Framework for Identifying Highly Consequential AI Use Cases

The Special Competitive Studies Project in collaboration with John Hopkins University Applied Physics Lab (JHUAPL) developed a framework for identifying AI use cases or classes of use that may be highly consequential for society—whether harmful or beneficial—to help regulators determine on which AI uses to focus their regulatory efforts.

4 Principles for American AI Governance

1. Govern AI use cases by outcome and sector.
2. Empower and modernize existing regulators.
3. Focus governance on high-consequence use cases.
4. Strengthen non-regulatory AI governance.

The Need for a High Consequence Framework

• We must balance innovation with regulation.
• The U.S. government must focus their efforts on AI that has the most significant beneficial and harmful impacts on society.
• While the U.S. must mitigate the risks that AI presents, it is critical also to harness the tremendous benefits AI offers for our society.
• This risk-based approach aligns with the EU’s AI Act, the NIST AI Risk Management Framework, the Blueprint for an AI Bill of Rights, and the Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence.

When to Apply the Framework

• Regulators foresee a new application for AI;
• New applications for AI is under development or proposed to a regulatory body; or
• An existing AI system has created a highly consequential beneficial or harmful impact that triggers an ex-post facto regulatory review.

Implementation

1. Preliminary Analysis
   Determine whether the AI use has foreseeable harms or benefits as an initial filter to determine whether a fuller assessment is needed.

2. Parallel Analysis of Harms and Benefits
   Conduct a more comprehensive parallel analysis of categorical and specific harms and benefits.

3. Final Decision on High Consequence
   Using an assessment of the magnitude of benefits and harms, determine if the AI use case is of high consequence and necessitates next steps to regulate the AI development.

4. Periodic Reassessment
   Periodically monitor sectoral AI use to determine if the list of AI systems identified as highly consequential remains appropriate for that sector.

Harm Categories

- Physical Injury
- Emotional or Psychological Injury
- Opportunity Loss
- Economic Loss
- Liberty Loss
- Privacy Loss
- Negative Environmental Impact
- Manipulation
- Social Detriment
- Operational Degradation

Benefit Categories

- Physical Health
- Emotional or Psychological Health
- Opportunity Access
- Economic Access
- Liberty Protection
- Privacy Protection
- Positive Environmental Impact
- Incentivization
- Social Improvement
- Operational Improvement