

Putting AI to Work

6

Generating Images and Visuals

Learning Objectives

- Compose effective prompts to generate images that include purposeful, legible text using AI tools
- Analyze and experiment with AI-generated visuals to interpret style, composition, and aesthetic choices
- Apply AI tools to modify, enhance, or add elements to existing images, including the use of recoloring or editing tools
- Observe and articulate detailed descriptions of visual content using AI-supported or manual tools
- Convert content across media formats, such as transforming text into visuals or diagrams, by using generative AI
- Design visual representations like charts, diagrams, or layouts to communicate abstract or complex information clearly

Module 6.1: Generating Images Containing Text

- AI-generated images are practical tools for marketing, branding, and professional applications.
- Image generation should be purpose driven (for example, to match visuals to brand, convey tone, or meet functional requirements).
 - Adobe Firefly excels at commercial marketing and brand-safe editable content.
 - DALL·E is good for concept art and photorealism but can struggle with fine details.
 - Midjourney specializes in artistic styles and rich textures but has a difficult user interface (UI).
 - Looka is specifically built for the creation of logo and branding designs with usable text.
- A major challenge is that AI generators have difficulty placing readable, accurate text in images.
- Most tools are trained on images with text, not text as a design element.
- There's no single best tool: Good results depend on trying multiple tools and refining prompts.

Module 6.1: Ethics in Action

- AI images with embedded text raise questions around authenticity and misinformation.
- Text in images can easily be misused or taken out of context.
- It's important to ensure clarity, avoid deceptive practices, and consider visual accessibility.
- Transparency about the use of AI-generated content builds trust with audiences.

Module 6.1: Techie Dive

- Most AI image generators struggle with readable text because they're not trained for typography.
- AI tools don't treat letters as discrete language units in image generation.
- DALL·E and Midjourney often render text as distorted characters.
- Some platforms use post-processing layers, but the results remain inconsistent.

Module 6.1: Business Lens

- Clear text in marketing content and branded visuals is often non-negotiable.
- There may be a need to supplement AI outputs with traditional design tools like Canva or Photoshop.
- Choosing the right tool and knowing its limitations saves time and maintains brand integrity.
- It's important to establish review workflows to ensure text accuracy and design consistency at scale.

Module 6.2: Exploring Artistic Ideas

- AI can serve as an artistic collaborator for brainstorming visual styles and moods.
- Applications:
 - Illustration
 - Concept art
 - Fashion
 - Game design
 - Décor
- Image-generation tools create entirely new images rather than pulling from existing sources.
 - Midjourney is known for cinematic, surreal, atmospheric visuals with intricate textures.
 - Adobe Firefly is good at interpreting mood, lighting, and visual atmosphere.
 - DALL·E offers balanced performance across realism, illustration, and creativity.
- Experimenting with abstract prompts can lead to unexpected visual outcomes.
- Different tools interpret the same prompt differently, so try multiple tools.

Module 6.2: Ethics in Action

- Emulating famous art styles or mimicking living artists raises originality and consent questions.
- Some tools block artist names in prompts, while others don't.
- Responsible users avoid copying a living artist's signature style without permission.
- Consider the impact on professional artists when AI replicates their distinctive work.

Module 6.2: Techie Dive

- Most AI image generators use diffusion models trained on large labeled image datasets.
- These models turn random noise into a coherent image by gradually refining based on prompt.
- Artistic styles are inferred from data patterns, not stored explicitly.
- Some tools allow prompt weighting or image input for more control.

Module 6.2: Business Lens

- Designers can rapidly prototype visual ideas without hiring illustrators for every iteration.
- Speed and cost effectiveness accelerate creative workflows significantly.
- AI art may not be commercially usable due to unclear copyright or licensing status.
- Check whether images are safe for commercial use: Look for content credentials and license indicators.

