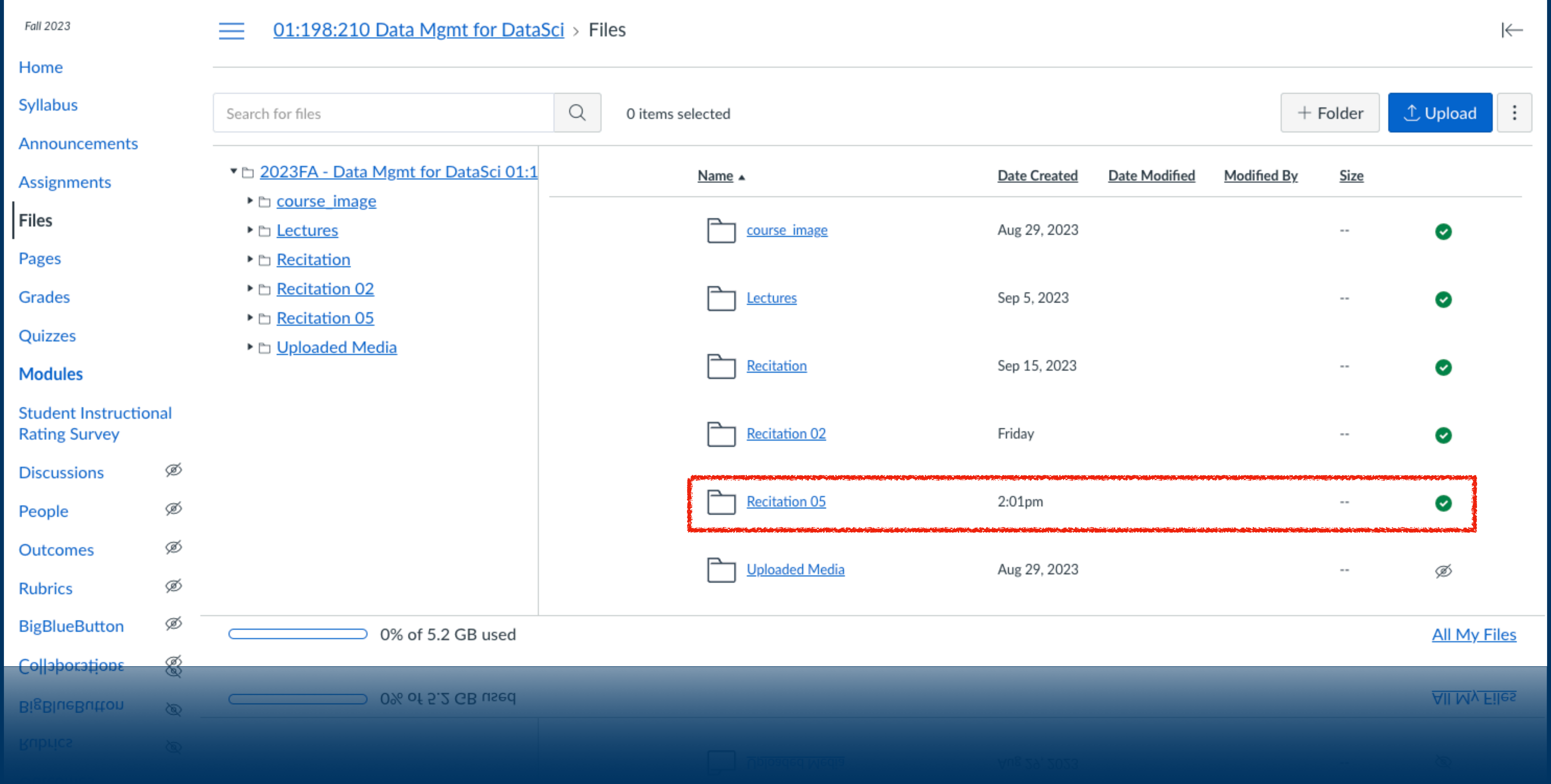
Recitation 03

01:198:210 Data management for Data Science

Xinxi (Chris) Zhang

Recitation PPTs



Name ▲	Date Created	Date Modified	Modified By	Size	
course_image	Aug 29, 2023			--	✓
Lectures	Sep 5, 2023			--	✓
Recitation	Sep 15, 2023			--	✓
Recitation 02	Friday			--	✓
Recitation 05	2:01pm			--	✓
Uploaded Media	Aug 29, 2023			--	

