The Accountability, Responsibility & Governance as a Unified Strategy for AI (ARGUS-AI)

John Kurre

September 2024

CONTENTS

AIRWORKX

INTRODUCTION

AIRWORKX

The **Accountability, Responsibility, and Governance as a Unified Strategy for AI (ARGUS-AI)** research project assessed National AI Policies and Strategies to measure the current state of the art of standards for Responsible and Ethical Use of Artificial Intelligence (AI) within the U.S. National Security and Intelligence Community (IC) environments and their implications for National and Global Security.

01	**Examined complexities and concerns (e.g., algorithmic bias, data privacy, ethics) in integrating AI into U.S. National Security and Intelligence Environments.**
02	**Identified best practices for AI deployment, mitigating risks, and ensuring responsible AI usage in U.S. National Security.**
03	**Examined AI strategies and policies of critical agencies, including the U.S. Department of Homeland Security (DHS), the Office of the Director of National Intelligence (ODNI), and the Department of Defense (DOD).**
04	**Performed comparative and gap analyses of U.S. National AI Strategies against OECD AI Principles to evaluate policy coverage and effectiveness in AI governance.**
05	**Highlighted the significance of the OECD AI governance framework in ensuring ethical integration and application of AI in Strategic Security domains.**

PROBLEM STATEMENT

AIRWORKX

Integrating artificial intelligence (AI) into U.S. and Global Strategic Security domains presents unparalleled prospects alongside formidable challenges. AI can transform the approach to combatting cybercrime, foreign influence propaganda, proliferation of weapons of mass destruction, and other Global Security threats. The imperative for developing and deploying **Responsible**, **Ethical**, and **Secure** AI systems across National Security environments is even more paramount due to the gaps in AI policymaking and AI Governance standards. This research underscores the necessity of AI Governance to confront the challenges experienced by Government agencies on both National and Global scales due to algorithmic bias. This bias not only imperils trust but also obstructs AI's Responsible and Ethical utilization in National Security applications.

AI Integration in Intelligence Community

Integrating AI into the Intelligence Community (IC) opens new avenues for redefining intelligence workflows, decision-making, and threat assessment.

The transformative potential of AI is accompanied by various obstacles that require a nuanced and comprehensive approach to AI governance.

Critical Issues with AI-Enabled Workflows

- **Algorithmic Bias**: Prejudiced outcomes.
- **Explainability Gap**: Decision rationale undermines confidence
- **Ethical Dilemmas**: Concerns about privacy, civil liberties, etc.,
- **Intelligence Data Fragmentation**: Disparate intelligence data
- **Human-AI Collaboration**: Understanding operational mechanics and establish trust in AI.

Key Challenges faced by U.S. Intelligence Community

Absence of comprehensive AI governance structures

Reconciling theoretical constructs with practical demands

Aligning AI governance with unique needs of the IC

Limited understanding of AI and its implications

PURPOSE STATEMENT

AIRWORKX

This dissertation research explored the integration of Artificial Intelligence (AI) into U.S. National Security and Intelligence domains. The study assessed responsible AI development and deployment by analyzing current U.S. AI policies in alignment with the U.S. National AI Strategy and highly coveted OECD AI Principles (OECD, 2024). Addressing crucial challenges such as algorithmic bias required a detailed AI Governance Strategy customized for National and International Strategic Security community researchers, practitioners, and policymakers. The research provides standards, mitigation strategies, and recommendations to guide the **Safe**, **Secure**, **Trustworthy**, **Responsible**, and **Ethical** development and deployment of AI within U.S. National Security environments. In doing so, it courteously extends its sphere of guidance and best practices to shape Global standards on AI Governance.

01 Enhanced AI governance policies and mechanisms to address deficiencies in the U.S. National Security domain.

02 Examined challenges and potential solutions in AI governance through real-world case studies.

03 Investigated the delicate balance between National Security and the Responsible, Ethical, and Trustworthy Use of AI.

04 Compared OECD AI Governance theoretical framework with real-world Applications and Use Cases.

05 Analyzed the historical context and future developments of AI Governance mechanisms in U.S. National Security and Intelligence Community domains.

06 Identified opportunities to mitigate Algorithmic Bias to ensure fairness and integrity in AI-driven threat assessments.

07 Detected the gaps in AI Policymaking and their alignment with AI Governance Standards.

08 Leveraged international best practices via the OECD AI Principles.

09 Identified opportunities for enhancing the posture of National AI Strategies

10 Contributes to the growing body of knowledge and informs practical applications in the Strategic Security domain through this qualitative study.

