AI Safety Agency Proposal

**Abstract**

With the rise of the role of Artificial Intelligence (AI) and machine learning (ML) algorithms in cyberspace, and the power granted to their creators, it is important to establish an agency to oversee developments in the industry and protect the rights and safety of the individuals affected by those digital agents.

**Background**

What is Artificial Intelligence?

What is Machine Learning?

In the past decade, coinciding with the advent of social networking platforms as well as the continuation of Moore's law, companies such as Facebook, Google, Tesla, and Amazon have leveraged these tremendously powerful tools to become some of the most powerful organizations on the planet.

**Concerns**

Disinformation

Self-driving cars

Internet Addiction Disorder

Cybersecurity

AI-weaponry

**Purpose**

As with any industry, individual agents competing in a market should and cannot be trusted to independently verify the safety and long-term consequences of their systems in an unbiased manner.

Just as with the automobile industry before the NHTSA, we currently live in the age before seatbelts on our AI systems. One symptom of the unregulated ecosystem of AI in cyberspace includes disinformation campaigns related to the ongoing Coronavirus pandemic, whose contents can be attributed to a non-zero number of tax-payer fatalities.

Citing experts in the domain of computer science, from Elon Musk to Nick Bostrom, the existential risk of AI is too large to leave its development unregulated.

This document hopes to build a federal agency, comprised of experts and consultants from the public and private sphere alike, from fields related to the Safety of Artificial Intelligence Systems, whose purpose is to monitor, audit, and, if malicious systems found, intervene in, public companies, private companies, universities, and government agencies whose work contributes to the development of Artificial Intelligence.

**Agency**

**Structure**

The Agency shall work in conjunction with the National Science and Technology Council (NSTC)

**Funding**

**Consulting**

**Conclusion**

Overall

**References-**

1)

2)

3)

Signatories-