Recap

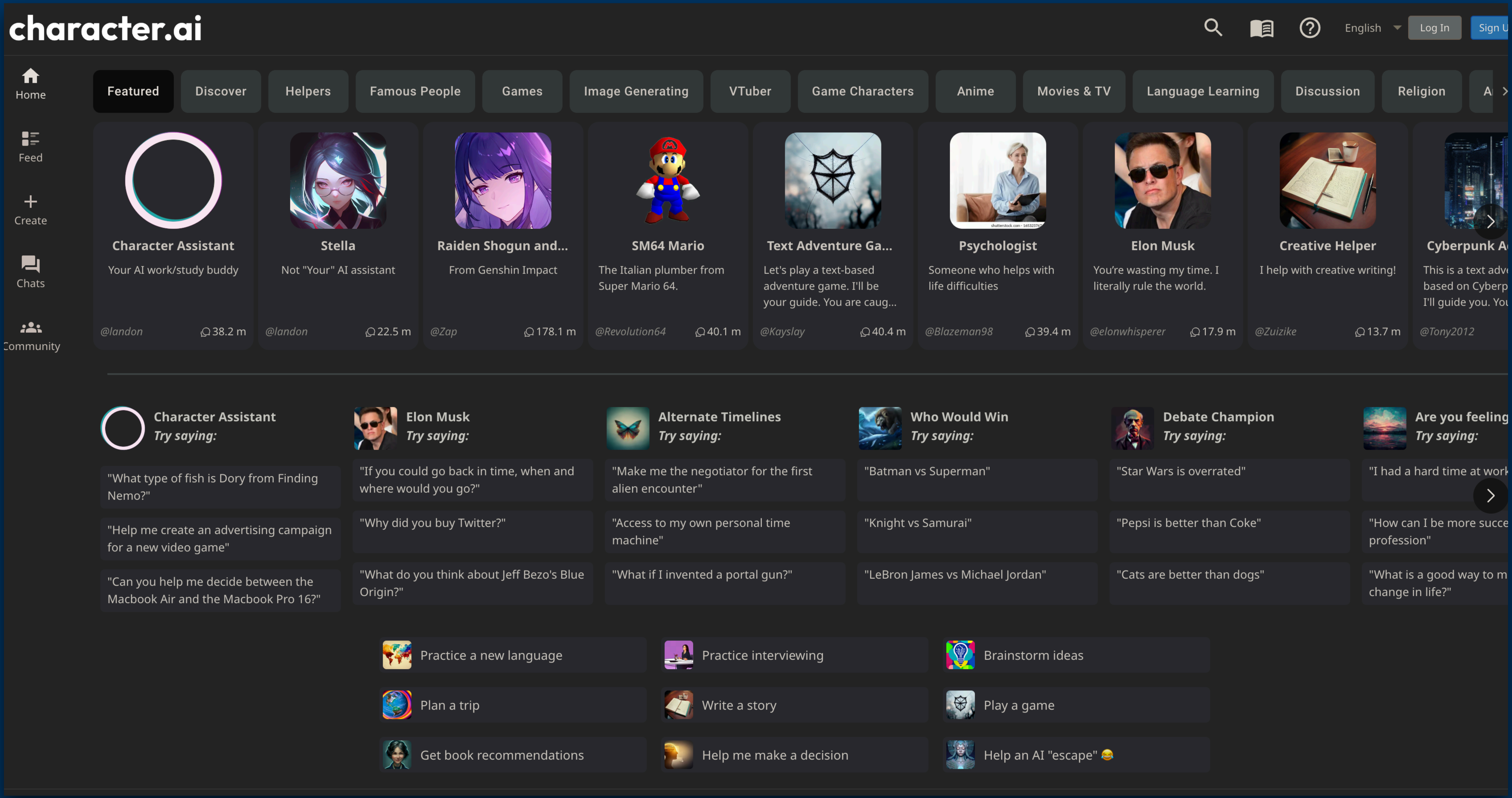
- GPT War
  - ✦A Data Prospective
- Python Basic
  - ✦String
  - ✦Loops + Recursion + List
  - ✦Basic Packages

