

Image generated with help of DALL E-3

**Title:** AI in 2100: A Governance Blueprint for BRICS in the Coming AI Age

**Place of study:** Indian Institute of Technology(IIT), Roorkee

**INDEX**

[**Executive Summary**](#_7j98v6uzxzkn) **2**

[**Introduction:**](#_dgw0n9p2hla8) **3**

[Why focus on BRICS?](#_gnwbh3qku9qv) 3

[Why focus on Artificial Intelligence?](#_1wkbbmdjpwhh)

**Forecast**5

[**Current Work in BRICS on AI**](#_rbfzptcoofou) **7**

[**Proposed Standards:**](#_thstdebygmin) **9**

[**Policy Recommendations:**](#_5hip6p1c0uoq) **12**

[**Limitations:**](#_1kfnjmhittq7) **14**

[**Conclusion:**](#_36homixuehba) **16**

[**References:**](#_v3ep6qy6d2um) **16**

## 

## **Executive Summary**

With the advent of artificial intelligence systems such as GPT-4 in the past few months, developing common implementable standards which promote safe AI while also using such tools for economic growth has become relevant. With AI's capabilities evolving at an unprecedented rate, there is a unique opportunity for BRICS nations to harness these technologies to address region-specific challenges and integrate AI solutions into the startup-government ecosystem. This approach can significantly contribute to solving pressing socio-economic problems unique to these nations. However, the potential risks associated with AI, particularly concerning data privacy and potential misuse, necessitate a comprehensive and cooperative approach to policy formulation.

This paper outlines a series of policy recommendations and implementable standards aimed at ensuring the safe, ethical, and inclusive development and deployment of AI technologies within BRICS countries. These recommendations encompass a range of critical areas, including data protection, algorithmic transparency, AI in education, and the establishment of a collaborative regulatory framework. By adopting these standards, BRICS nations can position themselves at the forefront of AI innovation, ensuring that AI development is aligned with the collective interests and values of their societies.

## **Introduction:**

## 

### Why focus on BRICS?

BRICS is an association of five major countries including Brazil, Russia, India, China, and South Africa. These emerging economies comprise 30% of the world’s GDP as per IMF nad nearly 46% of the world’s population as per the United Nations.

