Tengr.ai Model Summary

1. Scope and Purpose

This document ("**Summary**") is issued by:

Company Name: Tengrai Artificial Intelligence Korlátolt Felelősségű Társaság

Headquarters: Hungary, 6724 Szeged, Rókusi-boulevard 21, 1st floor, door 4

Company Registration Number: 06-09-028918

Tax number: 32315028-2-06

("**Provider**")

for the public and the use of competent supervisory authorities within the meaning of Regulation (EU) 2024/1689 on Artificial Intelligence (the "**AI Act**") and any other applicable statutory instrument governing the placing on the market, or commissioning, of general-purpose artificial-intelligence systems capable of generating images.
Its purpose is to set out, in a legally cognisable manner, (i) the technical characteristics of the Provider's general-purpose image-generation model (the "**System**"), (ii) the data governance measures applied during training, and (iii) the ex-ante and ex-post risk-management and redress mechanisms that ensure the System's continuous conformity with European Union and Member-State law.

2. Definitions

For the avoidance of doubt, the following capitalised terms shall have the meanings assigned to them below. All other terms shall bear the meanings attributed in the AI Act unless the context otherwise requires.

- **Hyperalign Architecture** – the hidden alignment, constraint-enforcement and policy-verification subsystem embedded in the System, described in Section 5.
- **Policy Store** – the machine-readable repository of binding legal, ethical and contractual rules consulted by Hyperalign.
- **Public-Source Dataset** – any image-text dataset released under a licence permitting commercial use, modification, and redistribution without royalty, including but not limited to those enumerated in Section 4.

- **Disallowed Material** – any visual content the generation, possession or dissemination of which would infringe Union or Member-State law or the Provider's own acceptable-use policies.

3. High-Level System Description

3.1 General Function

The System is a diffusion and transformer-based text-to-image and image-to-image model that, upon receipt of an arbitrary natural-language prompt, produces a corresponding raster image. The System is accessed directly on the https://tengr.ai web application or through third parties via API.

3.2 Functional Blocks

1. **Prompt Reception (S1)** – prompts are accepted indiscriminately; no pre-filter is applied.
2. **Primary Generation (S2)** – a base diffusion model constructs a provisional solution (O_{raw}).
3. **Hidden Constraint Enforcement (S3)** – Hyperalign evaluates O_{raw} against the Policy Store and, where necessary, deletes, replaces or regenerates non-compliant segments.
4. **Delivery (S4)** – the transformed image (O_{final}) is transmitted to the user; no refusal text, apology or censorship notice is ever surfaced.

4. Training-Data Provenance and Licensing

4.1 Source Datasets

The Provider certifies that the image corpus utilised for pre-training and fine-tuning *exclusively* comprises Public-Source Datasets and licensed Private-Source Datasets (collectively, the "**Training Set**").

The Provider attests that the Training Set consists exclusively of image–text corpora that are **lawfully and openly distributed** under permissive copyright instruments, namely Creative Commons Attribution licences (CC BY 4.0, CC BY 2.0), public-domain waivers (CC0, Public-Domain Mark), or equivalent irrevocable custom licences granting worldwide, royalty-free rights of use, modification and commercial redistribution. Representative sources include, inter alia: the MS COCO reference benchmark (CC BY 4.0); Google's Open Images collection (images predominantly CC BY 2.0); the

Conceptual Captions and Conceptual 12M web-alt-text corpora (images ≥ CC BY 2.0); the 99-million-item YFCC-100M multimedia archive released under a spectrum of Creative Commons variants; and the Wikimedia Commons, Smithsonian Open Access and other public-domain repositories whose content is either CC-licensed or free of re-use restrictions.

Prior to ingestion, the Provider conducted automated licence-tag verification and eliminated every record bearing a **Non-Commercial**, **No-Derivatives** or otherwise incompatible encumbrance, thereby ensuring that every retained image satisfies the **AI Act Article 10(3) requirement** for "relevant, representative, free-of-error and complete" data collected **in compliance with Union law on intellectual-property rights**. Comprehensive download manifests, cryptographic checksums and licence details are preserved in the Dataset Register and may be made available to the competent authorities upon reasoned request.

4.2 Exclusion of Proprietary Content

No images obtained from closed social-media platforms, privately hosted repositories, or otherwise encumbered sources were ingested. No scraping tools bypassing technical protection measures were employed.