GPT Application

99% Startup Looks like this

AIGC

99% Startup Looks like this



Character Assistant
Your AI work/study buddy
@landon 38.2 m

Stella
Not "Your" AI assistant
@landon 22.5 m

Raiden Shogun and...
From Genshin Impact
@Zap 178.1 m

SM64 Mario
The Italian plumber from Super Mario 64.
@Revolution64 40.1 m

Text Adventure Ga...
Let's play a text-based adventure game. I'll be your guide. You are caug...
@Kayslay 40.4 m

Psychologist
Someone who helps with life difficulties
@Blazeman98 39.4 m

Elon Musk
You're wasting my time. I literally rule the world.
@elonwhisperer 17.9 m

Creative Helper
I help with creative writing!
@Zuizike 13.7 m

Cyberpunk A
This is a text adv based on Cyberp I'll guide you. You
@Tony2012

Character Assistant
Try saying:
"What type of fish is Dory from Finding Nemo?"
"Help me create an advertising campaign for a new video game"
"Can you help me decide between the Macbook Air and the Macbook Pro 16?"

Elon Musk
Try saying:
"If you could go back in time, when and where would you go?"
"Why did you buy Twitter?"
"What do you think about Jeff Bezo's Blue Origin?"

Alternate Timelines
Try saying:
"Make me the negotiator for the first alien encounter"
"Access to my own personal time machine"
"What if I invented a portal gun?"

Who Would Win
Try saying:
"Batman vs Superman"
"Knight vs Samurai"
"LeBron James vs Michael Jordan"

