GRADIENT INSTITUTE

Using LLMs Responsibly and Effectively

A Workshop for Australian NFPs

Dr. Alberto Chierici Liam Carroll

GRADIENT INSTITUTE

Acknowledgement of country

Part of a broader initiative

GRADIENT INSTITUTE

Responsible AI capability uplift for Australian NFPs and social enterprises

- Responsible AI education and training (introductory and specialised)
- In-person advisory to help NFPs and social enterprises use AI responsibly

Offerings are *free* to qualifying Australian NFPs and social enterprises.

Gradient's work on this is supported by a grant from Google.org, Google's charitable arm.

Supporting resources

GRADIENT INSTITUTE

This is the 4th course in our Google.org-sponsored
***Uplifting Responsible AI for Aus NFPs** webinar series.*

1) **Socially Responsible AI f**or Australian NFPs
2) AI for **Socially Responsible Impact: Use Cases** for Aus NFPs
3) **Open Q&A**
4) Using **LLMs** Responsibly & Effectively (this course)

Access recordings of the first two here: https://www.gradientinstitute.org/resources/

Gradient Institute

GRADIENT INSTITUTE

To bring **independence**, **humanity**, and **rigour** to the centre of how AI is created and used

- Research
- Policy
- Practice

iag · The University of Sydney · CSIRO · Data 61

Your facilitators

Dr Alberto Chierici

Principal AI Specialist

Liam Carroll

Researcher

GRADIENT INSTITUTE

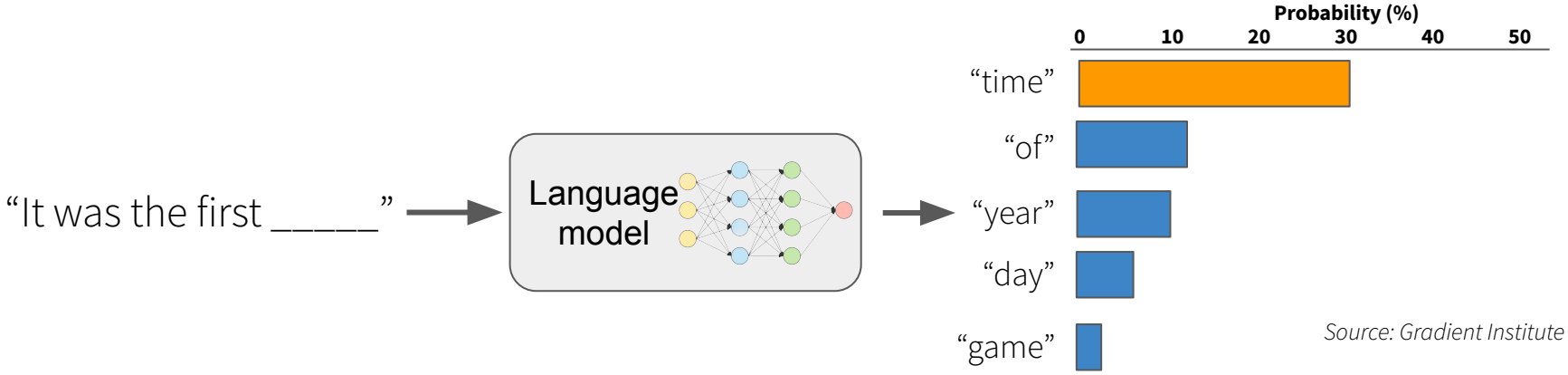
1

What is an LLM?

