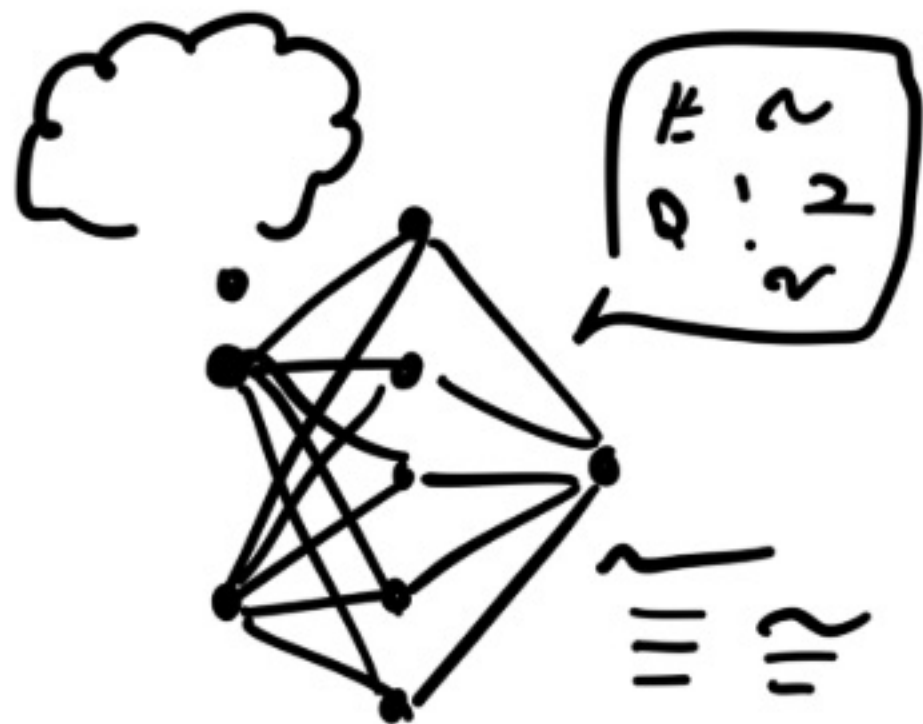# Learning goals

- develop your understanding of:
  - what LLMs are,
  - different types of "wild caught data" tasks that LLMs can help with,
  - how to use and check LLMs for specific data prepration tasks,
  - how to interact with LLMs in R using {ellmer}

How do we verify what LLMs are doing?

For a dishwasher, we consider:

- **...are the dishes clean?**

- **is there any dirt on the dishes?**

For LLMs...?

It depends on the (data) task!

Using LLMs for Data Preparation Tasks

1. Generating data wranging code

2. Transformating and generating new data

3. Modifying and augmenting existing data

- filling in missing data

- correct typos or inconsistencies

- creating new columns based on existing ones

lecture_llms-preparing-data-ellmer - RStudio

⊕ ▾   ⊕ ▾   📂 ▾   💾   🖨   → Go to file/function   ⊞ ▾   ⊞ ▾ Addins ▾   📄 lecture_llms-preparing-data-ellmer ▾

lecture_ellmer-demo.R ✕

↩   🔲   ☐ Source on Save   🔍   🪄 ▾   ▯   → Run   ↪→   ⊹ Source ▾   ▤

```r
1   library(ellmer)
2
3   # ---- session-are-chat ---
4
5   session <- chat_openai()
6
7   question <- "
8     How can I pick a random letter from A-Z.
9   "
10
11  session$chat(question)
12
13  session$chat("Return R code only")
14
15  session
16
17  # --- add-system-prompt ---
18
19  session_tidy_expert <- chat_openai(system_prompt = "
20    You are an expert R programmer
21    who prefers the tidyverse.
22    Only return code without explanation.
23  ")
24
25  session_tidy_expert$chat(question)
26
```

17:1   (Top Level) ⌄   R Script ⌄

Console   Terminal ✕   Background Jobs ✕

Ⓡ ▾ R 4.4.3 · ~/Dropbox/WORK/PROJECTS/MISC/lecture_llms-preparing-

```
```r
random_letter <- sample(LETTERS, 1)
print(random_letter)
```


This code utilizes the `sample` function to
randomly select one letter from the `LETTERS`
constant, which contains all uppercase letters of
the alphabet.
> session
<Chat turns=4 tokens=262/280>
── user ──────────────────────────────────────

    How can I pick a random letter from A-Z.
── assistant ─────────────────────────────────
To pick a random letter from A-Z, you can use
different methods depending on the tools or
programming languages you have available. Here are
a few examples:

Using Python
```python
import random
import string

random_letter =
random_choice(string.ascii_uppercase)
```

Environment   History   Connections   Git   Tutorial   Files   Plots   Packages   Help   Viewer   Presentation

📌 Remove pin

📶 Cynthia Huang (she/her)