Strathmore Law Review (SLR)
eISSN: 2414-8164 – Volume 9, Issue 1 (2024) p.37-62
https:doi.org/10.52907/slr.v9i1.495

BIT by BIT: Regulation of Highly Capable AI Through the Remodelling of Bilateral Investment Treaties

*Esther Omanga**

Abstract

This study addresses the pressing need for regulatory frameworks to govern highly capable Artificial Intelligence (AI) systems, which pose significant risks such as potential misuse to create or enhance biological threats. The research investigates the indirect role of Bilateral Investment Treaties (BITs) in regulating highly capable AI, particularly focusing on redesigning BIT clauses in light of the risks associated with the technology. Through qualitative research and examination of BIT clauses, the study finds that while current BITs do not explicitly address the risks posed by highly capable AI, there is potential for redesigning BIT provisions to do so by: i) amending the definition of 'investment' to explicitly include data centres, which play a significant role in the development of highly capable AI, and ii) broadening the scope of scenarios where police powers can be exercised to include AI risks. This study subsequently invites further research on the implementation of the proposed redesigned BITs to mitigate the risks associated with highly capable AI.

Key words: *AI governance, highly capable AI, bilateral investment treaties (BITs), data centres, expropriation, police powers, International Investment Law*

* The author is a Bachelor of Laws (LLB) graduate from Strathmore University (Nairobi, Kenya). This author dedicates this piece to all those who boldly try new things, knowing they will not be perfect but embrace imperfection as part of the journey towards progress, Bit by Bit.

I. Introduction

With the evolving nature of AI, the definition of what constitutes high risk AI is increasingly subject to change.[1] There is a possibility that highly capable foundation models could constitute high-risk AI due to their capacity to inflict substantial harm.[2] An example of such harm includes enabling non-experts to create and produce new biological weapons.[3] Large language models (LLMs)[4] can support the development and strategy for acquiring, altering, and spreading biological agents.[5] Recently, LLMs aided non-scientist students in not only pointing out four possible pandemic pathogens but also explaining how they can be made. As if this was not enough, the LLM went further to specify companies that deal in synthetic DNA without the process of reviewing customers and their orders.[6] This is dangerous because, without proper screening, there is a risk that individuals with malicious intent could use synthetic DNA to engineer harmful biological agents.[7]

The risks presented by highly capable AI lie in its ability to surpass human performance across various cognitive tasks, which is the key defining characteristic of the technology. Humans are unique because they can plan, learn, communicate,

---

1 Anderljung M, Barnhart J, Korinek A, Leung J, O'Keefe, Whittlestone J, Avin S, Brundage M, Bullock J, Cass-Beggs D, Chang B, Collins B, Collins T, Fist T Hadfield G, Hayes A, Ho L, Hooker S, Horvitz E, Kolt N, Schuett J, Shavit Y, Siddarth D, Trager R, and Wolf K, 'Frontier AI regulation: Managing emerging risks to public safety' *ArXIV,* 2023, 3 – arXiv:2307.03718v3 on 6 July 2023.

2 Anderljung M *et al*, 'Frontier AI regulation: Managing emerging risks to public safety' *ArXIV*, 2023, 7 – arXIV:2307.03718V3 on 6 July 2023.

3 Mirsky Y, Demontis A, Kotak J, Shankar R, Gelei D, Yang L, Zhang X, Lee W, Elovici Y and Biggio B, 'The threat of offensive AI to organisations' *ArXIV*, 2021, 5 – arXiv:2106.15764 on 30 June 2021.

4 LLMs could be considered to serve as foundational components for the development of highly capable AI, hence why they are used as an example here. Large language models (LLMs) sit at the forefront of AI development today, meaning they are among the most advanced systems in AI research, capable of handling complex tasks like language understanding, generation, and multi-modal processing. Their continued evolution is expected to lead to advanced AI, as improvements in computational power, data, and algorithms drive better performance. This could amplify the risks it presents. See Department for Science, Innovation and Technology, 'Capabilities and risks from frontier AI: A discussion paper on the need for further research on AI risk' AI Safety Summit, Buckinghamshire, 1-2 November 2023, 5-15.

5 Sandbrink J, 'Artificial intelligence and biological misuse: Differentiating risks of language models and biological design tools' *ArXIV*, 2023, 3 – https://arxiv.org/pdf/2306.13952.pdf on 24 June 2023.

6 Soice E, Rocha R, Cordova K, Specter M and Esvelt K, 'Can large language models democratize access to dual use biotechnology?' *ArXIV,* 2023, 2 – https://arxiv.org/pdf/2306.03809.pdf on 6 June 2023.

7 Sandbrink J, 'Artificial intelligence and biological misuse: Differentiating risks of language models and biological design tools' *ArXIV*, 2023, 3 – https://arxiv.org/pdf/2306.13952.pdf on 24 June 2023.

remember, imagine, reason and explain in ways that other species cannot.[8] These cognitive abilities matter because they are useful for problem-solving and other daily life activities.[9] They empower individuals to engage in pursuits like scientific research and engineering which play a crucial role in shaping society.[10] By surpassing human performance in these cognitive tasks, highly-capable AI could gain power in fields that humans currently dominate, possibly affecting the long term potential of humanity.[11] Among other factors, highly capable AI is likely to perform better than humans in these tasks because unlike the human brain, AI is not constrained by cell count, energy, communication speed and memory.[12]

Even though there exist compelling reasons to believe that highly capable AI could surpass human level intelligence and pose grave threats to humanity, there are arguments to the contrary. Notably, there is the argument that it may not be as easy for highly capable AI to perform better than humans because human intelligence is a collective matter rather than an individual one. What this means is that humans owe their success to a large community of other intelligences in the past.[13] A counterargument to this would be that in as much as we owe a lot to collective intelligence, it is built upon individual intelligences, which AI has already demonstrated that it can surpass in some specific tasks. For instance, AI models have defeated the world's best players in the game of Go. This indicates that AI is rapidly evolving to match or exceed human capabilities in tasks traditionally considered exclusive to human intelligence.[14] However, there are certain things such as legal rights and access to political positions that AI currently does not have which could prevent it from gaining power as rapidly as we think it will.[15]

---

8 Carlsmith J, 'Is power-seeking AI an existential risk?' *ArXIV,* 2022, 5 – arXiv:2206.13353 on 16 June 2022.

9 Skagerlund K, 'Decision-making competence and cognitive abilities: which abilities matter? 35(1), *Journal of Behavioural Decision Making*, 2021, 1.

10 Carlsmith J, 'Is power-seeking AI an existential risk?' *ArXIV,* 2022, 7 – arXiv:2206.13353 on 16 June 2022.

11 Carlsmith J, 'Is power-seeking AI an existential risk?' *ArXIV,* 2022, 7 – arXiv:2206.13353 on 16 June 2022.

12 Carlsmith J, 'Is power-seeking AI an existential risk?' *ArXIV,* 2022, 8 – <arXiv:2206.13353> on 16 December 2023.

13 Grace K, 'Counterarguments to the basic AI c-risk case' LessWrong, 14 October 2022 - <https://www.lesswrong.com/posts/LDRQ5Zfqwi8GjzPYG/counterarguments-to-the-basic-ai-x-risk-case> on 18 December 2023.

14 Silver D, Schrittwieser J, Simonyan K, Simonyan K, Antonoglou L, Huang A, Guez A, Hubert T, Baker L, Lai M, Bolton A, Chen Y, Lillicrap T, Hui F, Sifre L, Van Den Driessche G, Graepel T and Hassabis D, 'Mastering the game of Go without human knowledge', *Nature 550*, 354, 2017 -<https://doi.org/10.1038/nature24270> on 18 December 2023.

15 Grace K, 'Counterarguments to the basic AI c-risk case' LessWrong, 14 October 2022 - <https://www.lesswrong.com/posts/LDRQ5Zfqwi8GjzPYG/counterarguments-to-the-basic-ai-x-risk-case> on 18 December 2023.

