Foundations

At its core, copyright law is intended to protect the original expression of ideas, granting creators exclusive rights to reproduce, distribute, and adapt their works. As articulated in the Berne Convention (1886), these protections arise automatically upon the creation of a work and do not require registration. Ownership, as discussed by Litman (1990) [9], is not only a legal designation but also a cultural construct that reflects broader societal norms about authorship, creativity, and labour. In the digital age, determining authorship can become ambiguous—especially in collaborative environments or where algorithmic processes play a significant role.

Within the domain of cultural heritage, ownership is further complicated by issues of custodianship versus authorship. Museums, archives, and libraries often hold physical or digital copies of works they did not create, prompting legal and ethical debates about who controls the reproduction and reinterpretation of such materials. As Hunter (2004) [9] notes, the digital encoding of cultural assets introduces tensions between preservation imperatives and legal constraints, especially when those assets enter the public domain or are subjected to new forms of digital reuse.

Derivative Copyrights and AI-generated Works

The concept of derivative works is well-established in copyright law, referring to new creations that are based on or incorporate pre-existing material. Under U.S. copyright law (Title 17, Section 101), a derivative work must involve substantial transformation of the original and still retain elements of protectable expression. Examples include adaptations, translations, remixes, and, increasingly, outputs from AI models trained on copyrighted data.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

Generative AI technologies, such as those based on deep learning, complicate traditional definitions of derivative authorship. In most jurisdictions, works created solely by machines without human intervention do not qualify for copyright protection (Gervais, 2019) [8]. However, when a human guides the input or curates the output of a GenAI system, the extent of their creative contribution becomes the basis for legal consideration. Elgammal et al. (2017) [7] argue that while AI can simulate creativity, it lacks intention and subjective agency—criteria that are often central to legal definitions of authorship.

This issue is especially relevant for cultural institutions using GenAI to animate historical paintings or create derivative interpretations of archival material. As McCormack et al. (2019) [3] observe, the distinction between the original artwork, the AI-generated output, and the human curatorial role must be carefully defined to assign rights appropriately. Failing to do so risks legal ambiguity over whether the derivative is a new work, a collaborative one, or a reproduction that infringes on the original creator's rights.

Frameworks for Managing Rights in Digital Derivatives

To address these complexities, researchers and standards bodies have proposed frameworks that accommodate digital and derivative rights. The ISO/IEC 21000-23 standard (Contract Expression Language) enables the formal representation of licenses, obligations, and permissions in a machine-readable format. This is particularly useful in contexts where GenAI-derived content must be governed by specific conditions, such as non-commercial use, attribution requirements, or expiration dates.

Further, the use of blockchain and smart contracts has been explored as a mechanism to anchor rights information immutably and enforce usage rules programmatically (Liu et al., 2020 [2]; Zhao et al., 2019 [17]). These technologies offer cultural institutions the ability to record the provenance of both original and derivative works, register licensing agreements, and enable auditability of use over time.

In the literature, ownership and copyright in the context of digital cultural assets and GenAI-derived works present significant legal and conceptual challenges. The literature underscores the need for clear authorship attribution, structured rights documentation, and technological solutions that respect both creative innovation and legal compliance. As digital transformation reshapes the boundaries of authorship and reuse, cultural institutions must adopt integrative frameworks—such as ISO/IEC 21000-23 and blockchain-based provenance systems—to ensure ethical, transparent, and legally sound practices in managing ownership and derivative copyrights.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

1.3 Semantic framework

The development of semantic frameworks for digital rights management (DRM) and provenance tracking has evolved in response to the growing complexity of digital content creation, distribution, and reuse. Beginning in the late 1990s and early 2000s, researchers and standards bodies recognized the need for structured, machine-readable metadata that could express ownership, licensing terms, and usage rights across distributed digital ecosystems. Early efforts included the Open Digital Rights Language (ODRL), which provided an XML-based model for specifying permissions and constraints. ODRL became a W3C recommendation in 2018 and has since been extended for applications ranging from media licensing to educational content.