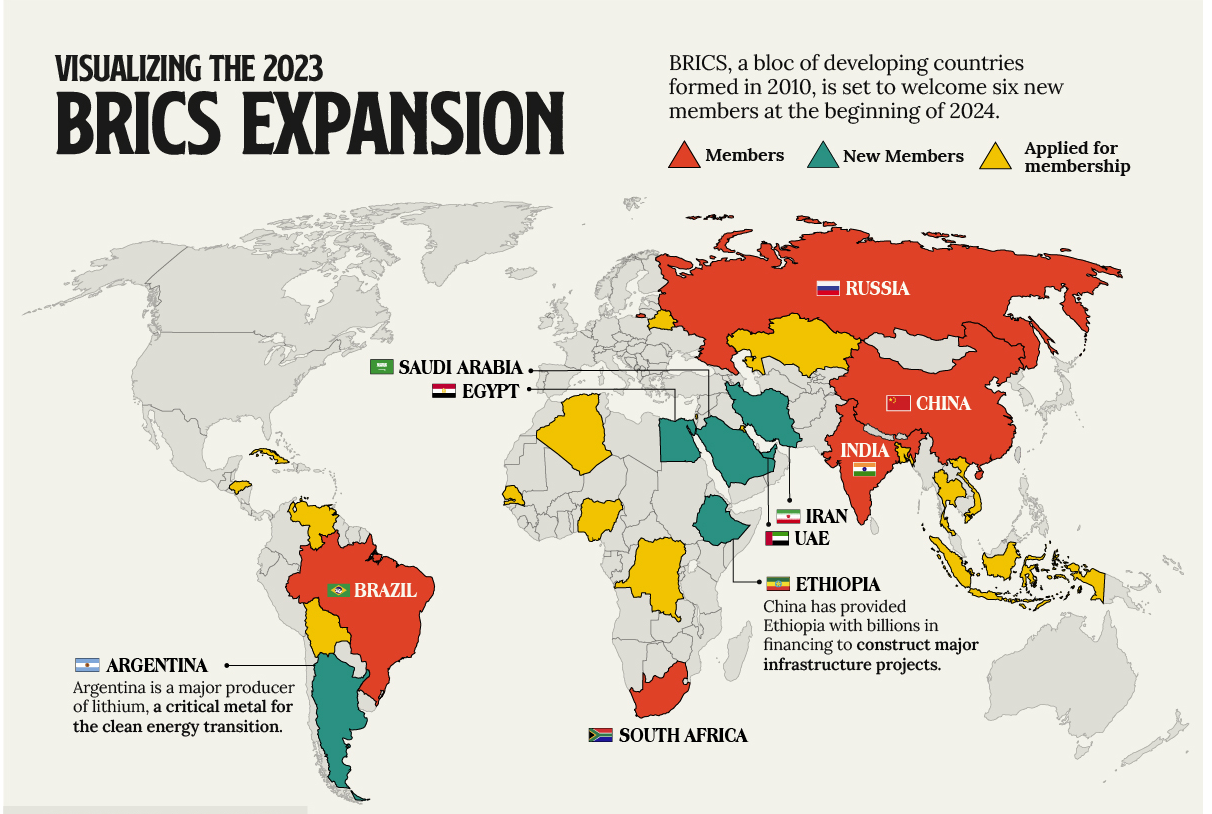


Image taken from Visualcapitalist.com

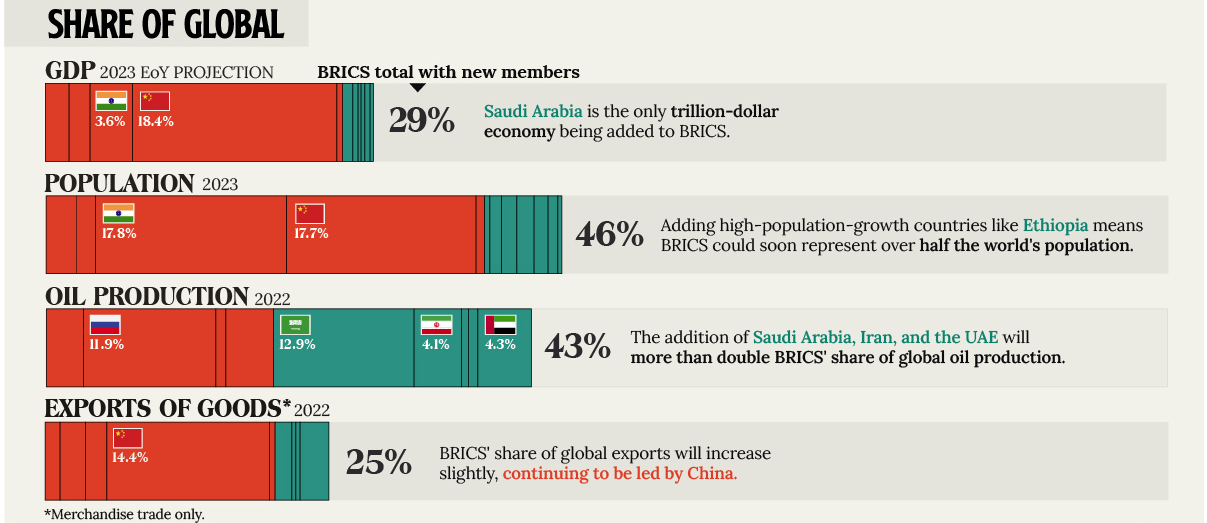
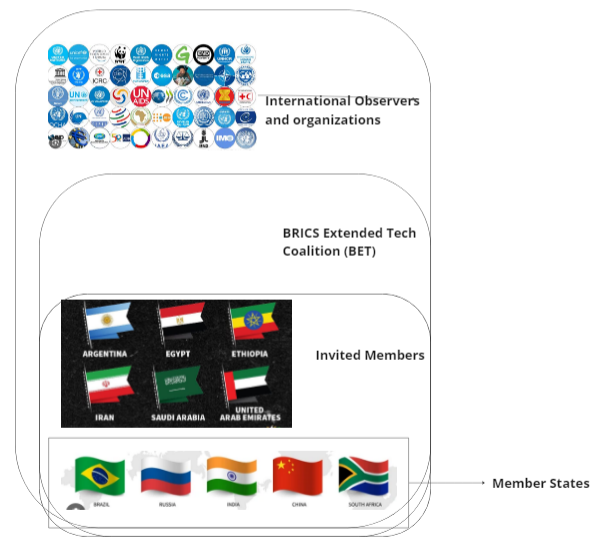


