CHAPTER 02

To What Extent Does Malaysia’s National Fourth Industrial Revolution Policy Address AI Security Risks?

02

To What Extent Does Malaysia's National Fourth Industrial Revolution Policy Address AI Security Risks?

JUN-E TAN

Abstract

The National Fourth Industrial Revolution (4IR) Policy was launched in July 2021 as a guiding document for Malaysia's direction in maximizing growth opportunities and mitigating potential risks arising from 4IR technologies. This chapter explores the policy to examine the extent to which Artificial Intelligence (AI) security risks are addressed, using the AI Security Map by Newman (2019) as a framework. In the policy, 4IR technologies including AI are seen through a techno-utopian lens, therefore its focus centres on rapid adoption rather than regulation and resilience. It is found that most of the policy initiatives focus on economic security and capacity building for the state, in order to keep up with the developmental race. Other areas of AI security such as the risks of unintended consequences or unsafe outcomes of AI, or risks of AI being used for malicious purposes, receive much less attention. However, as the N4IRP is still in its nascent stages of implementation, there is still room for its cross-ministerial governance structure to work on providing safeguards across different domains and sectors to achieve holistic and sustainable development.

Introduction

Artificial intelligence (AI) and the Fourth Industrial Revolution (4IR) have become the next big thing in the developmental race, as countries attempt to harness technology to get ahead, or at least to not be left behind. The Fourth Industrial Revolution describes a transformation in the ways we live, work, and communicate through the application of a range of technologies fusing the physical, digital, and biological worlds.[1] As much as some breakthroughs in technology had powered previous industrial revolutions—such as mechanization with the steam engine, mass production with electricity, and computerization with the semiconductor—AI (defined in this context as algorithms generating algorithms[2]), is one of the foundational technologies that will open up a new era of industrialization.

The National Fourth Industrial Revolution Policy[3] (N4IRP) of Malaysia was launched in July 2021, with the aim of "driving coherence in transforming the socioeconomic development of the country through ethical use of 4IR technologies". The key foci of the policy are on maximizing growth opportunities and on mitigating potential risks arising from 4IR. The policy includes AI as one of five foundational 4IR focal areas, the others being the Internet of Things (IoT), blockchain, cloud computing and big data analytics, and advanced materials and technologies. Among these technologies, AI "is expected to create the most impact", and is considered "the 'electricity' of the 4IR".[4]

Within this chapter, the lens of AI security is used to scrutinize the pathway towards 4IR in Malaysia. AI as a transformative technology has the potential to bring not only societal benefits, but also harms to society that may negate developmental gains. Already in other parts of the world, we see some of these harms in the form of unintentional consequences such as amplified systemic biases resulting in the marginalized being further marginalized,[5] or weaponized AI which efficiently surveil entire populations[6] or carry out automated cyberattacks.[7] The evolution of the technology outpaces the speed in which legal and regulatory safeguards are put in

[1] Klaus Schwab, The Fourth Industrial Revolution (Geneva: World Economic Forum, 2016).

[2] Internet Society, "*Artificial Intelligence and Machine Learning: Policy Paper*" (Internet Society, April 2017), https://www.internetsociety.org/resources/doc/2017/artificial-intelligence-and-machine-learning-policy-paper/.

[3] Government of Malaysia, "*National Fourth Industrial Revolution (4IR) Policy*" (Economic Planning Unit, Prime Minister's Department of Malaysia, July 2021), https://www.epu.gov.my/sites/default/files/2021-07/National-4IR-Policy.pdf.

[4] N4IRP, p59

[5] Ed Pilkington, "*Digital Dystopia: How Algorithms Punish the Poo*r", The Guardian, 14 October 2019, sec. Technology, https://www.theguardian.com/technology/2019/oct/14/automating-poverty-algorithms-punish-poor.

[6] Yael Grauer, "*Surveillance of Uyghurs Detailed in Chinese Police Database*", The Intercept, 29 January 2021, https://theintercept.com/2021/01/29/china-uyghur-muslim-surveillance-police/.

[7] Center for Security and Emerging Technology, Micah Musser, and Ashton Garriott, "*Machine Learning and Cybersecurity: Hype and Reality*" (Center for Security and Emerging Technology, June 2021), https://doi.org/10.51593/2020CA004.

place,[8] and also the ability of the average citizen to understand the implications of the technology and how best to protect oneself against possible dangers.

Situating AI technologies within the application contexts of 4IR brings some advantages. On one hand, the perspective of risks and harms is anchored in applications and implications instead of focusing only on technological limitations and errors. On the other hand, as 4IR covers a wide range of emerging technologies at various stages of maturity, a focus on AI narrows down the possible risks into a smaller array of known issues, which helps in concretizing problems and imagining solutions. That being the case, even though 4IR technologies would have a larger set of security risks, the concern of this chapter is on the ones that are associated with AI.

