



D3.2

Data sharing infrastructure, neural content description, rights management and monetisation v2

Project Title	XR mEdia eCOsystem
Contract No.	101070250
Instrument	Innovation Action
Thematic Priority	HORIZON-CL4-2021-HUMAN-01-06
Start of Project	1 September 2022
Duration	36 months

Deliverable title	Data sharing infrastructure, neural content description, rights management and monetisation v2
Deliverable number	D3.2
Deliverable version	V1.0
Previous version(s)	n/a
Contractual Date of delivery	30.06.2025
Actual Date of delivery	30.06.2025
Nature of deliverable	Report
Dissemination level	Public
Partner Responsible	FINC
Author(s)	Marcello Scipioni (FINC), Alessandro Annese (FINC), Werner Bailer (JRS), Maria Pegia (CERTH), Stamatis Samaras (CERTH), Sotiris Diplaris (CERTH), Antonis Karakottas (CERTH), Helmut Neuschmied (JRS), Cláudia Marinho (MOG) Antonio Calvo (i2CAT) Rubén Ramiro (Eviden), Maurizio Montagnuolo (RAI), Theodoros Chiou (IPR), Leander Stähler (KUL), Ralph Gasser (UNIBAS)
Reviewer(s)	Daniel Berjón (UPM), Francisco Morán (UPM), Nico Heise (DW)
EC Project Officer	Andreea Popescu

Abstract	This document describes the final version of the services for content analysis and search, rights management and monetisation. It also provides the legal analysis concerning compliance and copyright aspects, and describes the user interface of the marketplace.
Keywords	Repository, search, retrieval, content description, data spaces, rights management, licensing, smart contract blockchain, content analysis, derivative work



Copyright

© Copyright 2025 XReco Consortium

This document may not be copied, reproduced, or modified in whole or in part for any purpose without written permission from the XReco Consortium. In addition to such written permission to copy, reproduce, or modify this document in whole or part, an acknowledgement of the authors of the document and all applicable portions of the copyright notice must be clearly referenced.

All rights reserved.

Revision History

VERSION	DATE	MODIFIED BY	COMMENTS
V0.1	24/04/2025	Werner Bailer (JRS)	First table of contents
V0.2	11/06/2025	CERTH, JRS, RAI, i2CAT, KUL, IPR, MOG	Initial input on different services
V0.3	16/06/2025	CERTH, KUL, i2CAT, JRS	Further input and consolidation of sections
V0.4	24/06/2025	KUL, IPR, FINC, JRS	Version for internal review
V0.5	27/06/2025	UPM, JRS, FINC	Revision after internal review
V1.0	28/06/2025	KUL, JRS	Final Document



Glossary

ABBREVIATION	MEANING
A	Affinity matrix
$A_{\{i,j\}}$	Affinity probabilities
API	Application Programming Interface
CDVA	Compact Descriptors for Video Analysis
CJEU	Court of Justice of the European Union
CLIP	Contrastive Language-Image Pre-Training
CMCL	Cross-Modal Centre Loss
CDSM	(Directive on) Copyright in the Digital Single Market
CSQ	Central Similarity Quantization
DBMS	Database Management System
d_c	Hash code length
DGCNN	Dynamic Graphic Convolutional Neural Network
DOCH	Discrete Online Cross-modal Hashing
dB	Database
ELK	Elastic Stack
FCMH	Fast Cross-Modal Hashing
Fscore@k	F-score at k
FSOD	Few shot object detection
GDPR	General Data Protection Regulation
GPAI	General Purpose AI
GPU	Graphics Processing Unit
GSPH	Generalized Semantic Preserving Hashing
GUI	Graphical User Interface
HCOH	Hadamard Codebook based Online Hashing
HMD	Head-mounted Display
JSON	Javascript Object Notation
JWT	JSON Web Tokens
KDLFH	Kernel-based Discrete Latent Factor Hashing
l_2	Euclidean distance
LAH	Label-Attended Hashing
M	Set of all existing modalities
mAP	Mean Average Precision
MeshCNN	Mesh Convolutional Neural Network
MTFH	Matrix Tri-Factorization Hashing
MODT	Moving Object Detection and Tracking
NMR	Neural Media Repository
P	Semantic space
$p_{\{i,j\}}$	Semantic probabilities
PointNet	Point cloud Network
precision@k	Precision at k
RAI	Radiotelevisione Italiana



recall@k	Recall at k
S3	Simple Storage Service (Amazon Web Services)
SSAH	Self-Supervised Adversarial Hashing Network
SSL	Secure Sockets Layer
TDM	Text and data mining
UI	User Interface
ViT	Vision Transformer (NN architecture)
WP	Work Package



Table of Contents

Revision History	4
Glossary	5
Index of Figures	9
Index of Tables.....	11
1 Executive Summary	12
2 Introduction.....	15
3 Repository Infrastructure	16
3.1 Overview (JRS)	16
3.2 Neural Media Repository (UNIBAS)	16
3.2.1 Changes to vitivr-engine	17
3.2.2 Changes to the NMR.....	17
4 Content description services	19
4.1 vitivr analysis services (UNIBAS)	19
4.2 Weakly-supervised landmark classification (JRS)	19
4.2.1 Service API for landmark classification	19
4.2.2 Web application for landmark classification	24
4.3 2D similarity descriptors (JRS)	25
4.4 Few-shot object detection (JRS)	26
4.4.1 Approach	26
4.4.2 Evaluation	27
4.4.3 Web application for incremental training of novel classes	27
4.5 Shot boundary detection (JRS)	29
4.6 Cross-modal descriptors (CERTH).....	30
4.6.1 State-of-the-art Methods	30
4.6.2 Datasets	30
4.6.3 Experiments.....	31
4.7 2D object detection and tracking (i2CAT).....	34
4.7.1 Approach	35
4.7.2 Implementation as a service.....	35
5 User and workflow management	37
5.1 Authentication services (i2Cat)	37
5.2 Content basket management (i2Cat)	37
6 Search services	38



6.1	Local search backend (UNIBAS)	38
6.2	Metasearch service (i2Cat)	38
6.2.1	Microservice Based Architecture	38
6.2.2	Ranking Agent	39
6.2.3	Garbage collection	42
6.2.4	Outlook	43
7	Legal Requirements for Rights Management (KUL, IPR)	44
7.1	Data and content ingestion	44
7.1.1	AI Act	47
7.2	Search and retrieval	50
7.3	XR Services: 3D Reconstruction	50
7.3.1	The relation between 2D and 3D: Facts and data use	51
7.3.2	The relation between 2D and 3D assets: reproductions and adaptations	53
7.3.3	3D reconstruction as a reproduction/adaptation	53
7.3.4	Implications for Exceptions and Limitations to Copyright Law	59
7.3.5	Remaining points of legal (un-)certainty	60
7.4	XR Marketplace	61
8	Rights management and licensing tools (FINC, KUL, IPR)	62
8.1	XReco marketplace licensing approach	62
8.1.1	Background from D3.1	62
8.1.2	XReco CC+ licensing scheme	63
8.1.3	Analysis: Opportunities and Risks	71
8.1.4	Addendum: Upstream XReco licensing & monetization scheme for 2D inputs	77
8.2	Implemented workflow	81
8.3	Rights Management and Monetization Architecture	82
8.4	Monetization Manager	84
8.5	Marketplace User Interface and Workflow (MOG)	86
8.5.1	Publish to Marketplace	87
8.5.2	Marketplace	92
8.5.3	User Options	97
8.6	Data valuation (CERTH)	100
9	Conclusion (JRS)	106



Index of Figures

Figure 1: NMR backend architecture.....	17
Figure 2: Web application for landmark classification.	25
Figure 3: Web application for few-shot object detection.	29
Figure 4: Unimodal results in terms of MAP on XRECO.Buildings.Monuments and dataset.	31
Figure 5: Unimodal results in terms of MAP on ModelNet40-Extended and dataset.	32
Figure 6: Cross-modal results in terms of MAP on XRECO.Buildings.Monuments dataset for (a) Image to Mesh and (b) Mesh to Image.	32
Figure 7: Cross-modal results in terms of MAP on ModelNet40-Extended dataset for (a) Image to Mesh and (b) Mesh to Image.	33
Figure 8: Multimodal results in terms of MAP on (a) XRECO.Buildings.Monuments and (b) ModelNet40-Extended dataset.....	33
Figure 9: Detection and tracking with YOLOX and BoT-SORT in 2 frames separated by about 2 seconds in a highway setting.	35
Figure 10: General application architecture. The user interacts through the FastAPI to send jobs and get jobs status. The celery worker consumes jobs from the redis job queue and downloads the target data from a S3 bucket located in the target AWS instance.	36
Figure 11 Complete metasearch Architecture	39
Figure 12 Ranking agent and GC deployment	39
Figure 13: Relationship between 2D visual content, 3D reconstruction and 3D visual content.....	51
Figure 14: Examples of 2D input content	52
Figure 15: Examples stills of 3D reconstructions.....	52
Figure 16: Need for authorisations based on reproduction and adaptation rights.	59
Figure 17. Anatomy of a CC license (https://wikieducator.org/Creative_Commons_unplugged/Anatomy_of_a_CC_license).	64
Figure 18: Permissible adapter's license under Non-commercial CC Licenses over 3D assets.	69
Figure 19: XReco CC+ licensing scheme decision tree.....	71
Figure 20: Sequence Diagram showing the publishing workflow of an asset to the marketplace	82
Figure 21: Rights Management Architecture described in D3.1	83
Figure 22: Current Rights Management and Monetization architecture.....	83
Figure 23: Diagram of the Stripe "collect then transfer" payment model.	85
Figure 24: Representation of the XReco Purchase Receipt, generated each time an asset is purchased.	86
Figure 25: Orchestrator: Asset preview.....	87
Figure 26: Marketplace: Licensing wizard - initial state.	88
Figure 27: Marketplace: Licensing wizard (expertise question).....	88
Figure 28: Marketplace: Licensing wizard – CC Plus example (attribution question).....	89
Figure 29: Marketplace: Licensing wizard - CC Plus example (commercial use question).....	89
Figure 30: Marketplace: Licensing wizard - CC Plus example (set asset price).	90
Figure 31: Marketplace: Licensing wizard – CC Plus example (final step).....	90
Figure 32: Marketplace: Licensing wizard - Pricing model selection.....	91
Figure 33: Marketplace: Licensing wizard - License selection.....	91
Figure 34: Marketplace: Licensing wizard - Expertise path.....	91
Figure 35: Marketplace: Landing page.	92
Figure 36: Marketplace: Search filters.	93
Figure 37: Marketplace: Asset preview (Free asset).	94



Figure 38: Marketplace: Asset preview (Free+Paid asset).	95
Figure 39: Marketplace: Stripe checkout page.....	96
Figure 40: XReco Platform: User options.	97
Figure 41: XReco Platform: My Profile.	97
Figure 42: XReco Platform: My Earnings page.	98
Figure 43: XReco Platform: My Transactions page.....	99
Figure 44: XReco Platform: My Purchases page.....	99
Figure 45: XReco Platform: Published Assets page.	100
Figure 46: Method overview: During NeRF training, we evaluate a small validation set after each training iteration i , recording the PSNR associated with image I_i . When image I_i is revisited, we compute the change in PSNR by subtracting the previous value from the current one. These PSNR differences are then aggregated to estimate each image's $DVpsnr$	101
Figure 47: Contribution score correlations on the Brandenburg Gate dataset. The x -axis shows contributions scores from one training run, and the y -axis from a different run with a different seed. Each (x, y) point compares scores for the same image. (a) Aggregated PSNR difference, (b) Aggregated L1 difference, (c) Last epoch PSNR difference, (d) PSNR difference.....	103
Figure 48: Rendered images trained with DVpsnr selected training set and NeRF-w selected training set. (a-left) DVpsnr trained render of Brandenburg Gate - PSNR 19.51, (a-right) NeRF-w trained render of Brandenburg Gate – PSNR 17.59. (b-left) DVpsnr trained render of Sacre Coeur – PSNR 17.31, (b-right) NeRF-w trained render of Sacre Coeur – PSNR 17.03. (c-left) DVpsnr trained render of Trevi Fountain – PSNR 21.3, (c-right) NeRF-w trained render of Trevi Fountain – PSNR 20.7. (d-left) DVpsnr trained render of Taj Mahal – PSNR 24.38, (d-right) NeRF-w trained render of Taj Mahal – PSNR 22.9.	104
Figure 49: Left: Image selected to be part of the training set by NeRF-w according to its NIMA score, while DVpsnr calculates a low score. Right: Image disregarded from NeRF-w's training set, while DVpsnr calculates a high score.....	104



Index of Tables

Table 1: Integrated content analysis services.	18
Table 2: Evaluation result (Topk=3) of the De-Vit few shot object detector using the LVIS dataset. The check of the image patch overlap with annotation mask during training improves the result for Topk=3 (s: small, m: medium, l: large, r: rare, c: common, f: frequent).	27
Table 3: Evaluation result (Topk=10) of the De-Vit few shot object detector using the LVIS dataset. The check of the image patch overlap with annotation mask during training does not improve the result for Topk=10 (s: small, m: medium, l: large, r: rare, c: common, f: frequent).	27
Table 4: Comparison of all methods based on Precision at k (k = 10, 25, 50) for different number of epochs or code lengths on XRECO.Buildings.Monuments dataset.	34
Table 5: Comparison of all methods based on Precision at k (k = 10, 25, 50) for different number of epochs or code lengths on ModelNet40-Extended dataset.	34
Table 6: PSNR achieved for the same pipeline with the NeRF-w training set and with the training set selected by our data valuation approach.	103



1 Executive Summary

This document provides an overview of the results achieved in WP3 from M17 to M34. The delivery date of this deliverable marks the completion of the development work in the technical work packages, with the rest of the technical work being dedicated to completing the integration and making improvements based on feedback from tests.

This deliverable provides an update of the content presented in D3.1. For some components already described in D3.1, this document explains the extensions and improvements implemented since. This includes functional extensions as well as wrapping components into services for integration. In addition, this document also describes components that have been created in the second half of the project, either as originally planned, or added due to results from tests or additional requirements from the second round of demonstrators.

The **repository infrastructure** forms the backbone of the work around data sharing, search and licensing. This document describes the Neural Media Repository (NMR), which serves as the backend component for indexing and search. Upon ingest of content, it invokes the content description services also presented in this document. In order to enable not only search in the local content repository, but also metasearch across other content sources, connectors are provided. These connectors federate queries to other content sources. This enables querying these content sources with the metadata provided by the sources, though not the full range of descriptors extracted when content is ingested into the NMR. A number of connectors have been implemented, complementing the set of connectors for example with one for Europeana. There is also a connector to an NMR instance, in order to enable search across multiple instances.

The **content description services** include both readily available analysis services that have been integrated (such as those already available in the *vitivr* backend) as well as research on new analysis methods and descriptors. The work on cross-modal descriptors for querying between text, 2D and 3D content has advanced and provided as a service for integration. A new service for few-shot object detection has been added.

User and workflow management provides an update of the work on infrastructure for the authentication of users. These components support both the other WP3 services as well as the Orchestrator described in D4.2.

Search services describe the components for local search in the NMR of an XReco deployment. This includes the metasearch component, which enables federated search across a set of connectors, and the reranking of the obtained search results. All search functionalities have been integrated in a single user interface (described in D4.2), which enables local search and metasearch, as well as search in the content set offered on the marketplace.

The section on **legal requirements for rights management** offers an updated overview of general principles and relevant applicable EU copyright law *acquis* in the context of XReco objectives and operation. This deliverable advances the analysis of questions related with the qualification of an AI-generated output as copyrighted work or as derivative creation, and discusses the potential exceptions that may be applicable. The findings of legal research and analysis can be summarized as follows:

Opportunities

- Data and content Ingestion akin to *LAION* seems to be capable of benefitting from the research TDM exception.



- A question concerning the definition of ‘text and data mining’ has been referred to the Court of Justice of the European Union for a harmonised interpretation, which means that practical implementation of TDM exceptions will probably be clarified in the near future.
- Data and content ingestion for purposes of 3D reconstruction does not necessarily require the same quantity of content and data as the development of a general-purpose AI model.
- Downstream 3D reconstruction techniques do not necessarily display the level of generality, nor are they necessarily capable of performing a range of distinct tasks, akin to a general-purpose AI model. This possibly sets aside certain AI Act obligations.
- A proactive approach in light of the AI Act best practices could ‘flesh out’ compliance with the TDM exceptions and the practices that benefit from the exceptions.
- The legal analysis of D3.1 concerning facts and data use remains valid. XReco 3D outputs will not routinely contain recognizable elements of 2D inputs, insofar that they employ unprotectable mere facts and data embedded within 2D inputs and connected with the *represented* object.
- Where only facts and data are used, the EU copyright *acquis* generally does not apply, 2D inputs are not reproduced within 3D output and 3D output is not a 2D input derivative.
- Pending the interpretation of the Court of Justice of the European Union, some applications of 3D reconstruction techniques that are still using the protected expression of an underlying 2D input may benefit from the pastiche exception.
- CC licenses are standardised, trustworthy, public licenses with extended implementation in various jurisdictions, apply to 3D content and enable interested stakeholders to be engaged in sharing media content under standard and, to some extent, familiar terms.
- The use of an additional licensing scheme based on the CC+ protocol addresses the need for monetization and allows adoption of licensing solutions featuring meaningful balance between standardization and flexibility.
- There is the possibility of exploring a licensing scheme for 2D inputs monetization, reliant on algorithmic training licensing and a revenue sharing approach, based on relevance measured upon informational contribution (facts and data) of 2D inputs in the reconstruction of 3D scene.

Risks

- The conflation of ‘mining’, ‘scraping’ and ‘training’ under the same legal notion of text and data mining.
- In case of data and content ingestion for commercial purposes, rightholders’ opt-outs need to be identified and respected, whereas standards and best practices in that field are still under development.
- Where ‘generative AI’ tools are economically preferred to specific 3D reconstruction techniques, the applicability of the AI Act’s requirements may cover such practices.
- The possibility of reproduction of recognizable elements from 2D inputs’ protected expression within 3D outputs cannot be eliminated and 3D output qualification as derivative cannot be entirely excluded. This is dependent on a case-by-case assessment and, in the absence of harmonized right of adaptation, it will be subject to national law assessment. Especially concerning notions of ‘derivatives’ or adaptations, EU Member States may pursue quite different approaches. In those cases, in the absence of rights clearance over 2D inputs employed, copyright infringement risk at 3D output level remains.
- There is a general lack of clarity regarding the applicability of emerging interpretations of copyright concepts vis-à-vis specific practices such as 3D reconstruction.
- Alignment between CC license and the terms for uploading a 3D asset at the marketplace may be needed.



- Copyrightability of 3D assets cannot be assessed from the outset and in case that the marketed 3D asset is not copyrighted, CC licenses could not apply.
- In case that marketed 3D asset is a derivative work, rights clearance over 2D content is needed in order to avoid copyright infringement.
- The implementation of a licensing scheme for the monetization of 2D inputs may require prior opt-out from TDM exception, and is heavily dependent on the adopted business model at platform level.
- A 2D inputs monetization scheme requires reliable measurement of informational contribution and needs to align with the principle of appropriate and proportionate remuneration.

The section on **rights management and licensing tools** describes the legal work on the licensing framework for XReco, as well as the technical work on rights management and the microservices that have been developed for this purpose. This includes a rights management service providing the main entry point and orchestrating the other services, which are the Smart Legal Contracts (SLC) engine for creating, validating and executing smart contracts, the blockchain service provider (used for notarising the contracts) and the monetization manager. The marketplace user interface, that serves as the graphical component to guide the user through the licensing and publishing process, is also described. This section also presents the results on data valuation methods, that are able to assess the contribution of source data to a new asset has been performed.



2 Introduction

This document provides an overview of the results achieved in WP3, focusing on the period from M17 until the completion of the WP in M34. It describes improvements of components already delivered with D3.1, as well as new components developed since. These new components include both those already scheduled in the original project plan as well as those added based on requirements identified from initial tests of the platform and the second round of demonstrators.

Covering the diverse and multidisciplinary tasks in WP3, this document is structured as follows.

The **repository infrastructure** (Section 3) forms the backbone of the work around data sharing, search and licensing. As this infrastructure was already delivered with D3.1, this section focuses on the changes and updates of components.

The **content description services** (Section 4) include the underlying research on algorithms for extracting structured metadata and feature descriptions from the different types of content supported by the XReco platform. This section describes both updates and additional services.

User and workflow management (Section 5) describes the updates to the infrastructure for the authentication of users. These components support both the other WP3 services as well as the Orchestrator described in D4.2.

Search services (Section 6) describes the components for search in the local repository of an XReco deployment as well as for the search across all connected data sources. It focuses on updates of these components. The separate 3D search UI described in D3.1 has been discarded after this service has been fully integrated with the orchestrator.

Legal requirements for rights management (Section 7) provides an update of the analysis of the legal frameworks affecting XReco. Apart from updates related to emerging legislation (such as the AI Act), this section also deepens the analysis on the status of assets obtained from automatic 3D reconstruction processes.

Rights management and licensing tools (Section 8) describes the legal approach to licensing and the technical work on rights management based on it. In addition to an update of the implementation of smart legal contracts and monetization management this section describes the marketplace UI and the approach for data valuation.

Finally, a conclusion and summary is provided (Section 9).



3 Repository Infrastructure

3.1 Overview (JRS)

This section provides an update of the components establishing the backend of the repository infrastructure. The previous 1:1 relation between XReco platform and an instance of the Neural Media Repository (NMR) has been extended, so that multiple NMRs with different access permissions may be present in one instance of the platform. This is also one option for scaling the system beyond physical limitations of the system hosting the NMR instance.

The NMR provides the indexing and search backend functionality for the content managed by the XReco instance. Note that such an instance of the XReco platform may be a dedicated one for a large organisation (such as a broadcaster) as well as hosted one serving multiple users, with access rights to different subsets of the content. The NMR provides functionality for ingest, which invokes the content analysis services described in Section 4 for extracting descriptors to be indexed. This enables the local search functionality described in Section 6.1.

In addition to searching the NMR of the local XReco instance, the repository infrastructure provides connectors to enable federated search across other content sources. These connected content sources may be other instances of the XReco platform, APIs of archives, content feeds or content marketplaces, or public content repositories (e.g. Wikimedia). This section describes the approach taken to designing these connectors, and describes specific types of connectors implemented so far. These connectors enable the metasearch functionality described in Section 6.2.

3.2 Neural Media Repository (UNIBAS)

The XReco NMR backend is powered by the *vitivr-engine*^{1,2} – an open-source multi-modal multimedia retrieval system. *vitivr-engine* is part of the *vitivr* stack at, which is developed at UNIBAS and has been around for several years. It provides basic backend functionality for the two main workflows: *asset ingest* and *asset search*. This includes but is not limited to media decoding, segmentation, feature analysis & extraction and the orchestration of retrieval queries. For details, we refer to D3.1.

From an architectural perspective, the *vitivr-engine* is used as a dependency for the NMR backend project, which acts as a thin wrapper and provides the relevant functionality, such as:

- A RESTful API, which can be used by other components (e.g., *Orchestrator* and the *XR Marketplace*)
- Extensions to *vitivr-engine* to enable connection to XReco infrastructure (e.g., Amazon S3)
- Specialised feature modules usable by *vitivr-engine*, that provide interfaces to the *content descriptor services* provided by other XReco partners (e.g., NewsTagger or Cross-modal 3D search).

The changes to the NMR have taken place in two forms: Firstly, changes to the NMR itself. And secondly, changes to *vitivr-engine* to support certain requirements provided by the NMR. By employing this two-pronged approach, we make sure that relevant functionality can be made available to all *vitivr-engine* users.

¹ See <https://github.com/vitivr/vitivr-engine>

² Gasser, R., Arnold, R., Faber, F., & Schuldt, H., Waltenspül, R. Rossetto, L. (2024, January). A new Retrieval Engine for vitivr. International Conference on Multimedia Modeling.



3.2.1 Changes to vitrivr-engine

The most important change to vitrivr-engine in the context of this project is support for *PostgreSQL* with *pgVector*³ as a database backend. *pgVector* is an extension to the DBMS *PostgreSQL*, which allows for Nearest Neighbour Search (NNS), a functionality widely used in multimedia retrieval. For this project, *PostgreSQL* is the successor of Cottontail DB⁴, which is still supported but has been discontinued. The advantages of this change are self-evident: *PostgreSQL* is a well-maintained, well-established and prevalent DBMS that can easily handle a wide range of large database workloads. Furthermore, with *pgVector*, we have access to a scalable vector similarity search and all the features that *PostgreSQL* provide. The change is reflected in the architecture provided in Figure 1.

Other changes to vitrivr-engine involve bug fixes, optimisations in media handling to speed-up data processing and minimize use of compute resources, and the implementation of experimental feature that are currently not used in XReco (e.g., late-filtering for certain types of queries, which did not make it into the XR Marketplace).

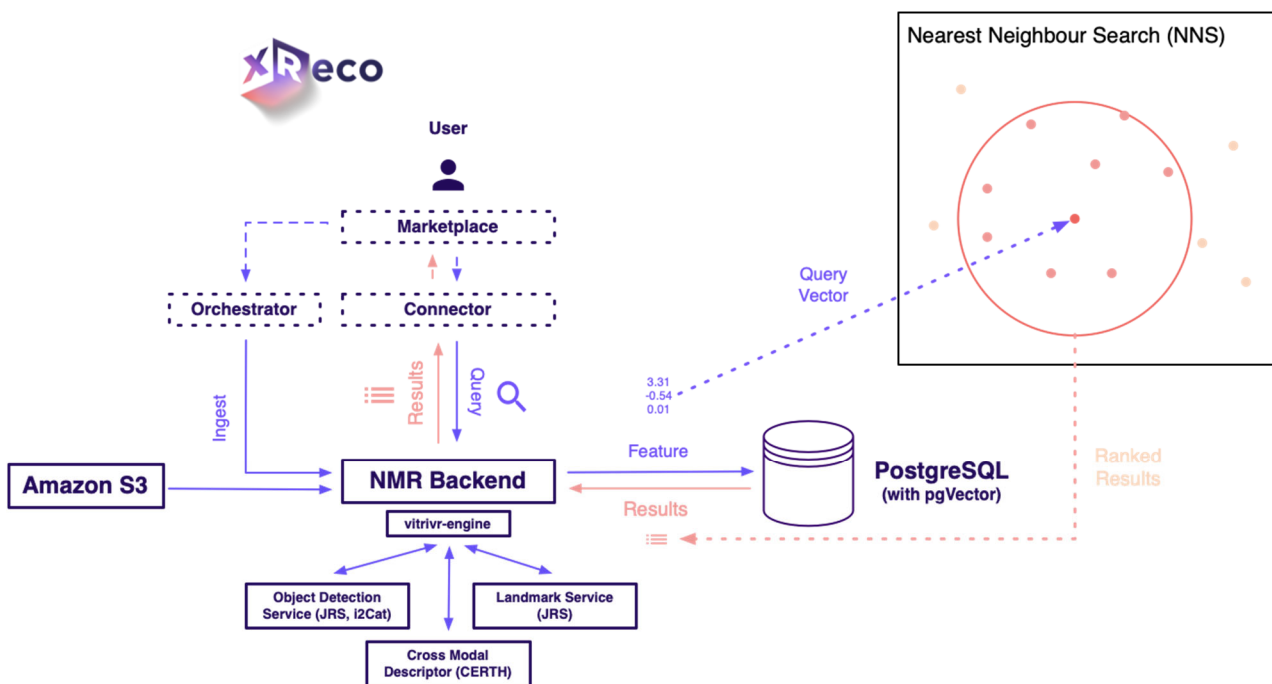


Figure 1: NMR backend architecture.

3.2.2 Changes to the NMR

The main changes to the NMR involve adjustments to the data model, data organisation and the APIs exposed by the backend. This process was continuously driven by the Orchestrator and XR Marketplace development.

Furthermore, we have integrated various external *Content Description Services* (see Section 4) provided by other consortium partners (see Figure 1 and Table 1). While the details vary from service to service, the general approach is the same for all the external content analysis services.

1. External services expose some sort of HTTP API.

³ See <https://github.com/pgvector/pgvector>

⁴ Gasser, R., Rossetto, L., Heller, S., & Schuldt, H. (2020, October). Cottontail DB: An Open Source Database System for Multimedia Retrieval and Analysis. In Proceedings of the 28th ACM International Conference on Multimedia (pp. 4465-4468).



2. Every such service is implemented as a dedicated **Analyser** with associated **Extractor** and **Retriever** in the NMR (part of *vitivr-engine* core facilities, we refer to the official documentation). The analysers define the *data fields* associated with the descriptor and classes that handle the analysis.
3. All the analysers are part of the *vitivr-engine* data schema (fields) and the data ingest pipeline.
4. The NMR takes care of storing incoming assets on Amazon S3, invoking the ingest pipeline and storing the descriptors that result from the ingest.
5. Individual services in the pipeline are called via an agreed-upon HTTP request – wherein the request points to the asset that requires analysis, and the response contains the derived feature descriptor.

In most cases, these four steps are only invoked during *data ingest*. As a result, the vectors are stored in PostgreSQL and available for subsequent queries. In some instances, the same invocations are also required at query time (e.g., for similarity search with an external baseline image).

The following *Content Description Services* have been integrated thus far:

Table 1: Integrated content analysis services.

Name	Partner	Description	Used For
CLIP	UNIBAS	Section 4.1	Similarity Search (More Like This)
Shot Boundary Detection	JRS	Section 4.5	Segments as hook for descriptors
Landmark Classification	JRS	Section 4.2	Fulltext Search
Few-shot Object Detection	JRS	Section 4.4	Fulltext Search / Tags
NewsTagger API	RAI	D4.2	Fulltext Search / Tags
Cross Modal Descriptors	CERTH	Section 4.6	Similarity Search
Object Detection and Tracking	i2CAT	Section 4.7	Fulltext Search / Tags

Another note-worthy change is the integration of the JRS Shot Boundary Detection service (see Section 4.5). This service determines the temporal segmentation of video material and replaces the fixed-length segmenter of *vitivr-engine*. Communication between that service and the NMR also takes place via a RESTful API. However, the resulting information does not lead to a feature descriptor but to a segmentation of videos.



4 Content description services

4.1 vitivr analysis services (UNIBAS)

We have described the analysis service provided by *vitivr-engine* (and thus the NMR) in Section 4.1 of D3.1. Currently, no other analysis services are provided by UNIBAS.

Nevertheless, the list of services supported by *vitivr-engine* has grown in the meanwhile. For example, we are now capable of interfacing with any type of inference model that can be hosted by TorchServe⁵. This is part of a concerted effort to make *vitivr-engine* more open to external services for media analysis and feature generation.

4.2 Weakly-supervised landmark classification (JRS)

We have described an approach for training landmark classifiers for visual content in Section 4.2 of D3.1. The motivation of the approach is the fact that archival content is often only weakly annotated, i.e. a landmark may be indicated in the description of a story or entire programme, but the specific temporal and spatial location where the landmark is depicted is not known. The proposed approach addresses this issue by allowing to train a classifier for such data.

In D3.1, the training approach was described. At that time, an inference service for applying the landmark classifier to content being ingested has been developed and deployed as a Docker container. Since then, also a training service has been developed and deployed, together with a web-based UI.

4.2.1 Service API for landmark classification

Once the Docker container for landmark classification is running, different service functions are accessible, which are used for:

- Configuration of the service
- Uploading images for inference and training
- Creating of inference and training jobs
- Querying of the job status

4.2.1.1 Service configuration

Images or videos can be uploaded from a MinIO or Amazon S3 storage service. The credentials for these services must be specified. This can be done by providing a configuration file or by environment variables (MINIO_USER, MINIO_PASSWORD, S3_USER, S3_PASSWORD) or by a web service function. The environment variable setting will override the configuration file setting at start of the service.

⁵ See <https://docs.pytorch.org/serve/>



Description of the configuration parameters:

Parameter	Description
VERSION	Version of the setting parameters.
SERVICE_PORT	Port of the landmark classification service.
MINIO_USER	Username for the MinIO storage
MINIO_PASSWORD	Password for the MinIO storage.
S3_USER	Username for the S3 storage.
S3_PASSWORD	Password for the S3 storage.
MAX_NUMB_WAITING_JOBS	Maximum number of video analysis jobs which can be queued at the server.
NUMB_OF_RESULTS	Number of landmark results per image ordered by confidence.

For defining the configuration parameters, a JSON file with the name *inf_server_config.env* has to be created. The initial parameters are loaded from this file when the Docker image is started.

Example of the configuration file:

```
{
  "VERSION": "1.2",
  "SERVICE_PORT": 8000,
  "MINIO_USER": "...",
  "MINIO_PASSWORD": "...",
  "S3_USER": "...",
  "S3_PASSWORD": "...",
  "MAX_NUMB_WAITING_JOBS": 10,
  "NUMB_OF_RESULTS": 10
}
```

The configuration settings can be queried using the service method "get_settings":

```
curl -X GET http://localhost:8000/get\_settings
```

Returned JSON data:

```
{
  "VERSION": "1.2",
  "SERVICE_PORT": 8000,
  "MAX_NUMB_WAITING_JOBS": 10,
  "NUMB_OF_RESULTS": 10
}
```

An update of these settings is possible with the POST request "change_setting". This sends the configuration settings as JSON data.

4.2.1.2 Check of Server Status

By using the 'status' endpoint, you receive information about the server's last response code, Docker image version, and queued video analysis jobs.

Example call:

```
curl -X GET http://localhost:8000/status
```

Result JSON data:

```
{
```



```

"code":200,
"version":"1.2.0",
"numb_waiting_jobs":0
}

```

4.2.1.3 Uploading images for inference or training

If several images have to be processed, they can be uploaded individually using the `send_image` endpoint. In addition to the upload path, the process type is specified in the `image_info` parameter with the value "extract" (images are used for inference) or "inc_training" (images are used for incremental training).

After the first image has been uploaded, a batch ID is obtained. This batch ID must then be specified in the `image_info` parameter for any additional images uploaded.

Example calls:

Uploading the first image:

```

curl -X POST http://localhost:8000/send_image /
-H 'accept: application/json' /
-H 'Content-Type: multipart/form-data' /
-F 'image=@/path/to/your/image' /
-F 'image_info={"process_type": "extract"}'

```

will cause an returned status information like:

```

{
  "status":"received",
  "batch_id": 1,
  "numb_files": 1
}

```

After the first image the next ones are uploaded by specifying the `batch_id`:

```

curl -X POST http://localhost:8000/send_image /
-H 'accept: application/json' /
-H 'Content-Type: multipart/form-data' /
-F 'image=@/path/to/your/image' /
-F 'image_info={"process_type": "extract", "batch_id": "1"}'

```

Uploading an image for incremental training (`batch_id` has only be specfied after the first uploaded image):

```

curl -X POST http://localhost:8000/send_image /
-H 'accept: application/json' /
-H 'Content-Type: multipart/form-data' /
-F 'image=@/path/to/your/image' /
-F 'image_info={"process_type": "inc_training", "batch_id": "2"}'

```

4.2.1.4 Analyzing an image

For analysing an image only the MinIO URL ("data") of the image has to be specified. The classification result is returned directly in JSON format. It contains an ordered list of recognized landmarks with their corresponding confidence values. The number of results per image is determined by the settings parameter `NUMB_OF_RESULTS`.

Example call:

```

curl -X POST http://localhost:8000/extract/landmark \
-H 'accept: application/json' \
-H 'Content-Type: application/json' \
-d @metadata_filename.json

```

The content of the file `metadata_filename.json`:

```

{

```



```
{
  "data": "http://192.168.20.100:9000/private/image_example.jpg",
  "model": "RAI"
}
```

The specification of the model is optional. The default model name is "RAI". If the MinIO server is on the same machine do not use "localhost" in the URL for the image.

Result JSON data:

```
{
  "Piemonte_Superga": 0.6049461364746094,
  "Trentino_Alto_Adige_Kurhaus_Merano": 0.0613950751721859,
  "Marche_Mole_Vanvitelliana": 0.03174290806055069,
  "Piemonte_Mole_Antonelliana": 0.02641391009092331,
  "Marche_Sferisterio_Macerata": 0.017989477142691612,
  "Veneto_Basilica_del_Santo": 0.01465498935431242,
  "Sicilia_Teatro_Politeama_Palermo": 0.013459655456244946,
  "Lazio_Palazzo_Madama": 0.011816633865237236,
  "Toscana_David_di_Michelangelo": 0.010778659954667091,
  "Lombardia_Castello_Sforzesco": 0.010563183575868607
}
```

4.2.1.5 Analyzing a video or an uploaded image batch

To create a video analysis job, the MinIO URL ("data") for the video, the segment "start" and "end" position, and a flag ("last") indicating whether this is the last analysis job for the video, must be sent to the server via a POST request to the "extract/landmark" endpoint. The JSON format of the attached data is shown in the example below. This adds the request to a job queue and returns a status message indicating whether or not the request can be handled. Requests are then processed one at a time. If the CALLBACK_COMPLETED_URL in the settings is specified, the result will be sent to that URL. Otherwise, the status of the job will be retrieved by the job_status call. If the job is completed, the status information returned will include the result of the analysis. The result will contain a list of classified landmarks with their corresponding confidence values for the corresponding video segment.

