

---

# Politicians on AI Safety<sup>1</sup>

---

Liam Elise  
Robins Racine  
George University  
Washingt of Oxford  
on  
University  
  
Manikanta Bhanu  
Revuri Reddy

With  
Johns Hopkins University & Apart Research

## Abstract

[Politicians on AI Safety \(PAIS\)](#) is a website that tracks U.S. political candidates' stances on AI safety, categorizing their statements into three risk areas: AI ethics / mundane risks, geopolitical risks, and existential risks. PAIS is non-partisan and does not promote any particular policy agenda. The goal of PAIS is to help voters understand candidates' positions on AI policy, thereby helping them cast informed votes and promoting transparency in AI-related policymaking. PAIS could also be helpful for AI researchers by providing an easily accessible record of politicians' statements and actions regarding AI risks.

*Keywords: AI policy tracker; AI ethics; geopolitical risks; existential risks*

## 1. Introduction

Artificial intelligence and technology policy is already one of the most important issues in Washington, and its importance is rapidly growing as technological

---

<sup>1</sup>AI Policy Hackathon at Johns Hopkins University 2024

progress accelerates. Voters, civil society groups, and other stakeholders need accurate information regarding what risks AI poses to their jobs, security, and way of life, as well as what actions their elected officials are (or are not) taking to mitigate those risks. Yet, it is difficult to keep track of candidates' stances on AI. Up until now, there has been no easy way for voters to compare different candidates' AI policies, and many candidates say little to nothing about AI. PAIS aims to solve this problem and bring transparency by creating an easily accessible database to track political candidates' quotes and votes regarding artificial intelligence and its risks.

PAIS is divided into three sections, based on the main categories of AI risk found in AI safety literature:

### **AI Ethics**

Sometimes called “mundane risks”, these are harms that AI currently poses to individuals or groups in society based on what values are built into the systems and how they are deployed. The word “mundane” is not meant to imply that these risks are unimportant, but merely that they involve different policy levers than the more extreme categories of risk.

Examples of ethical risks include robot workers causing unemployment, AI-generated misinformation, nonconsensual deep-fakes, and algorithmic discrimination.

### **Geopolitical risks**

These are risks that would cause America to lose power or security to geopolitical rivals. The most significant geopolitical risk involves fear that China will outpace the West in AI development, and thereby threaten the American-led international order. Other geopolitical risks involve the use of AI by terrorist groups, and the fear that AI will exacerbate a rise in worldwide authoritarianism.

### **Existential risks**

These are risks to the very existence of humanity, involving worst-case scenarios like AI-induced nuclear war or total human extinction. Many leading AI researchers believe that if we invent superintelligent machines before learning how to align those machines with human welfare, then the results could be catastrophic.

(These categories are not meant to be hard divisions, and some policies/risks will fall into multiple categories.)

Any time a candidate or elected official makes a statement about AI safety, it will be listed on their page based on what category of AI risk it belongs to. Any time a candidate or elected official releases a report about AI safety – like the Senate's Bipartisan AI Roadmap – it will be listed on their page. Any time a candidate or elected official issues a policy directive on AI safety – like Joe Biden's Executive Order on AI – it will be listed on their page. If a candidate has not articulated a clear position on AI safety, that will also be noted on their page.

PAIS focuses on AI safety and risks because we believe that AI risk management is an under-examined field. We do not mean to imply that AI's impacts will be entirely or mostly negative, nor that risk mitigation is the only relevant area of AI policy. Some policy decisions, such as utilizing AI to make government agencies more efficient, are important but beyond the scope of this website to document.

To be clear, PAIS is not partisan, and it does not promote any particular agenda or set of policies. It exists with the sole purpose of educating the public about AI safety policies, thereby allowing voters to decide for themselves which candidates are worth supporting.