Malaysia's N4IRP explicitly acknowledges that there will be potential risks arising from 4IR technologies, and states the government's commitment to address them. The objective of this chapter is therefore to review Malaysia's N4IRP, its goals, and in particular its outlined initiatives, to understand the types of AI-related risks it addresses. The chapter also aims to provide a perspective of technology governance from a developing country's context through delving into Malaysia's priorities in balancing the need to be competitive at an international level, yet protect its citizens' well-being locally.

To what extent does Malaysia's N4IRP address AI security risks? This question requires us to first unpack what AI security risks are, which we will do via the types of potential AI security risks from the AI Security Map proposed by Jessica Newman (2019).[9] We then go through a background of developmental policies of Malaysia to situate the N4IRP, and describe the structure of the policy's content. An analysis is provided on the types of AI security risks covered by the policy and the gaps in risk mitigation. The chapter ends with a discussion on assumptions behind the policy direction, and possible implications on technology governance on the country.

Types of AI Security Risks

What are the security risks of AI? In this section, we explore a framework by Jessica Newman, of the Center for Long-Term Cybersecurity, which lists out twenty types of such risks, organized into digital/physical, political, economic, and social domains (see Table 1). Within her 2019 paper, *Toward AI Security: Global Aspirations for a More Resilient Future*, Newman provides comprehensive examples of the risk areas, and also uses the AI Security Map to analyse national AI strategies and policy responses

[8] Gary Marchant, "*Governance of Emerging Technologies as a Wicked Problem*", Vanderbilt Law Review 73, no. 6 (1 December 2020): 1861.

[9] Jessica Cussins Newman, "*Toward AI Security: Global Aspirations for a More Resilient Future*", CLTC White Paper Series (Berkeley: Centre for Long-term Cybersecurity, February 2019), https://cltc.berkeley.edu/wp-content/uploads/2019/02/CLTC_Cussins_Toward_AI_Security.pdf.

of ten countries to determine their preparedness in handling AI security threats and opportunities.

Table 1: AI Security Map (Newman, 2019)

AI SECURITY DOMAINS			
Digital/Physical	Political	Economic	Social
Reliable, value-aligned AI systems	Protection from disinformation and manipulation	Mitigation of labour displacement	Transparency and accountability
AI systems that are robust against attack	Government expertise in AI and digital infrastructure	Promotion of AI research and development	Privacy and data rights
Protection from the malicious use of AI and automated cyberattacks	Geopolitical strategy and international collaboration	Updated training and education resources	Ethics, fairness, justice, dignity
Secure convergence / integration of AI with other technologies (bio, nuclear, etc.)	Checks against surveillance, control, and abuse of power	Reduced inequalities	Human rights
Responsible and ethical use of AI in warfare and the military	Private-public partnerships and collaboration	Support for small businesses and market competition	Sustainability and ecology

Newman defines AI security "as the robustness and resiliency of AI systems, as well as the social, political, and economic systems with which AI interacts", and looks beyond the narrow scope of national security to cover a more comprehensive landscape of security issues. The AI Security Map was chosen as a point of reference because it provides a comprehensive (but, as Newman emphasizes, not exhaustive) overview of the breadth of issues that can be included as AI security risks and risk mitigation. The systemic nature of the risks highlighted by the framework is suitable for national-level analyses; Newman goes beyond risks and accountability issues at a technical level that are often focused upon in discussions on AI governance,[10] and looks at a more holistic range of potential harms on society. That Newman has used the framework to conduct analyses on other countries also helps to provide some global context and different national priorities for comparison.

[10] Thilo Hagendorff, "*The Ethics of AI Ethics: An Evaluation of Guidelines*", Minds and Machines 30, no. 1 (1 March 2020): 99–120, https://doi.org/10.1007/s11023-020-09517-8.

The digital/physical domain of security risks focuses on various aspects of AI systems design and use that can threaten the security of intertwined digital and physical spaces. The political domain focuses on different actors and their interactions within the AI landscape, between state, market, and society. Relationships between actors reflect power imbalances and priorities that are at times aligned (such as between public and private entities), or at times conflictive (such as government surveillance on populations). The economic domain of AI security risks considers on one hand the impacts of AI technologies on the economy, and on the other, the importance of dedicating resources to drive the technology sector in order to not be left behind. For the social domain, security risk mitigation comes in the form of building in principles, rights, and obligations into AI technologies so that negative impacts on society can be minimized.

This researcher takes the liberty to simplify the 20 security areas into the mitigation of three types of risks: 1) the risks or opportunity costs of not implementing AI, missing out on potential benefits; 2) the risks of unintended consequences or unsafe outcomes of AI; and 3) the risks of AI being used for malicious purposes. We will return to the AI security risks later on, and now introduce Malaysia's N4IRP and its policy landscape related to AI.

