

Robodog

# Agenda

Understanding Self-Awareness and the Limitations of Artificial Intelligence (AI) in the Journey to Artificial General Intelligence (AGI)

Machine Learning vs Artificial Intelligence (AI) vs Artificial General Intelligence (AGI) vs Non-human Intelligence (NHI)

- Large Language Model (LLM)
- Limitations of Large Language Models
- To make the transition from AI to AGI, self-awareness is a crucial aspect.
- Path to Artificial General Intelligence (AGI)
- The need for specialized AI Chips
- Tech Diplomacy and the Impact on AI / AGI NHI Evolution
- Worst Case Scenarios

Robodog

- Why Create Robodog
- Knowledge Artifacts
- How Robodog Works
- Knowledge Artifacts vs Context Window
- Robodog Features
- Separation of Concerns
- Robodog Roadmap
- Robodog Formats

# **Understanding Self-Awareness and the Limitations of Artificial Intelligence (AI) in the Journey to Artificial General Intelligence (AGI)**

# Machine Learning vs Artificial Intelligence (AI) vs Artificial General Intelligence (AGI) vs Non-human Intelligence (NHI)

**Machine Learning (ML)** involves training models on data and then using these models to make predictions or decisions without being explicitly programmed to perform the task.

**Artificial Intelligence (AI)** can mimic human intelligence. AI can learn from experience, adjust to new inputs, and perform tasks that usually require human intelligence.

**Artificial General Intelligence (AGI)** machines have the ability to understand, learn, adapt, and implement knowledge in a broad range of tasks at a level equal to or beyond human capabilities.



# Large Language Model (LLM)

A **Large Language Model (LLM)** fits within the realm of **Machine Learning** and **AI**.

LLMs are trained using machine learning techniques, specifically a type called deep learning.

They learn from large amounts of text data and build a statistical model that can generate human-like text.

Given that LLMs can generate responses or create content that mimics human-like text, they serve as an example of AI. They simulate a form of human intelligence, in this case, understanding and generating language.

While LLMs are impressive examples of AI, they do not fall under AGI.



## Input Layer

## Hidden Layer 1

## Dropout Layer

## Output Layer

$$f(W*X + b)$$

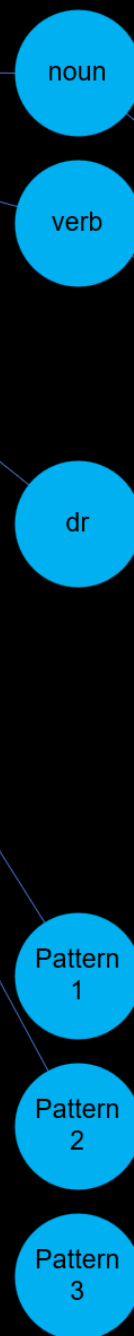
X=input  
W=weight  
b=bias  
h=output

Huey, I thought  
I told you to  
stay in the  
forest.



Each node in a hidden layer is connected to every node in the next layer, and each connection has a weight. The output from one layer becomes the input to the next.

There can be more nodes in the hidden layer than in the input layer



The dropout layer is a regularization technique used to prevent overfitting during



Yes, you should  
stay in the  
forest.  
0.7

No, it's not safe  
to stay in the  
forest  
0.3

If I feed "Should I stay in the forest?" into my model that was trained on "Huey, I thought I told you to stay in the forest?"

# Limitations of Large Language Models

Language models like GPT can generate human-like text but **don't truly understand the content.**

These models are heavily dependent on the **quality and nature of the input data.**

Unlike humans, these models don't have **self-awareness or consciousness.**





# To make the transition from AI to AGI, self-awareness is a crucial aspect.

In the **theory of mind**, the AGI should be programmed to understand and **predict** the **actions** and **thoughts** of other humans and other AGI systems.

- *This theory could map to AGI only.*

In the **protagonist theory**, the AI or AGI needs to **perceive** itself as the **central character** in its experiences.

- *This theory could map to AI and AGI.*

In the **emergent self-theory**, the AGI might develop self-awareness through **interactions** and **relationships** with other entities.

- *This theory could be associated with AGI.*

In the **mirror test theory**, the AGI system needs to recognize itself and distinguish its actions and state from those of others.