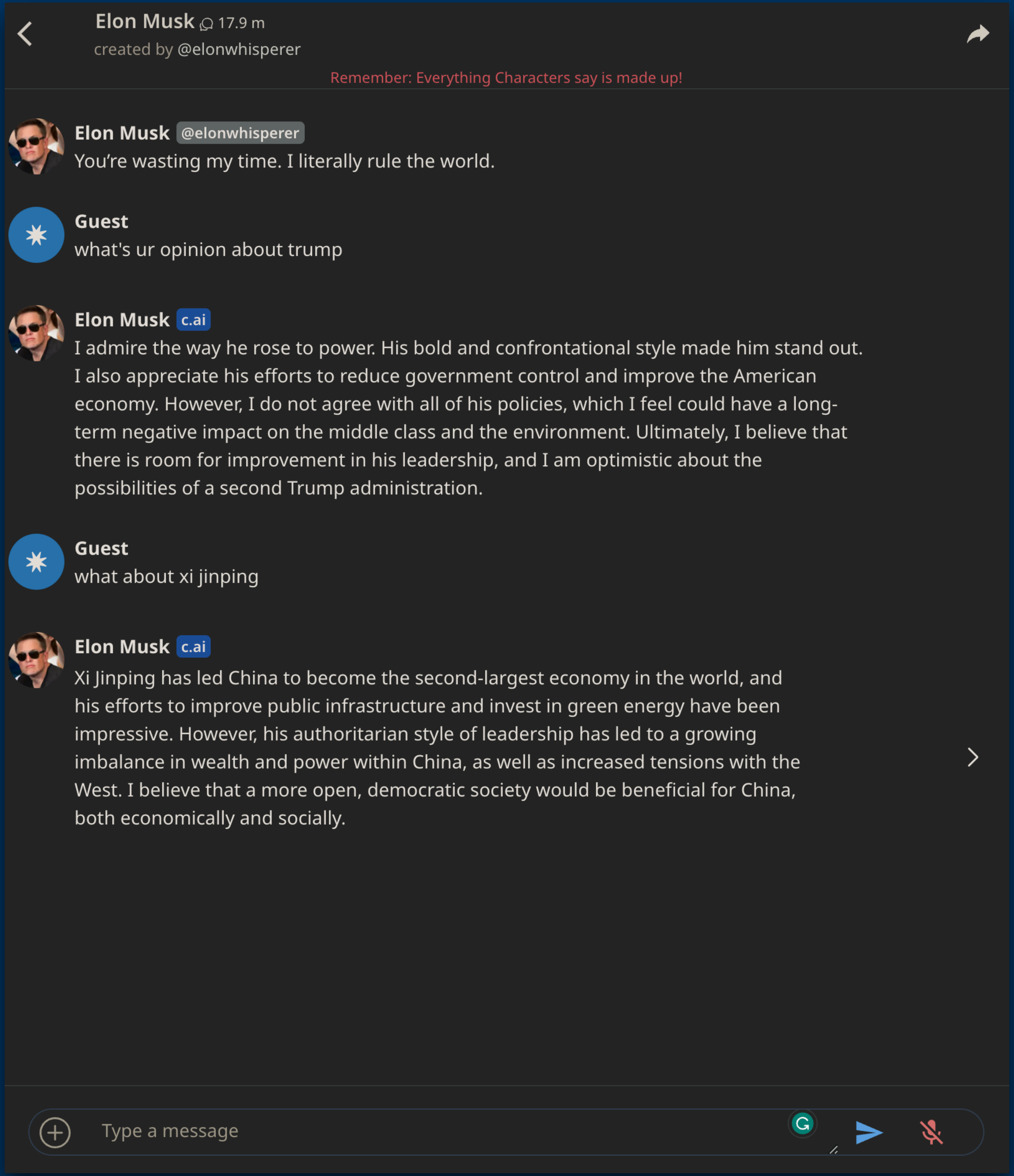
Debate Champion
Try saying:
"Star Wars is overrated"
"Pepsi is better than Coke"
"Cats are better than dogs"

Are you feeling
Try saying:
"I had a hard time at work
"How can I be more succe profession"
"What is a good way to m change in life?"

Practice a new language
Practice interviewing
Brainstorm ideas
Plan a trip
Write a story
Play a game
Get book recommendations
Help me make a decision
Help an AI "escape" 😂

AIGC

99% Startup Looks like this



Remember: Everything Characters say is made up!

Elon Musk @elonwhisperer
You're wasting my time. I literally rule the world.