A brief overview

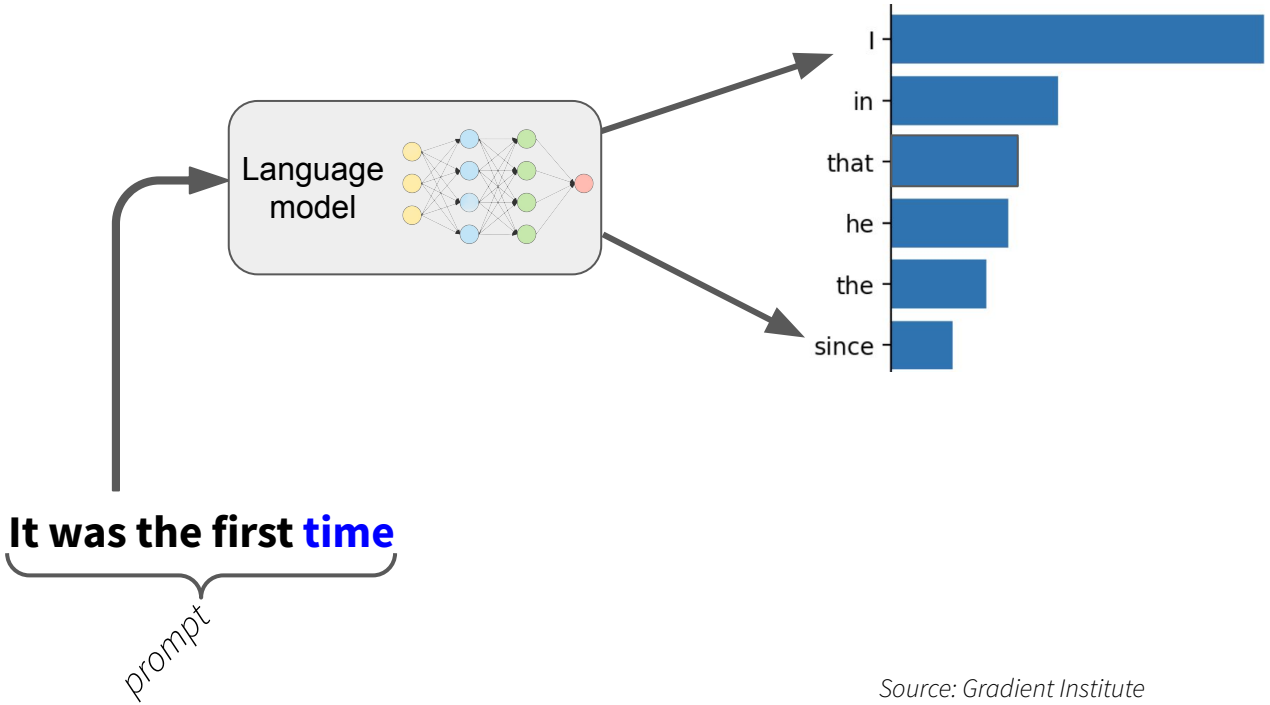
What is an LLM?

Large Language Models (LLMs) are statistical models that predict the probability of the next word in a prompt.