In parallel, the World Wide Web Consortium (W3C) developed the PROV-O ontology as part of its Provenance Working Group. Published in 2013, PROV-O offers a standardized vocabulary for describing the origins and transformations of digital resources, enabling systems to track how content is derived, altered, or influenced by other entities. This model has been widely adopted in scientific data management, archival systems, and cultural heritage projects, where maintaining a verifiable chain of custody is essential.

In the cultural informatics domain, CIDOC CRM (Conceptual Reference Model) emerged as a comprehensive ontology for describing cultural heritage information. Developed by the International Committee for Documentation (CIDOC) of ICOM, this framework facilitates semantic interoperability between museums, libraries, and archives by providing a shared language for cataloguing events, actors, and objects. CIDOC CRM has been instrumental in bridging traditional cataloguing practices with digital provenance frameworks.

As multimedia content and AI-generated derivatives became more prevalent, the MPEG group introduced the ISO/IEC 21000 series, particularly Part 23 (Contract Expression Language), to address contractual rights and obligations using formal semantic structures. CEL extends traditional metadata models by incorporating detailed descriptions of parties, actions, conditions, and validity periods in digital contracts. This standard aligns with the need for automated rights enforcement and supports integration with blockchain and smart contract systems.

Together, these semantic frameworks represent a progressive layering of models that address different facets of digital provenance and rights management. From foundational ontologies like PROV-O and CIDOC CRM to rights-focused standards

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

like ODRL and CEL, the evolution reflects a trajectory toward greater automation, legal interoperability, and cross-domain reuse of digital assets.

1.4 Blockchain technology

The emergence of blockchain technology has introduced novel mechanisms for securing and verifying the provenance of digital content. As a decentralized, tamper-resistant ledger, blockchain enables the transparent recording of asset creation, modification, and ownership, making it a powerful tool for addressing long-standing challenges in digital content authenticity, rights enforcement, and auditability. Over the past decade, academic and industry literature has increasingly examined how blockchain can support the lifecycle management of digital assets, particularly in domains where provenance is crucial, such as cultural heritage, media licensing, and academic publishing.

Zyskind et al. (2015) [18] were among the early proponents of using blockchain to support decentralized privacy and ownership for personal data, proposing architectures that allowed users to control access and usage terms. Building on similar principles, researchers in the digital arts sector began to explore blockchain as a tool for recording the authorship and transfer history of artworks. Benet (2014) [14] introduced the InterPlanetary File System (IPFS) as a peer-to-peer protocol for content-addressable storage, which, when paired with blockchain, allows for scalable and verifiable asset tracking. This synergy has enabled artists and institutions to anchor asset metadata on-chain while hosting the content off-chain to optimize efficiency.

In the context of cultural heritage, Liu et al. (2020) [11] proposed a blockchain-based provenance model specifically for museum and archival content. Their model integrates immutable metadata records with timestamped transformation histories, ensuring that any derivative use of an artifact can be traced back to its original source. This model addresses the need for multi-institutional collaboration by supporting shared governance structures and interoperable metadata schemas.

Hardjono et al. (2018) [15] further explored the role of blockchain in establishing trustworthy digital identity frameworks, highlighting the need for provenance and consent verification in distributed systems. This has implications not only for user-generated content but also for institutional stewardship of digital artifacts. Blockchain’s potential in enforcing provenance and rights has also been demonstrated in the publishing industry, where platforms like ScienceMatters and

Funded by the European Union

This project has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement no 101060660.

ARTiFACTS leverage distributed ledgers to timestamp scholarly contributions and verify citation chains.

More recently, Zhao et al. (2019) [13] evaluated the application of blockchain in artwork trading platforms, illustrating how smart contracts can automate licensing, revenue sharing, and resale tracking. Their findings underscore blockchain's capacity to embed programmable rules into content usage agreements, thereby reducing disputes and enhancing transparency. This has direct relevance for generative AI applications, where the provenance and permitted usage of outputs can be dynamically governed by smart contracts anchored to the original data and transformation history.

Despite these advances, several limitations remain. Scalability and interoperability across chains and metadata standards continue to pose challenges, particularly in high-volume environments. Additionally, the legal recognition of blockchain records in cross-border contexts is still evolving. Nevertheless, the literature broadly supports the integration of blockchain as a complementary layer within broader digital provenance architectures, particularly when aligned with semantic standards like PROV-O or ISO/IEC 21000-23.

