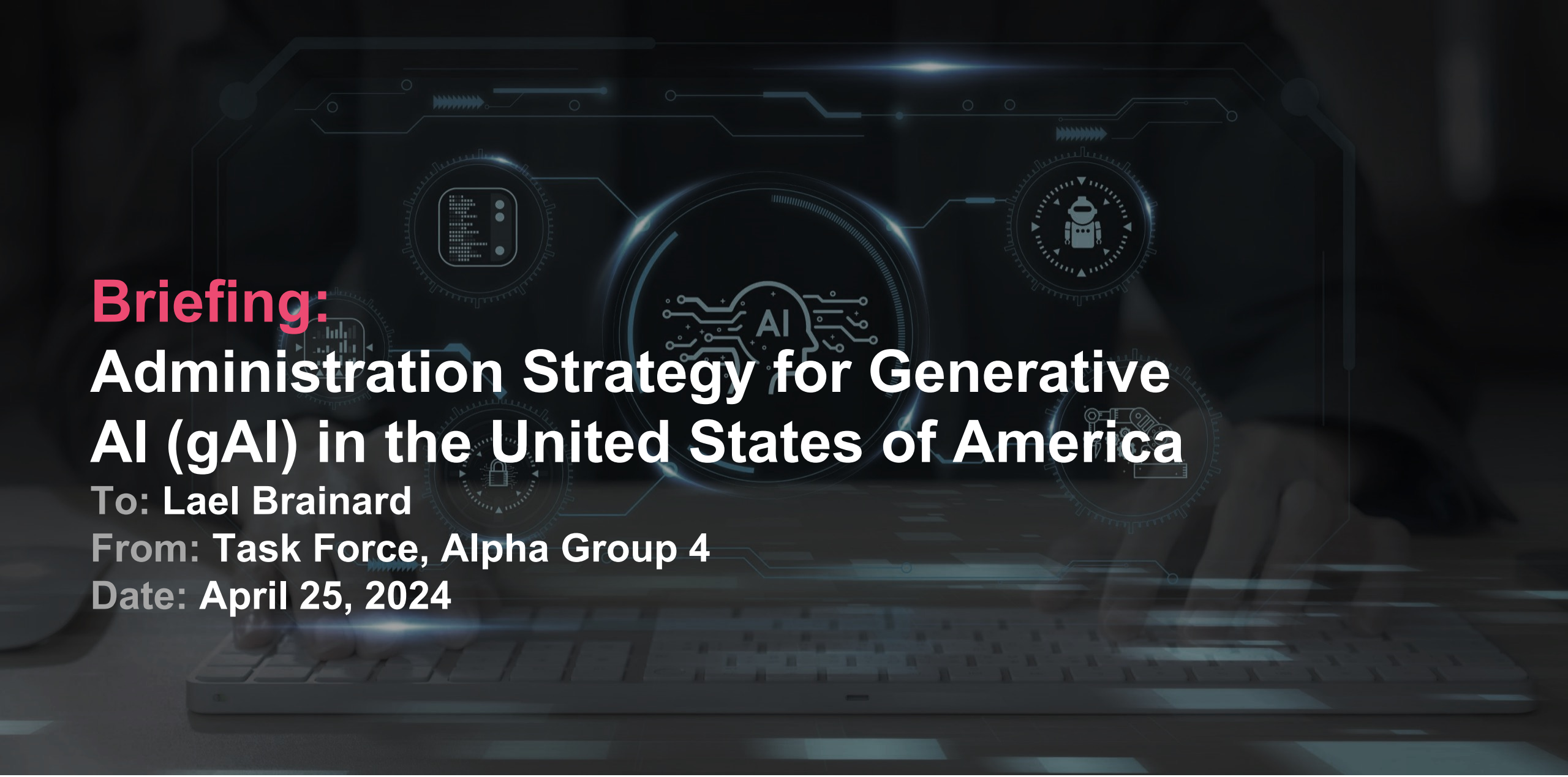



# Briefing:
# Administration Strategy for Generative AI (gAI) in the United States of America

**To:** Lael Brainard

**From:** Task Force, Alpha Group 4

**Date:** April 25, 2024

HKS

# Agenda and key takeaway

**High-level recommendation:**

Risk-based sectoral approach to gAI regulation, managed by a new agency, the U.S. Office of Artificial Intelligence (USOAI).