Malaysia's National 4IR Policy

Background

The N4IRP was unveiled in early July 2021, as a sister policy to the Malaysia Digital Economy Blueprint[11] (MDEB) which was launched in March 2021. The scope of the MDEB is broader, aiming to "transform Malaysia into a digitally-driven, high income nation, and a regional leader in digital economy", whereas the N4IRP zooms in a little closer into transforming the country's socio-economic development through the use of 4IR technologies, by providing key guiding principles and strategic direction, as well as guidelines to addressing risks.

The MDEB and N4IRP are expressly built to support and enable national development, as their goals are aligned with the objectives of Malaysia's developmental master

[11] Government of Malaysia, "*Malaysia Digital Economy Blueprint*" (Economic Planning Unit, Prime Minister's Department of Malaysia, March 2021), https://www.epu.gov.my/sites/default/files/2021-02/malaysia-digital-economy-blueprint.pdf.

plans (the Shared Prosperity Vision 2030[12] and the 12th Malaysia Plan[13] are referenced directly). These two policies are intertwined in that both are administered by the National Digital Economy and 4IR Council which is chaired by the Prime Minister, therefore they share a governance structure. The N4IRP also includes a page on how both policies complement each other. The two policies extend Malaysia's past efforts in developing its digital economy and high-tech ecosystem, notably through the Multimedia Super Corridor (MSC) initiative from the 1990s, to create an IT hub within the country in a version of Silicon Valley. Malaysia's path towards digitalization has been lined with several other policies, such as the National eCommerce Roadmap, the National Industry 4WRD Policy, the National IoT Framework, the National Big Data Analytics (BDA) Framework, the National Fiberisation and Connectivity Plan 2019-2023 (NFCP), and so on.

Malaysia also has policy documents that are focused on AI specifically. There are at least two: the National AI Roadmap (AI-RMap) that was launched in March 2021 by the Ministry of Science, Technology, and Innovation (MOSTI), and the National AI Framework by the Malaysia Digital Economy Corporation (MDEC)[14] which does not appear to have been released publicly.[15] The AI-RMap project was conducted by professors from Universiti Teknologi Malaysia (UTM) and industry experts from the National Tech Association of Malaysia (PIKOM), who were awarded a grant by MOSTI to study and propose paths forward in the area of AI. It was launched in a virtual town hall (because of movement control under the COVID-19 pandemic), introducing the current situation of AI in Malaysia, strategies to diffuse the technology, and proposed national AI projects.[16]

From the AI-RMap website and the few available media reports, it is not readily apparent if the Roadmap is at the stage of being proposed or it is already under implementation.[17] Even though the Roadmap offers specific timelines and action plans between 2021 and 2025, there are very few media reports covering the Roadmap

[12] The Shared Prosperity Vision (SPV) 2030 was launched in 2019 by then Prime Minister Mahathir Mohamad. A key aspirational document that is referenced repeatedly in other policies, SPV 2030 provides a longer term direction, with the primary aim of providing a "decent standard of living to all Malaysians by 2030", elaborated within its three objectives: 1) providing development for all; 2) addressing wealth and income disparities; and 3) building a united, prosperous and dignified nation. SPV 2030 continues the tradition of Malaysia's policy formula of growth, distribution, and unity, from previous grand plans such as the New Economic Policy (1971-1990), Vision 2020 (1991-2020), and the New Economic Model (2010-2020).

[13] Malaysia has five-year plans which steer the direction of the country's development. The 12th Malaysia Plan covers the period of 2021 to 2025.

[14] MDEC is the lead government agency instrumental in developing Malaysia's ecosystem for information and communication technologies and digital economy since the 1990s.

[15] There was no launch media article or announcement found within MDEC's database of press releases. However, the framework was referenced within the AI-RMap website with a snapshot of its cover.

[16] The contents of the Roadmap are available in https://airmap.my, in the form of slides and also videos of presentations given during the town hall, which happened on March 15, 2021.

[17] Attempts were made to reach out to the project leader, with no response.

itself, the virtual town hall event, or its proposed activities.[18] The Roadmap was launched by the Secretary General of MOSTI during the virtual town hall, but there is no mention of the Roadmap in the ministry's website. That it is addressed as a "living document"[19] adds to the tentativeness of the initiative. Erring towards the side of caution, analyses within this chapter focus on N4IRP to indicate Malaysia's priorities and direction when it comes to AI adoption and governance, within a larger context of the Fourth Industrial Revolution.