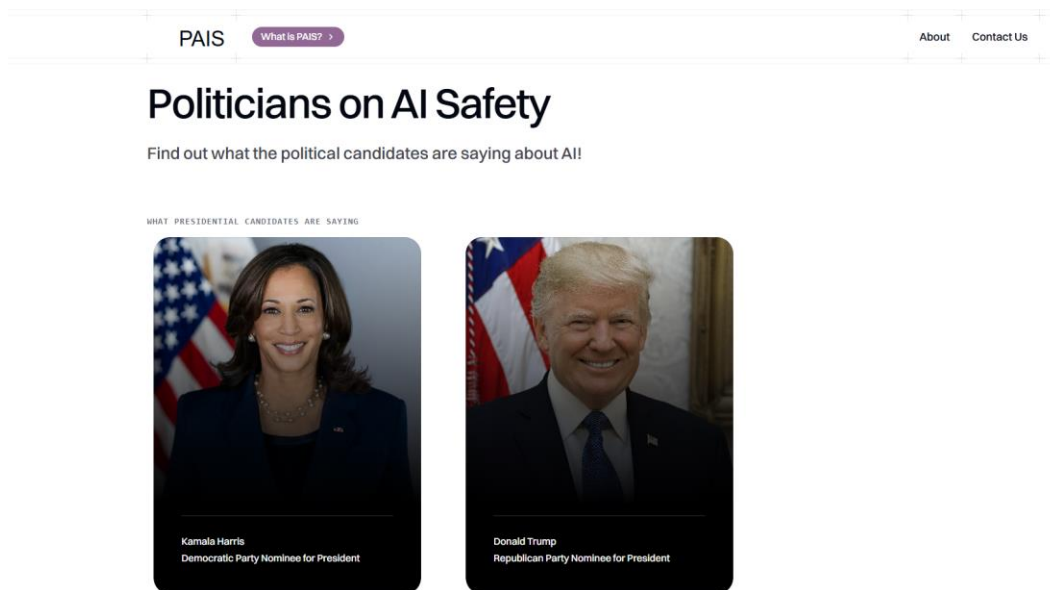
## 2. Current Status / Future Plans

PAIS is currently live. It has pages for each of the major parties' presidential candidates and running mates, as well as pages for "What is PAIS?", "About", and "Contact Us".

Over the next few weeks until Election Day, we intend to launch a publicity campaign on social media to spread the word about AI policy to voters. After Election Day, we hope to gather together a larger team and expand our candidate pages to include members of Congress.


## 3. Results

This is what PAIS currently looks like:



*Figure 1 – Front page*

PAIS [What is PAIS?](#) [About](#) [Contact Us](#)



**Donald Trump**  
REPUBLICAN PARTY NOMINEE FOR PRESIDENT  
 Businessman, television personality, and 45th president of the United States.

All categories ▾

---

Tuesday, September 24, 2024  
[Geopolitical Risks](#)

**Has pledged to “double our electric capacity” to compete for AI dominance with China**

Donald Trump has repeatedly stressed the need for America to increase its electric capacity given the vast amounts of energy required to power artificial intelligence systems and stay ahead of rivals like China. He has promoted environmental deregulation, and an expansion of both nuclear energy and fossil fuel drilling to increase American energy production.

[Read more >](#)

---

Thursday, June 13, 2024  
[Existential Risks](#)

**Acknowledges the seriousness of existential risks**


In a podcast with Logan Paul, Donald Trump described AI risks as “disconcerting”, “alarming”, and “scary”. In particular, Trump mentioned the possibilities of AI “[taking] over the human race” and accidentally triggering a nuclear war.

*Figure 2 – Example of a candidate page (Donald Trump)*

PAIS [What is PAIS?](#) [About](#) [Contact Us](#)

TUESDAY, SEPTEMBER 24, 2024

# Has pledged to “double our electric capacity” to compete for AI dominance with China



Donald Trump

[Geopolitical Risks](#)

Donald Trump has repeatedly stressed the need for America to increase its electric capacity given the vast amounts of energy required to power artificial intelligence systems. In August 2024, Trump stated “This is shocking to me, but AI requires twice the energy that the country already produces [...] You are gonna need tremendous amounts of electricity, like almost double what we produce now for the whole country, if you can believe it.”