**Topics for today:**

- Stakeholder insights
- Policy options for consideration
- Evaluation criteria
- Proposed policy
- Key action items
- Budget
- Risks & mitigation


# Issue statement

- Generative artificial intelligence (gAI) is poised to **radically reshape all sectors of society**
- No **clear strategy in place** to leverage the benefits
- Many risks
  - Geopolitical and economic competition
  - Job displacement
  - Privacy loss
  - Discrimination
  - Disinformation
  - Human extinction?
- Need to both **encourage innovation and mitigate risks**

# Stakeholder insights | Several key groups of varying influence are deeply invested in gAI policy decisions

| Group | Detail | Interest | Power |
|---|---|---|---|
| **Biden White House** | • Under pressure to not leave regulation to other countries<br>• Wishes to be seen as proactive and effective, especially in election year | High | High |
| **Private companies** | • Anchor of global AI innovation<br>• Likely looking for regulation like U.S. Code 230, Communications Decency Act (1996): promoted free market in Internet / social media<br>• Likely to form interest group for lobbying; risk that larger players propose anti-competitive regulation | High | Medium |
| **US federal government agencies & departments** | • Competing views on how regulated AI should be in their respective areas<br>• *E.g., DoD accelerating use of AI, DHS concerned about deepfakes* | High | Medium |
| **International allies** | • Wish to reap benefits of U.S. AI innovation but fear negative implications of under-regulation<br>• Some have already taken regulatory action | High | Low |
| **American public** | • Generally low understanding of implications of AI for everyday life<br>• Concerned about job losses and harms of Generative AI | Medium | Low |

# Policy options | Five major US gAI policy options were considered to promote innovation and mitigate risks

| # | Policy | Details |
|---|---|---|
| 1 | **Libertarian** | • Allow market forces to shape emerging gAI industry and avoid over-regulating to preserve US dominance<br>• Leave companies to largely self-regulate, and develop policy reactively |
| 2 | **Risk-based sectoral** | • Develop sector-specific regulations based on risk level of gAI applications in field<br>• Create AI-specific regulatory agency to determine risk levels, set guidelines<br>• Promote AI innovation through light industrial policy |
| 3 | **Protectionist libertarian** | • Limit regulation within US borders and defer to market forces<br>• Heavily regulate AI exports to box out foreign actors |
| 4 | **Nationalization** | • Government exerts control over AI development in high-risk sectors through contracting with private companies<br>• Technology released for civilian applications at discretion of government |
| 5 | **Participatory democracy** | • Treat gAI like a public good and decide how resources are used democratically<br>• Profits returned as dividends or reinvested into programs to benefit the public |

# Criteria | Policy options were scored against six criteria






| Geopolitical competitiveness | Political feasibility | Existential risk | Cost & financial feasibility | Bias, inclusion, & fairness | Enforceability |
|---|---|---|---|---|---|
| Risk posed to US' dominant position in international AI market, especially relative to major competitors | Potential of receiving bipartisan support in Congress, uptake by federal agencies and support within voter base | Risk of unchecked development with destructive consequences | Cost of implementation borne by private and public sectors | Ability to promote nondiscrimination and fairness | Ease of monitoring and enforcement |
| **25%** | **25%** | **20%** | **10%** | **10%** | **10%** |

# Proposed policy | Recommend adopting a risk-based sectoral approach

| Rationale | Precedent | Key components |
| --- | --- | --- |
| gAI is not a single product but rather a family of technologies with a diverse set of applications<br><br>More nuance in approach achieves balance between promoting innovation and managing risk | The US already applies risk-based regulation to<br>• **Chemicals**: regulated by the EPA<br>• **Pharmaceuticals**: regulated by the FDA<br>The US provides loan guarantees for **clean energy** projects and **trains** workers for digital jobs | • Establish **new AI regulatory body**: US Office of Artificial Intelligence (USOAI)<br>  • Define risk levels by sector based on potential threat to US (e.g., defense is high-risk)<br>  • Develop regulations specific to risk levels<br>  • Modify risk level of specific products as needed<br>• **Subsidize costs** for AI startup companies<br>• **Invest in educational programs** for gAI literacy |

