



GOVERNING ARTIFICIAL GENERAL INTELLIGENCE IN A NATION

A Common Approach to Managing Disruption and Displacement

John Rodman

March 2026





GOVERNING ARTIFICIAL GENERAL INTELLIGENCE IN A NATION

A Common Approach to Managing Disruption and Displacement

CONTENTS

FOREWARD	2
ACKNOWLEDGEMENTS	3
OVERVIEW	4
TECHNOLOGY AND SOCIAL CHANGE	6
AGI WILL BE DISRUPTIVE AND ADAPTATION MESSY	7
THE CROWDING OUT OF HUMAN COMPARATIVE ADVANTAGE	10
THE FRAMEWORK COMPONENTS FOR NATIONAL AGI GOVERNANCE	12
PUBLIC PERCEPTIONS OF SAFETY AND SECURITY WITH AGI	12
AGI DISRUPTIONS IN MARKETS AND OCCUPATIONS	16
DETERMINING THE LIKELIHOOD OF SUCCESSFUL AGI GOVERNANCE	22
COMBINING THE FRAMING INTO A MULTIDIMENSIONAL PLANNING TOOL	25
A COMPLICATED TRADE SPACE EVOLVES TO SOLVE	30
IMPLEMENTATION PLANNING SHORT GUIDE	37
CONTINGENCY PLANNING SHORT GUIDE	37
HIGHLIGHTED GOVERNANCE CHALLENGES FOR MANAGING AGI	38
THE LIMITS OF TRADITIONAL TECHNOLOGY PROTECTIONIST MEASURES	38
GOVERNMENT AND GOVERNANCE USUALLY LAG TECHNOLOGY CHANGE	40
SYSTEMIC RISKS IN DATA AND A COMPETITIVE AGI ENVIRONMENT	45
THE POTENTIAL FOR ERRONEOUS DATA AND MALIGNED INFORMATION	45
IS IT TOO LATE TO PREVENT A COMPETITIVE NATURE IN AGI?	49
CLOSING THOUGHTS ON CONSIDERATIONS FOR AGI GOVERNANCE	51
APPENDIX: IMPLEMENTATION AND CONTINGENCY PLANNING	52
ABOUT THE AUTHOR	54



FOREWARD

Artificial General Intelligence (AGI) is a transformative technology requiring novel, nation-state efforts to manage its effects on individual societies and global interactions. Unlike major industrial and technological changes in the past, adaptation to its potential for societal disruption may only have months, not years or decades. Planning for this technology is imperative to reduce profound economic and societal disruption. This is a paper for generalists interested or concerned about the future impact of AGI and planners of AGI governance for their nation. The issues and challenges of managing AGI's progress and integration are many and complex. Any approach needs to be structured and digestible so conversations can be held at all levels regarding AGI development and the role that it should have in the future. This paper does not presume these discussions are not currently happening but rather hopes that this approach furthers the discussion on standard approaches to governance that are transparent to the public, national planners, and technologists.

When AGI becomes part of our daily lives it will reach into the home, occupations, and societal fabric – forever changing safety, security, and stability. More than a tool for research, AGI may change how we engage with each other personally, locally, nationally, and internationally. This paper brings together disparate themes like digital feudalism, the danger of technology disruption, and the loss of human comparative advantage to illustrate the enormity of the issue and the need for comprehensive planning at the nation-state level in support of international rules.

AGI will be a major inflection point for humanity that will require an informed and involved public to safely leverage its potential. Its development and how a nation-state integrates AGI will affect the lives and livelihoods of its populace significantly. AGI is more than a tool. It challenges individual and collective futures. This guide is for nations developing governance and prioritizing efforts related to AGI but also is a way for the public to gain sufficient understanding of the hurdles so that they may follow progress and engage with their respective decision-makers on the issue.

This paper will explain why this technology is different than what humankind has adapted to before, what the key issues are that representational governments need to resolve, and possible approaches to creating appropriate governance of AI and AGI for a nation. It includes some examples from the author's home country but is meant to be generally applicable to any nation or international organization. This framework represents the author's opinion but hopefully will resonate and spur engagement at all levels on how to make the most of this technology.



ACKNOWLEDGEMENTS

The kernel of this model, or framework, was formed from discussions with Jerome Glenn of The Millenium Project during the early stages of the initiative that led to their recent, excellent publication, *International Governance Issues of the Transition from Artificial Intelligence to Artificial General Intelligence*. My thanks to him for those early conversations. My technical review of the Report from a policy perspective further reinforced my perspective on AGI governance and led to my expanding and completing this framework. From that effort, and discussions with experts more knowledgeable than me on AI issues, I became concerned that progress on managing AGI will be sporadic and ineffectual without a common methodological approach to governance development at the state-level. This document does not presume to represent their views and has not been reviewed by them.

During my decades of experience developing national strategies and planning processes around the world, I have been integral to the development of policy from numerous projects on data management and governance, as well as the challenges of integrating AI-related technology into national production and government operations. These engagements span decades and the experts are too numerous to acknowledge here. I give my thanks to Eric Bloedorn and Heather Mahardy for putting up with my questions over the last few years on a variety of AI-related issues. Material referenced in this paper also contributed to development of the framework.

I would be remiss to not acknowledge that, at times, I used various AI tools to clarify or distill concepts that I had already developed or to ensure I was considering any counterpoints to my position.

The concept for the model and text in this paper were created by the author, except for quotations or other annotated acknowledgements. The views expressed in this paper are those of the author and do not imply endorsement by individuals acknowledged or referenced.



OVERVIEW

International debates on the impact of Artificial General Intelligence (AGI) encompass government to academia, layman to technologist, and philosopher to scientist. They propose various solutions, from leaving it to the market and user to decide its development and integration to a range of government measures from draconian to light-handed. If the strength of these debates and positions were not reason enough, then the history of social upheaval from rapid and impactful technological innovations should tell us that planning and protocols for future AI-enhanced technologies are needed. This includes the coming age of AGI, which may be one of the most disruptive technologies to humanity in centuries. Since the harnessing of nonhuman power (water, steam, fossil fuels, and nuclear), civil societies and governments have struggled with the democratic governance of technology and its impact on societal issues. Unlike the rapid changes in the 18th to 20th Centuries, we have the means to monitor the global introduction of AGI, how it is impacting occupations and the economy, and whether efforts to govern its development and adoption are beneficial.

Managing AGI may seem overwhelming due to the pace of AGI development, the vast amounts of data involved, and the increasing extent to which AGI is affecting our lives. The scale, scope, and sophistication of AI usage will continue to increase, perhaps exponentially, and any regulation must account for this with a flexible assessment framework that supports decision-making and governance implementation. A plan for artificial general intelligence (AGI)¹ governance should not overregulate some sectors or underprovide the right support to other sectors in a nation. For regulation to not appear as an infringement on personal rights, or ill-conceived and implemented, the policies should reflect what the public wants, including their level of concern, with respect to AGI integration. Occupations may functionally change as AI/AGI integration expands with some traditional jobs disappearing and new jobs or vocations being added. There may also be a significant gap between new jobs and jobs lost, as well as increases of underemployment and misalignment of workforce location and skills. Regulation by specific jobs or occupations would require huge amounts of data acquisition and analysis by the government to constantly reevaluate the workforce makeup as AGI eliminates, restructures, and gives birth to new occupations. A flexible assessment

¹ Artificial general intelligence is defined for the purposes of this paper and policy development as artificial intelligence with the ability to understand, learn, and apply knowledge across occupations or tasks in a manner that is comparable, or replicates, human abilities and intelligence.



tool and shareable approach to managing the current and emerging effects of AGI on public safety and security, the impact on occupations and the economy is critical. However, it should also be realistic about the likelihood of managing AGI. Governance is about managing impact. Applied by government institutions, it is fallible and the rules and regulations can be challenging to implement. Nations and international organizations must find ways to temper these characteristics so that provinces, regions, and states within a nation, or nations within an international body, can find ways to adapt the protocols for their respective situations.

The factors that may affect AI regulation are often highly technical, many-sided, fast moving, and subject to interpretation. Nations enacting policies to manage the disruptive effects of AGI should consider the following factors that are important to develop appropriate and sound governance. As they relate to AGI specifically:



Cost/Benefit: If the cost of AI and AGI continues to drop and drives down the cost of complex computations (DNA sequencing, drug simulations, encryption, etc.), how should a nation determine the limits to economics that override societal concerns regarding the risks?



Scale: AI is already in more sectors and businesses than can be accounted for, should there be a national or international organization keeping tabs on AI integration to assess and make recommendations for regulation to protect general economic and societal interests?



Scope: Where do you stop the oversight and regulation? How do we determine when and where in the AI development process to insert brakes on AI's progress and integration, and will a nation decide the refresh rate on governance or policy to maintain responsible oversight?



Timing: Technology adoption by societal sectors and the lag time for regulation on new technologies may differ greatly, and the various participants can withstand differing levels of regulation. So how can you segment control with damaging businesses and individuals?



TECHNOLOGY AND SOCIAL CHANGE²

There is strong debate in many nations regarding how disruptive AGI might be in their respective countries, which can strongly influence decisions to manage its development and integration. Paradoxically, the anticipation of sweeping changes and widespread disruption is paralyzing decisions about governance in some governing bodies when it should be spurring action. Additionally, there are many private entities developing AI related technology and moving towards AGI to gain market advantages and profits that want to see as little regulation as possible.

Adapting to technology and its impact on how work is done is a recurring issue, so a good place to start is the relationship between technology and disruption in general. Management of any revolutionary change in industry eventually occurs as a society adapts, experiences the disruptions and impacts, and should elicit early, programmatic governance to minimize the historical patterns of disruption associated with sweeping technological changes. Unfortunately, there is usually a significant lag in implementing appropriate government actions. This is an important point and foundational to why framing and processes for AGI governance need to happen now, because the rate of development for AI associated technologies is greatly reduced compared to technology and its impact historically. Large blocks of change, which can include the concept of technology waves, have embedded within them a cycle of adaptation which previously occurred over years or decades instead of months as with AI and AGI developments. Even with those longer timelines for change, there were still great social disruptions associated with mercantilist practices, industrial revolution, mass media revolution (printing press, radio, television, the internet), and information-based societies combined with automation in industry.

The shortened timelines for major occupational, societal, and technological shifts that may occur from broad integration of AGI complicate creating sound governance. In nations with democratic norms, technological adaptation is influenced by the many competing interests and narratives that may be exposed by commercial interests, open debate in the government, or freedom of expression (accurate and inaccurate) by the populace. The adaptation also is dependent on the freedom of movement for essential resources (capital, labor, ideas, etc.) and the capital investment market.

² This section on how democratic institutions impact adaptation to technological change draws on multiple documents in addition to the author's own analysis. Those works are identified separately in the bibliography under the heading for this section.



Democratic processes, like open debate, can create polarization, uneven adoption (perhaps exaggerating income disparity and concentration of wealth), and ad hoc or poorly conceived government intervention to bandage the national fragmentation. Despite this, encouraging this debate is essential to socializing the impact of AGI and identifying essential data, processes, and solutions to its disruptive effects on industry and workers. The marketplace and society adapt and move faster than government institutions, which means that during that lag between industry and society adopting AGI and responsible governance damaging impacts or harmful effects can occur. A major technology shift like AGI warrants rapid action and preparation to manage.ⁱ

AGI WILL BE DISRUPTIVE AND ADAPTATION MESSY

A first step is to understand and anticipate the coming AGI disruption so that appropriate governance can be developed to reduce the negative aspects of this technology on industry and societal cohesion. There are thousands of likely variables across nations to consider, so we are going to start very generally to illustrate the importance of having governance discussions now. This will be done by superimposing 400 years of democratic evolution over a trend line of the accumulation of wealth in the U.S.

