

Lesson 02 — Types of AI

Four lenses to tell every "AI" apart — so headlines stop sounding like magic.

AI for Dummies

Module 1

A spam filter, ChatGPT, a chess engine, Midjourney, and Tesla Autopilot are all called **"AI."**

They work completely differently. This lesson gives you four questions to tell them apart.

Lesson arc: Capability → Output → Modality → Where it runs → Recap

Four lenses — one for each question

1

How capable?

Narrow → General → Super

Where we are today? Narrow only.

2

What does it produce?

Predictive (classifies & forecasts)

Generative (creates new content)

3

What kind of content?

Text · Image · Audio · Video

Multimodal (several at once)

4

Where does it run?

Cloud (company servers)

Local (your own machine)

Any AI tool you encounter — including ones invented next year — can be described with these four questions.

Lens 1: How capable? — Three tiers

Narrow AI (ANI)

EXISTS TODAY — every tool you can name

Good at *one kind of task*. "Narrow" does not mean weak — it means specialised. A chess engine is superhuman at chess. Ask it to write a poem: blank.

General AI (AGI)

HYPOTHETICAL — does not exist yet

As flexible as a human across *any* task. Seriously debated: some researchers say 5 years away, others say decades, others say the concept is ill-defined.

Superintelligence (ASI)

SCIENCE FICTION — makes great movies

Smarter than all humans at everything simultaneously. Nothing to download, nothing to interact with. Exists only in speculation and movie scripts.

Narrow AI is everywhere right now

Every one of these is Narrow AI — each brilliant at its specific job, clueless outside it.

♔ Stockfish / AlphaZero

Chess & Go. Superhuman at the board. Cannot book you a flight.

🔑 Gmail Spam Filter

Classifies billions of emails / day. Terrible at writing a reply.

💬 Claude / ChatGPT

"Broad" language models. Still narrow: they can't drive your car.

🚗 Tesla Autopilot

Reads cameras, radar, maps. Cannot compose a text message.

🎧 Suno / Udio

Composes music from a text prompt. Cannot play chess.

🖼️ Midjourney

Paints images from words. Cannot answer a customer email.

The rule: whenever a headline says "AI can now..." — fill in "narrow AI got better at one specific thing."

The capability timeline — where we actually are



Every AI product you can subscribe to or download today sits in the **green zone**. The orange and red zones are speculation.

Lens 2: What does it produce?

Predictive AI

Answers *"which bucket?"* or *"what number?"*

Input → classification or forecast.

This is the **old, quiet, everywhere** kind. Running silently in every major digital service for 20+ years.

spam/not-spam

fraud detection

weather forecast

churn prediction

medical diagnosis AI

Generative AI

Creates new content that didn't exist before.

Prompt → text / image / audio / video / code.

This is the **new, loud** kind. Exploded publicly in 2022 with DALL-E and ChatGPT.

ChatGPT

Claude

Midjourney

Suno

Sora

Copilot

Important: Both types are still Narrow AI. "Generative" describes the output format, not the capability level.
A music-generating model cannot also write code.

Predictive AI — it's been around for decades

You encounter predictive AI dozens of times a day without noticing it.

Gmail & Outlook spam filter

Classifies each incoming email: spam or not-spam. Processes billions of messages per day. Model answer is a single bit.

Netflix & Spotify recommendations

Predicts probability you'll like each title based on your history and similar users. One giant "will-you-enjoy?" classifier.

Stripe & PayPal fraud detection

Scores each transaction as fraudulent or legitimate in <100ms. Saves billions of dollars annually.

Hospital radiology AI

Flags potential tumours in X-rays or MRIs for a radiologist to review. Output: "possible malignancy at (x,y)" — not a story.

Weather forecasting (ECMWF, GFS)

Predicts temperature, rain probability, wind speed — numbers, not prose. Google's GraphCast beats classical models.

Tesla Autopilot vision

Classifies objects in camera frames: car, pedestrian, lane marking, traffic light. Runs ~45 times per second per camera.

Generative AI — by medium

Text & Code

ChatGPT (OpenAI)

Claude (Anthropic)

Gemini (Google)

GitHub Copilot

Llama (Meta)

Write prose, answer questions, summarise documents, generate code. The most discussed category.

Music

Suno

Udio

Text prompt → full song with vocals & instruments in seconds.
Output is an audio file.

Images

Midjourney

DALL·E 3

Stable Diffusion

Flux

Turn a text description into a photorealistic image in seconds.
Used in advertising, concept art, product mockups.

Video

Sora (OpenAI)

Runway Gen-3

Pika

Text or image → short video clips. Quality improving fast —
previously a VFX studio job now runs in a browser.

Lens 3: What kind of content? — Modality

Modality = the type of data the model works with. It matters when choosing a tool for a job.

Text

LLMs: Claude,
ChatGPT, Gemini,
Llama, Mistral, Qwen

Image

Generators:
Midjourney, DALL·E,
Stable Diffusion, Flux
Recognisers: Google
Lens

Audio

Speech-to-text:
Whisper
TTS: ElevenLabs
Music: Suno, Udio

Video

Generators: Sora,
Runway, Pika
Analysis: Gemini 1.5
Pro

Multimodal

Handles several types
at once. Photo + text
→ answer about the
image. Leading
models in 2025–26.

