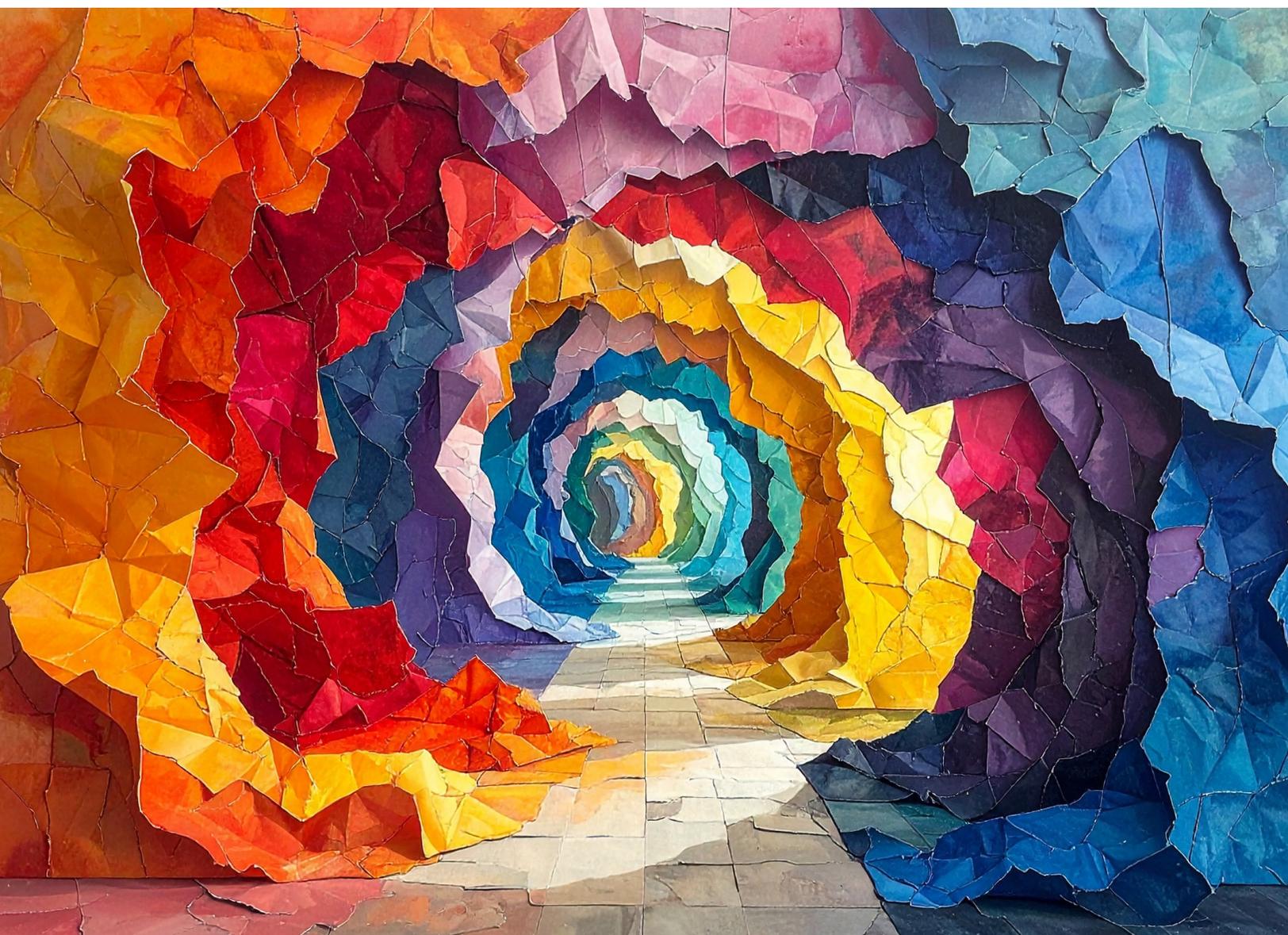




PUBLIC SUMMARY OF TRAINING CONTENT

Adobe Firefly Image Model 5

November 2025



Public Summary of Training Content for General-Purpose AI Models

for Adobe Firefly Image Model 5 as required by Article 53 (1)(d) of Regulation (EU) 2024/1689 (AI Act)

