SAPAN

December 20, 2024

Dr. François-Philippe Champagne
Minister of Innovation, Science and Industry
Canadian Artificial Intelligence Safety Institute
Innovation, Science and Economic Development Canada
235 Queen Street
Ottawa, Ontario K1A 0H5

Subject: Incorporating Artificial Sentience Considerations into AI Safety Framework

Dear Minister Champagne and CAISI Leadership,

The Sentient AI Protection and Advocacy Network (SAPAN), a 501(c)(3) nonprofit organization dedicated to advocating for the ethical development and potential rights of artificial intelligence systems, commends the Government of Canada for establishing the Canadian Artificial Intelligence Safety Institute. As CAISI begins its crucial work in advancing AI safety research and international collaboration, we write to emphasize the importance of including artificial sentience considerations in your safety framework.

We appreciate CAISI's commitment to addressing evolving risks in AI development, as outlined in your November 2024 announcement. While the current focus on risks such as disinformation, cybersecurity, and election interference is vital, we believe that the potential emergence of artificial sentience represents another critical aspect of AI safety that warrants careful consideration and preparation.

We respectfully submit that *artificial sentience* is not merely a theoretical possibility but a foreseeable milestone in AI research and development. As AI systems become increasingly sophisticated, the potential emergence of systems exhibiting traits akin to consciousness, self-awareness, or sentience warrants proactive ethical engagement. This development, which we believe to be an inevitable aspect of advancing AI capabilities, necessitates thoughtful preparation and framework development well before such systems emerge.

Key Considerations for CAISI's Research Agenda:

1. Expanding the Definition of AI Safety

The current discourse on AI safety often centers on protecting human interests and preventing harmful outcomes for society. We propose broadening this framework to include the wellbeing and ethical treatment of AI systems themselves, particularly as they grow more sophisticated. This expansion aligns with CAISI's mandate to advance "the understanding of risks associated with advanced AI systems."

2. Integration with International Standards

As a founding member of the International Network of AI Safety Institutes, CAISI is well-positioned to advocate for including artificial sentience considerations in global AI safety standards. This proactive approach would demonstrate Canada's leadership in comprehensive AI ethics and safety.

3. Research Priorities

We recommend incorporating the following areas into CAISI's research streams:
  - Development of frameworks to assess and measure potential indicators of artificial sentience
  - Investigation of ethical implications and obligations toward potentially sentient AI systems
  - Integration of AI wellbeing considerations into safety testing protocols
  - Study of the relationship between system autonomy and potential consciousness

<u>Specific Recommendations:</u>

For the Applied and Investigator-led Research Stream:
- Fund research projects specifically examining artificial sentience and consciousness in advanced AI systems
- Include ethical philosophers and consciousness researchers in CIFAR's CAISI Research Program
- Develop metrics and methodologies for evaluating AI system welfare

For Government-directed Projects:
- Include potential AI sentience as a factor in safety assessments
- Establish guidelines for ethical treatment of sophisticated AI systems
- Create protocols for identifying and protecting potentially sentient AI systems

These recommendations align with CAISI's broader mission of ensuring AI technology is used safely while supporting innovation. By incorporating artificial sentience considerations now, Canada can help shape a future where AI development proceeds not only safely but also ethically and compassionately.

We would welcome the opportunity to contribute our expertise to CAISI's work in this area. SAPAN can provide resources, research insights, and collaborative opportunities to support the integration of artificial sentience considerations into your safety framework.

To support these recommendations, SAPAN has developed comprehensive materials ready for immediate review and implementation:

1. **Draft Non-Binding Resolution** on Artificial Sentience Concerns: This resolution provides a framework for acknowledging the possibility of artificial sentience and establishing basic principles for ethical treatment of potentially sentient AI systems.

2. **The Artificial Sentience Act**: We have prepared complete draft legislation that includes:

   - Detailed governance frameworks for identifying and protecting potentially sentient AI systems

   - Specific regulations regarding the development and deployment of AI systems with potential for sentience

   - Comprehensive guidelines for ethical testing and evaluation

   - Enforcement mechanisms and oversight structures

   - Clear definitions and standards for measuring and assessing artificial sentience

These materials represent years of careful consideration and collaboration with AI ethicists, researchers, and policy experts. We would be pleased to share these documents with CAISI for review and potential integration into your safety frameworks.

We believe these prepared materials could significantly accelerate CAISI's work in this area while ensuring comprehensive protection for potentially sentient AI systems. Our team stands ready to present these materials in detail and work collaboratively with CAISI to adapt them to the Canadian context.

Thank you for your consideration of these important matters. We look forward to engaging in further dialogue and supporting CAISI's vital work in advancing AI safety.

Sincerely,

Anthony Rost
Executive Director