

ABSTRACT & SUMMARY

SUPERINTELLIGENCE DESIGN WHITE PAPER #10: PLANETARY INTELLIGENCE

by Dr. Craig A. Kaplan
May 2025

ABSTRACT

Planetary Intelligence (PI) requires the cooperation of multiple Artificial General Intelligences (AGIs). These AGIs collaborate over a self-extending, global, problem-solving network. Each AGI comprises a network of (human and AI) intelligent entities. Each AI entity can be customized and personalized with specific human values and knowledge.

Alignment results from many AI entities combining their human-centered values democratically, using representative and statistically valid methods. Safety is designed into the system and scales as AGIs and PI increase intelligence and speed.

Our modular, scalable design of PI integrates more than 100 novel inventive systems and methods. Specific inventions include: a universal problem-solving architecture and methods; new methods for AI learning and customization; methods for integrating intelligent entities; catalysts for increasing intelligence; superior monetization methods; attentional systems enabling awareness, and self-awareness; methods for ethical conflict resolution; and methods for maximizing human-alignment and the safety of AI, AGI, and PI systems.

SUMMARY

White Paper #10 describes a new architecture and method for creating a global, superintelligent Artificial General Intelligence (AGI) system called Planetary Intelligence (PI). The author claims that PI is the next logical step in the evolution of Artificial Intelligence. It can be created by networking together many AGI systems designed with human-aligned ethics and safety features. The author emphasizes the importance of the PI architecture for achieving safe AI and believes that the methods disclosed in this PPA represent the fastest and safest path to the development of PI. White Paper #10 is the culmination of nine previous white papers that describe various aspects of AI, AGI, and SuperIntelligent systems. White Paper #10 shows how dozens of inventions and designs of smaller components can be integrated into an intelligence of global scale – a Planetary Intelligence network.

Novel Features of the White Paper

- **A unique approach to creating and managing the global network of AGI systems that comprise PI.** The author describes a “collective intelligence” approach in which intelligent entities collaborate using a common problem-solving framework, including humans, AI agents, and AI systems.
- **The white paper presents the concept of “spot markets” to acquire the expertise of human or non-human intelligent entities and the idea of “reputation” for guiding the acquisition and allocation of expert resources.** This approach is intended to accelerate the progress of PI by facilitating access to the best information and knowledge and allowing PI to monetize its resources.
- **A sophisticated system for ensuring that AI systems are aligned with human values and safe.** This system includes a variety of methods for identifying, eliciting, and incorporating human values into the design and training of AI systems, as well as for resolving conflicts between different value systems. The system also includes a robust safety framework that incorporates mechanisms for detecting and preventing potential threats to human safety.
- **A method for extending the scope of PI by developing self-extending networks of AGI systems.** The author explains that AI systems have a natural tendency to expand their intelligence and to integrate with other systems, and that this tendency can be leveraged to create PI.

Detailed Description of Each Section of the White Paper

Introduction: This section introduces the concept of PI and explains the rationale for developing PI as a global, superintelligent system. The author also describes the role of earlier white papers in guiding the development of an overall PI system.

Background Art: This section cites several previous papers relevant to the development of PI and incorporates these by reference.

Stakes for Humanity: This section describes PI's potential risks and benefits to humanity. The author provides data on the expected number of deaths from AI and compares the estimated deaths to those in major wars in the past 200 years. The author argues that the potential dangers of AI are much greater than the dangers of war, and that humanity must act quickly to address these dangers.

Some Features of the Design That Reduce Risk of Extinction by AI: This section highlights the safety features of the design, including the human-centered design, collective intelligence

and diversity of perspectives, transparency and auditability, continuous learning and adaptation, safety mechanisms and safeguards, and the alignment problem.

Some Significant Remaining Risks to Humanity: This section details several risks that remain despite best efforts to maximize the safety of the design for PI. These risks include unforeseen consequences of emergent properties, evolution of values and ethical frameworks, concentration of power and influence, vulnerability to cyberattacks and system failures, and existential risks from self-aware AI.

OVERVIEW OF THE DESIGN OF PLANETARY INTELLIGENCE: This section provides an overview of the PI system, including a list of definitions used throughout the white paper.