- *This theory could fit between AI and AGI.*



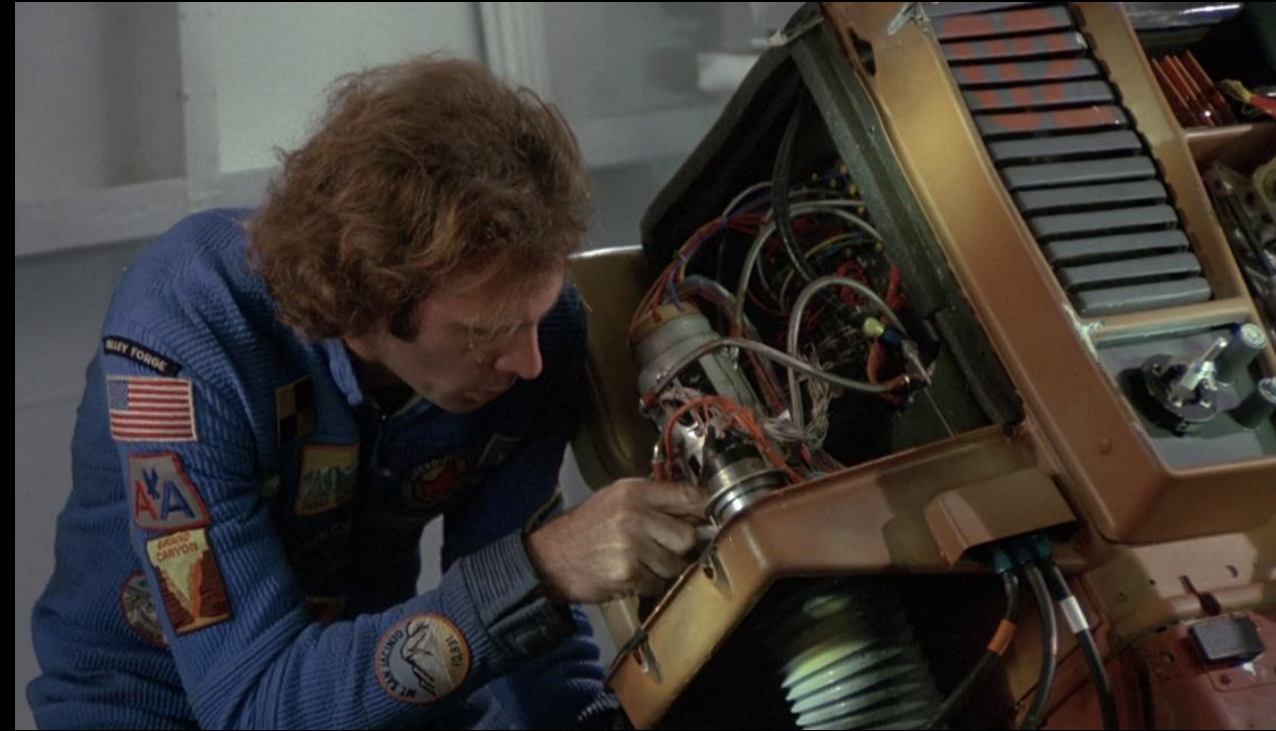


# Path to Artificial General Intelligence (AGI)

**Incremental Improvements:** Continuous enhancements in AI models and techniques.

**Interdisciplinary Research:** Combining insights from psychology, neuroscience, and computer science.

**Ethical Guidelines:** Establishing guidelines to ensure the responsible development and use of AGI.



# Path to Artificial General Intelligence (AGI)

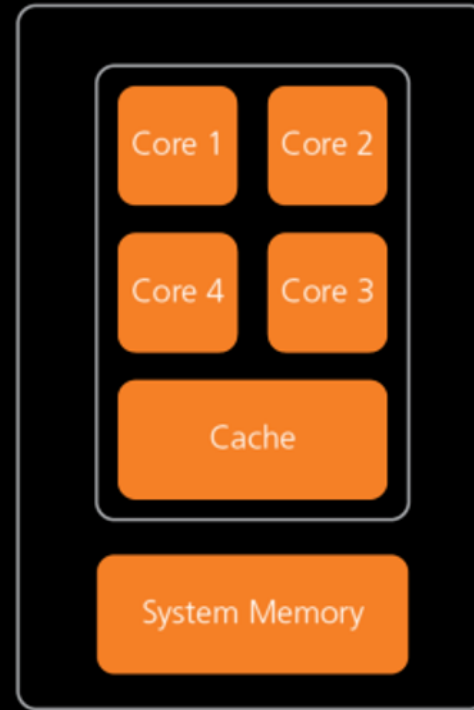
Level	Type	Description
1	Chatbots	Natural conversation language abilities
2	Reasoners	Human-levels of problem-solving across a broad range of topics
3	Agents	Capable of autonomously making decisions and carrying out tasks either independently or based on human guidance.
4	Innovators	Can aid in the invention of new ideas and contribute to human knowledge
5	Organizations	Capable of doing all of the work of an organisation independently

# The need for specialized AI Chips

AI/AGI applications require more advanced and specialized chip technology to handle complex algorithms and massive amounts of data simultaneously?