Since the 1940s, it has been expected that machines matching human intelligence will be made[16] and now, more than ever, these kinds of machines are likely to be developed. This can be attributed to the increasing availability of data and compute,[17] key factors driving AI advancements today,[18] which were not as accessible in the past. With regard to compute, improvements in computing techniques and specialised computing hardware (graphic processing units and tensor processing units) over the past 10 years have led to an increase in computing power by about 10 times each year, making it easier to develop advanced AI systems.[19] Turning to data, AI companies like OpenAI have managed to collect huge amounts of high-quality data from the internet in recent years to train their models.[20] Despite concerns about the depletion of high-quality data, companies continue to find new (sometimes controversial) ways to source the data needed for training.[21]

Due to the pivotal role played by data and compute in the development of highly capable AI, data centres are becoming increasingly important. Data centres are specialised facilities that host extensive hardware needed to process

---

16 Bostrom N, *Superintelligence: paths, dangers, strategies*, Oxford University Press, Oxford, 2014, 4.

17 Compute is a term used to refer to processing power, memory, networking, storage, and other resources required for AI. AWS 'What is compute', 2024 -<https://aws.amazon.com/what-is/compute/> on 21 May 2024; Pilz K and Heim L, 'Compute at scale: a broad investigation into the data enter industry' *ArXIV*, 2023, 2 – <arXiv:2311.02651> on 5 November 2023.

18 Open AI, 'AI and compute' 16 May 2018 -< https://openai.com/index/ai-and-compute/> on 5 November 2023. Also see Sastry G, Heim L, Belfield H, Anderljung M et al., 'Computing power and the governance of artificial intelligence', *ArXiV*, 2024, 8 - < 2402.08797> on 5 November 2023. Compute has played a prominent role in recent AI progress. Computing power (from the associated hardware and software) is used to execute algorithms and serves as the 'substrate' for the information processing involved in AI. Also see Aldoseri A, Al-Khalifa K, Hamouda A, 'Re-thinking data strategy and integration for artificial concepts, opportunities and challenges', *Applied sciences*, 2023, 12 - <Re-Thinking Data Strategy and Integration for Artificial Intelligence: Concepts, Opportunities, and Challenges>. Vast amounts of data are required to build highly capable AI systems because the method used to train highly capable AI (deep learning), requires lots of datasets to enable AI to 'know' about the world and to perform its tasks.

19 Before 2012, using Graphics Processing Units (GPUs) for AI was rare. Between 2012 and 2016, larger-scale training using ten to one hundred GPUs became more common. By 2016, improvements in computing techniques and specialized hardware, like GPUs and Tensor Processing Units (TPUs), made it easier to train more powerful AI models. This has led to the development of more advanced AI models over the years. See Open AI, 'AI and compute' 16 May 2018 -< https://openai.com/index/ai-and-compute/> on 5 November 2023.

20 Lambert N, 'Synthetic data: anthropic's CAI from fine tuning to pretraining, OpenAI's superalignment, tips, types and open examples', Interconnects, 2023 - <Synthetic data: Anthropic's CAI, scaling, OpenAI's Superalignment, tips, and open-source examples> on 15 September 2024.

21 Marwala T, Fournier-Tombs E, Stinckwich S, 'The use of synthetic data to train AI models: opportunities and risks for sustainable development, United Nations University, Number 1 2023, 3 - <UNU_PUBLICATION_Technology Brief 1-2023 – The Use of Synthetic Data to Train AI Models.indd> on 23 March 2024.

large amounts of data. These centres are important because they house the infrastructure necessary to train AI models. Data centres are essential for managing both the storage of vast datasets and the compute power needed for high-speed processing in deep learning.[22]

Additionally, data centres provide the infrastructure needed for cloud computing, which involves the delivery of computing resources over the internet. Cloud computing allows researchers to leverage remote systems. [23] Experts have noted that it makes sense to train AI at a single location because of the large amounts of data that need to be transferred between computing cores. This means that in a single location there is low latency,[24] which describes a computer network that is optimised to process a large volume of data messages with minimal delay.[25] It is plausible that AI could be trained in one location but operated in another. For example, GPT-4 was trained in Des Moines, Iowa, before being deployed for use globally.[26] Given these advancements in training and deployment, the potential risks and harms caused by AI are increasingly pertinent.[27]

This paper proposes that one way to address the risks of highly capable AI is through regulatory mechanisms such as Bilateral Investment Treaties (BITs), which can influence how AI-related infrastructure (data centres) can be regulated. Bilateral Investment Treaties (BITs) are treaties between two countries aimed at promoting and protecting investments made by investors.[28] The rationale behind BITs is to protect investor's investments from unfair exercise of power from host states. Behind BITs is the assumption that these treaties contain clear and

---

22 Zhang M, 'How data centres are enabling artificial intelligence', Digital Infra Real Estate 2.0, 15 March 2023 - <How Data Centers are enabling Artificial Intelligence (AI) - Dgtl Infra> on 18 June 2024.

23 Rescale, 'What is cloud high performance computing' -<https://rescale.com/cloud-hpc/> on 21 November 2023.

24 Informatica, 'What is low latency' -<https://www.informatica.com/nl/services-and-training/glossary-of-terms/low-latency-definition.html#:~:text=Low%20latency%20is%20desirable%20in%20online%20gaming%20as%20it,a%20more%20realistic%20gaming%20environment > on 21 November 2023.

25 Heim L, 'Compute at scale: a broad investigation into the data center industry' *ArXiV*, 2023, 17 – <arXiv:2311.02651> on 5 November 2023.

26 O'brien M and Fingerhut H, 'AI technology behind ChatGPT was built in Iowa - with lots of water' Bristol Herald Courier, 15 September 2023 -<https://heraldcourier.com/news/nation-world/business/ai-technology-for-chatgpt-was-built-in-iowa-with-lots-of-water/article> on 24 November 2023.

27 Bengio Y, 'AI and catastrophic risk', Journal of Democracy, September 2023 - <https://www.journalofdemocracy.org/ai-and-catastrophic-risk/> on 24 November 2023.

28 Kerner A, 'What can we really know about BITs and FDI?' 33(1), *Foreign Investment Law Journal*, 2018, 1.

enforceable rules that safeguard an investor's right and encourage investment.[29] BITs contain substantive standards that protect against dispossession of property such as nationalisation and expropriation.[30] Before substantive provisions can apply, two things must be established. First, that the purported investment constitutes a valid form of investment under the BIT and second, that there is a valid investor.[31]

While BITs aim to protect investors from unfair actions by host states, they also acknowledge that states have legitimate regulatory interests. States can exercise their regulatory interests through police powers. These regulatory powers are invoked by states for a public purpose such as to protect public morals, human, animal or plant life or health and safety.[32] Currently, there are no BITs that envision AI labs or data centres as investments that would allow for states to invoke the right to regulate when highly capable AI exhibits dangerous capabilities. While police powers have been used to address environmental harm or threats to human life, there is currently little discussion or consideration of using these powers in cases where highly capable AI poses dangers, such as assisting in the creation of biological weapons or providing malicious actors with advanced cyberattack capabilities.[33] The crucial role of police powers, in this context, lies in their ability to serve as an ex-post mechanism for regulating AI once something goes wrong. Since some risks, such as biological threats, may not be easy to detect with current detection mechanisms,[34] police powers become essential for addressing unforeseen issues after they arise.

Based on the foregoing, this study explores the potential for reshaping BIT provisions on the definition of an investment and scenarios where police can be exercised for the indirect regulation of highly capable AI. It argues that BITs can be a useful regulatory tool for highly capable AI if: i) we include data centres as an 'investment' in definitional clauses in order to allow the substantive clauses to

---

29 Salacuse J, 'BIT by BIT: The growth of bilateral investment treaties and their impact on foreign investment in developing countries' 24(3) *The International Lawyer*, 1990, 663.

30 Article 13(1), *Energy Charter Treaty*, December 1994, 2080 UNTS 100, Article 6, *US Model Bilateral Investment Treaty*, 2012, Article 6(2), *France-Seychelles BIT*, 29 March 2007, Article 1110 NAFTA. Salacuse J, 'BIT by BIT', 670.

