

September 27, 2024



NEWS ANALYSIS

Adobe Creates Ethical, Responsible AI Offerings

Adobe Enhances Firefly, a Family of Generative AI Models Introduced Last Year With a Responsible, Commercially Safe AI Approach



Andy Thurai
VP & PRINCIPAL ANALYST



Produced exclusively for Constellation Research clients

TABLE OF CONTENTS

Executive Summary3

Adobe Firefly: Adobe’s Family of GenAI Models for Image Generation4

Adobe Does Responsible Innovation in the Age of GenAI7

As a Founding Member, Adobe Forms the Content Authenticity Initiative and Coalition for Content Provenance and Authenticity 10

The Bottom Line 13

Endnotes 15

Analyst Bio. 16

About Constellation Research 17

EXECUTIVE SUMMARY

A few select software vendors are taking a thoughtful approach to ethical and responsible generative artificial intelligence (GenAI) offerings, and Adobe is one of them.

With its set of GenAI offerings, first introduced in March 2023, Adobe is looking to establish a major presence in the AI area—GenAI in particular—and differentiate itself from competitors in two ways: by embedding GenAI capabilities directly into existing creative, marketing, and operations teams' tools and workflows instead of offering them as stand-alone solutions; and by making its family of GenAI models commercially safe.

Despite rising investments in AI by enterprises, relying on what is delivered by AI can be a hit-or-miss process. For people to trust AI, it needs to be reliable, unbiased, ethical, accountable, and able to provide transparent outputs that will help instill confidence in those systems over time.

Adobe has introduced many AI-based tools across its platforms. Adobe Firefly is its family of creative GenAI models for content and media. The first of its kind from Adobe, it has generated more than 9 billion images and counting, according to the company's Q2 2024 earnings report. Adobe Firefly brings a governed, predictable, and secure GenAI to the unpredictable world of GenAI applications. Adobe has done a lot of work in ethical AI, responsible AI, and the ability to provide a transparent and validated approach to content provenance.

This research report evaluates Adobe's Firefly initiatives. It identifies key differentiators, examines functional capabilities, considers the offering's strengths and weaknesses, and suggests use cases. Technology buyers looking to implement solutions powered by Adobe Firefly should use this report along with other Constellation Research reports to evaluate their maturity and readiness for such applications.

BUSINESS THEMES



New C-Suite



Digital Marketing &
Sales Effectiveness



Technology
Optimization



Digital Safety and Privacy



Data to Decisions



Future of Work



Next-Generation
Customer Experience

ADOBE FIREFLY: ADOBE'S FAMILY OF GENAI MODELS FOR IMAGE GENERATION

GenAI is a form of AI that can “generate” new content such as text, images, video, audio, code, and even music. These foundation models, which are the underlying technology used in creating new content, are trained on large amounts of data to teach the machines how to achieve their intended generative goal (such as creating imagery).

Adobe released its own GenAI family of models, Firefly, in a beta version in March 2023 with general availability in August 2023. Adobe Firefly is a diffusion model initially designed to create images and text effects from text prompts. The company continues to innovate, having already introduced Image Model 3 with enhanced photorealism and image quality and new enterprise capabilities, such as Custom Models and Firefly Services.

Here's a rundown of some key capabilities in Adobe Firefly:

- **Text to Image.** This makes it possible to generate images from a simple text description (or prompt). For example, the prompt “A male lion with a fierce-looking face, mouth open showing teeth, on a cliff over an ocean with waves crashing on rocks, golden sunset, high definition, shot from above, good depth of field” produces the image shown in Figure 1. Although there are many AI tools that can produce images from prompts, Adobe takes a unique and responsible approach by automatically attaching Content Credentials and Firefly outputs, tamper-evident metadata¹ that acts like a “nutrition label,” enabling creators to attach valuable information about the content and their creative process—information that can be viewed later using the Content Credentials Verify website.²
- **Accessible via APIs.** Firefly Services are a comprehensive set of GenAI and creative APIs that automate workflows. Creating hundreds of asset variations, personalization, and localization for different customer segments, channels, and regions can all be done at scale with Firefly Services

Figure 1. Firefly-Produced Image Based on Text Prompting



Source: Adobe, Firefly Image Model 3

via APIs. All of the Firefly capabilities such as text to image, generative expand, generative fill, generative match, background removals, Photoshop actions, product cropping, and creating product masks can be accessed via APIs.