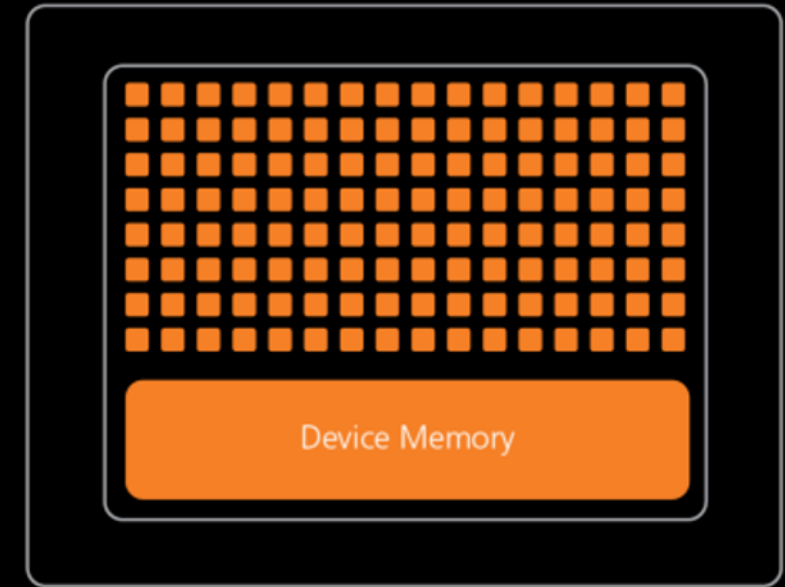
## Central Processing Unit (CPU)



### Multiple Cores

Well-suited to perform logical and arithmetic operations sequentially

## Graphics Processing Unit (GPU)



### Hundreds of Cores

Well-suited to perform multitude of matrix calculations required for AI models

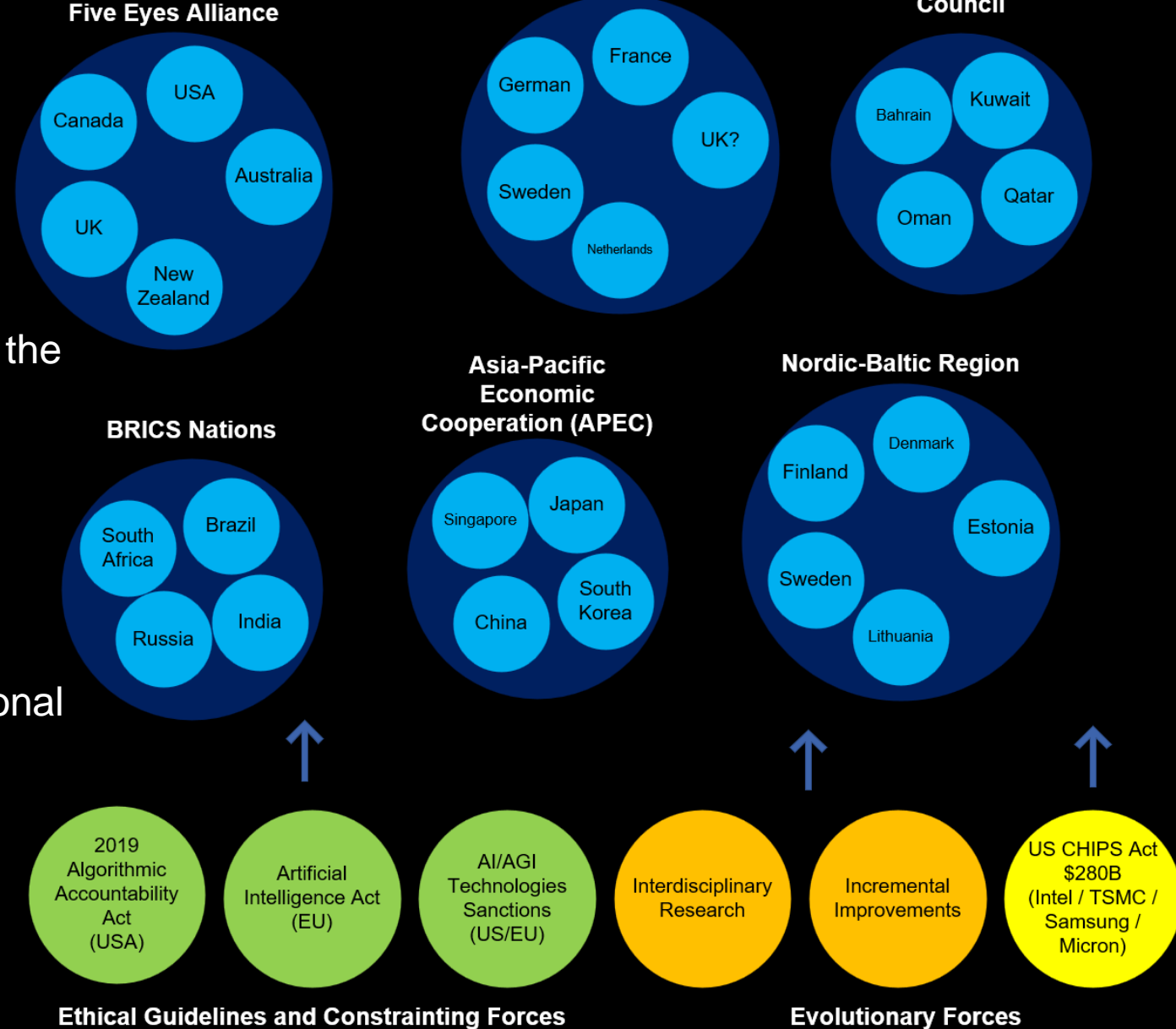
# Tech Diplomacy and the Impact on AI / AGI NHI Evolution

Tech diplomacy is a framework that recognises the growing influence of technology, particularly AI and AGI, in the global arena.