HKS

# Proposed policy | Sectors will be split across three risk levels and subjected to risk-specific policies

| Risk level | Low | Medium | High |
|---|---|---|---|
| **Example policies** | Companies encouraged but not required to publicly disclose data for gAI products<br><br>Must secure and protect personal data<br><br>Products must adhere to principles within Blueprint for AI Bill of Rights | Companies using gAI products must disclose risks specific to products and safety testing in SEC filings<br><br>Products generated using gAI must be tagged as such | New products, major upgrades must be submitted to USOAI for approval prior to launch<br><br>Must gain USOAI approval to export to other countries<br><br>Training data and generated data made available to USOAI<br><br>Systems must be accessible to humans during use |
| **Example sectors** | Entertainment | Financial Services<br>Media<br>Education | Healthcare<br>Defense |

# Key action items:
# Immediate and long-term goals

## Immediate term

*Within 6-12 months*

- Institution Building
- Risk Framework
- Risk Mitigation Strategy
- Value Alignment and Ethics Framework

## Long-term

*12+ months*

- Research and Development Efforts
- AI Literacy Programs
- International and Cross-sector Collaboration
- Government Incentives for Innovation

# Delivery plan | Annual Proposed Budget

| Expenses | Items | Estimate |
|---|---|---|
| **USOAI** | Staff | $500 million |
| | Technology and Infrastructure | $5 million |
| | Research & Framework Development | $2 million |
| | Travel and Auditing | $1 million |
| | Employee Training | $1 million |
| | Facilities | $2 million |
| | Operations | $5 million |
| **Innovation** | Loans and Tax Rebates | $200 million |
| **Education** | Public School Curriculum | $500 million |
| | Job Retraining | $1 billion |
| **Total Cost** | **$1.77 billion** | |

HKS

**Note:** For comparison, the Office of the Comptroller of the Currency, a banking regulatory agency within the Treasury Department with 3,509 employees, had a budget of **$1.2 billion** in FY 2023. The Office of Regulatory Affairs (ORA), which leads all regulatory field activities within the FDA, had 5,121 employees and a budget of **$1.4 billion** in FY 2023. The Loan Program Office at the Department of Energy made a budget request for **$180 million** in FY 2023.

# Risks & mitigation | Administration will need to control for high-priority risks

| Risk | Impact | Mitigation strategy |
|---|---|---|
| Risk level classification becomes outdated as tech develops | Over-classification as high-risk stifles innovation or high-risk products overlooked | Case-by-case risk assessments, upskill regulators, and constantly update the classification method |
| Partisanship leads to no Congressional action / consensus | Policy risks unenforceable and potentially temporary | Start early, mobilize bi-partisan support |
| Companies choose to leave the U.S. due to compliance costs (esp. high-risk sectors / products) | Lose global competitiveness | Proactively reach out to companies to assist them with compliance |
| Small and medium-sized enterprises struggle to compete with Big Tech | Innovation stifling | Level the playing field with government incentives |

# Appendix

**Further detail across:**

1. Current state of play
2. Legislative starting point
3. Design of the USOAI
4. Risk assessment
5. Features of policy options considered
6. Option scoring table
7. Objectives and key results
8. Delivery plan detail

*References*

# State of play | United States leads in the emerging and fast-growing gAI market

- During 2023, US-based institutions developed 61 AI models, compared to 21 in the EU, 15 in China
- In late 2023, President Biden signed an Executive Order to manage AI risks in US
- Lawmakers working to keep up with pace of AI development; legislation reactionary to-date
- EU and South Korea have adopted broad and specific AI regulations
- Estimates of economic impact of AI vary; $500B - $1T annual value for US in the next decade


HKS *Brendan Smialowski | AFP | Getty Images*

# Starting point | Policy builds on existing regulatory infrastructure in United States

October 2023 White House's Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence

**New Standards for AI Safety and Security**

- Require that developers of the most powerful AI systems share their safety test results and other critical information with the U.S. government.
- Develop standards, tools, and tests to help ensure that AI systems are safe, secure, trustworthy
- Protect against the risks of using AI to engineer dangerous biological materials
- Protect Americans from AI-enabled fraud and deception by establishing standards and best practices for detecting AI-generated content and authenticating official content.
- Establish an advanced cybersecurity program to develop AI tools to find and fix vulnerabilities in critical software
- Order development of a National Security Memorandum that directs further actions on AI / security