- **Custom Models and Style Kits for brand compliance.** Another unique feature of Adobe Firefly Services are Custom Models and Style Kits. Custom Models enable enterprises to train and customize models using enterprise-specific styles and subjects, which can then generate images at scale, and on demand, that are specific to their products and brand—including campaign styles and backgrounds, objects, characters, or even subjects. This can be very useful, especially if the work is outsourced to people outside their organization, because the desired output can be controlled by providing overall guidance and control. It also makes it easier to ensure brand consistency across

creative and marketing teams. Style Kits allow them to save templates, files, and generative presets to ensure brand consistency.

- **Generative Fill and Generative Expand.** This is an area in which Adobe, with its Photoshop picture-manipulation tools, excels. Without AI assistance, it requires manual work to do such photo manipulation, which could take hours. The use of AI has automated and reduced these repetitive and manual tasks to a matter of minutes. AI editing has added much more value in this area. Adobe's Generative Fill and Generative Expand features allow users to add or remove content from any image with simple text prompts.

Generative Expand allows for clicking and dragging to expand the image borders to seamlessly fill the expanded canvas with new, matching content. Generative Fill enables users to create dozens of new content concepts quickly. Integrated into Photoshop, both Generative Fill and Generative Expand allow users to make quick but powerful edits and additions to images, and then get them to the next level with the Adobe Creative toolset.

- **Object Composites.** This option allows users to upload an asset, such as a product photo, and then use a text prompt to generate a harmonized and composited scene with the uploaded object. This capability is exposed via the Text to Image module on the Firefly website as well.
- **Text Effects.** With this option, users can apply styles or textures to words and phrases to make them look cool via text prompts instead of painstaking labor.
- **Generative Recolor.** This option allows users to generate color variations of the vector artwork with ease via text prompting.
- **Text to Vector Graphic.** This option allows users to generate customizable, scalable vector graphics via simple text prompts, enabling them to create icons, patterns, social graphics, and other brand assets with ease instead of spending hours creating them in a Scalable Vector Graphics (SVG) format.

ADOBE DOES RESPONSIBLE INNOVATION IN THE AGE OF GENAI

Before its commercial release, Adobe Firefly was reviewed by Adobe's Ethical Innovation, Trust & Safety and Product Equity teams so the company could work to address any concerns regarding harm and bias. Adobe reviews the models on an ongoing basis and when new data is used to train the model.

The Firefly generative models are built on AI Ethics principles of accountability, responsibility, and transparency.³ Adobe states its core principle for AI Ethics as follows: "At Adobe, our purpose is to serve the creator and respect the consumer, and our heritage is built on providing trustworthy and innovative solutions to our customers."

To that end, Adobe is constantly working to eliminate unwanted and undesirable results by building AI responsibly and basing its AI on these core foundational principles.

On the responsibility front, Adobe is designing and maintaining its AI technology with thoughtful evaluation and careful consideration of the impact and consequences of its deployment. This includes, but is not limited to, design for inclusiveness and assessment of the impact of potentially unfair, discriminatory, or inaccurate results that might perpetuate harmful biases and stereotypes. This includes actions such as anticipating potential harms, taking preemptive steps to mitigate such harms, measuring and documenting system performance throughout the technology lifecycle, and establishing systems to monitor and respond to unanticipated harmful outcomes. Adobe also works toward mitigating harmful content, such as removing violent content and so forth.