In summary, blockchain technology is rapidly gaining traction as a solution for tracking and tracing digital content provenance. Its core attributes—immutability, decentralization, and programmability—make it well-suited for applications requiring trust, transparency, and accountability in the management of digital assets. As the ecosystem matures, future research and implementation will likely focus on bridging blockchain systems with existing semantic frameworks and storage architectures to ensure scalable, standards-compliant, and ethically grounded provenance solutions.

1.5. STRUCTURE OF THE REPORT

The rest of the deliverble is structured as follows. In Section 2, an overview of the Media Value Chain Ontology (MVCO) is presented along with a brief description of the Contract Expression Language (CEL). In Section 3, the detailed implemetnation of setting up a private blockchain using geth is presented, followed by conclusion in Section 4.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

2. Semantic framework for tracking and tracing media value chain

The **Media Value Chain Ontology (MVCO)** is a semantic framework developed to represent the roles, actions, and relationships involved in the creation, distribution, and consumption of digital media. Initially introduced by the MPEG-21 standardization committee (ISO/IEC JTC 1/SC 29/WG 11), MVCO is designed to support the modeling of complex media value chains by formalizing how entities interact within the digital content ecosystem. It enables interoperable rights management and provenance tracking by explicitly encoding the responsibilities and activities of actors (e.g., authors, producers, distributors, users) and the transformations applied to digital content throughout its lifecycle.

At its core, MVCO provides a set of **ontological classes and relationships** that describe:

- **Entities** such as content items, media components, or digital objects.
- **Roles** such as creator, modifier, consumer, or rights holder.
- **Actions** like create, modify, annotate, consume, distribute, or derive.
- **Events** which bind together roles, actions, and digital content instances at specific points in time.

MVCO builds upon foundational ontologies and rights expression languages such as ODRL and integrates them with MPEG-21 components, notably Digital Item Declaration (DID) and Digital Item Identification (DII). It facilitates provenance tracking by enabling systems to record how a piece of content was created, by whom, under what terms, and how it has been altered or re-used. MVCO is particularly relevant in environments where digital content undergoes multiple transformations—such as in broadcast media, streaming services, or AI-generated content pipelines.

The utility of MVCO becomes even more powerful when extended through the **ISO/IEC 21000-24 standard**, which formalizes the **MPEG-21 Contract Expression Language (CEL)**. While MVCO focuses on the semantics of content interaction and provenance, CEL provides a machine-readable syntax to express **digital contracts** that govern those interactions. CEL defines contracts as structured sets of clauses involving parties (e.g., creators, licensees), actions (e.g., reproduction, adaptation), and constraints (e.g., time limits, geographical restrictions, usage context).

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

The integration of CEL with MVCO addresses a crucial gap in media value chain management: the linkage between **provenance semantics** and **enforceable legal agreements**. With CEL, contracts can be defined that specify:

- The **rights granted** (e.g., right to display or modify content)
- The **obligations of parties** (e.g., attribution or non-commercial use)
- The **validity conditions** (e.g., temporal or territorial constraints)
- The **trigger events** (e.g., user access, derivative creation) that activate or terminate rights

These contract expressions are represented in XML or JSON, enabling interoperability with rights management systems, content delivery platforms, and blockchain infrastructures. When combined with MVCO, CEL allows systems to model not only what has happened to content but also what is permitted to happen under specific legal and contextual constraints.

This integrated approach is especially valuable in the context of AI-generated content, where distinguishing between the original work and its transformations is essential for ethical and legal accountability. For instance, using MVCO, one could trace the derivation path of a video generated by an AI model from an original painting. CEL would then define the licensing terms under which that video may be shared, remixed, or commercialized, making the rights governance both transparent and machine actionable.

The semantic model for the **Contract Expression Language (CEL)** is built upon the need to formally capture and enforce contractual agreements in digital environments. CEL provides a conceptual and syntactic structure for expressing rights and obligations associated with digital content. This semantic model includes a well-defined vocabulary and schema for expressing who can do what, to which content, under what conditions, and for how long.