Example call:

```
curl -X 'POST' \
  'http://localhost:8000/extract/landmark' \
  -H 'accept: application/json' \
  -H 'Content-Type: application/json' \
  -d @metadata_filename.json
```

The content of the file metadata_filename.json:

```
{
  "data": "http://192.168.20.100:9000/private/Piemonte_Superga_4.wmv",
  "start": 0,
  "end": 0,
  "last": false,
  "model": "RAI"
}
```

The specification of the model is optional. . For the XReco MVP, a model named "RAI" is provided. If the MinIO server is on the same machine do not use "localhost" in the URL for the video. The status of the analysis process can be polled by the job_status endpoint. Once the status is "completed" also the analysis result will be returned.

If uploaded images are to be analysed, the JSON data must include the batch_id:

```
{
  "batch_id": 1,
  "model": "RAI"
}
```



```
}
```

Returned JSON data:

```
{
  "jobID": "1",
  "status": "ongoing"
}
```

4.2.1.6 Incremental training

The `inc_training` endpoint can be used to incrementally train additional landmarks into the landmark recognition model. In order to do this, a ZIP file containing subdirectories with a set of training images for each landmark has to be provided. Additionally a JSON list of the landmark name and the corresponding subdirectory name in the ZIP file has to be specified. The model name is required for future functionality

Example call for incremental training:

In the following example call the ZIP file contains a directory "stephansdom_images" with the images of the landmark "Stephansdom":

```
curl -X POST http://localhost:8000/inc_training \
  -H 'Content-Type: multipart/form-data' \
  -F 'model="RAI"' \
  -F 'zip_training_images=@D:\project\XRECO\Landmark_Detection\landmark_service_data\stephansdom.zip' \
  -F 'new_class_info={"Stephansdom":"stephansdom_images"}'
```

This call returns the job ID and the status "ongoing" if all parameters has been correctly specified. The status of the training can be polled by `job_status` endpoint.

If the images have already be uploaded then no zip file has to be specified:

```
curl -X POST http://localhost:8000/inc_training \
  -H 'Content-Type: multipart/form-data' \
  -F 'model="RAI"' \
  -F 'new_class_info={"Stephansdom":"stephansdom_images"}'
```

4.2.1.7 Model reset

The network model is changed after each incremental training step. The `reset_model` endpoint can be used to reset the network to the originally trained model. The model name is required for future functionality.

Example call:

```
curl -X POST http://localhost:8000/reset_model/RAI
```

4.2.1.8 Polling of the job status

The status of the video analyses, the incremental training, and the model reset jobs have to be polled. For that the endpoint `job_status` has to be called. A status information will be received, which could have the following values:

- ongoing
- completed
- cancelled



In case of an analysis job if the status of the job is "completed" then additionally the analysis result is returned.

Example call for polling of the job with the ID 0:

```
curl -X POST http://localhost:8000/job_status/0
```

Returned JSON data:

If only one image has been analysed, then the following result will be returned:

```
{
  "status": "completed",
  "result": [
    {
      "label": "Piemonte_Superga",
      "confidence": 0.7979607582092285
    },
    {
      "label": "Marche_Palazzo_Ducale_di_Urbino",
      "confidence": 0.04206463694572449
    },
    ....
  ]
}
```

If an image batch has been analysed then a list of results (for each image) will be returned:

```
{
  "status": "completed",
  "result_list": [
    {
      "image_file": "image_1.jpg",
      "feature": [
        {
          "label": "Lazio_Colosseo",
          "confidence": 0.013099242001771927
        },
        {
          "label": "Toscana_Torre_di_Pisa",
          "confidence": 0.006271745543926954
        },
        ...
      ]
    },
    {
      "image_file": "image_2.jpg",
      "feature": [
        {
          "label": "Piemonte_Superga",
          "confidence": 0.7979607582092285
        },
        {
          "label": "Marche_Palazzo_Ducale_di_Urbino",
          "confidence": 0.04206463694572449
        },
        ...
      ]
    },
    ...
  ]
}
```

4.2.2 Web application for landmark classification

The web application uses the landmark classification inference and training service is based on an approach for incremental training using Swin+API-Net, a fine-grained image classification network. The usage scenario involves acquiring training data for a landmark, which can be done using integrated web search engines like DuckDuckGo or a metasearch service from the XReco project. Users can filter these images, semi-



automatically or by manually rejecting irrelevant ones. Users select from two pretrained base models, one based on a RAI dataset (163 landmarks) and another on the Weakly Annotated Video Landmark (WAVL) dataset (141 landmarks). Various settings, such as "Inc Train All" (using pretrained class examples as None class), "Bias Correction" (adding a bias correction layer after the classification layer), and "Balanced Inc Train" (using data augmentation to create a balanced training dataset for novel classes), can be configured before starting the training. Once training is finished, the new landmark appears in the list of recognized landmarks.

To test the trained model, users can upload and analyze a video. The application analyzes video frames, tracks and combines recognition scores for identified landmarks. Frames where a landmark is recognized with high scores are displayed along with their timestamps.

The application is implemented as a client-server application in Python using the Gradio package (see Figure 2). It is highlighted as providing a ready-to-use application suitable for domain experts. The interactive nature of the application allows users to experiment with different parameter settings to refine model performance.

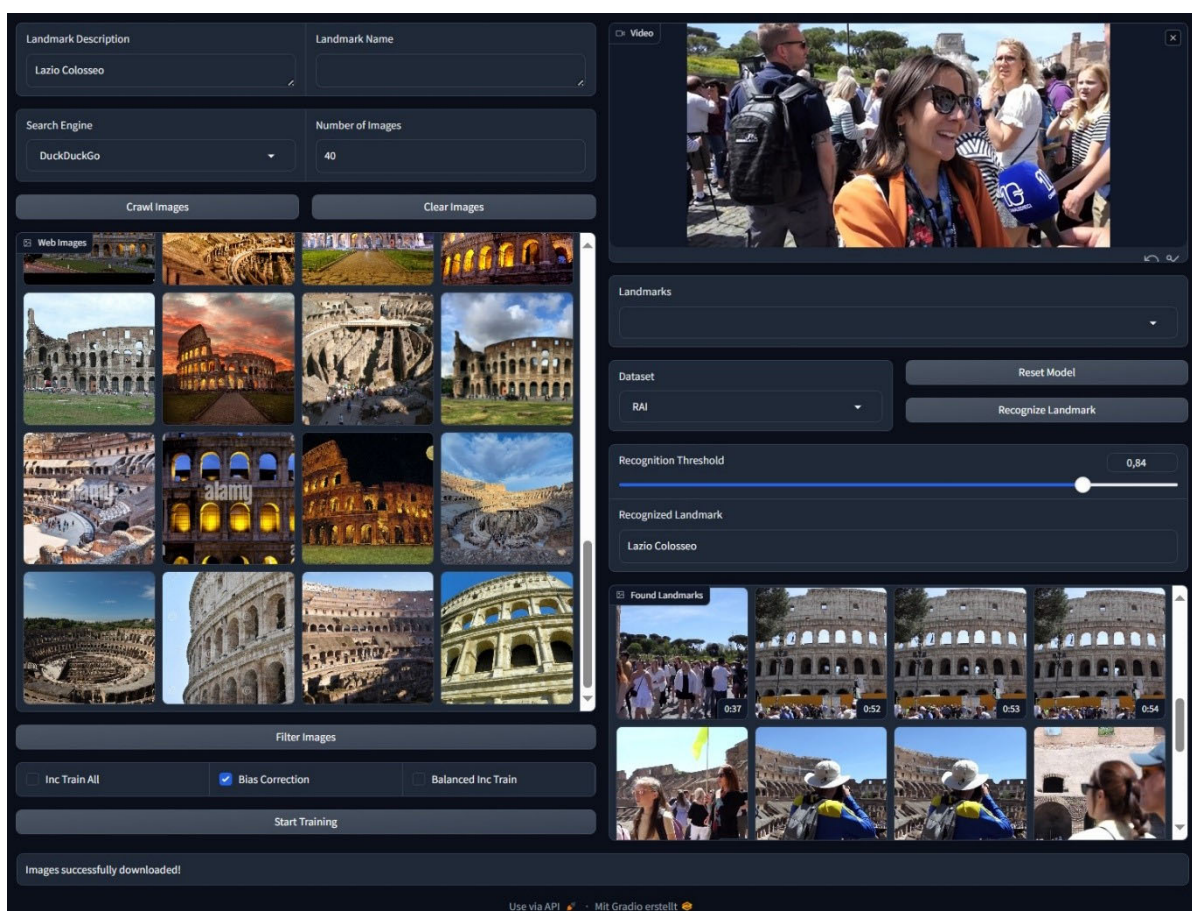


Figure 2: Web application for landmark classification.

4.3 2D similarity descriptors (JRS)

A service for extracting compact descriptors of images or video segments (based on the MPEG CDVA standard) has been implemented and deployed as a Docker container as described in Section 4.3 of D3.1. The descriptors are binarised and can be efficiently matched using Hamming distance. No further updates have been made since the version described in D3.1.



4.4 Few-shot object detection (JRS)

4.4.1 Approach

Few-shot object detection (FSOD) combines few-shot learning and object detection techniques to quickly adapt to novel object classes with limited annotated samples. Given an existing object detector that has been trained on abundant data for some categories, termed *base categories*, we wish to learn to detect *novel categories* using only a few annotations. If we want also to maintain the performance on the original base categories, we talk about generalized few-shot object detection (G-FSOD).

The model for base categories is naturally trained against detecting instances from novel classes, as they are unlabelled and treated as background. A further challenge is that the FSOD problem involves learning from extremely unbalanced data, so the number of training samples for base categories can be much larger than that for novel categories. A model that overfits a small number of novel instances naturally lacks the ability to generalize.

The generation of robust visual features by Visual Transformer Networks (ViT), which are trained in self-supervised manner on large image datasets, enables advances in FSOD methods. Kaul et al.⁶ use such a ViT network to generate pseudo-labels for new classes in the base training images. This increases the training data for new classes.

We select the recently developed method DE-ViT from Zhang⁷ and Fu⁸, which use DINOv2 Oquab⁹ as the backbone in a two-stage object detection framework similar to Mask R-CNN. Their approach outperforms other state-of-the-art methods especially for low-shot counts.

An off-the-shelf region proposal network (RPN) is used to generate the initial region proposals, as class-agnostic proposals. After that a region propagation network is applied that gradually propagates the proposal region to accurately cover and fit the object by refining an object mask. Object classification is performed using the visual features of the DINOv2 network derived from the object mask areas.

Networks trained on basic classes tend to overfit patterns specific to those classes. A common technique to mitigate overfitting is to represent data in a low-rank subspace Schittenkopf¹⁰. This approach explores the construction of a subspace of pre-trained Vision Transformer (ViT) features to reduce the accuracy gap between base and novel classes. Prototypes, which are class representatives built from support images, are computed by averaging the ViT features clipped with object mask areas. A prototype for a class is defined by an array of 10x1024 float values.

There is no need to retrain and fine-tune the backbone. Therefore, the training process is very fast. Different model sizes for DINOv2 can be chosen. By using the large model, the object detector has a memory demand of about 16GB GPU memory.

⁶ Kaul, P., Xie, W., Zisserman, A., Label, Verify, Correct: A Simple Few-Shot Object Detection Method, IEEE Conference on Computer Vision and Pattern Recognition, (2022)

⁷ Zhang, X., Wang, Y., Boularias, A.: Detect Every Thing with Few Examples, arXiv:2309.12969, <https://arxiv.org/abs/2309.12969>, (2024)

⁸ Fu, Y., Wang, Y., Pan, Y., Huai, L., Qiu, X., Shangguan, Z., Liu, T., Fu, Y., Gool, L., Jiang, X.: Cross-Domain Few-Shot Object Detection via Enhanced Open-Set Object Detector. ArXiv, abs/2402.03094, (2024)

⁹ Oquab, M., Darcet, T., Moutakanni, T., Szafraniec, H. Vo, M., Khalidov, V., Fernandez, P., Haziza, D., Massa, F., El-Nouby, A. et al.: Dinov2: Learning robust visual features without supervision, arXiv preprint arXiv:2304.07193 (2023)

¹⁰ C. Schittenkopf, G. Deco, and W. Brauer.: Two strategies to avoid overfitting in feedforward networks, *Neural Networks*, 10(3):505–516, ISSN 0893-6080. doi:[https://doi.org/10.1016/S0893-6080\(96\)00086-X](https://doi.org/10.1016/S0893-6080(96)00086-X), URL <https://www.sciencedirect.com/science/article/pii/S089360809600086X>, (1997)



A client-server architecture is used to allow outsourcing of inference and training processes. A web service allows direct processing of single images. Image sequences and videos can be processed by defining analysis jobs. Once an analysis job has been created, the results can be retrieved via a job status query.

4.4.2 Evaluation

We observed that the detection results were lower when the training images mainly contained narrow areas of the object. When we analyzed this behavior, we found that patch image regions were also being taken into account during training, even if they only had a small area of overlap with the object. To reduce the influence of the background during training, we ensured that only image patches covering more than 50% of the object were used (patch overlap check). To analyze the impact of these changes on a larger dataset, we performed an evaluation with the LVIS dataset¹¹, which contains 1203 object classes (see Table 2 and Table 3).

The Topk parameter must be specified for the evaluation. In the DE-ViT object detector, the Topk setting specifies the number of top-scoring region proposals or final detections to be kept at a given stage in the detection pipeline. This helps to balance accuracy with computational efficiency. A high Topk value is often used in papers for performance comparisons. In practice, however, a lower value is usually preferred, as this significantly reduces calculation time. We performed the evaluation for Topk=10 and Topk=3 for the LVIS dataset. The De-ViT object detector takes approximately 0.7 seconds to process an image for TopK=3 and around 1.6 seconds for TopK=10 (on a NVIDIA GeForce RTX 3090 GPU). The patch overlap check only improves the results for Topk=3. For Topk=10, the patch overlap check has a significant impact on rare (fewer than 10 training examples) object classes. This may be because the patch overlap check rejects too many training examples for these object classes.

Table 2: Evaluation result (Topk=3) of the De-ViT few shot object detector using the LVIS dataset. The check of the image patch overlap with annotation mask during training improves the result for Topk=3 (s: small, m: medium, l: large, r: rare, c: common, f: frequent).

Method	AP	AP50	AP75	APs	APm	APl	APr	APc	APf
Original	28,12	43,10	29,35	18,40	35,46	42,20	28,11	27,81	28,46
Patch overlap check	29,11	44,45	30,37	18,42	37,05	44,85	29,12	29,29	28,91

Table 3: Evaluation result (Topk=10) of the De-ViT few shot object detector using the LVIS dataset. The check of the image patch overlap with annotation mask during training does not improve the result for Topk=10 (s: small, m: medium, l: large, r: rare, c: common, f: frequent).

Method	AP	AP50	AP75	APs	APm	APl	APr	APc	APf
Original	30,94	47,13	32,47	20,21	39,46	45,57	33,81	30,08	30,65
Patch overlap check	30,48	46,52	31,96	19,65	39,19	47,08	30,40	30,68	30,29

4.4.3 Web application for incremental training of novel classes

A GUI for incremental training of new object classes was provided by a web application (see Figure 3). To obtain training images for a new object class, the web application offers the possibility to search for corresponding images on the Internet. The search term for the object class and the object name can be

¹¹ Gupta, A., Dollar, P., Girshick, R.: Lvis: A dataset for large vocabulary instance segmentation, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 5356–5364, (2019).



different. Additional images can be found by using different or extended search terms. All images found by the web search are displayed. To use an image for training, you must select one. The image is then shown in the annotation window. The image regions of the object can then be specified by drawing bounding boxes. Based on this data, a segmentation method based on visual prompting¹² can be applied to the image to determine the exact object boundaries. The “Use Annotation” button adds the image and the segmentation mask image to the training data. Once enough training data has been collected, training can begin and takes approximately two seconds per training sample (with a NVIDIA GeForce RTX 3090 GPU).

For testing, an image, an image sequence, or a video can be uploaded. After starting the detection process using the "Start Object Detection" button, the video will be analysed at a rate of one frame per second. Once the analysis process is complete, all analysed frames containing any of the objects listed in the "Object selection" text box will be displayed. New objects can be added to this text input field by typing or by selecting an object name from the "Object" drop-down list.

Both the training data with the computed object feature vectors and the object detection results can be analysed using the integrated data analysis tool FiftyOne. FiftyOne provides an interactive interface to explore image datasets and associated annotations or model predictions. You can easily view images alongside bounding boxes, segmentation masks, classifications, and other metadata. This allows for quick visual inspection of your data. This helps in identifying problematic training samples or interesting subsets for further analysis.

¹² Pan, T., Tang, L., Wang, X., Shan, S.: Tokenize anything via prompting, in Computer Vision – ECCV 2024: 18th European Conference, Milan, Italy, September 29–October 4, 2024, Proceedings, Part XLVII, ISBN 978-3-031-72969-0, https://doi.org/10.1007/978-3-031-72970-6_19, (2024)



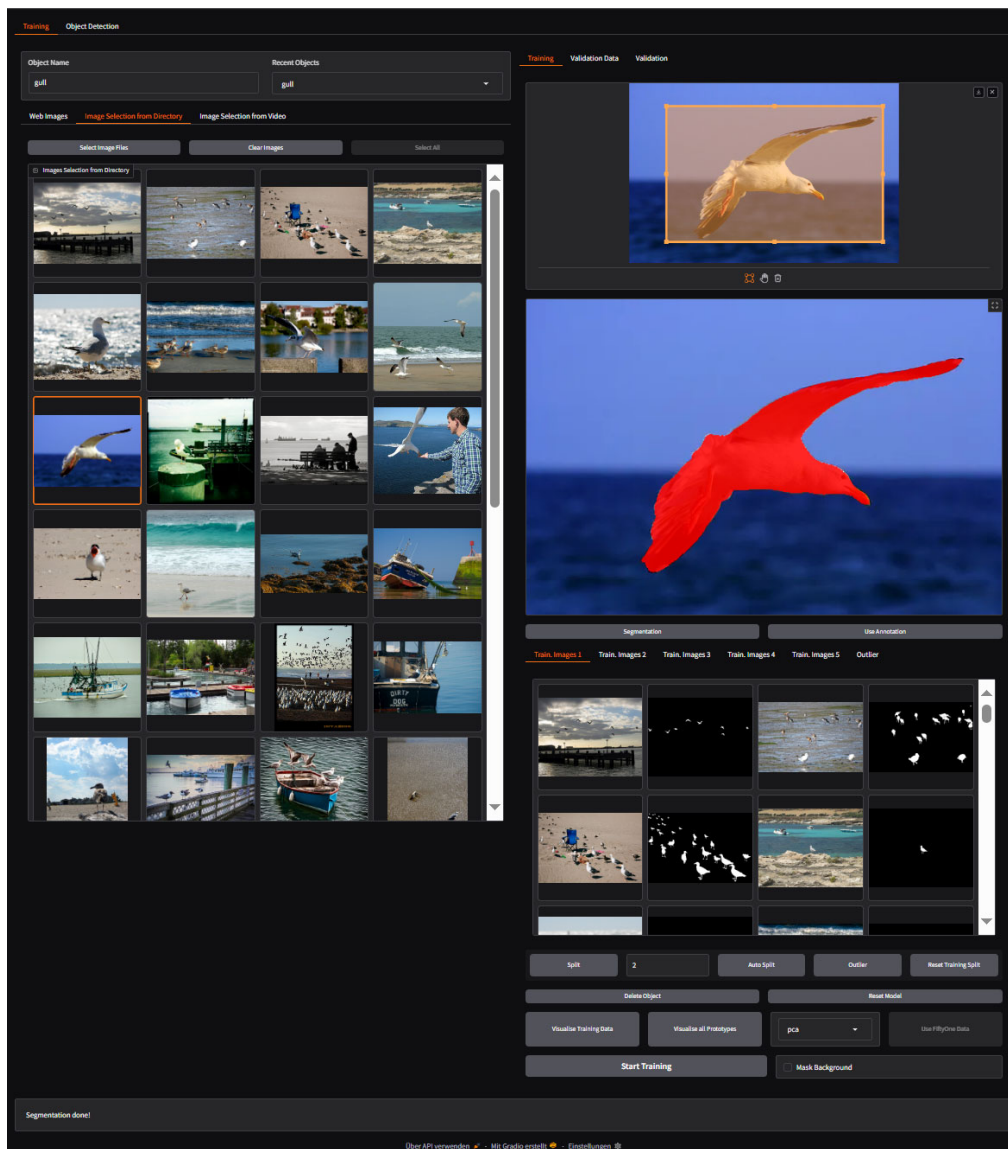


Figure 3: Web application for few-shot object detection.

4.5 Shot boundary detection (JRS)

The shot boundary detection service was not originally planned, but replaces the temporal segmentation in the vitivr analysis service. Based on the requirement to make temporal segments found in the search results an object that can be fed into other services or published on the marketplace, it was decided that a better temporal segmentation as needed.

The shot boundary detection algorithm is work done outside XReco¹³ (in the AI4Media project). It is a two-stage approach, performing a fast candidate check between consecutive frames, and a deep check in a four-frame time window around candidates. For robustness, the detector uses low-resolution optical flow estimation, which can be done efficiently on CPU. The algorithm also extracts keyframes based on visual activity in the content.

¹³ Fassold, H. Faster than real-time detection of shot boundaries, sampling structure and dynamic keyframes in video. Proceedings of the International Conference on Imaging, Signal Processing and Communication (ICISPC), 19-21 July 2024, Fukoka.



The algorithm is included in as a preprocessing step in the video to text service (described in D4.2). Due to the need for a standalone shot boundary detection service, a REST API similar to that of the video to text service has been provided. The service is deployed as a Docker container.

4.6 Cross-modal descriptors (CERTH)

The cross-modal descriptor framework was initially introduced in section 4.4 of deliverable D3.1. The main goal of this task is to support cross-modal search—meaning, searching using one type of data and getting results from another. In this context, a modality refers to different ways of representing an input. In this work, we focus on three types: images (snapshots of 3D models from specific viewpoints), meshes (the full 3D structure using points, edges and triangles), and point clouds (set of points that capture the shape of 3D model).

In the next subsections, we give a brief overview of current methods and our approach, MuseHash, for 3D retrieval. The MuseHash architecture and its technical details were described in detail in subsection 4.4.2 of D3.1. We also mention some of its pros and cons. Then, we describe two datasets used in the XRECO project for the training and testing of MuseHash for the second half of the project. The first one is used as it is, while the second is based on a known dataset, ModelNet, but with two new classes added. Finally, we show results from running tests using one or more modalities on these datasets.

4.6.1 State-of-the-art Methods

In D3.1, we looked at two types of searches: unimodal (using one type of data) and multimodal (using more than one type). For unimodal, we used mesh data because it gives more detail. We chose MeshNet and MeshCNN as the main methods¹⁴. A key advantage of the unimodal approach is that it is simpler and faster to implement, but it only works well, when both the query and the data are in the same format. For multimodal, we used CMCL¹⁵, CMIC¹⁶, MuseHash¹⁷, and LAH¹⁸ which work well for combining different types of 3D data. This approach allows for more flexible and accurate retrieval, but it is more complex and requires well-aligned data from multiple sources.

4.6.2 Datasets

In the previous deliverable (D3.1), we tested the methods on two public datasets: ModelNet40 and BuildingNet_v0, chosen because they relate to Demonstrator 1 (News Media) and Demonstrator 2 (Tourism & Automotive). In this deliverable, we used two new or updated datasets created during the XRECO project: XRECO.Buildings.Monuments and ModelNet40-Ext.

¹⁴ Jing, L., Vahdani, E., Tan, J., Tian, Y.: Cross-Modal Center Loss for 3D Cross-Modal Retrieval, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 3142-3151, openaccess.thecvf.com, (2021).

¹⁵ Jing, L., Vahdani, E., Tan, J., Tian, Y.: Cross-Modal Center Loss for 3D Cross-Modal Retrieval, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 3142-3151, openaccess.thecvf.com, (2021).

¹⁶ Lin, M. X., Yang, J., Wang, H., Lai, Y. K., Jia, R., Zhao, B., Gao, L.: Single image 3d shape retrieval via cross-modal instance and category contrastive learning, in Proceedings of the IEEE/CVF international conference on computer vision (pp. 11405-11415), (2021)

¹⁷ Pegia, M., Jónsson, B. P., Moutzidou, A., Diplaris, S., Gialampoukidis, I., Vrochidis, S., & Kompatsiaris, I. (2024, January). Multimodal 3D Object Retrieval. In International Conference on Multimedia Modeling (pp. 188-201). Cham: Springer Nature Switzerland.

¹⁸ Xie, Y., Liu, Y., Wang, Y., Gao, L., Wang, P., Zhou, K.: Label-attended hashing for multi-label image retrieval, IEEE Transactions on Cybernetics, <https://doi.org/https://doi.org/10.1109/TCYB.2021.3059886>, (2020)



XRECO.Buildings.Monuments¹⁹: This dataset contains 201 textured 3D mesh models of buildings and monuments, mainly from Europe, created using photogrammetry. The models, cleaned to focus only on the main structure, are labelled into 12 categories, including castle, church, mosque, palace, and villa, similar to BuildingNet_v0.

ModelNet40-Extended: This is an extended version of the ModelNet40 dataset²⁰, with 21 new 3D mesh models of radios and vintage electronics provided by RAI.

4.6.3 Experiments

As in D3.1, we used the same experimental setup. We tested how different hash code lengths and different numbers of training epochs affect performance. The methods were evaluated on the two datasets created for the project. We compared them using several metrics, such as Mean Average Precision (MAP), precision@k, recall@k and f-score@k.

4.6.3.1 Unimodal Retrieval Results

First, we tested using just one type of data, particularly meshes. Figure 4 and Figure 5 show the mAP results on the XRECO.Buildings.Monuments dataset and ModelNet40-Extended, using different hash code lengths and number of epochs.

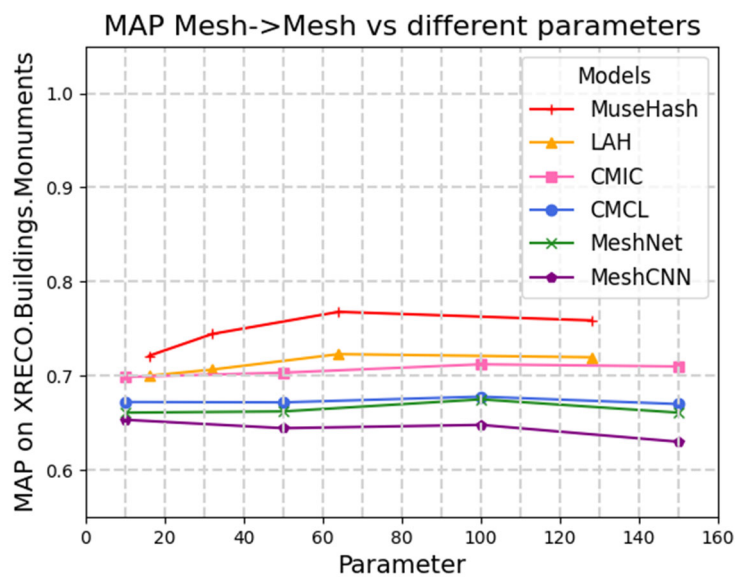


Figure 4: Unimodal results in terms of MAP on XRECO.Buildings.Monuments and dataset.

¹⁹ <https://zenodo.org/records/10809451>

²⁰ Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., Xiao, J.: 3D ShapeNets: A Deep Representation for Volumetric Shapes, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 1912-1920, cv-foundation.org, (2015)



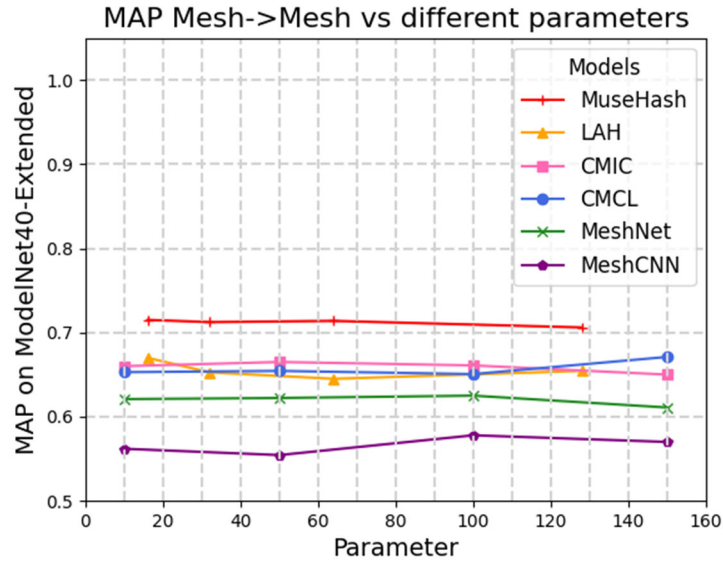


Figure 5: Unimodal results in terms of MAP on ModelNet40-Extended and dataset.

The hashing approaches, specifically MuseHash and LAH outperforms the other state-of-the-art methods. All methods present a stable behavior on both datasets as the hash code length or number of epochs is increased.

4.6.3.2 Cross-modal Retrieval Results

In this section, we test cross-modal retrieval, where one type of data is used to search for a different type. Figure 6 (a) and Figure 6 (b) show the results for the XRECO.Buildings.Monuments dataset. Figure 7 (a) and Figure 7 (b) show the results for the ModelNet40-Extended dataset.

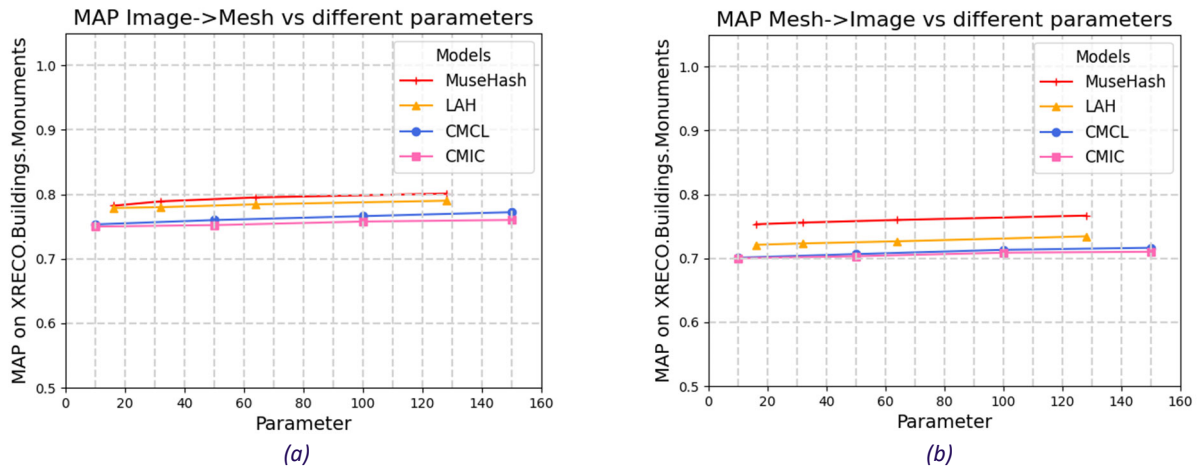


Figure 6: Cross-modal results in terms of MAP on XRECO.Buildings.Monuments dataset for (a) Image to Mesh and (b) Mesh to Image.

Similarly, MuseHash performs better than all the other state-of-the-art methods. Interestingly, the results for Image-to-Mesh retrieval are higher than those for Mesh-to-Image. This may be because image features, especially when extracted from multiple views, capture more distinctive visual patterns, making it easier to find matching 3D mesh structures.



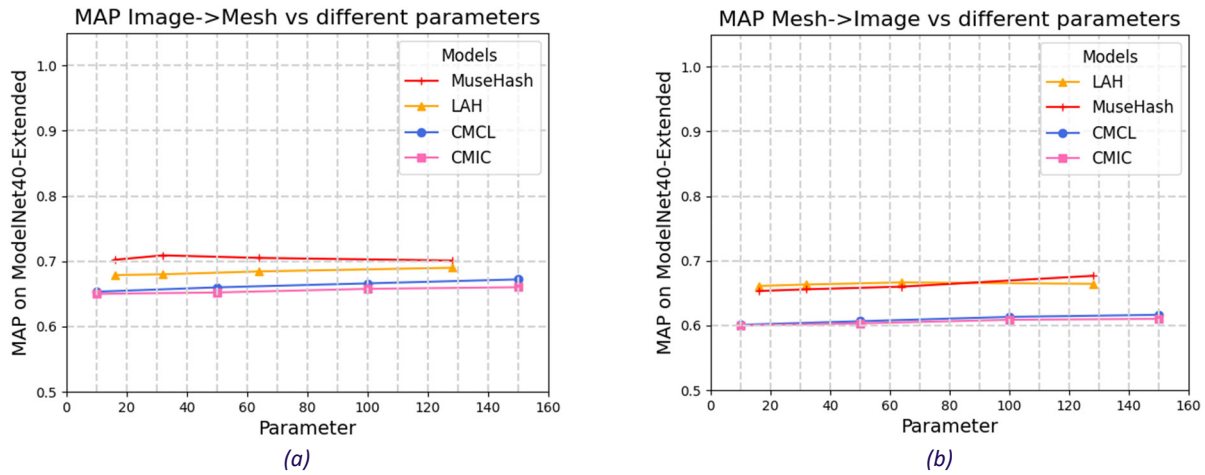


Figure 7: Cross-modal results in terms of MAP on ModelNet40-Extended dataset for (a) Image to Mesh and (b) Mesh to Image.

4.6.3.3 Multimodal Retrieval Results

In this section, we focus on multimodal retrieval, where all available data types are used for both searching and retrieving. Figure 8 (a) and Figure 8 (b) show the mAP results on the XRECO.Buildings.Monuments and ModelNet40-Extended datasets, respectively.

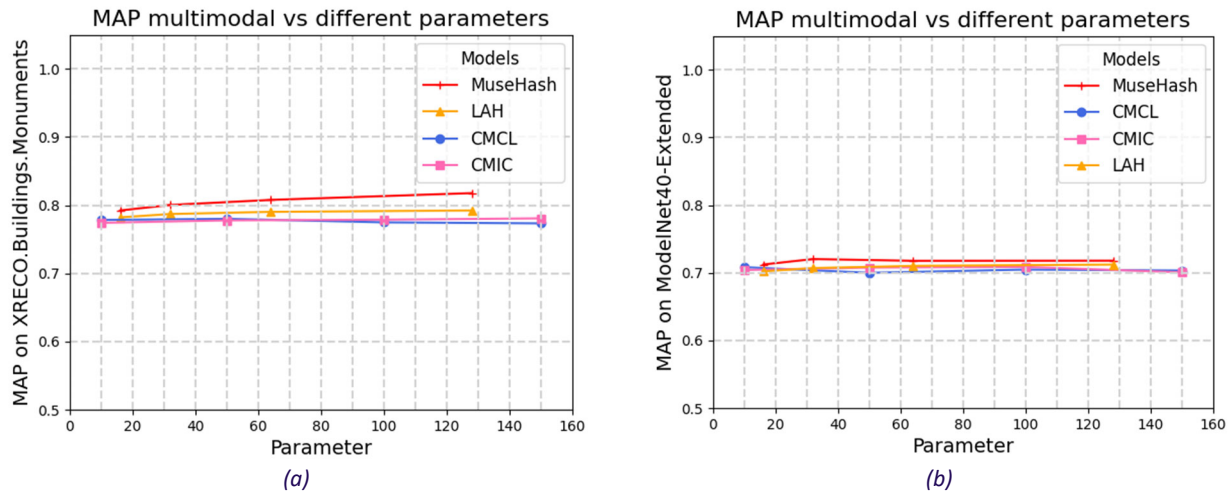


Figure 8: Multimodal results in terms of MAP on (a) XRECO.Buildings.Monuments and (b) ModelNet40-Extended dataset.

Next, we present the Precision@k, Recall@k, and F-score@k with k values of 10, 25, and 50. for the multimodal scenario across all methods. Results are shown in Table 4 for XRECO.Buildings.Monuments and in Table 5 for ModelNet40-Extended.



Table 4: Comparison of all methods based on Precision at k ($k = 10, 25, 50$) for different number of epochs or code lengths on XRECO.Buildings.Monuments dataset.