Image taken from Visualcapitalist.com

Moreover, the BRICS countries have a young population, are spread out over the world and are soon going to have new members.

Undoubtedly, these countries are major players in the geopolitical landscape on a regional and international level with a lot of potential. In light of this, other countries are bound to pay attention to what this grouping does as it has far-reaching consequences.



Proposed Structure of BRICS through Recommendations

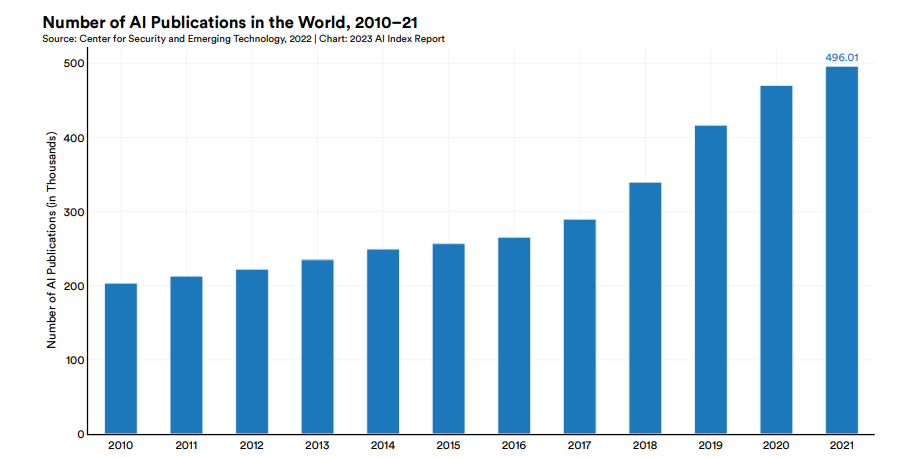
### **Why focus on Artificial Intelligence?**

At the recent BRICS summit, Chinese Premier Xi Jinping said, “We need to jointly fend off risks and develop AI governance frameworks and standards with broad-based consensus so as to make AI technologies more secure, reliable, controllable and equitable.” China “has set a national goal of investing US$150 billion in AI and becoming the global leader in this area by 2030. Countries like the United States, UK and the EU are also moving to regulate and focus on using AI for their economies. Recently, we have seen that the US president signed an executive order on AI as a matter of national importance and the EU negotiating its landmark EU AI Act.

AI is progressing at breakneck speed and is forecasted by many to reach human-level intelligence soon.

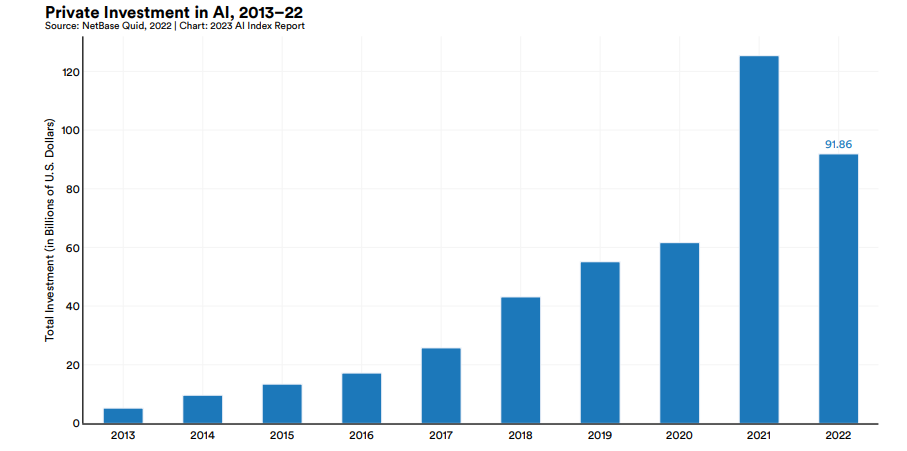
As tools using advances in natural language processing work their way into businesses and society, they could drive a 7% (or almost $7 trillion) increase in global GDP and lift productivity growth by 1.5 percentage points over a 10-year period as per Goldman Sachs.