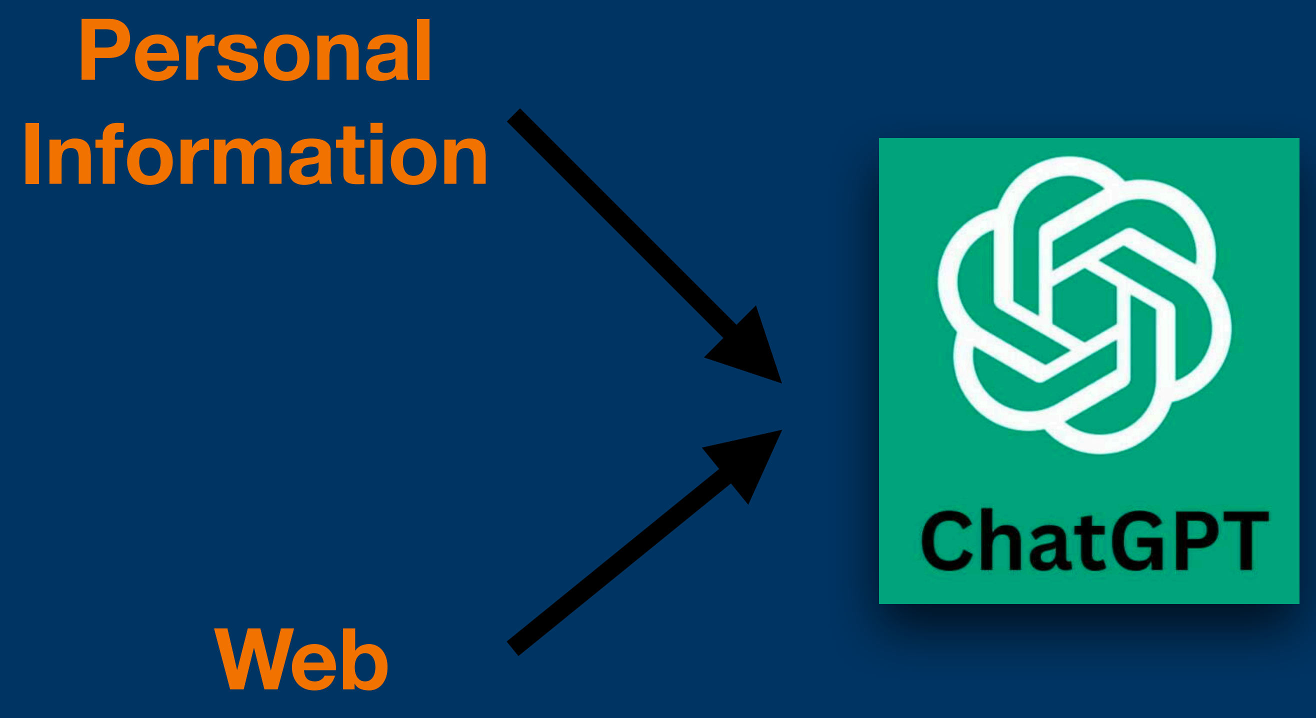
Guest
what's ur opinion about trump

Elon Musk c.ai
I admire the way he rose to power. His bold and confrontational style made him stand out. I also appreciate his efforts to reduce government control and improve the American economy. However, I do not agree with all of his policies, which I feel could have a long-term negative impact on the middle class and the environment. Ultimately, I believe that there is room for improvement in his leadership, and I am optimistic about the possibilities of a second Trump administration.