# Design | US Office of Artificial Intelligence (USOAI)

- **Tasks:**
  - Drafting of specific regulation
  - Evaluating compliance
  - Enforcing rules and taking corrective action
  - Fostering the development and use of trustworthy AI
  - Enhancing international cooperation to align standards
- **Bring AI community / expertise together**: government, companies, universities, NGOs to help translate technological breakthroughs into widespread benefits, while mitigating risks
- **Recruitment:** Staff from both the private and public sectors, including those from companies and government departments. They need to possess expertise in government regulation while simultaneously being aware of the impact for and the developments in the private sector

# Risk assessment | To standardize ratings, USOAI experts will develop methodology for risk-rating sectors

*Note: potential for overrides with justification.*

| Criteria | Example |
|---|---|
| Is gAI likely to be widely used in this sector to deliver critical services or make significant decisions in place of humans? | **Education**: The widespread use of gAI programs to develop personalized lessons for students as a complement or substitute for human teachers |
| Is the use of gAI likely to lead to large disruptions in the sector's workforce? | **Media**: The use of gAI at news organizations could have a large impact on the number of journalists and copywriters required |
| Will gAI models in this sector collect and store critical personal data from users? | **Healthcare**: gAI models used to create personalized healthcare plans are likely to use highly sensitive data regarding patients' medical history |
| Do gAI models in this sector have a high potential for producing biased information? | **Criminal Justice**: Using gAI models to generate training simulations for police officers may leverage data from prior encounters shaped by bias towards racial minorities and perpetuate this bias |
| Could gAI models developed in this sector be easily applied for malevolent purposes? | **Biotech**: gAI programs designed to help scientists find ways to synthesize new molecules for medicines could be repurposed by terrorist organizations to invent new bioweapons |

**Policy options | Eleven categories of features conceptualized across each option**

1. Privacy
2. Workforce
3. Environmental impact
4. Scope
5. Transparency
6. Institutional / Governance
7. Human oversight
8. Accountability
9. Innovation
10. Disinformation & Fairness
11. Technical robustness

HKS

# Option Scoring System

| Criteria | Weight | Libertarian | Risk-based / sectoral | Nationalization | Participatory democracy | Protectionist libertarian |
|---|---|---|---|---|---|---|
| Geopolitical competitiveness | 25% | 3 | 2 | 2 | 3 | 3 |
| Political feasibility | 25% | 3 | 3 | 2 | 1 | 3 |
| Existential risk | 20% | 1 | 3 | 3 | 2 | 1 |
| Costs/ Financial feasibility | 10% | 3 | 2 | 1 | 1 | 2 |
| Bias, inclusion, fairness | 10% | 1 | 3 | 2 | 3 | 1 |
| Enforceability | 10% | 3 | 2 | 2 | 2 | 1 |
| **Weighted total score** | 10% | 2.40 | **2.55** | 2.10 | 2.00 | 2.10 |
| | | | *Proposed option* | | | |