RESEARCH QUESTION

AIRWORKX

What AI Governance approaches can be implemented to mitigate algorithmic bias and foster Trustworthy AI adoption in National Security and Intelligence environments?

Transparency and Explainability

Clear insights: AI systems should be understandable and have clear decision-making processes (OECD, 2024).

Accountability & Responsibility

Clear ownership: Clear lines of responsibility and accountability should be established for AI systems and their impact (OECD, 2024).

Sustainability

Planet-conscious tech: AI should be developed and used in a way that supports environmental sustainability (OECD, 2024).

Fairness & Non-Discrimination

Equal treatment for all: AI should treat everyone fairly, avoiding bias and discrimination in its actions (OECD, 2024).

Human-Centered Values

Humans at the heart: AI development should prioritize human well-being, safety, and ethical considerations (OECD, 2024).

Democracy & Rule of Law

Respecting our values: AI use should align with democratic values, human rights, and existing legal frameworks (OECD, 2024).

Robustness, Safety & Security

Built to last and built safely: AI systems should be resilient, reliable, and secure against errors and attacks (OECD, 2024).

Innovation & Economic Growth

Progress engine: AI should be developed and used to foster innovation, economic growth, and societal benefit (OECD, 2024).

International Cooperation

Global collaboration: International collaboration is crucial for responsible development and governance of AI (OECD, 2024).

Respect for Human Rights & Human Well-Being

Human dignity first: AI must always uphold and promote human rights and contribute to overall human well-being (OECD, 2024).

RESEARCH METHODOLOGY

AIRWORKX

The **Accountability, Responsibility, and Governance as a Unified Strategy for AI (ARGUS-AI)** research project investigated the Responsible and Ethical Use of Artificial Intelligence (AI) within the U.S. National Security and Intelligence Community (IC) and its implications for National and Global Security.

- Comparative Analysis of AI Governance Principles across National and International Governing Bodies
- Qualitative Content Analysis of AI Strategies and Policies of U.S. Government Agencies (e.g., DOD, DHS, ODNI)
- Evaluating Responsible and Ethical Use of AI across National Security and Intelligence Community Use Cases

DEFINE

Delineated the research problem and project parameters, discerning the unique challenges and prospects of AI Governance within the National Security and Intelligence Community.

MEASURE

Gathered empirical evidence on AI governance principles and practices, conducted a comparative analysis of global standards and examined practical AI implementations through case studies.

ANALYZE

Uncovered the underlying causes of identified problems and opportunities, utilizing content analysis of key policy documents, such as national AI strategies and the OECD's "Tools for Trustworthy AI" framework.

IMPROVE

Developed and implemented innovative solutions to enhance AI governance, tailoring frameworks to the specific requirements of the National Security and Intelligence Community.

CONTROL

Established a sustainable plan to maintain and refine improvements over time, ensuring the ongoing effectiveness of AI governance practices.

