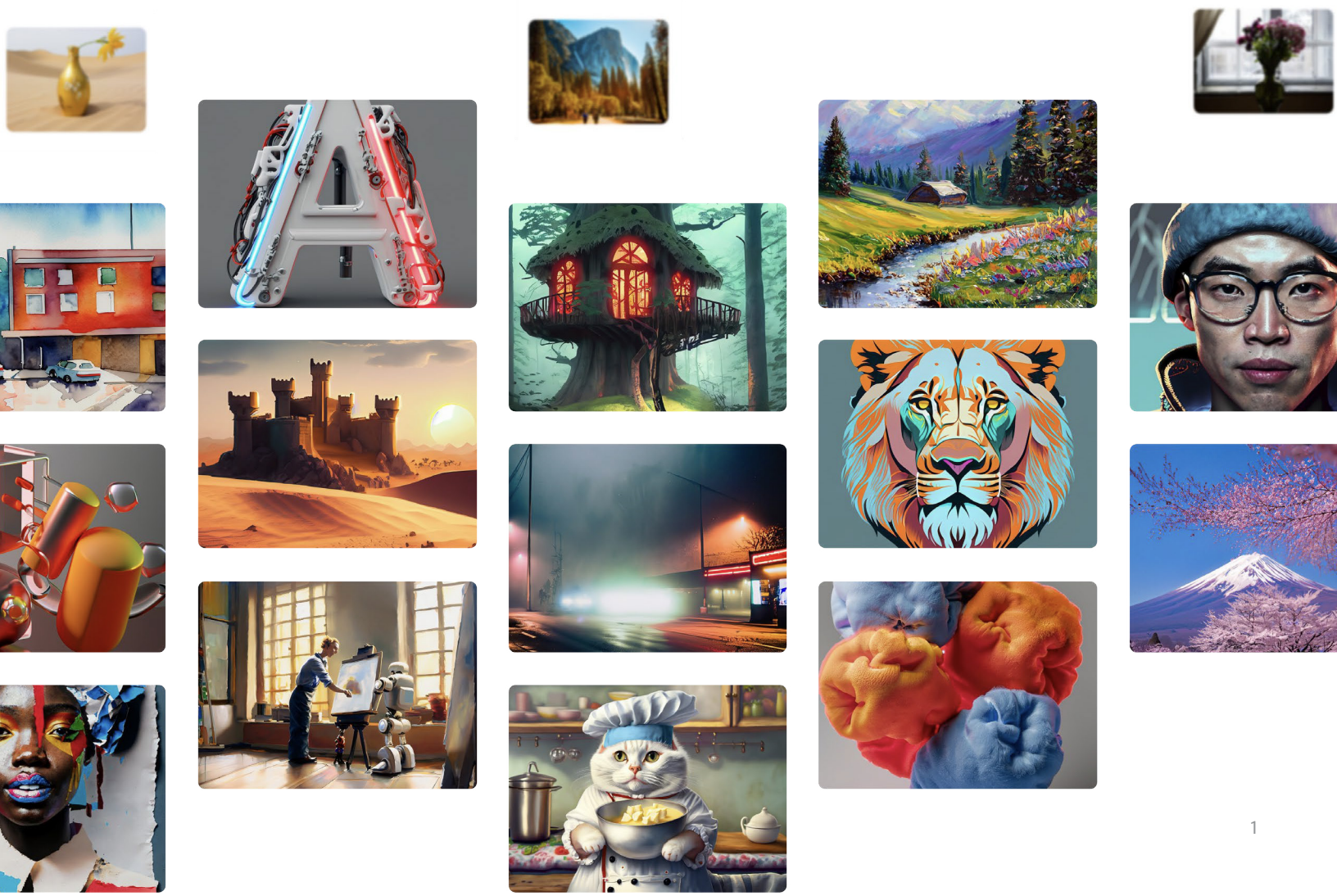
Adobe

SECURITY FACT SHEET

Adobe Firefly Custom Models

March 2025



About Firefly Custom Models

Firefly Custom Models enables brands to fine-tune the Firefly generative AI foundational model by training on their signature brand style, campaign style, character, or object. Using Custom Models, organizations can consistently create on-brand creative assets at scale, transforming their style or subject to explore new ideas, visualize different surroundings, generate innovative content, and tailor content to specific segments.