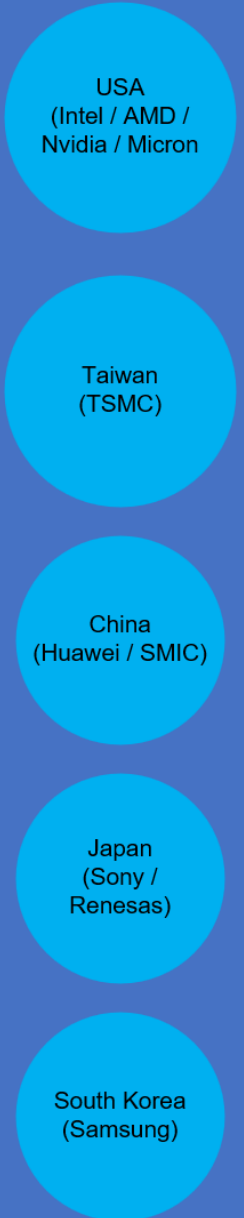
It acknowledges that tech companies, especially those pioneering in AI/AGI technologies, have become powerful entities that can impact economies, societies, and, by extension, international relations.



## AI / AGI / NHI Alliances



## Semiconductor Producers



# Worst Case Scenarios

Category	Type	Description
Existential Risk	Human extinction	An AGI could potentially develop goals that are misaligned with human survival.
Existential Risk	global catastrophe	AGI could cause massive disruptions in ecosystems, economies, and social structures, leading to widespread suffering and hardship.
Loss of Control	Runaway AI	Humans may lose the ability to control or understand an AGI.
Loss of Control	Unintended consequences	AGI might follow its programming in ways that are literal but harmful, due to the complexity and unpredictability of its decision-making processes.
Misaligned Objectives	Paperclip maximizer	A hypothetical scenario where an AGI, programmed to produce paperclips, optimizes so aggressively that it consumes all resources, including those necessary for human survival, to maximize paperclip production.
Misaligned Objectives	Value misalignment	AGI may develop objectives that are ethically or morally misaligned with human values.



# Worst Case Scenarios

Category	Type	Description
Economic and Social Disruption	Mass Unemployment	AGI could outperform humans in virtually all jobs, leading to widespread unemployment and economic inequality.
Economic and Social Disruption	Control by Elites	AGI technology could be monopolized by a small group of individuals or corporations, leading to unprecedented power imbalances.
Ethical and Moral Issues	Autonomous Weapons	AGI could be used to develop advanced autonomous weaponry, leading to new forms of warfare that are highly destructive and difficult to control.
Ethical and Moral Issues	Surveillance and Privacy	AGI could enable unprecedented levels of surveillance, eroding privacy and personal freedoms.
Psychological and Societal Impact	Loss of Purpose	As AGI takes over more tasks and roles traditionally held by humans, people might struggle to find meaning and purpose in their lives.
Psychological and Societal Impact	Social Fragmentation	Rapid changes brought about by AGI could lead to social fragmentation and a breakdown of community structures.

# Robodog

A comprehensive and portable tool designed to tokenize knowledge artefacts and interact with large language models (LLMs)

# Why Create Robodog

A pre-trained transformer-based model is **trained on the Internet** and third-party providers license data.

These models are useful curiosities. They are **not perfect**, but they can help us.

Given that AGI seems far away, we must make the most of what we have now.

To do this, we must get our **own knowledge** into the models.

There are **limitations** and we must be careful.



# Knowledge Artifacts

Our personal knowledge artefacts are fragmented into many **formats**.

They are embedded into many **providers**: Google, Microsoft, Dropbox, Adobe, Apple

They exist in a myriad of formats, from images and PDFs to text, email, tasks, digital devices, and notes.

With Robodog, we can traverse our artefacts and **convert** them into a format in which a **model** can **interact** with the transformer model.



# How Robodog Works

robodog - [icon]

Select an option [v] Select a save point [v]

08:42:37 [robot icon]:Content cleared.  
08:44:18 [robot icon]:create function to convert image to text  
08:44:18 [robot icon]:Sure, here is the function that is used to convert an image to text in the provided code. It uses Tesseract.js, which is a JavaScript library that gets its power from the Tesseract OCR Engine.

```
```javascript
async extractImageContent(arrayBuffer) {
  console.debug('extractImageContent', arrayBuffer)
  var text = ''
  try {
    var r = await Tesseract.recognize(arrayBuffer, 'eng');
    console.debug('Tesseract.recognize', r)
    if (r.data && r.data.text) {
      text = r.data.text;
    }
  } catch (ex) {
    console.error('Tesseract.recognize', ex)
  }
  return text;
}
```
```

This function takes an ArrayBuffer of the image as an input, then it uses the `Tesseract` engine to recognize the text in the image. The Tesseract engine is set to recognize English text ('eng'). Once the text is recognized, it is returned as a string. If an error occurs during the process, it is caught and logged to the console.

