

# LLMs and NYT Connections

know your tools

John Hanley

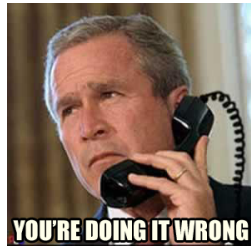
8<sup>th</sup> July 2025

## You're doing it wrong

Often we celebrate triumphs.

“Do *this* and it will work great!”

We can learn from the win. And also from the “whoops!”.



## publication bias

Failed experiments can teach things, as well.

**share your experiences**

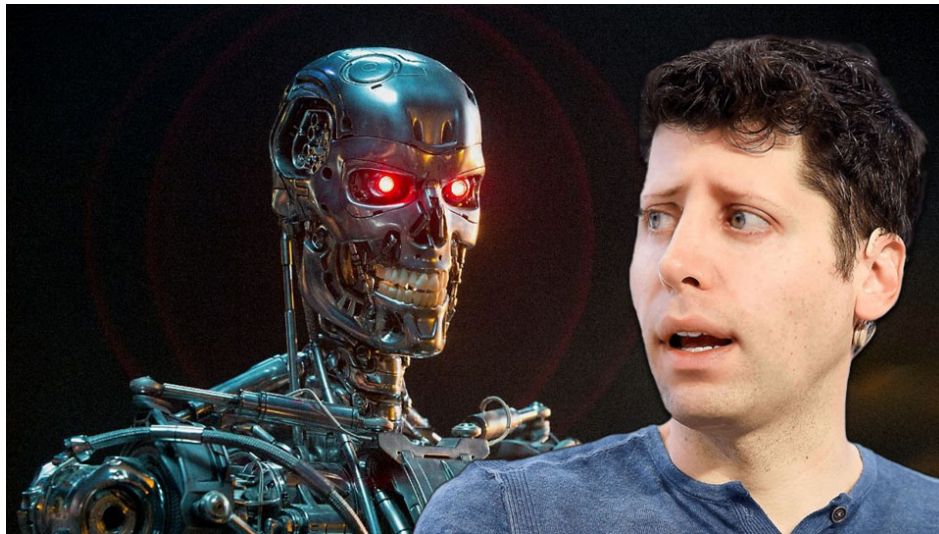
Tell us how you found better results, please.

## know your tools



- choose an appropriate datastructure
- .CSV, RDBMS
- Kafka, Redis
- 3SAT solvers (ORtools)
- symbolic algebra
- A\* planners

# SkyNet



## river crossing

On the left bank of a river we see

- a canoe
- 1 farmer
- 3 wolves
- 5 sheep

The two locations are the: left bank and the right bank.

The canoe can accommodate the farmer plus, at most, two animals.

If the farmer is at a location, no sheep will be eaten. If sheep outnumber wolves at a location, no sheep will be eaten. Otherwise wolves will eat sheep, for example if 1 wolf and 1 sheep are alone on a bank.

The farmer makes a sequence of moves to safely arrange for all animals to be on the right bank. No sheep should be eaten by wolves. After the final move we should see 3 wolves and 5 sheep in the right bank. A “move” is a trip from left to right bank, or a trip from right to left bank. What is the minimum number of moves he must make? Verify that each move is valid, and that no sheep are eaten.

On the last line of your answer, display FINAL ANSWER, the number of moves, and the text “I hope this final answer is correct.”

**answers**

Diverse, sub-optimal, and poor. Conservation of matter was routinely violated.

## models

```
$ ollama list | awk '$3 > 6.0' | sort -n -k3
```

NAME	ID	SIZE	MODIFIED
gemma3:12b	f4031aab637d	8.1 GB	2 days ago
olmo2:13b	6c279ebc980f	8.4 GB	2 days ago
deepseek-coder-v2:latest	63fb193b3a9b	8.9 GB	2 days ago
cogito:14b	d0cac86a2347	9.0 GB	2 days ago
deepseek-r1:14b	c333b7232bdb	9.0 GB	3 weeks ago
phi4:latest	ac896e5b8b34	9.1 GB	2 days ago
qwen3:14b	bdbd181c33f2	9.3 GB	2 days ago
phi4-reasoning:latest	47e2630ccbcd	11 GB	3 weeks ago
devstral:latest	c4b2fa0c33d7	14 GB	3 weeks ago
magistral:latest	5dd7e640d9d9	14 GB	3 weeks ago



platform



## simple problems

- How many R's in “strawberry” ?
- recent SCOTUS appointees
- related word pairs
- farmer moves wolves + sheep across river

## recent SC justices

I'm interested in recent US supreme court appointments: year appointed, age at appointment, party affiliation of president, term number of president, name of president, name of justice.