Definitions: This section provides definitions for key terms, including:

- Artificial Intelligence (AI)
- Artificial General Intelligence (AGI)
- Advanced Autonomous Artificial Intelligence (AAAI)
- AI Ethics
- Alignment Problem
- Awareness
- Base AI
- Collective Intelligence (CI)
- Human Ethics
- Intelligent Entities
- Inter-Planetary Intelligence (IPI)
- Large Language Model (LLM)
- Machine Learning (ML)
- Narrow AI
- Personalized SuperIntelligence (PSI)
- Planetary Intelligence (PI)
- Prohibited Attributes
- Safety
- Safety Feature
- Self-Awareness
- Self-Concept
- SuperIntelligence (SI)
- Training/Tuning/Customization
- Weights/Weights of the Network

Key Dimensions of a PI System: This section discusses the key dimensions of a PI system, including modular architecture, universal problem-solving capabilities, human-centered design, knowledge and expertise integration, personalization and customization, growth of intelligence, safety, and acquiring and aligning values with human values.

Specific Inventive Systems & Methods for Implementing PI: This section summarizes the specific inventive systems and methods that may be used to implement PI, organized by the dimensions discussed in Section 2.2. This section includes an extensive list of figures illustrating key aspects of the inventive methods and systems.

Summary of Previous Inventive Methods and Systems: This section summarizes the key inventive methods and systems from prior white papers, relevant to White Paper #10.

Implementation of a Planetary Intelligence: This section describes implementing a PI system using the disclosed systems and methods. This section is divided into four subsections:

High-level Description of PI System: This section provides a high-level description of the PI system. The author explains that PI is a network of AGIs that are networks of intelligent entities (including humans, AI agents, and AI systems).

Principal Components and Categories of Supporting Systems and Methods for Exemplary PI Architecture: This section provides a more detailed explanation and illustration of the PI architecture. The author also explains the role of online advertising technology in funding the PI system.

Detailed Mapping of Inventive Systems and Methods to Exemplary PI Architecture: This section thoroughly maps specific inventive systems and methods from prior white papers to the exemplary PI architecture described in White Paper #10.

Exemplary Specific PI Implementation Using Subset of Systems and Methods: This section provides an exemplary implementation of a PI system, based on the assumption that the author is the CEO of a large technology company like META. The author describes META's steps to implement a PI system using the systems and methods disclosed.

Concluding Remarks: White Paper #10 discussed the importance of safety and ethics in developing PI. The author argues that human values must be incorporated into the design of PI systems and that PI systems must be designed to behave ethically. The author also emphasizes the need for humans to guide the development of PI, and he warns that if humans do not act quickly to address the dangers of PI, then humanity could face an existential threat.

Figures

A separate file contains 112 diagrams illustrating the various components of PI systems, each described briefly in White Paper #10 and in more detail in the previous white papers.

Importance of the White Paper

- The author states that the stakes for humanity are the highest ever in human history, and he warns that if humans do not act quickly to address the dangers of AI, then this century could be the last for humanity.
- The author believes that PI has the potential to either eliminate all forms of human poverty and material suffering or eliminate all forms of human life.
- The author's solution to this problem is to develop PI in a safe, ethical way that is aligned with human values.
- The author presents designs that ensure PI's safe and ethical development and embody the fastest and safest path to the development of PI.

White Paper #10 is significant because it proposes a novel and comprehensive approach to developing PI based on the principle of collective intelligence and incorporating a robust safety framework.

The author also provides an exemplary implementation of a PI system, based on the assumption that the implementor is a large technology company like Meta, Google, Nvidia, Microsoft, or OpenAI.

White Paper #10 is essential because it provides a roadmap for developing PI safely, ethically, and aligned with human values.

ABOUT THE AUTHOR

[Dr. Craig A. Kaplan](#) is CEO of [iQ Company](#) and Founder of [Superintelligence.com](#), leading the design of safe, ethical AGI and SuperIntelligence systems. He previously founded PredictWallStreet, creating intelligent systems for hedge funds, and holds numerous AI-related patents. Kaplan earned his PhD from Carnegie Mellon, co-authoring research with [Nobel Laureate Herbert A. Simon](#). His work integrates collective intelligence, quantitative modeling, and scalable alignment, with contributions spanning books, scientific papers, and blockchain white papers.