The accumulation of wealth is a major driver in the development and integration of new technology into a nation's industry. It can also add to the social fragmentation by compounding the income disparity resulting from occupational disruption from the industrial and technological changes. Without government intervention, democracy,

AGI Governance Challenge

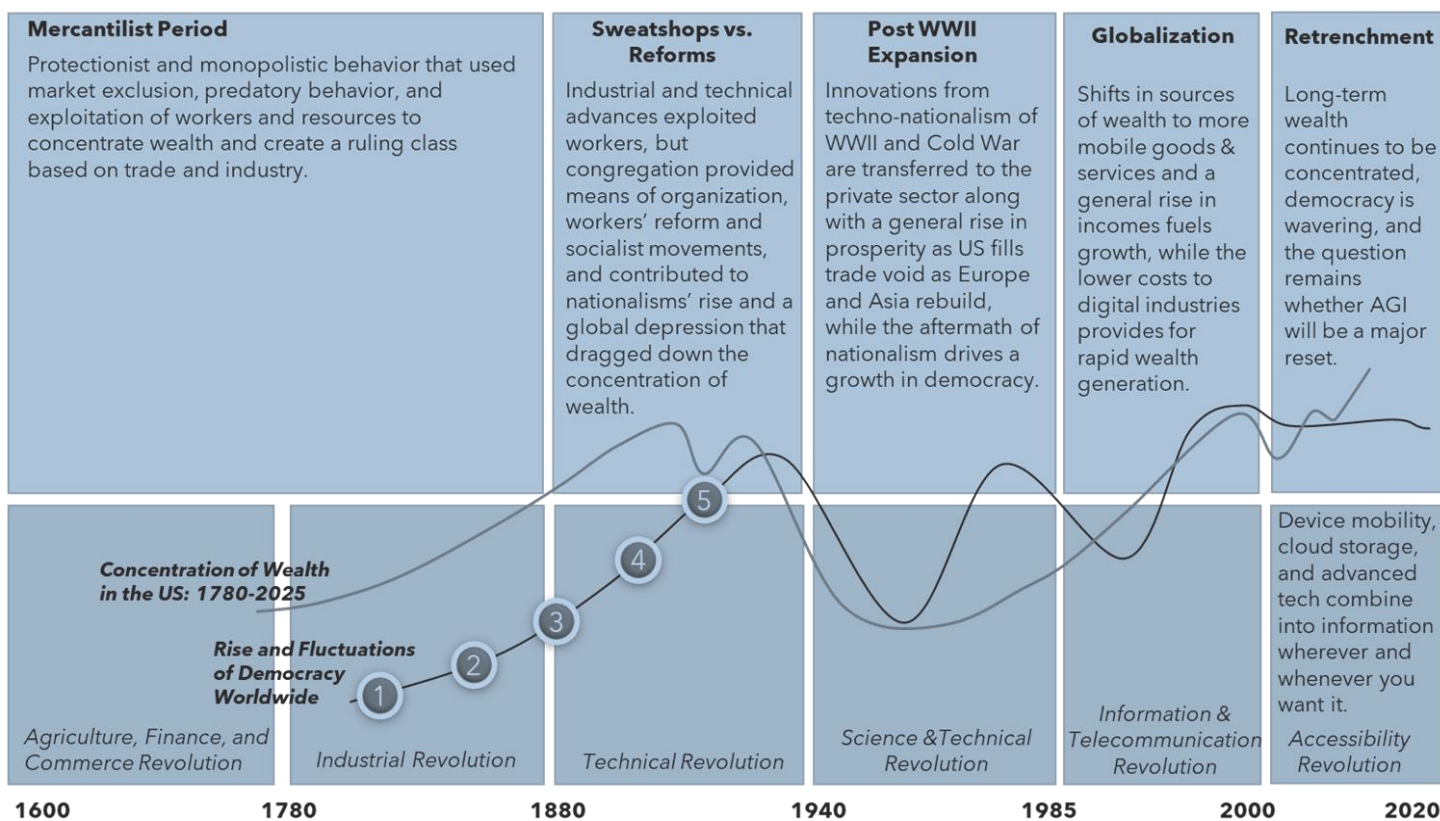
Governance will require a multi-dimensional, anticipatory process that covers labor markets and occupations, personal and national security, and more adaptive regulatory processes to increase the likelihood of mitigating the disruptive nature of technological change. It must be able to deal with the differing adaptation rates of different production sectors, disparate public positions on the dangers of AGI, lessen the lag time that normally occurs between innovation and appropriate government control. Unlike centralized control of AGI development, democratic governance of AGI often is formed around public debate, market forces, societal fragmentation and concerns, corporate/industrial suasion, and international competition. In those nations, this will take time and may add to the natural lag between disruption and beneficial governance processes.



capitalism, and accumulation of wealth may be reinforcing and lead to economic growth, innovation, and the expansion of the middle-to-upper class because they often provide predictable economic conditions and secure private property ownership.ⁱⁱ Democratic institutions, whether fledgling or fully formed, can foster open discourse about how to deal with technological changes that have a significant impact on a nation. Although sometime erratic, the U.S. has been a major influence on democratic movements around the world since its founding. That influence, whether direct or indirect, potentially became more significant as wealth in the U.S. increased and its economy as a share of the global economy. Global democracy, though increasing, oscillates around this upward trending wealth accumulation.

In the figure below, the line showing the path of democracy worldwide is annotated with the phases of how more democratic nations might adapt to major technological changes (phases 1-5 are detailed after the figure). The same phases, though condensed and not shown, likely exist within the subsequent upward sides of each

Long-term Wealth in the U.S., Oscillation of Democracy Worldwide, and the Phases of Adaptation



It is unclear how much AGI will disrupt society, but we do know that in the past major technology change has evoked distinct adaptation phases in nation states. Using democracy as a stand-in for measuring societal disruption in nations with representational governments, the waves are closer, shortened, and condensed with each technology wave. Suggesting AGI's impact will be more rapid and disruptive.



Since technology accessibility, representational governments have decreased compared to wealth concentration.



democracy curve. Both curves are generalized to fit account for four hundred years of change, but the upward trends of wealth accumulation and democracy align until recently.

Each upward part of the democracy curve is assumed to have five distinctive phases.

1. Early on reactions from various sectors that may be fragmented due to competing narratives, uneven adoption, and politicalization of the issue.
2. Perceptions of sufficient, perceived opportunities for innovators, researchers, producers, and early adopters. They became mercantilists, industrialists, and the middle class progressing through the 17th and 20th centuries.
3. Special interest groups, non-profit organizations, and politically motivated activists that raised their respective issues and advocate for their agendas, including a class of wealthy and/or influential individuals.
4. Government policy and planning functions that eventually take notice of the changes and disruption but struggle to implement beneficial regulations and keep pace with technological and social changes.
5. New laws, norms, economic and occupational structures, and oversight institutions established to balance and correct the disruption of the industrial and technological changes, sometimes too heavy-handed at first.

In general, it can be inferred that despite significant changes in commerce, industry, and technology over the centuries, representational governments with democratic processes and the aggregation of wealth have benefited from the adaptation of society to industrial and technological change. The Retrenchment period introduced by the author is our launching point for the AGI governance framework. The world is currently seeing a wavering of the trend toward democracy, democratic institutions, and a questioning of globalization that has contributed to the concentration of wealth but subverted national sovereignty. At the same time, the monetization of data over the previous decades has continued to concentrate wealth, at least in the United States. These two trends may be pulling apart as we are on the verge of AI associated technologies changing how a nation achieves and maintains economic security. This divergence coupled with AGI's seismic technology change is why governance is being debated and struggled with around the world. Contributing to the conflict are vested interests from influential parties and the risk of government oversight inhibiting innovation or disadvantaging the nation in the AGI race.



- Capital and wealth preservation are now often bound to the ability to control data and how it is used, governance of AGI could slow or chip away at large profits tied to achieving regular development and breakthrough capabilities in AI.
- If the processes of public discourse and transparent governance are beginning to erode, then those with an interest in unregulated AGI may be more persuasive than constituents within a less democratic government.
- The conflicting hype surrounding AGI says it will be the savior of humanity, or the downfall, or something in between. For a policy maker, this can cause a type of paralysis of whether to be restrictive or not, and whether to enact regulation and adapt it as technology changes or wait it out until the hype and integration settles down.

THE CROWDING OUT OF HUMAN COMPARATIVE ADVANTAGE

Added to this decision-making challenge is the speed at which technology is changing and being adopted, plus the potential that this is a game changer for many nations with respect to how much their workforce is displaced and whether those workers' skills are still necessary human functions. Governments, and the societies that they support, are facing the possibility of the loss of human comparative advantages in many, if not most, occupations, especially when robotics are supplemented with AI-related capabilities. The decision to create a national governance process is not just about controlling AGI. It is also about designing a response process that is agile enough to handle high levels of unemployment, underemployment, and widespread dissatisfaction that leads to increased social instability or fracturing within the nation. As illustrated in the following diagram, to develop effective AGI governance nations need to consider the marginal – or

A 1960s Visionary

In the early 1960's, Dr. Herbert Simon put forth the theory that humans will be "crowded into" tasks that require our biological evolution-equipped cognition, while machines will be crowded into tasks with well-defined structures. It was a vision, sixty-five years ago, of an economy that relied on machine-aided production and management that played to the strengths of humans and machines. He saw humans retaining a comparative advantage if they could retain a "flexible general-purpose problem-solving capability". Without directly addressing the morality or philosophy of his prediction, this paper prepares national planners for cognitive crowding out and the crowding in of humans to a narrower set of functions.



additional - benefit of human employment in various sectors. We will look at this from the perspective of effective, fair, and transparent governance that accounts for occupational displacement and safety or security.

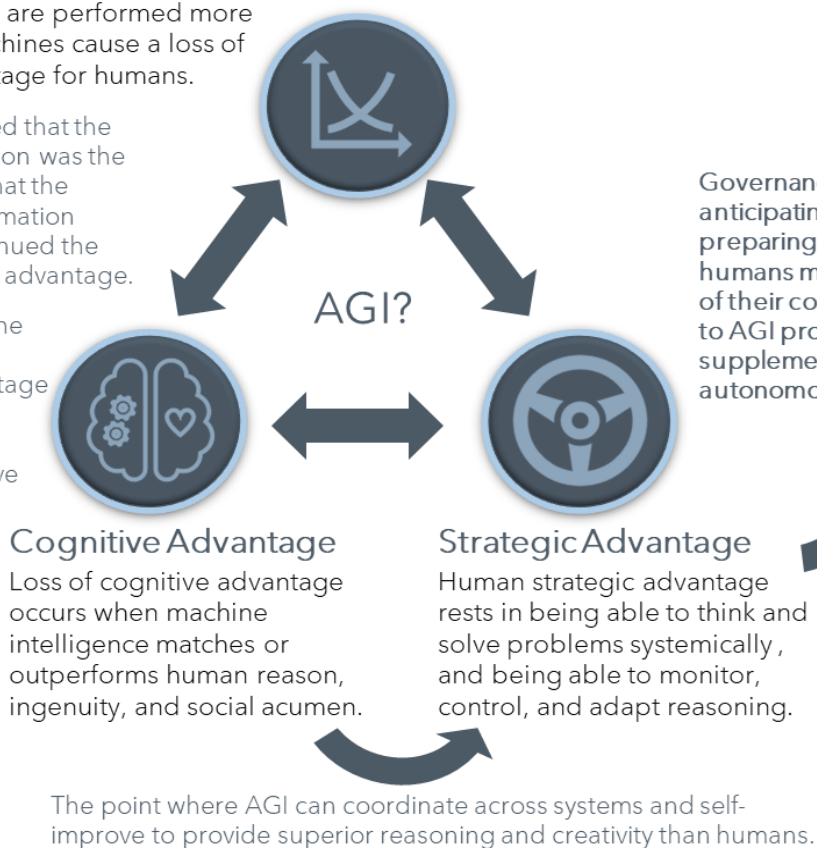
Preparing for a Loss of Human Comparative Advantage

Economic Advantage

Occupations that are performed more efficiently by machines cause a loss of economic advantage for humans.

It could be argued that the industrial revolution was the beginning and that the technology/information revolutions continued the loss of economic advantage.

If so, then does the complete loss of economic advantage come when AGI supplements or replaces Cognitive Advantage?



Governance is essential to anticipating, monitoring, and preparing for the possibility that humans may lose some, or all, of their comparative advantage to AGI productive output, AI supplemented labor, and autonomous mobile entities.

It is critical to establish government roles in alleviating occupational disruptions and public concerns for safety and security associated with AGI. Establishing functions and boundaries of AGI governance can determine the likelihood of success for governance across a spectrum of circumstances and nations. In any planning process, identifying measures of effectiveness and success are essential to right-sizing the regulation of industries and technologies but may be difficult due to uneven AGI integration. The likelihood of being able to manage AGI's development and integration can be an enabler or obstacle to technology transition depending on the accuracy of its actions. The governance likelihood factor characterizes a government's capacity for informative, innovative, and adaptive regulation that keeps



up with the rate of change in AI and related technology or applications. The likelihood of success in managing AGI depends on mobilizing resources, regulating development and/or deployment, warning public and private users, and protecting the public and nation from adverse effects of AGI. It is probable that the effects of AGI will be rapid, widespread, and provide little time to adapt and manage without national planning. Key factors in minimizing disruption are captured in the figure below. Beyond managing the current environment, governance has a role in remediation of AI and AGI effects that have already occurred and trying to anticipate future needs and requirements by the public and within the workforce specifically.

