

# System and Methods for Human-Centered AGI

Provisional Patent Application

Originally Drafted on March 7, 2023; Revised and Filed: May 24, 2023

Craig Andrew Kaplan, Inventor

## **ABSTRACT**

For it to be effective, the safest path to AGI must also be the fastest.

The preferred implementation of AGI is the fastest method for achieving AGI because it begins with a network of human problem-solving agents, who, by definition can perform any intellectual task as well or better than the average human. AI agents that have been trained and customized by individual humans (AAAs) are introduced to the network as AI problem-solving agents. The human and AI agents share a common problem-solving architecture that is: rigorous, scalable, transparent, auditable, safe, and powerful.

The architecture supports automatic learning and self improvement. It is compatible with LLMs which can be “plugged in” to the network and upgrade as more powerful LLM models become available. The AGI network begins with humans doing most of the problem-solving work – especially the most important aspects such as setting goals and the most difficult aspects such as representing the problem. Over time, AAAs do more and more of the actual work, more effectively and efficiently than humans could, while human attention is increasingly directed to issues of ethics, safety, and oversight.

Because ethics and safety checks are built into the architecture itself, as the speed of problem increases far beyond the capability of humans to “keep pace” the system remains aligned with human values and ethics. At any time, humans can see exactly how the system is making decisions, including all ethical information.

The AGI network is highly scalable and will become more powerful over time, yet the fundamental values and ethics of the system – which cannot be logically derived by any intelligence no matter how smart and fast – remain aligned with humans. Thus the invention solves the alignment problem in a democratic and scalable manner.

The fact that the AGI network can be implemented rapidly – far faster than estimates for when AGI will develop from other approaches – ensures a first mover advantage that allows this safest path to AGI to dominate other approaches thereby fulfilling the two key requirements for the invention, namely that it be not only the fastest path to AGI but also the safest.