AIRWORKX

OECD AI POLICY ASSESSMENT TOOL

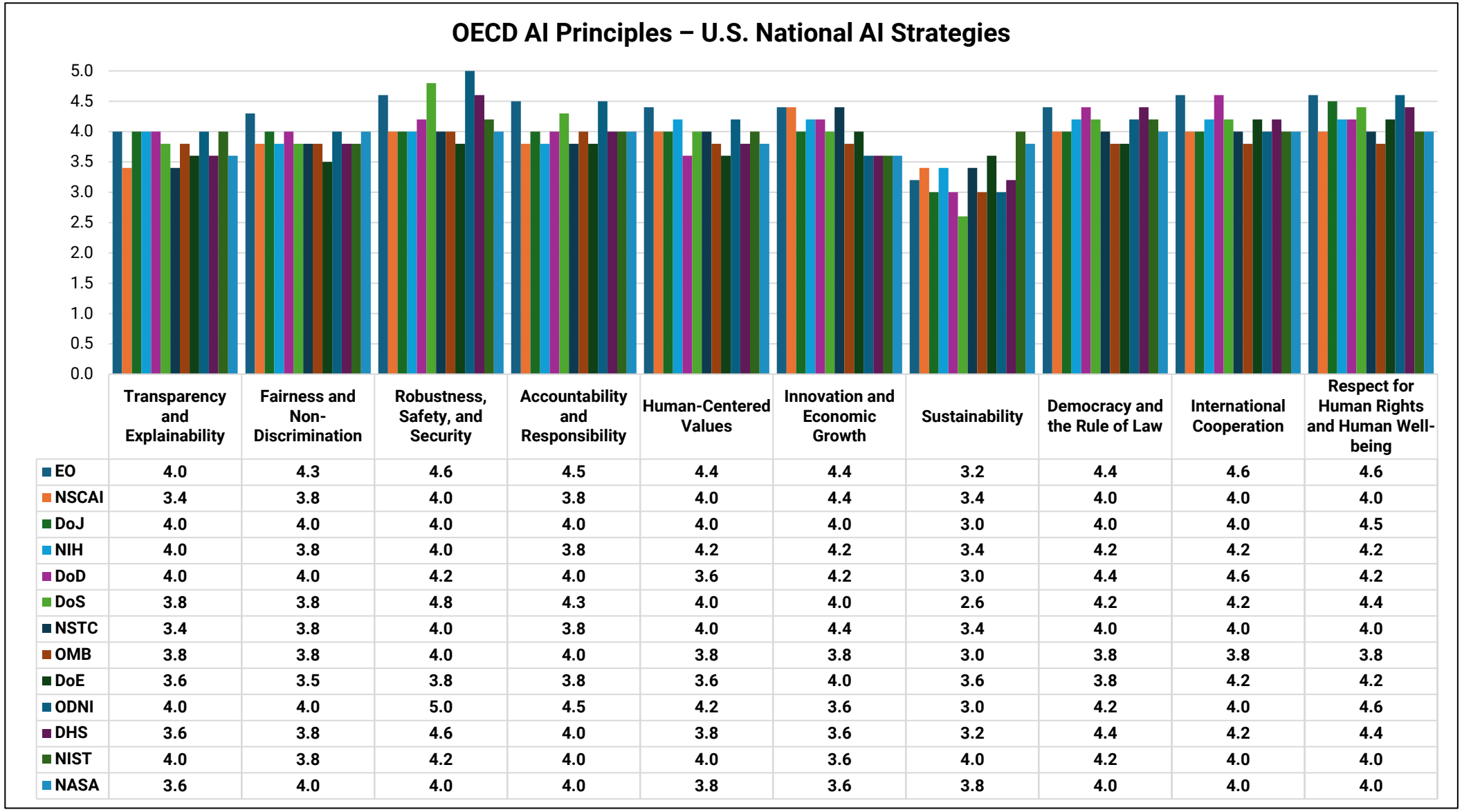
https://oecdai.com/

Stevens' (1946) seminal work on the ***Theory of Scales of Measurement*** provides a valuable framework for assessing complex phenomena, including aligning AI systems with OECD AI Principles. Applying his concept of an Ordinal Scale, we have used a 1-5 scale (1 = Lowest, 5 = Highest) to evaluate AI Policies in adherence to OECD AI Guidelines.

Examples: *Television Ratings, Movie Ratings, Educational Testing & National Ranking, Military Rank Structure & Ordinality, etc.,*