Bounding National Governance Functions for Planning



Later, in the framework, the center portion designated as management functions will be elaborated on and are being called the Governance Frontier. It is being called a Frontier because the governance mechanisms must consider two AGI boundaries on each side of what is happening now; what has already happened vs what may happen in the future. The Governance Frontier links the AGI function of remediation to management and management to anticipation so that all the framing can consider past, present, and future impacts across its effected society.

THE FRAMEWORK COMPONENTS FOR NATIONAL AGI GOVERNANCE

PUBLIC PERCEPTIONS OF SAFETY AND SECURITY WITH AGI

Sound governance requires a thorough understanding of societal awareness and concern regarding how AGI may impact their safety and security. Often, when developing policy, it is helpful to think of the major themes or issues that must be addressed. With AGI, the influence and needs of the various groups will be different for each society based on how they view AI - individually, as groups, or nationally -



and the risk it presents. When confronted with a new, life-changing technology, individuals are torn between the logic based on the knowledge they have of the potential impact on their lives and the emotional concerns they are experiencing. Though national planning cannot account for every individual, it can attempt to aggregate different levels of knowledge and concern into groups or blocs that must be considered. In the framework, the impact on a population's safety and security are divided into four categories that provide context, personalize the large-scale impacts on society, and make more tangible the public perceptions of how AGI may impact their lives. The categories are named to characterize the differences and make them more memorable. The categories in each nation should be determined through data collection and analysis but are generalized here to aid the framework explanation.

- Discounted Danger: There are many reasons why some segments of society might discount the danger and impact of AGI. We are seeing this with an AI-driven market bubble that is focused on short-term results from integrating AI. Some individuals might believe their occupations are not at risk from AGI. Others might be technologists or developers enamored with perceived benefits of greater AGI integration into a nation's production and workforce.
- Informed Distrust: This may be another segment of society that is very familiar with the field of AI and is more measured about the potential, timeline, and value of AGI to society. Individuals in this segment may be educated professionals who believe the benefits will be more limited than espoused, or individuals that believe AGI will never replace essential humanistic qualities.
- Fearful Future: There are always some segments of a society that are fearful of fast-moving technology, especially when it sounds like something out of science fiction. Media has long characterized a future where human-like robots replace many social interactions and occupations. Unlike the Informed Distrust category, this could be a messy mix of people that fear change, feel like they are being left out, or have a distrust of big tech companies in general.
- Divergent Response: Some governance needs to mitigate the impacts of AGI on society that may include disparate occupations experiencing upheaval from gained efficiencies or loss from replacement. To manage this complex and emotional segment of society, some rules will need to be minimal in low-risk areas, more stringent in high-risk areas, and some form of hybrid along a "governance frontier" that is dealing with technology that is constantly changing.

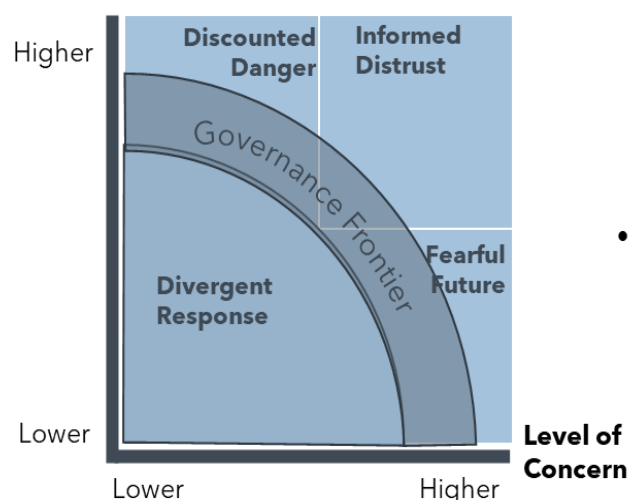


These categories broadly summarize what many societies are experiencing as AI becomes more integrated into their daily lives. AGI with its potential for displacing many human functions and occupations may accelerate and increase the concerns represented in the four categories. The figure below illustrates the four categories aligned against relative levels of knowledge and concern, which are used to frame the approach. Planning should account for each of the categories through national discourse and more formal methodologies to measure levels of knowledge and concern. Each category has its own unique challenges and will have significance based on data gathered. The Divergent Response region represents the difficulty of developing governance that does not seem overreaching for a group of people that may be uninformed or unaware of the impact AGI has had on their lives.

AGI Safety/Security Concerns

The Governance Frontier for safety and security must thread a mix of awareness and concern that may be difficult for nations to map and monitor over time. The influence and needs of the various groups will likely be different for each society based on how they view AI – individually, as groups, or nationally – and the risk it presents. Formalized methods of assessment that are continually updated may be necessary to adequately manage AGI, especially in the lower awareness – high concern area as the behavior and responses here may be more emotional and less able to predict or address.

AGI Awareness



- Discounted Danger: Individuals in this area may believe their occupations are not at risk from AGI. Others might be technologists or developers enamored with perceived benefits of greater AGI integration.
- Informed Distrust: Individuals in this segment may be educated professionals in disciplines who believe the benefits will be more limited than espoused, or individuals that believe AGI will never replace essential humanistic qualities.
- Fearful Future: There are always some segments of a society that are fearful of fast-moving technology, especially when it sounds like something out of science fiction. Unlike the Informed Distrust category, this could be a messy mix of people that fear change, feel like they are being left out, or have a distrust of big tech companies in general.
- Divergent Response: Individuals in this area have a mixed level of knowledge and concern regarding the safety of AI-related technology. Some will want minimal rules for perceived low-risk occupational areas. Some may already be experiencing displacement but are not aware that it is due to AI-related technology, which increases the need for an educational role in governance.



A planning effort collecting societal data might encounter concerns like the following.

Risk Model	Public Concerns	Policymaker Focus	Governance Role
Misuse	Scammers, deepfakes, fraud, personal harm; threat feels individual and immediate	National security, cyber offense & defense, biothreats, geopolitical competition; threat framed as strategic	Outer edge of a Governance Frontier to mitigate current threats and anticipate signs of misuse
Misalignment	Misalignment due to AI narratives; concern is intuitive but abstract	Misalignment of compliance, auditability, liability, and control; focuses on enforceable oversight	Inner edge of a Governance Frontier to set rules, provide oversight, and establish responsibilities. Governance links mitigate and remediate functions .
Accidents / Mistakes	Glitches, hallucinations, product failures, or personal inconvenience	Systemic failures in infrastructure, finance, healthcare, transportation; focuses on cascading risk	Outer edge of Governance Frontier focuses on anticipating systemic failure and mitigating damage from platform or product failure
Structural / Systemic Risks	Job loss, inequality, loss of meaning or status	Macroeconomic shocks, social unrest, institutional stress, and political destabilization	Inner portion of Governance Frontier to remediate occupational disruption and mitigate system shocks
Centralization & Governance	Corporate overreach, surveillance, lack of transparency	Regulatory authority, cross-border data issues, competition with other governments, sovereignty challenges	Middle portion of Governance Frontier to manage and regulate, establish boundaries and rules, and coordinated across domains
Existential Risk	Skeptical or divided; public may see as remote or unless amplified by media	Low-probability, high-impact national-risk category like nuclear or biosecurity tail risks	Outer edge of Governance Frontier and beyond to anticipate future risks and threats to supply chains, infrastructure, and critical economic sectors
Responsible Development	Reassurance, ethics, fairness, visible safety practices	Standards, audits, reporting requirements, liability frameworks, and partnerships	Mid-to-outer edge of the Governance Frontier to develop and anticipate and planning, policies, and regulations



AGI DISRUPTIONS IN MARKETS AND OCCUPATIONS

One of the primary debates about the coming age of widespread AGI integration is whether it will have an overall positive or negative impact on the workforce. For the purposes of the framework, that determination is less the focus than the inescapable disruption it will cause. Yes, some new jobs will be created, but many will be disrupted or subsumed by AGI. As with historical periods of new technology being introduced, there will be a period of disruption with the potential to cause alienation, economic strife, and dislocation for individuals. Such stressors are the fertile ground for political and social upheaval. Unlike in the past, laborers may not be the most impacted. Industrialization, machination, and automation changed agriculture, mass production, and consumer behaviors by introducing lower costs, consolidation of the workforce, and a redistribution of wealth that shaped our modern world. These technologies, often depicted as waves, redistributed of wealth as highlighted in the introduction.

The occupational disruption caused by AGI will likely be different. An important study by the Department of Treasury in 2024 looked at occupational exposure to AI in the U.S. by region and education. Several important findings emerged for our modern workforce; in high-density areas (metropolitan centers) the occupations exposed to AI are four times greater than other areas, individuals with bachelor's degrees or above will have much higher rates of AI exposure, and STEM majors will be a highly exposed area.ⁱⁱⁱ This aligns

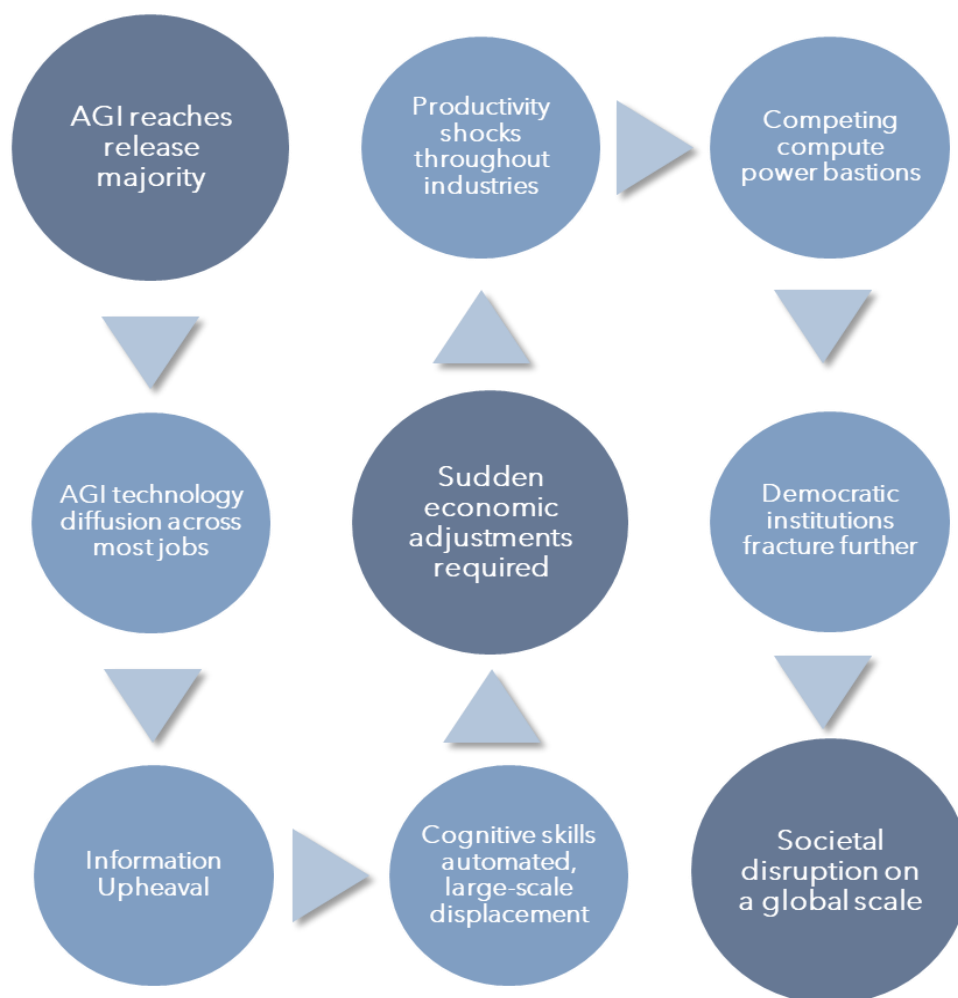
AI Workforce Disruption

There are more workforce disruptions studies for AI than AGI, since it is not widely deployed. Some takeaways:

- Goldman Sachs looked at 800 occupations and found the repetitive, rule-based, clerical tasks most at risk.
- The same study estimates AI could affect (disrupt in some way) a 300 million workforce globally.
- The Artificial Intelligence Policy Institute estimates that in the U.S. 35 million jobs could be at risk of being automated.
- Mustafa Suleyman, in an interview with Financial Times, stated that "professional-grade AGI" could replace most white-collar jobs.
- The impact of AGI on the U.S. workforce and economic stability is captured a strong methodological study by Satyadhar Joshi predicts that net job loss to AGI could be as high as 20% by 2040. (ISSN: 2583-7117, IJISEM, p. 345)



with earlier studies like the one from RAND Corporation that determined routine and data-intensive tasks are at risk whereas nonroutine cognitive tasks (teachers, nurses, technicians, etc.) may be less at risk.^{iv} AGI may have a much greater impact on those individuals that make up a significant portion of the upper middle class, are college educated, professionals in urban areas. Other studies have confirmed that unlike previous disruptive technologies that impacted more rural and routine work, AGI will have a greater impact on highly educated, high-paying, white-collar jobs in metro areas.^v The impact of rapid occupational disruption in this workforce might follow a path or something like the following due to their economic and political influence.