An AI Index analysis of the legislative records of 127 countries shows that the number of bills containing “artificial intelligence” that were passed into law grew from just 1 in 2016 to 37 in 2022. An analysis of the parliamentary records on AI in 81 countries likewise shows that mentions of AI in global legislative proceedings have increased nearly 6.5 times since 2016. (Stanford HAI report,2023)



2023 AI Index Report by Stanford HAI -Number of AI publications

It is also notable that investment in AI is 18 times higher now in 2013 which seems to suggest that the expected economic returns are very high.

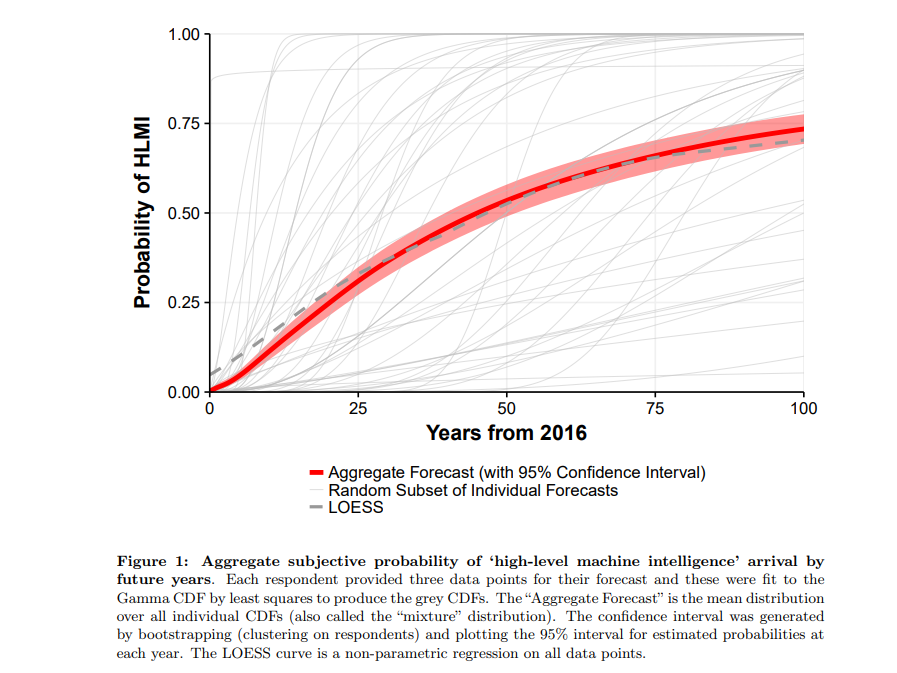


2023 AI Index Report by Stanford HAI

## Forecast

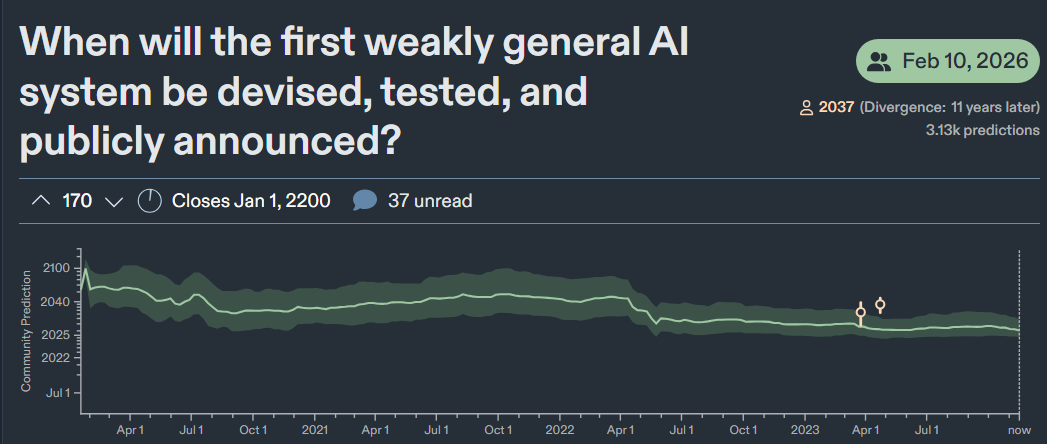
AI Impacts conducted a survey focusing on trends in AI development, targeting researchers who had papers published in NeurIPS or ICML, two of the leading AI conferences, in 2021. Out of the 4,271 researchers initially contacted, 738 responded, resulting in a 17% response rate. These researchers provided their forecasts on various AI-related topics, including the anticipated arrival dates and the expected impacts of AI on different professions(Epoch AI,2023). It is important to note that the results of this survey may exhibit some degree of sampling bias. For instance, it's possible that researchers who have more aggressive predictions about AI's timeline and impact might be more likely to participate in such surveys.