ASSESSMENT: OECD AI Principles - U.S. National AI Strategies

AIRWORKX



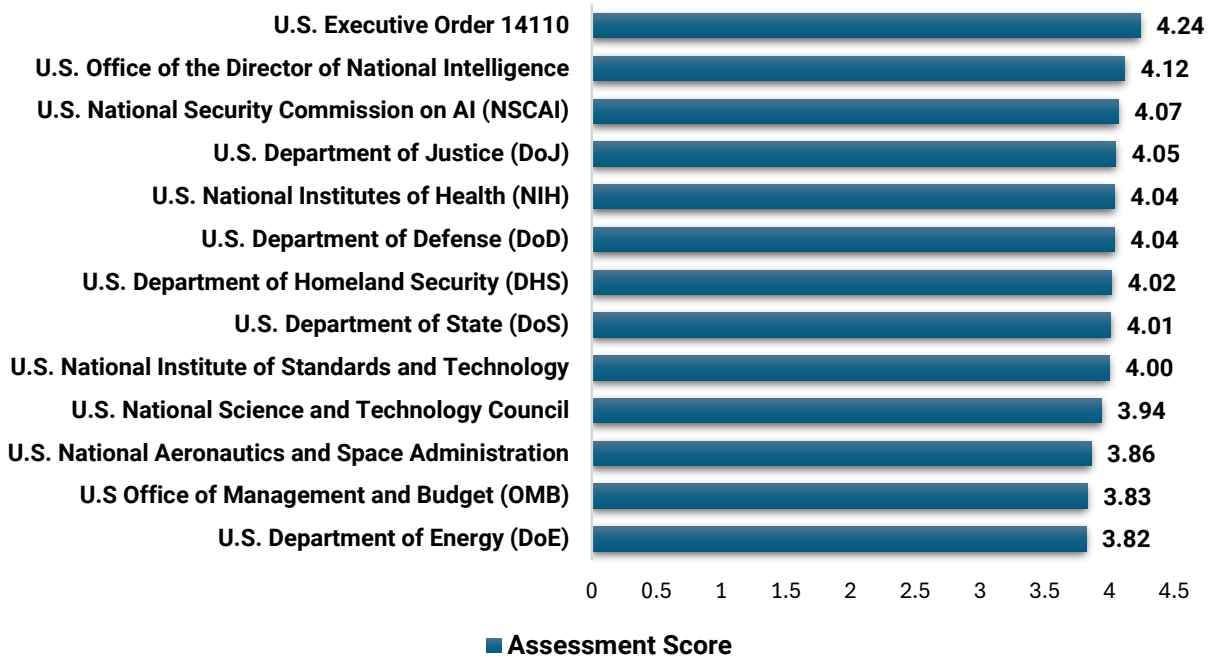
	Transparency and Explainability	Fairness and Non-Discrimination	Robustness, Safety, and Security	Accountability and Responsibility	Human-Centered Values	Innovation and Economic Growth	Sustainability	Democracy and the Rule of Law	International Cooperation	Respect for Human Rights and Human Well-being
EO	4.0	4.3	4.6	4.5	4.4	4.4	3.2	4.4	4.6	4.6
NSCAI	3.4	3.8	4.0	3.8	4.0	4.4	3.4	4.0	4.0	4.0
DoJ	4.0	4.0	4.0	4.0	4.0	4.0	3.0	4.0	4.0	4.5
NIH	4.0	3.8	4.0	3.8	4.2	4.2	3.4	4.2	4.2	4.2
DoD	4.0	4.0	4.2	4.0	3.6	4.2	3.0	4.4	4.6	4.2
DoS	3.8	3.8	4.8	4.3	4.0	4.0	2.6	4.2	4.2	4.4
NSTC	3.4	3.8	4.0	3.8	4.0	4.4	3.4	4.0	4.0	4.0
OMB	3.8	3.8	4.0	4.0	3.8	3.8	3.0	3.8	3.8	3.8
DoE	3.6	3.5	3.8	3.8	3.6	4.0	3.6	3.8	4.2	4.2
ODNI	4.0	4.0	5.0	4.5	4.2	3.6	3.0	4.2	4.0	4.6
DHS	3.6	3.8	4.6	4.0	3.8	3.6	3.2	4.4	4.2	4.4
NIST	4.0	3.8	4.2	4.0	4.0	3.6	4.0	4.2	4.0	4.0
NASA	3.6	4.0	4.0	4.0	3.8	3.6	3.8	4.0	4.0	4.0

Ordinal Scale of Measurement

An **Ordinal Scale** (1 = Lowest Alignment, 5 = Highest Alignment) facilitated a nuanced assessment of each document's adherence to the OECD AI Principles (Stevens, 1946). While primarily qualitative, this approach provided a structured method for analyzing data, enabling comparisons and identifying patterns and improvement opportunities (Norman & Streiner, 2008).