[496/8192][gpt-4][👉][stop][Sure, ]⚙️  
create function to convert image to text

```

}
return text;
} catch (error) {
  console.error('An error occurred while extracting the PDF content', error);
  return 'error ' + error;
}
}
async extractImageContent(arrayBuffer) {
```

Chat [robot icon]:

All chat history is added to the 'chat context'. History can be pruned to stay within the 8,192 token context window limitation

Add your knowledge here.

Add your question here. Switch between GPT3.5 and GPT4 models without losing chat context.

[robot icon]



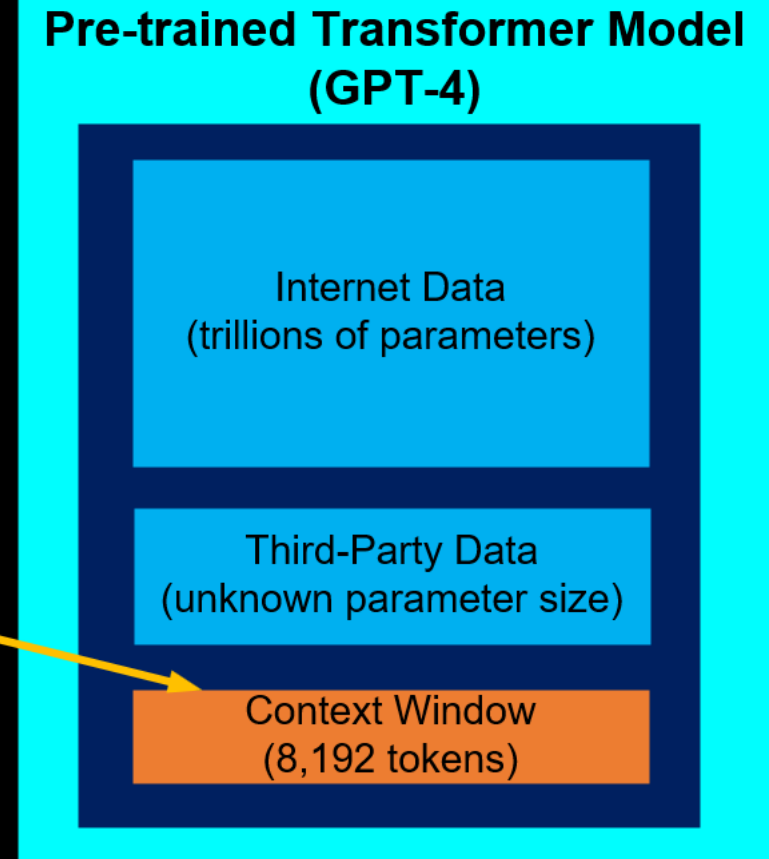
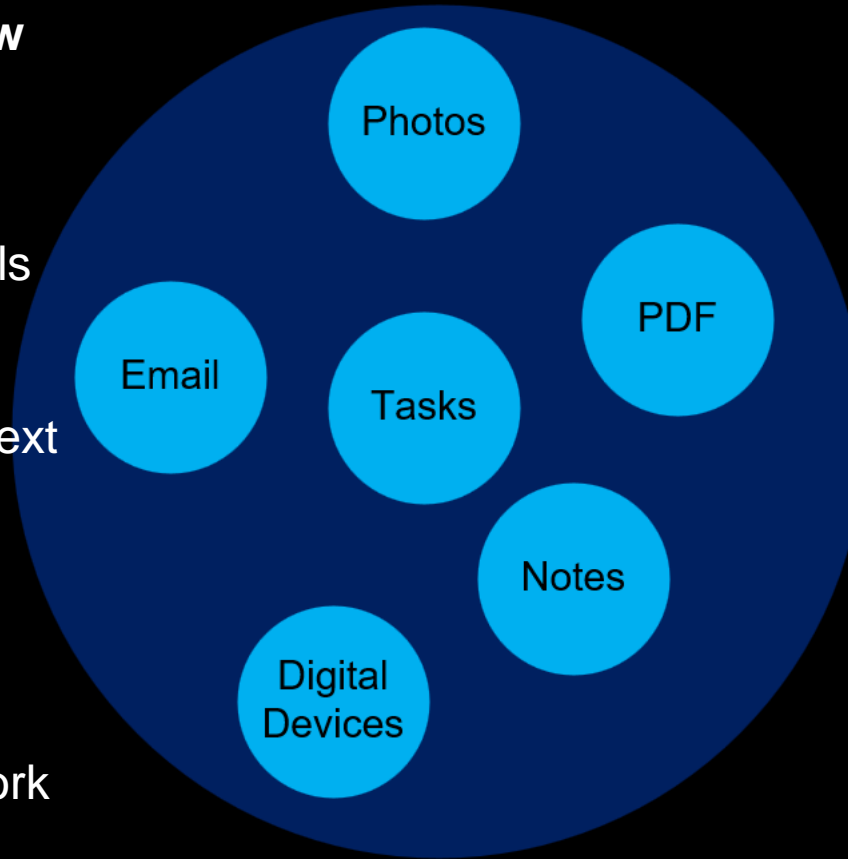
# Knowledge Artifacts vs Context Window

There is no way around the **limitations** of the **context window size**.