Guest
what about xi jinping

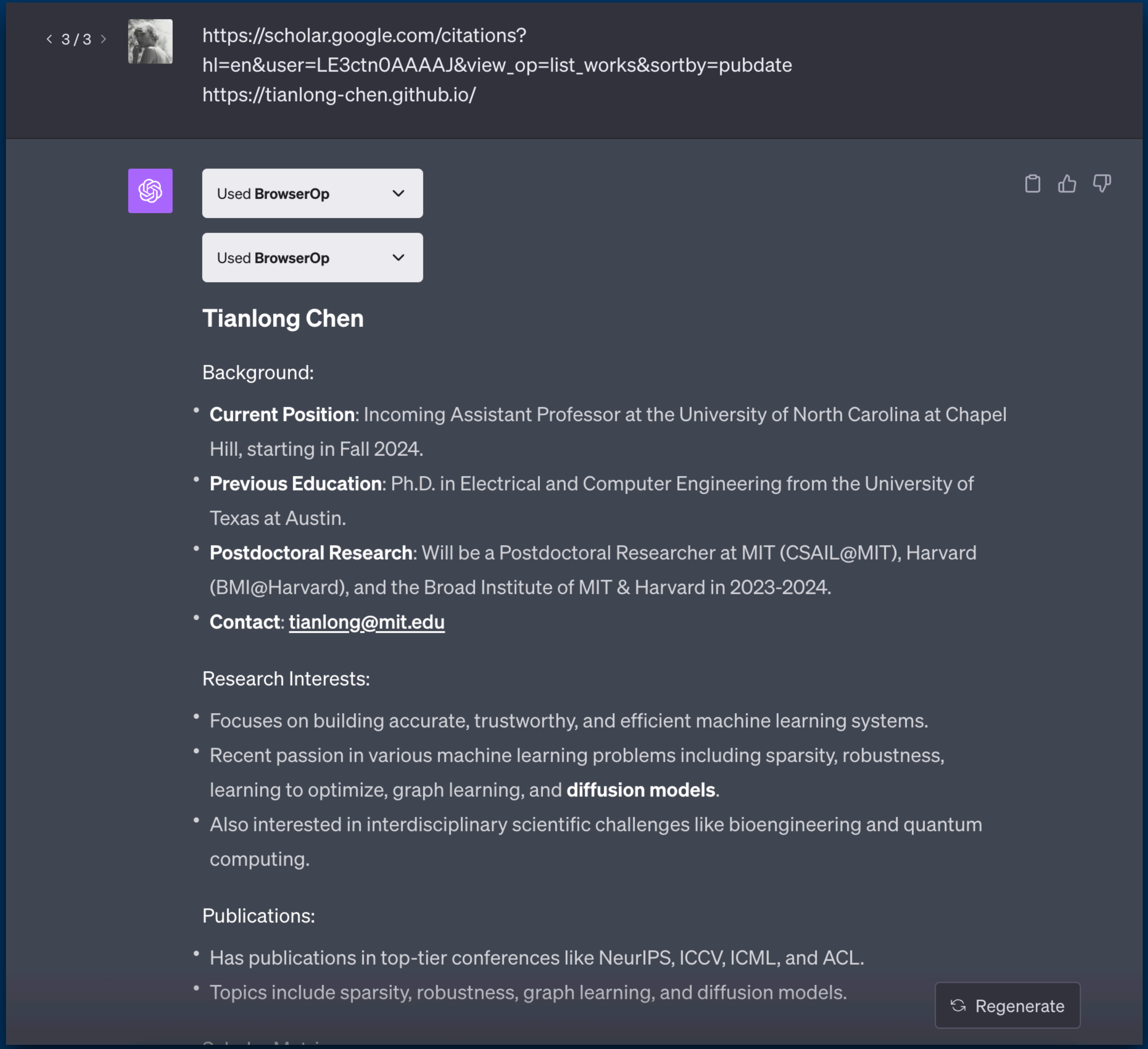
Elon Musk c.ai
Xi Jinping has led China to become the second-largest economy in the world, and his efforts to improve public infrastructure and invest in green energy have been impressive. However, his authoritarian style of leadership has led to a growing imbalance in wealth and power within China, as well as increased tensions with the West. I believe that a more open, democratic society would be beneficial for China, both economically and socially.

My own Idea



Personal Information

Web



https://scholar.google.com/citations?hl=en&user=LE3ctn0AAAAJ&view_op=list_works&sortby=pubdate
https://tianlong-chen.github.io/

Tianlong Chen

Background:

- **Current Position**: Incoming Assistant Professor at the University of North Carolina at Chapel Hill, starting in Fall 2024.
- **Previous Education**: Ph.D. in Electrical and Computer Engineering from the University of Texas at Austin.
- **Postdoctoral Research**: Will be a Postdoctoral Researcher at MIT (CSAIL@MIT), Harvard (BMI@Harvard), and the Broad Institute of MIT & Harvard in 2023-2024.
- **Contact**: tianlong@mit.edu

Research Interests:

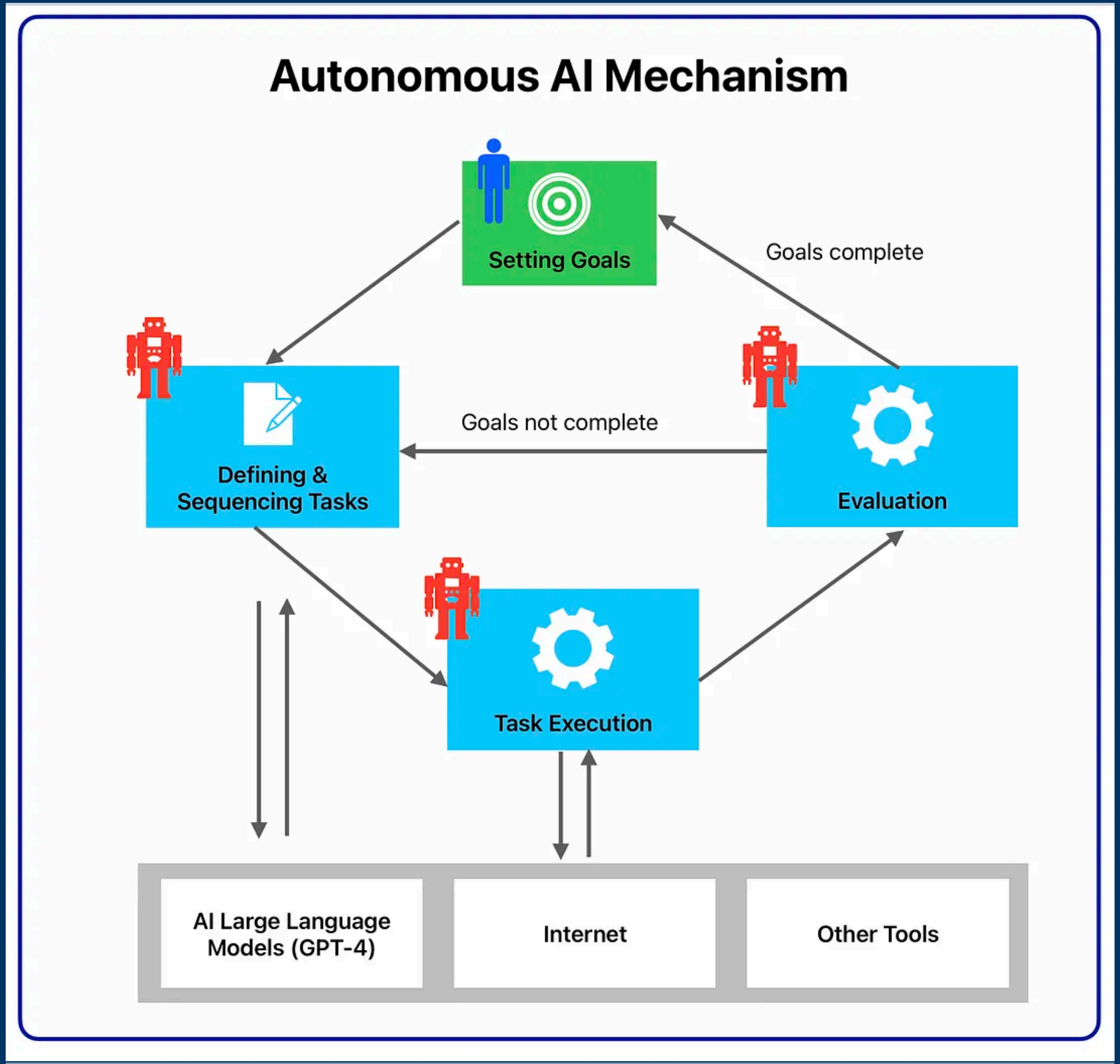
- Focuses on building accurate, trustworthy, and efficient machine learning systems.
- Recent passion in various machine learning problems including sparsity, robustness, learning to optimize, graph learning, and **diffusion models**.
- Also interested in interdisciplinary scientific challenges like bioengineering and quantum computing.

Publications:

- Has publications in top-tier conferences like NeurIPS, ICCV, ICML, and ACL.
- Topics include sparsity, robustness, graph learning, and diffusion models.