High-levels of rapid occupational disruption in a workforce of well-paid specialists may have a much larger, longer, and more negative impact on many countries' economies than previous technology disruptions. From industry to technology revolutions, those changes displaced workers, created misaligned labor, and incentivized labor migration. AGI will have two key differences; it will impact almost every job to some extent, and the changes may occur without sufficient time to avoid mass societal disruption.



The impact timeline suggests that the milestones for AGI's effects are differentiated and sequential. That is likely more representational than accurate, which rigorous risk and strategy methodologies implemented through good governance may help us discover and manage. The interdependencies between the factors, sectors, and workforce in a nation's economy may make this diagram look more like an influence diagram, which can be more like a bowl of spaghetti to look at. This timeline, in some ways, parallels the adjustment process nations go through in representational governments as shown by the steps along the global democracy curve in the paper's first figure. The milestones along the impact timeline will drive a government's response to rapid and disruptive technology changes. Because AGI will affect almost every industry and occupation to some extent in a very short period, this clear delineation and sequencing may not exist. The adjustment period and opportunity for governments to respond may contain all these concurrently and require much work to disentangle so that appropriate remediation, management, and anticipation steps can be taken.

So how can nations characterize and assess AGI's disruption to their workforce and develop sound policy to manage and hopefully minimize that impact? The framework focuses on two occupational attributes to represent contrasting traits of a future workforce confronted by widespread AGI adoption. On one hand, higher knowledge-based jobs may be in jeopardy, but higher ingenuity jobs may not. This ties back into our discussion on human cognitive advantage. Is the rate of acquired AGI traits rapid enough to "crowd out" workers from those innately human problem-solving positions or will the result be to "crowd in" workers to those few positions that require this ability. They may be the same problem, just different sides that governance will have to determine how to manage. This is an important juxtaposition for how policy may help alleviate or manage many societal issues from the integration of AGI like education to reeducation and workplace satisfaction to career changes.

- At one extreme are the highly knowledge-based jobs that may be less dependent on spontaneous creativity or intuitive solutions. They also may be less likely to engage with others for problem-solving or resolution of issues. The level of knowledge needed is used because many studies have shown that the coming AGI disruption will likely impact on our educated workforce.
- The other extreme is represented by occupations that require more ingenuity than formal education. Ingenuity is an ability to think creatively, develop innovative solutions, and often work with others to see them implemented. This could occur at personal, one-to-one level or within larger, organizational



problem-solving. These are skills that applied AI, and likely early AGI, will likely remain less capable than humans.

What constitutes the higher and lower knowledge or ingenuity-based jobs will be unique to a nation's economic, political, and social constructs. To determine the overall disruption, the entire spectrum of workforce disruption should be considered, from the minimally affected to the highly affected. The following descriptions are a starting point for risk discussions for the spectrum of knowledge and ingenuity jobs that may be disrupted.

- Knowledge-based Jobs: At one end of the spectrum are those tasks that may be lower in formal education but are not manual or trade occupations. On the other end of the spectrum are the jobs that take more specialized skills or formal education, including jobs in finance, law, engineering, and research.
- Ingenuity reliant Jobs: At one end of the spectrum are low ingenuity jobs that are routine, rules-based, or predictable. At the other end of the spectrum are the high ingenuity jobs that require creative, adaptive, and complex problem-solving skills that often involve engaging or partnering with others.

As previously noted in the Safety and Security section, the Governance Frontier is fluid with differing areas of emphasis for each nation, but it serves as a tool to frame priorities and resources needed to address past, current, and future occupational disruptions.

Ingenuity as a Determinant

Ingenuity is a critical determinant in understanding the evolutionary impact of AI on our societies. As used herein, it is that ability humans must have to create novel solutions, often from incomplete information requiring new rules or reframing an entire premise.

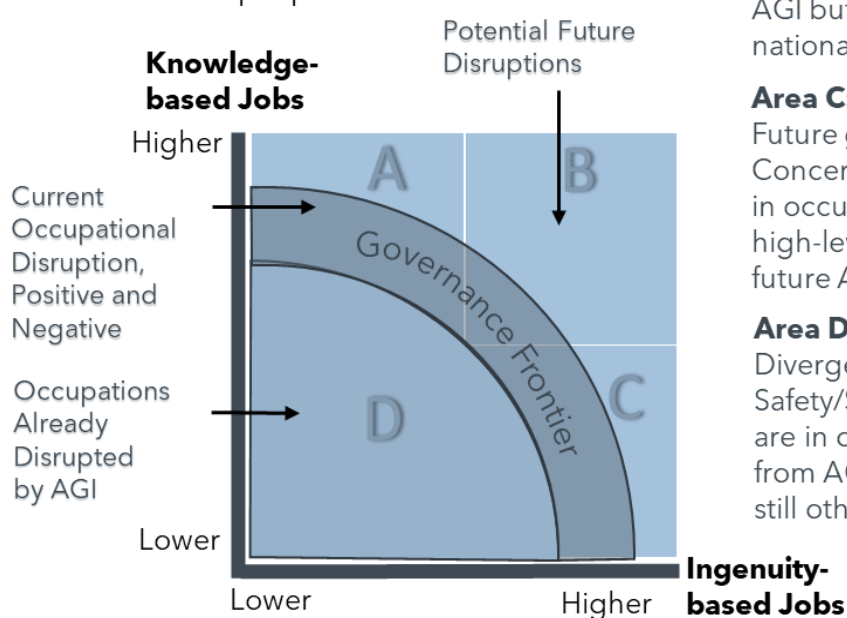
AGI promises to be as ingenious as very capable humans, but at machine speed. We are already seeing this in coding, program development, testing, etc. If true, then we are accounting for the human-side, and risk, of wide-spread AGI integration. The knowledge measurement is quantitative AGI skills and ingenuity a qualitative human skill.

There are fears that mature AGI will no longer just solve things faster but in ways that humans no longer understand as it transitions to Artificial Super Intelligence (ASI). Including ingenuity as foundational helps prepare us for the possibility of disruption from an AGI that redesigns that unique human trait.



Occupational Disruption

The Governance Frontier will be different for each society, wider in some, perhaps oriented more towards occupations that will be displaced, or focused on taking preventative measures for the higher knowledge and ingenuity-based occupations. The proper weighting can only be determined by incorporating formal means for assessing past, current, and future impacts of AGI on the nation and its people.



Area A: This roughly corresponds to the Discounted Danger group in the Safety/Security Concerns graphic. People in this area are in occupations that may have a high interaction with AGI, but they believe it will be beneficial or benign.

Area B: This corresponds to the Informed Distrust group in the Safety/Security Concerns graphic. People in this area are in occupations that may have a high interaction with AGI but are unlikely to have faith in national governance success.

Area C: This corresponds to the Fearful Future group in the Safety/Security Concerns graphic. People in this area are in occupations presumed safe due to the high-levels of human interaction, but future AGI capabilities are uncertain.

Area D: This corresponds to the Divergent Response group in the Safety/Security Concerns graphic. Some are in occupations with limited impact from AGI, others will be threatened, and still others will have to retrain.

This illustration of the effects on knowledge and ingenuity jobs also infers an important issue that has not been discussed, the time dimension of AGI integration and corresponding governance into occupations and the economy. While it was briefly touched on in the introduction – that the disruption may occur require adaptation in months versus years – from a planning perspective there are three periods of time that create specific governance functions. In addition to nation specific characteristics, the functions of remediation, management, and anticipation of AGI's impact will likely have common, foundational attributes that transcend national boundaries and are a basis for cooperation.

- **Remediation Function:** Already there are many occupational areas that are being disrupted by the introduction of AI-related technologies. While true AGI



may not be ready for adoption yet, it is coming. Identifying the steps necessary to remediate the displacement and disruption for AI-related technologies over the last decade will offer valuable insights to prepare for similar, perhaps greater disruption from AGI.

- Management Function: This is the present period where there are rapid changes occurring in the workforce and workplace. In the past, the workforce and workplace often retroactively created procedures and processes for safe and effective technology adoption. Similarly, consumer protections were often instituted because of negative events through lawsuits, recalls, and eventually legal protections. The faster rate of change with AI-related technology may accentuate this lag in mitigating the negative impacts with dire results. Governance agility, equity, and transparency will be essential, as will:
 - Transitions at the Inner Edge of the Governance Frontier: The inner edge of the Frontier bridges remediation policies and practices with current government efforts to manage the dynamic developments happening in AGI and leverage AGI's advantages. Monitoring the current industry and occupation effects will be challenging without a formalized process and responsible organization. Remedying previous occupational disruptions of those already affected by AI will have ramifications for AI and AGI management, requiring careful monitoring, planning, and implementation.
 - Transitions at the Outer Edge of the Governance Frontier: The outer edge of the Governance Frontier presents a different challenge to national planning. The plans, processes, requirements, and rules will play out in the future and potentially change trajectories of AGI development, integration, and adoption. The governance protocols put in place to manage the current AI-related challenges will need to also consider what are the second and third order effects, as well as, how those effects are related to the pace of AGI development and widespread adoption by consumers, industry, and in national institutions.
- Anticipation: Often the most difficult and necessary aspect of governance is anticipating what functions, regulations, and rules will be necessary to protect society and the workforce. Being able to identify future risks and threats those sectors cushions the overall impact on a national economy.



The historical lag for governance keeping up with technology is covered later in the paper. We know that the pace of technology will likely outpace the ability of most government institutions to draft, approve, and finalize policy. This does not consider the time it takes for such changes to be incorporated by government organizations, public and private institutions, and the commercial sector that sees governance as restricting the development and distribution of AGI capabilities. This lag may include legal formalities and obstacles enacted by industry or special interests to preserve any commercial or proprietary advantage in AGI development. However, industry and special interests' entities are also essential to developing supportive, responsive, responsible, and transparent governance through a collaborative approach that involves them in the entire process. Government-led consortiums can have success with managing disruptive technology.

- A good example of that is the Cybersecurity and Infrastructure Security Agency (CISA) in the U.S., which strives through multiple national and international collaborative efforts to identify risks and threats, harmonize cybersecurity standards, and bridge gaps between regional and private/public efforts.
- Similar initiatives on managing AI-related technology are underway in the European Union through action plans, strategies, and localized efforts to foster innovation, data-management and talent development.^{vi} This includes the European Union's AI Act, with a primary focus on high-risk AI systems posing risks to health, safety and fundamental rights of EU citizens.

DETERMINING THE LIKELIHOOD OF SUCCESSFUL AGI GOVERNANCE

Guardrails to how AI is used in future governance are developed and maintained through the thoughtful application of a framework. In the three categories introduced in this paper – Occupational Disruption, Security and Safety, and the Likelihood of Managing AGI – the innermost sector of each dealt to some extent with remediation of the effects of AGI. This region deals primarily with the data that characterizes the disruptions that have occurred, whereas the outmost portions of the framing are more future oriented and predictive.



A common framework is a first step to common understanding and solutions of the multidimensional impact of AGI in the future. It is important for each nation or group of nations to determine the appropriate level and type of governance as AGI becomes embedded in their economic, political, and social constructs. The first two components (occupational disruption and AGI safety/security concerns) embrace this by setting up a relationship between core issues on a spectrum (each axis) that differentiates AGI's impact on occupations and the economy and the concerns in various societal groups.