In addition to bias mitigation and fairness in models, Adobe goes a step beyond owning the responsibility of digital content created via its digital media tools. Adobe is committed to contributing to provenance tools and solutions that address the authenticity of digital content produced using Adobe tools and to contributing to solutions that address the issue of manipulated

media and forged/pirated content. One such initiative, for example, is Adobe's work with the Content Authenticity Initiative.⁴

On that note, Adobe is also very clear that the following is out of the scope of its current initiatives:

- **Ensuring egalitarian fairness.** The value of AI is in its ability to differentiate, rather than providing identical experiences to all users. As long as the AI is free of harmful bias, enabling customization and personalization of tools, recommendations, and technologies is valuable, important, and welcomed by users.
- **Certifying that inbound training data is unbiased.** All training data encodes bias in one direction or another, because humans create the data point, and humans have an inherent bias. Adobe does not believe it is reasonable or useful to commit to a zero-bias starting point; rather, it is more important and effective to commit to having an outcome that mitigates harmful bias.

Adobe specifically makes the following assertions on the accountability front:⁵

- Establishing governance processes to evaluate and track the performance of AI algorithms, data, and designs, including labeling datasets and models for any identified bias, to ensure remediation can occur at the product design stage
- Requiring an AI impact assessment (as part of Adobe's services development process) to ensure an AI ethics review happens before deployment of new AI technologies
- Maintaining an AI Ethics Advisory Board to oversee the promulgation of AI development requirements and be a place where any AI ethics concerns can be heard
- Maintaining processes to ensure remediation of any negative AI impacts that are discovered after deployment
- Educating engineers and product managers via training courses on AI ethics issues

On the transparency front, Adobe aims to be open about and explain the use of AI to its customers, so they have a clear understanding of Adobe AI systems and their applications. Adobe also aims to be transparent about the use of AI in its tools and processes, so its customers can understand how Adobe uses AI, the value AI-assisted tools bring to them, and what controls and preferences they have available when they engage with and utilize Adobe's AI-enhanced tools and services.

There is no doubt that AI will enhance human creativity, drive value, and personalize digital experiences in the digital ecosystem. However, many GenAI providers are rushing their solutions to the market without due consideration to ethical principles, which can do more harm than good to their users and to society at large.

Transparency is the reasonable public disclosure, in clear and simple language, of how a company responsibly develops and deploys AI within its tools. Adobe values its trusted relationship with its customers and feels that transparency is integral to that relationship. This includes sharing information on how or whether Adobe collects and uses customer assets and usage data to improve its products and services.

Adobe specifically lists the following best practices on the transparency front:

- Disclosure of data collection practices, including revealing when and if an individual's data will be collected for AI training and what controls a user will have over the collection
- Providing prior notice if human review of customer data for AI training will be implemented
- Describing how datasets are used in building AI models
- Defining how AI tools are tested for issues related to harmful bias and how those issues are resolved
- General disclosure of how data and AI are used in Adobe's tools and services
- External and internal feedback mechanisms to report concerns with Adobe's AI practices

AS A FOUNDING MEMBER, ADOBE FORMS THE CONTENT AUTHENTICITY INITIATIVE AND COALITION FOR CONTENT PROVENANCE AND AUTHENTICITY

Digital content is often at risk of being miscredited, plagiarized, stolen, or not credited at all when shared online, resulting in lost opportunities for creators to get proper credit or revenue. But more importantly, digital content can spread disinformation that can influence decisions—whether during election time or to influence buyers to make decisions based on inaccurate content. To avoid this situation, Adobe co-founded, and leads, the Content Authenticity Initiative (CAI), which was formed in November 2019 with the goal of developing a digital content attribution industry standard and tools that can increase trust and transparency in online content, as well as combating misinformation and disinformation.

This initiative is focused on promoting widespread adoption of Content Credentials—tamper-evident metadata that can be attached to digital content, based on the Coalition for Content Provenance and Authenticity (C2PA) standard. Content Credentials are akin to a “nutrition label” that creators can attach to their digital content and use to track content provenance. The CAI has more than 3,000 members today and growing—made up of media and tech companies, nongovernmental organizations (NGOs), academics, and others—and produces open source code for implementing Content Credentials. In addition to the founding members, Associated Press, Canon, Getty Images, Nikon, The New York Times, NVIDIA, Qualcomm, Reuters, Shutterstock, The Wall Street Journal, The Washington Post, TikTok, and many others have joined the CAI. This can enable publishers, creators, and consumers to trace the origin and evolution of a piece of media, including images, videos, audio, and documents. Because this is an open standard that can be adopted by any online platform, it can be used as a baseline for trust across the internet even by noncoalition members.