Enterprise Access and Controls

Customers can manage user access to Firefly Custom Models via the Adobe Admin Console. The customer's Adobe Admin assigns users to the Trainer role in the console, which allows a user to train or fine-tune Firefly models with their brand assets on the Firefly web app (firefly.adobe.com). While users with the Trainer role may share custom models to collaborate with other users, the customer's Adobe Admin can restrict custom models from being shared outside of the organization.

Users can only generate assets or images using Firefly Custom Models if they've been explicitly invited to do so by an existing collaborator via the Share menu or if the custom model belongs to a "project" to which they have been given access. For more information, please see custom model projects.

Model Training Security Architecture and Data Flow

Figure 1 below illustrates the data flow when a user with a Trainer entitlement initiates a model training activity:

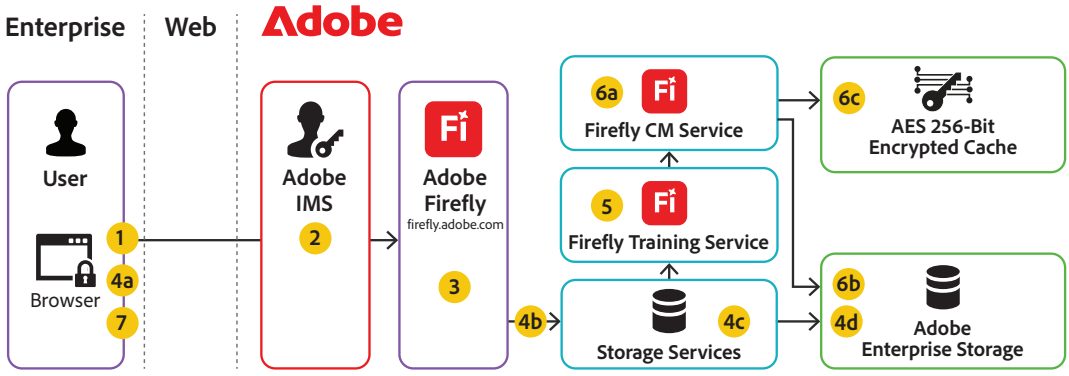


Figure 1: Firefly Custom Models model training data flow diagram

Step 1: In a web browser, the user signs into the Firefly web app using their Trainer credentials and selects "Custom models."

Step 2: Adobe Identity Management Services (IMS) validates the user and their Trainer entitlements. IMS returns a user token that authenticates the Trainer with both the Firefly Custom Models Service and the Firefly Training Service.

Step 3: The Trainer selects "Train a New Model" in the Firefly web app along with one of two training modes:

- **Style** trains the model on the colors, shapes, and background aesthetic.
- **Subject** trains the model on an object or character.

The Trainer also selects an existing or creates a new project folder in which to store the custom model.

Note: The Trainer can only store a custom model in a project to which they have appropriate permissions. For more information, please see Manage Custom Model Project Access.

Step 4: The Trainer (4a) selects at least 10 images they want to use to train the custom model. These images are uploaded to Storage Services, which (4b) scans the images for viruses, (4c) encrypts the images using a customer-managed encryption key (CMK), and (4d) stores the images in the organization's AES 256-bit encrypted S3 bucket.

Step 5: The Firefly Training Service uses the uploaded images to train a custom model. Once the training is complete, the model will appear on the "Your models" page in the Firefly web app with a Ready status.

Step 6: The Firefly Custom Models Service (a) creates the custom model and stores it, along with the delta weights and other metadata in both (b) the organizational storage for long-term storage (as noted in Step 4) and (c) in an encrypted cache to speed review and use of the model.

Step 7: The Trainer can preview and test the custom model to confirm it matches their intention before publishing it (described in the next section) The Trainer can also share the custom model with other users by granting them "Review" access, which enables others to preview and test the custom model and validate that the output matches the intention.