The third component in the framework is the likelihood of managing AGI. Managing is much different than controlling and is used deliberately to be realistic about what governments may be able to with respect to "right-sizing" their governance actions.

Success with the necessary appropriateness requires governance that assesses risks across many dimensions in society at each stage of AGI development and integration so that oversight can be scaled with AGI capabilities and applications. Connecting governance to risk prevents overreach and static rules for a dynamic technology while also disconnecting governance from political incentives. The table on the following page offers some notional measures of success.

Successful AGI Governance

This is a general definition for nations to use as they begin discussions on what is successful AGI governance means for their country specifically.

"A measure of success in AGI governance is when it maintains an adaptive balance between innovation, safety, economic competitiveness, social trust, and geopolitical stability – while preventing catastrophic misuse to ensure that AGI's integration into society is beneficial and aligned with national values."

The accompanying text in this section unpacks the many facets of this definition, but the key factor is that if accomplished, a nation will have a multi-dimensional, dynamic, and context-aligned process for managing AGI in their society.



POTENTIAL MEASURES OF SUCCESS FOR NATIONAL MANAGEMENT OF AGI

Economic competitiveness while ensuring safety	• <u>Success</u>: Domestic AI industry remains globally competitive with safety requirements that do not create excessive burden to high-risk R&D and provide support to lower risk innovation through government mechanisms – a new version of techno-nationalism. • <u>Failure</u>: Creating excessive or misaligned restrictions that might drive innovation abroad or underground. Conversely, such lax or outdated regulations that unsafe development and deployment occurs.
Public trust that fosters social stability	• <u>Success</u>: Oversight transparency with clear accountability for misuse to provide sufficient public trust that social disruption from AI and AGI development will be minimized. • <u>Failure</u>: A concentration of AI systems or control in a few leads to corporate or political misuse without sufficient mitigation processes to alleviate occupational shocks and widespread mistrust.
Institutions that adapt at the speed of technology	• <u>Success</u>: Rapid policy iterations by government bodies with sufficient expertise through strong coordination with industry and academia to ensure the ability audit and evaluate AGI models. • <u>Failure</u>: A fragmented government approach that contributes to an inability to enforce rules, creates contradictory regulations and oversight, and adds to the policy lag for technology.
International collaboration that reduces AGI-induced risk	• <u>Success</u>: Consistent AGI leadership communication resulting in agreements on compute tracking, model and safety standards, and actions for misuse to avoid a destabilizing AGI “arms race”. • <u>Failure</u>: Misaligned incentives and objectives between nation states that lead to a proliferation of dangerous capabilities and a zero-sum development environment without public transparency.

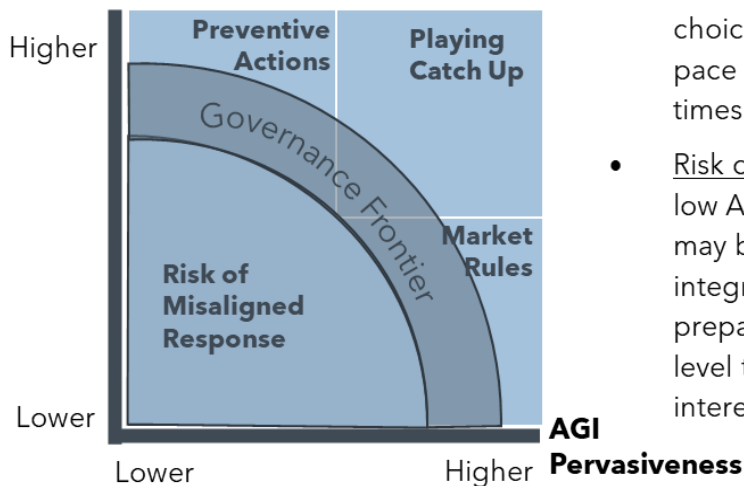


The framing postulates that responsive governance should be designed and scaled to AGI's development and risk, the framework pairs levels of regulation with how pervasive AGI has become throughout the various sectors of a nation on the two axes. Managing AGI can then be assessed by how well it keeps pace with monitoring and responding to AGI-enabled detrimental events and emerging threats.

Likelihood of Managing AGI

Managing AGI creates many challenges for policy and regulation. Success depends on a government's ability to regulate AGI compared to the extent to which AGI integration (pervasiveness) has impacted individuals and the means of production that fuel economic growth. At the lower end of the scale remediation may be necessary on occupational areas suffering loss. Above the Governance Frontier may require constant review to keep abreast of AGI impacts and to create mitigation strategies.

**AGI
Regulation**



- Preventive Actions: National governance policies in this area are leveraging lessons learned from sectors of the economy previously and currently affected by AGI to develop proactive management efforts.
- Playing Catch Up: National planners unfortunately find themselves constantly trying to keep up with technology changes to create beneficial governance policy.
- Market Rules: In this part of the management framework, the market rules outcomes because AGI regulation is almost nonexistent, whether by government choice, the inability to create policy at the pace of developments, or continual lag times.
- Risk of Misaligned Response: This area has low AGI pervasiveness and regulation. This may be the start of widespread AGI integration that exposes a lack of preparedness, an inability at the national level to design governance, or commercial interests that prohibit policy making.

COMBINING THE FRAMING INTO A MULTIDIMENSIONAL PLANNING TOOL

The components in the framework have been presented without a discussion of how they may interact with each other. How AI-related technologies manifest in the components of Occupational and Economic Disruption, Safety and Security



Concerns, and the Likelihood of Managing AI will be happening simultaneously, and some aspects and people may be common to all three components. That is a very challenging planning space. Additionally, we should be thinking about how to merge the concepts and drivers of each component with those governance functions of Remediation, Management, and Anticipation. Managing AGI is one of the most important roles that any representative government may face in the next few years.

To start thinking about AI and the coming AGI multivariate and multidimensional world we need to expand on each of the previously presented components so that the axes now represent a more detailed high and low bookending of their respective issue. The components of that describe occupational disruption and safety/security concerns are drivers of policy design because together they will be decisional for a government trying to foster innovation in the field of AI

How well a nation can assess occupational disruption and the safety or security concerns within in their society will be important determinants of how likely that government is to manage the changes that will be brought about by advances in AI and AGI over time.

while also trying to reassure its populace that continued development and integration of AI and AGI can be accomplished safely and securely. For each of these, planners will need to determine the characteristics of the landscape for each of the drivers. In other words, how can we expand with more detail each of the drivers while also “unfolding” the two-dimensional graphs so that they can become planes, or landscapes, that allow planners to accurately define the drivers of occupational disruption and AI safety/security concerns with detailed data and information?

For each of the drivers we can now think of the original axes as major themes that defined that issue but need further definition to be useful for developing research, data collection, and drafting policy for each driver. This expansion allows governments that take on this difficult task to develop more detailed understanding of the occupational disruption and safety/security concerns over time so that appropriate governance can be established. Redefining the two drivers on the next page keeps the concepts and their importance while creating the beginning detail for decision-making on governance. This is the first step for determining the type of data needed and the governance boundaries.

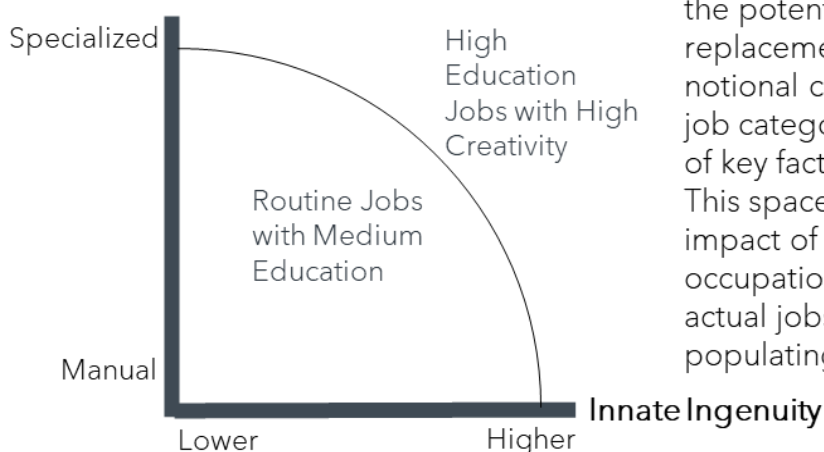


Drivers: Public Safety or Security vs. Occupational Displacement

Characterizing decision drivers is essential for building appropriate regulation. Simplified here are the drivers in two dimensions and key factors that may impact decision-making with more defining terms for the axes to help frame the type of data that needs to be collected and decisions on governance.

Occupational Disruption Space

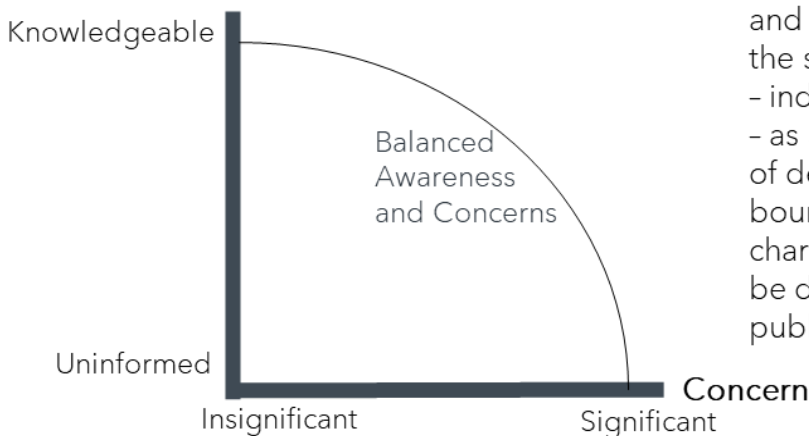
Formal Education



The driver in the graph represents the potential displacement or replacement of individuals with a notional center area illustrating job categories with some balance of key factors shown on the axes. This space represents the societal impact of AI on types of occupations and helps identify actual jobs at risk when populating this space with data.

AI Safety/Security Concerns Space

Awareness



This driver is public AI awareness and the concerns created regarding the safety or security of relying on AI – individually, as groups, or nationally – as a stand-in for risk. In the process of determining governance boundaries for this driver, the characteristics of this space would be determined by data collection or public engagement, or both.

Landscapes, in strategic planning, have a particular and valuable role. They help planners “map” the internal and external issues so they can create a decision space (terrain) that accounts for the relatively rougher and easier paths and help in the determination of realistic actions. The previous Occupational Disruption Space



becomes a landscape with by “opening it up”, as described in the following sidebar. This helps with governance design and transparency by better defining occupational categories at risk. The broad groups identified in the original Space for each component are still used for each pie area that now has more specific boundaries. Those groups of workers represent the impact of AGI on past, present, and future jobs bound by the axes for each pie. The density and location of jobs in each pie identifies occupations that need protecting and where AGI may be helpful in filling labor and resource gaps.

The same process is used for the previous Spaces developed for AI Safety and Security Space and the Likelihood of Managing AGI. The terms AI and AGI are not being used interchangeably in the model. For many people outside of the AI development field that are concerned about their jobs or safety and security, the difference is not as important or well understood. The unfolding of these two components creates two new landscapes with new, more encompassing delimiters. Once again, the planners and practitioners should populate each space with examples that fit into the breakdown of each pie quadrant using the previously defined groups for each of the original decision spaces of each component. This keeps placement of examples consistent with respect to the time and functions that governance can address. It also creates an internal consistency in the analysis and populating of each pie quadrant despite each being bound by slightly different axes.

Space to Landscape

How did we go from the previous representation of an Occupational Disruption Space to this Landscape? Imagine opening the previous space by unfolding to the left and downward so that the two axes become four quadrants separated by the two defining drivers, Formal Education and Innate Ingenuity. The landscape can now accommodate more detailed information that differentiates better the effected groups.