The Structure of the National 4IR Policy

The N4IRP is published by the Economic Planning Unit of the Prime Minister's Department. Its vision is to harness the power of 4IR technologies to enhance socio-environmental well-being and economic growth. Three missions are outlined: to improve quality of life by leveraging technological advancement, to enhance local capabilities to embrace 4IR across sectors, and to use the technologies to enhance the preservation of ecological integrity. In other words, the N4IRP aims for 4IR technologies, the chief of which is AI,[20] to be used "for good", from social, economic, and environmental points of view. The objectives stated are to seize growth opportunities arising from the 4IR, to create a conducive ecosystem to cope with the 4IR, and to build trust in an inclusive digital society.

The range of technologies covered by the N4IRP is broad, described as new technology that is characterized by "the fusion of physical, digital, and biological worlds, impacting all disciplines, industries and the economy". It covers building capacities in five foundational technologies: 1) artificial intelligence; 2) Internet of Things; 3) blockchain; 4) cloud computing and big data analytics; and 5) advanced materials and technologies, and capabilities in these are expected to be applied across ten key economic sectors[21] and six supporting sectors.[22]

As can be seen in Figure 1, the four *policy thrusts,* or thematic foci of the N4IRP revolve around 1) capacity development and skills training; 2) digital infrastructure development; 3) regulation; and 4) accelerating 4IR technology innovation and adoption. These are broken down into 16 *strategies*, colour coded by "beneficiary groups", which are businesses, government, and society. The 16 strategies are expanded into 32 national initiatives, which have specific timelines assigned to each: initiatives within Phase One to be completed by 2022, Phase Two by 2025, and Phase Three by 2030. For the ten key economic sectors, there are 60 *sectoral initiatives*

[18] The most comprehensive report found was one in the Newshub section of Universiti Teknologi Malaysia (https://news.utm.my/2021/07/ahibs-experts-entrusted-for-ai-roadmap-and-talent-development-in-malaysia/). No mention was found in mainstream news media.

[19] Malaysia Artifical Intelligence Roadmap. https://airmap.my/ai-roadmap-overview