At the heart of the CEL semantic model are four key constructs:

1. **Party**: Represents an entity involved in the contract. Parties may be individuals, institutions, or organizations taking on roles such as creator, rights holder, distributor, or consumer.
2. **Action**: Defines the permissible operations on content, such as view, reproduce, distribute, adapt, or create derivatives. These actions correspond closely with intellectual property rights under legal frameworks.
3. **Asset**: Refers to the digital object that is subject to the contract. This can include original works (e.g., images, videos, documents) or their

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

derivatives, with each asset being uniquely identifiable and describable through metadata.

4. **Condition**: Specifies the circumstances under which actions may or may not be performed. Conditions can include temporal constraints (e.g., start and end dates), geographical limits, attribution requirements, or commercial use restrictions.

Contracts in CEL are structured hierarchically, allowing for complex rights scenarios such as sublicensing, conditional delegation of rights, and bundled rights packages. CEL expressions can be serialized in XML or JSON, making them interoperable with digital rights management systems, content distribution networks, and blockchain-based provenance records.

The semantic rigor of CEL supports integration with ontologies like MVCO, enabling a linked data approach to digital content management. For example, a CEL contract may reference MVCO-defined roles and actions, ensuring consistency between rights enforcement and provenance tracking. This tight semantic coupling facilitates automated reasoning about rights—for instance, enabling a system to determine whether a given user can lawfully generate a derivative work based on contract terms and recorded provenance.

CEL also supports **event-driven conditions**, which can be linked to real-world or system-triggered events. Examples include triggering rights expiration upon a specified date, activating a new license upon derivative publication, or revoking access after a certain number of downloads. These mechanisms make CEL especially useful in dynamic digital ecosystems such as streaming platforms, collaborative media production, or AI-generated content pipelines.

In conclusion, the semantic model for CEL provides a robust foundation for managing and automating digital rights. By clearly defining the roles, actions, assets, and conditions involved in digital contracts, and linking these constructs to broader ontologies like MVCO, CEL enables legally grounded, machine-actionable governance of digital content. This approach enhances transparency, reduces compliance risks, and supports scalable rights management in complex digital environments.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

3 Lifecycle management of cultural assets

The adoption of blockchain technology in digital rights management and provenance tracking has become increasingly relevant for ensuring the integrity, transparency, and auditability of metadata linked to content assets. Among the various blockchain platforms available, **Geth** (short for Go-Ethereum) stands out as a widely adopted, open-source implementation of the Ethereum protocol written in the Go programming language. Geth enables the deployment and operation of private Ethereum-based networks, making it suitable for institutional use cases where data privacy, control, and customization are required. This document provides a detailed three-page overview of how a Geth-based private blockchain network can be configured and used to store and manage metadata derived from the ISO/IEC 21000-24 Contract Expression Language (CEL) ontology.

Geth Architecture and Features

Geth supports the Ethereum Virtual Machine (EVM), allowing users to deploy smart contracts, manage accounts, and execute transactions in a secure and decentralized environment. In the context of a **private blockchain**, Geth provides:

- **Full node control** with isolated chain configurations (genesis block, consensus mechanism).
- **Smart contract deployment** for custom logic related to rights and content metadata.
- **Peer-to-peer networking** for trusted institutions to participate.
- **Mining capabilities** to validate transactions without economic incentives, suitable for permissioned networks.

Each node in the Geth network maintains a full copy of the blockchain ledger, ensuring distributed validation of asset metadata without depending on external or public services.

Genesis Configuration for a Private Network

To initiate a private Geth network, a custom genesis.json file is defined. This file configures the initial state of the blockchain, including the consensus protocol (e.g., Proof of Authority or Clique), pre-allocated accounts, and network ID.