Model Testing Security Architecture and Data Flow

Figure 2 illustrates the data flow when a user who has (a) Trainer entitlement and has created a new custom model or (b) been granted Edit or Review access to a model initiates a model testing activity:

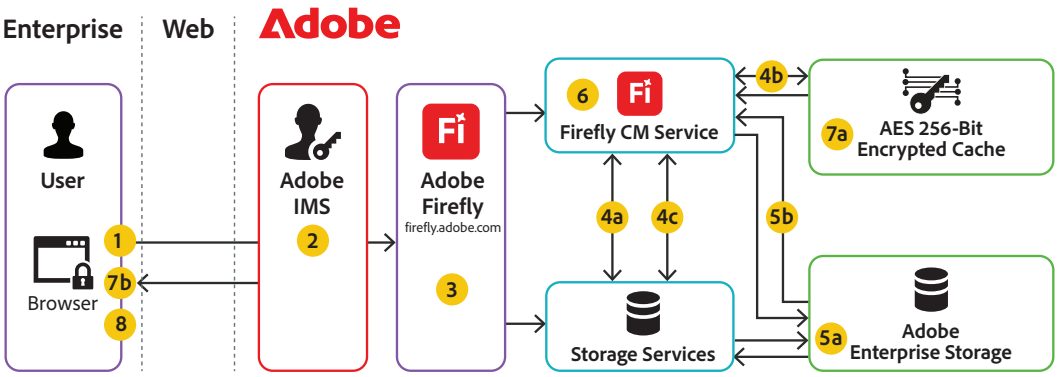


Figure 2: Firefly Custom Models model testing data flow diagram

Step 1: In a web browser, the user signs into the Firefly web app using their credentials and selects "Custom models."

Step 2: Adobe Identity Management Services (IMS) validates the user and their entitlements. The Firefly Custom Models Service relies on the IMS token to authenticate and authorize a user.

Step 3: The user opens a custom model and selects “Preview and test” in the Firefly web app along with one of the two available training modes.

Step 4: The Firefly Custom Models Service (4a) checks user permissions with Storage Services and requests the delta weights from (4b) the encrypted cache or (4c) Storage Services, if the cache has expired.

Step 5: If the cache has expired, Storage Services (a) retrieves and decrypts the delta weights from Adobe Storage for business using the customer’s managed encryption key (CMK) and (b) sends the content and metadata to the Firefly Custom Models Service.

Step 6: The Firefly Custom Models Service generates images by combining the delta weights with the base Firefly model for inference.

Step 7: The generated images are (a) temporarily stored in an application-managed cache storage and (b) a pre-signed URL for the cached images is returned to the Firefly web app.

Step 8: When the custom model is ready to use, the user can publish it and share it with other collaborators.

Content Generation Security Architecture and Data Flow

Figure 3 illustrates the data flow when a user generates new content with a custom model:

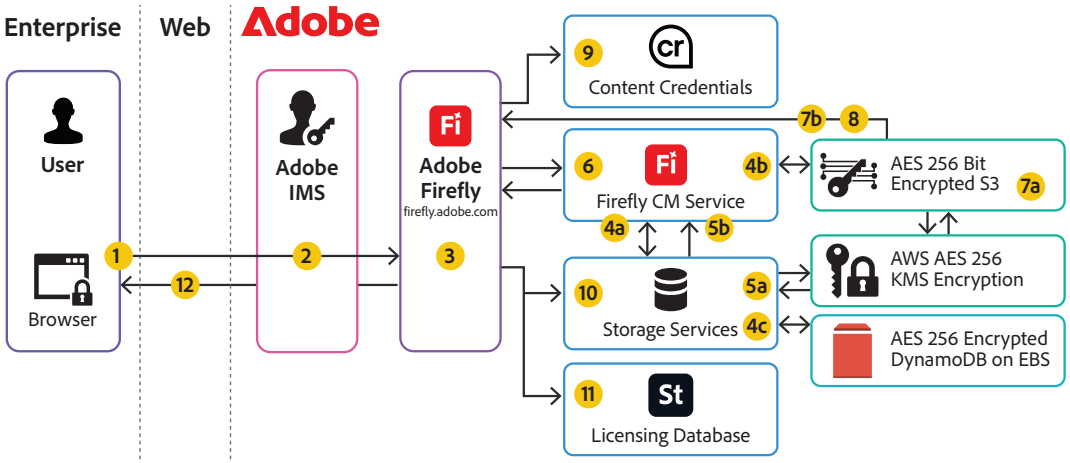


Figure 3: Firefly Custom Models content generation data flow diagram

Step 1: In a web browser, the user signs into the Firefly web app using their credentials.

Step 2: Adobe Identity Management Services (IMS) validates the user and their entitlement for Firefly Custom Models as well as their permissions to access certain custom models.

Step 3: The user initiates a "Text-to-Image" workflow and selects a custom model from the "Your Models" drop-down list, which also populates the prompt field with user-supplied sample prompt.