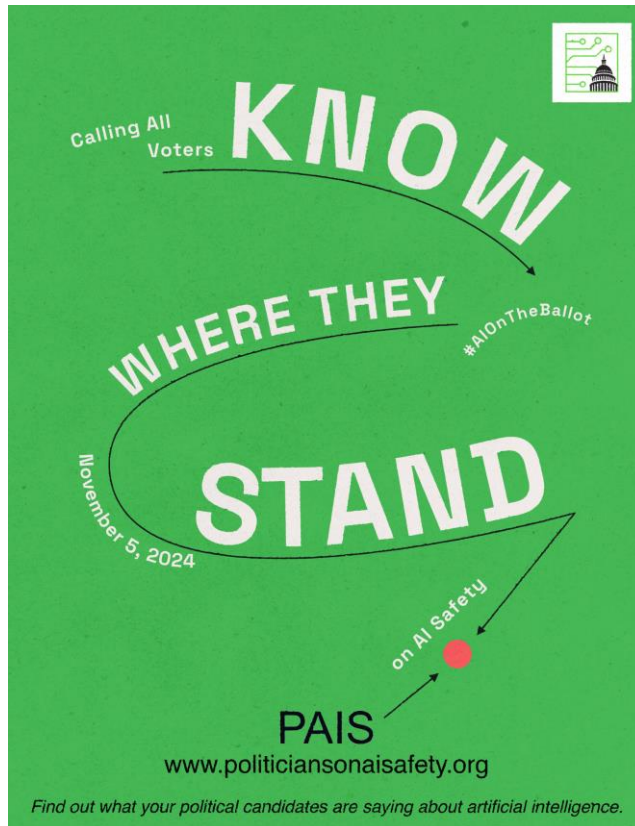
During a rally in September 2024, Trump stated, “We will be able to double our electric capacity, which will be needed to compete with China and other countries on artificial intelligence. [...] So we have to produce massive electricity that we don’t have, but the environmental impact statements won’t allow us to do that. The rules and regulations that we currently have won’t allow us to do any of it, not even a tiny fraction of it. But if I’m president, we’ll be able to do it.”

Trump has promoted environmental deregulation, and an expansion of both nuclear energy and fossil fuel drilling to increase American energy production.

Sources: [1](#), [2](#), [3](#)

[← Back to Donald Trump Policies](#)

*Figure 3 – Each candidate page links to policy pages.*



*Figure 4 – An example of the advertising that we are doing in advance of the election*

#### **4. Discussion and Conclusion**

On a technical level, PAIS is purposefully as simple as possible, since the goal is to translate complex AI risks into understandable terms, which is critical for bridging gaps between AI experts and policymakers. Real-world applications include facilitating voter education, aligning political agendas with AI safety priorities, and encouraging responsible AI governance. AI researchers and even companies will also be curious to know where elected officials stand on pressing AI policy debates.

Challenges include simplifying AI topics without losing critical nuances, a task managed by structuring issues into mundane, geopolitical, and existential risks. This approach enhances accessibility but limits the depth of discussion for each risk type, demanding continuous refinement as the field advances. It is difficult to interpret some candidates' statements, as many lack comprehensive AI strategies. Continually tracking and updating policy stances will require much vigilance. We must also be sure to not come across as partisan or biased, since PAIS strives to be as objective as possible. On a technical level, challenges include website malfunctions, ability to handle web traffic, and managing SEO.

From a market perspective, the need for informed AI governance creates strong growth potential for PAIS. Scalability is achievable with a larger research team,

which could be accomplished through partnerships with think tanks or AI research institutions.

As stated above, we believe that projects like this could inspire increased concern for AI safety and more thoughtfulness on the part of elected officials regarding what kinds of AI policy they seek to pursue.