Method	Variable	Precision@k			Recall@k			Fscore@k		
		10	25	50	10	25	50	10	25	50
CMCL	10	0.7722	0.7745	0.7780	0.8901	0.9011	0.9122	0.8269	0.8330	0.8398
	50	0.7745	0.7751	0.7795	0.8956	0.9023	0.9132	0.8307	0.8339	0.8411
	100	0.7710	0.7731	0.7721	0.8912	0.9021	0.9051	0.8268	0.8326	0.8333
	150	0.7701	0.7694	0.7634	0.8723	0.8834	0.8890	0.8180	0.8225	0.8214
CMCI	10	0.7623	0.7645	0.7677	0.8731	0.8742	0.8801	0.8139	0.8157	0.8201
	50	0.7631	0.7661	0.7672	0.8820	0.8845	0.8878	0.8183	0.8211	0.8231
	100	0.7670	0.7682	0.7734	0.8811	0.8823	0.8831	0.8201	0.8213	0.8246
	150	0.7704	0.7745	0.7754	0.8901	0.8945	0.8953	0.8259	0.8302	0.8310
Method	Code Length	10	25	50	10	25	50	10	25	50
Muse Hash	16	0.7990	0.8004	0.8089	0.8655	0.8756	0.8761	0.8309	0.8363	0.8412
	32	0.8002	0.8056	0.8123	0.8671	0.8733	0.8771	0.8323	0.8381	0.8435
	64	0.8102	0.8144	0.8152	0.8678	0.8803	0.8841	0.8380	0.8461	0.8483
	128	0.8123	0.8153	0.8177	0.8692	0.8782	0.8891	0.8398	0.8456	0.8519

Table 5: Comparison of all methods based on Precision at k ($k = 10, 25, 50$) for different number of epochs or code lengths on ModelNet40-Extended dataset.

Variable		Precision@k			Recall@k			Fscore@k		
Method	Epochs	10	25	50	10	25	50	10	25	50
CMCL	10	0.6622	0.6645	0.6680	0.6901	0.6011	0.6122	0.6759	0.6312	0.6389
	50	0.6645	0.6751	0.6795	0.6956	0.6023	0.6132	0.6797	0.6366	0.6446
	100	0.6610	0.6731	0.6621	0.6912	0.6021	0.6051	0.6758	0.6356	0.6323
	150	0.6701	0.6694	0.6634	0.6723	0.6834	0.6890	0.6712	0.6763	0.6760
CMCI	10	0.6523	0.6645	0.6677	0.6731	0.6742	0.6801	0.6625	0.6693	0.6738
	50	0.6631	0.6561	0.6572	0.6820	0.6845	0.6878	0.6724	0.6700	0.6722
	100	0.6670	0.6582	0.6634	0.6811	0.6823	0.6831	0.6740	0.6700	0.6731
	150	0.6504	0.6545	0.6654	0.6901	0.6945	0.6953	0.6697	0.6739	0.6800
Method	Code Len	10	25	50	10	25	50	10	25	50
Muse Hash	16	0.7490	0.7404	0.7389	0.7655	0.7556	0.7461	0.7572	0.7479	0.7425
	32	0.7402	0.7456	0.7423	0.7671	0.7533	0.7471	0.7534	0.7494	0.7447
	64	0.7302	0.7444	0.7452	0.7678	0.7503	0.7441	0.7485	0.7473	0.7446
	128	0.7523	0.7453	0.7477	0.7692	0.7582	0.7491	0.7607	0.7517	0.7484

4.7 2D object detection and tracking (i2CAT)

In the preceding deliverable, we delineated the object detection and tracking system developed for the XRECO Multi-Object Detection and Tracking (MODT) service. The system architecture comprised two core



Artificial Intelligence (AI) modules: a multi-object detection component and a multi-object tracking component, enabling 2D MODT capabilities. Our development methodology incorporated a comprehensive state-of-the-art analysis of existing object detection and tracking techniques to identify optimal algorithms and solutions for subsequent implementation and rigorous testing.

4.7.1 Approach

Following this state-of-the-art review, YOLOX²¹ and BoT-SORT²² models were selected and deployed for object detection and tracking, respectively. The performance validation of these algorithms was meticulously conducted through a series of empirical tests utilizing real-world video sequences acquired from a highway operational environment. Figure 9 shows detection and tracking with YOLOX and BoT-SORT in 2 frames separated by about 2 seconds in a highway setting.



Figure 9: Detection and tracking with YOLOX and BoT-SORT in 2 frames separated by about 2 seconds in a highway setting.

After the last deliverable, the service was seamlessly integrated with our partners' infrastructure. System design and feature sets underwent iterative refinement in close collaboration with stakeholders to ensure accurate and robust deployment, leveraging data sourced from AWS S3 buckets.

4.7.2 Implementation as a service

Deployment of the 2D object detection and tracking models is facilitated through a dockerized Application Programming Interface (API). This method ensures a secure and isolated operational environment for the models and their requisite dependencies. The dockerized framework significantly streamlines application deployment and scalability, while also providing the agility to dynamically allocate computational resources tailored to the distinct demands of each model.

This MODT service is capable to identify and track 80 different classes of objects. For more information on the classes, please visit the repository's documentation. This application is built using docker-compose and 3 different Docker containers:

- **xreco-2dmodt-api-app**: a FastAPI container that acts as a producer for the system, it puts the user's queries to the queue and retrieve information about the state of the jobs.

²¹ Ge, Z., Liu, S., Wang, F., Li, Z., & Sun, J. (2021). YoloX: Exceeding yolo series in 2021. arXiv preprint arXiv:2107.08430.

²² Aharon, N., Orfaig, R., & Bobrovsky, B. Z. (2022). BoT-SORT: Robust associations multi-pedestrian tracking. arXiv preprint arXiv:2206.14651



- **xreco-2dmodt-celery-app**: a Celery container responsible for consuming jobs from the queue and performing the actual object detection and tracking inferences.
- **redis**: message broker and database for the queue and jobs.

Figure 10 depicts the overall architecture of the system. Although the service can run without using a GPU, It is desirable that the system has an NVIDIA GPU, and NVIDIA drivers and the NVIDIA Container Toolkit installed. Also, the multimedia files which have to be analyzed have to be provided by an AWS instance. The system should have valid credentials to access the data.

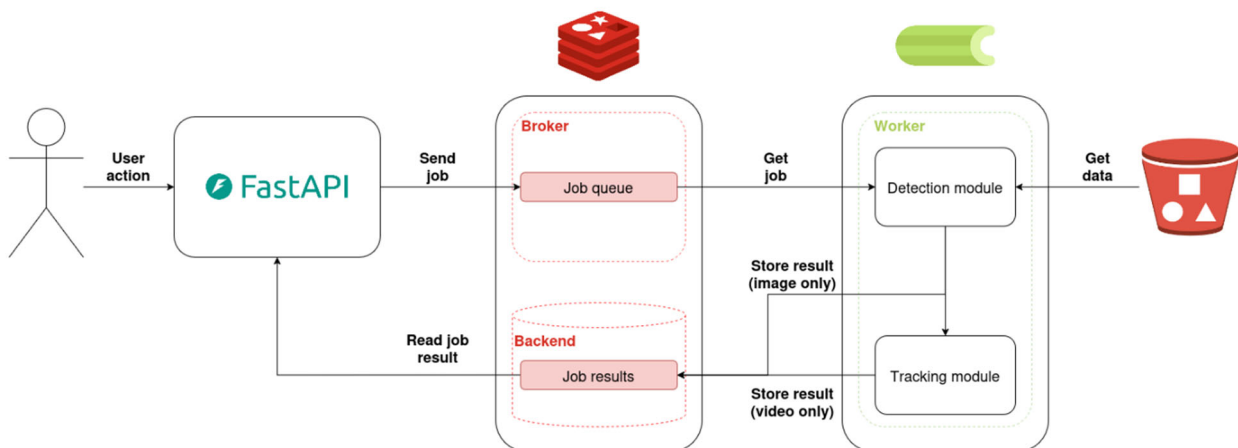


Figure 10: General application architecture. The user interacts through the FastAPI to send jobs and get jobs status. The celery worker consumes jobs from the redis job queue and downloads the target data from a S3 bucket located in the target AWS instance.

The API exposes several endpoints to interact with the application. For a complete documentation, please visit the repository's README.md or the automatic documentation generated by FastAPI, with interactive test and more detailed documentation, using the /docs/ endpoint. The most important interactions for the general user are:

Method	Endpoint	Description
POST	/inference/image/	For analyzing an image, only the target URL of the image has to be specified. This endpoint returns the job_id of the query once the query has been accepted in the queue.
POST	/inference/video/	To create a video analysis job, the target URL for the video has to be specified. Optionally, you can set an initial and/or end frame to only run the MODT in a segment of the video. This endpoint returns the job_id of the query once the query has been accepted in the queue.
GET	/status/{job_id}/	The status of a query can be checked using the following endpoint. The system returns the status of the job and the job result if available. The status of a job can be one of the following list: - I FAILURE: Job failed. - II PENDING: Job waiting in queue or job state is unknown (assumed pending as you have provided an ID). - III STARTED: Job was started by a worker. - IV SUCCESS: Job correctly succeeded.

5 User and workflow management

This section discusses infrastructure functionalities that are needed to implement both the content search functionalities as well as link them to the reconstruction services. One is user authentication, which is particularly relevant if the XReco platform is deployed as a hosted solution with users from multiple parties, but also, when remote services are invoked from the platform. The other are content baskets, as a concept of collecting search results related to a production, select a set of third-party items to be ingested or define a set of items to be passed to a particular reconstruction service.

5.1 Authentication services (i2Cat)

The authentication services continue to rely on the implemented Keycloak²³ identity and access management system. The latest changes to the JWT data model are reported and explained in detail in Deliverable 2.3; however, to support the platform's expansion, certain organizational configurations in Keycloak are detailed in this deliverable:

- Two dedicated realms have been defined: one for administrators, and one for the global platform organization, based on the Metasearch.
- Keycloak User Groups are used to configure organization-specific spaces. This information is extracted from the JWT token by the Orchestrator and later used in the marketplace UI as a rights management reference for assets.
- Realm roles are used to assign organization-specific access to the platform's various services (e.g. `marketplace_access`, `local_assets_access`, `metasearch_access`, `services_access`). Attributes are assigned to these roles for later use by the Orchestrator.
- Three backend instances of the NMR are currently deployed, creating separate spaces for assets: Marketplace, XReco Consortium Internal, and XReco External for external testers. User redirection to the appropriate backend is managed through specific attributes assigned via the realm roles for each organization.
- A single organization group has been created for each XReco consortium organization, along with a global group named "XReco External Testers" that includes all external testers.

No personal information from users is stored on the platform. All external tester accounts follow the generic format (`tester_<tester_number>@xreco.eu`), and the passwords are shared with external testers privately and individually. The information contained in the JWT tokens is not modifiable by third parties; the JWT technology is used as a secure and reliable mechanism for transmitting critical information across the deployed microservices and the platform backend.

5.2 Content basket management (i2Cat)

No relevant changes have been made to the content baskets integration or infrastructure details since Deliverable 3.1. The only specific changes to the data model, introduced to support authoring tools, are reported and explained in detail in Deliverable 2.3.

²³ <https://www.keycloak.org/>



6 Search services

6.1 Local search backend (UNIBAS)

The search backend is an integral part of the *Neural Media Repository (NMR)* and therefore described in Section 3.2.

6.2 Metasearch service (i2Cat)

The first version of the XReco Metasearch framework D3.1 demonstrated that a micro-service architecture can federate queries across heterogeneous audiovisual repositories and return a single ranked list. This deliverable document two major extensions introduced over the last development cycle:

- **AI-based Ranking Agent:** A multimodal re-ranking service that integrates scores produced by an AI based model with the existing BM25 based text ranker.
- **Garbage-Collection Workflow:** an event-driven maintenance micro-service that automatically removes stale connector responses, cached embeddings and orphaned thumbnails, guaranteeing deterministic latency and bounded storage.

This section explains the architectural changes, justifies the need for the new components, derives the fusion formula used by the Ranking Agent and presents the first evaluation results.

6.2.1 Microservice Based Architecture

The figure below illustrates the final schematic of the metasearch service. The main changes to this architecture are its complete integration with the XReco Orchestrator, now validated by both internal and external platform testers, and the addition of the Ranking Agent and Garbage Collector services. These two services run periodically and connect directly to the Elasticsearch backend. Both services rely on a Celery-based²⁴ queuing mechanism, and their execution frequency can be tuned to the host machine's capacity, making the system adaptable to various infrastructure sizes.

²⁴ <https://docs.celeryq.dev/en/stable/>



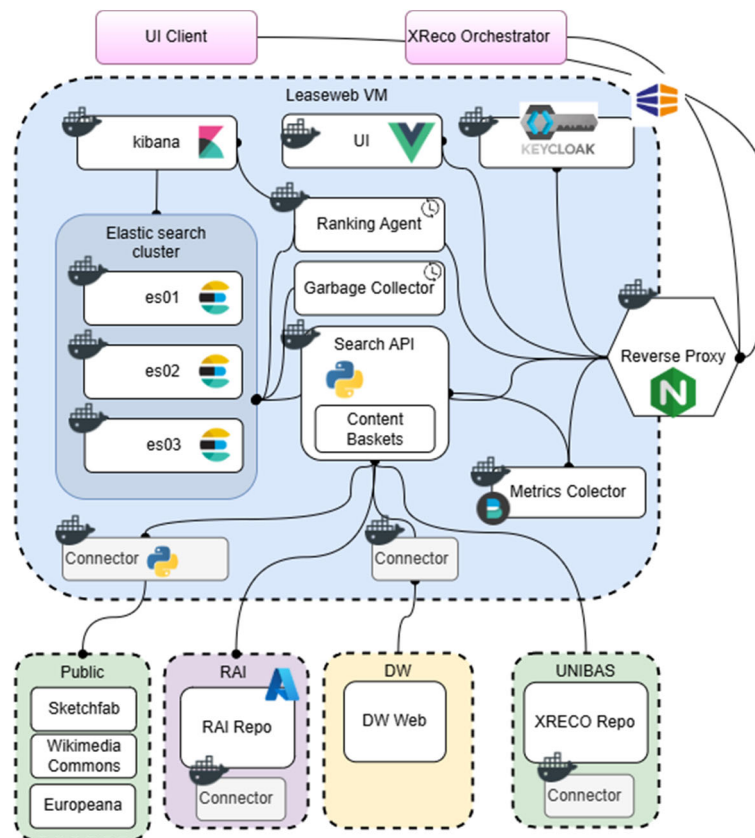


Figure 11 Complete metasearch Architecture

No changes have been imposed on connectors base. The Ranking Agent consumes the existing JSON schema produced by the currently integrated adapters (RAI-Solr, DW API, UNIBAS NMR, Wikimedia, Sketchfab API, Europeana API). This design decision preserved backward compatibility while allowing us to deploy the new service without redeploying the connectors themselves.

6.2.2 Ranking Agent

6.2.2.1 Architecture

The Garbage Collector (GC) and the Ranking Agent are packaged as autonomous Python microservices, each consisting of a lightweight Celery Worker combined with Celery Beat. Both are configurable via environment files, which are later translated into a cron scheduling format. Redis has been selected as the Celery pipeline in-memory backend, and Flower UI is deployed to monitor worker health and task execution. Additional workers can be seamlessly added as the infrastructure scales.



Figure 12 Ranking agent and GC deployment



6.2.2.2 Motivation

Real-world media archives seldom provide complete, well-structured metadata. Titles may be generic ("IMG_20210405"), descriptions can be absent, and controlled vocabularies differ across institutions. Under such conditions purely text-based retrieval underperforms relevant items remain buried because their sparse metadata shares no tokens with the query.

6.2.2.3 Model Choice

CLIP²⁵ (Contrastive Language–Image Pre-training) maps images and natural language into a shared vector space by maximizing the cosine similarity of matching pairs during training. This ability to measure cross-modal semantic similarity makes it ideal for rescuing items whose visual content is pertinent even when their textual fields are silent or misleading.

After benchmarking several checkpoints, we adopted openai/clip-vit-base-patch32(ViT-B/32)²⁶, Key facts relevant to our use case:

- Architecture – Vision Transformer “Base” (12 encoder layers, 768 hidden units, 12 attention heads) with 32 × 32 patch size.
- Input Resolution – 224 × 224 px images; textual queries are tokenized with a 49 408-token BPE vocabulary.
- Embedding Size – 512 floating-point dimensions for both modalities, enabling efficient cosine similarity and affordable storage.
- Training Data – CLIP is trained on 400 million (image, text) pairs scraped from the web, providing strong zero-shot capabilities across domains and languages.
- Latency vs. Quality Trade-off – ViT-B/32 is ~30 % faster than its Patch16 sibling on NVIDIA T4 GPUs while yielding only a minor drop in recall, making it the best fit for an online service.

6.2.2.4 Two Stage Ranking

The original engine relied exclusively on a BM25²⁷ with controlled weights calculation calculated over connector-specific textual fields (title, description, tags). Although effective when metadata is rich and homogeneous, this approach failed whenever descriptions were missing, multilingual or semantically distant from the query. The revised pipeline therefore splits the ranking task into a two phases process: a BM25 pre-rank that answers first-time queries instantly, and a CLIP-powered re-rank executed later in batch.

6.2.2.5 Offline Batch Re-ranking Workflow

CLIP inference remains the most expensive step, but we avoided introducing yet another storage technology. Instead, the existing Elasticsearch cluster serves double duty:

- The searches index keeps one document per user query (fields: term, _lastUpdated, _timesSearched, etc.).

²⁵ Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021). Learning transferable visual models from natural language supervision (arXiv:2103.00020). arXiv. <https://arxiv.org/abs/2103.00020>

²⁶ <https://huggingface.co/openai/clip-vit-base-patch32>

²⁷ https://lucene.apache.org/core/7_0_1/core/org/apache/lucene/search/similarities/BM25Similarity.html



- The `connectors_responses` index stores every connector hit and—when available—its 512-element CLIP embedding in a `dense_vector` field.

The system delivers relevance in two distinct phases:

1. **Immediate response (BM25-only):** When a user submits a query that the platform has never seen before, Metasearch fan-outs the call to the connectors and returns the text-based BM25 ranking right away. No multimodal processing is involved at this point, so the latency remains close to the connector round-trip time.
2. **Deferred multimodal re-rank (Clip-powered):** A background worker, launched via Celery Beat, periodically scans the `searches` index for the most-searched yet-to-be-reranked queries (see the endpoint `/tasks/add_rank_tasks`). For each qualifying `query_id` it enqueues a `tasks.query_results_rank` job. The job:
 - *Retrieves the previously cached connector responses.*
 - Computes CLIP embeddings for every preview image.
 - Applies the fusion formula and connector weights described later.
 - Persists the sorted list, including the new `Sfinal` score into the `(ranked_connectors_responses)` index.
 - Updates bookkeeping fields (`_lastRanked`, `_lastSentToRank`) in the originating document inside `searches`.

On the second and subsequent user requests the API can therefore serve a fully re-ranked list instantly without GPU involvement, because the heavy computation already happened in the parallel worker. Embeddings themselves are ephemeral: once the new ranking plus scalar `Sfinal` scores are stored, raw 512-D vectors are discarded. This keeps the index small and obviates the need for an external vector store while still reaping the benefits of CLIP.

6.2.2.6 Extended Score-Fusion Formula

To integrate heterogeneous similarity signals we now apply a three-stage procedure:

1. **Per-Modality Normalisation:** Previous BM25 based weighted formula, scores and cosine similarities are first rescaled to the 0-1 interval using adaptive min-max, adding a small epsilon (1e-8) to prevent division-by-zero.
2. **Logistic Calibration of CLIP Similarities:** Visual similarities below 0.25 rarely indicate relevance. A logistic transform:

$$S_{clip_{cal}} = 1 \div \left(1 + e^{(-\gamma \times (S_{clip_{norm}} - \tau))} \right)$$

where $\gamma = 14$ and $\tau = 0.32$ sharpen the separation between noise and signal.

3. **Weighted Linear Fusion:** Textual, visual and tag-based signals are blended into a single relevance score that the front-end finally displays.



$$S_{final} = w_c \times [\alpha \cdot S_{text_{norm}} + (1 - \alpha) \cdot S_{clip_{cal}}]$$

where α = text-vs-vision blend (default 0.6), w_c = connector trust weight.

If a connector returns a tag array, we compute the Jaccard overlap between the query tokens (after stop-word removal) and the tag set and extend the fusion as:

$$S_{final} = w_c \times [\alpha \cdot S_{text_{norm}} + (1 - \alpha) \cdot S_{clip_{cal}} + \beta \cdot J]$$

A preliminary grid-search on annotated pairs chose $\alpha = 0.6$ and $\beta = 0.15$ as the best trade-off between mean average precision and runtime. We plan that future iterations will learn both coefficients online from implicit feedback.

6.2.3 Garbage collection

The previous version relied on manual clean-ups triggered by system administrators. The new Garbage-Collector runs as a Cronjob. Applies retention policies derived from HTTP cache headers or a 30-day fallback, deletes expired raw hits, thumbnails and CLIP vectors, and pushes an updated JSON to Kibana.

6.2.3.1 Query results Refresh

Consist of a task that periodically updates results for the most consulted queries, the workflow can be summarized as follows:

- Selection of stale queries: The endpoint `add_garbage_collection_tasks` (see `ranking-agent/worker.py`) retrieves the N most searched queries from the `searches` index whose `_lastUpdated` is older than `GARBAGE_DAYS_AGO_THRESHOLD` days.
- Task fan-out: For every selected `query_id` a Celery job `tasks.update_results` is enqueued.
- Update phase in Metasearch API:
 - Fetches the query document from the `searches` index.
 - Calls all configured connectors asynchronously (`call_connectors_async`).
 - Deletes previous connector responses for that query (`delete_documents_by_attribute` on the `connectors_responses` index).
 - Indexes the fresh connector output and computes a new text-only ranking (`rank_connectors_responses`).
 - Updates bookkeeping fields: `_lastUpdated`, `_lastSentToRank`, `_timesSearched`.

6.2.3.2 Low-Frequency Query Eviction

Consist of a companion task that prune rarely searched and stale queries. This task is scheduled by Celery Beat but it uses its own configuration parameters. The workflow for the task can be summarized as follows:

1. Run an Elasticsearch Boolean query that filters by both usage count and age as defined above.
2. Enqueue one Celery `tasks.delete_query` job per matching `query_id`; each job removes connector responses, ranked responses and finally the corresponding `searches` document.

This mechanism clean-up will stop seldom-used, outdated queries from clogging the indices while ensuring that high-value searches remain fully cached.



6.2.4 Outlook

The CLIP-powered Ranking Agent improves recall in the presence of thin or inconsistent metadata, while the automated Garbage-Collector maintains predictable performance. Future extensions could address the refinement of the ranking workflow, possibly including:

- Exposing ranking formula weights as parameters and adjusting learning dynamically.
- Extending multimodal fusion to audio embeddings (OpenAI Whisper) or 3-D descriptors.
- Extend the Lexical BM25 core formula with Elasticsearch semantic capabilities (kNN), upgrading the base search formula to a hybrid-search.²⁸

²⁸ <https://www.elastic.co/what-is/hybrid-search>



7 Legal Requirements for Rights Management (KUL, IPR)

This section further develops certain key legal issues that emerge from the legal framework applicable to the relevant stages of XReco as a platform for the creation and sharing of XR content. It follows the stages that are elaborated in the context of the first WP3 deliverable (D3.1), i.e. the distinction of the XR production chain involving: (1) data and content ingestion; (2) search and retrieval tools; (3) XR services, and; (4) a marketplace for XR content. Specific additions to the legally-relevant technical aspects of the workflow are also made concerning the development of XR services. Relevant licensing approaches are discussed in Section 8.1 below.

In line with deliverable D3.1, the focus of the legal analysis is based on EU law. For illustrative purposes regarding certain aspects of the analysis, some national legal sources are also addressed.

7.1 Data and content ingestion

In line with deliverable D3.1, data and content ingestion is understood as the sourcing of pre-existing two-dimensional (2D) content and data. As highlighted, the ingestion process is generally understood to comprise the sourcing of content and data, and the potential storing of metadata about this content and data²⁹.

As highlighted³⁰, text and data mining (TDM) – “any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations”³¹ – may be carried out, subject to the specific conditions, for specific research purposes or generally. TDM exceptions have been transposed into national law by the EU Member States, as required by the CDSM Directive³², albeit in part with some significant delay³³.

The three main cases have addressed various aspects of the requirements of the TDM exceptions contained in Articles 3 and 4 of the CDSM Directive. At time of writing, three cases have been decided by national courts, namely in Germany, the Netherlands and in Hungary. Of these cases, the case in Hungary has been referred to the Court of Justice of the EU for a preliminary ruling.³⁴ In relying on the judgments at stake and scholarly commentary thereon, their potential impact on relevant TDM practices is briefly outlined. The cases in question address different fact patterns, and don't raise the same legal issues within the broader are of conflict concerning TDM. Formally, the three main cases address the following legal provisions:

- *LAION* (Germany): rules on the **research** TDM exception (Art. 3 CDSM Directive, section 60d German Copyright Act); *considers* – without ruling thereon – aspects of the general TDM exception (Art. 4 CDSM Directive, section 44b German Copyright Act)³⁵.

²⁹ D3.1, 78.

³⁰ D3.1, 84.

³¹ Art. 2(2) CDSM Directive.

³² By 7 June 2021 (Art. 29(1) CDSM Directive).

³³ Ana Lazarova, ‘The Last in Line: Bulgaria Implements the CDSM Directive’ (*Kluwer Copyright Blog*, 27 December 2023) <<https://copyrightblog.kluweriplaw.com/2023/12/27/the-last-in-line-bulgaria-implements-the-cdsm-directive/>>.

³⁴ Case C-250/25 *Like Company* [2025] (Request for preliminary ruling), published in OJ C/2025/3039.

³⁵ The court also ruled that the acts in question do not benefit from the temporary reproduction exception, as the reproduction was not transient nor incidental (*LAION* (Judgement of 27 September 2024, Landgericht Hamburg) openJur 2024, 9199 <<https://openjur.de/u/2495651.html>>, paras. 58-66).



- *Howardshome* (Netherlands): rules on general TDM exception (Art. 4 CDSM Directive, Art. 15o Dutch Copyright Act)³⁶.
- *Like Company* (Hungary): rules on general TDM exception (Art. 4 CDSM Directive, Art. 35/A Hungarian Copyright Act).

Within this case law, and in addition to the analysis carried out in D3.1, a specific legal issue that should be further elaborated is the dynamic role of ‘opt-outs’ from AI-related uses of data and content, which is relevant for XReco instances. As previously noted³⁷, the general TDM exception does not apply where rightholders enumerated reproduction, extraction and adaptation rights³⁸, have expressly reserved the use of their works or other subject matter in an appropriate manner³⁹. This is the case for rightholders of any of the enumerated reproduction, extraction and adaptation rights⁴⁰. This is an important consideration in the context of data and content ingestion, as the implementation of ingestion techniques may require a lawful authorisation to perform acts of reproduction under the EU copyright *acquis*⁴¹. In light of recent developments concerning AI and the risk of AI encroaching on creative practices, the two introduced TDM exceptions introduced have been especially controversial. The exact mechanism for opting-out, including also the developments concerning compliance with the AI Act discussed below, has led to concerns how exactly the general TDM exception will function⁴².

The first decision that addressed questions concerning the new TDM exceptions is the case referred to as *LAION*⁴³. This case, decided by the Regional Court of Hamburg (*Landesgericht Hamburg*) on 27 September 2024, concerned the dataset of LAION, a registered association that makes available a “so-called dataset for image-text pairs is”, contained in a “table document that contains hyperlinks to publicly accessible images or image files on the Internet as well as further information on the corresponding images, including an image description (also known as alternative text) that provides information on the content of the image in text form”⁴⁴. In line with Article 3 of the CDSM Directive, the court rules that the conduct of LAION in fact benefits from the German TDM exception for purposes of scientific research⁴⁵. The conclusion of the German court reaches an assessment of regarding whether the act is carried out by a research organisation for the purposes of scientific research⁴⁶. The activity carried out by LAION coincides with the general understanding of ‘TDM’,

³⁶ It should be noted that the court also ruled on exceptions for reporting, news, and quotation (*Howardshome* (Judgement of 30 October 2024, Rechtbank Amsterdam) C/13/737170 / HA ZA 23-690 <<https://uitspraken.rechtspraak.nl/details?id=ECLI:NL:RBAMS:2024:6563>>, paras. 4.24-28).

³⁷ D3.1, 85.

³⁸ Specifically, Art. 5(a) Database Directive; Art. 2 InfoSoc Directive; Art. 4(1)(a) and (b) Software Directive; Art. 15(1) CDSM Directive.

³⁹ I.e. “the use of works and other subject matter referred to in that paragraph has not been expressly reserved by their rightholders in an appropriate manner, such as machine-readable means in the case of content made publicly available online” (Art. 4(3) CDSM Directive).

⁴⁰ Specifically, Art. 5(a) Database Directive; Art. 2 InfoSoc Directive; Art. 4(1)(a) and (b) Software Directive; Art. 15(1) CDSM Directive.

⁴¹ D3.1, 84-85.

⁴² Issues include the univocality of such an opt-out, costs incurred by multiple opt-outs, remuneration, temporal effects of the opt-out, role of opt-outs also in third countries (extraterritoriality), and the practices yet to be established by the EU AI Office, see *infra* (Cf. Thomas Margoni, ‘TDM and generative AI: Lawful access and opt-outs’ (2024, forthcoming) Auteurs & Media, accessible at:

<https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5036164>, 11-12.

⁴³ *LAION* (Judgement of 27 September 2024, Landgericht Hamburg) openJur 2024, 9199 <<https://openjur.de/u/2495651.html>>.

⁴⁴ Translated from the original text: “Der Beklagte stellt unter der Bezeichnung “...” ein sogenanntes Dataset für Bild-Text-Paare öffentlich kostenfrei zur Verfügung. Es handelt sich dabei um eine Art Tabellendokument, das Hyperlinks zu im Internet öffentlich abrufbaren Bildern bzw. Bilddateien sowie weitere Informationen zu den entsprechenden Bildern enthält, darunter eine Bildbeschreibung (auch Alternativtext genannt), die Auskunft über den Inhalt des Bildes in Textform gibt” (*LAION* (Judgement of 27 September 2024, Landgericht Hamburg) openJur 2024, 9199 <<https://openjur.de/u/2495651.html>>, para. 5).

⁴⁵ Section 60d German Copyright Act, transposing Art. 3 DSM Directive.

⁴⁶ Cf. D3.1, 84.



which is common to both the TDM exception for research and the general TDM exception⁴⁷, but the court reached no conclusion regarding the applicability of general TDM exception.

Regarding opt-outs, and without reaching a conclusion on the general TDM exception, the German court in *LAION* addresses the general TDM exception's requirement that opt-outs should be 'machine-readable', highlighting that whether an opt-out expressed in natural language can be understood to comprise a machine-readable will always depend on the technical development at the time of the use of the work⁴⁸.

The *LAION* case was followed up by a case before the court of Amsterdam (*Rechtbank Amsterdam*), in the case referred to as '*Howardshome*'⁴⁹, which addresses the general TDM exception. Howardshome offered a commercial 'alert' service regarding news items publicly available online, and was sued by publishers of news media alleging infringement of their press publisher's right⁵⁰. Unlike the *LAION* case above, the court considered whether Howardshome benefitted from the commercial TDM exception⁵¹. The court found that the claimants had not sufficiently demonstrated that Howardshome was not using only publicly-available information, with Howardshome thus having lawful access. Further the court held that the publishers had not appropriately opted-out from Howardshome's bot. The court also rejected the idea of an 'implicit' opt-out.⁵²

The third and most recent case to address TDM is a decision by a Hungarian court, in which *Like Company*, the publisher of multiple online news sites, alleges infringement of its copyright by Google's training of its AI chatbot, Gemini (formerly Bard).⁵³ The district court of Budapest (*Budapest Környéki Törvényszék*) on appeal has subsequently requested a preliminary ruling from the Court of Justice of the EU (CJEU). In outlining the facts of the case, the court reiterates that Google's "chatbot provided a detailed response which included a summary of the information appearing in the news media belonging to the applicant",⁵⁴ i.e. Like Company. The CJEU is specifically asked to clarify whether "the process of training an LLM-based chatbot constitutes an instance of reproduction, where that LLM is built on the basis of the observation and matching of patterns, making it possible for the model to learn to recognise linguistic patterns", and whether "such reproduction of lawfully accessible works fall within the exception provided for in Article 4 of Directive 2019/790, which ensures free use for the purposes of text and data mining"⁵⁵.

⁴⁷ Cf. Section 60d(1) German Copyright Act, referring to section 44b(1) German Copyright Act.

⁴⁸ "Allerdings wird man die Frage, ob und unter welchen konkreten Voraussetzungen ein in "natürlicher Sprache" erklärter Vorbehalt auch als "maschinenverständlich" angesehen werden kann, stets in Abhängigkeit von der zum jeweils relevanten Werknutzungszeitpunkt bestehenden technischen Entwicklung beantworten müssen." (*LAION* (Judgement of 27 September 2024, Landgericht Hamburg) openJur 2024, 9199 <<https://openjur.de/u/2495651.html>>, para. 102).

⁴⁹ *Howardshome* (Judgement of 30 October 2024, Rechtbank Amsterdam) C/13/737170 / HA ZA 23-690 <<https://uitspraken.rechtspraak.nl/details?id=ECLI:NL:RBAMS:2024:6563>>.

⁵⁰ Art. 15 DSM Directive.

⁵¹ Article 15o of the Dutch Copyright Act, transposing Art 4 DSM Directive.

⁵² Etienne Valk and Iris Toepoel, 'DPG Media et al vs. HowardsHome – A National Ruling on DSM's Press Publishers' Rights and TDM Exceptions' (*Kluwer Copyright Blog*, 16 January 2025) <<https://copyrightblog.kluweriplaw.com/2025/01/16/dpg-media-et-al-vs-howardshome-a-national-ruling-on-dsms-press-publishers-rights-and-tdm-exceptions/>>.

⁵³ Peter Mezei, 'Third European Court Decision on the General Purpose TDM Exception Is Out' (*Kluwer Copyright Blog*, 8 May 2025) <<https://copyrightblog.kluweriplaw.com/2025/05/08/third-european-court-decision-on-the-general-purpose-tdm-exception-is-out/>>.

⁵⁴ Case 250/25 *Like Company* [2025] (3 April 2025, Request for a preliminary ruling) <<https://curia.europa.eu/juris/showPdf.jsf?text=&docid=300681&pageIndex=0&doclang=EN&mode=req&dir=&occ=first%E2%88%82=1&cid=5279466>>.

⁵⁵ Case 250/25 *Like Company* [2025] (3 April 2025, Request for a preliminary ruling) <<https://curia.europa.eu/juris/showPdf.jsf?text=&docid=300681&pageIndex=0&doclang=EN&mode=req&dir=&occ=first%E2%88%82=1&cid=5279466>>, 2.



There thus remains a significant level of uncertainty regarding the exact interpretation of the TDM exceptions, as well their applicability to specific cases such as the forms of data and content ingestion carried out in the context of a platform such as XReco. As already commented upon by Mezei⁵⁶, the Hungarian court seems to conflate web scraping and search engine indexing as two ‘forms’ of TDM. Opt-outs will need to be respected where the general TDM exception is relied upon, but *what* is being opt-out from. Critically, Mezei highlights that a failure to disaggregate steps of AI development risks “a mishmash of distinct technological activities and [subjecting] them to a single legal provision”⁵⁷.

In addressing the request for a preliminary ruling, and from the text of the DSM Directive, it is unlikely that the CJEU will entertain the idea that Article 4 “ensures free use for the purposes of text and data mining”, as insinuated by the Hungarian appeals court⁵⁸. Article 4 comprises principally an exception to the right of reproduction⁵⁹, and not for all uses covered by exclusive rights under copyright. Nevertheless, the case will be the first opportunity for the CJEU to clarify the scope and content of the new TDM exceptions in a turbulent time of AI development in a variety of contexts⁶⁰.

Opportunities

- Ingestion akin to *LAION* seems to be capable of benefitting from the research TDM exception
- Question concerning the definition of ‘text and data mining’ has been referred to the Court of Justice of the European Union for a harmonised interpretation

Risks

- Conflation of ‘mining’, ‘scraping’ and ‘training’ under the same legal notion

7.1.1 AI Act

Alongside the interpretation of the TDM exceptions, new practices are emerging in line with the EU’s AI Act that may affect forms of data and content ingestion. For clarity, the AI Act does not intervene in the EU copyright *acquis*. Rather, and among a range of other new rules, the AI Act also lays down specific requirements aimed at ensuring the compliance of providers of general-purpose AI (GPAI) models with EU copyright law⁶¹. These requirements are requirements for the providers of such models aimed at (proactively)

⁵⁶ Péter Mezei, ‘Third European Court Decision on the General Purpose TDM Exception Is Out’ (*Kluwer Copyright Blog*, 8 May 2025) <<https://copyrightblog.kluweriplaw.com/2025/05/08/third-european-court-decision-on-the-general-purpose-tdm-exception-is-out/>>.

⁵⁷ Péter Mezei, ‘The Multi-Layered Regulation of Rights Reservation (Opt-out) Under EU Copyright Law and the AI Act -For the Benefit of Whom? (V3.0)’ (Social Science Research Network, 31 March 2025) <<https://papers.ssrn.com/abstract=5064018>>, 17.

⁵⁸ It should also be noted that the Court of Justice has previously rejected a ‘right of free use’ (Case C-476/2017 *Pelham* [2019] ECLI:EU:C:2019:624, para. 22).