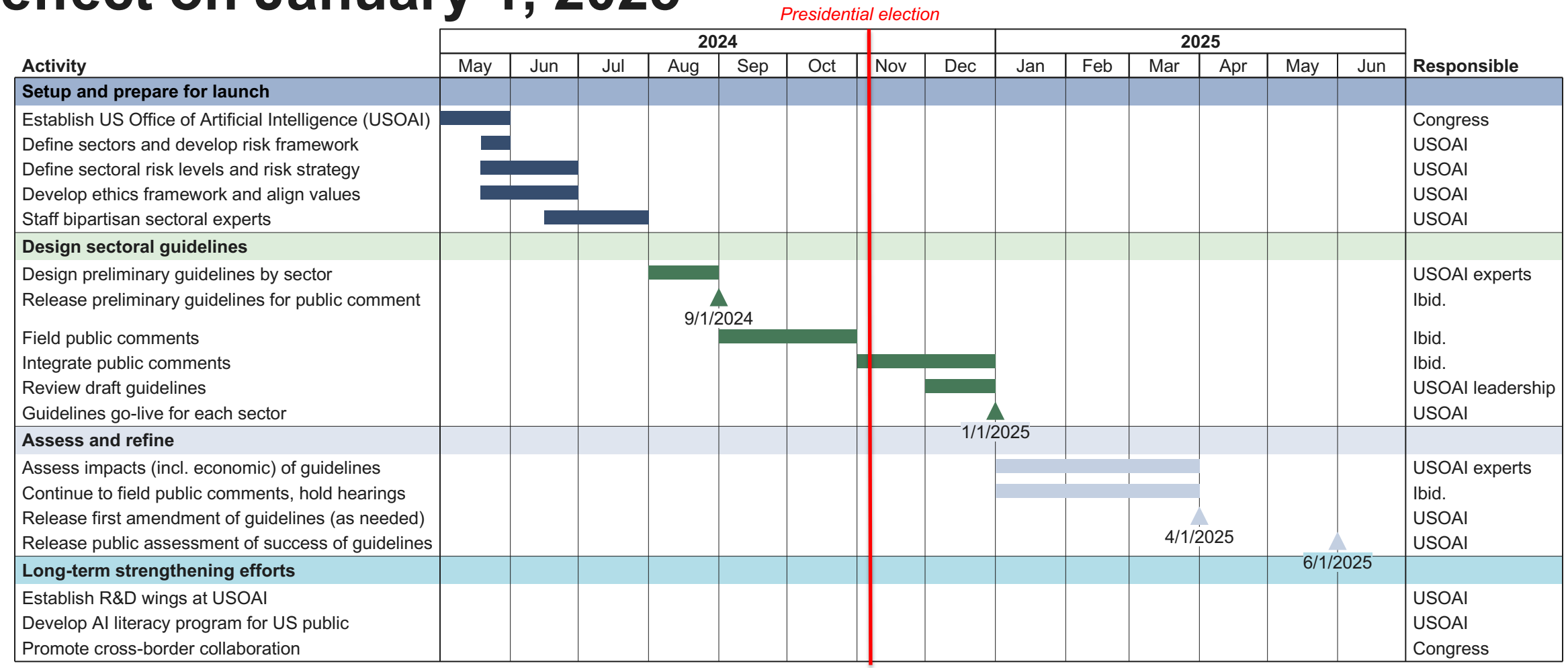
*1 = Low (e.g., low transparency, limited accountability, major privacy concerns)*
*2 = Medium*
*3 = High (most appealing)*

HKS

# Objectives and Key Results

| Objectives | Key Results |
|---|---|
| **Objective 1:** Establish the United States Office of Artificial Intelligence (USOAI) | **KR1:** Publish the organizational structure and the necessary staff profiles by May<br>**KR2:** Fill 50% of all staff positions by May and 100% by July<br>**KR3:** Finalize comprehensive set of mandates and operational objectives |
| **Objective 2:** Determine the risk levels associated with AI and develop a framework | **KR1:** Define a framework that classifies risks into high, medium, and low by June<br>**KR2:** Conduct a stakeholder consultation and collaborate on the draft frameworks with at least 10 AI industry leaders, 10 academic institutions and all necessary government departments by December<br>**KR3:** Pilot and finalize guidelines by March |
| **Objective 3:** Strengthen AI research and development | **KR1:** Establish 3 pilot R&D programs focusing on innovation until December<br>**KR2:** Refine programs and start 5 new programs and initiatives until summer 2025<br>**KR3:** Develop guidelines on AI Values and Ethics, Transparency, and Accountability by end of September |
| **Objective 4:** Align international standards for AI | **KR1:** Establish an international wing in the USOAI by July and designate “AI Ambassadors”<br>**KR2:** Host one international conference on AI after the guidelines have been adopted and published<br>**KR3:** Negotiate one data share and standards agreement until 2030 |
| **Objective 5:** Develop AI literacy and engagement programs | **KR1:** Design and issue AI curricula in public schooling by December<br>**KR2:** Provide capacity building programs in a phased manner for the next year across government departments and agencies<br>**KR3:** Establish a public forum and feedback system for engaging citizens directly. |