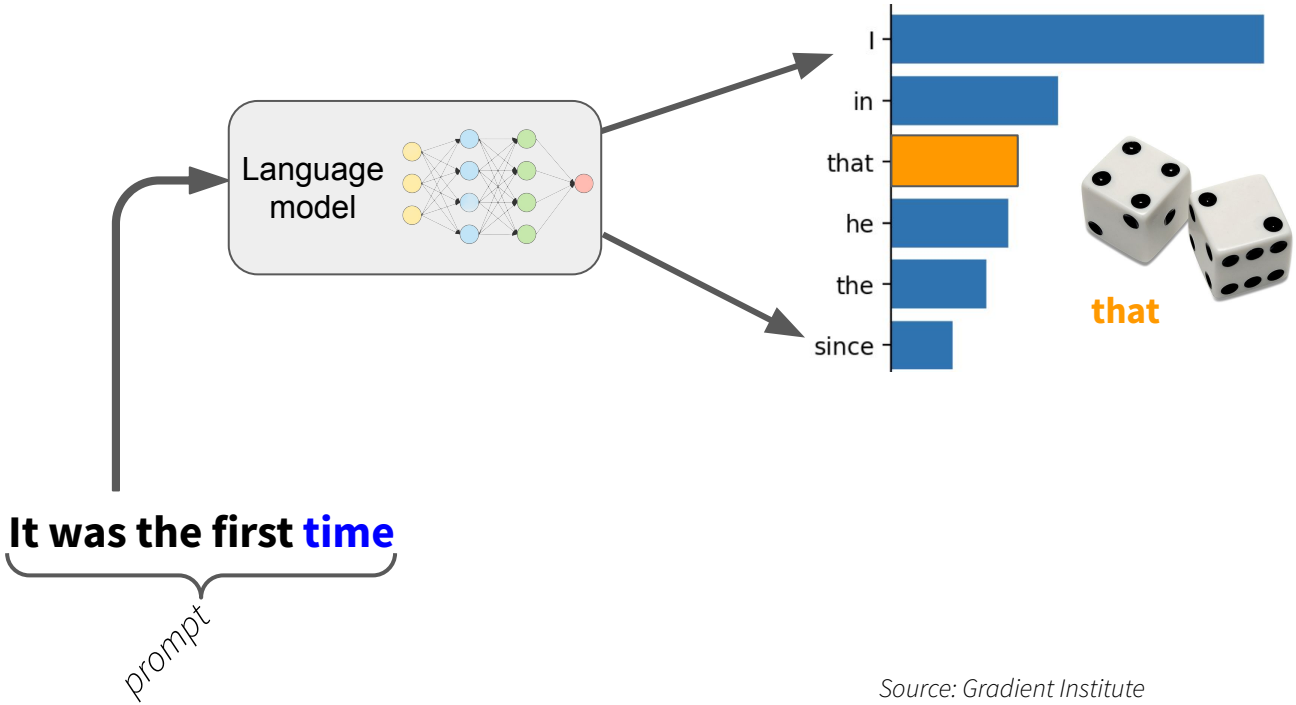
Source: Gradient Institute

Predicting the next word



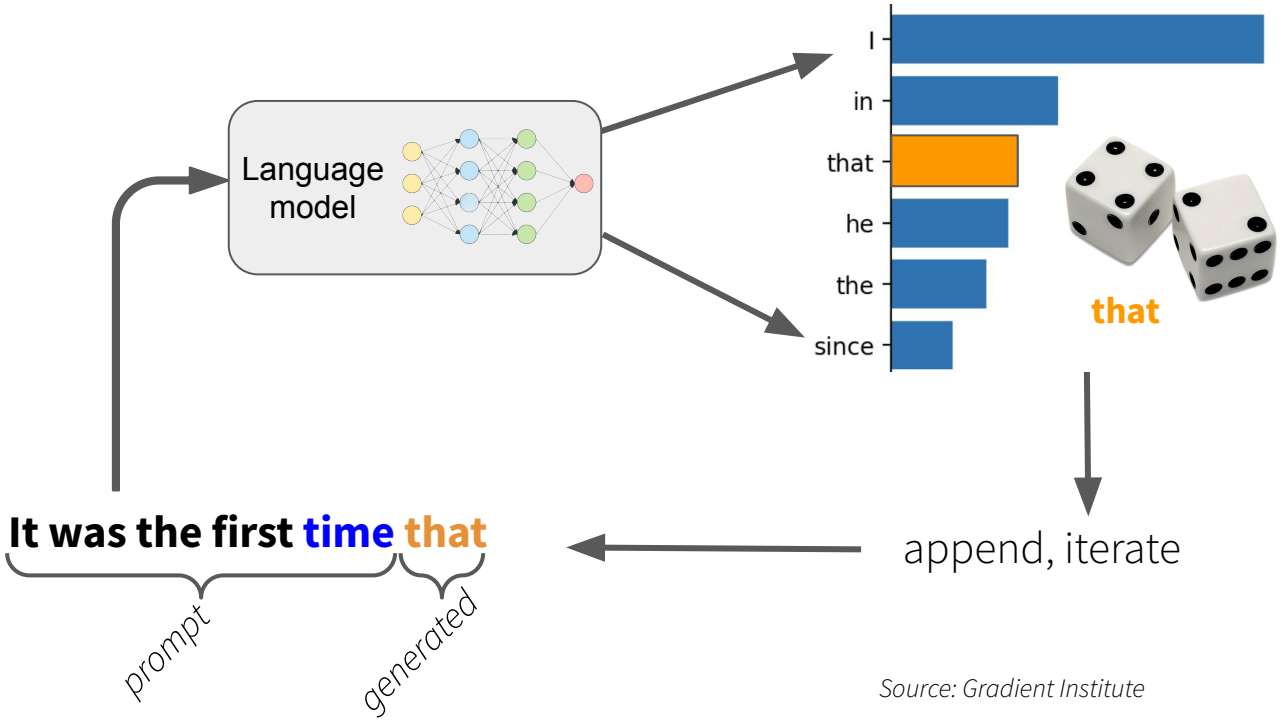
Source: Gradient Institute

Predicting the next word



Source: Gradient Institute

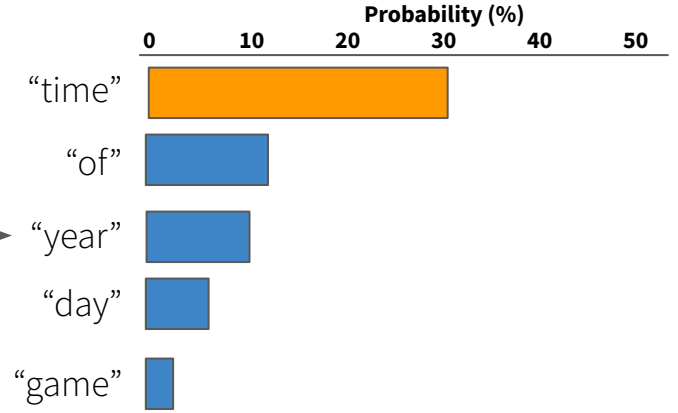
Predicting the next word



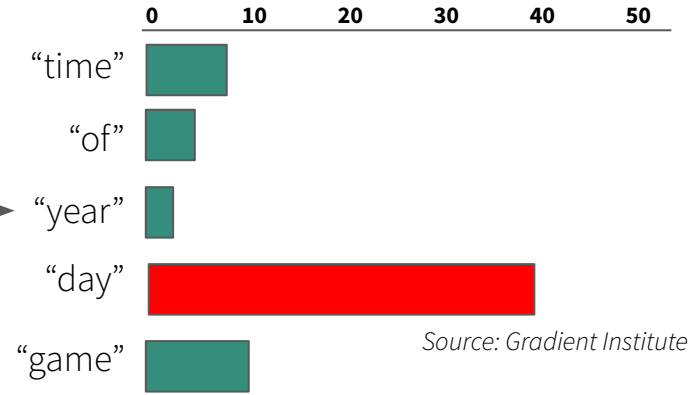
Source: Gradient Institute

Different model, different output

"It was the first _____" → ChatGPT o3-mini →



"It was the first _____" → Claude Sonnet-3.7 →



Source: Gradient Institute

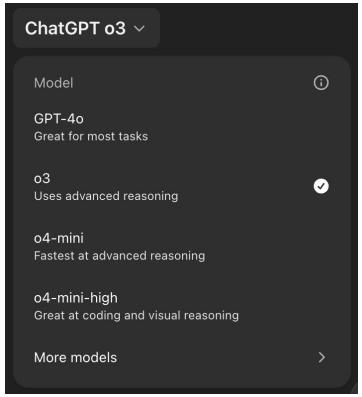
How do I choose a model?

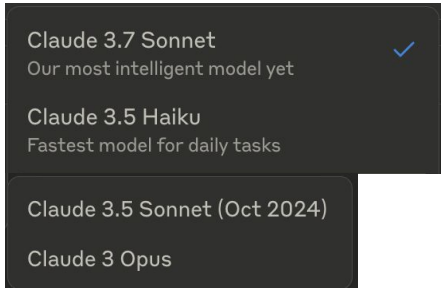
OpenAI

Claude BY ANTHROPIC

Gemini Google

Meta AI





deepseek

Copilot

?!?!?!?!?!

How LLM products differ

Model Capabilities

- **General intelligence** and coherence
- **Domain expertise and specialised skills** e.g. good at code
- **"Reasoning" ability** for hard problems
- **Access to tools** like running code or browsing the web

User Experience

- **Personality**: How does it feel to interact with?
- **Sycophancy**: Is it trying to flatter you?
- **Reliability**: Is it giving trustworthy answers, or is it *hallucinating*?

Ethics and Safety

- **Safety fine-tuning**: Does it obey nefarious commands, or give offensive outputs?
- **Ethical development**: Is the environment or IP considered?
- **Data privacy**: What happens to your data?

Practical factors

- **Context window**: How much information can it process?
- **Data recency**: How up-to-date is its knowledge?
- **Integration with other apps**
- **Usage limitations and quotas**

GRADIENT INSTITUTE

Better models & products = $$$$$

But the free versions are still very capable!

(With considerations)

GRADIENT INSTITUTE

2

What LLMs are good at