Example snippet of a genesis.json file:

```
{
  "config": {
```


Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

```
    "chainId": 2025,
    "clique": {
      "period": 15,
      "epoch": 30000
    },
    "homesteadBlock": 0,
    "eip150Block": 0
  },
  "difficulty": "1",
  "gasLimit": "8000000",
  "alloc": {
    "0xYourInstitutionAddress": {
      "balance": "1000000000000000000000"
    }
  }
}
```

The chain can be initialized using the command:

```
geth --datadir ./mydata init genesis.json
```

Smart Contracts for CEL Metadata

To store and manage CEL-derived metadata, a smart contract is deployed to the network. This contract allows the registration of digital asset metadata, the association of parties, and retrieval of contract URIs or hash references.

Example smart contract:

```
pragma solidity ^0.8.0;


contract CELRegistry {
    struct CELMetadata {
```


Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

```
        string assetId;
        string metadataHash;
        string contractURI;
        string rightsHolder;
        uint timestamp;
    }

    mapping(string => CELMetadata) public registry;

    function registerAsset(
        string memory _assetId,
        string memory _metadataHash,
        string memory _contractURI,
        string memory _rightsHolder
    ) public {
        registry[_assetId] = CELMetadata(
            _assetId,
            _metadataHash,
            _contractURI,
            _rightsHolder,
            block.timestamp
        );
    }
}
```

Smart contracts like the above can be compiled and deployed using tools such as Remix, Truffle, or Hardhat.

Metadata Storage Workflow

The workflow for integrating CEL metadata with the Geth blockchain includes:

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

1. **Generate metadata**: A JSON or XML file conforming to CEL ontology is generated for each asset.
2. **Compute metadata hash**: The file is hashed using SHA-256 to ensure integrity.
3. **Upload contract file**: The metadata file is uploaded to MinIO or IPFS, generating a persistent URI.
4. **Register on-chain**: A smart contract function is invoked with asset ID, metadata hash, and storage URI.
5. **Verify integrity**: On-chain metadata can be used to verify the authenticity of files through hash comparisons.

This method ensures that the metadata and contractual terms expressed through CEL remain tamper-evident and verifiable.

Benefits and Use Cases

Using Geth for CEL metadata management brings several advantages:

- **Data Sovereignty**: Institutions can host their own blockchain network, controlling access and validation.
- **Auditability**: Transactions are immutably recorded, allowing for complete histories of metadata updates.
- **Interoperability**: Contract and metadata structures can be aligned with broader standards like PROV-O, MVCO, or ODRL.
- **Smart Contracts**: CEL metadata can be linked with programmable rights enforcement, such as usage counters or time-based access.

Key use cases include:

- Managing provenance and licensing for AI-generated media.
- Registering transformations and derivative rights of digital artworks.
- Institutional archiving of multimedia licensing contracts.

A private blockchain network powered by Geth offers a robust, customizable platform for managing and verifying digital rights metadata derived from ISO/IEC 21000-24 (CEL). When integrated with semantic models and decentralized storage, it enables cultural and creative institutions to enforce and document rights in a transparent, standardized, and legally secure manner. This setup supports

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

emerging needs around AI content governance, multi-party licensing, and long-term digital stewardship.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

4. CONCLUSIONS

In this deliverable, a detailed description of the digital rights management framework as developed in the SHIFT project is presented. The proposed framework brings together key technologeis on use to keep track of the digital content and digitally transormed content using GenAI. In this deliverable, the technical details of the software implementation integrating MinIO, geth and CEL metadata formats is described in detail. The proposed implemetnation will be integrated into the SHIFT platform for validation.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