Content Credentials can help build credibility and trust by providing important information about how digital content was created, modified, and published, and even who created the content. Content

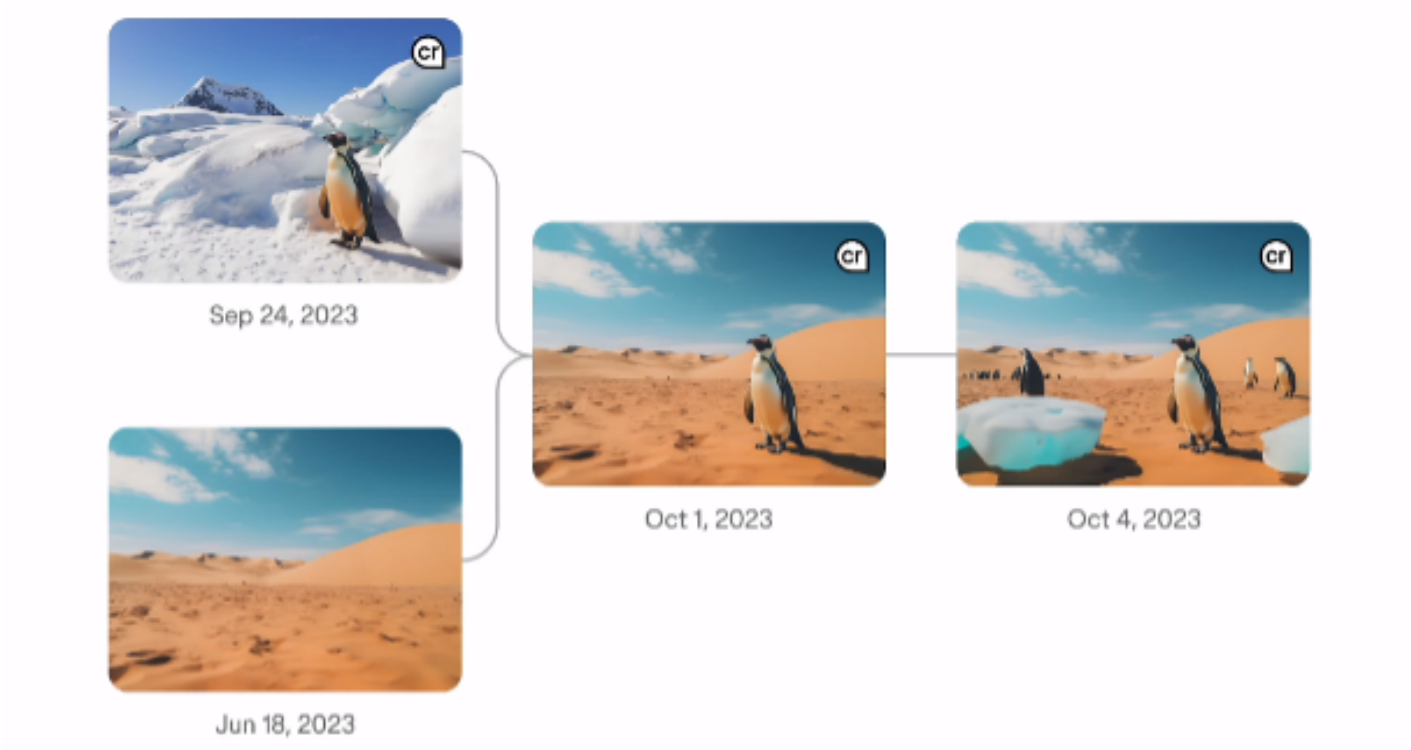
Credentials can also be used by creators to get recognition for their work online, by attaching information such as their name or social media accounts directly to the content they create. To engage with Content Credentials data, viewers can click on the “Cr” pin attached to a piece of content or use the Content Credentials Verify website to see the chain of Content Credentials history tied to an asset. These tools let users validate the relevant information, including the creator and modifier, if available, and a detailed history of changes over time.