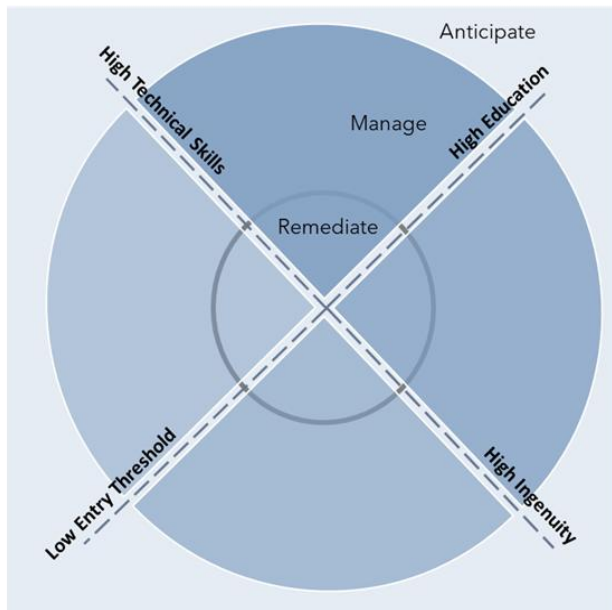
For this framework, the landscape rotates the quad to break the standard practice of thinking above and below the line and right or left of some arbitrary center point. Also, our real goal is to think of the spectrum of occupational traits in all directions, from corner point to corner point, and along the occupational requirements on each side of the pie quads. Inside the pie shapes are AGI disrupted job combinations in a nation’s economy. Some countries may be more heavily weighted in some areas than others, but that too is important to understand for developing appropriate governance.



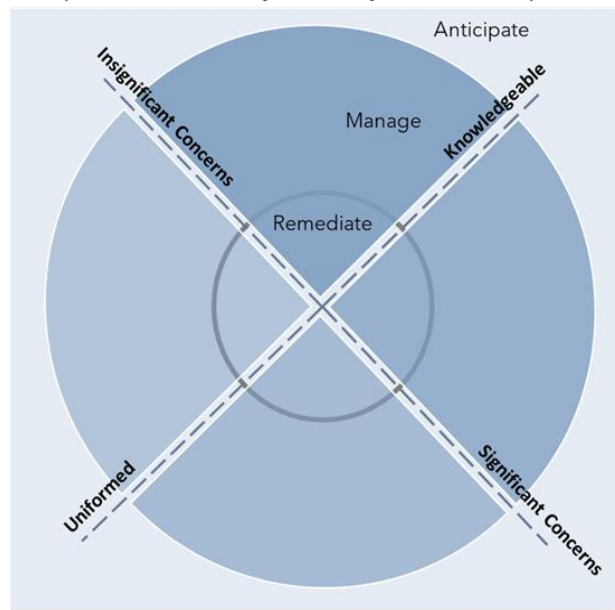
Expanding Each Decision Space into a Planning Landscape by Unfolding

Each quadrant in each new landscape uses the assessing criteria from their original space. For instance, in each AI Safety/Security quadrant what groups in society fit into the following criteria: Discounted Danger, Informed Distrust, Fearful Future, and Divergent Response?

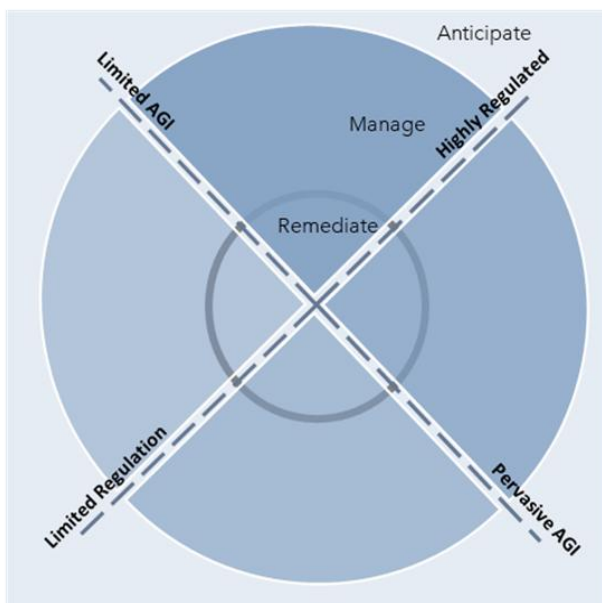
Expanded Occupational Disruption Space



Expanded AI Safety/Security Concerns Space

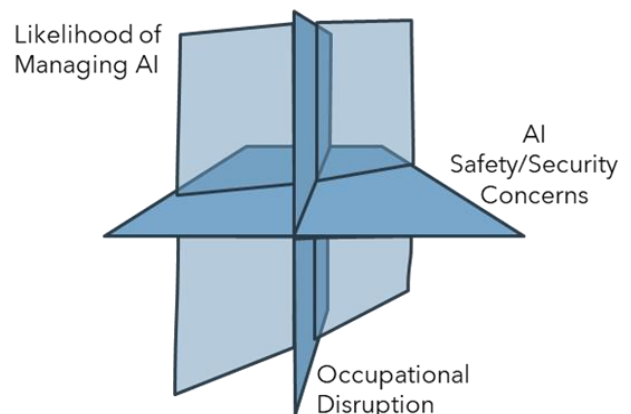


Expanded Likelihood of Managing AGI Space



The populated landscapes are constantly interacting with the Likelihood of Managing AI associated technology added to the primary issues of occupational disruption and safety and security. It gets difficult to plan for, so the next step is to simplify it into strategy themes.

Interactive Governance Landscapes





A COMPLICATED TRADE SPACE EVOLVES TO SOLVE

At this point, data collection and analysis has identified primary factors contributing to AGI's potential impact on a nation's occupations, categorized them, and separated them into quadrants so applicable and appropriate governance can be developed. We did the same for safety and security concerns. We then thought realistically about where governments might be able to have impact and manage AI and AGI developments and disruptions. Using the pivoted, four-square categorization landscapes to expand the original component spaces ensures that the original issues are considered and replicated across all quadrants with less bias.

There is one final, missing piece to developing an applicable and appropriate governance strategy and structure for AI and AGI for a nation. The goal of the dissection of the problem and creating framing tools was to make a seemingly unimaginably large and complex problem digestible by breaking it into pieces that can be assessed. Governance at the national level always involves resource limitations and trade-offs. To make those as palatable as possible, there is a need to prioritize national governance efforts. That concept, first introduced early in the paper, recognizes that representative governments make decisions with limitations on what they can accomplish, afford, and have been accepted by public and private institutions, commercial enterprises, and the populace. To help identify what can realistically be

Landscape Based Governance

The original Spaces developed for Occupational Disruption and AI Safety and Security Concerns provide a common, repeatable way to assess potential impacts. The unfolding process that was just covered allows us to create delimiters, or bookends, to determine the data needed to populate each component's Landscape.

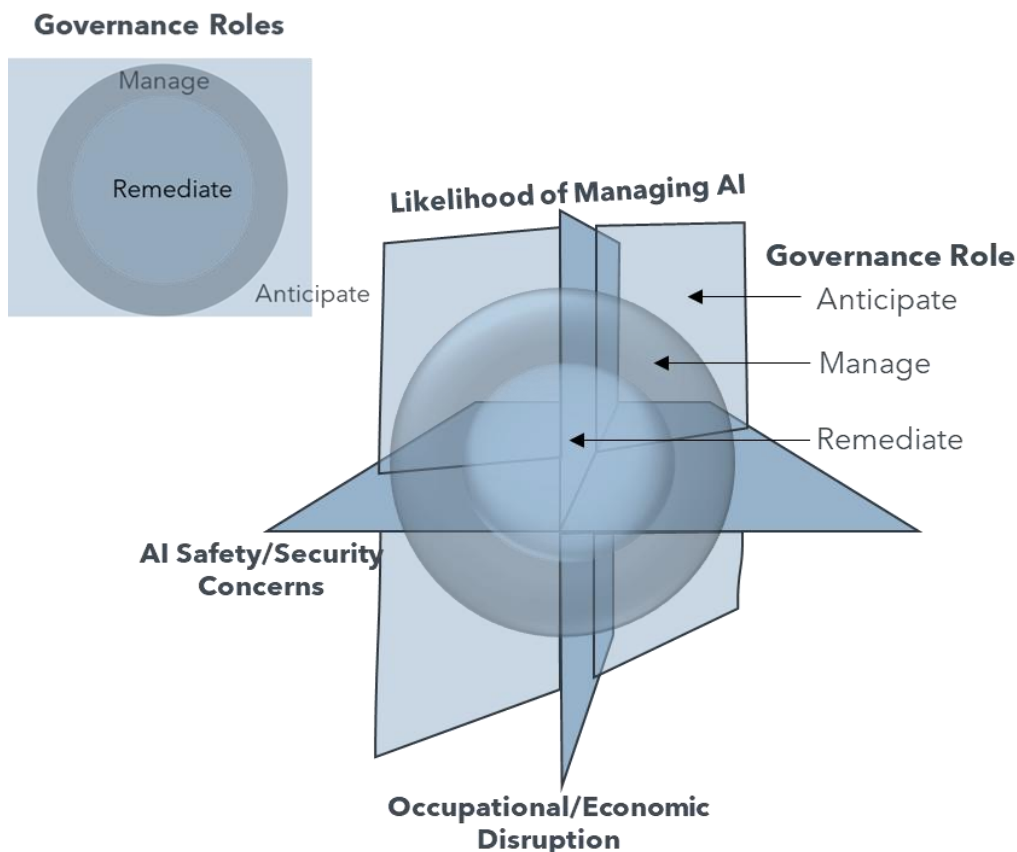
The original Spaces provide context for discriminating AI and AGI outcomes in each component. The Landscapes allow governance development and implementation to be responsive across the entire range of a society's needs. The Likelihood of Managing AI Landscape takes the government roles from that component's originally developed Space and compares those across the cross-referenced Landscapes of the other two components. We are qualitatively assessing the risk from AI and AGI by combining potential threats to occupations and safety with the likelihood of mitigating the impact.



accomplished the three landscapes are merged and the governance roles of remediation, management, and anticipation are overlaid. This is used to help government planners prioritize by governance functions the occupational and safety/security risks previously identified and used to populate each quadrant in the landscapes.

The prioritization and response for each government function (Remediate, Manage, Anticipate) results from a final organization of the risks in each quadrant. Recap:

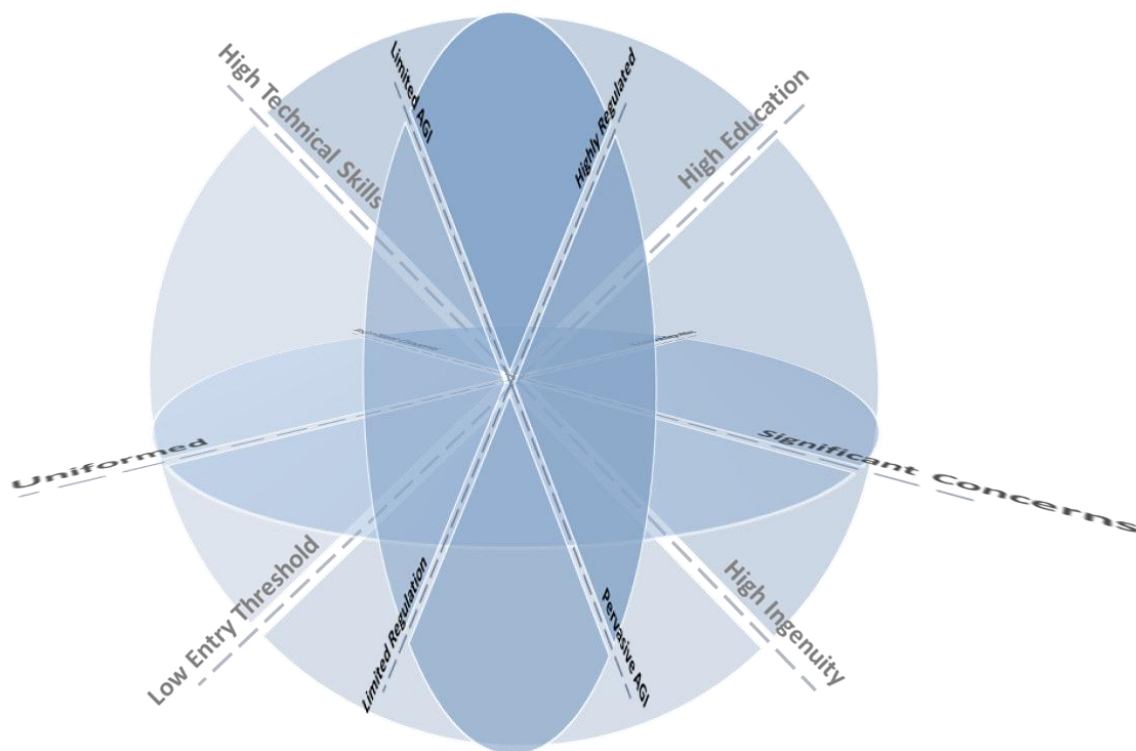
1. Expanded (unfolded) the original Spaces for the components to account for a wider, more detailed range of impacts from AI and AGI. These represent a nation's risks across as broad a spectrum that can be identified.
2. We then had three Landscapes, two representing risks and one the likelihood of managing those risks, which we integrated and populated through data collection, research, and debate. The latter cutting through the needs of the occupation and economic landscape and the safety and security landscape.
3. Prioritization and resource management was then introduced by rearranging the contents of each quadrants by whether they require remediation (past), management (current), or anticipation (future) policy actions.