⁵⁹ As well as the right of extraction under the Database Directive and the right of adaptation under the Software Directive.

⁶⁰ Specifically, independent of consideration of the TDM exception contained in Article 4 CDSM Directive, the CJEU will need to clarify whether the process of training an LLM-based chatbot constitutes an instance of reproduction in the first place (Case 250/25 *Like Company* [2025] (3 April 2025, Request for a preliminary ruling) <<https://curia.europa.eu/juris/showPdf.jsf?text=&docid=300681&pageIndex=0&doclang=EN&mode=req&dir=&occ=first%E2%88%82=1&cid=5279466>>, 2).

⁶¹ “Any provider placing a general-purpose AI model on the Union market should comply with this obligation [to put in place a policy to comply with Union law on copyright and related rights], regardless of the jurisdiction in which the copyright-relevant acts underpinning the training of those general-purpose AI models take place. This is necessary to ensure a level playing field among providers of general-purpose AI models where no provider should be able to gain a competitive advantage in the Union market by applying lower copyright standards than those provided in the Union.” (Recital 106 AI Act); “This Regulation does not affect the enforcement of copyright rules as provided for under Union law.” (Recital 108 AI Act).



achieving compliance, but compliance with the AI Act’s rules of course do not guarantee compliance with copyright law as such.

These rules under the AI Act also only apply to the providers of a specific types of models. Namely, these general-purpose AI models are defined as “AI model (...) that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market”.⁶² In effect, such GPAI models can be understood to typically cover what is colloquially referred to as generative AI, such as those allowing the “flexible generation of content, such as in the form of text, audio, images or video”⁶³.

Certain aspects of the rules for GPAI model providers under the AI Act may be crucial for content ingestion practices. Notably, the AI Act specifies that such providers must “put in place a policy to comply with Union law on copyright and related rights, and in particular to identify and comply with, including through state-of-the-art technologies, a reservation of rights expressed pursuant to Article 4(3) of Directive (EU) 2019/790 [Copyright in the Digital Single Market Directive]”⁶⁴. In the absence of a harmonised standard, such providers “may rely on codes of practice (...) to demonstrate compliance” with this obligation⁶⁵. These codes of practice should also cover compliance with the above obligation concerning copyright⁶⁶, which should be finalised by 2 August 2025⁶⁷. In that regard, this code of practice and, once available, relevant standards may serve as an important form of guidance for the use of various AI tools.

Certain (downstream) 3D reconstruction techniques may utilise the data and content ingested and may permit such a flexible form of generation of 3D visual content. Whether this is the case depends on the technique in question is sufficiently general and capable of performing distinct tasks. In light of the techniques considered in deliverable D3.1⁶⁸, it is very unlikely that those techniques, in isolation, qualify as GPAI models. At the same time, some 3D reconstruction techniques may be (increasingly) integrated with such models, or alternatively, such models may be capable of carrying out forms of 3D reconstruction as part of their suit of ‘distinct tasks’. In that regard, it will remain important to consider what content has been sourced, including where it has been ingested in certain ways, including where 3D reconstruction techniques are integrated within the capabilities of a GPAI model.

Since the entry into force of the AI Act, the Commission has consulted iterations of the draft codes of practice, with three drafts being published⁶⁹. At the time of writing, the third draft of this code of practice addresses multiple relevant aspects that may have important implications for the training of general-purpose AI models.

⁶² Art. 3(63) AI Act.

⁶³ João Pedro Quintais, ‘Generative AI, Copyright and the AI Act’ (2025) 56 Computer Law & Security Review 106107, 5.

⁶⁴ Art. 53(1)(c) AI Act.

⁶⁵ Art. 53(4) AI Act.

⁶⁶ Art. 56(2) AI Act.

⁶⁷ Or, failing that, the Commission is empowered to adopt an implementing act (Art. 56(9) AI Act).

⁶⁸ Specifically the workflow described in: D3.1, 77-80.

⁶⁹ First: Commission, ‘First Draft of the General-Purpose AI Code of Practice published, written by independent experts’ (European Commission, 14 November 2024) <<https://digital-strategy.ec.europa.eu/en/library/first-draft-general-purpose-ai-code-practice-published-written-independent-experts>>; Second: Commission, ‘Second Draft of the General-Purpose AI Code of Practice published, written by independent experts’ (European Commission, 19 December 2024) <<https://digital-strategy.ec.europa.eu/en/library/second-draft-general-purpose-ai-code-practice-published-written-independent-experts>>; Third: Commission, ‘Third Draft of the General-Purpose AI Code of Practice published, written by independent experts’ (European Commission, 11 March 2025) <<https://digital-strategy.ec.europa.eu/en/library/third-draft-general-purpose-ai-code-practice-published-written-independent-experts>>.



Some of these measures under consideration may be informative for similar content ingestion practices, including⁷⁰:

- Drawing up, keeping up-to-date and implementation of a copyright policy in a single document;
- Non-circumvention of effective technological measures⁷¹, such as paywalls;
- Making reasonable efforts to avoid “piracy domains”;
- Ensuring that machine-readable reservations of rights⁷², such as robots.txt, are identified and complied with if using web crawlers⁷³;
- Enabling rightholders to obtain information about the web crawlers used;
- Making reasonable efforts to obtain adequate information about protected content not web-crawled;
- Mitigation of risk of producing copyright-infringing output, including the risk that the model memorises copyright training content, and by prohibiting copyright-infringing uses in acceptable use policies;
- Designation of a point of contact for rightholders.

Especially relevant is how these practices may be seen as a way of approaching compliance with the above TDM exceptions, including beyond the category of GPAI models. From the perspective of other techniques relying on these exceptions for their development, certain measures may be especially relevant. Measures such as those addressing web crawlers may be more specific to GPAI development, whereas those enabling rightholders to obtain information can flesh out how TDM ‘opt-outs’ can be respected at a general level⁷⁴. Rules concerning non-infringing outputs may be especially relevant at the level of AI services provision, but these rules may nonetheless have significant implications for content ingestion practices, e.g. where certain forms of ‘memorisation’ are prevented. Most importantly, however, is perhaps the repeated ‘reasonable efforts’ standards, especially regarding the obtaining of adequate information, which may be fulfilled by established data management practices.

A curious inclusion, however, is that concerning effective technological measures (also ‘technological protection measures’ or TPMs): the regulation of such measures is complex, yet – in principle – TPMs should not interfere with certain exceptions and limitation, and should not override the TDM exception⁷⁵.

At the same time, it needs to be underlined that data and content ingestion remains within the realm of case-by-case analysis. Data and content ingestion *can* be a method as part of AI development, but it can also occur in a manner entirely independent from the rest of the AI value chain. This is true also regarding other components of a platform such as XReco, where the level of integration of the separate tools provided is variable. This has been highlighted in the legal literature as well, as the level of web crawling implemented may not always be directly linked to training of GPAI models, or their downstream implementation within various forms of AI systems⁷⁶.

⁷⁰ Commission, ‘Third Draft of the General-Purpose AI Code of Practice published, written by independent experts’ (European Commission, 11 March 2025) <<https://digital-strategy.ec.europa.eu/en/library/third-draft-general-purpose-ai-code-practice-published-written-independent-experts>>.

⁷¹ As defined in Art. 6(3) InfoSoc Directive.

⁷² As expressed pursuant to Art. 4(3) CDSM Directive.

⁷³ Specifically, “if they use web-crawlers or have such web-crawlers used on their behalf to crawl, scrape and/or otherwise compile data for the purpose of text and data mining” (Third CoP, 3).

⁷⁴ Martin Kretschmer, Thomas Margoni and Pinar Oruç, ‘Copyright Law and the Lifecycle of Machine Learning Models’ (2024) 55 IIC - International Review of Intellectual Property and Competition Law 110, 130.

⁷⁵ Cf. Art. 6 CDSM Directive; Thomas Margoni and Martin Kretschmer, ‘A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology’ (2022) 71 GRUR International 685, 695-697.

⁷⁶ João Pedro Quintais, ‘Generative AI, Copyright and the AI Act’ (2025) 56 Computer Law & Security Review 106107, 6.



Opportunities

- Content and data ingestion for purposes of 3D reconstruction does not necessarily require the same quantity of content and data as the development of a general-purpose AI model.
- Downstream 3D reconstruction techniques do not necessarily display the level of generality, nor are they necessarily capable of performing a range of distinct tasks, akin to a general-purpose AI model.
- A proactive approach in light of the AI Act best practices could ‘flesh out’ compliance with the TDM exceptions and the practices that benefit from the exceptions.

Risks

- Where ‘generative AI’ tools are economically preferred to specific 3D reconstruction techniques, the applicability of the AI Act’s requirements may expand.

7.2 Search and retrieval

As highlighted in D3.1, search and retrieval is understood to comprise the component that enables finding and, where relevant, the use of suitable content and data in the context of XR services⁷⁷. As highlighted above, changes to the legal analysis of search and retrieval components depends especially on the interpretation of the copyright TDM exceptions.

7.3 XR Services: 3D Reconstruction

XR services and the law applicable thereto have been the subject of certain changes. In addition, certain clarifications should be made regarding the analysis of XR services made previously. A general overview of the process of 3D reconstruction – in amended form⁷⁸ – can be seen below. In brief⁷⁹, 3D reconstruction allows input 2D visual content of a singular underlying object (generally, from multiple angles and viewpoints – the ‘pose’) to output a singular piece of 3D visual content.

⁷⁷ D3.1, 78.

⁷⁸ Cf. D3.1, 89 et seqq.

⁷⁹ Cf. descriptions in D4.2.



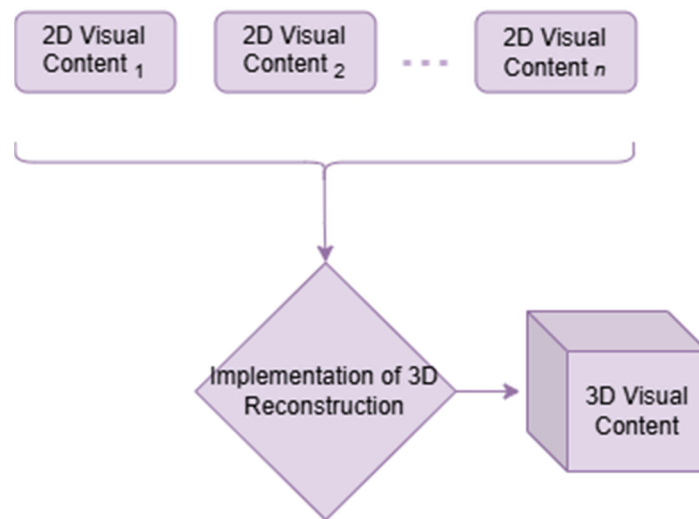


Figure 13: Relationship between 2D visual content, 3D reconstruction and 3D visual content

Some 3D reconstruction techniques may retain features of the underlying *2D visual content* in the output 3D visual content⁸⁰. Others may be designed and operated in such a manner to maximise the potential range of inputs⁸¹, meaning that as long as the underlying object (such as a known landmark) remains consistent, timing, lighting and other aspects ('occlusions') can be reduced in their impact on the output 3D visual content. This is typically done by supervising the 3D reconstruction techniques, e.g. via supervision signals. In effect, this means that, depending on the specific technique and how they are applied, the relationship between a singular piece of input 2D visual content on the one hand, and the output 3D visual content on the other, can be quite heterogeneous. Whether or not this amounts to a copyright-relevant act interfering with the exclusive rights of the relevant rightholders has been discussed in the D3.1. Further details relevant to case-by-case analysis are offered here.

7.3.1 The relation between 2D and 3D: Facts and data use

From the perspective of EU law, no reproduction occurs when elements reconstructed by a 3D reconstruction technique are not the author's own intellectual creation, i.e. they are not part of the *original expression* of the work. This will be the case, among others, when these elements have been dictated by technical considerations, rules or other constraints which have left no room for creative freedom⁸², and, as a consequence, they cannot be regarded as possessing the originality required for it to constitute a work.

Crucially, and as already underlined⁸³, copyright generally does not cover ideas. In terms of the CDSM Directive, this is also recognised as 'mere facts and data'⁸⁴, or more broadly, *facts and ideas*. Even if EU

⁸⁰ This is the standard case for NeRF, cf. Ben Mildenhall and others, 'NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis' (arXiv, 3 August 2020) <<http://arxiv.org/abs/2003.08934>> accessed 24 April 2023.

⁸¹ Dongbin Zhang and others, 'Gaussian in the Wild: 3D Gaussian Splatting for Unconstrained Image Collections' (arXiv, 14 July 2024) <<http://arxiv.org/abs/2403.15704>> accessed 2 June 2025; Jonas Kulhanek and others, 'WildGaussians: 3D Gaussian Splatting in the Wild' (arXiv, 31 October 2024) <<http://arxiv.org/abs/2407.08447>> accessed 2 June 2025; Weining Ren and others, 'NeRF On-the-Go: Exploiting Uncertainty for Distractor-Free NeRFs in the Wild' (arXiv, 2 June 2024) <<http://arxiv.org/abs/2405.18715>> accessed 2 June 2025.

⁸² See Case C-833/18 *Brompton* [2020] ECLI:EU:C:2020:461, para. 24.

⁸³ D3.1, p. 160.

⁸⁴ Cf. regarding TDM exceptions: "Text and data mining can also be carried out in relation to mere facts or data that are not protected by copyright, and in such instances no authorisation is required under copyright law." (Recital 9 CDSM Directive); as well as regarding the rights of press publishers: "They should also not extend to mere facts reported in press publications." (Recital 57 CDSM Directive); this principle is also recognised by the Software Directive in regard to computer programs: "For the avoidance of doubt, it has to be



Copyright *acquis* does not explicitly horizontally exclude them from copyrightable subject-matter, as international instruments do⁸⁵, this stems from the idea-expression or fact-expression dichotomy that characterizes the EU copyright *acquis*⁸⁶. However, they may be excluded from scope of protection by the implementation of the originality criterion. In fact, authors' intellectual creation cannot be mere facts and data that reflect the reality, because there is no author's own intellectual creation on them, reflecting the author's personality⁸⁷. This determination can be made by national courts⁸⁸, logically on a case-by-case basis.



Figure 14: Examples of 2D input content

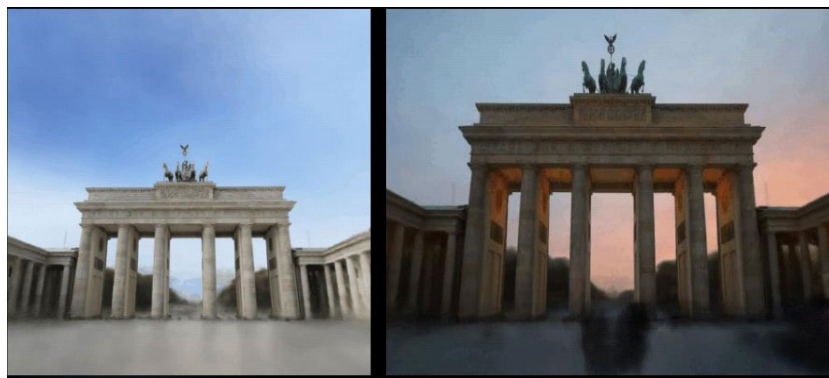


Figure 15: Examples stills of 3D reconstructions

It should be underlined, however, that for a potentially-significant number of applications of 3D reconstruction techniques, the new 3D asset contains unrecognisable fragments of the 2D content⁸⁹. A Depending on the applicable law, the example above in Figure 14 and Figure 15 may show this, i.e. where a photograph, its aspects views, angles and perspectives as original expressions are not necessarily recognisable in the 3D reconstructed asset. It should be noted that the examples above compare raw inputs with still views of a 3D reconstruction, which can be manipulated in 3D space. Beyond the role of originality,

made clear that only the expression of a computer program is protected and that ideas and principles which underlie any element of a program, including those which underlie its interfaces, are not protected by copyright under this Directive.” (Recital 11 Software Directive).

⁸⁵ See e.g. Art. 2(8) Berne Convention: “The protection of this Convention shall not apply to news of the day or to miscellaneous facts having the character of mere items of press information.” Cf. Case C-406/10 *SAS Institute Inc. v World Programming Ltd* [2012] ECLI:EU:C:2012:2592, para. 33: “With respect to international law, both Article 2 of the WIPO Copyright Treaty and Article 9(2) of the TRIPS Agreement provide that copyright protection extends to expressions and not to ideas, procedures, methods of operation or mathematical concepts as such.”

⁸⁶ See among others, Thomas Margoni & M. Kretschmer, ‘A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology’ (2022) 71 *GRUR International* 685, pp. 689-690; P Bernt Hugenholtz, ‘Copyright and the Expression Engine: Idea and Expression in AI-Assisted Creations’ (2024, forthcoming) *Chicago-Kent Law Review*, accessible at: < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4982516>.

⁸⁷ Case C-145/10 *Painer* [2011] ECLI:EU:C:2011:798, para. 94.

⁸⁸ Case C-5/08 *Infopaq I* [2009] ECLI:EU:C:2009:465, para. 48.

⁸⁹ Cf. the analysis in D3.1, 89-95.



which can be interpreted to have been harmonised⁹⁰, the applicable law may differ on certain aspects of interpreting the relationship between assets, which is outlined below.

7.3.2 The relation between 2D and 3D assets: reproductions and adaptations

The phenomenon that connects one piece of data or content (work) with another, where the second is in some manner based on certain information contained in the first, can be considered from multiple practical and legal perspectives. A few clarifications concerning this relationship between two pieces of content, including in the context of 3D reconstruction, are warranted.

Practically speaking, some techniques that are applied to the first may always lead to a second piece of data or content that reflects elements of the first. A classic example of this is a photocopy of a written piece of text. Independent of whether the text itself is protected or not, the second piece of data or content (the content of the photocopy) clearly resembles the underlying first piece. Where the written text is protected under copyright, the photocopy entails a reproduction thereof⁹¹. This consideration becomes more complex when the technique employed aims to reflect only certain specific aspects. An example of this could be the quoting of a text in another text. For such specific situations as quoting, the EU copyright *acquis* has therefore provided a specific exception for copyright and certain related rights pertaining to quotations⁹², subject to the fulfilment of further requirements, a quotation is permitted by law even in the absence of a license between rightholder and quoter. For acts of 3D reconstruction, there is no specialised legal treatment akin to that of quotation. This means that a further consideration of other existing rules is needed, bearing in mind the interpretational uncertainties that persist in EU law as it pertains to copyright.

In lay person's terms, a '3D reconstruction' that comprises a 3D output that is in some manner based on information contained within an underlying 2D input can be understood to be exactly that – a reconstruction, e.g. "[the] rebuilding of something natural, artificial, or abstract"⁹³; something that can be understood as a synonym of "revision", "reconversion" or "redesign"⁹⁴. At the same time, a reconstruction understood akin to a 'simulation', reconstruction can be understood to comprise also a "reproduction", a "copy" or a "recreation". This might give rise to the impression that 3D reconstruction techniques are always in some manner a reproduction in legal terms, or failing that, some type of derivative of the underlying 2D content, which is still under the legal control of the rightholder of that asset. For the sake of completeness, however, this impression is inaccurate when taken to the general understanding of 3D reconstruction, which refers to the production of 3D content based on *certain information contained* in said 2D content⁹⁵.

7.3.3 3D reconstruction as a reproduction/adaptation

As utilised in the first deliverable, the expression derivative works "refers to those works that are based on pre-existing works"⁹⁶. This expression may, however, fail to align with the general understanding of 3D reconstruction, especially as being based on the information contained in that work is different than being

⁹⁰ D3.1, pp- 89-90; cf. Thomas Margoni, 'The Harmonisation of EU Copyright Law: The Originality Standard' in Mark Perry (ed), Global Governance of

Intellectual Property in the 21st Century: Reflecting Policy Through Change (Springer International Publishing 2016).

⁹¹ Art. 2 InfoSoc Directive.

⁹² Art. 5(3)(d) InfoSoc Directive.

⁹³ OED def. 1.a. 2009

⁹⁴ MW 1 <https://www.merriam-webster.com/thesaurus/reconstruction>

⁹⁵ Cf. above regarding 'mere facts and data' use in section 7.3.1.

⁹⁶ D3.1, 90; referring to: Thomas Margoni, 'The digitisation of cultural heritage: originality, derivative works and (non) original photographs' (IViR, 2015), available at: <<https://www.ivir.nl/publicaties/download/1507.pdf>>, 18.



based on that work – the protected expression – as such. The term ‘derivative work’ can also be found in the text of the Berne Convention, arguably the most significant treaty of international law addressing copyright. Therein, it is provided that where the original work (e.g. a piece of 2D content) is protected by copyright⁹⁷, the author of that original work enjoys the exclusive right to authorise “adaptations, arrangements and other alterations”⁹⁸, as well as to authorise translations⁹⁹.

Despite being distinctly enumerated in the Berne Convention, however, the rights of adaptation, arrangement and other alteration and of translation are expressed differently across members of the Berne Union, including Member States of the EU and the EU itself. For instance, whereas the US provides an exclusive right to prepare derivative works based upon copyright work¹⁰⁰, this is based on an “expansive” definition of ‘derivative work’¹⁰¹. In other jurisdictions, the right of adaptation, arrangement and other alteration and the right of translation may be subsumed under the right of reproduction¹⁰².

The concept of a ‘derivative’ can be a lay person’s way to describe the relationship between ‘works’ (irrespective of whether either actually or potentially enjoy copyright or related rights protection)¹⁰³. Yet, as a matter of EU law, this lay understanding is not wholly accurate. EU law has horizontally harmonised certain rights, but not all. Most importantly for the purposes of 3D reconstruction remains the horizontally harmonised right of reproduction¹⁰⁴, whereas antecedent EU law also harmonise the right of reproduction vertically (i.e. in regard to specific protected subject matter)¹⁰⁵. By contrast, and as highlighted in deliverable D3.1, there is no horizontally-harmonised right of adaptation. This right is only addressed under EU law by the Database Directive¹⁰⁶ and the Software Directive¹⁰⁷.

By contrast, and especially in light of jurisdictions where acts that would be covered by the right of adaptation comprises a component of the right of reproduction¹⁰⁸, 3D reconstruction could be considered in light of the right of reproduction. As outlined briefly below, this would enable a comprehensive EU law analysis, e.g. where ‘derivatives’ are considered to necessarily interfere with the right of reproduction.

Copyright	EU law	Berne Convention
Reproduction	Recognised (horizontally: InfoSoc Directive; vertically: computer programs (Software Directive), databases (Database Directive), photographs (Term Directive))	Recognised
Adaptation	Only vertically (Database Directive, Software Directive)	Recognised

⁹⁷ The Berne Convention enumerates protected works in Article 2; it should be noted that it also protects ‘derivative work’ as works, i.e. “[translations], adaptations, arrangements of music and other alterations of a literary or artistic work (...) without prejudice to the copyright in the original work” (Art. 2(3) Berne Convention).

⁹⁸ Art. 12 Berne Convention.

⁹⁹ Art. 8 Berne Convention.

¹⁰⁰ 17 USC Section 106(2).

¹⁰¹ Goldstein and Hugenholtz, International Copyright, 299-300.

¹⁰² Or an “aspect” thereof: Goldstein and Hugenholtz, International Copyright, 299.

¹⁰³ This is one way to conceptualise a derivative also in the context of AI, machine-learning or neural network techniques, e.g. where AI models are understood as ‘derivatives’ of underlying data used to train them.

¹⁰⁴ Art. 2 InfoSoc Directive.

¹⁰⁵ Software Directive; Database Directive; Term Directive.

¹⁰⁶ Stipulating that “[in] respect of the expression of the database which is protectable by copyright, the author of a database shall have the exclusive right to carry out or to authorize: (...) translation, adaptation, arrangement and any other alteration” (Art. 5(b) Database Directive).

¹⁰⁷ Stipulating that it is the exclusive right of the rightholder to do or to authorise: “the translation, adaptation, arrangement and any other alteration of a computer program and the reproduction of the results thereof, without prejudice to the rights of the person who alters the program” (Art. 4(1)(b) Software Directive).

¹⁰⁸ Cf. Jongsma.



Irrespective of potential licensing solutions that may be available, exceptions and limitations may differ across the Member State copyright frameworks. Specifically, as highlighted e.g. by Rosati¹⁰⁹, any exception specific to the right of adaptation must be “limited to what can be considered as pure adaptations, not transformative uses of a work that nonetheless also involve its simple reproduction”.

MSs with separate rights of reproduction and adaptation	MSs where the right of adaptation is an aspect of the right of reproduction
Exceptions and limitations for reproductions	Exceptions and limitations for reproductions (also cover adaptations, unless specified)
Exceptions and limitations for adaptations	

As the case may be, some technologies may interfere with the rights of adaptation and/or reproduction, depending on the applicable law. Where they do so, the use of the underlying 2D visual content must be authorised, either by agreement (license) or by statute (exception or limitation to copyright). The following provides a few illustrations concerning how the right of adaptation is treated across EU Member States, and specifically, how it interacts with the right of reproduction thereunder. It should be noted that this is not a full comparative analysis of the legal state of the art. The below demonstrate the potential complexity of considering adaptations and derivative works from a cross-EU perspective.

7.3.3.1 Germany

In Germany, the right of adaptation is a distinct economic right (*Verwertungsrecht*) of authors. Section 23 of the German Copyright Act stipulates that “adaptations or other transformations of a work (...) may be published or exploited only with the author’s consent”¹¹⁰. Importantly, this section also clarifies when a ‘newly created work’ does not constitute an adaptation or other transformation.

The German Copyright Act was changed markedly in this respect after the CJEU’s decision in *Pelham*, wherein the Court held that the previous ‘free use’ exception¹¹¹ was incompatible with the exceptions and limitations outlined in Article 5 InfoSoc Directive¹¹². As of 7 June 2021, the free use exception was abolished, and the right of adaptation was appended with the present clarification stating that “[if] the newly created work maintains sufficient distance to the work used, this does not constitute adaptation or transformation (...)”¹¹³ (the ‘sufficient distance’ criterion). In line with the CJEU’s *Pelham* ruling, and going beyond the right of phonogram producers that was rules upon therein, the German Supreme Court (*Bundesgerichtshof*) has subsequently clarified that the sufficient distance criterion is to be evaluated by way of *recognisability*¹¹⁴.

¹⁰⁹ <https://ipkitten.blogspot.com/2014/05/the-right-of-adaptation-has-not-been.html>

¹¹⁰ Section 23(1) first sentence GCA; for the specific subject matter of film versions of a work, the execution of plans and drafts of an artistic work, the reproduction of an architectural work or the adaptation or transformation of a database work, the production of the adaptation/transformation requires the author’s consent also (section 23(2) GCA).

¹¹¹ Now repealed: Section 24 GCA.

¹¹² Specifically, “a Member State cannot, in its national law, lay down an exception or limitation other than those provided for in Article 5 of Directive 2001/29 to the phonogram producer’s right provided for in Article 2(c) of that directive” (Case C-476/2017 *Pelham* [2019] ECLI:EU:C:2019:624, para. 65).

¹¹³ Section 23(1) second sentence GCA.

¹¹⁴ “Für die Beurteilung des hinreichenden Abstands soll maßgeblich sein, inwieweit auch nach der Bearbeitung oder Umgestaltung noch ein Ausdruck der eigenen geistigen Schöpfung des Urhebers des vorbestehenden Werks erkennbar ist.” (BGH, Urteil vom 7.4.2022 – I ZR 222/20 – Porsche 911 – OLG Stuttgart, para 48).



Recognisability, as already highlighted in Deliverable 3.1¹¹⁵, is the standard outlined by the CJEU in *Pelham* to determine whether a new work infringes the copyright of a sampled work¹¹⁶. This is interpreted by the German Supreme Court to be consistent also with the previous interpretation of the free use exception, subject to which use of another work is non-infringing where the ‘impression’ of the expressive elements of the underlying work have ‘faded’ (lit. “verblassen”) to such a strong extent, that the underlying work is no longer recognisable¹¹⁷. Some have argued that the previous ‘fading’ jurisprudence regarding the free use exception is, however, less strict than the new recognisability standard, especially as even cases in which recognisable expression was retained in the newly created work would benefit from the exception, as long as the underlying work was superimposed by characteristics of the newly created work, when the overall impression (*Gesamteindruck*) of the two works was compared¹¹⁸. Overall, the delineation between “unfree adaptation” and “free use” remains open under German law¹¹⁹. It should also be noted that where the newly created output fails to qualify as a work protected by copyright, the use may still be non-infringing where the elements that give ground to the copyright protection of the underlying work, fade in the overall view of the newly created output¹²⁰.

Even though Germany thus provides for a separate right of adaptation, some German scholars have argued that the full harmonising effect of the economic rights under the InfoSoc Directive also capture the right of adaptation¹²¹. The above makes clear that the CJEU’s *Pelham* ruling has indeed significantly how the right of adaptation is to be interpreted. Nevertheless, it is not accurate to speak of the right of adaptation as an aspect of the right of reproduction under German law, akin to other EU jurisdictions. However, one potential avenue would be to consider the right of adaptation *after* an assessment of the right of reproduction has been made¹²².

7.3.3.2 Belgium

In Belgium, the right of adaptation is not a distinct right, but can be seen as an aspect of the right of reproduction. Specifically, Art. XI.165 s 1, second sentence clarifies that the author’s right of reproduction

¹¹⁵ D3.1, 91.

¹¹⁶ “In exercising [the freedom of the arts], the user of a sound sample, when creating a new work, may decide to modify the sample taken from a phonogram to such a degree that that sample is unrecognisable to the ear in that new work” (Case C-476/2017 *Pelham* [2019] ECLI:EU:C:2019:624, para. 36), cf. para. 39.

¹¹⁷ “Wie nach der bislang geltenden Rechtslage unter § URHG § 24 UrhG aF soll dann von einem hinreichenden Abstand ausgegangen werden können, wenn die aus dem vorbestehenden Werk entlehnten eigenpersönlichen Züge entsprechend der bisherigen Rechtsprechung des Bundesgerichtshofs dem Gesamteindruck nach gegenüber der Eigenart des neuen Werks so stark verblassen, dass das vorbestehende Werk nicht mehr oder nur noch rudimentär zu erkennen ist (BT-Drs. 19/27426, 78)” (BGH, Urteil vom 7.4.2022 – I ZR 222/20 – Porsche 911 – OLG Stuttgart, para 48).

¹¹⁸ Hartwig Ahlberg and Anne Lauber-Rönsberg, ‘§ 23’ in Horst-Peter Götting, Anne Lauber-Rönsberg and Nils Rauer (eds), *BeckOK Urheberrecht* (C.H. Beck 2025, 45th ed), para. 35

¹¹⁹ Felix Stang, ‘Freie Benutzung und unfreie Bearbeitung in der urheberrechtlichen Praxis’ (2024) 2024 Gewerblicher Rechtsschutz und Urheberrecht 176, 181; cf. regarding the impact of this on ‘user’s rights’: Haimo Schack, ‘Schutzgegenstand, „Ausnahmen oder Beschränkungen“ des Urheberrechts’ (2021) 2021 Gewerblicher Rechtsschutz und Urheberrecht 904.

¹²⁰ “die den Urheberrechtsschutz des älteren Werks begründenden Elemente im Rahmen der Gesamtschau in der neuen Gestaltung verblassen, also nicht mehr wiederzuerkennen sind” (BGH, Urteil vom 7.4.2022 – I ZR 222/20 – Porsche 911 – OLG Stuttgart, para. 58).

¹²¹ Joachim von Ungern-Sternberg, ‘§ 23: Bearbeitungen und Umgestaltungen’ in Gerhard Schricker and Ulrich Loewenheim (eds), *Urheberrecht* (C.H. Beck 2020, 6th ed), para 143; cited in Hartwig Ahlberg and Anne Lauber-Rönsberg, ‘§ 23’ in Horst-Peter Götting, Anne Lauber-Rönsberg and Nils Rauer (eds), *BeckOK Urheberrecht* (C.H. Beck 2025, 45th ed), para. 3.

¹²² Felix Stang, ‘Freie Benutzung und unfreie Bearbeitung in der urheberrechtlichen Praxis’ (2024) 2024 Gewerblicher Rechtsschutz und Urheberrecht 176, 179; cf. the argument that a broad interpretation of the pastiche exception to the right of reproduction may be a way to ‘revive’ free use: Helmut Haberstumpf, ‘Die freie Benutzung lebt!’ (2022) 2022 Zeitschrift für Urheber- und Medienrecht 795.



comprises (*omvat, comporte*) also the exclusive right to authorise adaptations or translations¹²³. There are further considerations of adaptations: In the special case of audiovisual works, subject to proof to the contrary, adaptors are also seen as authors of a collaborative work¹²⁴. In terms of copyright contract law, the granting of an audiovisual adaptation right must also be addressed by a contract separate from that addressing the publication of the work¹²⁵. However, it should be noted that the right of reproduction of the author is notably broad under Belgian law, covering also the author's right of distribution and rental and lending rights, with scholars commenting that it can be seen as an "assignment right" (*bestemmingsrecht*) for articles of reproduction¹²⁶. Generally, under Belgian law, an adaptation must, on the one hand, take essential, concrete and original elements from the original work, and on the other hand, must add its own design (*eigen vormgeving*)¹²⁷. Netherlands

7.3.3.3 Netherlands

The Dutch Copyright Act (*Auteurswet*) considers (whole or partial) adaptations as well as any "imitation in a changed form" (*nabootsing in gewijzigden vorm*) as a form of reproduction (*verveelvoudiging*)¹²⁸. This notion of reproduction (*verveelvoudiging*) should be distinguished from the 'primary' understanding of reproduction (*reproducen*), which is already covered by the overarching notion of *verveelvoudiging*. In that regard and as a matter of factual assessment, adaptations are considered reproductions only if copyright-protected elements are taken over from the underlying work¹²⁹. The scope of protection also depends on the originality of the underlying work¹³⁰. Alongside this, the Dutch Copyright Act restates the Berne Convention rule regarding derivative works as independent works of¹³¹.

7.3.3.4 France

Under French law the right of adaptation is not a separate economic right but it is rather included under the orbit of the right of reproduction (art. L. 122-3 Intellectual Property Code)¹³². Adaptation is expressly covered by exclusive powers of the author given that, according to art. L. 122-4 Intellectual Property Code any translation, adaptation or transformation, arrangement or reproduction by any art or process whatsoever, in whole or in part without the consent of the author or his successors or assigns is unlawful. There is no specific regulation of that right in the law, including the definition of the concept of adaptation or the criteria for assessing its subsistence. It has been suggested by the French Superior Council of Literary and Artistic Property, even prior to the Pelham jurisprudence, that the transformation of the original work would be

¹²³ Original Dutch text: "Dat recht omvat onder meer het exclusieve recht om toestemming te geven tot het bewerken of het vertalen van het werk."; original French text: "Ce droit comporte notamment le droit exclusif d'en autoriser l'adaptation ou la traduction." (Book XI of the Code of Economic Rights).

¹²⁴ "Behoudens tegenbewijs worden geacht auteurs te zijn van een in samenwerking tot stand gebrachte audiovisueel werk: (...) (b) de bewerker" "Sont présumés, sauf preuve contraire, auteurs d'une oeuvre audiovisuelle réalisée en collaboration (...) (b) l'auteur de l'adaptation" (Art. XI.179 CER)

¹²⁵ Art. XI.184.

¹²⁶ Jeff Keustermans and Peter Blomme, *Auteursrecht – Capita selecta* (Intersentia 2021, 2nd ed), 145; cf. Frank Gotzen, *Het Bestemmingsrecht van de Auteur* (Larcier 1974).

¹²⁷ Hendrik Vanhees, *Handboek intellectuele rechten* (Intersentia 2020), 40.

¹²⁸ Art. 13 Dutch Copyright Act.

¹²⁹ Spaanse tegels in Dirk Visser 2025.

¹³⁰ Heertje/Hollebrand in Dirk Visser 2025.

¹³¹ "Verveelvoudigingen in gewijzigde vorm van een werk van letterkunde, wetenschap of kunst, zoals vertalingen, muziekschikkingen, verfilmingen en andere bewerkingen, zomede verzamelingen van verschillende werken, worden, onverminderd het auteursrecht op het oorspronkelijke werk, als zelfstandige werken beschermd." (Art. 10(2) Dutch Copyright Act).

¹³² Audrey Lebois, 'Droit de reproduction (CPI, ar. L. 122-3)' (2021) *Juris-Classeur Propriété littéraire et artistique*, Fasc. 1246,



subject to authorization by the authors of the initial works only if it is communicated to the public in a form that enables its identification¹³³.

7.3.3.5 Greece

Under Greek Copyright law (Law No 2121/1993), the right of adaptation is a distinct exclusive economic right that is separate from both the right of fixation and reproduction and the right of translation. Under the terms of Art. 3(1)(c) Law No 2121/1993, the economic rights shall confer upon the authors notably the right to authorize or prohibit [...] c) the arrangement, adaptation or other alteration of their works. Greek law does not provide a definition of “derivative works” and no specific criteria for assessing the existence of a derivative work which triggers the right of adaptation. It is argued by scholars that the right in question covers the creation of new works by means of modification of any kind¹³⁴ of a preexisting work and that it presupposes a) the taking of the original expression of the preexisting work within the new work plus b) a creative (original) contribution from the side of the adapter¹³⁵. This right is understood as generally covering the creation of derivative works¹³⁶. A neighbouring concept is that of “composite work” (Art. 7(3) Law No 2121/1993), i.e. a work which is composed of parts created separately, such as a collage. The right of adaptation is not recognised in favour of related rights rightholders.