31 The Organisation for Economic Cooperation and Development (OECD), *Definition of Investor and Investment in International Investment Agreements*, 2008, 10 and 17.

32 Fraga M and Reetz R, *Public purpose in international law: Rethinking regulatory sovereignty in the global era*, Cambridge University Press, 2015, 155.

33 Anderljung M, 'Frontier AI regulation: Managing emerging risks to public safety' *ArXiV,* 2023, Footnote 5 – <arXiv:2307.03718v3> on 6 July 2023.

34 Egan J and Rosenbach E, 'Biosecurity in the age of AI: What's the risk? Harvard Kennedy School Belfer Center, 6 November 2023 < https://www.belfercenter.org/publication/biosecurity-age-ai-whats-risk> on 19 May 2024.

apply to them and ii) broaden the scope of the clauses in BITs containing police powers of states to include some of the risks that highly capable AI may present.

It is worth noting that the findings of this paper are especially relevant to Africa. Power and spatial requirements influence the location of data centres thus many of these warehouses are in remote areas.[35] Resultantly, there has been a notable increase in the presence of data centres in global south countries such as South Africa, where data centres owned by Amazon and Microsoft now operate.[36] Despite the cross-border potential for AI risks, this growing trend underscores the need for African countries to rethink their BITs, ensuring these agreements are adapted to address the unique challenges and opportunities posed by AI development and foreign investment in the region.

This article proceeds as follows: Section II, with Section I being this introduction, considers the role of data centres in building advanced AI and explains the risks that highly capable AI could pose. Section III discusses what constitutes an investment in international investment law and the scenarios where police powers can be invoked. Additionally, the section proposes that BITs can be indirect regulators of highly capable AI through the inclusion of data centres as an 'investment' in definitional clauses and broadening scenarios where police powers are invoked to include AI risk. Section IV steps back to analyse the instances that have influenced BIT remodelling in the past. This section argues that the factors that drove past BIT redesigns are comparable to the context of highly capable AI and its associated risks, suggesting that redesign is possible. Section V is the conclusion that ties up the arguments of this paper.

II. The Role of Data Centres in Building Highly Capable AI and Potential Dangers

i. The Influence of Scaling, Compute and Hardware Advancements on AI Progress

Compute has played a prominent role in recent AI progress. Compute refers to processing power, memory, networking, and storage. Computing power

[35] Isberto M, 'Are there benefits of a rural data center?' Colocation America, 9 June 2021 - <https://www.colocationamerica.com/blog/rural-data-center-benefits> on 22 November 2023.

[36] Microsoft Azure, 'Microsoft opens first data centers in Africa with general availability of Microsoft Azure', 6 March 2019 - <https://azure.microsoft.com/en-us/blog/microsoft-opens-first-data-centers-in-africa-with general-availability-of-microsoft-azure/> on 2 March 2024, Rachel England, 'Amazon opens its first cloud data center in Africa', Engadget, 22 April 2020 - <https://www.engadget.com/amazon-opens-its-first-cloud-data-center-in-africa-101534093.html> on 18 April 2024.

(from the associated hardware and software) is used to execute algorithms and serves as the 'substrate' for the information processing involved in AI.[37]

This study assumes the scaling hypothesis, which suggests that increasing the size of models, the amount of compute and the volume of data leads to better performance and the ability to solve more complex problems.[38] Contrary to a normal outcome of research, the blessing of scale suggests that harder problems are easier to solve than simple problems. For AI, the larger the neural network, compute, data, or problem the faster the AI can learn. A good example of this is GPT-2 vis-a-vis GPT-3. Many researchers were impressed with the strides in natural language tasks such as text summarisation that GPT-2 could do. GPT-3 surpassed GPT-2's capabilities due to great increase in size and compute power.[39]

In light of the above, a reason to believe that developers are likely to develop highly capable AI is the availability of compute. Additionally, it is likely that compute will continue to be available. This is hinged on three justifications. First is Moore's law, which is the idea that computing power increases approximately every two years due to better technology. Second, parallelized computing through high performance computing (HPC) allows many chips to work on the same task at the same time, which speeds up the process. Third, specialised chips such as graphic processing units (GPUs) have been developed to run algorithms. A GPU is a specialised chip for graphics rendering and machine learning. GPUs are essential for training AI because they can process large amounts of data in parallel, speeding up the complex calculations needed for deep learning models.[40] These chips have led to significant improvements in AI.[41]

---

37 Sastry G, Heim L, Belfield H, Anderljung M et al., 'Computing power and the governance of artificial intelligence', *ArXiV*, 2024, 8 - < 2402.08797> on 18 April 2024.

38 Aryeh Englander, 'How would the scaling hypothesis change things?', LessWrong, 13 August 2021, < https://www.lesswrong.com/posts/ED28KSXKc4j8CNoi8/how-would-the-scaling-hypothesis-change-things> on 18 May 2024.

39 Branwen G, 'The Scaling Hypothesis', Gwern.net, 29 May 2020 - <https://gwern.net/scaling-hypothesis> on 10 January 2024.

40 Pilz K, Heim L, 'Compute at scale a broad investigation into the data centre industry', *ArXiV,* 2023, 17 - <https://arxiv.org/pdf/2311.02651.pdf> on 18 April 2024.

41 Buchanan B, 'The AI triad and what it means for national security strategy', Center for Security and Emerging Technology, August 2020 – 17-19 < https://cset.georgetown.edu/publication/the-ai-triad-and-what-it-means-for-national-security-strategy/ > on 18 April 2024.

ii. The Role of Data Centres in Building Highly Capable AI

Data centres are facilities that contain hardware at scale and provide computational resources.[42] A data centre can also be described as a data hall, warehouse, or AI laboratory. There are two main types of data centres. First, enterprise data centres which are private facilities that are owned by a company for its own use. Second, colocation data centres are those facilities that are rented out to individual companies to host their hardware.[43]

As noted earlier, training highly capable AI requires a lot of computational resources. This has led to the increased demand for GPUs.[44] Due to this growing demand, there has been a scarcity of chips (GPUs). Consequently, this has put cloud infrastructure firms such as Microsoft Azure and Google Cloud at an advantage because they own large data centres equipped with extensive GPU resources. These firms lease GPU capacity to AI developers for remote access via the cloud, enabling them to train and deploy advanced models without needing to invest in the scarce and expensive physical hardware themselves.[45]

Moreover, data centres are attractive due to their ability to support HPC. Cloud infrastructure firms usually offer HPC solutions as part of their services.[46] HPC is the practice of aggregating computing resources to gain performance greater than that of a single computer.[47] HPC can be carried out in two ways, either using supercomputers or individual computers known as clusters. The three main components of HPC are compute, networking, and storage. HPC is desirable because it allows organisations to solve large computational tasks at high speeds by bringing together multiple computers that produce more computing power.[48] Sometimes HPC is also referred to as supercomputing since some of

---

42 Pilz K, Heim L, 'Compute at scale a broad investigation into the data center industry', *ArXiV,* 2023, 2 - <https://arxiv.org/pdf/2311.02651.pdf> on 18 April 2024.

43 Pilz K, Heim L, 'Compute at scale a broad investigation into the data center industry', *ArXiV,* 2023, 17 - <https://arxiv.org/pdf/2311.02651.pdf>.

44 A GPU is a specialised chip for graphics rendering and machine learning. GPUs are essential for training AI because they can process large amounts of data in parallel, speeding up the complex calculations needed for deep learning models. Pilz K, Heim L, 'Compute at scale a broad investigation into the data centre industry', *ArXiV,* 2023, 17 - <https://arxiv.org/pdf/2311.02651.pdf>.

45 Vipra J and West S, 'Computational power and AI', AI Now Institute, 27 September 2023, < <https://ainowinstitute.org/publication/policy/compute-and-ai> on 20 December 2023.

46 'High performance computing (HPC)', Amazon Web Services, 2024, < https://aws.amazon.com/hpc/> on 16 May 2024.