6. REFERENCES

1. Hardjono, T., Lipton, A., & Pentland, A. (2018). Towards a digital infrastructure for identity. Blockchain in Healthcare Today, 1(1).
2. Liu, Y., Li, X., & Mei, Q. (2020). Blockchain-based provenance for cultural informatics. Journal on Computing and Cultural Heritage, 13(2), 1-21.
3. McCormack, J., Gifford, T., & Hutchings, P. (2019). Autonomy, Authenticity, Authorship and Intention in Computer Generated Art. Digital Creativity, 30(1), 52-68.
4. Vetro, A., Eleftheriadis, A., & Christopoulos, C. (2003). MPEG-21: Multimedia framework. IEEE Signal Processing Magazine, 20(2), 80-87.
5. Wang, X., & Vetro, A. (2011). A framework for supporting rights management of multimedia content in MPEG. Multimedia Tools and Applications, 55(3), 545–562.
6. Zhao, Z., Fan, W., Yan, J., & Wu, H. (2019). Applying Blockchain Technology to Artworks Trading. Procedia Computer Science, 147, 271-277.
7. Elgammal, A., Liu, B., Elhoseiny, M., & Mazzone, M. (2017). CAN: Creative Adversarial Networks, Generating "Art" by Learning About Styles and Deviating from Style Norms. arXiv preprint arXiv:1706.07068.
8. Gervais, D. (2019). The Machine As Author. Iowa Law Review, 105(5), 2053-2085.
9. Hunter, J. (2004). Rights Markup Extensions for the Protection and Management of Cultural Heritage Assets. Proceedings of the 4th International Symposium on Information and Communication Technologies in Cultural Heritage.
10. Litman, J. (1990). The Public Domain. Emory Law Journal, 39(4), 965–1023.
11. Liu, Y., Li, X., & Mei, Q. (2020). Blockchain-based provenance for cultural informatics. Journal on Computing and Cultural Heritage, 13(2), 1–21.
12. McCormack, J., Gifford, T., & Hutchings, P. (2019). Autonomy, Authenticity, Authorship and Intention in Computer Generated Art. Digital Creativity, 30(1), 52–68.
13. Zhao, Z., Fan, W., Yan, J., & Wu, H. (2019). Applying Blockchain Technology to Artworks Trading. Procedia Computer Science, 147, 271–277.
14. Benet, J. (2014). IPFS - Content Addressed, Versioned, P2P File System. arXiv preprint arXiv:1407.3561.
15. Hardjono, T., Lipton, A., & Pentland, A. (2018). Towards a digital infrastructure for identity. Blockchain in Healthcare Today, 1(1).
16. Liu, Y., Li, X., & Mei, Q. (2020). Blockchain-based provenance for cultural informatics. Journal on Computing and Cultural Heritage, 13(2), 1–21.
17. Zhao, Z., Fan, W., Yan, J., & Wu, H. (2019). Applying Blockchain Technology to Artworks Trading. Procedia Computer Science, 147, 271–277.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

18. Zyskind, G., Nathan, O., & Pentland, A. (2015). Decentralizing privacy: Using blockchain to protect personal data. IEEE Security and Privacy Workshops, 180–184.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

ANNEX – Software implementation

```
from minio import Minio
from web3 import Web3
import hashlib
import json
import os
from datetime import datetime
from flask import Flask, request, jsonify


# ---- CONFIGURATION ----
MINIO_ENDPOINT = "localhost:9000"
MINIO_ACCESS_KEY = "minioadmin"
MINIO_SECRET_KEY = "minioadmin"
BUCKET_NAME = "cultural-assets"


BLOCKCHAIN_RPC = "http://127.0.0.1:8545"  # Ethereum Private Network
CONTRACT_ADDRESS = "0xYourContractAddress"
CONTRACT_ABI = [...]  # Paste your smart contract ABI here


UPLOAD_DIR = "uploads"
CONTRACT_DIR = "contracts"


# ---- SETUP ----
os.makedirs(UPLOAD_DIR, exist_ok=True)
os.makedirs(CONTRACT_DIR, exist_ok=True)
```


Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

```
app = Flask(__name__)

# MinIO client
minio_client = Minio(
    MINIO_ENDPOINT,
    access_key=MINIO_ACCESS_KEY,
    secret_key=MINIO_SECRET_KEY,
    secure=False
)
if not minio_client.bucket_exists(BUCKET_NAME):
    minio_client.make_bucket(BUCKET_NAME)

# Web3 client
web3 = Web3(Web3.HTTPProvider(BLOCKCHAIN_RPC))
contract = web3.eth.contract(address=CONTRACT_ADDRESS, abi=CONTRACT_ABI)

# ---- HELPERS ----
def hash_file(filepath):
    with open(filepath, 'rb') as f:
        return hashlib.sha256(f.read()).hexdigest()

def upload_to_minio(filename):
    object_name = os.path.basename(filename)
    minio_client.fput_object(BUCKET_NAME, object_name, filename)
    return f"{MINIO_ENDPOINT}/{BUCKET_NAME}/{object_name}"

def store_metadata_on_chain(asset_id, metadata):
```


Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

```
    metadata_json = json.dumps(metadata, sort_keys=True)
    metadata_hash = hashlib.sha256(metadata_json.encode()).hexdigest()
    tx_hash = contract.functions.registerAsset(
        asset_id,
        metadata_hash,
        metadata['data_uri'],
        metadata['created_by']
    ).transact({'from': web3.eth.accounts[0]})
    return tx_hash.hex()

def generate_cel_contract(asset_id, creator, licensee, asset_uri):
    contract = {
        "contract_id": f"c-{asset_id[:8]}",
        "parties": [
            {"role": "creator", "name": creator},
            {"role": "licensee", "name": licensee}
        ],
        "asset": {
            "id": asset_id,
            "title": os.path.basename(asset_uri),
            "uri": asset_uri
        },
        "rights": [
            {
                "action": "generate_derivative",
                "target": "AI video",
                "conditions": {
```


Funded by
the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

```
                "requires_attribution": True,
                "non_commercial": True
            },
            "validity": {
                "start": datetime.utcnow().strftime("%Y-%m-%d"),
                "end": "2026-01-01"
            }
        }
    ]
}
path = os.path.join(CONTRACT_DIR, f"contract_{asset_id}.json")
with open(path, 'w') as f:
    json.dump(contract, f, indent=2)
return path

# ---- ROUTES ----
@app.route("/upload", methods=["POST"])
def upload_asset():
    file = request.files['file']
    creator = request.form['creator']
    licensee = request.form.get('licensee', 'Unknown')
    asset_type = request.form['type']  # original / derivative
    derived_from = request.form.get('derived_from', None)

    save_path = os.path.join(UPLOAD_DIR, file.filename)
    file.save(save_path)
```


Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

```
asset_hash = hash_file(save_path)
asset_id = asset_hash[:16]

minio_uri = upload_to_minio(save_path)

contract_path = generate_cel_contract(asset_id, creator, licensee, minio_uri)
contract_uri = upload_to_minio(contract_path)

metadata = {
    "asset_id": asset_id,
    "type": asset_type,
    "filename": file.filename,
    "hash": asset_hash,
    "created_by": creator,
    "created_on": datetime.utcnow().isoformat() + 'Z',
    "data_uri": minio_uri,
    "contract_uri": contract_uri
}
if derived_from:
    metadata["derived_from"] = derived_from

tx_hash = store_metadata_on_chain(asset_id, metadata)

return jsonify({
    "status": "success",
    "asset_id": asset_id,
    "minio_uri": minio_uri,
```


Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

```
        "contract_uri": contract_uri,
        "blockchain_tx": tx_hash
    })


# ---- MAIN ----
if __name__ == "__main__":
    app.run(port=5000, debug=True)
```


Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

ANNEX – Contract template output


```
{
  "contract_id": "c-8492a1f3",
  "parties": [
    { "role": "creator", "name": "User A" },
    { "role": "licensee", "name": "AI Research Org" }
  ],
  "asset": {
    "id": "1234-abcd",
    "title": "sunset.png",
    "uri": "http://localhost:9000/cultural-assets/sunset.png"
  },
  "rights": [
    {
      "action": "generate_derivative",
      "target": "AI video",
      "conditions": {
        "requires_attribution": true,
        "non_commercial": true
      },
      "validity": {
        "start": "2025-03-31",
        "end": "2026-01-01"
      }
    }
  ]
```


Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

}

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.

ANNEX – Metadata storage in Blockchain

```
contract ProvenanceRegistry {
    struct Asset {
        string assetId;
        string metadataHash;
        string dataURI;
        string createdBy;
        uint timestamp;
    }
    mapping(string => Asset) public assets;

    function registerAsset(
        string memory _assetId,
        string memory _metadataHash,
        string memory _dataURI,
        string memory _createdBy
    ) public {
        assets[_assetId] = Asset(_assetId, _metadataHash, _dataURI, _createdBy, block.timestamp);
    }
}
```


Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101060660.