In another survey of top researchers conducted by Grace et al. in 2018, researchers anticipate that AI will surpass human capabilities in several tasks. They predict that AI will excel in language translation by 2024, crafting high-school-level essays by 2026, truck driving by 2027, and managing retail tasks by 2031. Looking further ahead, they foresee AI writing a bestselling book by 2049 and performing surgeries by 2053. Additionally, these researchers estimate a 50% likelihood that AI will outperform humans in all tasks within the next 45 years and automate every human job within 120 years.



Forecast of top researchers, Grace et al.,2018

Popular crowd forecasting website, Metaculus has the forecast for a weakly general intelligent AI system at February, 2026.

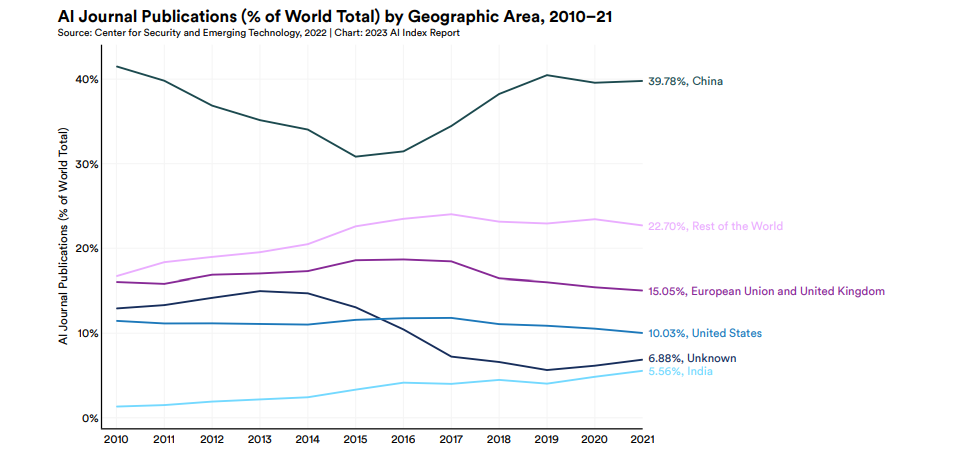


Metaculus Prediction

## **Current Work in BRICS on AI**

BRICS countries have agreed to launch the AI Study Group of BRICS Institute of Future Networks at an early date. BRICS countries have their National AI strategies and data protection agencies, which are vital for the protection of human rights and the safe development of AI. BRICS countries have some of the youngest people in the world who are skilled in IT and can contribute to the economy by means of AI.