One reason for the lack of progress in AI and AGI governance design or implementation is that it is such a complex, all-encompassing, messy issue to manage at the national and international level. The inner dimensions of this model, when combined, begin to get confusing. A short description of each possible environment that will require remediation, management, or anticipatory planning might be helpful.

- Each regulation area should consider the possibility of a governance structure for the future that includes the extreme outer risks and solutions, which eventually can be applied towards contingency planning.
- The governance region includes the more nuanced middle ground where compromises are being made on jobs to protect, levels of perceived safety and security, and realistic assessments of what the government can do in that slice or segment of the sphere in the accompanying graphic.
- Finally, a description of the innermost part of sphere for each slice or segment will need to be oriented towards remediating entrenched public positions, jobs already displaced by AI, and corrective measures that can be put in place quickly by the government.



A suggested approach to simplifying this complex governance problem follows with the caveat that each nation employing this or any other framework for dealing with AI



associated occupational displacement and public concerns may combine these characteristics differently based on the quantitative and qualitative data that is collected and the planning techniques used. Nations will have unique AI/AGI areas to protect and to expand even if a common framework like the one provided is used.

Simplifying the multiple dimensions created by our components requires pulling back from some details so planners can begin thinking about strategic orientation, objectives, and implementation paths. All the detailed work so far was to help planners and the public break down the complexity so they could identify, isolate, and prioritize the impacts of AI and AGI on their country. If the steps covered so far are undertaken there should be enough detail to begin building the governance documents and protocols for AGI. The populating of the pie quads and depicting their interaction was to develop AGI impacted areas based on assessment. One of the difficult parts of building national plans is determining where to focus and facing the reality of restrictions. Compromises will have to be made because a government must work with limited resources on a myriad of issues. That is the application stage and next planning step. It is time to start thinking about how the issue of trade space in each of the components will be handled.

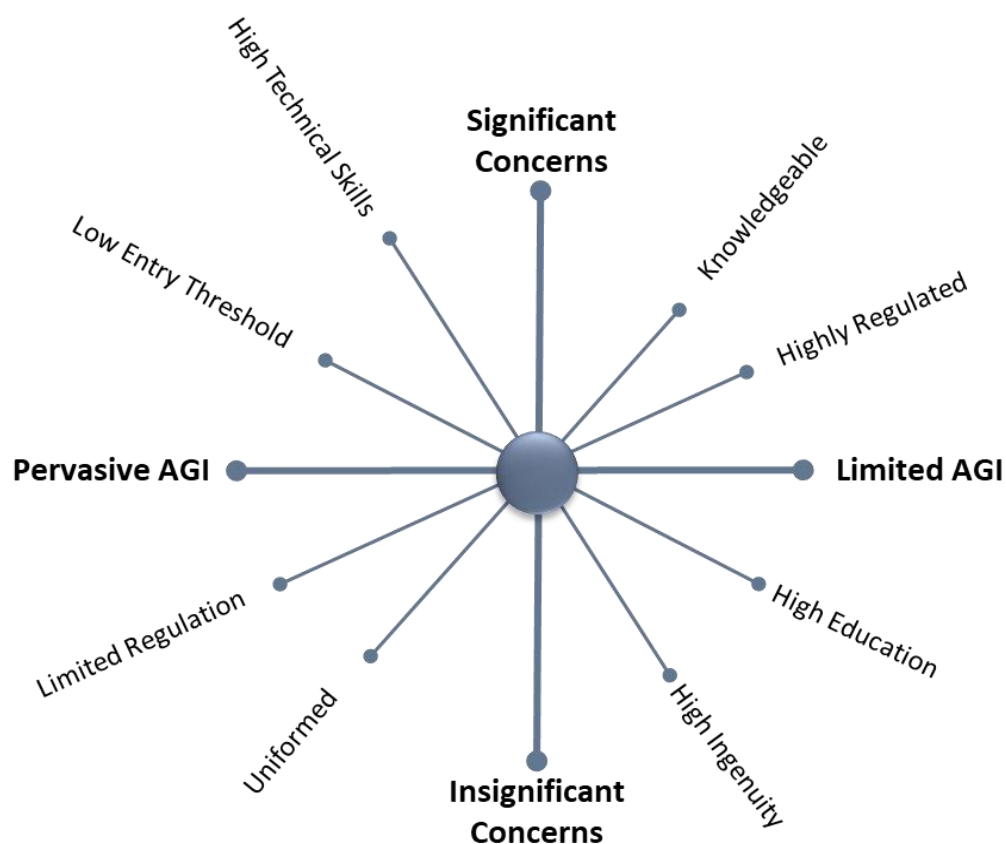
- The pervasiveness of AI and AGI means that technology will permeate most occupations and compromises in who and what to protect in the labor market will have to be made. It might be worth a revisit to the cognitive comparative advantage section to discuss what occupational areas you want to protect.
- In the safety and security component, our work with the model suggests that there will be a very wide spectrum of responses based on the public's knowledge, biases, and that even those may change as AI and AGI become more pervasive and the risk to personal safety and security changes. This may require a regular polling or survey process by a reputable academic institution.
- Lastly, beyond the challenging decisions about how or what a government should try and control there is the real competition for government resources in defense, infrastructure, health, and security. Somewhat separate from the AGI strategy, implementation, and contingency plans, this is an internal policy and resource discussion within each respective government.

To simplify the multi-dimensional space on the previous page we can leave the detailed, populated sections behind and focus on what the axes are trying to show us. Each line has an endpoint, which is characterized as an extreme position on that issue. Some are clear opposites, i.e., high education requirements versus low entry thresholds, and some characterize the difference in issues, i.e., jobs that rely on high



technical skills versus jobs that are more dependent on ingenuity and problem solving. Establishing trade space on multiple issues is challenging but can be done. Reimagining the world of impacts and positions contained in the three-dimensional sphere requires thinking about the function of each line, its relationship to other issues, and how to reconcile actions across several of them at once. The challenges for a country using this process were identified and mapped into a three-dimensional landscape and combining those was necessary to fully understand that the issues and problems unique to a nation are constantly interacting with many other factors that will determine governance.

To build accurate and appropriate governance policies and procedures, the three-dimensional world of AGI impacts needs to be simplified so that broader themes can be addressed and policy written for them. Policy developed separately for each of the three component landscapes may be disconnected and inflexible. Due to its detailed look at each component, it also might require more regular review and amending than is possible with limited government resources. Inflexible governance will be problematic for managing a dynamic, pervasive technology like AGI. Elevating the issues may provide a way to focus governmental actions.

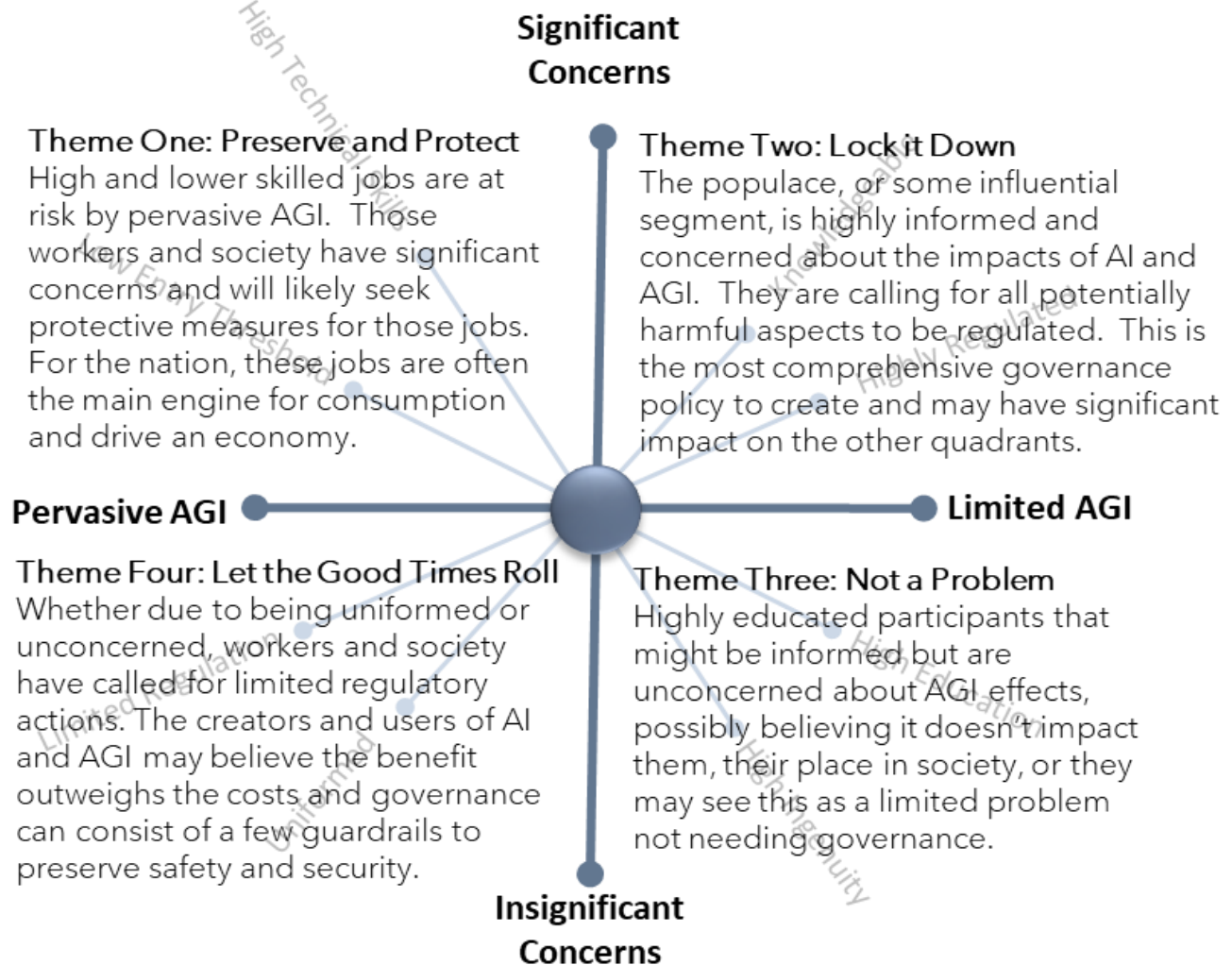




The two-dimensional depiction includes a rearranging of the axes for each of the landscapes. This is a notional arrangement, and each nation may want to highlight different counterbalancing axes. In this model, the axes dealing with societal concerns and the pervasiveness of AGI have been highlighted as being fundamental to a country deciding the scale and scope of oversight on AGI. The axes that are between these are not bold because in the model they define and provide characteristics for that quadrant and contribute to the populating of that space. The primary axes define decision space drivers and the axes between help define the impact of AGI.

Creating a national governance structure for AI and AGI is a balancing act between demands for greater restrictions to protect occupations and security and fewer restrictions to prevent undue government oversight and a loss of innovation. The benefit of pulling back from the detailed work done in the first part of the paper is we can identify groups within a population that need supportive governance and issues that need, or will need, addressing. Although AI and AGI may automate many thinking functions, this paper promotes extensive debate on what should be the boundaries of AGI. Humans relate better to stories and visualization, so it can be helpful to characterize the quadrants containing the descriptive axes and bounded by the primary drivers in a way that is more story-like. They can be characterized in a way that helps the planner focus and message their policy decision of the objectives of a national AGI plan in a way that the public can absorb.