7.3.3.6 General

Of the Member States jurisdictions briefly outlined above, the only one that has a significant demarcation between the right of reproduction and the right of adaptation is Germany. Even in Germany, the right of adaptation has been significantly impacted by EU jurisprudence concerning the right of reproduction, as shown by the changes to German copyright law’s treatment of ‘free use’ and reformulation of the ‘sufficient distance’ criterion. As this is not an exhaustive discussion of the right of adaptation across the EU. However, it is likely that for those Member States that, like Germany, treat it as a separate right of authorship, the impact of the *Pelham* judgement may be similar, if not near identical.

In that regard, the considerations of exceptions and limitations carried out in D3.1 focus on those that are statutory authorisation regarding *reproductions*. Adaptation – while a separately considered economic right of authorship by the Berne Convention and under the law of some Member States – may have a differing levels of relevance across the Member States.

¹³³ Valérie Laure Benabou, ‘RAPPORT DE LA MISSION DU CSPLA SUR LES « ŒUVRES TRANSFORMATIVES »’ (CSPLA, 2014), 36.

¹³⁴ D. Kallinikou, *Copyright and related rights*, (P.N. Sakkoulas 2024, 5th ed), 217.

¹³⁵ K. Christodoulou, *Copyright Law*, 2nd ed., (Nomiki Vivliothiki 2023, 2nd ed), 114-115; Theodoros Chiou, The non-creative alteration of a work as form of plagiarism EEMPD, 463-472.

¹³⁶ Irini Stamatoudi, ‘Article 3’ in Irini Stamatoudi (ed), *Copyright law* (Nomiki Vivliothiki 2025), 94.



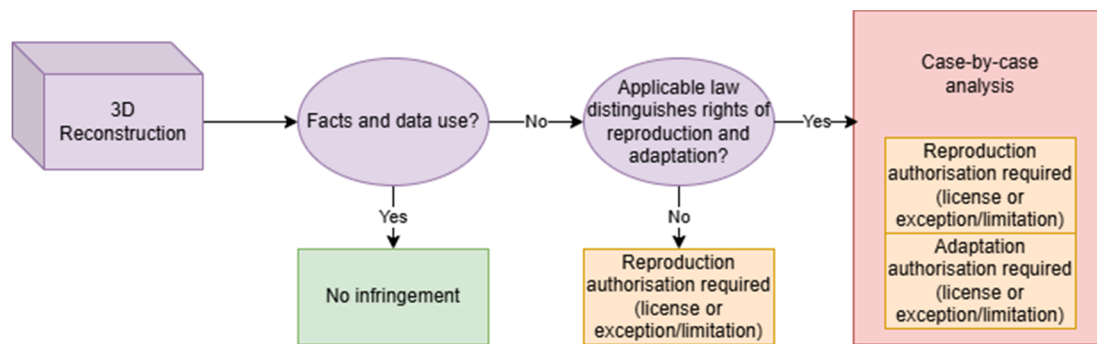


Figure 16: Need for authorisations based on reproduction and adaptation rights.

7.3.4 Implications for Exceptions and Limitations to Copyright Law

As discussed in D3.1, copyright law does not only the rights of rightholders, but also provides a number of exceptions and limitation to those rights, which are permissions or authorisations for certain actors to perform certain acts on a work or other subject matter¹³⁷. The above shows that reproductions and adaptations may be structured differently across the EU Member States, which may impact the availability of exceptions and limitations, because these are *exceptions and limitations to specific rights*. The below briefly addresses the TDM and pastiche exceptions.

7.3.4.1 TDM

The TDM exceptions to copyright address the right of reproduction, including in regard of certain related rights, the right of extraction in regard of *sui generis* databases, and the right of adaptation for computer programs. As highlighted above concerning data and content ingestion, the applicability of TDM exceptions under the EU copyright *acquis* remains uncertain in regard of the case law and the scholarly commentary therein. As highlighted, there remains a risk that TDM is conflated with the subsequent processes and uses of the mined data and content.

7.3.4.2 Pastiche

One exception that has not been interpreted by the CJEU thus far is the exception for ‘pastiche’. Specifically, under Article 5(3)(k) of the InfoSoc Directive, EU Member States may provide for exceptions to the rights of reproduction, of communication to the public and of making available to the public¹³⁸, in case of “use for the purpose of (...) pastiche”. In a request for preliminary rule of 2023, the German Supreme Court has posed two questions to the CJEU concerning the specific provision regarding pastiche¹³⁹. First, it asked whether pastiche is “a catch-all clause at least for artistic engagement with a pre-existing work or other object of reference, including sampling” and whether pastiche is subject to “limiting criteria, such as the requirement of humour, stylistic imitation or tribute”. Secondly, it asked whether the pastiche exception requires a “determination of an intention on the part of the user to use copyright subject matter for the purpose of a pastiche”¹⁴⁰.

¹³⁷ D3.1, Annex VIII.

¹³⁸ Artt. 2 and 3 InfoSoc Directive.

¹³⁹ Case C-590/23 *Pelham*, also referred to as “Pelham II”.

¹⁴⁰ Péter Mezei and others, ‘Oops, I Sampled Again ... the Meaning of “Pastiche” as an Autonomous Concept Under EU Copyright Law’ [2024] IIC - International Review of Intellectual Property and Competition Law.



There are varying understandings of the notion of pastiche across EU Member States, yet it is unlikely that pastiche will emerge as the EU's 'catch-all clause' akin to the US copyright system's approach to fair use¹⁴¹. The Advocate General has supported a "contextual and purposive interpretation"¹⁴², proposing that a uniform interpretation of the concept of pastiche as "an artistic creation which (i) evokes an existing work, by adopting its distinctive 'aesthetic language' while (ii) being noticeably different from the source imitated, and (iii) is intended to be recognised as an imitation."¹⁴³ Whether the CJEU will follow the advice of the Advocate General is yet uncertain. As the case may be, the CJEU's interpretation may bear relevance for the application of 3D reconstruction techniques.

7.3.5 Remaining points of legal (un-)certainty

It should be reiterated here that EU law has only partially harmonised copyright law across its Member States, with important implications for the legal analysis of 3D reconstruction techniques. This is a consequence of EU law as a creation of international law, while it is demarcated from 'traditional' international law due to its "distinctive supranationality"¹⁴⁴. This puts EU law in a complex relationship with international law, including international copyright law such as the Berne Convention.

Of special consequence for a dynamic set of technologies such as 3D reconstruction and its categorisation under the EU copyright *acquis*, is the lacking hierarchy of interpretation followed by the CJEU. For instance, as demonstrated by Rosati, whereas international instruments such as the Berne Convention plays a role in the CJEU's rulings concerning economic rights, it ranks below the standards of a 'high level of protection' and 'interpretation in light of objectives pursued by legislation at issue'¹⁴⁵. Such findings are consistent with general findings of the CJEU's interpretational practice, which may tend to "over-use teleological interpretation", follow an inconsistent and unforeseeable pattern apart from the consolidation of EU law, which is especially recurrent in subject-specific areas¹⁴⁶. In that context, it remains uncertain how the EU copyright framework will address issues specific to 3D reconstruction. In view of the ambiguities and lacunae connected to the contours of the concepts of adaptation and derivative works, including under national law, the harmonisation of the exclusive right of adaptation as secured by the Berne Convention needs to be further explored as a matter of EU law.

The latest developments concerning pastiche have spurred on further discussion on the flexibility of the EU copyright system in light of the settled interpretation of the *acquis* by the CJEU. For instance, the Advocate General clarified, in light of the balance of copyright between the 'two creators' at stake in a pastiche¹⁴⁷, that "it could be desirable to increase the flexibility of [the copyright system] with respect to artistic (or even communicative) reuse of protected material".¹⁴⁸ Indeed, scholarly studies have shown that the present regulation of copyright flexibilities is fragmented with remaining lacunae, is legally uncertain and outdated, making it difficult to gauge the impact in new settings.¹⁴⁹

¹⁴¹ Péter Mezei and others, 'Oops, I Sampled Again ... the Meaning of "Pastiche" as an Autonomous Concept Under EU Copyright Law' [2024] IIC - International Review of Intellectual Property and Competition Law, 1248-1253.

¹⁴² Case- C-590/23 *Pelham* [2025] ECLI:EU:C:2025:452, Opinion of AG Emiliou, para. 57.

¹⁴³ Case- C-590/23 *Pelham* [2025] ECLI:EU:C:2025:452, Opinion of AG Emiliou, paras. 81 and 133.

¹⁴⁴ Such as the granting of rights to individuals and the primacy of EU law over national law (Rudolf Streinz 161)

¹⁴⁵ Eleonora Rosati, *Copyright and the Court of Justice of the European Union* (Oxford University Press 2019), 62-69.

¹⁴⁶ Marcella Favale, 'The role of the Court of Justice in the development of EU copyright law: an empirical experience' in Irini Stamatoudi and Paul Torremans (eds), *EU Copyright Law* (Elgar 2021, 2nd ed), 1078-1079.

¹⁴⁷ Case- C-590/23 *Pelham* [2025] ECLI:EU:C:2025:452, Opinion of AG Emiliou, para. 128.

¹⁴⁸ Case- C-590/23 *Pelham* [2025] ECLI:EU:C:2025:452, Opinion of AG Emiliou, para. 131.

¹⁴⁹ Caterina Sganga, 'The Past, Present and Future of EU Copyright Flexibilities' (2024) 55 IIC - International Review of Intellectual Property and Competition Law 5, 11-12.



In that regard, the phenomena studied in the context of XReco should form part of the policy considerations in the domain of copyright, as the specific needs of the creative sector utilising 3D reconstruction techniques, in due respect of the rights of authors of 2D material, need to be adequately addressed. Whereas the practices emerging under the AI Act may be relevant on a case-by-case basis, they are an imperfect fit for the specific cases of 3D reconstruction techniques.

Opportunities

- The legal analysis of D3.1 concerning facts and data use remains valid. Where only facts and data are used, the EU copyright *acquis* generally does not apply.
- Pending the interpretation of the Court of Justice of the European Union, some applications of 3D reconstruction techniques that are using the protected expression of an underlying 2D work may benefit from the pastiche exception.

Risks

- Especially concerning notions of ‘derivatives’ or adaptations, EU Member States may pursue quite different approaches.
- There is a general lack of clarification of the applicability of emerging interpretations of copyright concepts vis-à-vis specific practices such as 3D reconstruction.

7.4 XR Marketplace

As outlined in D3.1, a marketplace for XR assets as proposed by the DoA can be understood to comprise a facility that enables the commercialisation of individual or multiple 3D content and content and data for their use in commercial applications in return for remuneration, which may be federated¹⁵⁰. In order to be an effective, such a marketplace would need to determine which aspects of it are regulated by the potential operator of a platform such as XReco, with relevant implications for the platform governance thereof¹⁵¹. From an open-ended perspective concerning the operator of the platform, approaches to licensing of 3D assets within the XR marketplace are discussed in greater detail in Section 8.1 below.

¹⁵⁰ D3.1, 78.

¹⁵¹ Cf. D3.1, 75-76.



8 Rights management and licensing tools (FINC, KUL, IPR)

8.1 XReco marketplace licensing approach

8.1.1 Background from D3.1

Any placement of digital visual assets, such as 3D assets in an online (web-based) XReco Marketplace (marketing or “sharing” digital assets) requires (1) storage of the file that embodies the asset, which, in turn will make possible (2) the display (visual appearance) and online dissemination of that asset via the marketplace website and, subsequently, its further use by buyers/end-users as marketed (purchased) asset (3).

As already discussed in D3.1¹⁵², both storage and display of a copyrighted 3D asset are copyright-sensitive acts. In particular, it has been established that the upload and storage of the digital file would be qualified as **reproduction** (art. 2(a) Infosoc Directive) and the public online dissemination would constitute an act of **communication to the public (making available to the public)** (Art. 3(1) Infosoc Directive). Accordingly, all relevant rights over marketed 3D assets should be sufficiently cleared (typically by means of a license granted by the relevant rightholder(s)-Licensor) prior to placement/sharing of the asset through the Marketplace. At the same time, the downstream use and exploitation of the marketed asset by any end-user, who wishes to publicly use the acquired purchased 3D asset, probably also involves the realisation of copyright-significant acts, such as reproduction, communication to the public etc.

Accordingly, already under D3.1¹⁵³, rights management has been seen as a core component when designing a copyright-compliant XReco Marketplace for 3D assets and licensing has been found as being the default solution for rights management in the two instances that have been determined:

- **Marketplace license** for the upstream use (Marketplace placement as such) of the asset, which would typically make part of the Terms of Service applicable to XReco Marketplace (Licensor-service provider relation), and
- **End-user license** for the downstream use (licensing conditions) of the marketed 3D asset by the buyer/end-user (Licensor-buyer/end-user relation).

Under both licensing instances, the Licensor should be entitled to license all relevant rights, which typically means that he should hold the relevant rights, either as an owner or as a (sub)licensee¹⁵⁴.

The below sections outline the main components of XReco licensing scheme that is designed in order to tackle Marketplace rights management needs, with main focus on **end-user licenses**.

¹⁵² See D3.1, Data sharing infrastructure, neural content description, rights management and monetisation v1, 9.2.2024, available at: <https://xreco.eu/wp-content/uploads/2024/02/XReco-Delivarable-3.1.pdf>, (hereinafter: D3.1), pp. 97 ff.

¹⁵³ See D3.1, p. 97.

¹⁵⁴ See also D3.1, p. 97.



8.1.2 XReco CC+ licensing scheme

8.1.2.1 Introduction

The determination of XReco licensing scheme aimed at both enabling monetization and incentivizing sharing of 3D assets¹⁵⁵. In that regard, it has been found important to build on **existing best practices** in the field of standardised, trustworthy, public licenses, which both incentivize the sharing and enable the commercial exploitation and monetization of 3D assets and that will be compatible with the smart legal contract (SLC)-enabled marketing of those assets (see Section 8.2). In that regard, Creative Commons (CC) licenses have been found to be an appropriate choice¹⁵⁶.

8.1.2.2 The choice of Creative Commons licenses

The Creative Commons suite of licenses (hereinafter: CC licenses)¹⁵⁷ is a public, standardised set of copyright licenses for use and exploitation of works (and other protectable subject matter), that are used in large scale with regards various types of creative content (e.g. texts, photographs, music and others) in several jurisdictions¹⁵⁸. It is important to note that CC licenses are public licenses, in the sense that the Licensor (i.e. the individual(s) or entity(ies) that own copyright) agrees to authorize the reproduction and sharing (i.e. provide material to the public by any means or process that requires permission, such as reproduction, public display, public performance, distribution, dissemination, communication and making available) of protected licensed material (i.e. artistic or literary work, database, or other material), in whole or in part, by any user who is recipient of the licensed material. All CC licenses are royalty-free, non-sublicensable, non-exclusive, irrevocable (2(a)(1)A of CC licenses) and impose the to the licensee the obligation attribution of the creator/Licensor (BY)¹⁵⁹. Additional conditions of non-commercial use (NC), no derivative works (ND) and the requirement to share any derivative work under the exact license terms (SA condition) may be imposed. The combination of those licensing elements results in 6 standardised CC licenses types under the current version of Creative Commons 4.0 International license¹⁶⁰:

1. [CC-BY \(Attribution 4.0 International\)](#)
2. [CC-BY-SA \(Attribution-ShareAlike 4.0 International\)](#)
3. [CC-BY-NC \(Attribution-NonCommercial 4.0 International\)](#)
4. [CC-BY-ND \(Attribution-NoDerivatives 4.0 International\)](#)
5. [CC-BY-NC-SA \(Attribution-NonCommercial-ShareAlike 4.0 International\)](#)
6. [CC-BY-NC-ND \(Attribution-NonCommercial-NoDerivatives 4.0 International\)](#)

The application of a CC license over the content is recognizable in human readable and machine-readable form.

¹⁵⁵ DoA, Objective 2, p. 3: “develop a system based on holistic trust among its participants, providing a secure data sharing ecosystem that overcomes legal and commercialisation difficulties”; DoA, p. 5: “clear incentives for the media industry to be engaged in *sharing* media content”. Cf. DoA, p. 34-34.

¹⁵⁶ Thomas Margoni & Diane Peters, Creative Commons Licenses: Empowering Open Access (March 10, 2016). Available at SSRN: <https://ssrn.com/abstract=2746044>; Thomas Margoni, CCPlusDesign.eu Or How to Apply Creative Commons Licences to 3D Printed Products in the Light of the Most Recent Developments of the European Court of Justice Case Law (May 27, 2015). Available at SSRN: <https://ssrn.com/abstract=2611152>.

¹⁵⁷ See generally, <https://creativecommons.org/licenses/>.

¹⁵⁸ For an overview, see <https://creativecommons.org/share-your-work/cclicenses/>.

¹⁵⁹ See Margoni (n1); Margoni and Peters (n1). See also [About CC Licenses - Creative Commons](#).

¹⁶⁰ *Ibid*.



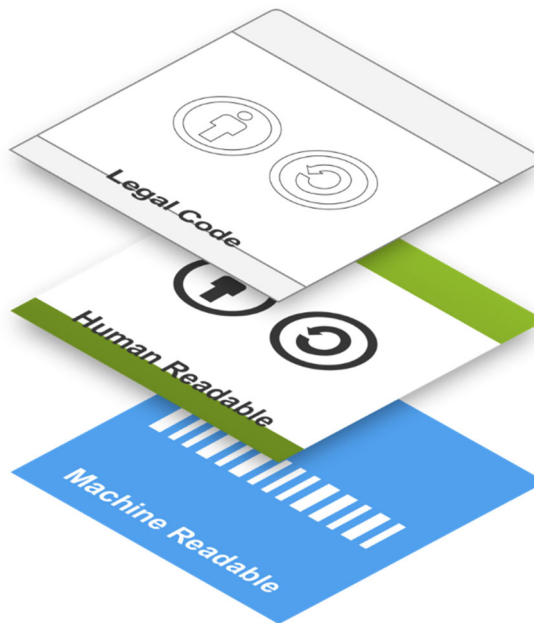


Figure 17. Anatomy of a CC license (https://wikieducator.org/Creative_Commons_unplugged/Anatomy_of_a_CC_license).

Application of a CC license over content is entirely voluntary. It is a discretionary choice of the Licensor that refers to both the choice of CC licensing as such, as well as the choice of the particular CC license type. By means of CC license application, every recipient of the Licensed Material automatically receives an offer from the Licensor to exercise the Licensed Rights under the terms and conditions of the applied CC license¹⁶¹ and the acceptance of the offered license is expressed by the licensee upon exercising the licensed rights¹⁶². It is at that point that a CC license agreement is concluded between the Licensor and this particular end-user/licensee.

8.1.2.3 Addressing the need for commercial licensing: the choice of CC+ protocol

In view of achieving the development of a data sharing ecosystem that offers a viable business model based on commercial exchange and monetization of 3D assets in the Marketplace¹⁶³, free CC licenses need to be combined with a **commercial licensing scheme** that enables monetization of the licensed work from a Licensor's perspective, i.e. allowing Licensor/uploader of the 3D asset to *receive remuneration* from the marketing of the asset through XReco marketplace and, in particular, from granting end-user licenses to potential buyers/end-users.

In that regard, an appropriate option would be to exploit **the CC-Plus or CC+ protocol**¹⁶⁴, in addition to the mere CC licensing solution, as described above. CC-Plus is a protocol allowing for the combination of a standardized Creative Commons license (in **unmodified** form) with an additional and independent (licensing) agreement granting more permissions of use/exploitation of the licensed material¹⁶⁵. The "Plus" implies the baseline permissions of the selected CC license are granted, plus more¹⁶⁶. However, the "Plus" component

¹⁶¹ See Section 2(a)(5)A of CC licenses.

¹⁶² See Preamble of CC licenses: "By exercising the Licensed Rights (defined below), You accept and agree to be bound by the terms and conditions of this Creative Commons [license] [...]".

¹⁶³ See e.g. DoA, p. 25.

¹⁶⁴ See generally <https://wiki.creativecommons.org/wiki/CCPlus>.

¹⁶⁵ *Ibid.*

¹⁶⁶ *Ibid.*



could not contain terms that contradict or modify the terms contained in the selected CC license. Accordingly, an harmonic and not conflicting coexistence of the two components should be safeguarded. In sum, the CC-Plus protocol is **not a license in itself, but it technically facilitates the acquisition of additional permissions**, on top of those already granted by the selected CC license¹⁶⁷.

Under the CC+ protocol, XReco licensing scheme would be composed by **two separate components**: a baseline permission granted under a type of CC license and a separate, **additional set** of licensing terms, which would cover **uses that are not licensed under the CC license**. The possibility of offering such additional agreement is based on the combination of Section 6(c) and Section 7(b) of CC licenses, according to which “the Licensor may also offer the Licensed Material under separate terms or conditions” and “[a]ny arrangements, understandings, or agreements regarding the Licensed Material not stated herein are separate from and independent of the terms and conditions of this Public License.”

8.1.2.4 Description of XReco CC+ licensing scheme

Given the need to implement a licensing scheme that addresses monetization of 3D assets, CC+ licensing scheme could build on a dividing line between noncommercial and commercial uses of the 3D asset by the potential end-user: Noncommercial uses of the 3D asset would be authorized under a CC license carrying the NC parameter (i.e. **CC-BY-NC, CC-BY-NC-ND, CC-BY-NC-SA**), which prohibits commercial uses of the marketed asset. Then, authorization of commercial uses could be “unlocked” through the offer (and “purchase”) of an **additional license with monetary consideration**. This additional license would then enable Licensors to **monetize** the commercial use of the asset by end-users.

As per the above choice, XReco licensing scheme would be comprised of the below separate licensing components:

- a) the **“Non-commercial CC license” component**, which needs to be comprised of a standard non-commercial CC-license, allowing non-commercial uses of 3D asset for free, and
- b) the **“Plus component”**, which will contain the licensing terms for allowing commercial use of the 3D asset against remuneration.

Under the CC+ protocol, Component 1 reflects a “sharing economy” approach, whereas the Component 2 reflects the “commercial economy” approach. Under that scheme, Component 1 CC license may not be modified or altered by means of Component 2. Accordingly, Component 2 is complementary to Component 1 and may not be in conflict with Component no 1. This means that, for instance, if Component 1 corresponds to CC-BY-NC license, Component 2 could not prohibit the creation of derivative works for non-commercial purposes.

In addition, in case of adoption of CC-Plus Protocol, there are some implementation instructions that may be followed.¹⁶⁸ Moreover, under CC+ scheme, the Component 1 is an indispensable component and common starting point for any design, meaning that, by default any marketed asset with monetization intention will necessarily be subject to the terms of the non-commercial CC license.

It becomes obvious that the definition of noncommercial uses under CC licenses becomes the critical element that defines the contours and interplay of each licensing component. As per Section 1 of relevant CC licenses

NonCommercial means [use, made by the licensee] not primarily intended for or directed towards commercial advantage or monetary compensation. For purposes of this Public

¹⁶⁷ *Ibid.*

¹⁶⁸ On the implementation of CC+ protocol in practical/technical terms, see generally <https://wiki.creativecommons.org/wiki/CCPlus>.



License, the exchange of the Licensed Material for other material subject to Copyright and Similar Rights by digital file-sharing or similar means is NonCommercial provided there is no payment of monetary compensation in connection with the exchange.

The following section offers a more detailed description of two components of XReco CC+ licensing scheme.

8.1.2.4.1 *Component 1: CC licenses for non-commercial uses*

Component 1 of the XReco CC+ licensing scheme would be comprised of one out of three available CC licenses that contain an NC element, namely:

1. CC-BY-NC

2. CC-BY-NC-ND, or

3. CC-BY-NC-SA

By setting this default licensing option as Component 1, marketing of a 3D asset via XReco Marketplace would entail, as a minimum, the following:

i) Free and non-monetized non-commercial use

By means of a non-Commercial CC license the Licensor would irrevocably offer a license and, once the license agreed, the licensee (end-user) would obtain accordingly a license that authorizes, **without remuneration** (royalty-free), the non-commercial reproduction and sharing (i.e. provision of the asset to the public by any means or process that requires permission under the Licensed Rights, such as reproduction, public display, public performance, distribution, dissemination, communication, or importation, and to make material available to the public including in ways that members of the public may access the material from a place and at a time individually chosen by them – Section 1(i) of CC-BY-NC, Section 1(k) of CC-BY-NC-SA and Section 1(h) of CC-BY-NC-ND) of the marketed 3D asset, as well as the generation and reproduction (**but not sharing**) of protected material that is derived from or based upon the marketed 3D asset, in a non-exclusive way, in all media and formats (including permission to make necessary technical modifications to this end-Section 2(a)(4) of CC licenses), without temporal (i.e. for the whole duration of legal protection- Section 6(a) of CC licenses) or territorial limitations (worldwide). The above could be equally be undertaken by every downstream recipient of the marketed asset, i.e. not only every user of the XReco marketplace that accesses the offered asset, but also any subsequent user that accesses the 3D asset when this is used/exploited by the licensee (end-user) under the terms of the CC license (Section 2(a)(5)(A) of CC licenses) ¹⁶⁹.

ii) Attribution

The licensee (end-user and any downstream recipient) will be bound under any type of non-commercial CC license to the **attribution obligation** (Section 3(a) CC licenses) which imposes:

A) the retention of the following elements with the marketed 3D asset:

- i. identification of the creator(s) of the Licensed Material and any others designated to receive attribution, in any reasonable manner requested by the Licensor (including by pseudonym if designated);
- ii. a copyright notice;

¹⁶⁹ See also Kacper Szkalej & Martin Senftleben, Generative AI and Creative Commons Licences The Application of Share Alike Obligations to Trained Models, Curated Datasets and AI Output, 15 (2024) JIPITEC 315, para 5.



- iii. a notice that refers to this Public License;
- iv. a notice that refers to the disclaimer of warranties;
- v. a URI or hyperlink to the Licensed Material to the extent reasonably practicable;

B) The indication of (technical) modification of 3D asset and the retention an indication of any previous modifications.

C) The indication that the 3D asset is licensed under the CC license, and include the text of, or the URI or hyperlink to, this License.

The retention obligation under A) is effective insofar these elements are supplied by the Licensor¹⁷⁰. Accordingly, it seems appropriate that the XReco Marketplace platform is designed in a way that the Licensor would be directed in providing attribution details, including Title of work, Creator of work, Link to work, Link to Creator Profile, Year of creation¹⁷¹.

iii) Sharing 3D asset derivatives depending on Non-commercial CC license type

According to Section 1(a) of CC licenses: “Adapted Material means material subject to Copyright and Similar Rights that is derived from or based upon the Licensed Material and in which the Licensed Material is translated, altered, arranged, transformed, or otherwise modified in a manner requiring permission under the Copyright and Similar Rights held by the Licensor. For purposes of this Public License, where the Licensed Material is a musical work, performance, or sound recording, Adapted Material is always produced where the Licensed Material is synched in timed relation with a moving image.”

In the context of 3D assets marketed via the XReco Marketplace, an adapted material could possibly stem from the downstream (end) use of the marketed 3D asset made by the buyer/end-user and would correspond to a derivative asset from such 3D asset created by the end-user, following the “purchase” of the original asset from the XReco Marketplace. This would correspond typically to a copyrighted transformed version of the 3D asset itself or to a broader creation or production of the end-user, where the 3D asset is integrated in a modified form as an element (e.g. an immersive production).

Even though the creation of an adapted material/derivative asset is generally allowed under CC licenses applicable to the marketed 3D asset, the end-user/licensee will not necessarily have the right to **share** this derivative asset¹⁷². Indeed, this depends on the selected Non-commercial CC license as Component 1. More specifically:

- Under **CC-BY-NC**, sharing of derivatives / adapted material from the marketed 3D asset is allowed but only for non-commercial purposes (as defined under the CC licenses) (Section 2(a)(1)(B)). The part of the derivative that corresponds to the initially marketed 3D asset will remain licensed under the initial CC-BY-NC license, while the protectable contribution of the licensee / creator of the adapted material will be subject to a separate license, whose terms would be basically **defined by the licensee** (“adapter’s

¹⁷⁰ Cf. <https://creativecommons.org/chooser/>: Attribution details (optional).

¹⁷¹ See for these details under <https://creativecommons.org/chooser/>: Attribution details. Cf. Recommended practices for attribution, https://wiki.creativecommons.org/wiki/Recommended_practices_for_attribution, where a recommended attribution contains Title, Author, Source, and License (TASL).

¹⁷² It is reminded that under CC licenses, sharing means provide the derivative asset to the public by any means or process that requires permission under the Licensed Rights, such as reproduction, public display, public performance, distribution, dissemination, communication, or importation, and to make material available to the public including in ways that members of the public may access the material from a place and at a time individually chosen by them.



license”, Section 1(b) of CC-BY-NC¹⁷³). Accordingly, derivative works (adapted material) based on licensed-marketed 3D asset need not to be bound by the same CC-BY-NC Licence terms. This means that the adapted material may be subject to a conventional (i.e. not an open) “proprietary” license determined by the licensee. Nonetheless, the adapter’s license in this case should still **not allow for commercial use** of the derivative (Section 2(a)(1)(B) of CC-BY-NC license), but could, however, contain an obligation of remuneration for authorizing those non-commercial uses of the adapted material. In addition, the adapter’s license could not prevent recipients of the adapted material from complying with the CC-BY-NC license (including the attribution obligation) of the original 3D asset (licensed under CC-BY-NC).

- Under **CC-BY-NC-SA**, sharing of derivatives / adapted material from the marketed 3D asset is again allowed only for non-commercial purposes (as defined under the CC licenses) (Section 2(a)(1)(B)). The part of the derivative that corresponds to the initially marketed 3D asset will remain licensed under the initial CC-BY-SA license, while the protectable contribution of the licensee / creator of the adapted material will have to be subject to a separate license, which, however, needs to be a CC-BY-NC-SA Compatible License, i.e. a (free) license listed at creativecommons.org/licenses/by-nc-sa/, approved by Creative Commons as essentially the equivalent of CC-BY-NC-SA license (Share alike condition- Section 3(b) of CC-BY-NC-SA license). Accordingly, the adapter’s license could not neither allow commercial use of the derivative asset, nor provide for onerous authorization of non-commercial uses of the derivative.
- Under **CC-BY-NC-ND**, sharing of derivatives / adapted material from the marketed 3D asset is forbidden for any purpose (Section 3(a) of CC-BY-NC-ND license, *in fine*: “For the avoidance of doubt, You do not have permission under this Public License to Share Adapted Material”). Accordingly, sharing of derivatives would remain exclusively reserved by the Licensor of the marketed 3D asset.

iv) The case of a sole predefined type of Non Commercial CC license as Component 1

For the sake of simplicity and in order to increase standardization, it could be possible to provide only one type of non-commercial Creative Commons licence **as sole option** for Component 1, defined by XReco Marketplace provider, with no possibility for the Licensor of choosing different non-commercial CC licence (or any other license). In that regard, all non-commercial CC licenses are fit for the implementation of a CC+ protocol licensing scheme, as described above. The selection of a sole non-commercial CC license is, of course, predominantly a business model decision. As per current XReco working model, the CC-BY-NC license is selected as predefined Component 1.

From a legal perspective, this approach should take into account the effects from the interaction between each non-Commercial CC license and the “Plus” component. More precisely, in case that CC-BY-NC-ND license is selected as Component 1, the “Plus” component could cover and set licensing terms for the sharing of 3D assets derivatives, even for non-commercial purposes. This means that sharing of 3D assets derivatives even for non-commercial purposes could also be monetized by the Licensor. In case that Creative Commons Attribution-Non commercial license (CC-BY-NC) is selected as Component 1, then the end-users will be able to share 3D assets derivatives for non-commercial purposes not only under non-commercial CC licenses but even under paid licenses determined by themselves. On the contrary, the selection of a CC-BY-NC-SA license would dismiss the above issues and, therefore, seems an appropriate option as predefined value of Component 1.

Adapter’s license over 3D assets derivatives (end-user selection)

¹⁷³ Section 1(b) of CC-BY-NC: “Adapter’s License means the license You apply to Your Copyright and Similar Rights in Your contributions to Adapted Material in accordance with the terms and conditions of this Public License.”



Component 1 (marketed 3D asset – BY Licensor's XReco selection)	CC license	BY-NC	BY-NC-ND	BY-NC-SA	BY-ND	BY-SA	Custom license
BY-NC							*Allowing Non-commercial uses only
BY-NC-ND							
BY-NC-SA							

Figure 18: Permissible adapter's license under Non-commercial CC Licenses over 3D assets.

8.1.2.4.2 Component 2: "Plus" component for Commercial uses

Given the non-commercial nature of CC licenses applied as per Component 1, the Licensor **still retains** the exclusive power over the exercise of all licensed rights for commercial purposes and has the ability to offer the licensed 3D asset under **separate (non-conflicting with Component 1) licensing terms or conditions** that refer to rights not covered by the CC license (Section 6(c) CC licenses)¹⁷⁴.

The possible scope of this separate licensing agreement will be mainly defined, *e contrario*, from the definition of non-commercial under CC licenses. Accordingly, Component 2, namely the "Plus" Component, may constitute an optional licensing scheme that would cover the **commercial use** of the licensed 3D asset and the (commercial¹⁷⁵) use (sharing) of 3D asset's derivatives, i.e. **any public use (reproduction and sharing) of the 3D asset or 3D asset's derivatives [made by the licensee or the user of 3D asset derivatives] primarily intended for or directed towards commercial advantage or monetary compensation**. This would be, for example, the case of trading the 3D asset or using the 3D asset as part of a commercial production or application, i.e. production or application which is primarily intended for or directed towards commercial advantage for the producer/developer-licensee.

Apart from the above limitation of scope (commercial use), the exact content of the terms of the Plus component may, in principle, be discretionally defined by the holder of rights over the marketed 3D asset. In that regard, freedom of contract remains the default rule, with the reserve of possible limits set by national law provisions on copyright contracts¹⁷⁶. However, in view of satisfying XReco objective for a secure data sharing ecosystem that **overcomes legal and commercialization difficulties**, rights management for commercial uses would preferably be organized under **standardized terms** that allow monetization of the asset, i.e. terms for a licensing agreement with monetary consideration (paid license)¹⁷⁷.

The XReco Marketplace could then provide for standard terms regarding the "Plus" component. The main content and architecture of these end-user licenses could be determined by means of calibration of the main licensing building blocks, as described in Annex VII of D3.1¹⁷⁸. In addition, their content could be fixed/predefined by the XReco platform provider. It could also be possible to leave certain variable licensing

¹⁷⁴ See also e.g. https://wiki.creativecommons.org/wiki/NonCommercial_interpretation#cite_note-1: "NonCommercial licenses are non-exclusive. Like all CC licenses, the NC licenses are non-exclusive. This means that an NC Licensor is free to offer the material under other terms, including on commercial terms. A frequently discussed use case for the NC licenses is a creator who wishes to allow NonCommercial use but also authorizes commercial uses in exchange for payment. (Additional permissions such as this may always be offered; Licensors may also use our CC+ protocol to offer these in a standardized manner.)".

¹⁷⁵ It would be possible that the Plus component sets the licensing terms for the sharing of 3D assets derivatives, even for non-commercial purposes, in case that CC-BY-NC-ND license is selected as Component 1 (see also above, under 8.1.2.4.(a)(iv)).

¹⁷⁶ See on that issue D3.1 pp. 98-99.

¹⁷⁷ DoA, Objective 2, p. 3.

¹⁷⁸ D3.1, pp. 172 ff.



components, in order to allow XReco Marketplace Licensors to customize their exact content, by means of activating/deactivating or by defining their value.

As per current XReco working model, a fixed-components approach is largely adopted, in a way that “Plus” licensing terms are generally fixed and the Licensor may only define the amount of the remuneration. The main outline of the “Plus” component licensing elements is the following:

1. Authorised uses: sharing the 3D asset as per CC licenses definition (see above).
2. Purpose of authorised uses: commercial use, namely use primarily intended for or directed towards commercial advantage or monetary compensation [by the licensee] (*e contrario* definition from Section 1 of non commercial CC licenses).
3. Duration of authorized use: unlimited in time.
4. Territory of authorized use: worldwide.
5. Remuneration of Licensor [payment clause]: custom (variable) appropriate and proportionate lump sum, **to be defined by the user**.
6. Derivative works: authorizing creation of adapted material, i.e. material that is derived from or based upon the Licensed 3D asset for commercial purposes.
7. Sublicensable: offering possibility for licensee to grant further licenses for use / exploitation of the asset to third parties.
8. Exclusivity in favor of licensee: Yes.
9. Attribution of the Author (and the Licensor, if distinct): Yes.
10. No obligation of exploitation of 3D asset imposed to the licensee.