Moreover, frameworks such as the UN Sustainable Development Goals, G20 Declaration on AI and OECD AI principles have been guiding lights for the BRICS countries when developing their own national policies.



AI Journal Publications, 2023 AI Index Report by Stanford HAI

Particularly notable is that BRICS countries India and China are major leaders in AI Journal publications with China having nearly 40% of all publications between 2010 and 2021.



Image generated with help of DALL E-3

## **Proposed Standards:**

The following are the standards and policy recommendations that the BRICS AI Study group can put into practice.

1. Principle of AI for Public Benefit

The Parties agree that AI technologies shall be designed, developed, and deployed to enhance public welfare, aligning with global sustainable development objectives.

Reference: United Nations Sustainable Development Goals, ISO/IEC 38507:2017 - Governance of information technology for the organization.

2. Edification Through AI

Each Party commits to utilizing AI to improve educational systems, ensuring accessibility and cultural adaptability.

Reference: UNESCO's Beijing Consensus on Artificial Intelligence and Education, ISO/IEC 23894:2020 - Information technology — Risk management for AI systems.

3. Autonomous AI Systems Regulation

The Parties shall ensure that autonomous AI systems operate within a framework of accountability and transparency.

Reference: European Commission’s Ethics Guidelines for Trustworthy AI, ISO/IEC 42010:2011 - Systems and software engineering — Architecture description.

4. Adoption of Universal AI Guidelines

The Parties shall adopt a common set of guidelines for AI, ensuring adherence to human rights, diversity, and non-discrimination principles.

Reference: OECD Principles on Artificial Intelligence, ISO/IEC 42001 - Information technology — Management system standards guidance.

5. Human-Centric AI Decision-Making

AI decision-making processes shall be subject to human oversight, especially in critical sectors impacting human rights.

Reference: EU’s White Paper on Artificial Intelligence, ISO/IEC 27001:2013 - Information technology — Security techniques — Information security management systems — Requirements.

6. Risk Tolerance in AI

Each Party shall define standards for assessing and managing AI-related risks.

Reference: IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, ISO/IEC TR 24028:2020 - Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence.

7. Data Protection and Privacy

The Parties commit to stringent data protection standards.

Reference: General Data Protection Regulation (GDPR), ISO/IEC 29100:2011 - Information technology — Security techniques — Privacy framework.

8. Prohibition of AI in Lethal Autonomous Weapons

The development and deployment of AI in lethal autonomous weapons are hereby prohibited.

Reference: International Committee of the Red Cross’s position on autonomous weapons, IEEE P7000 - Model Process for Addressing Ethical Concerns During System Design.

9. Mitigation of AI Bias

Methodologies to identify and mitigate biases in AI systems shall be developed, ensuring alignment with human dignity.

Reference: UNESCO's Recommendation on the Ethics of Artificial Intelligence, ISO/IEC TR 24029-1:2021 - Information technology — Artificial intelligence — Assessment of the robustness of neural networks.

10. Harmonization of AI Standards

The Parties shall collaborate to harmonize AI standards, facilitating interoperability and shared best practices.

Reference: IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, ISO/IEC 30141:2018 - Internet of Things (IoT) — Reference architecture.

11. Knowledge Exchange in AI

A platform for knowledge exchange among the Parties shall be established to share advancements and solutions in AI.

Reference: World Economic Forum’s Centre for the Fourth Industrial Revolution, ISO/IEC TR 24368:2020 - Information technology — Artificial intelligence — Use cases.

12. Algorithmic Explainability and Transparency

Standards ensuring the explainability and transparency of AI algorithms shall be implemented.

Reference: AI Now Institute’s Report on Algorithmic Accountability, ISO/IEC TR 24027:2020 - Information technology — Artificial intelligence — Bias in AI systems and AI-aided decision making.

## **Policy Recommendations:**

Recommendations for AI Policy Among BRICS Nations

*Whereas, recognizing the need for a collaborative approach in the development and governance of Artificial Intelligence (AI) technologies among BRICS nations, and aiming to establish a framework for cooperation, ethical standards, and mutual growth in the field of AI, the following recommendations are proposed:*

Cross-Border Safety Toolings

The Parties shall develop and implement infrastructure support for monitoring and auditing AI systems, along with inclusive datasets to foster a thriving AI ecosystem.