Module 6.3: Transforming Images

- Image transformation modifies or remixes existing images rather than creating them from scratch.
- Types:
 - Inpainting modifies a specific part of image and allows object swapping.
 - Outpainting extends an image beyond its original borders.
 - Style transfer applies an artistic filter to change the visual style.
 - Background replacement changes the context while keeping the subject.
 - Colors and textures can be changed.
 - Image-to-image remixing generates a new image inspired by an existing one.
- These are used in design, advertising, education, marketing, and entertainment.

Module 6.3: Transforming Images (cont.)

- Common pitfalls:
 - Overediting to the point of distortion misrepresents products and undermines trust.
 - Failing to credit source images violates copyright even when the images have been altered.
 - Ignoring licensing restrictions can lead to legal issues.
 - AI doesn't know intended use without clear prompts and may lose context.

Module 6.3: Ethics in Action

- Transformation tools can dramatically alter reality with ethical and legal implications.
- Misleading edits like removing people or changing expressions require transparency.
- Context is key when presenting transformed images in journalism, politics, and education.
- Always consider whether transformations could deceive or harm viewers.

Module 6.3: Techie Dive

- Image transformation relies on models trained on visual-text datasets.
- Inpainting uses masked image regions and context clues.
- Style transfer involves neural networks mapping style onto content.
- Image transformation requires substantial GPU power and pixel-level prediction.

Module 6.3: Business Lens

- Brands use transformation to localize campaigns, repurpose content, A/B test visuals.
- The speed and cost effectiveness of AI allows businesses to produce more assets at scale.
- It's important to ensure brand integrity and consistency across transformed assets.
- Establish quality control processes to maintain professional standards.

Module 6.4: Describing Image Content

- AI tools can describe what's in an image (image captioning).
- Captioning supports accessibility for users with visual impairments and enables automatic tagging for organization and search features.
- There are applications in autonomous vehicles, smart surveillance, and e-commerce.
- Different prompting approaches:
 - Simple summary
 - Detailed breakdown
 - Alt text
 - Tagging
 - Storytelling
- AI can identify people, activities, objects, brands, settings, and emotions.
- The results aren't always accurate, especially with abstract or cluttered images.

Module 6.4: Describing Image Content (cont.)

Common pitfalls:

- Trusting AI blindly: Always verify the accuracy of descriptions.
- Over-reliance on generic outputs: It's good to customize for specific needs.
- Forgetting context: It's important to clarify marketing, legal, or accessibility requirements in prompts.

Module 6.4: Ethics in Action

- Describing images raises ethical concerns when people are involved, as the descriptions may infer gender, emotion, or identity inaccurately or offensively.
- Avoid harmful assumptions: AI should support, not replace, human interpretation.
- This is critical in accessibility and legal contexts where accuracy matters.

Module 6.4: Techie Dive

- Tools like CLIP and BLIP power visual understanding.
- Systems use paired text-image datasets to associate visual features with language.
- Multimodal models like GPT-4o analyze images alongside text prompts.
- This enables rich interactions (for example, ingredient identification and chart analysis).

Module 6.4: Business Lens

- Descriptive AI streamlines image tagging, improves SEO, and automates compliance.
- It helps organize visual databases efficiently.
- The fields of real estate, fashion, and publishing benefit from automated visual summarization.
- Human review is essential to ensure brand alignment and accuracy.

Module 6.5: Transforming Media Types

- AI can be used to transform content from one format to another (for example, from text to image, from image to text, and from text to audio).
- Types of transformations:
 - Text-to-image: Create visuals for marketing, pitch decks, storyboarding.
 - Image-to-text: Automate product descriptions, which saves time and improves consistency.
 - Text-to-audio: Create voiceovers for videos, explainers, and accessibility tools.
 - Audio-to-text: Generate transcripts for meetings, podcasts, and videos.
 - Text-to-video: Create explainer videos, animations, and video ads.
 - Image-to-video: Animate still images for social media or storytelling.
- The goal is to repurpose content for different audiences and platforms without starting over.