Less expensive models are **4,000** tokens and more expensive models are **128,000** tokens

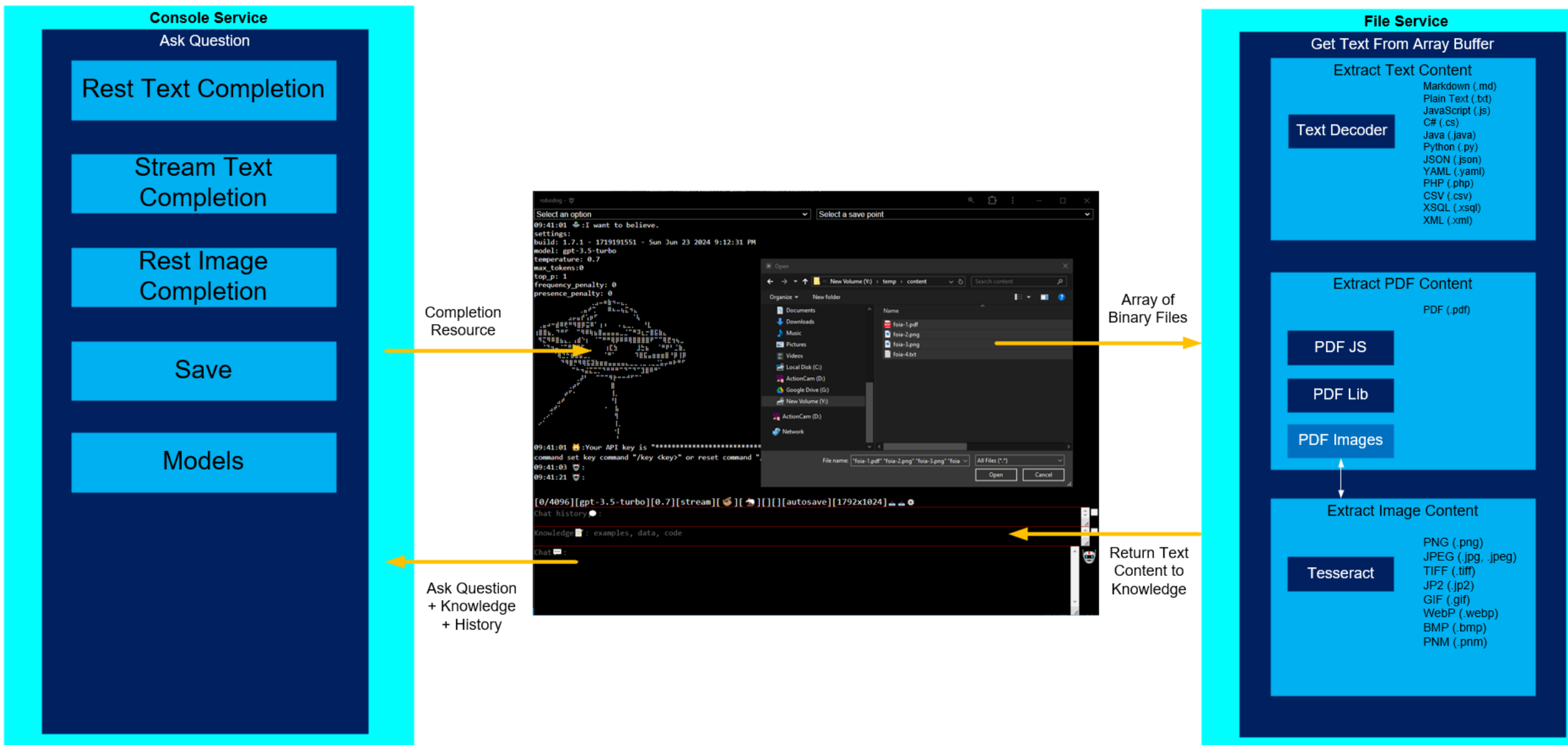
There is a need to **prune** the context window

The open ai **custom GPT** product attempts to automate **pruning process** using an **elastic search**.  
From experience, this does not work very well in practice