According to the above, the “Plus” component would correspond to a **worldwide, exclusive, perpetual, sublicensable, paid licence to use and exploit the 3D asset (including to reproduce, distribute, modify, display and perform it), without Licensor’s attribution, for commercial purposes**, that would be applied on top of a Creative Commons license that covers non-commercial uses of the same asset.

From a technical point of view and in order to show consistency with the downstream effects of Creative Commons licenses, in case that Plus Component is enabled for the marketing of an asset, the copy of the asset which is subject to that component should be distinct from the copy that is subject to the Component 1 (Non commercial CC license). Decision tree

Based on the previous analysis, a decision tree for implementing the **XReco CC+ licensing scheme** would look as follows:



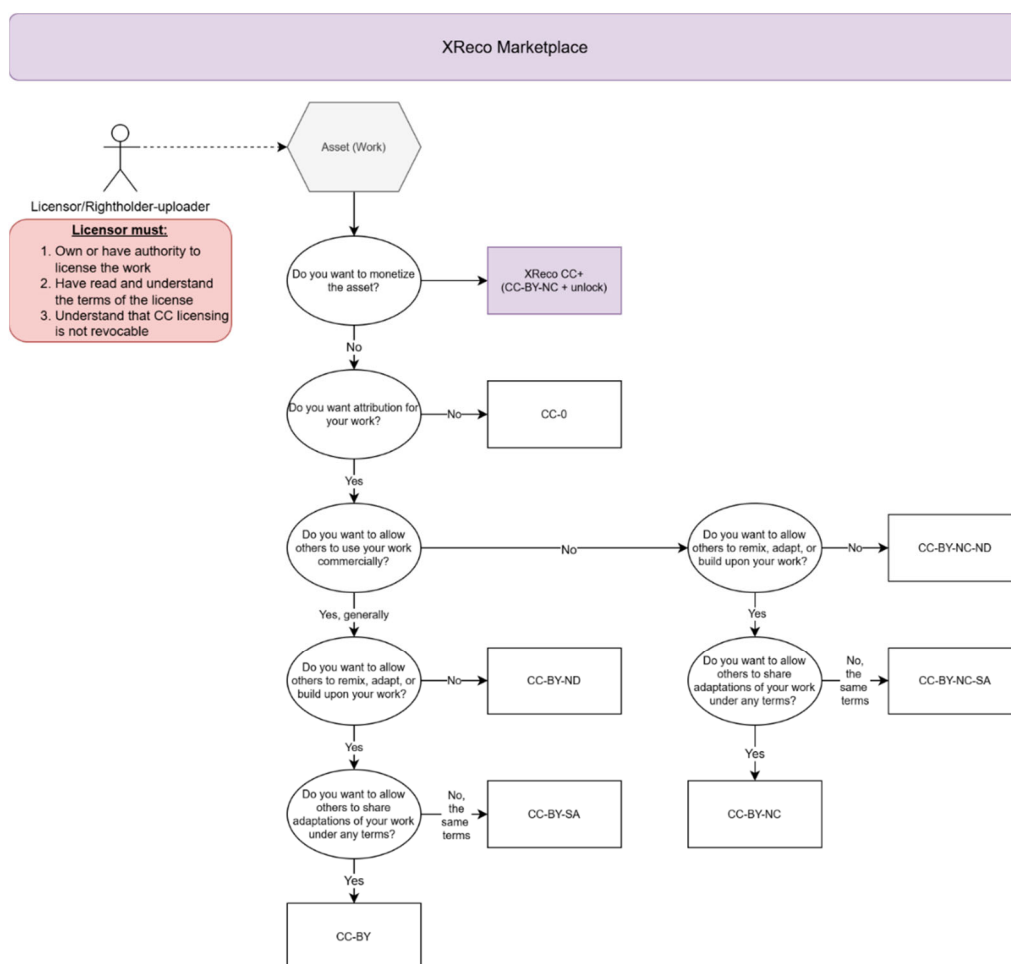


Figure 19: XReco CC+ licensing scheme decision tree.

8.1.3 Analysis: Opportunities and Risks

8.1.3.1 Opportunities

i) Standardized multifaced and open-ended licensing solution

The option of CC licenses seems to be an adequate licensing scheme for the XReco Marketplace end-user license. First, CC licenses are standardised, trustworthy, public licenses with extended implementation in various jurisdictions around the world and are designed to be compatible with different applicable laws, including EU law. Due to their widespread use for more than two decades, CC licenses are amongst the most popular free licenses for works other than software. This multiplies the chances that both content owners and potential users will be familiar with this type of licensing. At the same time, CC licensing would certainly underpin sharing, reuse and, under some types of CC licenses, creative reuse of 3D assets. CC licensing would render XReco platform a 3D content sharing platform, enabling interested stakeholders to be engaged in sharing media content under standard and, to some extent, familiar terms¹⁷⁹, eliminating the difficulties associated with drafting bespoke private licensing agreements and reducing search&transaction costs for both rightholders and users¹⁸⁰.

¹⁷⁹ Cf. DoA, KPI 2.1.

¹⁸⁰ Guido Russi, Creative Commons. CC-Plus, and Hybrid Intermediaries: A Stakeholder's Perspective, 7 BYU Int'l L. & Mgmt. R. 102 (2011). Available at: <https://digitalcommons.law.byu.edu/ilmr/vol7/iss2/5> p. 117.



At the same time, the benefits of standardized licensing terms would equally apply in case of a standardized “Plus” component, with fixed licensing components, such as the one described above. Moreover, the “Plus” component remains flexible enough in order to accommodate various degrees of standardization and calibration, to be shaped also under the influence of business model decisions, including the addition of further licensing components, such as e.g. a specific variable regarding the commercial use of 3D assets for AI training purposes¹⁸¹.

ii) Applicability of CC licenses on 3D content

CC licenses are sufficiently flexible to cover 3D content as licensing material. According to CC licenses’ definition, “Licensed Material means the artistic or literary work, database, or other material to which the Licensor applied this Public License”. Under this broad definition, any artistic work or material may be subject-matter of a CC license. If seen in combination with the CC licenses definition of “Licensed Rights”¹⁸², licensed material under CC licenses needs to be material that is protected under Copyright and Similar [related] rights.

3D assets in general may be eligible subject matter of protection under EU copyright law. Indeed, the EU copyright law does not provide for a *numerus clausus* of protectable subject matter, nor holds any exclusion of artistic or creative domain from its scope¹⁸³. Digital 3D content may be classified as form of digital visual art. Visual works are indirectly recognized as protectable subject matter under Directive 2012/28/EU of the European Parliament and of the Council of 25 October 2012 on certain permitted uses of orphan works¹⁸⁴ (see Annex, no 3). This conclusion is also supported by the application of Article 2 (1) of the Berne Convention for the Protection of Literary and Artistic Works (Paris Act of 24 July 1971), as amended on 28 September 1979 and also ratified by the EU (The expression “literary and artistic works” shall include every production in the literary, scientific and artistic domain, whatever may be the mode or form of its expression). As a consequence, 3D models could be in principle qualified as copyrighted works under EU law and, thus, they can validly be licensed under CC licenses as licensed material.

3D assets are eligible protectable subject matter under EU law as works, insofar the conditions of protection are satisfied, namely: a) they are original creations (expression of free and creative choices of the author that reflect his personality¹⁸⁵), even if their realization is dictated by technical considerations¹⁸⁶, that do not prevent the author from reflecting his personality, and b) whose existence is identifiable with sufficient precision and objectivity¹⁸⁷. The second condition is easily satisfied to the extent that 3D assets are perceivable digital artefacts by human sight¹⁸⁸. The first condition raises concerns of generic nature, that apply in all form of works and needs to be assessed on a case-by-case basis. It raises, however, particular questions in case of 3D output produced via AI-based systems¹⁸⁹. In addition, a particular concern could be

¹⁸¹ See responses collected during evaluation in D5.2.

¹⁸² Section 1 CC licenses: “Licensed Rights means the rights granted to You subject to the terms and conditions of this Public License, which are limited to all Copyright and Similar Rights that apply to Your use of the Licensed Material and that the Licensor has authority to license.”

¹⁸³ Cf. rec. 10 Infosoc Directive (creative and artistic work of authors).

¹⁸⁴ OJ L 299, 27.10.2012, p. 5–12.

¹⁸⁵ See on that topic D3.1, Annex VI, Section 15.2.4.

¹⁸⁶ ECJ, Judgment of the Court (Fifth Chamber) of 11 June 2020, *SI and Brompton Bicycle Ltd v Chedech / Get2Get*, Case C-833/18 (Brompton), para. 26.

¹⁸⁷ Brompton, para. 25.

¹⁸⁸ Cf. the distinction between “more objective mechanical senses” (sight, hearing) and more subjective chemical senses (taste, smell) in J. McCutcheon, *The Concept of the Copyright Work under EU Law: More Than a Matter of Taste* (July 30, 2019). Forthcoming, 2020, *European Law Review*, Available at SSRN: <https://ssrn.com/abstract=3428844>, p. 19.

¹⁸⁹ D3.1, p. 94. See also below 8.1.3.2.(d)(ii).



related with the impact of technical considerations that are related with the production of a 3D output by means of the implementation of XReco 3D reconstruction techniques.

iii) Filling the monetization gap

One potential benefit of the use of an additional licensing scheme based on the CC+ protocol is that the sharing economy promoted by public CC licenses coexists and can be integrated with commercial economy¹⁹⁰. The use of the CC-Plus framework would have positive impacts on all actors within the XReco ecosystem. It has moreover been argued that, by allowing users to “sample” the works for free for certain uses, the CC-Plus model may increase users’ willingness to pay for additional rights/uses and lead to potentially higher revenues for creators¹⁹¹. Beyond the above practical and economic benefits, the free uses allowed by CC licenses would also promote participation and engagement in sharing 3D assets, crucial in building a thriving ecosystem within XReco¹⁹².

The implementation of CC-Plus framework within the context of the XReco Marketplace seems a viable and attractive choice, to the extent that it could retain the benefits of standardization offered by the CC suite of licenses, while allowing Licensors to monetize (commercial) uses of their assets, on the basis of an additional and independent agreement (the ‘Plus Component’). Several examples of CC+ implementation demonstrate the viability of such an approach¹⁹³.

8.1.3.2 Risks

a) CC licenses adequacy (Component 1) as Marketplace license

CC licenses are in principle an adequate licensing mechanism for authorizing also the upstream use of the 3D asset, i.e. the acts that are necessary for lawfully placing the protected 3D content in the XReco Marketplace. In other words, the applied CC license as end-user license could also play the role of Marketplace license (see above 8.1.1). In fact, the scope of licensed rights (reproduction and sharing) typically covers upload and display of the marketed asset in the XReco Marketplace, which both require permission under EU copyright law (since they are covered by reproduction right and communication to the public right). Accordingly, by means of the application of a CC license by the Licensor, the XReco Marketplace provider would also be a “recipient” of the licensed 3D asset as per section 2.a.5.A section of CC licenses. Then, CC license would be effective starting from the exercise of licensed rights by the XReco Marketplace provider, i.e. by the storage of the asset for display purposes.

However, noncommercial types of CC licenses may not be an adequate licensing scheme for a Marketplace license (marketplace upload and display), in case that upload and display at the XReco Marketplace would qualify as commercial use according to the relevant CC licenses definition, i.e. a use that is primarily intended for or directed towards commercial advantage or monetary compensation (for the user, here: the XReco Marketplace provider). This will depend on the business model adopted and it derive from e.g. a paid service for merely enabling the upload and display of the 3D asset in the Marketplace or an ad-based service primarily intended for or directed towards commercial advantage (for the XReco Marketplace provider).

¹⁹⁰ See CC and CC+ Overview for the World Wide Web, CREATIVE COMMONS, http://wiki.creativecommons.org/images/3/37/Creativecommons-ccplus-overview-for-the-worldwide-web_eng.pdf.

¹⁹¹ *Ibid.* p. 118; Martin Peitz & Patrick Waelbroeck, Why the Music Industry May Gain from Free Downloading—The Role of Sampling, 24 INT’L J. INDUS. ORG. 907, 912 (2006)

¹⁹² Michael Jonathan Todosichuk, Understanding Musical Artists’ Motivation to Share Creative Commons Licensed Musical Works: Applying Social Capital and Social Cognitive Theory (June 2009).

¹⁹³ See e.g. examples referenced under <https://wiki.creativecommons.org/wiki/CCPlus#Examples>.



In those cases, it would be necessary that the Marketplace license takes the form of an agreement separate to Component 1 CC license, concluded between the 3D asset rightholder and the XReco Marketplace provider (typically the Terms of (Marketplace) Service), that would grant, among others, the necessary rights for lawfully performing acts of reproduction and communication to the public. Nonetheless, such separate agreement seems preferable, as it could also contain an appropriate clause regarding warranty of title and authority to license, given by the Licensor over the licensed rights. This could tackle the risk related with the absence of warranty, and especially “warranties of title”, and the limitation of liability from Licensor’s side under CC licenses (Section 5 of CC licenses). However, such clause would not be enforceable by the end-user against his Licensor, unless if such clause is explicitly agreed between those parties as a separate clause to the end-user CC license (as per Section 7(b) of CC licenses).

b) CC licenses (Component 1) and SLC

A note can be made with regards the compatibility of the CC license with the SLC mechanism¹⁹⁴ and, in particular, with regards the time of the conclusion of the CC license agreement between the Licensor (uploader of the asset at the marketplace) and the end-user. In particular, as per CC licenses terms, the offer of the license is directed towards any user of the XReco marketplace who receives (i.e. accesses) the displayed 3D content (Section 2.a.1.5. CC licenses) and the conclusion of the agreement will take place by the moment the licensee *exercises* the licensed rights (typically: proceeds with acts of reproduction or sharing of licensed 3D asset) (see preamble of CC licenses). On the contrary, the SLC featuring the CC license terms is concluded prior to this exercise, by means of acceptance of the transaction by the user, which is further notarized in the Blockchain thanks to a XReco Receipt, as detailed in Section 8.4. This shift would amount to a different term regarding the conclusion of the agreement, whose treatment should be based on Section 7(a) of the CC licenses. According to this term, “the Licensor shall not be bound by any additional or different terms or conditions communicated by [the end-user] unless expressly agreed”. The acceptance of transaction, by which the end-user agrees to be bound by the terms of the CC license could be seen as an “additional or different term or condition communicated by the [end-user]”, which should be expressly agreed by the Licensor. Given that the Licensor is not in position to expressly agree with this term at the moment of the transaction acceptance, this agreement should be provided at an upstream level, when agreeing for the upload of the asset to the marketplace¹⁹⁵.

c) Other inherent limitations

The XReco licensing scheme faces some inherent limitations. As per suggested organisation of licensing components, **it is not possible for Licensors to monetize non-commercial uses** of the licensed 3D asset. Indeed, all CC licenses do not allow for monetization, whereas in case of the CC+ licensing scheme, only commercial uses are subject to monetization, as per the terms of the “Plus” component. On the contrary, as seen above, it remains possible for the Licensor to monetize the noncommercial use of 3D assets derivatives, in case that a CC-BY-NC-ND license is selected as Component 1.

In case that the Licensor has selected a self-standing noncommercial CC license (with no use of the CC+ licensing scheme), then it remains possible for the Licensor to authorize (and monetize) commercial uses of the 3D asset *outside* XReco Marketplace. This could be detrimental for XReco platform to a variable extent, depending on the business model adopted (e.g. XReco commission over onerous 3D asset transactions that take place via the Marketplace). A solution for that would be the elimination of non-commercial CC licenses

¹⁹⁴ See on that issue also D3.1, Annex IX, pp. 183 ff.

¹⁹⁵ This situation does not amount to “waiver” of a term or condition and, therefore, Section 8(c) should not apply (Section 8(c) of the CC licenses: No term or condition of this Public License will be waived and no failure to comply consented to unless expressly agreed to by the Licensor.)



from the list of available CC licenses under the non-monetizing licensing scheme, but this could render the decision tree for license selection more complex.

Last, some limitations are related with the contours of some CC licenses definitions and especially the definition of “Noncommercial”. Indeed, the flexible definition provided under CC licenses recognizes that no activity is completely disconnected from commercial activity. Whether a use is commercial will then depend on the specifics of the situation and the intentions of the user¹⁹⁶. The delineation between the two components of the CC+ licensing scheme based on that concept may create insecurity in cases where the qualification of a use as non-commercial is not straightforward. This limitation could be possibly tackled with guidance at Marketplace interface of both Licensors and end-users, based on examples of usual commercial and non-commercial uses of the 3D asset according to the applicable definition.

d) Risks associated with the copyright status of marketed 3D assets

The risks that are associated with the copyright status of marketed 3D assets merit closer examination. In fact, the status of the marketed asset has serious impact, especially under the light of AI-generated 3D assets and especially 3D assets that are produced as an output from AI-based systems or technologies that are trained over preexisting copyrighted material, either externally or via the implementation of XReco 3D reconstruction techniques.

i) Derivative 3D assets

As per current XReco working model, marketed 3D assets **are assumed to constitute new and independent works**, i.e. at the same time a) copyrighted works that b) do not reproduce protected expression from preexisting works. New and independent works (3D assets) may correspond either to a 3D output that is generated via the implementation of XReco 3D reconstruction techniques¹⁹⁷, or to an asset that is directly uploaded to the XReco Marketplace as an external asset, without any use of XReco services. Under this assumption, the XReco Marketplace licensing scheme will be applied in a straightforward way, to the extent that the 3D asset will constitute the sole object of licensed rights and the sole licensing material. In addition, the risk of copyright infringement resulting from the marketing of the 3D asset would arise only in case that the Licensor has not the legal authority to grant the license (e.g. in case that the Licensor is not the owner of rights over the asset). However, this risk is inherent to any copyright transaction. The only difference under the XReco licensing scheme relates with the absence of warranties that characterize CC licensing (see Section 5 of CC licenses).

Analysis needs to be differentiated in case that the marketed 3D asset is not an independent work but **a derivative work**¹⁹⁸, i.e. where the 3D content derives or is based on preexisting **copyrighted** content (namely preexisting 2D content, but also possibly 3D content) and the applicable law determines that the rights held in the relevant preexisting content cover also 3D content that is based on this content. Legal challenges related with the licensing of derivative 3D assets are three-fold: a) the qualification of a 3D asset as derivative work, as per applicable law (see above, under Section 7.3.2) and, in the affirmative scenario, b) the granting of a license from the holder of rights over preexisting works and c) the eventual payment of remuneration against the granting of such license.

¹⁹⁶ Creative Commons, Frequently asked question, available at: <https://creativecommons.org/faq/#does-my-use-violate-the-noncommercial-clause-of-the-licenses>.

¹⁹⁷ See D3.1, Status no 3, pp. 93-94.

¹⁹⁸ See D3.1, Status no 1, pp. 92-93.



In a situation like this, the Licensor of the derivative 3D asset may opt for the marketing of the 3D asset without prejudice to the copyright in the original (initial) preexisting work¹⁹⁹. In other words, marketing of this 3D asset requires **rights clearance** of all rights over preexisting works. At the same time, the end-user license shall cover equally both layers of rights: those which are relevant with the derivative 3D asset as well as those pertaining to the preexisting work(s) (namely preexisting 2D content, but also possibly 3D content). If not, then both the upload in the marketplace and the further downstream use of the asset by the end-user risks to **infringe** the rights over the preexisting works and render both the marketing of the asset and its subsequent use **unlawful**. At the same time, uncertainties regarding the qualification of derivative works remain at EU level (see above, under Section 7.3.3.).

“External” derivative 3D assets

In case of “external” derivative 3D assets, i.e. assets that are directly uploaded to the XReco Marketplace by the Licensor, without any interference of XReco 3D reconstruction services, the above challenges echo the general risk connected with legal defects of the licensed rights and material uploaded at the Marketplace. Typically, it should be up to the Licensor to warrant against both the XReco Marketplace provider and the end-user that all third-party rights over the derivative 3D asset are properly cleared. This could take the form of a warranty term within the Marketplace license and/or of within the end user license, which, in the latter case, would be a separate agreement vis-à-vis the CC licenses (see above, Sec. 8.1.3.2.(a)) and an additional term for the “Plus” component. Any remuneration due to the rightholders of preexisting works would be agreed and paid on the basis of the internal relationship between those rightholders and the rightholder over the derivative 3D asset, i.e. the Licensor of marketed asset.

3D outputs produced by means of XReco 3D reconstruction services

Further focus needs to be put on cases of 3D assets that are produced by XR Services users “thanks to the platform and the ecosystem”²⁰⁰, i.e. through the implementation of XReco AI-based 3D reconstruction techniques by users themselves. As analyzed under D3.1²⁰¹, under EU law, a 3D asset that is generated by means of XR Services would be considered as derivative insofar this 3D asset contains (recognizable) elements from concrete exteriorized protectable (original) expression of the preexisting 2D/3D content that have been subject to a “dimension shifting”. The above analysis, as well as the analysis contained in D.3.1. has shown that, according to EU law, **copyrighted 2D content** (hereinafter: 2D inputs) employed as training material for the XR Services models (e.g. NeRFs) **will not routinely be reproduced** in a legal sense within 3D outputs, as long as the output stage of 3D Reconstruction employs unprotected metainformation (such as, e.g. datapoints) embodied in the 2D inputs, related with the represented object. In that regard, in a potentially-significant number of applications of 3D reconstruction techniques, the new 3D asset would contain *unrecognisable* fragments of protected 2D inputs elements. As a consequence, it may be asserted that **3D outputs produced by users, by means of XR Services, would not routinely constitute derivative works of employed 2D inputs**. Accordingly, rights clearance of rights over preexisting 2D inputs as described above would not be necessary. In this case, preexisting 2D inputs would still be **relevant** vis-à-vis 3D output, since it would form the immediate source of necessary metainformation employed for the generation of the 3D output²⁰², but not in a legal (copyright law) sense.

¹⁹⁹ Cf. already D3.1, p. 90.

²⁰⁰ DoA, p. 3.

²⁰¹ D3.1, pp. 91 ff.

²⁰² See in that regard DoA, p. 9: “There is currently a lack of regulation and policy making regarding the model itself, and the products of the model, **which can all be considered as derivatives of the shared data that were used to train it.**” [emphasis added] and *ibid.* pp. 8-9: One of the main objectives of the project is to create XR 3D media assets by combining sets of 2D and 3D content. This not only implies that a created new media asset is **a derivative work of preexisting media** [...]” [emphasis added].



Nonetheless, it cannot be entirely excluded that such copyright significant reproductions of preexisting 2D inputs may take place within 3D outputs produced by means of XReco 3D reconstruction services. Indeed, some technologies may interfere with the rights of adaptation and/or reproduction, depending on the applicable law (see above, under Section 7.3.2. Under this scenario, infringement risks persist and the underlying 2D inputs' downstream use within 3D outputs would need to be licensed, unless if the application of an exception could be invoked (see above, under Section 7.3.2).

ii) Unprotected 3D assets, including (XReco) AI generations

There is no possible way to anticipate the **absence of copyrightability of a 3D asset** from the outset. This may be of particular risk in case of 3D content which is AI-generated, either by means of XReco 3D Reconstruction techniques or by means of third-party GenAI applications²⁰³. The absence of warranties as to the legal vices of licensed rights ("warranties of title", Section 5(a) of CC licenses) renders this question more complex. In case that no human authorship is established, the output will not be an original copyrighted work²⁰⁴ and, thus, the CC license will not be applicable.²⁰⁵ The CC license may still apply in case that not-copyrighted 3D content is protected under a related/neighbouring rights regime, given that the relevant definition of CC licenses for "Copyright and Similar Rights" covers "similar rights closely related to copyright". In case that such regime is recognized in the future, the CC license would be applicable from the date of its entry into force and for the duration of its legal protection.

Apart from that, it can be asserted that the "Plus" component (as well as any other separate agreement other than CC license) could remain solely applicable, even in case that the 3D asset is not copyrightable. In the absence of protected subject matter, the Component 2 would not function as "Plus" component, since the CC license would not be applicable. Depending on applicable contract law, it could still function, however, as an *inter partes* agreement between the Licensor and the end user that would set the rules for use of an unprotected intangible object. Any violation of its terms could then amount to breach of contract, according to applicable contract law rules, but would not implicate infringement of exclusive and absolute rights of the Licensor. At the same time, this agreement would not deter unlimited free use and exploitation of the unprotected 3D asset by any third party.

8.1.4 Addendum: Upstream XReco licensing & monetization scheme for 2D inputs

8.1.4.1 Shifting focus from the output stage

As per XReco's objective, rights management and monetization schemes could possibly expand at the upstream stage and cover this use of preexisting content 2D inputs as training material for the models (e.g.

²⁰³ An additional risk is related with the possible enactment of and *ad hoc* related right over non-copyrighted AI-generated by future legislation. See e.g. the sui generis right to non-original objects generated by a computer program established under article 33 of the Ukrainian Copyright Act (Law of Ukraine No. 2811-IX of December 1, 2022, on Copyright and Related Rights). On that topic see e.g. Theodoros Chiou, Copyright ownership challenge arising from AI-generated works of art: A time to stand and stare in I. Stamatoudi, Research Handbook on Intellectual Property and Cultural Heritage, (Edward Elgar 2022), 267-268 and 270.

²⁰⁴ See in more detail, D3.1, p. 90.

²⁰⁵ See particularly Section 8(a) of the CC licenses: "For the avoidance of doubt, this Public License does not, and shall not be interpreted to, reduce, limit, restrict, or impose conditions on any use of the Licensed Material that could lawfully be made without permission under this Public License". See however <https://creativecommons.org/2023/08/18/understanding-cc-licenses-and-generative-ai/>: "If you create works using generative AI, you can still apply CC licenses to the work you create with the use of those tools and share your work in the ways that you wish. The CC license you choose will apply to the creative work that you contribute to the final product, even if the portion produced by the generative AI system itself may be uncopyrightable."



NeRFs) that will generate 3D outputs. Under those cases, measuring the **relevance**²⁰⁶ as enabler for monetization of preexisting 2D inputs remains a critical aspect, but, in this case, relevance needs to be also **copyright-significant** in the first place.

Assessing copyright-significant relevance at the output stage requires, as a first step, the legal evaluation of copyright-significant reproduction of protectable elements of original expression of preexisting 2D inputs within 3D output. This, however, corresponds to a case-by-case purely legal assessment which is dependent on applicable law, especially in the absence of a harmonized adaptation right and derivative works concept in the EU. In that regard, further research could be directed towards the possible automation of this assessment by means of technical solution that would assess both originality criterion and the reproduction of original expressions within the 3D output, possibly by means of a type of *recognizability* assessment, which, however, goes beyond the scope of the current project²⁰⁷. Even if such assessment would still be useful for tackling the marginal risks associated with derivative 3D outputs (see above under 8.1.3.2(d)(i)), it would not constitute a preferable basis for a licensing/monetization scheme, insofar it seems that 3D outputs produced by users, by means of XR Services, would not routinely constitute derivative works of employed 2D inputs.

8.1.4.2 Monetizing reproductions of 2D inputs for algorithmic training purposes

Given the above limitations and in view of tackling XReco objectives, instead on focusing on the output stage, licensing and monetization of preexisting 2D inputs in the context of marketing of 3D outputs produced by means of XReco 3D reconstruction services could focus on acts of reproduction that take place during the 3D reconstruction training stage (namely, on the “AI-generation license” as per D3.1²⁰⁸). This approach would have the merit of being implementable independently of the status of the generated 3D output (derivative or not), given that the algorithmic training stage will be subject to the scope of copyright protection, to the extent that its deployment requires **reproduction** of preexisting works²⁰⁹ (namely transient RAM copies as well as temporary storage of the ingested content in the respective server)²¹⁰ and to the extent that rightholders are expected to routinely opt-out from the TDM exception, under Art. 4(3) CDMSD. At the same time, such license could also anticipate and cover downstream uses, including, among others, any eventual reproduction of recognizable protected elements of 2D inputs within 3D outputs.

8.1.4.2.1 The concept of relevance as monetization metric

The “relevance” of preexisting works could be still used as a criterion for developing rights management and licensing approaches that enable fine-grained monetization of preexisting works (2D inputs) used as training

²⁰⁶ Cf. DoA, p. 9: “Measuring the “relevance” of the media assets is a critical aspect and it is a matter of specific research and investigation in the context of the project [...]”.

²⁰⁷ Cf. D3.1 p. 102: “Focus could be made on the technological identification of traces of expression of pre-existing works within new 3D assets as well as on defining metrics of significance of contribution, allowing machine-to-machine management and monetization of rights under the light of EU acquis.”

²⁰⁸ D3.1, p. 88.

²⁰⁹ Cf. for a similar approach Martin Senftleben, Remuneration for AI Training - A New Source of Income for Journalists? (August 04, 2024). Martin Senftleben, Kristina Irion, Tarlach McGonagle and Joost Poort (eds.), The Cambridge Handbook of Media Law and Policy in Europe, Cambridge University Press, Forthcoming, Available at SSRN: <https://ssrn.com/abstract=4963458> or <http://dx.doi.org/10.2139/ssrn.4963458>, p. 21: “Qualifying copies made for AI training purposes as relevant reproductions, the lawmaker can create a legal basis for a remuneration claim in copyright law. With regard to AI output, the copyright basis for remuneration is less clear. Instead of reproducing individual expression – protected free, creative choices by a human journalist – AI output may merely reflect unprotected news of the day, facts, concepts and styles.” For a reference to machine-readable remuneration protocols connected with TDM opt-out of art. 4(3) CDMSD and use of works for AI-training purposes, see M. Senftleben, Generative AI and Author Remuneration. *IIC* 54, 2023, 1546.

²¹⁰ See D3.1, p. 89.



data for 3D reconstruction purposes. However, instead of measuring the relevance of each 2D input with regards the expression of the 3D output, which remains largely casuistic, to some extent subjective and possibly insignificant from a legal perspective, focus should be put on the “contribution” of each training work at the upstream stage of algorithmic training and, in particular, on the **contribution** of each training work in the **provision of metainformation** (such as, e.g. datapoints of the scene- the “mere data”) that is used for 3D reconstruction purposes, i.e. for model training²¹¹.

8.1.4.2.2 Scenarios of implementation

There are at least two possible scenarios of implementation of such monetization scheme for 2D inputs algorithmic training. However, the overall question of implementation and, in particular, the actual mode of implementation, is strongly related to the governance scheme of the platform, which may vary, depending on the selected business model. Accordingly, the below scenarios and analysis may only serve as tentative/indicative examples.

- **Scenario 1. Marketing of XReco 3D output via the XReco Marketplace.** A first scenario would be the allocation of a share of the amount paid by an XReco Marketplace end-user, in case of marketing of the XReco 3D output by the XReco services user via the XReco marketplace under the CC+ licensing scheme²¹².
- **Scenario 2. Paid generation of XReco 3D output.** Another scenario would be the allocation of a share of the amount paid by the XReco platform user for using XReco 3D reconstruction services in order to generate the XReco 3D output that is based on the metainformation contributed by certain third-party 2D content that is uploaded at the XReco platform via the Orchestrator.

In both cases, such amount would be paid as remuneration for the **AI-generation license agreement** (“**XReco AI-generation license**”) **concluded** between 2D input rightholders and XReco platform provider, to the extent that licensable acts (reproduction for training purposes) are technically undertaken **by the platform itself** and not by the deployer of 3D reconstruction services (XReco user). At the same time, this remuneration should be collected by the XReco platform provider within the framework of its contractual relation with the user of its services (3D reconstruction services and/or Marketplace, e.g. the Marketplace license).

8.1.4.3 Opportunities

XReco platform could safeguard transparency by tracking the use of preexisting 2D content during the AI training by means of XR services deployment and by providing the Licensor of 2D inputs information regarding the calculation of relevance of his/her licensed 2D content in the generation of 3D output and the calculation of royalty share. In that regard, it would be recommendable that 2D content upload should be combined with the provision of appropriate metadata, such as information on the work and on the

²¹¹ Cf. Kacper & Senftleben, *op. cit.* p. 17: “If the trained model is primarily seen as a giant collection of data points and vectors, it can be assumed that it does not contain copyright-protected traces of works used for training.” Cf. also EUIPO, The Development of Generative Artificial Intelligence from a Copyright Perspective, 2025, available at: https://euipo.europa.eu/tunnel-web/secure/webdav/guest/document_library/observatory/documents/reports/2025_GenAI_from_copyright_perspective/2025_GenAI_from_copyright_perspective_FullR_en.pdf, p. 109 (tokenization as a pricing metric).

²¹² Cf. Valerie Benabou, Comment associer les créateurs à la valeur produite par l’utilisation de leurs œuvres par les systèmes d’intelligence artificielle?, *Communication Commerce Electronique*, 4/2025, p. 14: « Un fait générateur de rémunération pourrait, le cas échéant, s’imputer sur les revenus issus de l’utilisation professionnelle de ces outils : il s’agirait alors de prendre en compte **la richesse créée** par la personne **qui vend** des « extrants » générés à l’aide des solutions d’IA ».



rightholder(s) in order to ensure proper identification of each preexisting work (see also above 8.1.2.4.(a)(ii))

213

The establishment of a licensing scheme such as the one outlined above that enables monetization of AI training works would incentivize the sharing of 2D content, such as photographs and videos, through XReco platform²¹⁴. This approach would constitute an interesting paradigm for remuneration systems based on AI training activities, according to which the provider of AI-based XR asset production services would safeguard rights clearance over training content by design and by default and would cater the participation of rightholders of training material in the revenue resulting from the commercialization of AI generated outputs, even in the absence of a copyright-sensitive reproduction of 2D inputs data within those outputs. This would reflect a fair approach towards AI training based on protected 2D content.

8.1.4.4 Challenges

Several challenges are related with the design and implementation of such licensing scheme. To begin with, the implementation of such licensing scheme requires the absence of any application of an exception covering AI training reproductions of 2D inputs. In that regard, given that the application of Text and Data Mining exceptions under Articles 3 and 4 of the CDSM Directive remains pertinent²¹⁵ it would be necessary that the rightholder of 2D inputs has priorly exercised the opt-out under article 4(3) CDSM Directive, before agreeing to the AI generation license. At the same time, AI-generation license could not correspond to a free license, such as a CC license, given that it could not accommodate the reception of remuneration for the licensed uses.

Another delicate issue is the one relating with the actual determination of the royalty share that the rightholder of 2D content should be entitled to. Again, freedom of contract remains the default rule under EU copyright contract law²¹⁶. However, the principle of **appropriate and proportionate** remuneration applies to licenses for the exploitation of 2D content signed by authors (initial copyright rightholders). In view of establishing a FRAND approach in this context, this principle could be voluntarily extended to any XReco AI-generation license, independently of the class of Licensor (namely, initial or secondary). As per art. 18 CDSMD, the adjective “appropriate” is connected with the *necessary adequacy* of a given remuneration under given circumstances²¹⁷ and the adjective “proportionate” requires that the economic value of the remuneration constitutes a fair proportion of the economic value of licensed right²¹⁸.

In case of algorithmic training in the context of 3D reconstruction services, the XReco 3D output stems directly from the training of *specific* 2D content used as training input. Within this context, it could be suggested that a 50/50 split of compensation received by the Licensor of marketed XReco 3D output via the XReco Marketplace could reflect a fair, reasonable and non-discriminatory approach and would result in an appropriate and proportionate remuneration of 2D rightholders against the grant of an AI-generation license. Then, in case of multiple 2D content use as training material, the 2D content 50% royalty share could be

²¹³ Cf. Martin Senftleben, Copyright Data Improvement for AI Licensing – The Role of Content Moderation and Text and Data Mining Rules (May 4, 2024). Available at SSRN: <https://ssrn.com/abstract=4817796> or <http://dx.doi.org/10.2139/ssrn.4817796>, p. 2.

²¹⁴ Cf. DoA, p. 9: “Concluding, XRECO will research, design, and develop a technology to allow for the monetisation of shared data when used to train data-driven models on their results, tracking the contribution of each data point in a meaningful manner. This will further incentivize organizations to share data, with no hidden royalties lost.”

²¹⁵ See D3.1, p. 96.

²¹⁶ See D3.1, Annex VII, p. 164 ff.

²¹⁷ See Theodoros Chiou, Data as (non-monetary) consideration in digital copyright licensing transactions: an essay on the principle of ‘appropriate and proportionate remuneration’ under art. 18 EU Directive 2019/790 in the data era, *see* Synodinou, T., Jougoux, P., Markou, C., Prastitou-Merdi, T. (Eds.), *EU Digital Law in the AI Era*, Springer (forthcoming), under 4.1.

²¹⁸ *Ibid.*, under 4.2.



further distributed within different 2D content rightholders on the basis of the contribution of each specific content in terms of **provision of metainformation** (such as, e.g. datapoints of the scene) that is used for 3D reconstruction purposes (a task for the component of Monetization Manager).