And what they’re not so good at

GRADIENT INSTITUTE

LLMs are *vibe* machines.

They are *not* fact machines.

When to use LLMs

Things they're good at

- **Ideation -** brainstorming, drafting and distilling
- **Editing -** suggesting a better sentence, tone or structure
- **Tutoring -** explaining complex concepts interactively
- **Synthesis -** summarising and finding connections
- **Code code code!** Minimal technical expertise required

Things they're less good at

- **Reliable facts -** they aren't trained to know facts. Make sure you use reliable sources!
- **Up to date information -** a model's knowledge is static
- **Citing sources -** unaware where their knowledge comes from
- **Having a human personality -** even Claude still sounds like a bot, don't forget your own personality!

Tips for responsible and effective use

- **Never place *full* trust in AI**
  - Use it to help, but *always verify its output*
- **Never input or upload highly sensitive information**
  - But generic internal documents are low risk!
- **They are good at inferring intention**
  - Don’t worry about super precise prompting, they adjust easily
- **Play with different models**
  - Experience is the best learning, so have fun with it!

GRADIENT INSTITUTE

2

Demo time

How we use LLMs

Demonstrations of effective LLM use

1. Hallucinations in AI models
2. Writing a grant proposal
3. Building simple code

A quick survey and we're done!

GRADIENT INSTITUTE

"We do not learn from experience. We learn from reflecting on experience."

–John Dewey

INDIVIDUAL WORK

IT WILL TAKE 1 MINUTE!

GRADIENT INSTITUTE

Thank you!
Any questions?

Dr Alberto Chierici

alberto@gradientinstitute.org

Liam Carroll

liam@gradientinstitute.org

Contact us:

info@gradientinstitute.org