Step 4: The Firefly Custom Models Service (a) checks user permissions with Storage Services and requests the delta weights (b) from the encrypted cache or (c) from Storage Services, if the cache has expired.

Step 5: If required, Storage Services (a) retrieves and decrypts the delta weights using the customer's managed encryption key (CMK) and (b) sends the content and metadata to the Firefly Custom Models Service.

Step 6: The Firefly Custom Models Service generates images by combining the delta weights with the base Firefly model for inference.

Step 7: The generated images are (a) temporarily stored in an application-managed cache storage and (b) a pre-signed URL for the cached images is returned to the Firefly web app.

Step 8: If the user performs an action on the generated images such as "Download," "Save to library," "Copy Image," or "Edit in Adobe Express," the Firefly web app loads the generated output from that pre-signed URL.

Step 9: The Firefly web app attaches a Content Credentials manifest to the downloaded, saved, or edited image and saves this manifest to the Content Credentials cloud. For more information on Content Credentials, please see the "Content Credentials" section below.

Step 10: If the user chooses "Save to library," or "Edit in Adobe Express," the Firefly web app sends the image and the attached Content Credentials to the respective application, which stores it in user-managed storage through Storage Services.

Step 11: If the enterprise has purchased output indemnification, the Firefly web app sends a full resolution copy of the generated image with embedded Content Credentials to the Licensing Database.

Step 12: If the user chooses to "Download" the image, the Firefly web app sends the image and the attached Content Credentials to user's desktop.

Content Credentials

Adobe automatically generates Content Credentials for certain Firefly-generated assets to help provide transparency that the asset was created using Generative AI. Content Credentials typically contain the following metadata:

- In certain cases, a thumbnail of the generated image
- The tool/tools used to generate the asset
- Whether the asset was completely generated by Firefly or combined with other content
- Summaries of the type of actions taken in Firefly (such as use of a reference file, edit activity, etc.
- A cryptographic hash of the image and its metadata in a verifiable, tamper-evident signature that provides proof that the image and metadata have not been altered. The cryptographic hash is irreversible.

Content Credentials are attached to the exported asset file and stored in the Content Credentials cloud repository, which allows recovery of the Content Credentials in the event it is stripped from the exported asset.

Note: Text prompts are never included in any automatically generated Content Credentials.

Content Storage and Processing

Uploaded images used to train the Custom Models and the associated delta weights are stored in Adobe storage for business, which is a secure cloud storage hosted in Amazon Web Services (AWS) data centers. (see "Storage Services" in data flow narrative above). More information on input and output storage and processing can be found in the Adobe Firefly for enterprise security fact sheet.

Note: Adobe does not train the foundation Firefly generative AI models on any Creative Cloud subscriber's personal content.

User Identity Information

Adobe uses named user licensing to uniquely identify users of any Adobe product, including Custom Models. Custom Models is fully integrated with Creative Cloud for Enterprise identity access and management using Adobe Identity Management Services (IMS), allowing multi-factor authentication (MFA) to any SAML2- compliant provider.

More information on named user licensing can be found in the Adobe Identity Management Services Security Overview.

Data Storage Locations

Figure 3: Firefly Custom Models data storage locations

Reference images uploaded for Firefly Custom Models, as well as corresponding delta weights and other metadata, are stored in the customer's assigned regional data center in US-East, EMEA-West, or APAC.

Adobe currently processes, caches, and stores additional Firefly input content (such as Generative Match reference images) in Amazon Web Services (AWS) data centers in the US-East and US-West regions, regardless of the user's location.

Data Types and Retention

In addition to the data types and retention periods noted in the Firefly for enterprise security fact sheet, Adobe also stores data to identify the model(s) used in inferencing and the associated delta weight/s. This data is stored permanently in the customer's Adobe-managed Enterprise Storage as well as cached for two (2) weeks in AES 256-bit encrypted S3 storage for performance reasons.

Testing

Adobe teams rigorously test our generative AI products to reduce the potential for biased and harmful outcomes. For more information on the development and testing processes for our generative AI solutions, please see the Generative AI Built for Business solution brief. For the annual Security Testing Report for Adobe Firefly, please see the Adobe Firefly Security Testing Report (NDA required).

Conclusion

If you have any additional questions about the security posture and capabilities of Firefly Custom Models, please contact your Adobe account manager. For all other questions about Adobe's security programs and processes and compliance certifications, please see the Adobe Trust Center.