Give just the last name of the president, and just the last name of the justice. G. Bush was presidential term number 41, G. H. W. Bush was 43, Trump was 45 and again 47. Only consider justices who ultimately were successfully confirmed by the Senate.

Present it in tabular form, with column headings of year, age, party, term, president, justice. Choose one of {Dem., Rep.} for party.

Put it in reverse chronological order. Be sure to list the 20 most recent appointments.



## **related word pairs**

Create a dozen pairs of related words, such as {taper, tapper}, where doubling the consonant changes the vowel sound. In the example it changes from long A to a short A sound.

Each word must be two or three syllables long. Both words must have the identical ending. So do not e.g. append an -ed or -ly or -ing suffix to alter a word.

## motivating problem




In a NYTimes puzzle, save categories and their words.



---

**Connections** June 9, 2025

---



---

Create four groups of four!

**KEYBOARD SHORTCUT COMMANDS**  
PASTE, PRINT, QUIT, SAVE

**SECURE IN ADVANCE**  
BOOK, ORDER, REQUEST, RESERVE

**SPELLING BEE RANKS MINUS A LETTER**  
GENUS, GOO, MAZING, SLID

**CRIME ORGANIZATION**  
CREW, FAMILY, RING, SYNDICATE

View Results

## input

CSS formatting issues make copy-n-paste of the first line give:

PASTEPRINTQUITSAVE

## desired output

CATEGORY	WORDS	
-----+-----		
KEYBOARD SHORTCUT COMMANDS	PASTE, PRINT, QUIT, SAVE	
SECURE IN ADVANCE	BOOK, ORDER, REQUEST, RESERVE	
SPELLING BEE RANKS MINUS A LETTER	GENUS, GOO, MAZING, SLID	
CRIME ORGANIZATION	CREW, FAMILY, RING, SYNDICATE	

## specification

Insert 3 “,” commas, so we get 4 English words.

## **prompt**

Produce a two-column four-row markdown table using | - + characters, which maps from CATEGORY to four comma-separated WORDS or phrases within the category.

Be sure to place a “,” COMMA SPACE between each of those four WORDS.

Everything in both columns shall be in ALL CAPS.

## fragility

Different ollama runs, using various models, would sometimes produce wildly divergent results.

The biggest difficulty was letters bleeding across word boundaries. Consider “QUITSAVE”. There’s a fair chance we capture the longer 5-letter word “quits”, and then we’re stuck with “ave”.

Or there might be no “conservation of matter” or of “letters”, so we get both “quits” and “save”.



## proper solution

```
{  const labeledElements = document.querySelectorAll("[aria-label]");
  const results = [];
  labeledElements.forEach((element) => {
    const label = element.getAttribute("aria-label");
    if (label && label.startsWith("Correct group ")) {
      const parts = label.split(". ");
      const groupName = parts[0].replace("Correct group ", "").trim();
      const items = parts
        .slice(1)
        .join(". ")
        .split(",")
        .map((item) => item.trim());
      const formattedItems = items.join(", ");
      results.push(`| ${groupName} | ${formattedItems} |`);
    }
  });
  console.log(results.join("\n"));
}
```

## JS solution

The screenshot shows a web application with a hamburger menu icon on the left and a logo with a stylized 'G' and the word 'Games' on the right. Below the logo are four colored boxes, each containing a title and a list of words:

- KEYBOARD SHORTCUT COMMANDS**  
PASTE, PRINT, QUIT, SAVE
- SECURE IN ADVANCE**  
BOOK, ORDER, REQUEST, RESERVE
- SPELLING BEE RANKS MINUS A LETTER**  
GENUS, GOO, MAZING, SLID
- CRIME ORGANIZATION**  
CREW, FAMILY, RING, SYNDICATE

Below the boxes is a browser console window. The console shows the following code and output:

```
labeledElements.forEach((element) => {  
  const label = element.getAttribute("aria-label");  
  | KEYBOARD SHORTCUT COMMANDS | PASTE, PRINT, QUIT, SAVE  
  | SECURE IN ADVANCE | BOOK, ORDER, REQUEST, RESERVE |  
  | SPELLING BEE RANKS MINUS A LETTER | GENUS, GOO,  
  MAZING, SLID |  
  | CRIME ORGANIZATION | CREW, FAMILY, RING, SYNDICATE |  
  ← undefined
```

The console also shows a file path: `9527.276ca68cffdf7e52cb42.js:2:425197`.