47 <https://cloud.google.com/discover/what-is-high-performance-computing> on 1 November 2023.

48 Jones K, 'Launching a new class of US supercomputing' Office of Science, 17 November 2022 -< https://arxiv.org/pdf/2312.04616> on 17 May 2024.

these computational tasks may be too large for standard computers to handle alone and may take too long to process.[49] High performance computing (HPC) makes it easier for an enterprise to run traditional applications and store data in an economical way compared to relying solely on supercomputers. Additionally, HPC uses compute to produce faster and more accurate results. HPC can be done both on-premises and off-premises or a hybrid of both.[50] Through HPC there is simultaneous execution of tasks across multiple processors which makes it possible to train highly capable AI from anywhere in the world.[51]

The release of GPT-4 in 2023 suggests that developers are making significant progress towards developing highly capable AI through on premises high performance computing. This model was trained using the cloud service of OpenAI's investor, Microsoft. OpenAI and Microsoft co-created supercomputers with tens of thousands of Nvidia GPUs connected to each other. From this, the cloud company planned to distribute the knowledge acquired during design to the AI market meaning that more AI companies may be able to build large-scale AI systems.[52]

iii. Highly Capable AI can Exhibit Dangerous Capabilities

In this section, the author focuses on the risks posed by highly capable AI arising from malicious use.[53] This is because these risks align with the areas traditionally governed by police powers, being public morals, health and safety. The potential for AI to be misused in various harmful ways such as through cyberattacks, bioterrorism or over-surveillance mirrors the types of threats police powers aim to prevent.[54]

---

[49] <https://cloud.google.com/discover/what-is-high-performance-computing> on 1 November 2023.

[50] Linthicum D, The cloud as supercomputer, 30 November 2021 - < https://www.infoworld.com/article/3642848/the-cloud-as-supercomputer.html> on 2 November 2023.

[51] What is cloud high performance computing? <https://rescale.com/cloud-hpc/> on 4 November 2023.

[52] Moss S, 'As OpenAI releases GPT-4, Microsoft details Azure AI infrastructure behind it, 15 March 2023 <https://www.datacenterdynamics.com/en/news/as-openai-releases-gpt-4-microsoft-details-azure-ai-infrastructure-behind-it/> on 5 December 2023.

[53] Hendrycks D, Mazeika M, Woodside T, 'An overview of catastrophic AI risks', *ArXIV*, 2023, 6-7 - < https://arxiv.org/pdf/2306.12001.pdf> on 3 February 2024.

[54] *Chemtura Corporation v Government of Canada* (2010), United Nations Commission on International Trade Law.

a. Cyberattacks

A key concern raised by Brundage et al is that advanced AI is likely to be used to commit cybercrimes. Currently, AI in the cybersecurity field is being used on the defensive side for activities such as spam and malware detection. However, according to data collected from a Blackhat conference, 62% of the researchers at the conference believe that there is a strong likelihood that AI will be used for cyberattacks soon. Research has shown that because these systems are trained on machine learning algorithms that may allow them to learn on their own, there is a chance that these systems may learn to evade human-authored malware. The danger lies in the fact that attackers could potentially rely on AI's ability to improve on its own through experience to create attacks that IT professions are not prepared for.[55]

Due to the threat machine learning algorithms pose on cybersecurity, researchers have developed ways such as adversarial machine learning to protect against deception attacks by AI. An adversarial attack is a deceiving technique that attempts to fool machine learning algorithms by using defective data. For example, an adversarial attack on a self-driving car will make the car take a stop sign as a speed limit.[56] However, with the likelihood of development of highly capable AI, there is also the likelihood that these attacks will become more dangerous causing severe attacks in many sectors. For example, in the medical sector, researchers have found that adversarial attacks can make machine learning algorithms to classify benign moles as malignant.[57]

b. Bioterrorism

A study done by Hendrycks reveals that highly capable AI is likely to facilitate bioterrorism. Currently, general purpose LLMs such as ChatGPT have the ability to amass expert knowledge about dangerous pathogens such as smallpox and can even provide a systematic guide on how a person can create them while disregarding safety protocols.[58] From this, there is reason to believe that more capable AI could be of great help to bioterrorists since they will be able to not only synthesise more information but also acquire information that

[55] Brundage M et al, 'The malicious use of artificial intelligence: forecasting, prevention, and mitigation, 2018, 31 - <https://arxiv.org/pdf/1802.07228.pdf> on 5 December 2023.

[56] Boesch G, 'What is adversarial machine learning? Attack methods in 2024' Viso.ai, 2024 - <https://viso.ai/deep-learning/adversarial-machine-learning/> on 20 January 2024.

[57] Finalyson S, 'Adversarial attacks on medical machine learning: Emerging vulnerabilities demand new conversations', 363(6433), *Science*, 2019, 1.

[58] Soice E, Rocha R, Cordova K, 'Can large language models democratize access to dual-use biotechnology?' *ArXIV*, 2023, 3 - < https://arxiv.org/pdf/2306.03809.pdf> on 2 February 2024.

is not accessible to the public. LLMs are likely to contribute greatly to biological misuse by decreasing the obstacles needed for one to access biological agents. Normally, such agents are not easily accessible to the public.[59]

There are arguments that LLMs hallucinate thus they can give false information to bioterrorists. According to Sandbrink, there is reason to believe that developers will develop more sophisticated AI systems that are likely to be more accurate thus contributing to bioterrorism. Soon, LLMs will be fine-tuned on biological data, and these will become the most powerful biological design tools (BDTs). Presently, BDTs can only perform simple functions such as creating proteins using simple functions, but soon, we are likely to see an evolution of BDTs which could also lead to biological misuse.

c. Highly capable AI can be Misused by Authoritarian Regimes for Surveillance and Control

Many law enforcement agencies around the world are looking for new technologies to ensure easy detection and prosecution of criminals in a bid to keep their citizens safe. To achieve this goal, governments have opted to use technology such as surveillance cameras, automated license plate readers, body cameras, drones and now facial recognition technologies.[60] Facial recognition technology (an AI-based technology)[61] works by identifying human features in an image and comparing these features to other images to determine whether it is the same person.[62] While facial recognition technologies have been praised for its advantages such as fraud detection, airport and border control, there is concern that these technologies can cause fear and censorship among citizens. Use of advanced technology for surveillance adds a layer of complexity as citizens will be constantly monitored.[63] There is also the possibility that these technologies could be manipulated by malicious actors which could lead to false convictions. In countries such as China, surveillance has stifled dissent and debate in society

---

59 Hendrycks D, Mazeika M, Woodside T, 'An overview of catastrophic AI risks', 7.

60 Almeida D, 'The ethics of facial recognition technologies, surveillance and accountability in an age of artificial intelligence: a comparative analysis of US, EU and UK regulatory frameworks, 2, *AI and Ethics*, 2021, 1.

61 Raposo V, Li Du 'Facial recognition technology: is it ready to be used in public health surveillance?' *International Data Privacy Law*, 2023, 1 - < https://doi.org/10.1093/idpl/ipad021> on 5 December 2023.