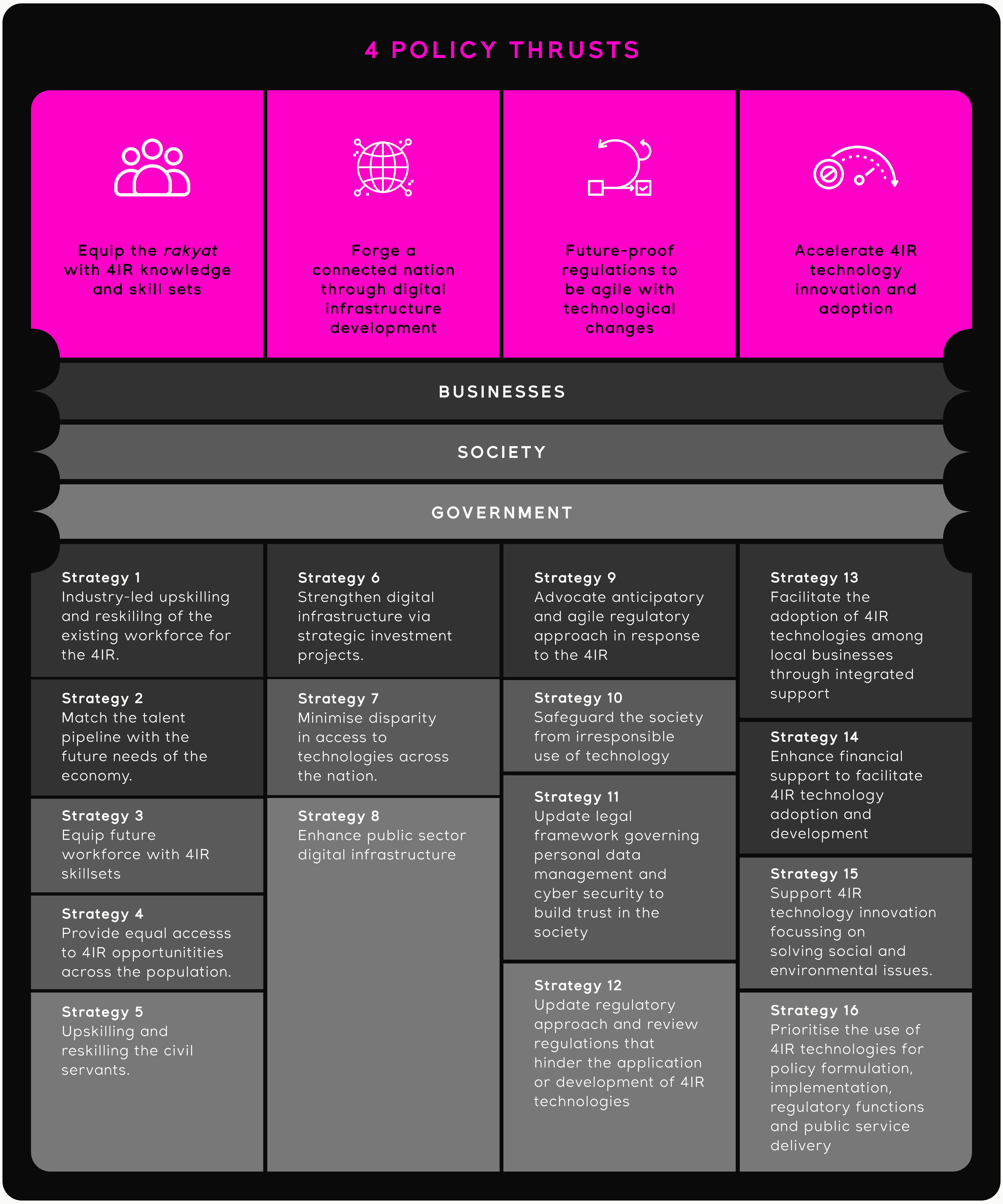
[20] Page 59 of the N4IRP

[21] Including 1) wholesale and retail trade; 2) transportation and logistics; 3) tourism; 4) finance and insurance; 5) utilities; 6) professional, scientific and technical services; 7) healthcare; 8) education; 9) agriculture; and 10) manufacturing.

[22] Including 1) construction; 2) real estate; 3) mining and quarrying; 4) information and communication services; 5) arts, entertainment and recreation services; 6) administrative and support services.

which also align with the four policy thrusts, with some sectoral nuances but mostly adhering to the same themes.

Figure 1: Screenshot of policy thrusts and strategies from the N4IRP



4 POLICY THRUSTS

Equip the *rakyat* with 4IR knowledge and skill sets

Forge a connected nation through digital infrastructure development

Future-proof regulations to be agile with technological changes

Accelerate 4IR technology innovation and adoption

BUSINESSES

SOCIETY

GOVERNMENT

Strategy 1
Industry-led upskilling and reskililng of the existing workforce for the 4IR.

Strategy 2
Match the talent pipeline with the future needs of the economy.

Strategy 3
Equip future workforce with 4IR skillsets

Strategy 4
Provide equal accesss to 4IR opportunitities across the population.

Strategy 5
Upskilling and reskilling the civil servants.

Strategy 6
Strengthen digital infrastructure via strategic investment projects.

Strategy 7
Minimise disparity in access to technologies across the nation.

Strategy 8
Enhance public sector digital infrastructure

Strategy 9
Advocate anticipatory and agile regulatory approach in response to the 4IR

Strategy 10
Safeguard the society from irresponsible use of technology

Strategy 11
Update legal framework governing personal data management and cyber security to build trust in the society

Strategy 12
Update regulatory approach and review regulations that hinder the application or development of 4IR technologies

Strategy 13
Facilitate the adoption of 4IR technologies among local businesses through integrated support

Strategy 14
Enhance financial support to facilitate 4IR technology adoption and development

Strategy 15
Support 4IR technology innovation focussing on solving social and environmental issues.

Strategy 16
Prioritise the use of 4IR technologies for policy formulation, implementation, regulatory functions and public service delivery

Mapping AI Security Risk Mitigation in the N4IRP

In this section, we provide an analysis of AI security risk mitigation in the N4IRP based on the framework of Newman's AI Security Map. As the scope of the N4IRP covers a wider range of technologies than only AI, the policy may be justifiably vague in some coverage of AI security risks. AI in the policy is also addressed from the angle of 4IR, therefore not all AI security risks within Newman's framework may fit within the context of the policy. For example, "protection from disinformation and manipulation" as listed in the map may not be considered as relevant to the 4th Industrial Revolution. However, there is still merit in the exercise of measuring Malaysia's mitigation of AI security risks according to Newman's framework, as most of the risks listed do still apply under the N4IRP, and we will still be able to identify gaps at the domain level.

To separate the rhetoric from the implementation priorities, emphasis is put on examining the national initiatives outlined under the N4IRP's strategies to be carried out in the next decade. These initiatives are concrete action plans with timelines attached, and represent stated commitment by the government to address certain issues. Table 2 provides an overview of Malaysia's plans to mitigate AI security risks, sorting the security areas into three categories: 1) a clear commitment by the N4IRP to address the issue, based on its inclusion in the planned initiatives; 2) indirect reference or acknowledgment within the policy document, which implies possible action; and 3) no mention of the security risk area, which implies a lower likelihood of the risk being managed. Since the initiatives are not described in detail within the policy, there are some ambiguities which require interpretation and assumptions, which are explained below the table, in the order of priority within the N4IRP.