In February 2021, Adobe, Arm, BBC, Intel, Microsoft, and Truepic cofounded the C2PA, the formal standards development organization dedicated exclusively to drafting technical standards and specifications as a foundation for universal content provenance. As of September 2024, C2PA steering committee members include Adobe, Amazon, BBC, Google, Intel, Meta, Microsoft, OpenAI, Publicis Groupe, Sony, and Truepic.

Content Credentials are defined by the C2PA open technical standard.⁶ What makes Content Credentials more compelling than other types of metadata is that they are more resilient and durable, as well as tamper-evident.

For example, in the pictures below, the origination of the contents (coming from two different images) as well as the modification information on who did what and when are accurately captured (see Figure 2).

Figure 2. A Composite Picture Produced by Adobe Firefly Showing the Content Authenticity and Provenance



Source: Adobe

THE BOTTOM LINE

Responsible Innovation Should Be the Future of GenAI Done Right

The proliferation of deepfakes and the ability of AI to produce strikingly realistic images have taken the world by storm. Although using GenAI for fun and educational purposes was the initial use case, bad actors are also using AI for malicious purposes. Media content provenance is one of the hardest things to prove in digital media today. By establishing the C2PA standard, an Adobe-developed invisible watermark, and digital fingerprint, Adobe is solving this difficult problem with ease.

Although many companies, both small and large, have rushed to provide tools, models, software, infrastructure, storage, memory, and a whole host of other services to cash in on this current wave of “AI gold rush,” very few have taken steps to promote AI solutions that are designed to be ethical, responsible, and safe for commercial use. Adobe has done not only this but also has joined other big names to make it a priority.

While most of the current text-to-image tools can do portions of what Adobe Firefly can do, there are a few things that are different in the way Adobe has done it. First, Adobe has embedded the GenAI models into the tools creators already use. Second, Adobe provides tools for the creators to produce artifacts while keeping brand consistency. Third, Adobe compensates creators via stock loyalty and bonus programs. Finally, content credentialing is automatic.

Most AI teams spend an inordinate amount of time preparing data formats and quality, but little time on eliminating biased data. Only very few companies, such as Adobe, are even trying to address this issue. To that end, Adobe has designed its GenAI tools and Firefly models to be safe for commercial use. In addition to its AI ethics board being actively engaged, Adobe also provides indemnification to the enterprises that use those tools and models.

In a recent Harvard Business Review paper,⁷ this author and independent researcher and writer Joe McKendrick discussed how many executives from large corporations still don't trust AI to make or

even assist in decision-making. For them to trust AI, morally and ethically, AI-assisted decisions need to reflect ethical and moral values as well—partly because they want to appear as a company with ethical and moral values and operate with integrity, and partly because of the legal liabilities that may arise from making wrong decisions that can be challenged in courts—especially given that, if the decision were either AI-made or AI-assisted, it will go through an extra layer of scrutiny.

Executives could be wary about absorbing AI decisions if there is a lack of transparency. Most AI decisions don't have explainability built into them, which makes them harder to trust. When a decision is made and an action is taken that risks millions of dollars for an enterprise, or it involves people's lives/jobs, saying AI made this decision so we are acting on it is not good enough.

It is about time that we start building ethical, responsible, and explainable AI. Adobe's initiatives are a good step in the right direction.

ENDNOTES

¹ “Authentic Storytelling Through Digital Content Provenance,” Content Authenticity Initiative. <https://contentauthenticity.org/>

² “Inspect Content to Dig Deeper,” Content Credentials. <https://contentcredentials.org/verify>

³ “Adobe’s Commitment to AI Ethics,” Adobe. <https://www.adobe.com/ai/overview/ethics.html>

⁴ “Authentic Storytelling Through Digital Content Provenance,” Content Authenticity Initiative. <https://contentauthenticity.org/>

⁵ “Adobe’s Commitment to AI Ethics,” Adobe. <https://blog.adobe.com/en/publish/2022/03/24/putting-principles-into-practice-adobes-approach-to-ai-ethics>

⁶ “C2PA Overview,” Coalition for Content Provenance and Authenticity. <https://c2pa.org/>

⁷ Andy Thurai and Joe McKendrick, “Overcoming the C-Suite’s Mistrust of AI,” Harvard Business Review, March 23, 2022. <https://hbr.org/2022/03/overcoming-the-c-suites-distrust-of-ai>

ANALYST BIO

Andy Thurai

Vice President and Principal Analyst

Andy Thurai is a vice president and principal analyst at Constellation Research. He is focused on the challenges facing IT teams in the “digital-native” era, including AI, ML, AIOps, edge, and cloud. He specializes in how to operationalize emerging technologies and in how to productionalize them. He advises CxO-level IT executives on how to digitize businesses.