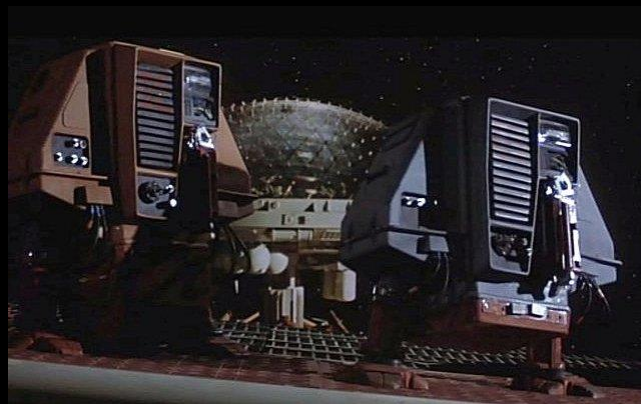
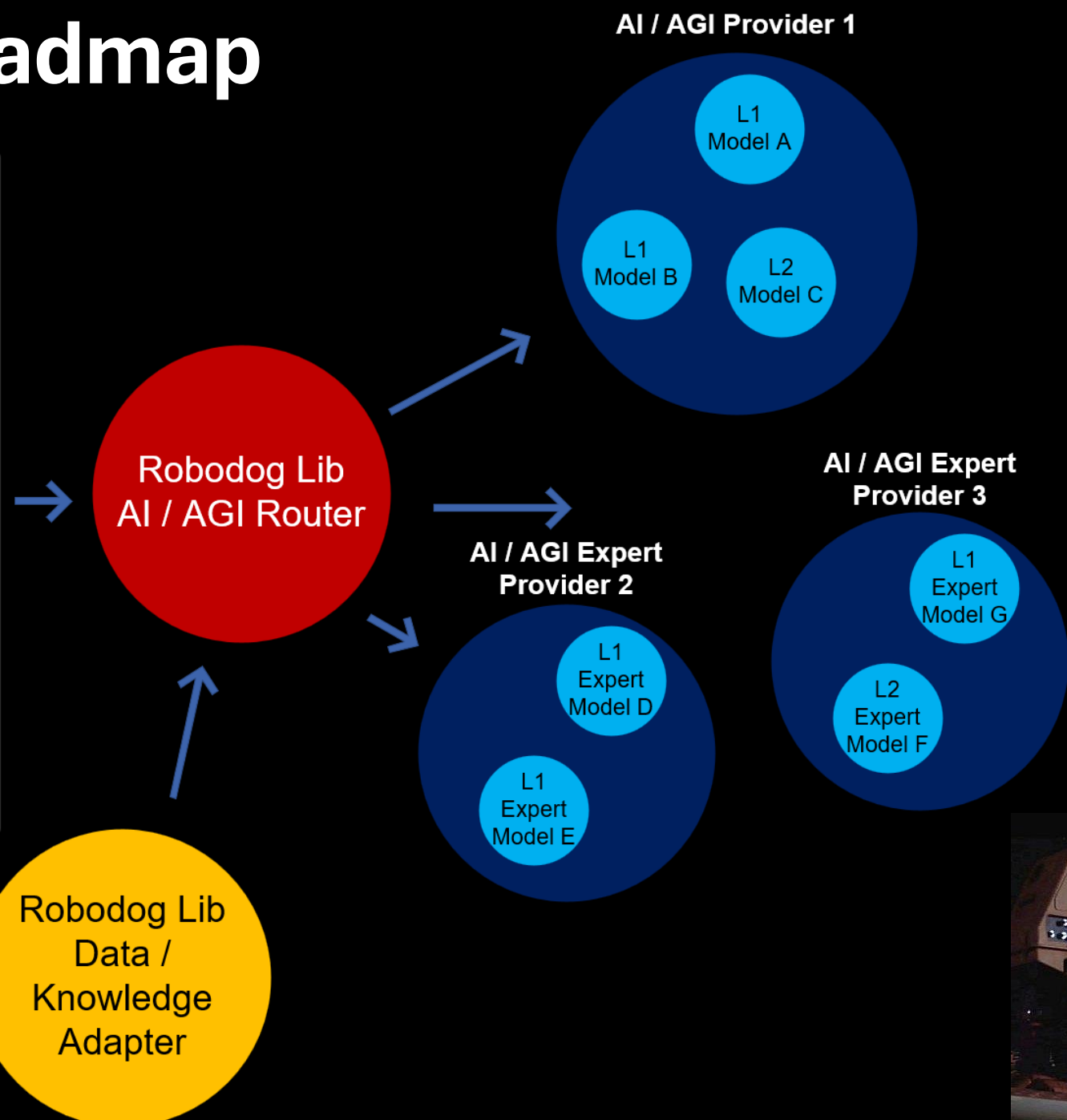
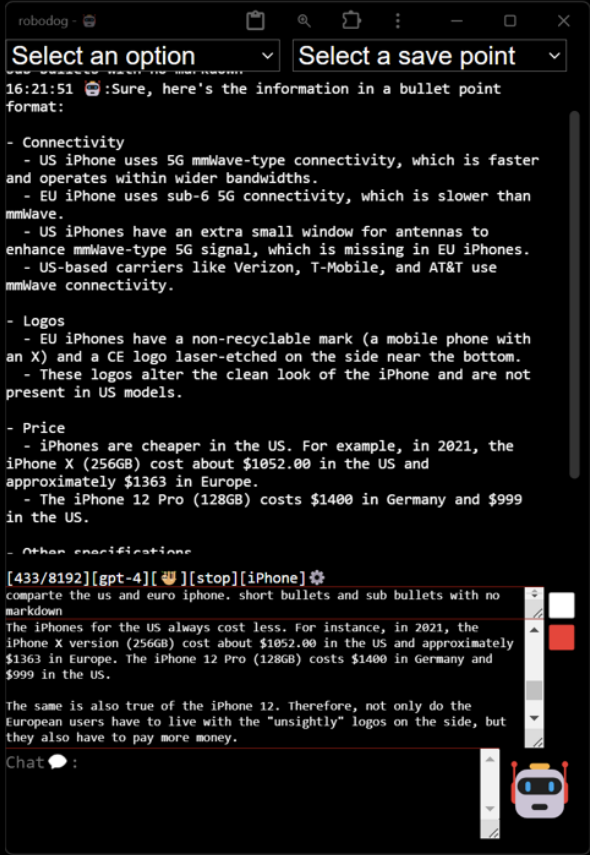