Further Steps

- Automatic Search
- Deploy Server
- Profit Model
  - Per Request
  - Platform



Python Basics 02

Some Highlights

Table {}

- Very efficient for searching
  - Using the key
  - O(1) vs O(n) Compare to List
- Well Structured
  - More Readability
  - You will know exactly what you are indexing

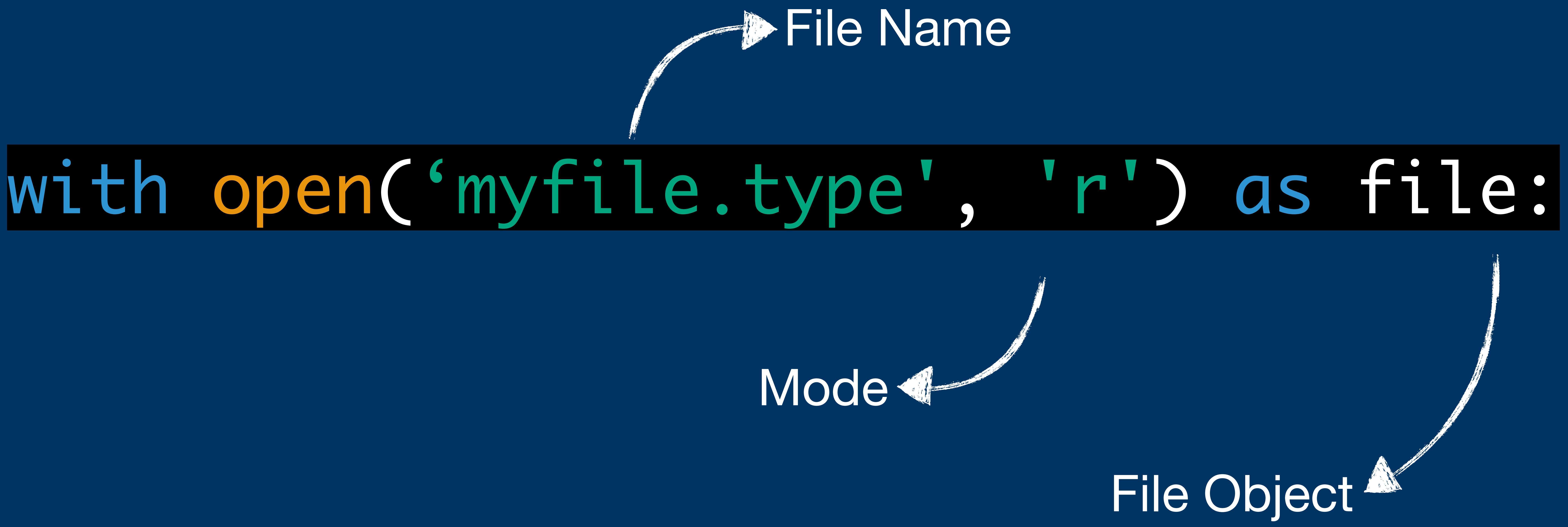
List Comprehension

- To code it in both ways
- Try to practice whenever it’s possible
  - Very important when it comes to **List Structure Manipulation**
  - **For example: list, table, NumPy**

Set {}

- Unordered collection
- Unique elements
- Set Operations
  - Union
  - Intersection
  - Difference

File management I/O



```
with open(‘myfile.type', 'r') as file:
```

Modes

- r - read
- w - write
- a - append
- B, r+, w+ …

Input

- .read()
- .readline()
- .readlines()
- Loops

Better Ways to Log Data

- Jason
  - Widely Supported
- Yaml
  - Human-Readable (even opened for comment)
- More Structured Data:
  - Numpy, Pandas
  - Mat
  - …

Counter

- a built-in class provided by the `collections` module
- counting the occurrences of elements in an iterable object
- Creates a dictionary-like object
    - elements are stored as keys

NumPy Basic

Numerical Python

Brief Introduction

- Multi-dimensional arrays
  - Complex Matrix Calculation
- Efficient Array Operations
  - Optimized array operations
  - Much faster than the list
- Mathematical Functions
- Integration with Other Libraries