62 What is facial recognition? - <https://aws.amazon.com/what-is/facial-recognition/> on 10 February 2024.

63 Crawford K, Dobbe R et al, 'AI now 2019 report', page 20, New York: AI now Institute 2019, < https://ainowinstitute.org/publication/ai-now-2019-report-2> on 5 December 2023.

which are crucial aspects for any democratic society.[64] Left in the hands of authoritarian governments, facial recognition technologies could be used to acquire and maintain power by making policies that may not be favourable to their citizens.[65]

III. Understanding Investments and Police Powers: Gaps in BIT Clauses in the Era of Highly Capable AI

i. What Constitutes a Covered Investment in BITs

BITs are important because they offer different standards of protection to foreign investors and their investments. These could be substantive or procedural guarantees. Some substantive guarantees include national treatment, most-favoured nation treatment, fair and equitable treatment, full protection and security and protection from unlawful expropriation. Procedural protections relate to dispute settlement.[66]

Investment agreements today acknowledge that the definition of what constitutes an investment is constantly changing. Hence these agreements prefer to give an open-ended definition of what includes an investment. These agreements often use the words, 'every kind of asset' followed by an illustrative list of what could qualify as an investment. Multilateral trade agreements and regional agreements such as the Energy Charter Treaty (ECT) and North American Free Trade Agreement (NAFTA) give broad definitions to investments. Article 1(6) of the ECT defines an investment as 'every kind of asset'. NAFTA provides a broad definition, but it adds an exhaustive list of assets with specific exclusions.[67]

BITs, a type of investment agreement, take four factors into consideration when defining an investment. First, the form of the investment, second, the sector in which the investment operates economically, third, when the investment was created and fourth, the relationship between the investor with the other

[64] Llaria Mazzacco, 'The AI-Surveillance Symbiosis in China', Big Data China, 27 July 2022, < https://bigdatachina.csis.org/the-ai-surveillance-symbiosis-in-china/> on 20 May 2024.

[65] Seger E, Dreksler N et al, 'Open sourcing highly capable foundation models: An evaluation of risks, benefits and alternative methods for pursuing open-source objectives', *Centre for the Governance of AI*, 2023, 13, on 15 December 2023.

[66] Polanco R, 'The impact of digitization on international investment law: are investment treaties analogue or digital?' 24(3), *German Law Journal*,2023, 574.

[67] OECD, Definition of Investor, and Investment in international investment agreements, 2008, 49.

contracting state.[68] Additionally, BITs usually include a list of what could constitute an investment. The list can be classified into five categories: i) assets that can be physically moved or fixed in place, ii) ownership stakes in businesses which involve investments in debt and equity, iii) claims to money and claims under a contract having a financial value, iv) intellectual property rights such as patents, trademarks, and industrial designs and v) the government granted rights for business operations such as licenses for extraction of natural resources.[69]

According to Article 1 of several BITs such as the United States (US)-Argentina BIT, US-Albania BIT, US-Mozambique BIT, US-Lithuania BIT, US-Croatia BIT, an investment includes every kind of asset in the territory of one party owned or controlled directly or indirectly by nationals or companies of the other party. This includes equity, debt and service and investment contracts.[70] These BITs provide a non-exhaustive list of what could fit into the definition of an investment. The BITs list tangible and intangible property which includes rights, mortgages, liens, and pledges as investments.[71]

Some recent BITs such as Angola-European Union (EU) Sustainable Investment Facilitation Agreement (SIFA)[72] define what an investment is, but do not provide a non-exhaustive list of what fits into the definition. Others, such as the Angola-Japan BIT, still go ahead to list what falls under an investment which is the usual list in most BITs. BITs such as the Turkey-United Arab Emirates (UAE) BIT include the list of what constitutes an investment but also include a list of what does not constitute an investment. The Turkey-UAE BIT notes that claims to money, commercial transactions for sale of goods and services and extension of credit for financing commercial trade do not include investments.[73] These varying approaches illustrate the different forms that investment definitions can take in BITs.

Article 25 of the Convention for the Settlement of Investment Disputes Between States and Nationals of Other States (ICSID Convention) states that,

---

68 Salacuse J, Sullivan N, 'Do BITs really work? An evaluation of bilateral investment treaties and their grand bargain', 46, *Harvard International law Journal*, 2005, 67.

69 OECD, Definition of Investor, and Investment in international investment agreements, 2008, 50.

70 Article 1, *US-Argentina BIT* 20 October 1994; Article 1, *US-Albania BIT* 4 January 1998; Article 1, *US-Mozambique BIT* 3 March 2005; Article 1, *US-Lithuania BIT* 22 November 2001; Article 1, *US-Croatia BIT*, 20 June 2001.

71 Article 1, *US-Argentina BIT* 20 October 1994; Article 1, US-Albania BIT 4 January 1998; Article 1, *US-Mozambique BIT* 3 March 2005; Article 1, *US-Lithuania BIT* 22 November 2001; Article 1*, US-Croatia BIT*, 20 June 2001.

72 Article 3, *Angola-EU SIFA* 18 November 2023.

73 Article 9, *Turkey-United Arab Emirates BIT*, 24 July 2011.

'the jurisdiction of the centre shall extend to any legal dispute arising directly out of an investment'.[74] This means that for a tribunal of the International Centre for Settlement of Investment Disputes (ICSID) to have jurisdiction over a matter, there must be a covered investment. While the ICSID Convention does not give a definition of 'investment', ICSID tribunals generally apply the Salini test[75] to determine whether a supposed investment is an investment. For a supposed investment to be considered an 'investment' under this test: i) it must involve significant contributions such as funds, equipment or expertise by the investor, ii) it must span a minimum length of time of two to five years, iii) the investor must participate in the risks of the investment, and iv) the investment should contribute to economic development of the host state.[76] There is some contention on how to apply the test. Some tribunals state that the factors are interdependent. Other tribunals have applied a modified version of the test which includes adding or removing one of the factors.[77]

Tribunals have also discussed the issue of territoriality of an investment. In *SGS v Philippines* and *BIVAC v Paraguay*, the tribunals stated that even if the investment is partially conducted outside the host state's territory, it still forms a key part of the overall investment.[78] On the other hand, some tribunals do not agree with this approach, in *Saba Fakes v The Republic of Turkey* have removed the fourth part of the test arguing that an economic contribution to the host state is difficult to establish.[79] In *Lesi v People's Democratic Republic of Algeria*, the tribunal argued that the fourth part of the test is covered in the other factors of the test.[80]

ii. Scenarios when Police Powers can be Invoked

The doctrine of police powers gives a state the inherent right to regulate for the sake of public interest and states that successfully invoke this doctrine are not obliged to pay compensation. For a state to invoke this doctrine, they

---

74 Article 25, *Convention on the settlement of investment disputes between States and nationals of other states*, 14 October 1966, 575 UNTS. The International Centre for Settlement of Investment Disputes (ICSID) is an institution established by the World Bank to provide arbitration for international investment disputes.

75 S*alini Costruttori S.p.A and Italstrade S.p.A v Kingdom of Morocco* (2001), ICSID. The Salini test named after the landmark case Salini v Morocco, is a set criteria used by ICSID tribunals to determine whether a transaction qualifies as an 'investment' under the ICSID Convention.

76 *Salini Costruttori S.p.A and Italstrade S.p.A v Kingdom of Morocco* (2001), ICSID 52-58.

77 Chan D, Lai J, 'Two decades after Salini v Morocco: The case for retaining the Salini test with modifications', Arbitration International, 39(1), 2023, 2.

78 *SGS Societe Generale de Surveillance S.A v Republic of Paragua* (2012), ICSID para 504.

79 *Saba Fakes v Republic of Turkey* (2010), ICSID para 203.

80 *Lesi S.p.A and Astaldi, S.p.A v People's Democratic Republic of Algeria* (2005), ICSID para 308.

ought to ensure that the measures imposed were bona fide, non-discriminatory, proportionate, and follow due process.[81]

When assessing the police powers doctrine, tribunals often disagree on the scope of legitimate purposes. Some tribunals are of the view that a narrow scope should be adopted while other tribunals are of the opinion that many scenarios could fall under a legitimate purpose. Some tribunals observe that only measures relating to tax, crime and public order constitute a legitimate purpose, but others note that a broader approach with measures relating to health, safety and morality should be adopted.[82] This captures the debate that is often had regarding the police powers doctrine. Many authors are of the opinion that this doctrine allocates to states too much power and is not investor friendly.[83]

BITs often include clauses that clarify non-discriminatory regulatory actions aimed at protecting public welfare do not constitute indirect expropriation. For instance, the Canada-Jordan BIT and the Canada-Peru BIT contain provisions stating that regulatory actions taken to safeguard public welfare are not considered indirect expropriation.[84] BITs also include a clause titled 'general exceptions' that lists scenarios that amount to a legitimate purpose.[85]

What amounts to a legitimate purpose has been a point of contention, but some tribunals have granted states the space to impose regulatory measures if they are taken in a bona fide manner. In *Methanex Corporation v United States of America*, the claimant was a Canadian based manufacturer of methanol which is an ingredient in methyl tertiary-butyl ether (MTBE). When making MTBE, a gasoline additive, a significant percentage of methanol is used. The State of California imposed a ban on MTBE because it was polluting surface water and groundwater. However, the claimant was of the view that the ban was imposed for political reasons in favour of its rival company that made ethanol that is also

---

81 United Nations Conference on Trade and Development, Expropriation, UNCTAD series on issues in international investment agreement II, xii -< https://unctad.org/system/files/official-document/unctaddiaeia2011d7_en.pdf> on 14 March 2012.