Module 6.5: Transforming Media Types (cont.)

Common pitfalls:

- Ignoring the tone and context: Transformed content may not match the original intent.
- Mismatched style: Ensure the output aligns with brand and audience expectations.
- Accessibility oversights: Include captions, transcripts, and alt text for inclusive design.

Module 6.5: Ethics in Action

- AI can make media more accessible and inclusive through alt text and captions.
- Always consider how people with different needs will experience your output.
- Accessibility isn't optional: It's fundamental responsibility in content creation.
- Ensure transformed content maintains accuracy and doesn't contain misleading elements.

Module 6.5: Techie Dive

- Media transformation relies on multimodal models trained across formats.
- Models convert audio, images, and text into shared data representations, which allow the system to process and regenerate content in new forms.
- Key tools:
 - Runway
 - ElevenLabs
 - Descript
 - OpenAI's Whisper
 - DALL·E

Module 6.5: Business Lens

- Repurposing content boosts marketing ROI (return on investment) and streamlines production.
- A single podcast can be used to generate a transcript, blog post, social post, and YouTube video.
- Learning to prompt for transformation provides competitive advantage.
- Efficiency gains allow teams to maximize existing content investments.

Module 6.6: Visualizing Content

- Convert complex information into visual formats for better understanding.
- Types:
 - Charts/graphs show proportions, trends, and comparisons in reports and presentations.
 - Infographics combine images, text, and icons for marketing and education.
 - Diagrams and flowcharts map workflows, processes, and relationships.
 - Mind maps organize ideas and brainstorm connections.
 - Timelines display chronological events or project milestones.
- Careful prompting is essential to avoid generic or misrepresentative visuals.

Module 6.6: Visualizing Content (cont.)

Common pitfalls:

- Vague prompting: Be specific about data, chart type, labels, audience.
- Wrong chart type: Match the visualization to the data story you want to tell.
- Missing labels: Ensure all elements are clearly identified.
- No context given: Make sure to specify the purpose, audience, and intended use.

Module 6.6: Ethics in Action

- Be careful when visualizing data: Misleading graphs spread misinformation.
- Always verify the source and accuracy of data you visualize.
- Check whether generated images contain biased or stereotyped illustrations.
- Responsible visualization requires honesty in how data is presented.

Module 6.6: Techie Dive

- ChatGPT with Code Interpreter converts prompts to Python code using matplotlib or seaborn.
- Canva and Adobe Express allow semiautomated creation using templates.
- Gemini can combine visuals with AI-generated summaries for hybrid outputs.
- Understanding technical capabilities helps a user choose appropriate tools.

Module 6.6: Business Lens

- Data visualization boosts understanding, speeds decisions, and impresses clients.
- Automated tools can create visuals that look right but convey the wrong messages.
- Always match visuals to the audience (for example, internal reports need detail, whereas client slides need clarity).
- Professional-quality visualizations enhance credibility and communication.

Key Takeaways

- AI-generated images are practical tools across industries for marketing, education, branding, and creative projects.
- Different AI tools have specific strengths: Firefly is best for commercial applications, Midjourney is best for artistic applications, and DALL·E is versatile.
- Text rendering in AI images remains challenging, and most tools struggle with the creation of legible, accurate text.
- Image transformation enables the modifying of visuals through inpainting, outpainting, and style transfer.
- AI can describe image content for accessibility and tagging but requires human verification.
- Media transformation allows for the repurposing of content across formats to maximize reach and efficiency.
- Data visualization with AI requires careful prompting to avoid misleading outputs.
- There are ethical considerations related to authenticity, copyright, accessibility, consent, and the avoiding of misinformation.
- Businesses benefit from the use of AI visual tools but must maintain quality control and legal compliance.
- Human oversight is essential: AI assists but doesn't replace thoughtful judgment in visual content creation.