Table 2:
Malaysia's priorities on AI security risk mitigation

		CLEAR COMMITMENT	INDIRECT REFERENCE/ POSSIBLE INCLUSION	NO MENTION
ECONOMIC	Mitigation of labour displacement	/		
	Promotion of AI research and development	/		
	Updated training and education resources	/		
	Reduced inequalities		/	
	Support for small businesses and market competition	/		
POLITICAL	Protection from disinformation and manipulation			/
	Government expertise in AI and digital infrastructure	/		
	Geopolitical strategy and international collaboration	/		
	Checks against surveillance, control, and abuse of power			/
	Private-public partnerships and collaboration	/		
SOCIAL	Transparency and accountability		/	
	Privacy and data rights	/		
	Ethics, fairness, justice, dignity	/		
	Human rights		/	
	Sustainability and ecology	/		
DIGITAL/PHYSICAL	Reliable, value-aligned AI systems		/	
	AI systems that are robust against attack		/	
	Protection from the malicious use of AI and automated cyberattacks		/	
	Secure convergence / integration of AI with other technologies (bio, nuclear, etc.)			/
	Responsible and ethical use of AI in warfare and the military			/

Economic Domain

Economic security is the highest in priority for the N4IRP. This is unsurprising, given that the document was launched by the Economic Planning Unit, and that AI is framed within the context of 4IR and the digital economy. From the 32 national initiatives, more than half (at least 17) are directly linked to the economy, mostly in supporting the promotion of AI research and development, and providing training and education. There are 4IR development centres and innovation parks planned, as well as initiatives to accelerate investment and adoption in businesses. Several training programmes have been proposed, aimed at a wide range of stakeholders, from students to civil servants.

In terms of mitigating labour displacement, there are initiatives to "provide incentives to minimise the risk of job displacements",[23] "enhance formal social protection mechanism for gig workers"[24] and "gradually reduce foreign labour dependency".[25] Micro, small and medium enterprises (MSMEs) are specifically mentioned as recipients of coordinated support and facilitation to accelerate innovation.[26] The only economic security area that is not directly addressed by the outlined initiatives is the reduction of inequalities, but it was acknowledged in the document that 4IR technologies can widen social and economic inequality.[27]

Political Domain

Many of the initiatives fall under the political domain, but most of them (at least nine) focus on the category of government expertise in AI and digital infrastructure. Within that are a number of services targeted at the government sector—such as MyGovCloud to promote cloud computing in the public sector,[28] a 4IR Innovation Accelerator to drive 4IR adoption at all levels of government,[29] and a Government Experience Lab to drive 4IR innovation.[30] The National Digital Identity programme is expected to catalyse more adoption of 4IR technologies at the state level.[31] In terms of geopolitical strategy and global collaboration, there is a WEF Centre for the 4IR planned, "as a hub of global stakeholders' cooperation to facilitate the development of policy frameworks".[32] 4IR development centres are meant to be "industry-led", so there is definitely public-private partnerships outlined.

[23] Initiative 9

[24] Initiative 10

[25] Initiative 4

[26] Initiative 26

[27] Page 21

[28] Initiative 16

[29] Initiative 11

[30] Initiative 32

[31] Initiative 31

[32] Initiative 23

For political security, mis/disinformation and the manipulation of the communication environments were not mentioned within the N4IRP, and neither were checks and balances for surveillance or limitations in power.

Social Domain

Under the social domain, the N4IRP pledges to safeguard society against possible harms by "introducing an ethics framework for technological development, deployment and utilisation",[33] "enhancing personal data protection law, regulations and guidelines",[34] and "introducing specific legislation for cybersecurity".[35] These initiatives are relatively limited in scope, as legal protections are only afforded to personal data protection and cybersecurity issues. The proposed ethics framework, which is not legally binding, seems to cover all other potential harms. In terms of sustainability and ecology, the N4IRP does not discuss the environmental footprint of technology and possible mitigation; what is offered is just support provided to businesses to leverage 4IR technologies to solve socio-environmental issues.[36]

Transparency and accountability of AI systems are not specifically mentioned but as most AI ethical frameworks do cover these,[37] presumably Malaysia's would as well. Human rights are not mentioned within the document. In particular, there is no reference to civil and political rights (CPR), or protections against surveillance. However, as much of the N4IRP focuses on delivering economic, social, and cultural rights (ESCR), it can be argued that the policy does aim to address some aspects of human rights.

Digital/ Physical Domain

Within the digital/physical domain of AI security threats, there are two initiatives that address the issue of cybersecurity, focusing on "introducing specific legislation on cybersecurity"[38] and "enhancing the existing cybersecurity framework by incorporating safeguard measures for the implementation and operationalisation of 4IR across the public sector, with a focus on IoT"[39] (Initiative 25). These do not spell out clearly the aspects of cybersecurity covered, and while "irresponsible use and manipulation of technology" was mentioned a few times in the document as a catch-all phrase for cyber threats, no further elaboration was given. Therefore, potential action could include or exclude any of the digital/physical security risk areas which AI systems can pose a threat to.