82 Mostafa B, 'The Sole effects doctrine, police powers and indirect expropriation under international law, 15(267), Australian *International Law Journal*, 2008, 274-277.

83 Mostafa B, 'The Sole effects doctrine, police powers and indirect expropriation under international law' 15(267), *Australian International Law Journal*, 2008, 269.

84 *Canada-Jordan BIT* 14 December 2009, Article 13, *Canada-Peru BIT* 20 June 2007, 'Neither party shall nationalise or expropriate a covered investment either directly, or indirectly through measures having an effect equivalent to nationalisation or expropriation (hereinafter referred to as 'expropriation'), expect for a public purpose in accordance with due process of law, in a non-discriminatory manner and on prompt, adequate and on prompt, adequate and effective compensation'.

85 Article 6.14, *India-Republic of Korea CEPA* 1 January 2010; Article 10 *Canada-Jordan BIT* 14 December 2009.

used to make MTBE. Based on this, the claimant argued that the ban amounted to indirect expropriation. The tribunal stated that the regulation by the State of California to ban the use of MTBE was a non-discriminatory regulatory measure taken to protect human, plant and animal health as MTBE.[86] On one hand critics of the decision such as Martini are of the view that the decision failed to consider the impact that the regulation had on the investment. On the other hand, the decision has been praised for granting states flexibility to impose measures that substantially deprive investors of their property rights for the sake of a legitimate purpose.[87]

In *Chemtura Corporation v Government of Canada* (Chemtura decision), the tribunal stated that Canada validly exercised its police powers when it phased out the use of lindane (pesticide) in agricultural products due to the dangers it presented on human health and the environment. The claimant had been using lindane on canola but in March 1998, it was declared that importing canola seeds treated with unregistered pesticides such as lindane was illegal. The claimant was allowed to import canola that was treated with lindane only until 1 June 1998 but continued to do so past the date. Ultimately, in February 2002, the Pest Management Regulatory Agency (PMRA) notified the claimant of the suspension of the claimant's company registrations due to continued use of lindane.[88] The claimant argued that this suspension amounted to indirect expropriation of their company. The respondent emphasised that they were well within their right to regulate to protect human health and the environment.[89] The tribunal acknowledged this stating that the ban of the use of lindane and suspension did not amount to indirect expropriation.[90] The Chemtura decision has not been without criticism. Critics argue that the tribunal failed to take the right approach when assessing the matter. They posit that the tribunal ought to have used a two-step analysis by first assessing the impact of the measure on the investment and then looking into whether the expropriation is lawful under the BIT. In essence, this argument centres on whether the police powers doctrine

86 *Methanex Corporation v United States of America* (2005), United Nations Commission on International Trade Law (UNCITRAL), para 300.

87 Martini C, 'Balancing investors' rights with environmental protection in international investment arbitration: An assessment of recent trends in investment treaty drafting' 50(3), *International Lawyer*, 2017, 553.

88 *Chemtura Corporation v Government of Canada* (2010), United Nations Commission on International Trade Law, para 6-30.

89 *Chemtura Corporation v Government of Canada* (2010), United Nations Commission on International Trade Law, para 253.

90 *Chemtura Corporation v Government of Canada* (2010), United Nations Commission on International Trade Law, para 261.

should be considered as an exception or after an analysis of the impact of the measure and whether it amounted to lawful.[91]

In *Philip Morris Brands Sarl, Philip Morris Products S.A and Abal hermanos S.A (claimants) and Oriental Republic of Uruguay*, Uruguay passed two tobacco control measures. First, a single presentation requirement and second, the 80/80 regulation. According to the respondent, these regulations were put in place to curb the high smoking rates in Uruguay. The first regulation required manufacturers to print health warnings on 50% of the area of the cigarette packages and the second regulation required manufactures to print health warnings on 80% of the package.[92] It was the claimant's argument that these regulations infringed on their legally protected right to use a trademark thus reducing the value of their investment.[93] The tribunal found that the regulations were not arbitrary or unjustifiable and they were put in place to address public health concerns. The regulations were also put in place to adhere to the WTO FCTC (Framework Convention on Tobacco Control) which Uruguay had ratified.[94]

iii. BITs as Indirect Regulators of AI risks

For a long time, tribunals preferred a deferential approach when deciding what constitutes an investment. Tribunals made this assessment on the meaning of 'investment' based on the circumstances of the case and the parties' agreement. Tribunals were liberal by respecting without serious inquiry states' ex ante decisions to cover all kinds of economic enterprises under the protection of the ICSID convention.[95] Recently, tribunals have shifted from a liberal approach to a stricter one, relying on the Salini test, which is often criticized as being strict and inflexible when determining what qualifies as an 'investment'.[96] Conversely, two cases *Biwater Gauff v Tanzania and Malaysian Historical salvors (annulment)* have rejected the use of the Salini test and instead the tribunals relied on state practice

---

91 Martini C, 'Balancing investors' rights with environmental protection in international investment arbitration: An assessment of recent trends in investment treaty drafting' 50(3), *International Lawyer*, 2017, 549.

92 *Philip Morris Brands Sarl, Philip Morris Products S.A and Abal hermanos S.A (claimants) and Oriental Republic of Uruguay* (2016), International Centre for Settlement of Investment Disputes, para 10 and 274.

93 *Philip Morris Brands Sarl, Philip Morris Products S.A and Abal hermanos S.A (claimants) and Oriental Republic of Uruguay* (2016), International Centre for Settlement of Investment Disputes, para 11.

94 *Philip Morris Brands Sarl, Philip Morris Products S.A and Abal hermanos S.A (claimants) and Oriental Republic of Uruguay* (2016), International Centre for Settlement of Investment Disputes, para 13.

95 Mortenson J, 'The meaning of "investment": ICSID's travaux and the domain of international investment law'*, Harvard International Law Journal*, 51(1), 2010, 259.

96 Mortenson J, 'The meaning of "investment": ICSID's travaux and the domain of international investment law'*,* 259.

to determine an 'investment'.[97] The concern with tribunals employing the Salini test is that they often rely on external criteria to determine what qualifies as an investment, granting tribunals broad discretion in this area.[98] The core purpose of BITs is to allow the parties to decide on key provisions, including what constitutes an investment. This is intended to decrease weighing and balancing by arbitrators and enhance predictability.[99] As economic realities evolve, more items are being explicitly defined as investments in BITs rather than being left under general clauses like 'any kind of asset' or 'every kind of asset'.[100]

For instance, Intellectual Property Rights (IPRs) are now commonly recognised as investments in BITs, even though few cases have been adjudicated on such rights. Another example is the ongoing debate over whether data should be explicitly included as a covered investment.[101] Subsequently, leaving data centres at the mercy of general clauses in BITs undermines the predictability that BITs ought to provide. Including data centres explicitly in the list of covered investments is useful. For example, in the event AI is linked to a risk such as over-surveillance then BITs can be used as an indirect form of regulation in allowing states to expropriate and exercise their police powers over the data centre as an investment.