In practical terms, under Scenario 1, the amount due would originate from the 3D asset Licensor and would be paid to XReco within the framework of the Marketplace license signed between the Licensor and XReco platform, when this asset is sold to a Marketplace end-user. It would be calculated on the basis of the remuneration fee determined by the Licensor under the CC+ licensing scheme (see above). In that case, the 3D asset Licensor would have to pay the remuneration for the copyright-sensitive reproduction that takes place for algorithmic training, which is triggered by his actions and proxied by XReco service provider (and not for acts realized by the 3D asset Licensor himself²¹⁹). Then, the amount received would be paid by the XReco platform provider to the Licensor of 2D content use for algorithmic training, within the framework of the XReco AI-generation license signed between those parties.

A similar approach could be adopted under Scenario 2, in case of generation of a 3D output by means of 3D reconstruction services which is not marketed through the XReco Marketplace. In this case, the basis for the 50/50 royalty share would be the amount paid by the XReco user to the XReco platform provider for using XReco's 3D reconstruction services. This approach could be easily implemented in case of a per-download pricing system, but would raise more difficulties in case of subscription-based provision of services.

8.2 Implemented workflow

Compared to what was outlined in D3.1, the licensing workflow has been significantly revised to better align with practical requirements and the architectural evolution of the XReco platform. In particular, the previous approach, where license selection and registration occurred during the asset upload phase via the Orchestrator, has been replaced by a more flexible model in which licensing decisions and formalization are deferred to the publication phase on the Marketplace.

This change stems from the impracticality of anticipating all possible licensing scenarios applicable to uploaded assets. To avoid excessive rigidity while supporting the creation of a sustainable ecosystem for content sharing and monetization, a two-step mechanism has been introduced: a simplified upload phase and a structured publication phase, which may also include monetization.

During the upload phase, the user can submit an asset through the Orchestrator and optionally associate it with a license selected from a limited set. This license is used solely as a metadata tag to indicate the asset's origin (e.g., "No License", one of the supported Creative Commons licenses, or a generic "Other" option). At this stage, no contractual document is generated.

Once uploaded, the asset becomes available for internal operations, including 3D model reconstruction, integration with authoring tools, or organization within content baskets. At any point, the user may choose to publish the asset through the Marketplace. When doing so, the interface presents a license selector, guiding the user in choosing a definitive license from those compatible with the XReco legal framework (Section 8.1). This includes standard Creative Commons licenses (such as CC-BY or CC-BY-NC) or hybrid models like CC-BY-NC combined with CC-Plus. The latter enables non-commercial use while allowing commercial reuse under a paid license agreement.

²¹⁹ Cf. Benabou, op. cit., p. 14: « Si le fournisseur du service d'IA s'est acquitté d'une rémunération selon la nature de l'outil commercialisé et/ou que l'usage des œuvres présentes dans les datasets a déjà fait l'objet d'un paiement, une rémunération due par l'utilisateur ne pourrait se justifier que si l'output présente des similitudes avec des œuvres humaines préexistantes identifiables. Dans ce cas, la licence individuelle retrouverait sa justification. »



Upon confirmation, an automated backend workflow is triggered, involving a coordinated chain of microservices. This workflow is summarized in the following sequence diagram, which illustrates the end-to-end publishing process on the Marketplace.

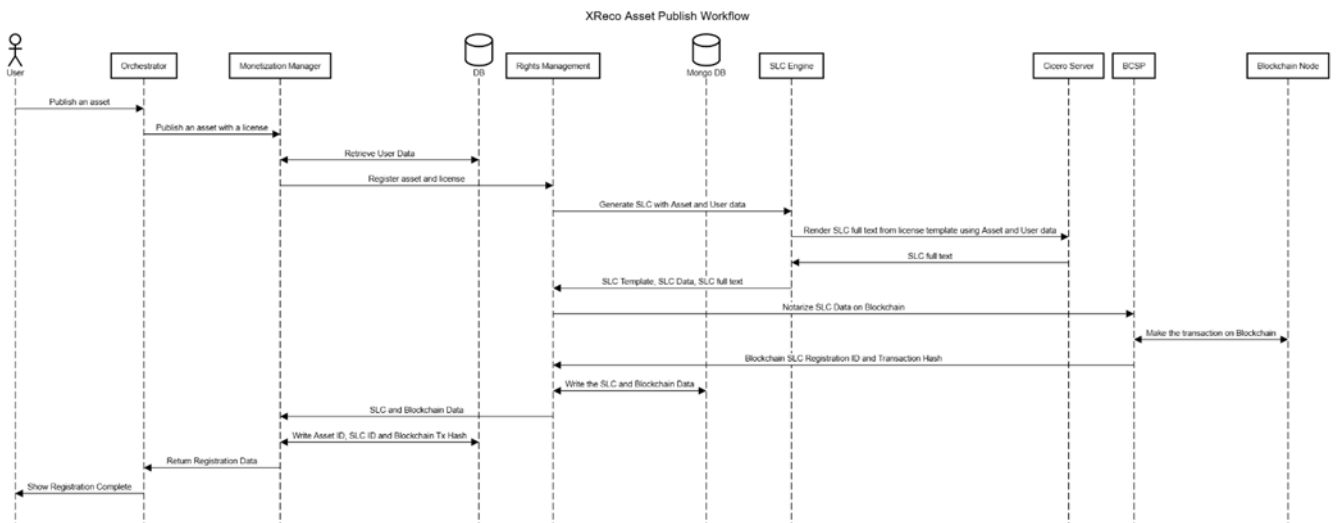


Figure 20: Sequence Diagram showing the publishing workflow of an asset to the marketplace

As shown in Figure 20, the workflow begins with the user interacting with the Orchestrator. By selecting an asset already uploaded to the platform, the user is prompted to choose a license through a wizard. Upon clicking “Publish”, the Orchestrator sends the asset metadata and selected license to the Monetization Manager (MM).

The Monetization Manager then forwards the request to the Rights Management (RM) Service, which initiates the Smart Legal Contract (SLC) generation and blockchain notarization. Rights Management first interacts with the SLC Engine to generate an SLC composed of two parts: the SLC Template and the SLC Data. This step leverages the Cicero Server, which takes the SLC Template and the SLC Data as input and compiles them into a complete, human-readable contract.

Following contract generation, Rights Management calls the Blockchain Service Provider (BCSP) to notarize the SLC Data on the XReco blockchain. The resulting SLC and notarized data are stored in MongoDB for persistence and future consultation. These results are then returned to the Monetization Manager, which registers the user, asset and SLC association in the internal database. Finally, the Orchestrator is notified with the updated information, enabling the frontend to reflect the asset’s published status and license details.

This workflow ensures full consistency between selected licenses, generated contracts, and monetization mechanisms. Blockchain-based notarization provides legal traceability and immutability of license records. This revised model clearly separates operational phases (upload, usage, publication), giving users greater control while supporting advanced licensing schemes, such as the CC-BY-NC + CC-Plus combination, that preserve openness for non-commercial use while enabling commercial reuse through contractual payment.

8.3 Rights Management and Monetization Architecture

As described in the previous section, the selection and notarization of licenses have been moved from the upload phase to the asset publication phase. Compared to the architecture illustrated in D3.1, a new microservice has been introduced: the Monetization Manager. This component acts as an intermediary



between the Orchestrator and the Rights Management service, thereby centralizing the handling of licenses and payments.

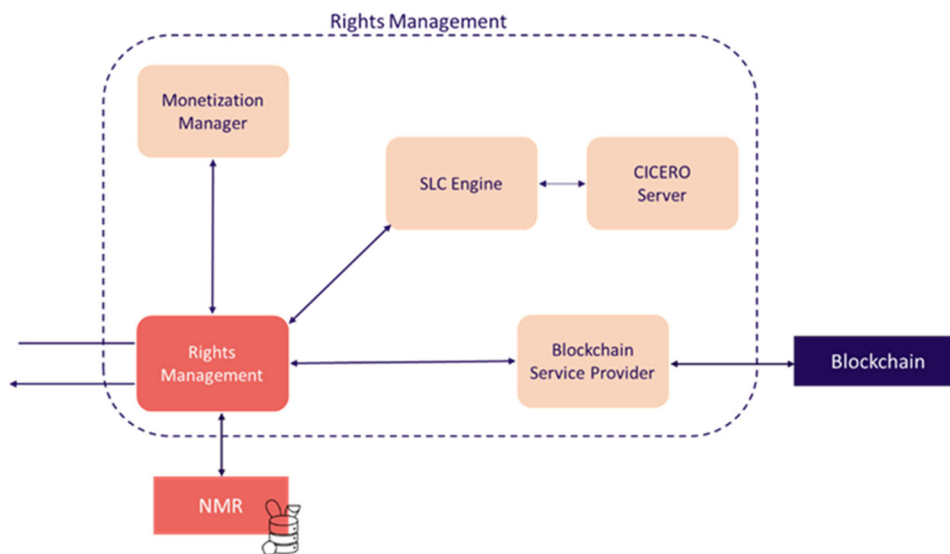


Figure 21: Rights Management Architecture described in D3.1

In the initial version of the architecture (Figure 21), the Orchestrator communicated directly with the Rights Management module to initiate the generation of the Smart Legal Contract (SLC) and its notarization on the blockchain. This process took place during the asset upload phase, before any decision was made regarding its publication on the Marketplace. At that stage, no backend component for monetization was yet available, nor was there a structured system for managing the commercial lifecycle of the asset.

With the introduction of the updated architecture (Figure 22), the licensing workflow has been shifted to the asset publication phase on the Marketplace. The Monetization Manager has been introduced as a new microservice that receives the publication request from the Orchestrator, triggers the SLC generation through the Rights Management service, handles payments via Stripe, and stores a persistent record of all transactions in a PostgreSQL database.

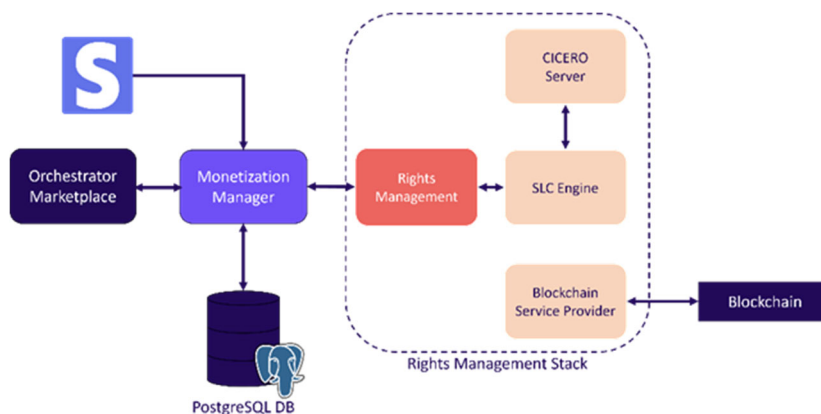


Figure 22: Current Rights Management and Monetization architecture.

The Rights Management service retains the core functionality described in D3.1, including the generation and notarization of SLCs. However, the addition of the Monetization Manager allows for a clear separation of responsibilities: the Rights Management module focuses on the legal aspects, while the Monetization Manager oversees the payment and monetization processes. The integration between these two



components ensures full consistency between selected licenses, generated contracts, and executed transactions.

This architecture defines the general structure of the rights management and licensing flow in the XReco platform, with the Monetization Manager now acting as the main point of interaction between the Orchestrator, Rights Management, and payment services. The next section focuses specifically on the Monetization component, detailing how it manages transactions and integrates with Stripe for payment processing.

8.4 Monetization Manager

The Monetization Manager is a backend component at the core of XReco's monetization infrastructure. It is responsible for automating the economic workflows tied to the publishing and commercial licensing of assets on the XReco Marketplace. Though invisible to end users, it plays a fundamental role in ensuring seamless and secure transactions by integrating with the Rights Management subsystem and external payment services.

As described in the previous sections, the licensing and monetization flow is triggered during the publication phase. When a user selects an asset and chooses to make it available on the Marketplace, the Orchestrator delegates the request to the Monetization Manager. At this point, the MM activates a coordinated backend workflow that includes license registration, payment setup, and transaction logging.

If the selected license allows for commercial reuse through payment, the Monetization Manager generates a SLC by interacting with the Rights Management system. The contract is composed and validated via the SLC engine, and its essential metadata is notarized on the blockchain through the Blockchain Service Provider. This ensures immutability and legal verifiability of the license terms associated with the asset.

In parallel, the MM integrates with Stripe to create a secure checkout session specific to the published asset. Buyers are redirected to a Stripe-hosted payment interface, where they can complete the purchase of the commercial license. This architecture ensures that XReco never handles sensitive payment data directly, thereby maximizing privacy, security, and compliance with financial regulations.



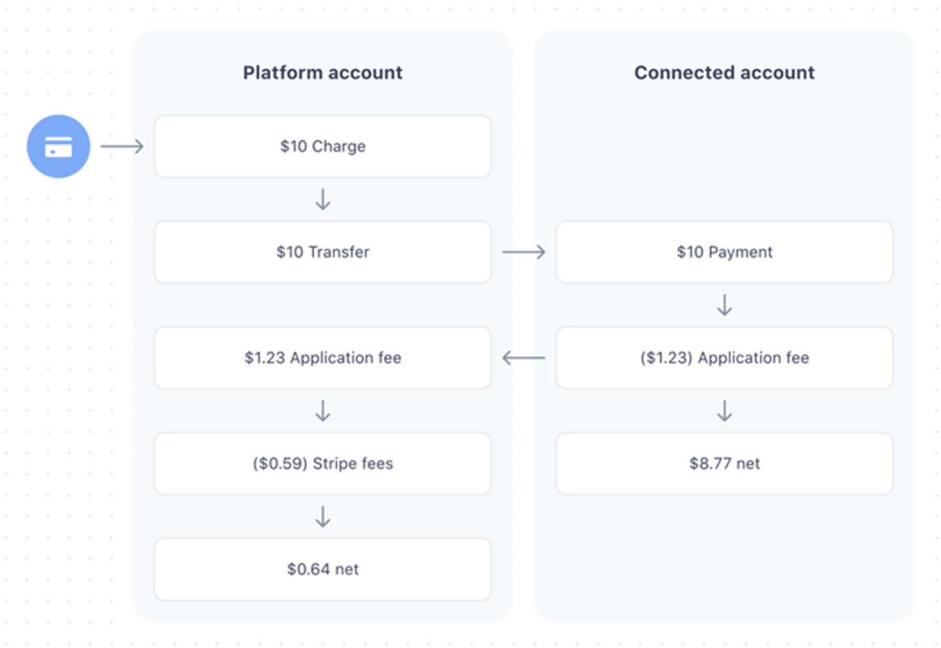


Figure 23: Diagram of the Stripe "collect then transfer" payment model.

Once the payment is confirmed, the Monetization Manager computes the revenue share for the asset’s rights holder, deducts the service fee retained by the platform (Figure 23), and proceeds with the fund distribution using Stripe’s “collect then transfer” model. This approach allows the platform to temporarily collect payments, apply platform-wide logic, and automate payouts to authors without manual intervention. The Stripe transaction fees are absorbed into the service fee retained by XReco, ensuring a predictable experience for both buyers and sellers.

Each purchase results in the automatic generation of an “XReco Purchase Receipt” (Figure 24). This receipt is a Smart Legal Contract that captures the key elements of the transaction: asset identifier, buyer and seller identities, timestamp, license terms, and amount paid. The receipt is also notarized to track the purchase of the asset on blockchain.



XReco Purchase Receipt

This document certifies that the **Licensee**, fincons_test (Email: finconstest@example.com), identified by XReco ID **ab9f217a-175d-45ab-8dbc-b17de52a4244**, has purchased the **CC-Plus License Unlock** for the asset named **Test**, identified by ID **84c06b9f-7983-4e77-82b5-9fe5530111b9**, through the XReco platform.

The asset has been published on the XReco Marketplace by the following Licensors:

- ID: ed9a586a-8c32-47dd-9b7d-223ff432308a
- Username: mog
- Email: undefined
- Ownership Percentage: 100.0%

The Licensee has completed the purchase with a total payment of **EUR 5.0**, processed via the platform on **12 Feb 2025 09:16:12**.

This receipt confirms that the Licensee holds the rights associated with the **CC-Plus License Unlock** as stipulated by the terms and conditions under which the asset was registered on the XReco platform.

Figure 24: Representation of the XReco Purchase Receipt, generated each time an asset is purchased.

All information related to transactions, license contracts, and payment outcomes is stored in a persistent PostgreSQL database. The Orchestrator is then notified with updated metadata, allowing the frontend to reflect the asset's published status and associated commercial license. From a user's point of view, these backend operations are abstracted behind a step-by-step publication interface. A detailed overview of this interface, including screenshots, is provided in the following section.

By consolidating financial operations, legal guarantees, and backend consistency, the Monetization Manager plays a foundational role in enabling a sustainable and transparent content economy within the XReco ecosystem.

8.5 Marketplace User Interface and Workflow (MOG)

The XReco web platform is a modular ecosystem designed to support the full lifecycle of XR content, with a strong focus on asset creation through advanced 3D reconstruction and optimization services. It offers a unified suite of tools for managing multimedia assets and producing immersive experiences, including:

- The Orchestrator allows users to upload, organize, and process media assets, enabling reconstruction workflows and seamless publishing to the Marketplace. It also integrates tools such as XRCapsules, Zaubar CMS, and the Authoring Tool for extended content creation.
- The XReco Marketplace is where users can discover, purchase, and download published assets, with a strong focus on XR-ready content.

A detailed overview of the Orchestrator, including asset management, repository integration, and service orchestration, is available in D4.2.

The Marketplace integrates with the Neural Media Repository (NMR), ensuring all published assets are stored and searchable with AI-generated metadata. It also integrates with the Monetization API and leverages the



Monetization stack described in the previous sections to enable licensing management, commercial transactions, and secure user accounts within a cohesive and robust framework.

8.5.1 Publish to Marketplace

To publish an asset, users must locate it in the Orchestrator. Only assets owned by the user or their organization can be published. These are typically stored in the Neural Media Repository (NMR) or in content baskets.

Clicking the "Publish to Marketplace" button (Figure 25) on the asset preview initiates the publishing workflow. A Stripe Seller account is required to complete the process. If not already configured, the user will be prompted to set it up.

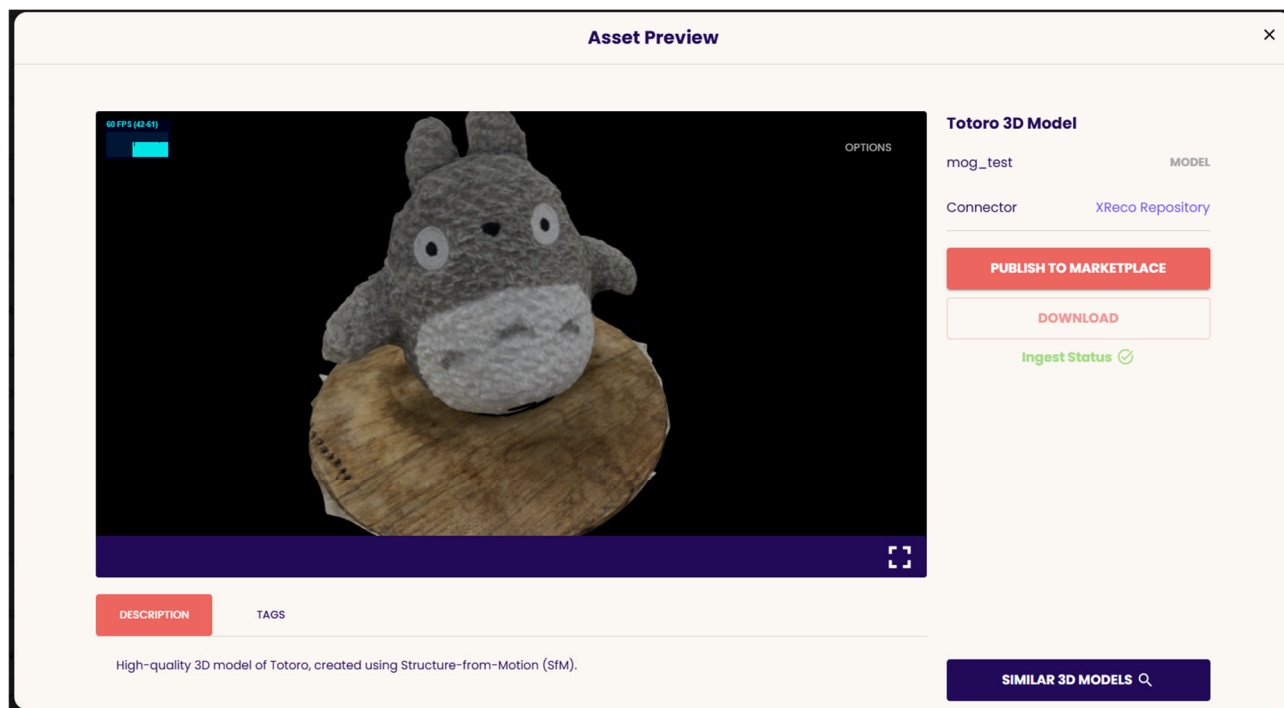


Figure 25: Orchestrator: Asset preview.

During publishing, users are guided through an interactive licensing wizard designed to help determine the appropriate licensing terms for their asset. Figure 26 illustrates the initial view of this licensing wizard. The interface is divided into two main sections:

- On the left side, the wizard displays a step-by-step progression of licensing questions and selectable answers, helping users navigate through license options based on their preferences and expertise.
- On the right side, the license overview dynamically updates to reflect the current selections. This panel provides a summary of the chosen license(s) and allows users to preview detailed legal texts associated with each license.

When users complete all the licensing steps and reach the final confirmation stage, the “Publish Asset” button on the right panel becomes active, allowing them to finalize and publish their asset to the Marketplace.



Publish Asset

Follow the steps below to select the most suitable license for your asset and publish it to the Marketplace.

1 DO YOU KNOW WHICH LICENSE YOU NEED?
1 of 6

☐ Yes. I know the license I need.

☐ No. I need help selecting a license.

NEXT

2 ATTRIBUTION
2 of 6

3 COMMERCIAL USE
3 of 6

4 DERIVATIVE WORKS
4 of 6

5 SHARING OF DERIVATIVE WORKS
5 of 6

6 CONFIRM THAT LICENSING CONDITIONS ARE APPROPRIATE
6 of 6

SELECTED LICENSE

You haven't selected a license yet. Use the wizard on the left to choose one!

PUBLISH ASSET

Figure 26: Marketplace: Licensing wizard - initial state.

Users begin by answering: “Do you know which license you need?”, which determines whether they follow the non-expert (guided) or expert (direct) licensing path (Figure 27).

1 DO YOU KNOW WHICH LICENSE YOU NEED?
1 of 6

☐ Yes. I know the license I need.

☒ No. I need help selecting a license.

NEXT

Figure 27: Marketplace: Licensing wizard (expertise question).

8.7.1.1. Non-Expert Path (Guided License Selection)

This path is designed for users who are unfamiliar with licensing and need step-by-step guidance to select the most appropriate license for their asset. The wizard poses a series of clear, easy-to-understand questions that help tailor the license to the user’s preferences which cover the topics outlined below:

- **Attribution:** Users decide whether others must credit them when using the asset. If the user selects "No," the asset is licensed under CC0 (public domain). Selecting "Yes" enables further customization.
- **Commercial Use:** Users define how their asset can be used commercially, choosing one of three options:
 - **Commercial use with payment:** This option activates the CC Plus model (CC BY-NC + UNLOCK). The asset is available free of charge for non-commercial use with attribution. For



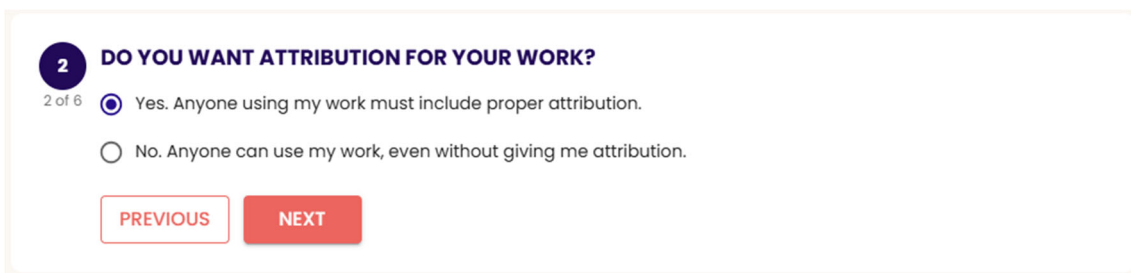
commercial use, users must pay a fee (e.g., €5.00), after which they receive a worldwide, non-exclusive, perpetual license to use the asset commercially without attribution. This path skips questions on derivative works, and users are prompted to set the asset price.

- **Free commercial use:** Allows others to use the asset for commercial purposes at no cost. Users continue to questions about derivative works and sharing terms.
- **No commercial use:** Prohibits all commercial use. Users proceed to define terms for adaptations and sharing.
- **Derivative Works:** The user decides if others may remix, adapt, or build upon their work. Selecting "No" ends the wizard, and a no-derivatives license is applied.
- **Sharing of Derivative Works:** If derivatives are allowed, the user chooses whether adaptations must use the same license or can be shared under any terms.
- **Final Confirmation:** The user confirms they have rights to license the work and understand the irrevocable nature of Creative Commons licenses.

Each step dynamically adjusts the license parameters to ensure the asset is published under a legally appropriate and clear framework, without requiring the user to have prior legal knowledge.

In the following walkthrough, we demonstrate the CC Plus licensing workflow, which allows the asset to be used commercially for a fee.

In the first step, the user indicates that attribution is required by selecting “Yes” in the attribution question (see Figure 28).



2 DO YOU WANT ATTRIBUTION FOR YOUR WORK?

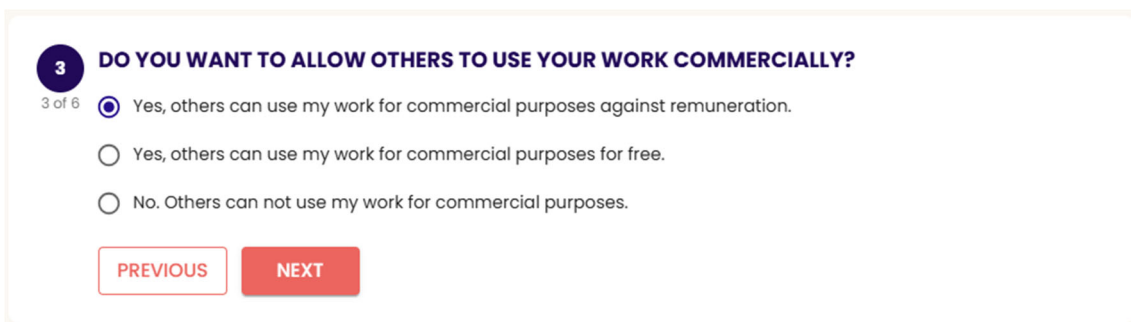
2 of 6 ☒ Yes. Anyone using my work must include proper attribution.

☐ No. Anyone can use my work, even without giving me attribution.

PREVIOUS **NEXT**

Figure 28: Marketplace: Licensing wizard – CC Plus example (attribution question).

Next, the user specifies that others may use the asset for commercial purposes, provided they pay a fee (see Figure 29).



3 DO YOU WANT TO ALLOW OTHERS TO USE YOUR WORK COMMERCIALY?

3 of 6 ☒ Yes, others can use my work for commercial purposes against remuneration.

☐ Yes, others can use my work for commercial purposes for free.

☐ No. Others can not use my work for commercial purposes.

PREVIOUS **NEXT**

Figure 29: Marketplace: Licensing wizard - CC Plus example (commercial use question).

The user is then prompted to set a price for commercial use of the asset (see Figure 30).



6 WHAT PRICE WOULD YOU LIKE TO SET FOR YOUR ASSET?

6 of 7

€ 15.00

This will be the cost others pay for commercial use of your asset.

PREVIOUS NEXT

Figure 30: Marketplace: Licensing wizard - CC Plus example (set asset price).

After clicking “Next” one last time, the user reaches the final step. Here, they can review the licensing terms applied to their asset: CC BY-NC for non-commercial use and the Unlock license for paid commercial. As shown in Figure 31, the user can now click “Publish Asset” to complete the process.

Publish Asset

Follow the steps below to select the most suitable license for your asset and publish it to the Marketplace.

- 1 LICENSE EXPERTISE**
1 of 7 I need help selecting a license.
- 2 ATTRIBUTION**
2 of 7 Anyone using my work must include proper attribution.
- 3 COMMERCIAL USE**
3 of 7 Others can use my work, even for commercial purposes.
- 4 DERIVATIVE WORKS**
4 of 7 This step is disabled because you require remuneration for commercial use.
- 5 SHARING OF DERIVATIVE WORKS**
5 of 7 This step is disabled because you require remuneration for commercial use.
- 6 ASSET PRICING**
6 of 7 €15.00
- 7 CONFIRM THAT LICENSING CONDITIONS ARE APPROPRIATE**
7 of 7
 - ☒ I own or have authority to license the work.
 - ☒ I have read and understand the terms of the license.
 - ☒ I understand that CC licensing is not revocable.

PREVIOUS

SELECTED LICENSES

CC Plus
(CC BY-NC + UNLOCK)

Non-Commercial Use Conditions: Attribution required. Free to share and adapt for non-commercial purposes only. [REVIEW LICENSING CONDITIONS](#)

Commercial Use Conditions: Requires a €15.00 payment. Upon payment, grants a worldwide, non-exclusive, perpetual, sublicensable license to use and exploit the asset (including to reproduce, distribute, modify, display and perform it), without licensor's attribution, for commercial purposes. [REVIEW LICENSING CONDITIONS](#)

PUBLISH ASSET

Figure 31: Marketplace: Licensing wizard – CC Plus example (final step).

8.7.1.2. Expert Path (Direct License Selection)

For users who know which license they want, the wizard provides a direct license selection path.

The first step in this path is selecting a pricing model: either Free or Free + Paid (see Figure 32).



2

2 of 4

2

HOW DO YOU WANT TO OFFER YOUR ASSET?

☒ Free: Offer your work at no cost under your selected CC license.

☐ Free+Paid: Free for non-commercial use, paid for commercial use (CC Plus).

PREVIOUS

NEXT

Figure 32: Marketplace: Licensing wizard - Pricing model selection.

If “Free” is selected, the user will then choose a Creative Commons license (see Figure 9). If “Free + Paid” is selected, the user must define the asset’s price (see Figure 33).

3

3 of 4

3

WHICH LICENSE DO YOU NEED?

License Selection

CC BY

CC0

CC BY

CC BY-SA

CC BY-ND

CC BY-NC

CC BY-NC-SA

CC BY-NC-ND

4

4 of 4

4

CONFIRM THAT LICENSING CONDITIONS ARE APPROPRIATE

APPROPRIATE

Figure 33: Marketplace: Licensing wizard - License selection.

Finally, the user reviews and confirms the licensing conditions. An example of a completed workflow is shown in Figure 34.

Publish Asset

Follow the steps below to select the most suitable license for your asset and publish it to the Marketplace.

1

1 of 4

1

LICENSE EXPERTISE

I know the license I need.

2

2 of 4

2

PRICING MODEL

Anyone can use your work for free, subject to the conditions you set.

3

3 of 4

3

LICENSE SELECTION

CC-BY

4

4 of 4

4

CONFIRM THAT LICENSING CONDITIONS ARE APPROPRIATE

☒ I own or have authority to license the work.

☒ I have read and understand the terms of the license.

☒ I understand that CC licensing is not revocable.

PREVIOUS

SELECTED LICENSE

CC BY

Attribution required. Free to use, modify, and distribute for any purpose, including commercial use. [REVIEW LICENSING CONDITIONS](#)

PUBLISH ASSET

Figure 34: Marketplace: Licensing wizard - Expertise path.



8.5.2 Marketplace

The Marketplace provides access to a wide range of published assets, including those shared by other users and assets published by the current user. Assets can be browsed or located using the built-in search and filter features. The Marketplace landing page is shown in Figure 35.

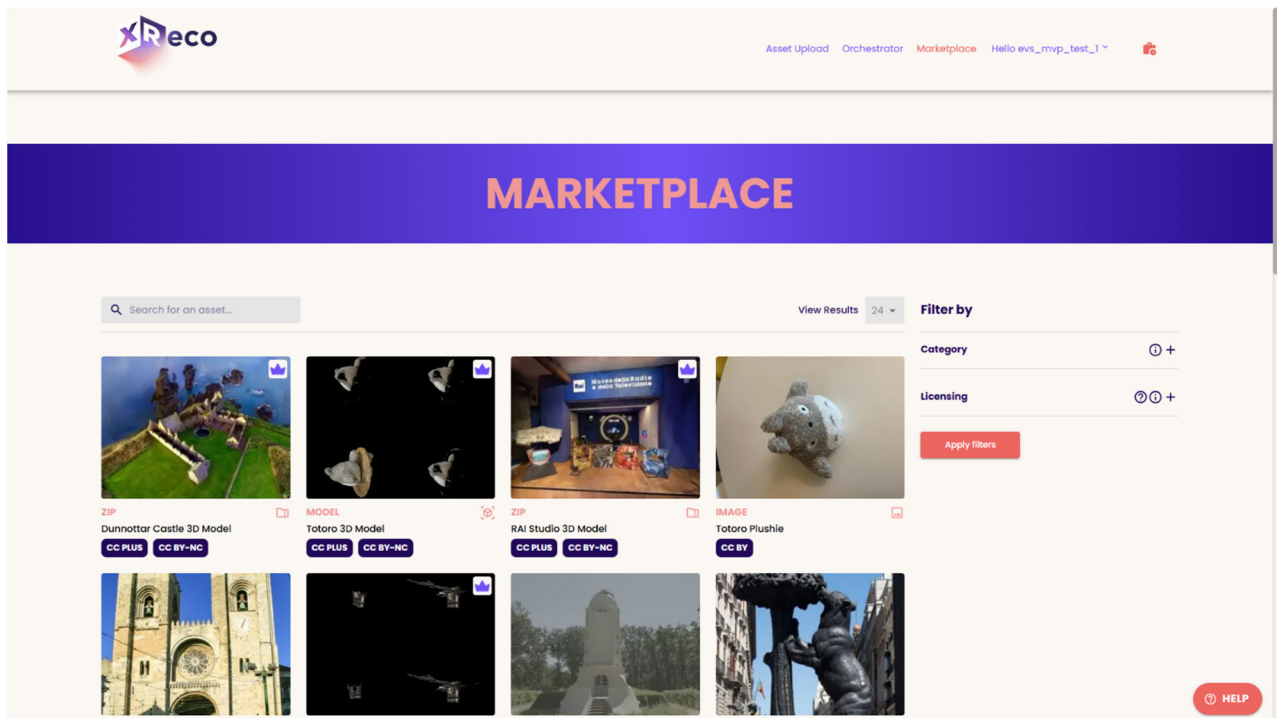


Figure 35: Marketplace: Landing page.

Each asset listing shows a thumbnail, title, asset type, and its license. Clicking the license badge reveals a preview of usage terms.

Filtering options include category and license type (Figure 36). To refine the results, the user selects one or more filters and clicks "Apply filters". Filters can be used in combination with or without a text-based search query. To reset the view, the user can click "Clear filters" to remove all selected filter options.



Figure 36: Marketplace: Search filters.

Assets in the Marketplace are distributed under one of two pricing models:

- **Free** – fully accessible at no cost, under standard Creative Commons terms.
- **Free+Paid** – available free for non-commercial use, with payment required for commercial usage under CC Plus licensing.

When accessing the asset preview (Figure 37), users can take advantage of the advanced search capabilities also found in the Orchestrator. These include AI-generated tags, displayed under the Tags tab, which function as clickable filters for exploring related content. The preview further supports visual similarity searches through dedicated buttons, enabling the discovery of related images, video segments, and 3D models. Cross-modal search functionality is also available, allowing to retrieve 3D models using an image as input. A full overview of these features is provided in D4.2, which details the Orchestrator’s capabilities.