# Delivery plan | Sector-specific guidelines will go into effect on January 1, 2025




| Activity | Responsible |
|---|---|
| **Setup and prepare for launch** | |
| Establish US Office of Artificial Intelligence (USOAI) | Congress |
| Define sectors and develop risk framework | USOAI |
| Define sectoral risk levels and risk strategy | USOAI |
| Develop ethics framework and align values | USOAI |
| Staff bipartisan sectoral experts | USOAI |
| **Design sectoral guidelines** | |
| Design preliminary guidelines by sector | USOAI experts |
| Release preliminary guidelines for public comment | Ibid. |
| Field public comments | Ibid. |
| Integrate public comments | Ibid. |
| Review draft guidelines | USOAI leadership |
| Guidelines go-live for each sector | USOAI |
| **Assess and refine** | |
| Assess impacts (incl. economic) of guidelines | USOAI experts |
| Continue to field public comments, hold hearings | Ibid. |
| Release first amendment of guidelines (as needed) | USOAI |
| Release public assessment of success of guidelines | USOAI |
| **Long-term strengthening efforts** | |
| Establish R&D wings at USOAI | USOAI |
| Develop AI literacy program for US public | USOAI |
| Promote cross-border collaboration | Congress |

HKS

# References


‘47 U.S. Code § 230 - Protection for Private Blocking and Screening of Offensive Material’, LII / Legal Information Institute, accessed 16 April 2024, https://www.law.cornell.edu/uscode/text/47/230.

‘AI in Elections: Where Republican Candidates Stand on Tech’, accessed 17 April 2024, https://www.newsnationnow.com/politics/debates/republican-candidates-artificial-intelligence-stances/.

"Commissions, Staffing, and Financing of AI Office Raises Eyebrows in Capitals," Euronews, February 1, 2024, https://www.euronews.com/next/2024/02/01/commissions-staffing-and-financing-of-ai-office-raises-eyebrows-in-capitals.

Dan Milmo and Dan Milmo Global technology editor, ‘AI Firms Must Be Held Responsible for Harm They Cause, “Godfathers” of Technology Say’, The Guardian, 24 October 2023, sec. Technology, https://www.theguardian.com/technology/2023/oct/24/ai-firms-must-be-held-responsible-for-harm-they-cause-godfathers-of-technology-say.

‘Department of Homeland Security Unveils Artificial Intelligence Roadmap, Announces Pilot Projects to Maximize Benefits of Technology, Advance Homeland Security Mission | Homeland Security’, accessed 16 April 2024, https://www.dhs.gov/news/2024/03/18/department-homeland-security-unveils-artificial-intelligence-roadmap-announces.

Faverio and Tyson. "What the Data Says About Americans' Views of Artificial Intelligence." Pew Research Center. November 21, 2023. https://www.pewresearch.org/short-reads/2023/11/21/what-the-data-says-about-americans-views-of-artificial-intelligence/.

Fred Pearce, "How Artificial Intelligence Could Help Scale Up Low-Carbon Energy and Cut Emissions," Yale Environment 360, March 3, 2020, https://e360.yale.edu/features/artificial-intelligence-climate-energy-emissions.

“Generative AI Could Raise Global GDP by 7%,” Goldman Sachs, April 5, 2024. https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html.

# References


Michelle Faverio and Alec Tyson, ‘What the Data Says about Americans’ Views of Artificial Intelligence’, Pew Research Center (blog), accessed 16 April 2024, https://www.pewresearch.org/short-reads/2023/11/21/what-the-data-says-about-americans-views-of-artificial-intelligence/.

Office of the Comptroller of the Currency, "OCC FY 2024 Congressional Justification," PDF file, accessed April 19, 2024, https://home.treasury.gov/system/files/266/24.-OCC-FY-2024-BIB.pdf"https://home.treasury.gov/system/files/266/24.-OCC-FY-2024-BIB.pdf.

PricewaterhouseCoopers, ‘PwC’s Global Artificial Intelligence Study: Sizing the Prize’, PwC, accessed 21 April 2024, https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html.

‘Section 230: An Overview’ (Congressional Research Service, 4 January 2024), https://crsreports.congress.gov/product/pdf/R/R46751#:~:text=Section%20230%20of%20the%20Communications,users%20of%20interactive%20computer%20services.

‘The U.S. Plans to “Lead the Way” on Global AI Policy’, LAWFARE, accessed 16 April 2024, https://www.lawfaremedia.org/article/the-u.s.-plans-to-lead-the-way-on-global-ai-policy.

U.S. Food and Drug Administration, "Justification of Estimates for Appropriations Committees," PDF file, accessed April 19, 2024, https://www.fda.gov/media/176925/download.

‘White House Wades into Debate on “Open” versus “Closed” Artificial Intelligence Systems | AP News’, accessed 16 April 2024, https://apnews.com/article/ai-executive-order-biden-opensource-models-1c42092e55729d731d246440094f7fed.