This shall include mechanisms for regular cross-border data exchange, and maintaining data integrity and security.

2) Periodic Peer Review Mechanism

A rotating periodic peer review system shall be established, where each Party reviews the AI policies of another, identifying opportunities for cooperation and offering recommendations for achieving the aforementioned standards.

3) National Data Storage and Cybersecurity Standards

Each Party agrees to develop a practice of national data storage, with shared cybersecurity standards pursuant to international guidelines, ensuring the protection of citizen data from unauthorized access and breaches.

4) Joint Technical Advisory Committee

A joint committee of technical experts from all member states shall be formed to advise governments on best practices in AI. This committee will also facilitate the exchange of research, development strategies, and ethical considerations in AI, working with the BRICS Business Council.

5) Annual Rotating AI Audit Among Members

An annual audit system is proposed, where each member state evaluates the AI systems and policies of another, ensuring compliance with agreed standards and promoting continuous improvement.

6) Openness to Independent Audits

All member states shall commit to transparency by being open to independent audits of their National AI strategies and implementations conducted by recognized international observer agencies.

7) AI Risk Classification Scheme

The Parties shall collaborate to develop a comprehensive risk classification scheme for AI systems, addressing different levels of impact and complexity.

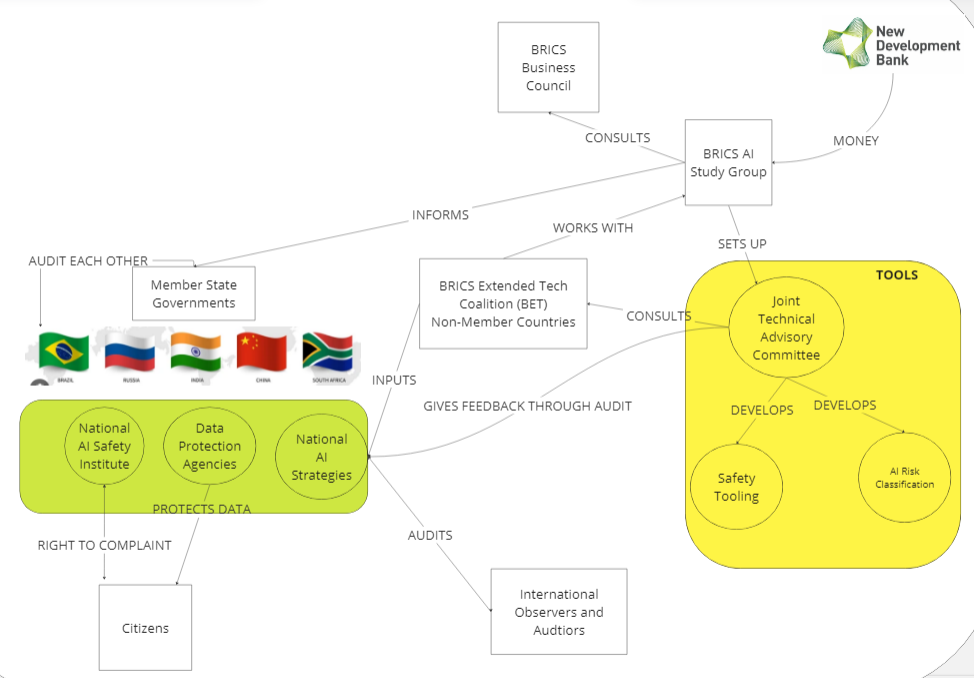
8) National AI Safety Institutes

Each Party is recommended to establish a National AI Safety Institute. These institutes shall provide mechanisms for citizens to report concerns related to AI bias or human rights violations, ensuring accountability in AI deployment.

9)BRICS Extended Tech Coalition (BET)

The establishment of the BRICS Extended Tech Coalition (BET) is recommended for engaging non-member states. BET will facilitate technical cooperation, input on standards, and implementation strategies in AI.