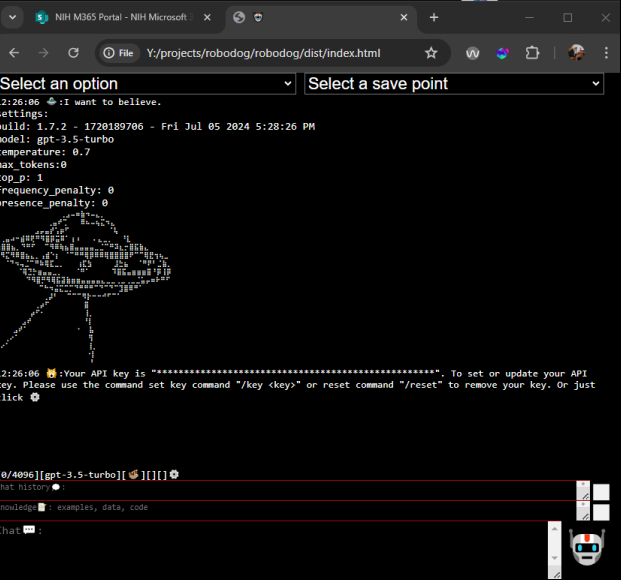
# Separation of Concerns



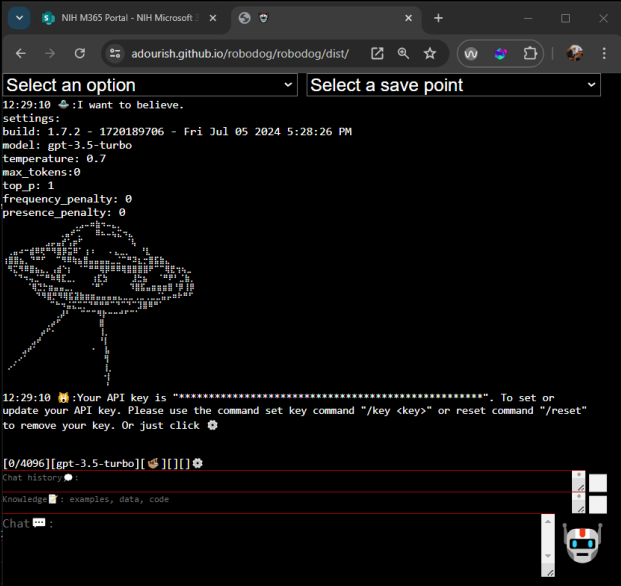
# Robodog Roadmap



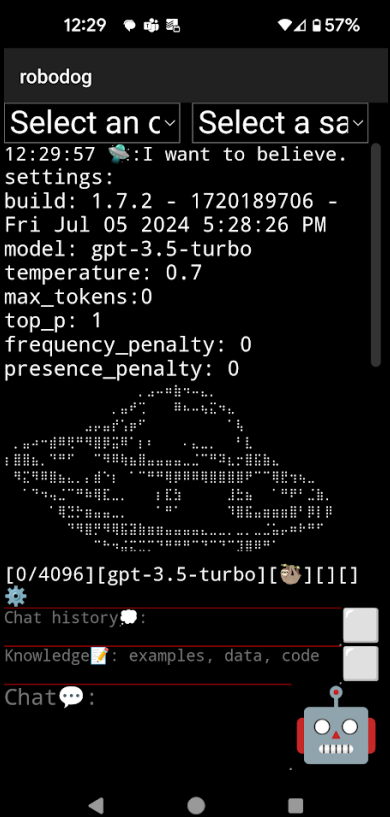
# Robodog Formats



HTML Bundle



Web



Android App

- [npm install robodoglib](#)
- [npm install robodogcli](#)
- [npm install robodog](#)

NPM Packages



# Appendix

# Theories of Self-Awareness

The **theory of mind** suggests that self-awareness stems from our ability to comprehend that others have different experiences and thoughts. *This theory could map to AGI only.*

The **protagonist theory** proposes that we see ourselves as the central character in our life's narrative, which leads to self-awareness. *This theory could map to AI and AGI.*

The **emergent self-theory** implies that self-awareness evolves from our interactions and relationships with others. *This theory could be associated with AGI.*

The mirror test theory posits that self-awareness can be measured by an entity's ability to recognize itself in a mirror. This theory could fit between AI and AGI.

