

POLICY 01

EU Copyright Compliance Policy for MOTHER Foundation Models

A policy for complying with Union law on copyright and related rights, including the identification and honouring of machine-readable reservations of rights under the text-and-data-mining regime, as required of general-purpose AI model providers.

PROVIDER

Media Stream AI Limited (MSAI), United Kingdom

MODELS IN SCOPE

MOTHER CORE V2 & V3 (open-weight, Hugging Face) and all MSAI model training

LEGAL BASIS

Reg. (EU) 2024/1689 Art. 53(1)(c); Dir. (EU) 2019/790 Arts. 3–4

VERSION

1.0 · Effective [effective date]

1. Purpose & commitment

Media Stream AI Limited (MSAI) develops the MOTHER family of foundation models in accordance with Union copyright law. This policy sets out the measures MSAI applies across the model lifecycle to ensure that text and data mining (TDM) for training is carried out only on lawfully accessible content and in a manner that honours rights holders’ reservations of rights. It applies to MOTHER CORE V2 and V3 (published as open-weight models) and to all subsequent MSAI training runs.

2. Legal framework

- **AI Act Art. 53(1)(c)** — obligation to put in place a policy to comply with Union copyright law, in particular to identify and comply with a reservation of rights expressed under Art. 4(3) of Directive (EU) 2019/790.
- **DSM Directive Art. 3** — the TDM exception for scientific research by research organisations and cultural-heritage institutions (no opt-out).

- **DSM Directive Art. 4** — the general TDM exception, which applies *only where rights have not been expressly reserved* by the rights holder in an appropriate machine-readable manner for content made publicly available online.

3. Operative commitments

3.1 Lawful access

MSAI mines only content to which it has lawful access — publicly available web content, openly licensed datasets, data licensed to MSAI, and MSAI’s own or contributed data. MSAI does not circumvent access controls or paywalls to acquire training data.

3.2 Honouring reservations of rights (opt-outs)

MSAI identifies and complies with machine-readable reservations of rights under Art. 4(3). MSAI’s crawling and data-preparation pipeline detects and excludes content whose rights have been reserved through the following mechanisms:

MECHANISM	HOW MSAI HONOURS IT
robots.txt	The MSAI crawler obeys Disallow directives and AI-specific user-agent rules; disallowed paths are not fetched for training.
TDM Reservation Protocol (TDMRep)	W3C TDMRep signals — tdm-reservation HTTP headers, .well-known/tdmrep.json, and in-page metadata — are parsed and reserved content is excluded.
ai.txt / AI-usage signals	AI-usage declaration files and equivalent site-level signals are respected where present.
Metadata & rights headers	Machine-readable rights metadata (in HTML meta tags, media metadata, and HTTP headers) indicating a TDM reservation is honoured.
Per-work licences	Where a work carries a licence prohibiting TDM/AI training, it is excluded even if otherwise publicly accessible.

Reservation detection is applied at acquisition time and re-checked at dataset-assembly time; excluded sources are logged.

3.3 Identifiable crawler

MSAI operates its training crawler under a documented, stable user-agent string and publishes the information rights holders need to apply a reservation effective against MSAI. Rights holders may reserve rights prospectively against this agent. Crawler identifier: `[MSAI crawler user-agent]`.

3.4 Licensed & permitted data

Where MSAI uses licensed datasets or data obtained under contract, it retains records of the licence terms and mines such data only within the scope those terms permit. Where MSAI uses publicly available third-party datasets, it relies on the terms under which those datasets are published.

3.5 Removal on substantiated request

MSAI operates a rights-holder contact channel. On a substantiated request or notice, MSAI will (a) exclude the identified content from future training runs, (b) record the request, and (c) where technically and legally appropriate for models still in development, take reasonable steps regarding affected material. Requests may be sent to the contact point in §6.

3.6 Downstream information

For open-weight releases, MSAI provides model documentation referencing this policy so that downstream deployers are aware of the copyright-compliance measures applied upstream.

3.7 Records & the training-content summary

MSAI maintains internal records of data sources and the reservation-handling measures applied, and publishes the public summary of training content required by Art. 53(1)(d) (companion document).

4. Governance

- **Policy owner:** `[Head of AI Governance / DP0]`, Media Stream AI Limited.
- **Review cadence:** reviewed at least annually and on any material change to data-acquisition practice, to the AI Act GPAI Code of Practice, or to European AI Office guidance.
- **Enforcement readiness:** the AI Office may request information and enforce the GPAI provisions; MSAI maintains the records

needed to respond.

5. Scope of application

This policy applies to all data-acquisition and text-and-data-mining activity undertaken by or on behalf of MSAI for the training of the MOTHER foundation models, including the publicly available datasets, licensed data, and MSAI-generated and internally curated data described in the companion training-content summary.

6. Contact for rights holders

Reservations of rights, licensing enquiries and removal requests: legal@mediastreamai.com, Media Stream AI Limited, United Kingdom.