Data centres play a crucial role in the development and operation of AI systems. They house the computing infrastructure necessary for training and deploying AI models, making them integral in building of highly capable AI. While some may argue that data centres fall under the category of 'movable property' and automatically qualify as investments, this perspective overlooks their unique characteristics, which sets them apart from ordinary labs or warehouses. First, data centres go beyond physical storage by enabling cloud computing, which is integral to developing highly capable AI. Second, they are equipped with specialised infrastructure, including high-capacity servers and advanced cooling

---

97 From 2006 to 2010, seven decisions had relied on the Salini test with only *Biwater* and *Malaysian Historical Salvors* rejecting it. Mortenson J, 'The meaning of "investment": ICSID's travaux and the domain of international investment law', 277, *Malaysian Historical Salvors II* annulment and *Biwater Gaufff (Tanz.), Ltd v United Republic of Tanzania* (2008), International Centre for Settlement of Investment Disputes.

98 Mortenson J, 'The meaning of "investment": ICSID's travaux and the domain of international investment law', 278.

99 Peterson E, 'Bilateral Investment Treaties and Development Policy-making', International Institute for Sustainable Development, 2004, 3 - <Bilateral Investment Treaties and Development Policy-Making> on 5 November 2023. BITs include absolute standards of treatment and relative standards of treatment for the benefit of investors.

100 McLaughlin M, 'Regulating artificial intelligence in international investment law', 275.

101 McLaughlin M, 'Regulating artificial intelligence in international investment law', 24(2) *Journal of World Investment and Trade*, 2023, 275.

systems to prevent overheating. Third, they house vast amounts of sensitive data, making them vulnerable to security breaches that could result in significant data leaks.

A challenging issue arises with colocation data centres which involve shared infrastructure used by multiple AI companies. Expropriating such a centre due to the risks posed by one AI system could inadvertently disrupt the operations of other AI companies that are not connected to the harmful AI. While this issue is significant, it falls beyond the scope of this paper.[102]

The narrow drafting of BIT clauses on police powers could present significant challenges in the age of highly capable AI. BITs often fail to address the specific risks posed by AI, making it difficult for states to justify expropriation when necessary. This gap is concerning given the potential dangers of AI such as enabling bioterrorism, increasing cybersecurity threats or allowing authoritarian regimes to exert greater control over their citizens.[103] States have also struggled to prove that certain measures were 'necessary' to protect the public, which has often led to the failure of actions taken in the public interest.[104] To address this, BITs could include AI-related risks as a general exception. A provision like this would give states a clear legal basis to justify measures taken to mitigate these risks.

IV. Redesigning BITs: Past Examples and Similarities with AI- Related Risks

i. BITs are Redesigned to Create a Balance Between Host States and Investors Interests

For a long period of time, BIT clauses read as though they were designed to protect investors over the host state. This was because of the view that protection of foreign investors would automatically lead to an increase in economic growth. This kind of design has been linked with US model BITs that restrict host states from imposing performance requirements on investors. It has been discovered that this assumption is not true as increase in foreign

[102] Pilz K, Heim L, 'Compute at scale a broad investigation into the data center industry', *ArXIV,* 2023, 6- <https://arxiv.org/pdf/2311.02651.pdf> on 5 November 2023.

[103] Mclaughlin M, 'Regulating artificial intelligence in international investment law' 24(2), *Journal of world investment and trade*, 2023, 2.

[104] Mostafa B, 'The Sole effects doctrine, police powers and indirect expropriation under international law, 274-277.

direct investment is caused by other factors such as quality of institutions within the host country.[105]

The substantive clauses of new generation BITs are being worded differently. The new language emphasizes that the standards of protection in BITs are high. The provisions for protection include those relating to the meaning of minimum standard of treatment in accordance with international law and the provisions relating to indirect expropriation. The language in these provisions has been redesigned to remove excess discretionary power given to arbitral tribunals when these provisions are vaguely worded. For these BITs, it is not enough for an investor to prove that a measure taken by the state has had an adverse effect on their asset for it to be indirect expropriation. These BITs have been keen to state that non-discriminatory regulatory actions by a host state to protect public health, safety and the environment do not amount to indirect expropriations.[106]

In 2017, Tanzania decided to end its investment agreement with the Netherlands. This was done because many of its agreements with investors sought to restrict control over its resources. This decision was heavily influenced by civil society groups that were urging the country to enter into agreements that were beneficial to the country's development rather than investors.[107] Similarly, Indonesia terminated and revised over 25 of its treaties with countries such as France, Germany, and India. Indonesia, like Tanzania, was of the view that the agreements' design favoured investors over its regulatory powers.[108] Similarly, the Morocco and Nigeria BIT that was signed in 2016 was praised for introducing provisions that recognize the regulatory powers of states. One of the key features of the Morocco-Nigeria BIT is the provision on regulatory powers that provides ample leeway for the state to decide on environmental matters.[109]

---

105 Hamdy M, 'Redesign as reform: a critique of the design of bilateral investment treaties' 51(2), *Georgetown International Law Journal*, 2020, 308.

106 United Nations Conference on Trade and Development, Recent Developments in international investment agreements, IIA Monitor Number 2, 2005, 1 - < https://unctad.org/system/files/official-document/webiteiit20051_en.pdf> on 13 January 2024.

107 Gathii J, 'Understanding Tanzania's termination of its BIT with the Netherlands in Context' AfronomicsLaw, 1 April 2019 - < https://www.afronomicslaw.org/2019/04/01/understanding-tanzanias-termination-of-its-bit-with-the-netherlands-in-context> on 15 December 2023.

108 Hamzah L, 'Bilateral Investment Treaties (BITs) in Indonesia: A paradigm shift, issues, and challenges' 21(1), *Journal of Legal, Ethics and Regulatory Issues*, 2018, 2.

109 Gazzini T, 'Nigeria and Morocco move towards a 'new generation' of Bilateral Investment Treaties, Blog of the European Journal of International Law, 8 May 2017 - <https://www.ejiltalk.org/nigeria-and-morocco-move-towards-a-new-generation-of-bilateral-investment-treaties/> on 16 January 2024.

ii. BITs have been Remodelled to Meet Environmental Concerns

Due to the many questions pertaining to environmental concerns that arbitral tribunals have had to grapple with, there has been a call to redesign BIT provisions considering evolving environmental standards. Recent BITs of 2022 such as Japan-Bahrain BIT, Oman-Hungary BIT have included unique provisions relating to environmental standards.[110]

Even though many BITs are still silent on environmental matters, there are BITs that have amended their clauses in light of environmental concerns raised by the international community. Clauses that provide policy space for environmental regulation are the forms of environmental texts that are preferred. While the scope of what qualifies as an environmental concern varies on a case-by-case basis, some concerns such as sanitary and phytosanitary issues and measures concerning exhaustible natural resources fit within the realm of environmental concerns.[111]

There has been a growing trend among new BITs to elaborate on the scope of the environmental clause in the BIT. These clauses include the beneficiaries of the protection for example human, animal and plant life or health or the protection of natural resources. Some BITs go as far as stating the elements that the regulation seeks guard against. For example, 'prevention or control of the release or emission of pollutants or environmental contaminants, the control of hazardous or toxic chemicals and wastes and the protection or conservation of wild flora or fauna'.[112]

Although in rare circumstances, some BITs have included clauses that explicitly state that environmental regulation is not indirect expropriation. This is to give states leeway to take on measures necessary for environmental protection without investors disregarding them with the claim of indirect expropriation. These clauses state, 'The parties confirm their shared understanding that [...] Except in rare circumstances, non-discriminatory regulatory actions by a Party that are designed and applied to protect legitimate public welfare objectives, such as public health, safety and the environment, do not constitute indirect expropriations'.[113]

---

110 Article 24, *Japan-Bahrain BIT*, 6 September 2023, Article 3, *Oman-Hungary BIT*, 24 August 2022.

111 United Nations Conference on Trade and Development, Recent Developments in international investment agreements, IIA Monitor Number 2, 2005, 1 - < https://unctad.org/system/files/official-document/webiteiit20051_en.pdf> on 13 January 2024.

112 *Canada-Jordan BIT*, 14 December 2009, *Canada-Latvia BIT*, 27 July 1995, *Canada-Peru BIT*, 20 June 2007. This language resembles GATS Article XIV and GATT Article XX.