**SCHEMATIC DIAGRAM:**



Schematic of Proposed Governance Structure

## **Limitations:**

A lot of this is an iterative process of ironing out details subject to negotiations and other considerations of each individual member state. Some of the limitations of the standards and proposed policy recommendations are given below:

1) The BRICS Study Group was formed with the purpose of monitoring development in AI and not drafting policy recommendations. However, the experts on the Group can recommend this for effective monitoring and use of AI technology to member governments.

2) Policies on data storage and cybersecurity must balance global cooperation with national interests and data sovereignty issues.

4) The establishment of new committees and institutions (like the National AI Safety Institute) may require significant resources. Clear guidance on funding and resource allocation is necessary.

5) The policies cover broad areas and might need more specific implementation strategies for different countries, considering their unique socio-political and technological landscapes.

The feasibility of harmonizing standards across diverse nations should be carefully assessed, especially in areas like data privacy and AI in weaponry.

## **Conclusion:**

The development of AI systems is occurring at a breathtaking pace, and there must be guardrails on these developments so that the societal fabric isn’t disturbed. Care must be taken by citizens and regulators that AI systems are aligned with human values, respect human rights and are used for maximizing social welfare. The BRICS nations have a key role to play in this growth and must set standards so that AI can be used for the maximum productivity possible and solve the pressing problems of emerging countries amidst the 21st century with a goal of being a leader by 2100.

## **References:**

1. ISO/IEC 38507:2022. (2022). Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations. ISO. Retrieved from <https://www.iso.org>

2. ISO/IEC 23894:2023. (2023). Information technology — Artificial intelligence — Guidance on risk management. ISO. Retrieved from <https://www.iso.org>

3. ISO/IEC/IEEE 42010:2011. (2011). Systems and software engineering — Architecture description. ISO. Retrieved from https://www.iso.org

4. OECD. (2019). AI Principles Overview. Retrieved from <https://oecd.ai>

5. ISO/IEC 27001:2013. (2013). Information technology — Security techniques — Information security management systems — Requirements. ISO. Retrieved from https://www.iso.org

6. IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. (n.d.). Retrieved from https://standards.ieee.org

7. ISO/IEC TR 24028:2020. (2020). Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence. ISO. Retrieved from <https://www.iso.org>

8. GDPR.eu. (2018). What is GDPR, the EU’s new data protection law? Retrieved from https://gdpr.eu

9. ISO/IEC 29100:2011. (2011). Information technology — Security techniques — Privacy framework. ISO. Retrieved from <https://www.iso.org>

10. Goldman Sachs. (2023, April 5). Generative AI could raise global GDP by 7%. Retrieved from https://www.goldmansachs.com [50†source].

11. Xinhua. (2023, August 24). Full Text: Remarks by Chinese President Xi Jinping at 15th BRICS Summit. China Daily. Retrieved from https://www.chinadailyhk.com [54†source].

12. UNESCO. (2019, June 25). First ever consensus on Artificial Intelligence and Education published by UNESCO. Last updated: April 20, 2023. Retrieved from https://www.unesco.org [58†source].

13. Digital Watch Observatory. (2023, August 23). BRICS announces formation of AI study group. Retrieved from <https://dig.watch>

14. Nestor Maslej, Loredana Fattorini, Erik Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons,

James Manyika, Helen Ngo, Juan Carlos Niebles, Vanessa Parli, Yoav Shoham, Russell Wald, Jack Clark,

and Raymond Perrault, “The AI Index 2023 Annual Report,” AI Index Steering Committee,

Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2023.

15. Grace, Katja, et al. (2018). When Will AI Exceed Human Performance? Evidence from AI Experts. Journal of Artificial Intelligence Research. Retrieved 2023, from https://arxiv.org/pdf/1705.08807.pdf.

16. Epoch AI (2023). Literature Review of Transformative Artificial Intelligence Timelines. Retrieved 2023, from https://epochai.org/blog/literature-review-of-transformative-artificial-intelligence-timelines. ​