Thurai’s coverage areas include AIOps, CloudOps, ITOps, observability, MLOps, ModelOps, DataOps, AI, ML, AI ethics, and AI bias.

Thurai is an accomplished IT executive, strategist, adviser, enterprise architect, and evangelist with more than 25 years of experience in executive, technical, and architectural leadership positions at companies such as IBM, Intel, BMC, Nortel, and Oracle. Thurai has written more than 100 articles on emerging technology topics for Forbes, The New Stack, AI World, VentureBeat, DevOps.com, GigaOm, and Wired.

His strength is selling technology to the CxO audience with a value proposition rather than the usual technology sales pitch. Find more details and samples of Thurai’s work on his website, thefieldcto.com.

 [@andythurai](https://twitter.com/andythurai)  constellationr.com/users/andy-thurai  linkedin.com/in/andythurai/

ABOUT CONSTELLATION RESEARCH

Constellation Research is an award-winning, Silicon Valley—based research and advisory firm that helps organizations navigate the challenges of digital disruption through business model transformation and the judicious application of disruptive technologies. Unlike the legacy analyst firms, Constellation Research is disrupting how research is accessed, what topics are covered, and how clients can partner with a research firm to achieve success. Over 350 clients have joined from an ecosystem of buyers, partners, solution providers, C-suite, boards of directors, and vendor clients. Our mission is to identify, validate, and share insights with our clients.

Organizational Highlights

- Named Institute of Industry Analyst Relations (IIAR) New Analyst Firm of the Year in 2011 and #1 Independent Analyst Firm for 2014 and 2015.
- Experienced research team with an average of 25 years of practitioner, management, and industry experience.
- Organizers of the Constellation Connected Enterprise—an innovation summit and best practices knowledge-sharing retreat for business leaders.
- Founders of Constellation Executive Network, a membership organization for digital leaders seeking to learn from market leaders and fast followers.



www.ConstellationR.com



[@ConstellationR](https://twitter.com/ConstellationR)



info@ConstellationR.com



sales@ConstellationR.com

Unauthorized reproduction or distribution in whole or in part in any form, including photocopying, faxing, image scanning, emailing, digitization, or making available for electronic downloading is prohibited without written permission from Constellation Research Inc. Prior to photocopying, scanning, and digitizing items for internal or personal use, please contact Constellation Research Inc. All trade names, trademarks, or registered trademarks are trade names, trademarks, or registered trademarks of their respective owners.

Information contained in this publication has been compiled from sources believed to be reliable, but the accuracy of this information is not guaranteed. Constellation Research Inc. disclaims all warranties and conditions with regard to the content, express or implied, including warranties of merchantability and fitness for a particular purpose, nor assumes any legal liability for the accuracy, completeness, or usefulness of any information contained herein. Any reference to a commercial product, process, or service does not imply or constitute an endorsement of the same by Constellation Research Inc.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold or distributed with the understanding that Constellation Research Inc. is not engaged in rendering legal, accounting, or other professional services. If legal advice or other expert assistance is required, the services of a competent professional person should be sought. Constellation Research Inc. assumes no liability for how this information is used or applied nor makes any express warranties on outcomes. (Modified from the Declaration of Principles jointly adopted by the American Bar Association and a committee of publishers and associations.)

Your trust is important to us, and as such, we believe in being open and transparent about our financial relationships. With our clients' permission, we publish their names on our website.

San Francisco Bay Area | Boston | Colorado Springs | Denver | Ft. Lauderdale | New York Metro
Northern Virginia | Portland | Pune | San Diego | Sydney | Washington, D.C.