[33] Initiative 20

[34] Initiative 22

[35] Initiative 21

[36] Initiative 29

[37] Anna Jobin, Marcello Ienca, and Effy Vayena, "*The Global Landscape of AI Ethics Guidelines*", Nature Machine Intelligence 1, no. 9 (September 2019): 389–99, https://doi.org/10.1038/s42256-019-0088-2.

[38] Initiative 21

[39] Initiative 25

This author takes the discretion to decide that the first three areas within the domain, i.e., reliable and value-aligned systems, systems robust against attacks, and protections against malicious use of AI, could be addressed as part of the cybersecurity initiatives and ethical framework as mentioned before, and therefore can be categorized as "possible inclusions". As for the secure convergence of AI and other technologies and AI in warfare, as they are more specific, the assumption is that they are not addressed at this moment.

Discussion

Within this section, we will discuss the assumptions behind the N4IRP and resulting implications on priorities and implementation of Malaysia's technology governance and AI security mitigation.

The Assumptions

The N4IRP's policy wording and slated initiatives point towards a few underlying assumptions. Firstly, even though risks are mentioned, most of the policy strongly suggests that outcomes of 4IR and its associated technologies, including AI, are largely beneficial. For example, stakeholder groups such as businesses, society and government are addressed within the policy as "beneficiary groups" (see Figure 1). Technology is celebrated as progress, its benefits necessarily outweighing the risks. For the most part, the N4IRP reads fairly typically as a policy document, with the language of visions, missions, strategies, and indicators. However, there is a moment in the text where it breaks character and imagines a techno-utopian scenario:

> “Let us take the agriculture sector as an example of the fusion of technologies. A 4IR-ready farmer will oversee a fleet of sensors and robots, and grow tailor-made crops packed with nutrition. The fresh produce will be purchased by consumers from the comfort of their own homes, enabled by the internet and peer-to-peer business models platform. Instead of in-person collection, autonomous vehicles will transport the goods without the need for human travel. Though this scenario may still be years away for some parts of the world, in many places, this is already commonplace.” (p.20) ”

The sense is that Malaysia needs to be heading towards the above scenario, or risk being left behind. Following that, the second apparent premise of the N4IRP is that 4IR is "an inevitable wave of change"[40] that countries will have to adapt to, with urgency. Success will bring about economic growth, competitive advantage, efficiency, and convenience; failure will result in the country losing the developmental race. In order to ride the wave, Malaysia has no choice but to invest heavily in its 4IR ecosystem in the short- and mid-term.

This brings us to the third underlying assumption of the policy: that, with sufficient resources rapidly invested into infrastructure and capacity-building, Malaysia would catch up with countries that are ahead in the technology race, and reap the fruit of its investments. However, this assumption downplays the overwhelming advantage held by other countries in success factors such as talent and innovation ecosystems in the United States, or oceans of data available in China to train and refine its AI models.

Lastly, as is typical in many developmental projects, economic growth is the main indicator and direction, with an implicit orientation towards trickle-down economics. While the rhetorics have shifted towards sustainable development, the majority of the action plans in the N4IRP focus on economic security, with a business-friendly, business-as-usual approach.

The Implications

The assumptions behind the N4IRP bring a set of implications to technology governance and security risk management. Firstly, technology is viewed from the lens of being a solution rather than a potential problem, and therefore most AI security measures within the policy fall within the bucket of mitigating the risk of being left behind, instead of risks connected to safety and unintended consequences, or abuse with malicious intent. While solutions are touted for sustainable development in rhetoric, the main focus remains to be economic competitiveness. 4IR technologies are not scrutinized for the social and environmental problems that they may bring; instead, great faith is placed on technological innovation which may not address systemic and structural causes to the problems.

Secondly, there is a limited approach towards regulation, with an emphasis on speed instead of safeguards. In the N4IRP, regulatory frameworks were mentioned but specifically within the areas of personal data protection and cybersecurity issues, but there was no mention of legislation in areas such as product safety, protection from AI discrimination and bias, algorithmic accountability and transparency in 4IR technology use, just to name a few areas. Throughout the policy, an "anticipatory and agile regulatory approach" was advocated, elaborated within Initiative 19 as regulations to "meet the needs of the digital economy businesses".[41] The proposed

[40] Subsection within Chapter One, p.20, N4IRP

[41] N4IRP, Page 52

ethics framework seems to be the proposed safety net to address safeguards, but it is not legally binding.

> "4IR technologies are not scrutinized for the social and environmental problems that they may bring; instead, great faith is placed on technological innovation which may not address systemic and structural causes to the problems."