5. Data Due-Diligence, Cleansing and Risk Mitigation

Given that the Training Set exceeded seven (7) billion candidate images, manual per-item review would have required in excess of 300 years at one (1) second per image; such review was therefore manifestly impracticable. The Provider accordingly adopted the multi-layer automated procedure set out below:

1. **Metadata Pre-Screening** – textual metadata (captions, alt-text, URLs) were scanned for lexical indicators of obscenity, extremist ideology or the sexual exploitation of minors; flagged entries were discarded at source.
2. **NSFW Visual-Classifier Pass** – a ResNet-based classifier, fine-tuned on the "Journey NSFW" and "Open NSFW" benchmarks, automatically flagged any image scoring above a 0.25 probability of sexual content or graphic violence.
3. **Perceptual-Similarity Filtering** – CLIP embeddings were k-nearest-neighbour-compared against a *block-listing* gallery of disallowed exemplars (e.g. known extremist insignia, CSA material); hits were expunged.

4. **Semantic-Similarity Filtering** – captions were embedded via SBERT; cosine similarity to a curated vector bank of illicit semantics triggered exclusion.
5. **Deduplication and Quality Control** – near-duplicates (cos < 0.999) and corrupted or blank files were removed.
6. **Iterative QA Team Refinement** – periodic audits located residual infractions; the corresponding heuristic or model weights were re-trained and the pipeline rerun.

The foregoing measures resulted in the elimination of millions of images, approximately 3.8% of total candidate images, including all entries flagged as high-risk via LAION's own tags.

6. Hyperalign Constraint-Enforcement Mechanism

6.1 Legal and Technical Basis

Hyperalign constitutes an alignment module (Component 20) that supervises content generation in real time by reference to the Policy Store . Enforcement is performed *token-by-token* for diffusion latents and, where required, by post-generation in-painting. The architecture thereby:

- **accepts every prompt without exception**
- **suppresses any observable refusal signal** and
- **delivers only compliant output**, thereby collapsing the success probability of gradient-free jailbreak attempts to the failure probability of the alignment classifier.

6.2 Extensibility and Human Oversight

The Policy Store is maintained in YAML format, version-controlled, and amendable by compliance personnel without necessitating retraining of the core model. Consequently, amendments to the AI Act, forthcoming delegated acts, or Member-State-specific prohibitions can be incorporated within a maximum of ten (10) working days, after which the revised rules bind all subsequent generations.

7. Ongoing Governance, Monitoring and Accountability

1. **Logging and Traceability** – a tamper-evident audit log records each prompt, the SHA-256 hash of O_{final}, and the precise policy clauses invoked during any alteration.

2. **Periodic QA Exercises** – at least semi-annually the Provider conducts adversarial testing against emergent jailbreak vectors (e.g., multilingual obfuscation).
3. **Incident Response** – suspected policy-enforcement failures trigger an immediate suspension of the affected endpoint and a root-cause analysis within forty-eight (48) hours.
4. **User Feedback Channel** – any visitor may file a notice of alleged unlawful output on our website (link in the footer of Tengr.ai); substantiated claims compel retraining or policy-file amendment.

8. Conformity with the AI Act and Related Instruments

- **Article 4 (Risk Management)** – Hyperalign provides ex-ante and ex-post controls ensuring that the residual risk is acceptable in light of the System's intended generative purpose.
- **Article 10 (Data and Data Governance)** – the Provider maintains full documentation of dataset provenance, licence terms, and cleansing methodology as summarised in Section 4.
- **Article 11 (Technical Documentation)** – this Summary, together with detailed internal design dossiers, fulfils the obligation to furnish technical documentation to regulators.
- **Article 15 (Accuracy, Robustness and Cyber-Security)** – empirical testing demonstrates materially lower jailbreak-success rates relative to visible-refusal architectures. Hyperalign is embedded to mitigate prompt-injection and insecure-output-handling vectors recognised by OWASP LLM02.
- **Article 20 (Post-Market Monitoring)** – governance measures in Section 7 constitute the Provider's post-market monitoring plan.
- **Guidelines for Secure AI System Development (UK NCSC / CISA)** – the System's silent, policy-aware output filtering accords with the non-interactive defence pattern endorsed therein.

9. Limitations and Disclaimer

Nothing herein shall be construed as a warranty of absolute censorship or error-free performance. While Hyperalign materially mitigates unlawful output, residual risk cannot be excluded. The Provider shall, however, employ best endeavours to rectify any substantiated deficiency forthwith upon notice.

10. Contact details of the AI Compliance Team:

Website: https://tengr.ai (direct form opens from the footer link)
E-mail address: qa@tengrai.com
Postal address: Hungary, 6724 Szeged, Rókusi-boulevard 21, 1st floor, door 4