



Input

What data should and can generative AI use as input?
Different AI methods & tools expect different types & formats of data - also referred to as "data modality".
Multimodal AI systems accept different types of input.

Text, Code & Markup

Examples: user input ("prompt"), articles, Python code or XML file.

Images, Video & Audio

Examples: images, interview-videos or podcast recordings.

Models, Graphs & Diagrams

Examples: technical drawings, sales statistics or flowcharts.

Other

Examples: sensor data, time series or presentations.

Function

What function(s) or operation(s) should the AI system perform on the input data to produce the output?
By combining different specialized AI models, even complex sequences of functions can be implemented or generate different outputs.

Extraction & Filtration

Examples: extract all people from photos or filter out photos with people.

Anonymization & Masking

Examples: replace names in text with random strings or pixelate people's faces in photos.

Augmentation & Extension

Examples: increase the resolution of photos or add previously removed areas to photos.

Detection & Prediction

Examples: recognize all people in photos or predict what will happen next in a photographed scene.

Summarization & Explanation

Examples: summarize an article in a few sentences or explain why a photo is funny or offensive.

Transformation & Translation

Examples: convert JavaScript to Python code or translate from German to English.

Output

What output should the AI system generate?
Specialized AI methods generate specific output types & formats.
So-called multimodal models can generate different types of output.

Text, Code & Markup

Examples: documentation, JavaScript code or HTML files.

Image, Video & Audio

Examples: photorealistic 3D visualizations, animations or synthesized speech.

Models, Graphs & Diagrams

Examples: 3D models, data visualizations or organizational charts.

Other

Examples: computer games, smartphone apps or synthetic data.