113 *United States Model BIT*, 1 November 2005; *Canada-Jordan BIT*, 14 December 2009; *Canada-Peru BIT,* 20 June 2006; *United States-Uruguay BIT*, 1 November 2006.

Over the years, there has also been inclusion of BIT clauses that discourage lowering of environmental standards for the sake of increasing investments. A sample of these types of clauses is, 'Contracting parties recognise that it is inappropriate to encourage investment by investors of other Contracting Party by relaxing environmental measures. To this effect each Contracting Party should not waive or otherwise derogate from such environmental measures as an encouragement for the establishment, acquisition, or expansion in its territory of investments by investors of the other Contracting Party'. The Japan-Korea BIT, Japan-Peru BIT, Japan-Uzbekistan BIT and Japan-Vietnam BIT are good examples of BITs with such clauses.[114]

This growing trend in BITs tries to strike a balance between not only safeguarding investor rights but also recognising state sovereignty in contributing to a sustainable environment.[115]

iii. BITs have been Remodelled to Include Human Rights Clauses

There is a debate about how far human rights should permeate into investment disputes and how far tribunals should go in considering human rights claims in investment disputes. The conventional view is that human rights and investment should exist separately, like oil and water, as such very few BITs have been redesigned to include human rights clauses.[116] According to Aatreya, human rights are a fundamental component of international law. Therefore, investment arbitrations should consider human rights, even though investment agreements typically lack provisions explicitly requiring tribunals to address them.[117] Scholars such as Jun Zhao propose that investors should receive investment treaty protection after complying with human rights. This would also empower states to introduce counterclaims on the grounds of human rights against investors, making it easier to hold investors accountable.[118]

---

[114] *Japan-Korea BIT*, 1 January 2003, *Japan-Peru BIT*, 10 December 2009, *Japan-Uzbekistan BIT*, 24 September 2009, and *Japan-Vietnam BIT*, 19 December 2004.

[115] Gordon K, Pohl J, 'Environmental Concerns in International Investment Agreements: A survey', OECD Working Papers on International Investment Number 1, 2011, 11, - < https://www.oecd.org/investment/investment-policy/WP-2011_1.pdf> on 15 February 2024.

[116] Mehmet T, Schultz T, 'The state a perpetual respondent in investment arbitration? Some unorthodox considerations in the backlash against investment arbitration: Perceptions and Reality' in Waibel M, Asha Kaushal, Kyo-Hwa Liz Chung, Balchin C (eds), *The backlash against investment arbitration: perceptions and reality,* Kluwer Law International, 2011, 577.

[117] Aatreya S, 'Human rights and the ISDS Regime - rethinking the bipartisan structure of international investment arbitration' 22(1), *Gonzaga Journal of International Law*, 2019, 30.

[118] Jhun Z, 'Human rights Accountability of Transnational Corporations: A potential response from bilateral investment treaties' 8(1), *Journal of East Asia and International Law*, 2015, 2.

The 2018 Draft Dutch Model BIT carries a provision that provides tribunals with the power to limit an award of compensation to an investor if they do not comply with the 2011 United Nations Guiding Principles on Business and Human Rights.[119] Cotula encourages the reform of BITs to meet human rights obligations. This would affirm investor obligations to adhere to human rights. This is because investors often go unscathed even when engaging in activities such as mining, petroleum, logging, and agribusiness that usually have negative impacts on small-scale farmers, people who live in forests, pastoral communities, artisanal fishers, and Indigenous peoples.[120]

iv. Similarities Between Past BIT Remodelling and AI-Related Risks

The examples above demonstrate two key points. First, they illustrate that remodelling BITs is possible. Second, they highlight specific instances where BIT remodelling has already occurred. This section of the paper will argue that there are similarities between the circumstances that prompted past remodelling efforts and the risks associated with AI, which could justify further remodelling of BITs.

Section IV demonstrates that redesigning has been guided by social issues such as the inclusion of human rights clauses and provisions addressing environmental concerns. In this context, the author argues that the risks associated with AI, such as bioterrorism or enabling authoritarian regimes to engage in mass surveillance, represent significant social issues that should similarly influence the remodelling of BITs. Just as environmental concerns and human rights have prompted the evolution of BIT provisions, the emerging risks of AI demand explicit recognition to empower states to address these risks through the protections within investment treaties.

Practically, this could involve redesigning BIT clauses to explicitly grant states the right to exercise their police powers in response to specific AI-related risks. For instance, BITs could include provisions prohibiting the use of facial recognition technologies (FRT) by authoritarian regimes for surveillance and control over their citizens. Alternatively, BITs could incorporate a general exception under police powers, allowing states to expropriate investments when necessary to protect human life from the risks posed by highly capable AI. BIT clauses on police powers and expropriation in the age of AI can be worded

[119] Article 3, *Draft Model Bilateral Investment Treaty of the Kindgom of Netherlands*, 22 March 2019.

[120] Cotula L, '(Dis)integration in global resource governance: extractivism, human rights and investment treaties' 23(2), *Journal of International Economic Law*, 2020, 432.

in various ways. For example, the BIT clause can read this way, 'The parties recognise the inherent right of each party to exercise its police powers to protect public welfare, including the protection of public morals, public order, health and safety. In cases where highly capable AI poses specific risks, such as the use of facial recognition technologies for over-surveillance, the party may expropriate provided that such measures are bona-fide, non-discriminatory, proportionate and follow due process'.

The remodelling of BITs to include data centres as covered investments align with the broader trend of making BITs more predictable and objective which is what was intended when remodelling BITs to employ a balance between host state and investor interests. Including data centres explicitly in the definition of 'investment' would provide states with the legal framework to expropriate harmful AI-related technologies. Given that many African countries are eager to attract foreign investment from cloud computing providers, one effective way to incentivise such investment is through measures like tax incentives to encourage the establishment of data centres.[121] Including data centres as covered investments in BITs would provide these countries with the legal certainty and protection that they need to guard their regulatory powers. This predictability is valuable because African countries are often at a disadvantage compared to more powerful countries.[122] In this way, by ensuring that data centres are explicitly recognised within BIT provisions, African countries can balance the need for foreign investment with the protection of their regulatory powers.

However, it is acknowledged that this shift comes with risks. States could misuse these provisions, citing AI risks as a reason to seize data centres from foreign investors unfairly. One solution to this would be to put a requirement for states to consult with AI governance bodies prior to expropriating data centres.

V. Conclusion

This study was sparked by the great likelihood of the existence of highly capable AI and the risks that it may present in the future. The study introduced bilateral investment treaties (BITs) as an indirect regulatory framework for states

[121] Abungu C, Malonza M, Adan S, 'Can apparent bystanders distinctively shape an outcome? Global south countries and global catastrophic risk focused governance of artificial intelligence, *ArXiV*, 2023, 11 - < arXiv:2312.04616**>.**

[122] Yimer A, Geda A, 'What drives foreign direct investment into Africa? Insights from a new analytical classification of countries as fragile, factor driven or investment driven', 15, Journal of the Knowledge Economy, 2024, 14199.

to use considering these risks. This study has investigated the intersection of highly capable AI and BITs and underscores the importance of adapting these BIT clauses to deal with the potential harms of advanced AI technologies.

As a result of the risks associated with highly capable AI such as, LLMs making it easier for biological information to be misused by malicious actors or the use of facial recognition technologies (FRTs) for surveillance and control causing fear in citizens, this study puts forth a few recommendations. First, this study suggests the redesigning of BIT clauses to explicitly recognise data centres as investments given their importance in the development of highly capable AI systems. Second, the author proposes redesigning the criteria for invoking police powers in BITs to take into account the risks associated with these AI systems. This is to ensure that states are well-equipped to address emerging threats by these systems through lawful expropriation. Third, the author also recommends that tribunals should ensure that these BIT provisions are updated to cater for the evolving risks and challenges posed by AI. Finally, the author calls for continuous monitoring of BITs in response to advancements in AI technology.