Thirdly, the positioning of 4IR as a key economic enabler to Malaysia's development has certain implications in the governance structure and implementation of the N4IRP. Spearheaded by the Economic Planning Unit which is central to Malaysia's development planning, 4IR and related technologies are elevated into high priority to be mainstreamed across the public sector and civil service. While the lead ministry on digital technologies is the Ministry of Communications and Multimedia (KKMM) and the National Policy on Industry 4.0 (Industry4WRD) focusing on the manufacturing sector is overseen by the Ministry of International Trade and Industry (MITI), these are now consolidated under the National Digital Economy and 4IR Council, led by the Prime Minister.

With six clusters (digital talent, digital infrastructure and data, emerging technology, economy, society, and government) chaired by line ministers and the chief secretary to the government, and relevant ministries slated under the individual clusters, the governance structure embeds a higher level of cross-ministerial and interdisciplinary coordination. Although some have commented that the bureaucracy of the Council may stifle innovation [42] and power dynamics within the Council are yet unclear, it can be argued that some level of friction and feedback loops from relevant ministries and agencies may be beneficial to bring in more holistic considerations and safeguards.

How may this look like in practice? While the N4IRP does not assign lead agencies to policy actions, its sister policy the Digital Economy Blueprint does go to that level of granularity. The MDEB provides an indication on how policy initiatives can be cascaded to ministries that have the mandate and the experience to handle challenges that arise from the digital economy, such as assigning the Malaysia Competition Commission (MyCC) to streamline competition policies and laws, MITI to incorporate digital economy elements into international trade arrangements and negotiations, and the Ministry of Finance to come up with a digital tax framework. These ministries are not traditionally involved in digitalization or technology, but are important for integrating the digital into policy and regulatory frameworks.

[42] Siew Yean Tham, "*Malaysia's Digital Economy Blueprint: More Is Not Better*", FULCRUM, 2 March 2021, https://fulcrum.sg/malaysias-digital-economy-blueprint-more-is-not-better/.

Conclusion

Malaysia's N4IRP follows a familiar playbook of investing in economic and human resources to catch up in the technological and developmental race. The policy plan lays out a ten-year plan of "enhancing 4IR awareness and adoption" (two years, in Phase One), "driving transformation and inclusivity" (three years, Phase Two), and "achieving balanced, responsible and sustainable growth by leveraging 4IR technologies" (five years, Phase Three). Mapped against Newman's AI Security Map, misuse and abuse of AI technologies do not weigh heavily in this trajectory, and much remains unsaid within the policy about safeguards, regulatory or otherwise.

AI security risks aside, a techno-utopian vision of Malaysia's future seems simplistic and divorced from realities on the ground. While supporting 4IR technologies is high up in its priorities, there are many local and global challenges that compete for attention and resources within the country. In August 2021, a month after the launch of the N4IRP, the then Prime Minister Muhyiddin Yassin had to resign and dissolve his cabinet following months of political instability. This was the third change of administration in Malaysia within the span of a little more than three years.[43] As such, Muhyiddin Yassin who provided the foreword in the N4IRP is no longer in power. Indeed, in the recent years Malaysia has been fraught with uncertainties including drastic disruptions by the COVID-19 pandemic and the resulting global economic downturn; it is also vulnerable to the climate and ecological crisis which requires much resources for mitigation[44] and adaptation.[45] The N4IRP which advocates a "whole of nation" approach does not mention how the above conditions faced by the public and private sectors in Malaysia, or indeed, the population in general, may hamper the country's abilities to invest in, coordinate on, and benefit from the Fourth Industrial Revolution.

As the N4IRP has just been announced, there is still much room to refine the country's 4IR pathway to focus on resilience rather than rapid adoption. While policy initiatives do focus on narrow economic gains, the interagency governance structure to implement and monitor the N4IRP has the potential to provide the bridging mechanism and expertise across domains to ensure adequate safeguards, so that the pursuit of technology for development does not come at the cost of sustainable development itself.

[43] In 2018, the 14th General Election of Malaysia saw an unprecedented defeat of the ruling coalition Barisan Nasional which had governed the country from its independence in 1957, and the regime changed hands. The opposition coalition, Pakatan Harapan, came into power, only to be overthrown two years later in 2020 because some members of parliament changed their party allegiance. The new Prime Minister Muhyiddin Yassin governed for 17 months, during which a state of emergency was announced, suspending parliament and all elections due to the worsening COVID-19 pandemic. The political instability continued towards the end of the emergency in August 2021, when Muhyiddin Yassin resigned after losing majority support of the MPs, paving the way for a new cabinet by current Prime Minister Ismail Sabri Yaakob.

[44] The country's plan to be carbon neutral earliest by 2050 was announced during the tabling of the 12th Malaysia Plan in September 2021.

[45] Reuters, "*Malaysia to Spend $335 Million for Flood Relief*", Reuters, 29 December 2021, sec. Commodities, https://www.reuters.com/markets/commodities/malaysia-spend-335-million-flood-relief-2021-12-29/.