General Information	
Provider identification	
Provider name and contact details	Adobe Inc. 345 Park Avenue, San Jose, CA 95110
Authorised representative name and contact details	Adobe Systems Software Ireland Limited Address: 4-6 Riverwalk, Citywest Business Campus, Saggart Dublin 24, D24 DCW0, Dublin, Ireland Email: euaiactauthorisedrep@adobe.com
Model identification	
Versioned model name(s)	Firefly Image Model 5
Model dependencies	N/A
Date of placement of the model on the Union market	October 28, 2025
Modalities, overall training data size and other characteristics	
Modality	Image and Video
Training data size	<ul style="list-style-type: none">• More than 1 billion images• 10,000 to 1 million hours of video
Types of content	<ul style="list-style-type: none">• Image — Photography, illustrations, vector images, multi-layered images• Video — Stock footage
Latest date of data: acquisition/collection for model training	September 2025
Description of the linguistic characteristics of the overall training data	The visual data include character and text elements from multiple languages.
Other relevant characteristics of the overall training data	N/A

Additional comments (optional)	N/A
--------------------------------	-----

List of Data Sources

Publicly available datasets

Have you used publicly available datasets to train the model?	Yes
If yes, specify the modality(ies) of the content covered by the datasets concerned	Image
List of large publicly available datasets	OpenImages v7 https://storage.googleapis.com/openimages/web/index.html OpenImages v7 is a large dataset of permissively licensed images with annotations.
General description of other publicly available datasets not listed above	Other datasets subject to permissive licenses.

Private non-publicly available datasets obtained from third parties

Datasets commercially licensed by rightsholders or their representatives

Have you concluded transactional commercial licensing agreement(s) with rightsholder(s) or with their representatives?	Yes
If yes, specify the modality(ies) of the content covered by the datasets concerned	Image and video

Private datasets obtained from other third parties

Have you obtained private datasets from third parties that are not licensed as described in Section 2.2.1, such as data obtained from providers of private databases, or data intermediaries?	Yes
If yes, specify the modality(ies) of the content covered by the datasets concerned	Image and video
If publicly known, list private datasets obtained from other third parties	N/A

Data crawled and scraped from online sources	
Were crawlers used by the provider or on behalf of?	No
Purposes of the crawler(s)	N/A
General description of crawler behaviour	N/A
Period of data collection	March 2023 – May 2024
Comprehensive description of the type of content and online sources crawled	Adobe did not crawl online sources. Instead, we searched specifically for content licensed under CCO (license dedicating content to the public domain) or in the public domain in limited amounts within select sites — such as photo hosting sites, educational websites, and museum websites — that allowed for content downloading.
Type of modality covered	Image
Summary of the most relevant domain names crawled	Flickr.com
User data	
Was data from user interactions with the AI model (e.g. user input and prompts) used to train the model?	No
Was data collected from user interactions with the provider's other services or products used to train the model?	No
Synthetic data	
Was synthetic AI-generated data created by the provider or on their behalf to train the model?	Yes
If yes, modality of the synthetic data	Image
If yes, specify the general-purpose AI model(s) used to generate the synthetic data if available on the market	Firefly Image Model 5
Information about other AI models, including provider's own AI model(s) not available on the market, used to generate synthetic data to train the model to which this Summary applies	Internal Adobe model
Other sources of data	
Have data sources other than those described in Sections 2.1 to 2.5 been used to train the model?	No

Data Processing Aspects

Respect of reservation of rights from text and data mining exception or limitation

Are you a Signatory to the Code of Practice for general-purpose AI models that includes commitments to respect reservations of rights from the TDM exception or limitation?

No

Describe the measures implemented before model training to respect reservations of rights from the TDM exception or limitation before and during data collection, including the opt-out protocols and solutions honoured by the provider or, as applicable, by third parties from which datasets have been obtained

N/A. TDM exception was not relied on for training data for this model.

Removal of illegal content

General description of measures taken

Adobe has a policy to scan all visual content used for AI training with classifiers to identify and remove CSAM. In addition, Adobe uses a combination of automated classifiers along with human moderation designed to ensure that content used for training meets Adobe's policies around harmful content (see <https://www.adobe.com/trust/transparency/content-policies/harmful-content.html> and <https://www.adobe.com/trust/transparency/content-policies/content-rights.html>).