A text-only LLM cannot "see" a photo, no matter how good it is at writing. Match the tool's modality to your task.

Multimodal in action — photo + question → answer

This is how Claude, GPT-4o, and Gemini handle mixed inputs. The model processes text and image together.



Most leading models in 2025–26 are multimodal by default. A text-only model (like an older Llama variant) simply cannot accept an image input — it's a hard limit, not a setting.

Lens 4: Where does it run?

Same idea, two very different places — with real consequences for privacy and cost.



Cloud (Closed)

Model runs on the company's servers.

You connect over the internet.

Claude

ChatGPT

Gemini



Local (Open)

Model downloaded to your computer.

Nothing leaves your machine.

Llama

Mistral

Gemma

Qwen

The distinction matters for **privacy**. Cloud: your prompts travel to a company server. Local: zero bytes leave your machine — conversations are completely private.

Cloud models — powerful, easy, connected

Example products

Claude (Anthropic)

claude.ai · Opus 4, Sonnet 4.5, Haiku 3.5 · Free tier + Pro/Max paid plans

ChatGPT (OpenAI)

chatgpt.com · GPT-4o, o3 · Free tier + Plus/Pro paid plans

Gemini (Google)

gemini.google.com · Gemini 2.0 Flash, 1.5 Pro · Free + Advanced paid

✓ Pros

- Extremely powerful — runs on thousands of GPUs
- Zero setup — sign up, start typing
- Always up to date, maintained for you
- Multimodal out of the box

⚠ Cons

- Your prompts leave your machine
- Heavy use costs money at scale
- Needs internet connection
- Subject to company data policies

Local models — private, free to run, offline

Popular open models

Llama 3.3 (Meta)

Mistral 7B

Gemma 3 (Google)

Qwen 2.5 (Alibaba)

DeepSeek-R1

How to run them

Ollama (terminal)

```
ollama run llama3.2
```

Downloads and runs in one command. Fast, scriptable.

LM Studio (GUI)

Graphical app, ChatGPT-style interface. No terminal needed.
Mac/Win/Linux.

✓ Pros

- 100% private — nothing leaves your machine
- Free per query after download
- Works fully offline
- You control the model version

🖥️ Hardware reality check

8 GB RAM — small models, slowly

16 GB RAM — comfortable for 7B models

32 GB+ — large models at good speed

GPU / Apple M-chip — makes it much faster

Modest laptop? Stay on cloud chat.

When closed vs open actually matters

☛ Choose Cloud when...

- You want the most capable model available
- You're starting out — zero friction
- You need multimodal (text + image) easily
- Your hardware is modest
- You're not working with sensitive data

🏠 Choose Local when...

- You're pasting confidential contracts or medical notes
- You need it to work offline (travel, lab, no internet)
- You run it thousands of times — per-query cost adds up
- You want to modify or fine-tune the model itself
- You're in a regulated environment (GDPR, HIPAA)

Example: A lawyer pasting a client contract should use local AI. A student asking for essay feedback is fine on cloud.

Quick check — classify this tool

You open **Spotify**. It shows you a "Recommended for you" playlist of songs it thinks you'll enjoy. What kind of AI is behind that feature?

Generative AI — it creates new music for you

Predictive AI — it forecasts which songs you'll like

AGI — it understands music like a human would

Local AI — it runs on your phone hardware

All four lenses — combined mental map

Apply all four to any AI story. Here are three examples:

Tool / story	Capability	Output	Modality	Where
Gmail spam filter	Narrow	Predictive	Text	Cloud
Claude (claude.ai)	Narrow	Generative	Multimodal	Cloud
Midjourney	Narrow	Generative	Image	Cloud
Whisper (local)	Narrow	Predictive	Audio	Either
Tesla Autopilot	Narrow	Predictive	Video	Local
Llama (Ollama)	Narrow	Generative	Text	Local

Notice: **every row has "Narrow" in the capability column.** Because all real AI tools today are narrow AI.
The theory page has an interactive grid to filter all 12 examples.

Recap — what you can do now

The four questions

1. **Capability:** All real AI is Narrow. AGI and ASI are hypothetical.
2. **Output:** Predictive classifies & forecasts. Generative creates content.
3. **Modality:** Text / Image / Audio / Video / Multimodal.
4. **Where:** Cloud (easy, powerful, data leaves) vs Local (private, free, needs hardware).

Try it now

Go to **claude.ai**. Sign up for free. That is: Narrow · Generative · Multimodal · Cloud.

Install **Ollama**, run `ollama run llama3.2`. That is: Narrow · Generative · Text · Local.

Next time you see an "AI does X" headline — ask all four questions. The answer is almost always: *"Narrow AI got better at one specific thing."*

Coming up: Lesson 03 — the four doorways into AI: Chat, API, CLI, and Code / agentic tools.

