Apple Intelligence and machine learning

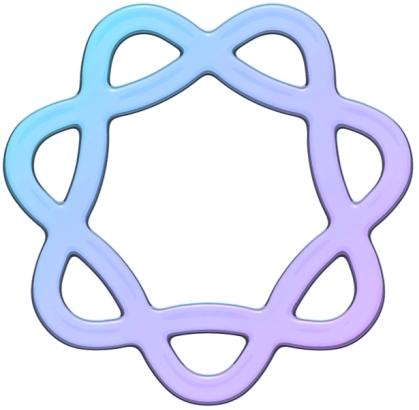
Add intelligent features with Apple Intelligence, machine learning, and related technologies.

Intelligent features make people's experiences on a device more personal. Apple incorporates intelligent features into many of its frameworks and technologies, making it easy to adopt those features quickly in your own apps. When a framework doesn't do what you need it to, you can customize its behavior by training a model with your own data — like training a new sound classification model to detect your own sound. Of course, if you have your own machine learning models, you can also incorporate them into your app in a way that works well with Apple hardware and the system's other intelligent features.

Apple Intelligence

Apple Intelligence is the personal intelligence system that helps you communicate, work, and express yourself. It combines generative models with your personal context to deliver intelligence that is most useful and relevant to people, while designed with privacy in mind. Now you have direct access to the on-device model at the core of Apple Intelligence, so you can build experiences that are smart, private, and work even without internet connectivity.

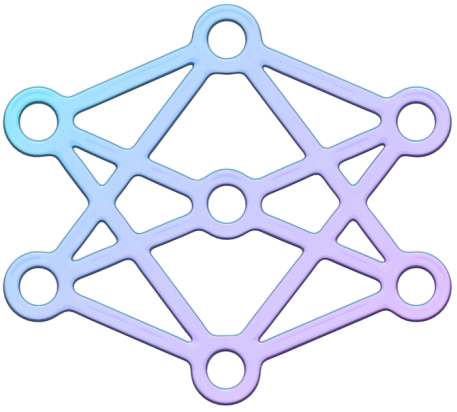
- Add intelligent tools to compose, rewrite, or proofread text.
- Provide image generation capabilities that allow people to personalize the images they want to use in your app.
- Support the addition of Genmoji in text-based content.
- Integrate your app's content and features to the system.



To learn more, read Apple Intelligence.

Foundation Models framework

Run your own prompts on the on-device model at the core of Apple Intelligence to enhance your existing features with model capabilities like text extraction and summarization. When you prompt the model, provide your Swift data type that it uses to avoid making structural mistakes — allowing you to focus on your prompts and not parsing JSON schemas. The models can also call your own code to help perform additional tasks that might need real-time information or content that's in a local database.

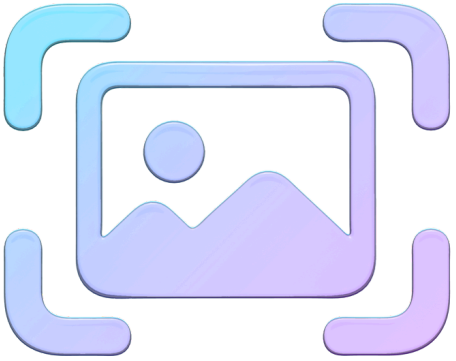


- Access a highly optimized on-device language model that's specialized for text-based prompting.
- Describe the output you want as a new Swift data type and get results using the custom data type you define.
- Extend the functionality of the on-device models by creating custom tools that help with app-specific tasks.

To learn more, read Foundation Models.

Intelligent frameworks

Many system frameworks use machine learning models to make difficult tasks easier, and make seemingly impossible tasks possible. You can use system frameworks that help with performing analysis of images, sound, and language.

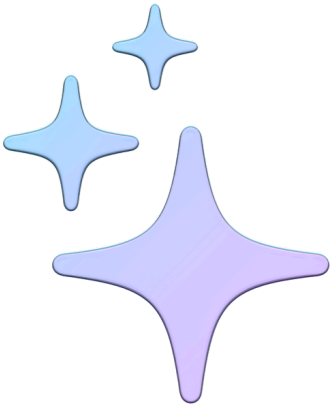


- Add image processing tasks like document recognition, object identification, and face detection.
- Perform speech recognition to convert recorded or live audio of speech into text.
- Match audio against the Shazam music library or a custom catalog that contains your own audio.
- Classify sounds in real time to identify environmental sounds like laughter, sirens, or dog barks.
- Analyze natural language to identify parts of speech or translate text into the language a person prefer.

To learn more, read Intelligent frameworks.

Custom machine learning models

Create your own machine learning models when you need more customization than what the system technologies provide. Use the Create ML app when you want a fast, user-friendly way to create and train models with your own data. If you already have a model you've built, you can convert it to the Core ML format with Core ML tools and deploy it into your app.



- Build and train a model using your own data with Create ML.
- Convert and optimize models you create with training libraries by using Core ML Tools and integrate them with Core ML.
- Download and explore models that are already in the Core ML format and ready for integration into your app.
- Train and fine-tune neural networks with Apple's MLX Python training library.
- Work with neural network layers for fine-grained control over task execution on the CPU and GPU.

To learn more, read Machine learning models.