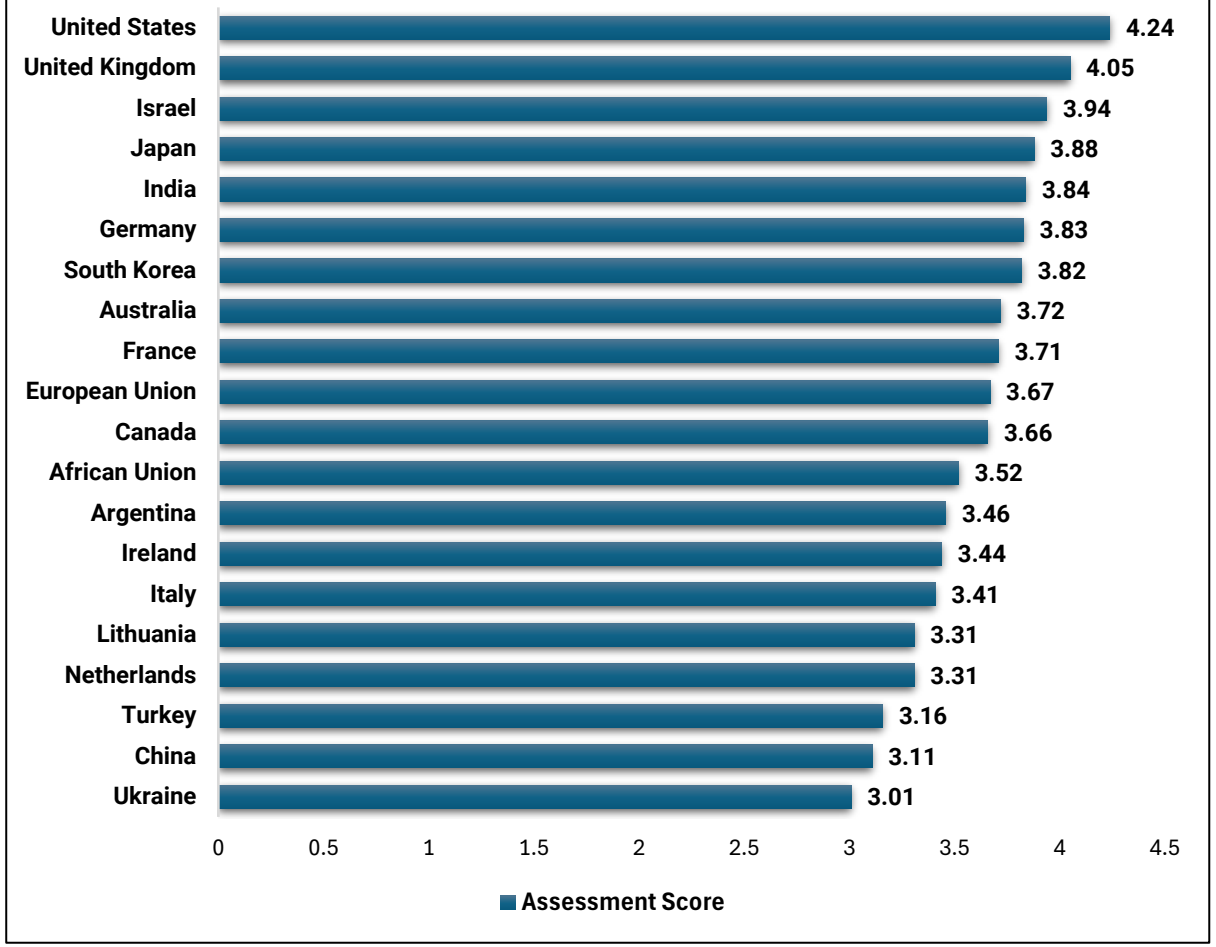
COMPARISON: National AI Strategies – U.S. & Global

U.S. Agencies - National AI Strategies Comparison



An **Ordinal Scale** (1 = Lowest Alignment, 5 = Highest Alignment) facilitated a nuanced assessment of each document's adherence to the OECD AI Principles (Stevens, 1946). While primarily qualitative, this approach provided a structured method for analyzing data, enabling comparisons and identifying patterns and improvement opportunities (Norman & Streiner, 2008).

Top 20 Countries - National AI Strategies Comparision



RESEARCH SIGNIFICANCE

AIRWORKX

Practitioners

- Embrace Documentation and Transparency
- Prioritize Alignment with OECD AI Principles
- Foster Proactive Community Engagement
- Embed Continuous Improvement Mechanisms

Policymakers

- Establish a Unified National AI Strategy
- Prioritize Human-Centered Values as Foundational Elements
- Mandate Robust Oversight and Accessible Redress Mechanisms
- Invest in Sustainability and Foster International Collaboration

Researchers

- Conducting Cross-National Comparative Analyses
- Investigating Sector-Specific Challenges and Opportunities
- Exploring Alternative Research Methodologies

References

Kurre (2024). *OECD AI Policy Assessment Tool*. https://oecdai.com/

OECD. (2019). *OECD principles on artificial intelligence*. Organisation for Economic Co-operation and Development (OECD) https://www.oecd.org/going-digital/ai/principles/

OECD.AI. (2021). *Database of national AI policies*. Organisation for Economic Co-operation and Development (OECD). https://oecd.ai

AIRWORKX

THANK YOU