FINXTER OpenAI Glossary Cheat Sheet 👨‍💻 ✺ OpenAI

🤖 **Artificial General Intelligence (AGI)**: AGI refers to a hypothetical AI that can perform any intellectual task a human being can do, demonstrating human-like cognitive abilities across diverse domains.

🚀 **Singularity**: A theoretical point in the future when AI advancements lead to rapid, uncontrollable, and transformative changes in society, potentially surpassing human comprehension.

◐ **AI Safety**: AI safety is the study and practice of building AI systems that operate securely and align with human values, ensuring that they benefit humanity without causing harm.

◉ **Alignment Problem**: The alignment problem is the challenge of designing AI systems that understand and act upon human intentions, values, and goals, rather than optimizing for unintended objectives.

✺ **OpenAI**: OpenAI is an AI research organization that focuses on developing artificial general intelligence (AGI) that benefits everybody.

💡 **Deep Learning**: Deep learning is a subfield of machine learning that uses artificial neural networks to model complex patterns and make predictions or decisions based on input data.

✵ **Artificial Neural Network**: An artificial neural network is a computational model inspired by the human brain's structure and function, consisting of interconnected nodes called neurons that process and transmit information.

🕹 **Supervised Learning**: Supervised learning is a machine learning approach where a model is trained on a dataset containing input-output pairs, learning to predict outputs based on new inputs.

🌐 **Unsupervised Learning**: Unsupervised learning is a machine learning approach where a model learns patterns and structures within input data without explicit output labels, often through clustering or dimensionality reduction.

🎮 **Reinforcement Learning from Human Feedback (RLHF)**: RLHF is a method that combines reinforcement learning with human feedback, allowing AI models to learn from and adapt to human preferences and values.

💬 **Natural Language Processing (NLP):** NLP is a field of AI that focuses on enabling computers to understand, interpret, and generate human language.

📖 **Large Language Models**: Large language models are AI models trained on vast amounts of text data, capable of understanding and generating human-like text.

✺ **Transformer**: The Transformer is a deep learning architecture designed for sequence-to-sequence tasks, known for its self-attention mechanism that helps capture long-range dependencies in data.

◎ **Attention mechanism**: Attention mechanisms in neural networks enable models to weigh the importance of different input elements relative to one another, improving their ability to capture context.

🔲 **Self-attention**: Self-attention is a type of attention mechanism used in transformers that allows the model to relate different positions of a single sequence.

📖 **BERT (Bidirectional Encoder Representations from Transformers)**: BERT is a pre-trained transformer-based model developed by Google for natural language understanding tasks, which can be fine-tuned for specific applications.

🖋 **GPT (Generative Pre-trained Transformer)**: GPT is a series of AI models developed by OpenAI, designed for natural language processing tasks and capable of generating coherent, contextually relevant text.

🌐 **GPT-3.5**: GPT-3.5 is an intermediate version of the GPT series, bridging the gap between GPT-3 and GPT-4 in terms of model size and capabilities.

💬 **GPT-4**: GPT-4 is a more advanced version of the GPT series, expected to have larger model size and enhanced capabilities compared to its predecessors.

🐾 **Pre-training**: Pre-training is the initial phase of training a deep learning model on a large dataset, often unsupervised

🎚 **Fine-tuning**: Fine-tuning is the process of adapting a pre-trained model for a specific task by training it on labeled data related to that task, refining its performance.

◉ **Zero-shot learning**: Zero-shot learning is a machine learning approach where a model can make predictions or complete tasks without being explicitly trained on that task's data.

✒ **Few-shot learning:** Few-shot learning is a machine learning approach where a model can quickly adapt to new tasks by learning from a small number of labeled examples.

📑 **Token**: A token is a unit of text, such as a word or subword, that serves as input to a language model.

✎ **Tokenizer:** A tokenizer is a tool that breaks down text into individual tokens for processing by a language model.

🖼 **Context window**: The context window is the maximum number of tokens that a language model can process in a single pass, determining its ability to capture context in input data.

💡 **Prompts**: Prompts are input text given to a language model to generate a response or complete a specific task.

⚙ **Prompt Engineering**: Prompt engineering is the process of designing effective prompts to elicit desired responses from language models, improving their utility and reliability.

🤖 **ChatGPT**: ChatGPT is a conversational AI model developed by OpenAI based on the GPT architecture, designed to generate human-like responses in text-based conversations.

🗨 **InstructGPT**: InstructGPT is an AI model developed by OpenAI, designed to follow instructions given in prompts, enabling it to generate more task-specific and accurate responses.

🔧 **OpenAI API**: The OpenAI API is a service provided by OpenAI that allows developers to access and utilize their AI models, such as ChatGPT, for various applications.

🎨 **DALL-E**: DALL-E is an AI model developed by OpenAI that generates images from textual descriptions, combining natural language understanding with image generation capabilities.

🗨 **LaMDA**: LaMDA is Google's conversational AI model designed to engage in open-domain conversations, understanding and generating responses for a wide range of topics.

◉ **Midjourney**: AI program and service that generates images from natural language descriptions, called "prompts", similar to OpenAI's DALL-E and Stable Diffusion

◢ **Stable Diffusion**: A deep learning, text-to-image model released in 2022 and used to generate detailed images conditioned on text descriptions. Also used for inpainting, outpainting, and generating image-to-image translations guided by a text prompt.

📈 **Diffusion models**: Diffusion models are a class of models that represent the spread of information, influence, or other phenomena through a network.

🔲 **Backpropagation**: Backpropagation is a widely-used optimization algorithm in neural networks that minimizes the error between predicted outputs and true outputs by adjusting the model's weights.