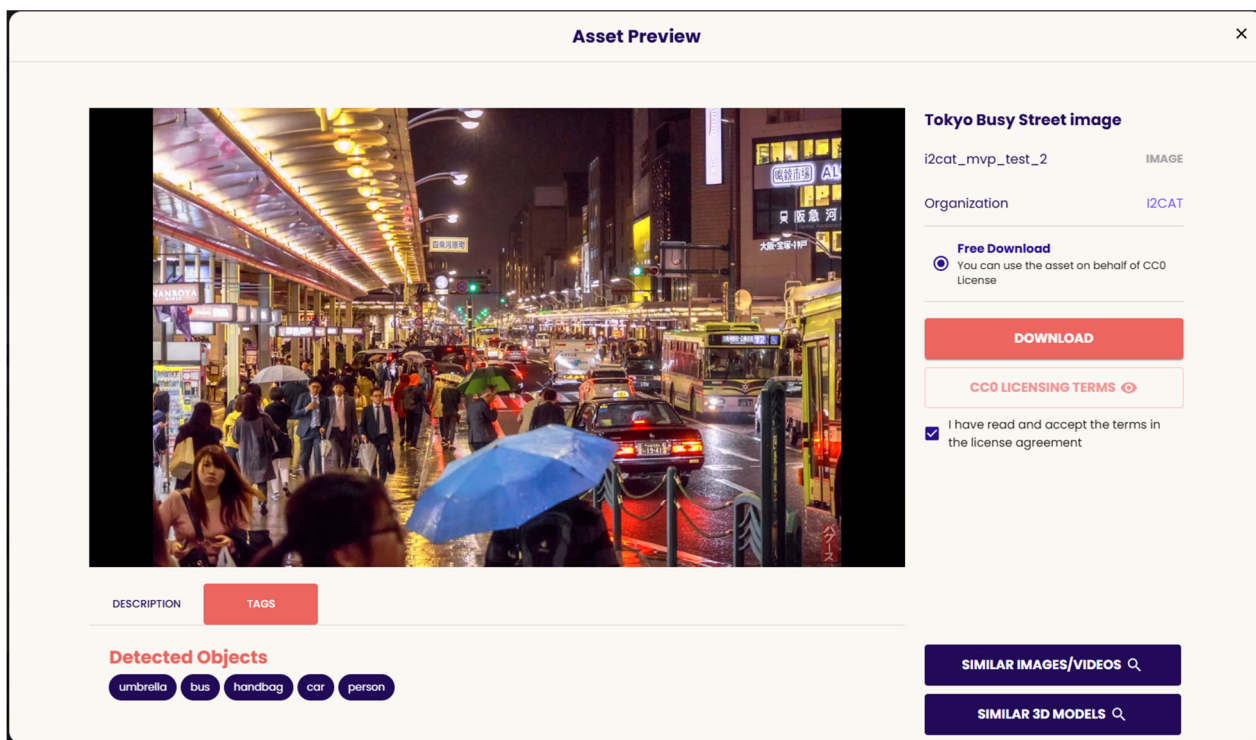


Figure 37: Marketplace: Asset preview (Free asset).

For assets under the Free pricing model (example shown in Figure 13), users must accept the license terms by checking a confirmation box before downloading. Terms can be reviewed via the "Licensing Terms" button. Once accepted, the "Download" button becomes active. If the asset was published by the user, it can be downloaded immediately without additional confirmation.

For assets offered under the Free+Paid pricing model (example shown in Figure 38), users are presented with two options:

- **Non-commercial use:** The user may review and accept the license terms to download the asset free of charge.
- **Commercial use:** The user must purchase the asset under the CC Plus licensing terms.



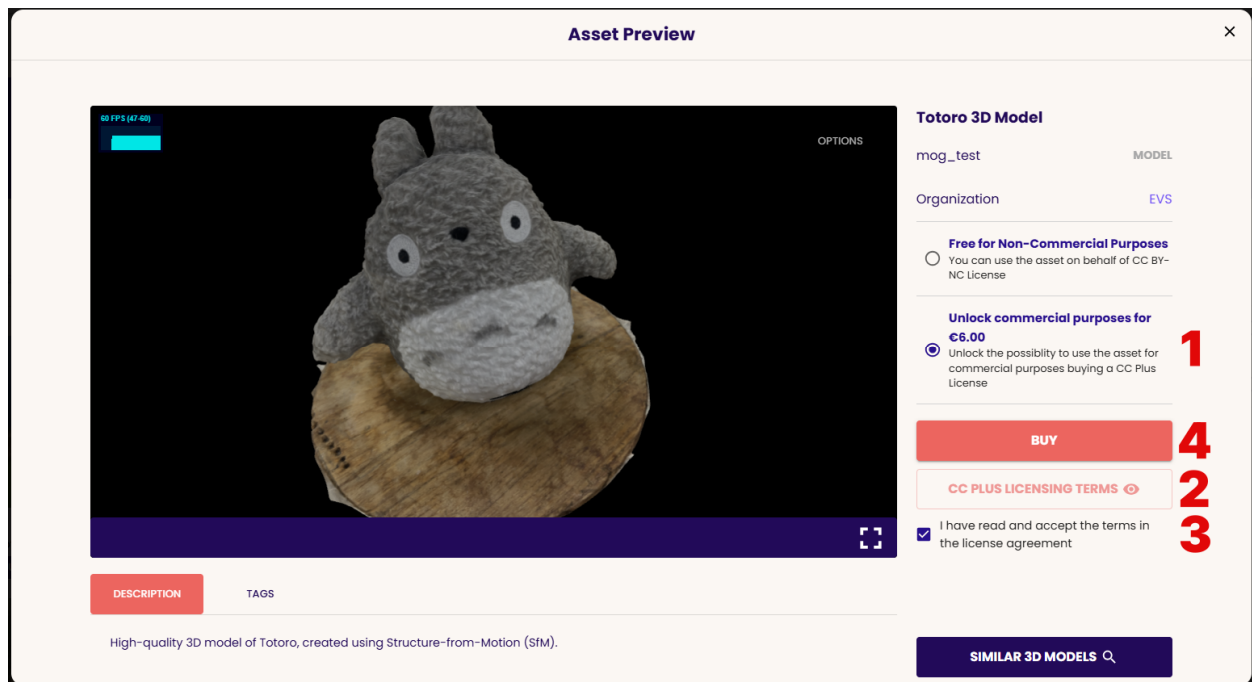


Figure 38: Marketplace: Asset preview (Free+Paid asset).

To purchase a Free+Paid asset for commercial purposes, users must:

1. Select the Commercial purposes option.
2. Read the CC Plus licensing terms.
3. Accept the licensing terms by ticking the checkbox.
4. Initiate the purchase process by clicking the “Buy” button.

This action redirects the user to the Stripe Checkout interface (see Figure 39), where personal details and payment information must be provided to complete the asset purchase.



XRECO DEVELOPMENT **TEST MODE**

Totoro 3D Model

€ 6,00

Pagar com link

Ou

E-mail

test@mail.com

Você quis dizer @gmail.com?

Forma de pagamento

☒ **Cartão**

Dados do cartão

4000 0000 0000 0077 visa

01 / 27 123 1234

Nome do titular do cartão

Test User

País ou região

Portugal

☐ **IDEAL**

☐ **Bancontact**

☐ **EPS**

☐ **Salvar meus dados com segurança para fazer checkout com 1 clique**
 Pague mais rápido em XRECO DEVELOPMENT e em todos os lugares que aceitam o LINK.

Pagar

Powered by Stripe
 Termos Privacidade

Figure 39: Marketplace: Stripe checkout page.

Upon successful payment, the system displays a Purchase Success page. After a brief delay, the user is automatically redirected to the “My Purchases” page, where the acquired asset becomes available.



8.5.3 User Options

The User Options section consists of a collection of pages primarily dedicated to managing Marketplace-related activities. Figure 40 shows the available user options on the XReco platform, accessible from the top navigation bar.

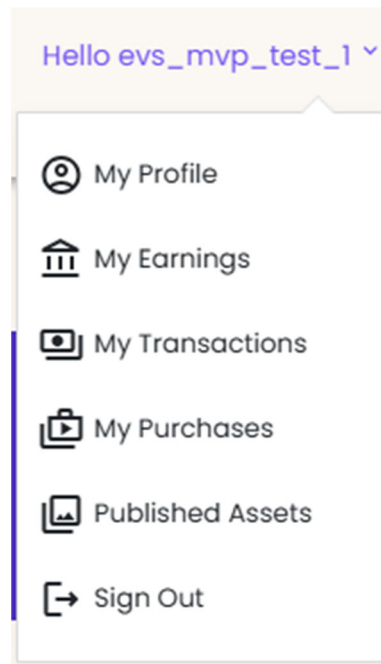


Figure 40: XReco Platform: User options.

The following provides a detailed description of each option available in the user dropdown menu:

1. My Profile

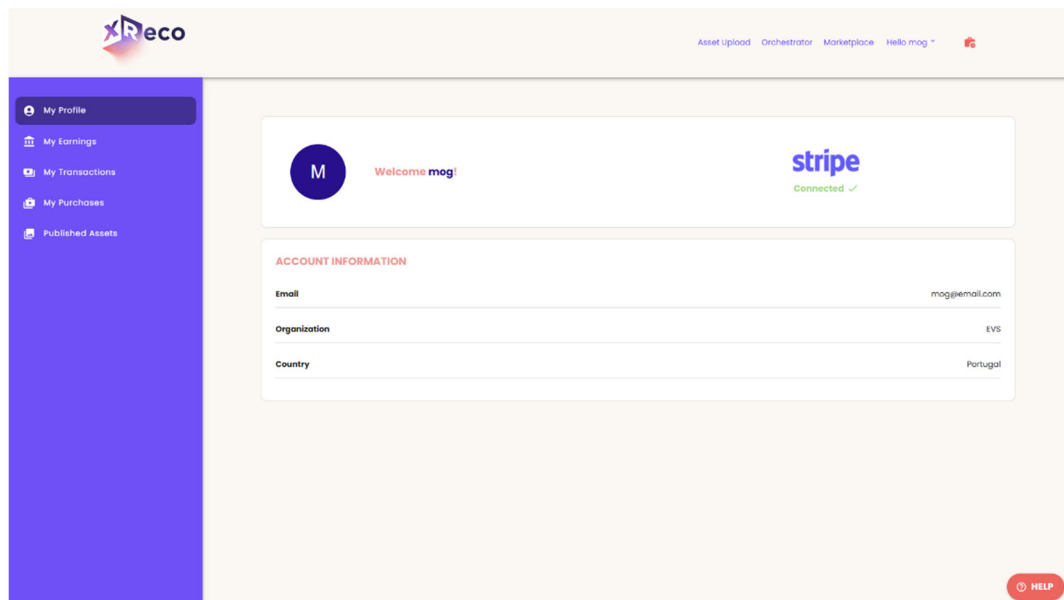


Figure 41: XReco Platform: My Profile.



The My Profile page (Figure 41) displays the user’s personal and account information. It also provides functionality to connect a Stripe account if this has not yet been configured. Establishing this connection is essential for publishing assets to the Marketplace and for receiving payouts from asset sales.

2. My Earnings

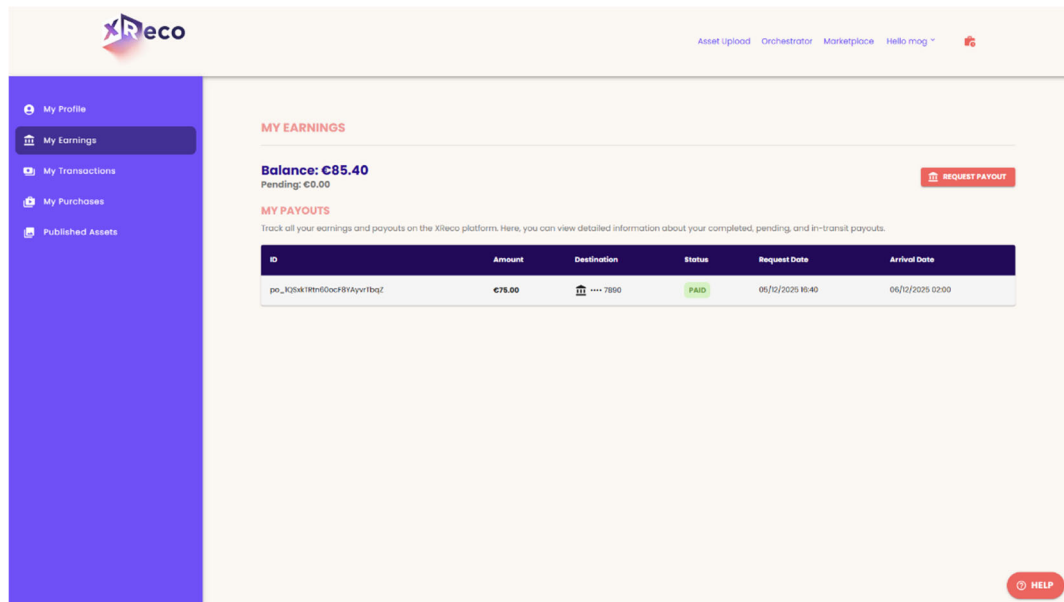


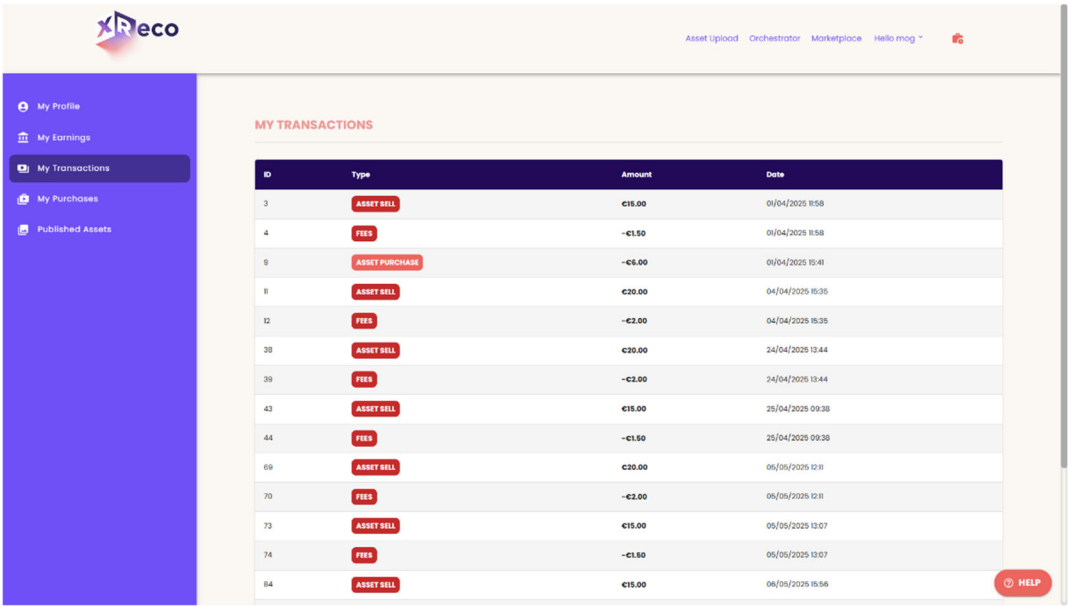
Figure 42: XReco Platform: My Earnings page.

The My Earnings page (Figure 42) offers a comprehensive summary of the user’s financial activities on the platform. It enables tracking of earnings and payout statuses, including completed, pending, and in-transit transactions. Users can initiate payout requests to their linked bank accounts. Key financial indicators displayed include:

- **Balance:** The total earnings currently available for payout.
- **Pending:** Earnings that are still undergoing the clearing process.

3. My Transactions



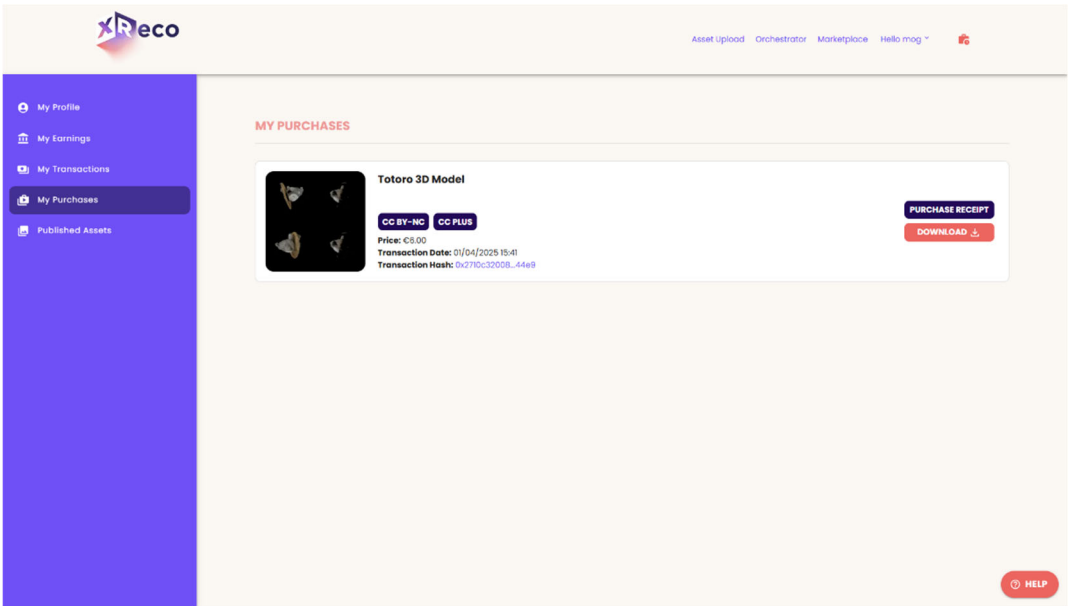


ID	Type	Amount	Date
3	ASSET SELL	€15.00	01/04/2025 15:58
4	FEES	-€1.50	01/04/2025 15:58
9	ASSET PURCHASE	-€6.00	01/04/2025 15:41
11	ASSET SELL	€20.00	04/04/2025 15:35
12	FEES	-€2.00	04/04/2025 15:35
38	ASSET SELL	€20.00	24/04/2025 13:44
39	FEES	-€2.00	24/04/2025 13:44
43	ASSET SELL	€15.00	25/04/2025 09:38
44	FEES	-€1.50	25/04/2025 09:38
69	ASSET SELL	€20.00	05/05/2025 12:11
70	FEES	-€2.00	05/05/2025 12:11
73	ASSET SELL	€15.00	05/05/2025 13:07
74	FEES	-€1.50	05/05/2025 13:07
84	ASSET SELL	€15.00	06/05/2025 15:56

Figure 43: XReco Platform: My Transactions page.

The My Transactions page (Figure 43) provides a detailed record of all blockchain transactions associated with the user account. This includes asset purchases, sales, XReco platform fees, and other relevant transaction entries, enabling users to monitor their complete transaction history.

4. My Purchases



MY PURCHASES	
	Totoro 3D Model CC BY-NC CC PLUS Price: €5.00 Transaction Date: 01/04/2025 15:41 Transaction Hash: 0x270c3200b...44e9 PURCHASE RECEIPT DOWNLOAD

Figure 44: XReco Platform: My Purchases page.

The My Purchases page (Figure 44) lists all assets acquired by the user through the Marketplace. It allows users to preview detailed information for each purchased asset, including asset preview, title, original licensing terms, purchase price, transaction date, and blockchain transaction hash. Users can also view the purchase receipts and download the assets directly from this page.



5. Published Assets

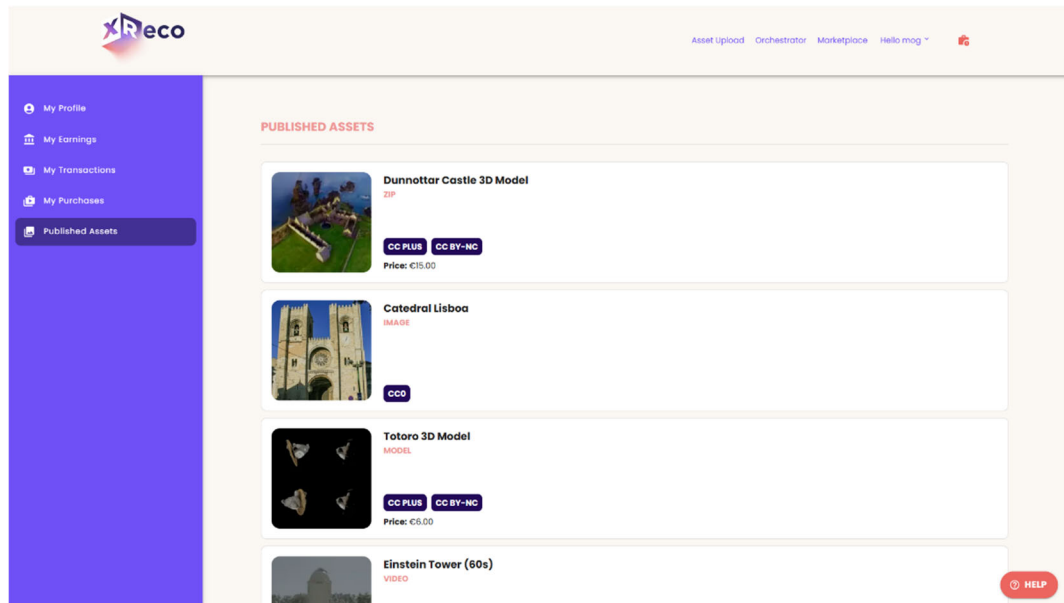


Figure 45: XReco Platform: Published Assets page.

The Published Assets page (Figure 45) displays a list of all assets the user has published to the Marketplace. Users can review license details for each asset by clicking on the corresponding license badge.

6. Sign Out

This option allows the user to securely sign out of the XReco platform, terminating the current session.

8.6 Data valuation (CERTH)

D3.1 established a theoretical framework for Data Valuation, specifically detailing its integration within a Data Monetization strategy. This framework involves quantifying the importance of individual data samples used in training machine learning models, particularly within the context of Neural Rendering-based scene reconstruction. This approach is especially relevant for large datasets comprising contributions from multiple parties, where compensation can be linked to the value each party's data provides in the final trained model. In D3.1 the methods that are relevant in the context of XReco were presented.

Initially, the theoretical underpinnings for Data Valuation were explored predominantly through the lens of Data Monetisation. However, subsequent insights from stakeholder feedback necessitated a recalibration of this contextual approach. The current investigative trajectory is therefore predominantly legalistic, focusing specifically on quantifying the degree to which a reconstructed scene, having been trained on an extensive dataset comprising contributions from various parties, can be legally categorised as a derivative work to the original source data.

In D3.1, Data-OOB²²⁰ was presented as a candidate adoption of an approach which was inspired by Data Valuation using Reinforcement Learning (DVRL), was proven suitable for the following reasons:

- It does not require any calculation after training the original model.
- It adds an acceptable computational overhead.

²²⁰ Kwon, Y. & Zou, J.. (2023). Data-OOB: Out-of-bag Estimate as a Simple and Efficient Data Value. Proceedings of the 40th International Conference on Machine Learning, in Proceedings of Machine Learning Research.



- It outputs contribution scores which are continuous values that correspond to either beneficial (positive), or damaging (negative) images. The values are proportional to their impact.
- Specific contribution scores were identified that tend to be reproducible with different training setups. Several tests we carried out to verify this, for different versions of contribution scores.
- The contribution scores were explained by visual inspection and by re-training the NeRF model by selecting or excluding images based in their contribution scores.

DVRL enables the valuation of data within a single training process. In each training step, DVRL calculates a reward, based on updates to the validation loss, which then feeds into a CNN. This CNN, acting as a valuator function, outputs the probability of selecting a specific datum (in this case, an image) from a batch. As the training process converges, this probability becomes proportional to the data's impact on the model.

The algorithm that was tested within XReco is a modification of the original algorithm, mainly decoupling the reward from the CNN valuator function for four main reasons:

- To reduce processing time.
- To avoid memory breakdown by selecting one image in each iteration step.
- To select only image permitted to select adequate number of pixels/rays from it.
- To examine various reward methodologies and contribution scores.

The overall process is presented in Figure 46.

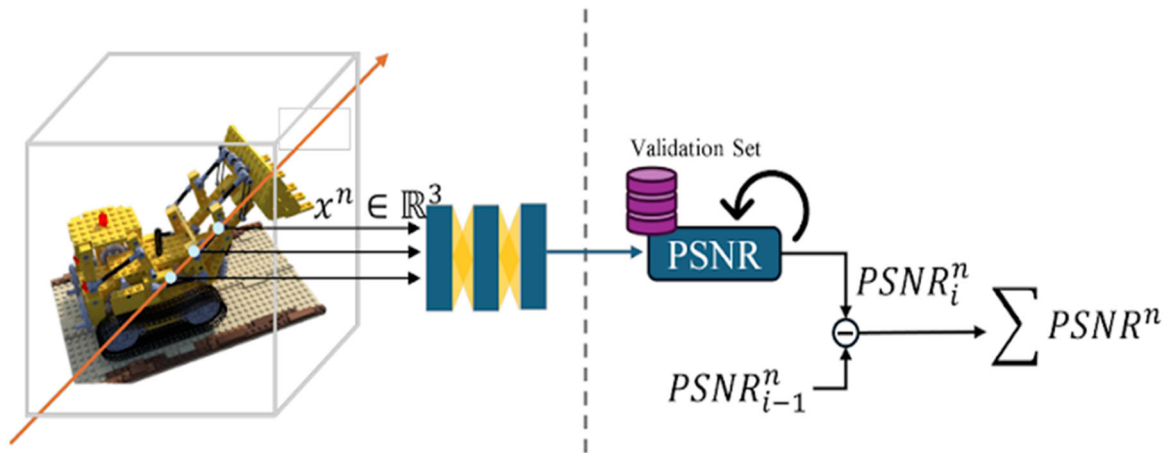


Figure 46: Method overview: During NeRF training, we evaluate a small validation set after each training iteration i , recording the PSNR associated with image I_i . When image I_i is revisited, we compute the change in PSNR by subtracting the previous value from the current one. These PSNR differences are then aggregated to estimate each image's DV_{psnr} .

To quantify the individual contribution of each image in the context of NeRF-based reconstruction, we leverage the mentioned DVRL-inspired method. Contribution is assessed via reconstruction quality metrics based on PSNR and MSE. At each training iteration, we evaluate the PSNR on a fixed test set and associate it with the currently sampled image I_i . For subsequent appearances of I_i , we compute the change in PSNR relative to its previous value. These PSNR deltas are then aggregated across iterations to quantify each image's overall influence on reconstruction quality, which we refer to as DV_{psnr} .

The instantaneous PSNR difference at training step t , is calculated via:

$$\Delta_{psnr_i}(t) = PSNR(I_i, \theta_{t+1}) - PSNR(I_i, \theta_t)$$



Here I_i represents the i -th training image. θ_t denotes the NeRF’s MLP parameters before training on I_i , and θ_{t+1} are the NeRF parameters after training on I_i . Δ_{psnr_i} quantifies the immediate effect of training image I_i on the validation-based PSNR. These delta values are aggregated for each image for all epochs, excluding the first one, and the outcome contribution score is referred to as DV_{psnr} .

Training configuration: Our NeRF training setup adheres to a conventional structure where each training step samples rays from a single image. All training images are utilised once per epoch, with their sampling order shuffled at the beginning of each epoch. This configuration aligns with standard NeRF training practices, enabling systematic tracking of each image’s influence on validation performance. Although we measured each image’s impact using both L1 and PSNR, we experimentally found that PSNR consistently provided more reliable results. This design offers two key advantages for image valuation. First, per-image sampling allows for direct attribution of performance changes to specific inputs; Second, the randomized order across epochs help mitigate ordering bias and supports robust aggregation of impact scores.

Identifying fair contribution scores: Since the model’s parameters evolve throughout training, the impact of a given image depends on the model’s current state. To ensure a fair comparison across images, we considered three complementary strategies: (a) weighting the contribution of each image by training progress, (b) aggregating impact estimates across epochs, and (c) reverting the model to its pre-update state before measuring the effect of each training image. Option (a) does not provide a principled framework for defining a weighting function, as any choice of weight might introduce arbitrary assumptions not grounded in theoretical or empirical validation. Option (c) was found computationally intensive and unstable in practice, with measurements exhibiting high variance across runs with different random seed. Therefore, Option (b) was selected as the main approach due to its practicality and empirical stability.

Experimental setup and results: We evaluated our method on four scenes from the Phototourism dataset²²¹, which comprises internet-sourced phot collections of well-known landmarks. These scenes exhibit significant variation in viewpoint, illumination, transient objects and subjects, and appearance – posing challenges such as inconsistent lighting, occlusions due to crowds, and a wide range of camera parameters. This makes the dataset a strong benchmark for testing robustness in unconstrained, real-world settings.

To assess the consistency of DV_{psnr} , we focused on the Brandenburg Gate scene, which includes 1312 images. NeRF-w²²² allocated 763 images for training, 96 for validation, and 10 for testing, while the remaining images were excluded based on an automated quality assessment (NIMA)²²³. Additionally, we computed contribution scores for three other scenes – Sacre Coeur, Taj Mahal, Trevi Fountain – to evaluate how training NeRF-w with images prioritised by our method influences performance.

Reproducibility evaluation with correlation: A key goal of our approach was to design contribution metrics that remain stable regardless of the image loading order. As illustrated in Figure 47, DV_{psnr} scores demonstrate high consistency across multiple runs, with a correlation coefficient of 0.8. During training, the PSNR was evaluated at each step using a small validation set of 10 images.

²²¹ Jin, Y., Mishkin, D., Mishchuk, A., Matas, J., Fua, P., Yi, K.M., Trulls, E.: Image matching across wide baselines: From paper to practice. International Journal of Computer Vision 129(2), 517–547 (2021)

²²² Martin-Brualla, R., Radwan, N., Sajjadi, M.S., Barron, J.T., Dosovitskiy, A., Duck-worth, D.: Nerf in the wild: Neural radiance fields for unconstrained photo collections. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 7210–7219 (2021)

²²³ Talebi, H., Milanfar, P.: Nima: Neural image assessment. IEEE transactions on image processing 27(8), 3998–4011 (2018)



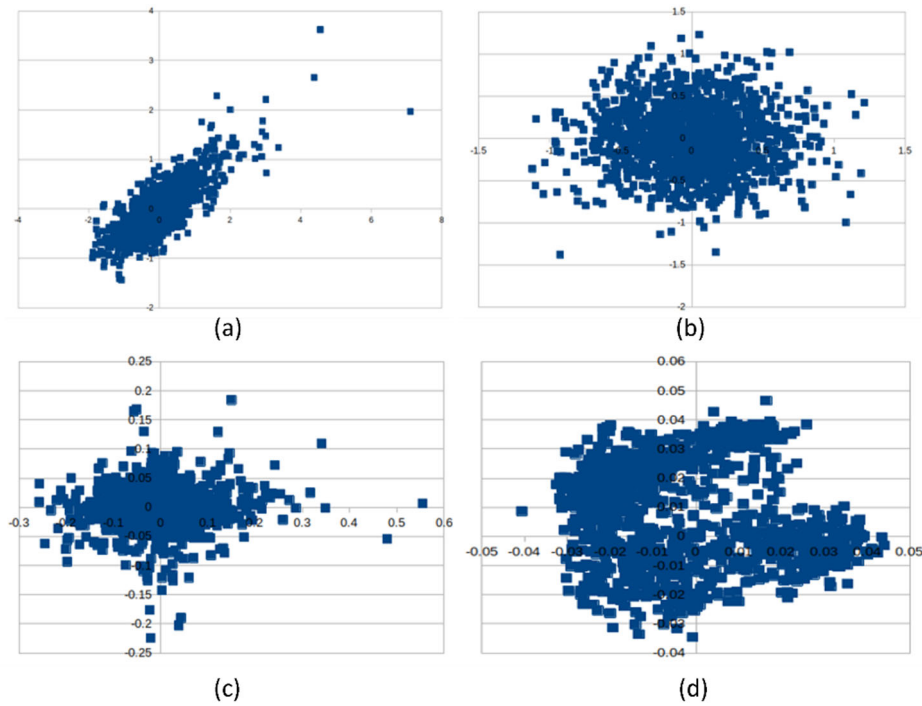


Figure 47: Contribution score correlations on the Brandenburg Gate dataset. The x-axis shows contributions scores from one training run, and the y-axis from a different run with a different seed. Each (x, y) point compares scores for the same image. (a) Aggregated PSNR difference, (b) Aggregated L1 difference, (c) Last epoch PSNR difference, (d) PSNR difference

Impact of Training set composition: Another central objective is to evaluate how different training subsets influence model performance. To efficiently compute contribution scores, the PSNR was recorded at each training iteration. For the Brandenburg Gate scene, 716 images were identified as positively contributing to PSNR improvements on the small validation set. To assess the effect of this selection, we compared two training sets using 43 held-out test images that were not involved in training or in the contribution score calculations. The training set selected by our method yielded higher PSNR on the 10-image validation set but lower performance on the 43-image test set, which contained more fine-grained details, as shown in Table 6.

Table 6: PSNR achieved for the same pipeline with the NeRF-w training set and with the training set selected by our data valuation approach.

PSNR	Brandenburg Gate		Sacre Coeur		Taj Mahal		Trevi Fountain	
	val	test	val	test	val	test	val	test
NeRF-w	19.27	17.72	16.94	15.62	17.93	16.22	17.51	17.33
DV_{psnr}	19.96	16.79	17.45	16.23	18.33	16.55	17.65	17.50

Additionally, as shown in Table 6, for the other three scenes, the training sets selected using our data valuation framework consistently outperformed those chosen by NeRF-w, achieving higher PSNR on the test sets. In these cases – Sacre Coeur (11 validation, 11 test), Taj Mahal (14 validation, 13 test), and Trevi Fountain (10 validation, 9 test), the validation and test sets are nearly equal in size, offering a more balanced and representative distribution of images.





Figure 48: Rendered images trained with DV_{psnr} selected training set and NeRF-w selected training set. (a-left) DV_{psnr} trained render of Brandenburg Gate - PSNR 19.51, (a-right) NeRF-w trained render of Brandenburg Gate – PSNR 17.59. (b-left) DV_{psnr} trained render of Sacre Coeur – PSNR 17.31, (b-right) NeRF-w trained render of Sacre Coeur – PSNR 17.03. (c-left) DV_{psnr} trained render of Trevi Fountain – PSNR 21.3, (c-right) NeRF-w trained render of Trevi Fountain – PSNR 20.7. (d-left) DV_{psnr} trained render of Taj Mahal – PSNR 24.38, (d-right) NeRF-w trained render of Taj Mahal – PSNR 22.9.

In contrast, for test views captured from more distant perspectives, the training set selected by DV_{psnr} delivers improved performance, as shown in Figure 48 (a-b). The model trained on our selected dataset more accurately reconstructs details – such as the statue atop the Brandenburg Gate - and achieves a higher PSNR. In the example test images from the Trevi Fountain (Figure 48c) sharper details are noticeable, particularly along the bottom.



Figure 49: Left: Image selected to be part of the training set by NeRF-w according to its NIMA score, while DV_{psnr} calculates a low score. Right: Image disregarded from NeRF-w's training set, while DV_{psnr} calculates a high score.

As illustrated in Figure 49, DV_{psnr} flags an image used in NeRF-w's training set despite containing many transient objects (left) – as potential detrimental. In contrast, it identified another image, which was excluded from NeRF-w's training, validation, and test sets, as highly valuable. Both images have distinct advantages and drawbacks: The left one is brighter and may reveal finer details in unoccluded regions, while the right one is darker with less clarity in some areas but features significantly fewer occlusions.

Discussion: We introduced DV_{psnr} , a data valuation method designed to quantify the contribution of individual training images to the final NeRF-based 3D reconstruction. The core ideal is to determine how



much each original image influences the resulting reconstruction, effectively tracing derivative aspects of the output back to specific inputs. We evaluated our method in comparison to NeRF-w, a state-of-the-art NeRF-in-the-wild approach, by systematically removing training images with low contribution scores and observing the impact on reconstruction quality.

A central focus of our study was the reproducibility of the contribution scores. Using PSNR as a performance metric, we demonstrated that the scores remained consistent across training runs with different random seeds – highlighting the method’s robustness. More importantly, we assessed the practical utility of the scores by using them to guide training set selection. This resulted in improved reconstruction quality in several benchmark scenes, underscoring the method’s potential to optimise data usage.

Beyond performance optimisation, our approach offers a pathway for analysing how much the final reconstruction constitutes a derivative work of specific images within a dataset. This opens up broader implications for understanding data provenance, intellectual property, and dataset influence in generative models.



9 Conclusion (JRS)

This document provides the final documentation of the work in WP3. It briefly summarises results already presented in D3.1, provides updates for components that have been developed further and presents components added in the second half of the project.

The Neural Media Repository (NMR) has advanced to a scalable multimodal search system, allowing multiple NMR instances (with potentially different visibility and permissions) to be grouped to one XReco platform instance. The connector infrastructure does not only allow to connect to external data sources (e.g., Wikimedia, Europeana, APIs of broadcast archives), but also enables a federated connection NMRs hosted at different XReco platform instances.

The content analysis services have been extended, in particular concerning cross-modal descriptors and few-shot object detection, and the analysis infrastructure provides also the basis for connecting some of the services developed in T4.1. Some of these analysis services are highly customisable, resulting in specifically fine-tuned AI models. The training and handling of these models is not fully covered in the current UI, but this approach has high potential for organisation or even project specific workflows.

The search services have been extended and harmonised, so that a single interface integrates the different types of content search. For the metasearch service, reranking of heterogenous results coming with different amount and granularity of metadata has been implemented in order to improve the relevance of search results.

The legal analysis has been updated based on emerging legislation such as the AI Act. In the first version of the deliverable, the question of the legal status of 3D assets created from 2D content using fully automatic reconstruction methods has been identified as a key question. This version of the document has thus deepened this analysis and provided a set of practical guidelines helping to assess whether such content could be considered an own creation or not.

Another result of the legal work is the licensing framework for the XReco marketplace, based on a CC-plus approach. The main goal is to allow the most relevant choices for content creators, while keeping the licensing framework as simple as possible. The user is guided with questions through the license selection process. The license is then created using the rights management tools, and notarised on the block chain.

This document provides also novel research on data valuation, which was initially intended to serve as an input to monetisation. However, based on the results of legal analysis, it was found not relevant for this purpose, but rather to provide guidance about the degree to which the resulting assets maintains the appearance of certain source content assets.