In the graphic on the next page, the elements of the segments that were pushed to the background are captured in thematic descriptions of the types of governance that might be required so that they are less technical and can be messaged to the public. The button in center represents a policy position where everyone accepts equal compromise. This position is never reached! National planners will have to determine which of the quadrants, representing different Themes, to prioritize and in what sequence. Here they are shown as having equal weight, but it is likely that once they are populated through data analysis and/or polling or survey they will not be equal. That is good. The work previously done to identify the types of occupations, groups of people, and positions on managing AI all contribute to the final size of each quadrant and help the governance designers prioritize resources.



Once a nation has identified a manageable number of themes upon which to center a strategy, an end-state and major objectives are easier to establish. Although the ones provided above are notional, they illustrate how going from detailed analysis of the drivers and the decisional spaces they create to broader, national planning level themes can occur. This process was also important to determine the scale of affected people and types of issues in Occupational and Economic Disruption, Safety and Security, and Likelihood of Managing Components for the nation undertaking this type of process. The detailed work that occurred prior to creating the three-dimensional trade space and then simplification is valuable for later in the planning process when a country is developing implementation and contingency planning for AGI and must decide which areas to focus on due to limited resources and sequential steps for success.



IMPLEMENTATION PLANNING SHORT GUIDE

The initial, detailed breakdown of the AI and AGI landscape that might require governance and the subsequent thematic map help set a nation's orientation with respect governing AI and AGI. To truly minimize the potential disruptions an Implementation Plan (I-Plan) will be needed for the remediation and management functions of governance. What an I-Plan covers and how it is implemented is dependent on the preceding steps, so a generalized plan that any nation can use is not possible. In fact, the broader the scope of governance the more general and less applicable an I-Plan may be. Implementation is essential for success but must be developed post a scanning of the problem, deciding on the governance orientation based the analysis and building out of the framework, and completing a national AGI strategy. Although highly specific to each nation, some basic attributes of an Implementation Plan are provided in the accompanying Appendix.

Important Caveat: Though the I-Plan and CONPLAN are different stages of framing and planning for AI and the coming age of AGI, they must be connected through bridging language, guidance, and requirements to successfully fulfill a nation's response to AGI disruptions.

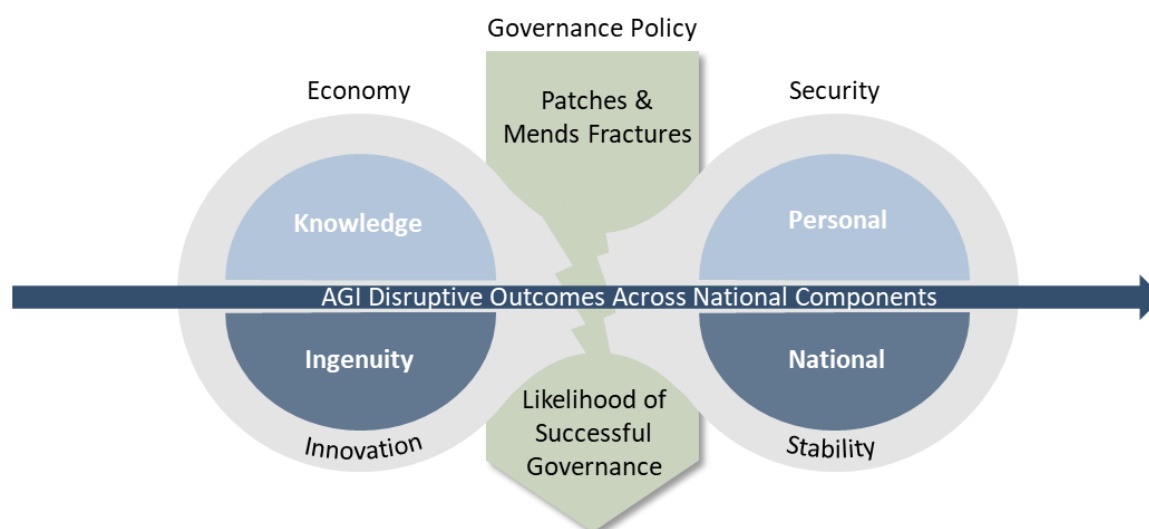
CONTINGENCY PLANNING SHORT GUIDE

The above recommendations and requirements for an I-Plan also apply to a Contingency Plan, sometimes referred to as a CONPLAN by government and defense planners. An AGI CONPLAN will be unique to a nation's requirements for governing AI and AGI. What is very different about a CONPLAN from the I-Plan is that it often focuses on what could be and what should be done if the Strategy and I-Plan are not successful. It is an insurance plan that can fill in unknowns, gaps in future capacity, and contingency actions should the direction of AI and AGI development, integration, and impact unfold differently than anticipated. In our combined component landscapes, it is primarily the area beyond the management ring that addresses the anticipation function of governance. It is dependent on the previous steps of scanning, strategy development, and implementation planning. Scenarios are often used to "play out" the unknowns and gaps so that future capacity building can be built into the CONPLAN. Like the I-Plan, it is highly specific to a nation, but some basics are provided in the Appendix. A methodologist in foresight processes can be very helpful in isolating actions for the contingency plan.



HIGHLIGHTED GOVERNANCE CHALLENGES FOR MANAGING AGI

The third component of national-level AGI governance planning introduced in the framework was the likelihood of managing AGI at the national level. Determining likelihood requires introducing risk, which is related to the ability of a government to anticipate a threat and respond accordingly. Whether direct governance or coordinated governance with the responsible entities, some key considerations for a government undertaking national AGI planning might include the historic challenges of governing technology, the lag time for governing technology, and the possible effects of zero-sum gain AGI. Another way to look at the role of governance is as a “patch” for inherent market imperfections.



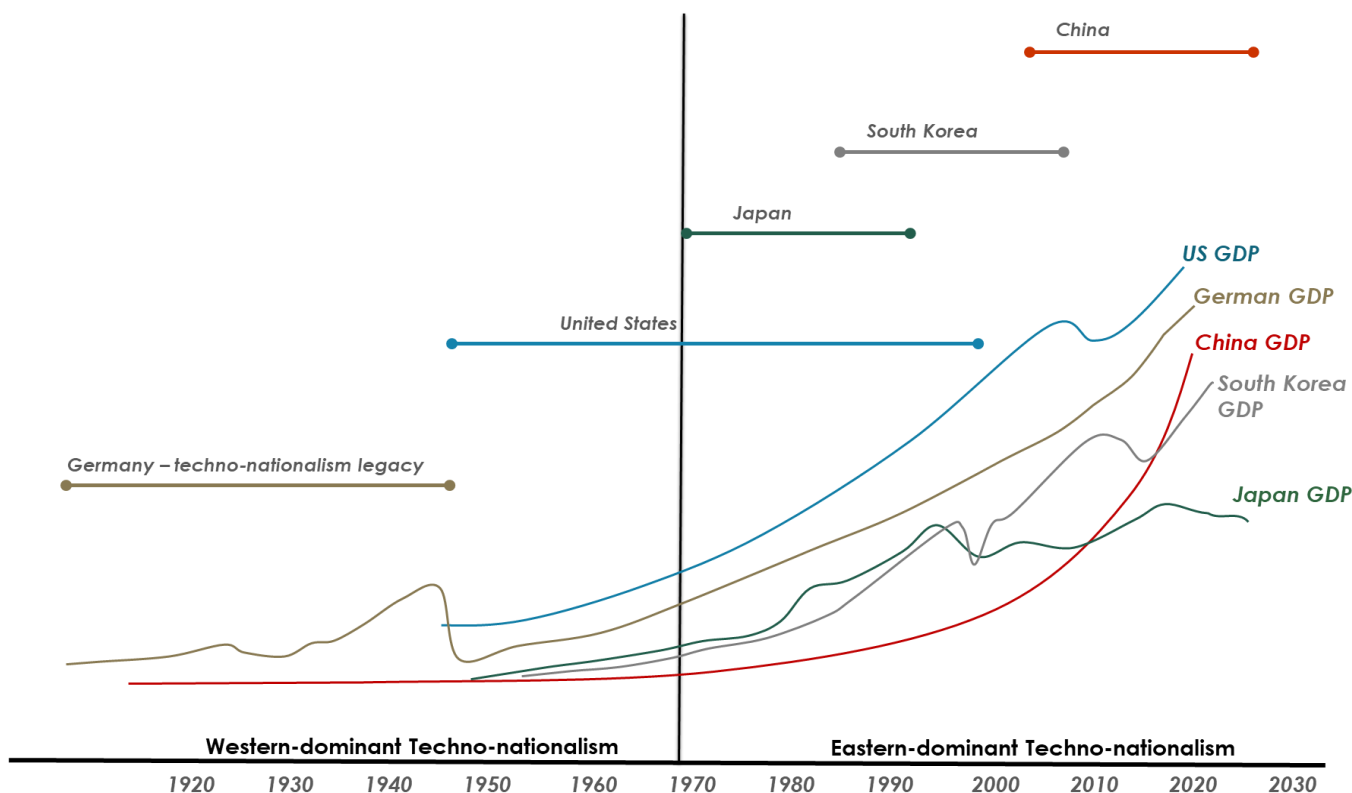
Likelihood of governance success in creating stability and capitalizing on AGI to promote an innovative economy.

THE LIMITS OF TRADITIONAL TECHNOLOGY PROTECTIONIST MEASURES

Traditionally, techno-nationalism used technology to advance nationalist agendas, which could include trying to create a stronger national identity.^{vii} Techno-nationalism is rooted in the concept that the strength of a nation is based on its ability to innovate and harness technology.^{viii} In the diagram above, government policy in the middle of the diagram is designed to foster innovation through protectionist practices that provide sufficient personal and national security for a new technology by protecting it from outward competition and market pressures. For AI or AGI, it might include focusing on incentives and government investment in development, or



protectionist policies, or both. It was, and is, how many nations seek advantage in industry and technology over the last century.



Techno-nationalism can be government incentives and protectionist measures like the United States' AI Action Plan and China's state funding of AI-related research, supercomputing, and quantum technologies. The European Union, Japan, and South Korea are among other nations practicing AGI techno-nationalism. Techno-nationalism can have lasting effects on the manufacturing and income base of a country if it can provide protection and subsidies to fledgling technology and build national production capacity.

Since many nations have practiced different forms of techno-nationalism over

AGI Techno-nationalism Gaps

Techno-nationalism may be used to develop, improve, and dominate physical and tangible technology through R&D investment, protecting industrial supply chains and infrastructure, and controlling critical material components. The uncertainty around the pace and scale of AGI progress and increasingly ubiquitous inclusion in our lives may make more traditional management approaches impractical. Techno-nationalism may work for the physical aspects of AGI, but may not be appropriate for managing a disruptive, innovative technology will work for less tangible parts like algorithms and data and open development platforms.



the last one hundred years, it is natural to turn to using it again to build up and protect national AI associated technology. This can provide a false reassurance to national planners that are thinking of AGI as a traditional technology instead of a transformative technology impacting almost every aspect of their economy, society, and security. Transformative technologies present important challenges to the effectiveness of techno-nationalism as an AI or AGI governance policy.

- The cost advantages associated with globalization and democratization of data reduced costs while creating more dependent up-stream aspects to most industries that reduced the benefits of techno-nationalism. Policy reactions often represent an excessive swing in the opposite direction, which allows some nations to justify even stronger competitive and protectionist behaviors. In the long run this may limit their future competitiveness in AGI development and benefit from its integration or even drive their own innovators to seek other markets.
- Techno-nationalism focusing on technologies for military or industrial uses often benefited from R&D funding, protectionist measures, restricting foreign investment, and setting industry standards or objectives in the past. That was before globalization. The average level of international components in consumer and manufacturing products has risen to.... And the level of cross-discipline patents has also risen...
- Previous techno-nationalism provided the breathing space for new firms and technology to mature and gain a market share due to high costs associated with building complex components or advanced technology. The cost structures around software, algorithms, compute capacity, and talent are very different for AI associated development.

Unlike past technology, AGI will likely change how individuals and society function. It may change our decisions on what constitutes national security and even be considered by some an existential threat like nuclear proliferation. Whether it assists or replaces national and social determination may depend on informed governance, especially if AGI becomes preemptive, paranoid, or focused on zero-sum gains - something we are already beginning to see.

GOVERNMENT AND GOVERNANCE USUALLY LAG TECHNOLOGY CHANGE

The potential for occupational and societal disruptions is high, and AGI may have an equal chance of displacing as many, or more, workers than it adds. Any attempts to manage its development with incomplete or poorly constructed national policy could



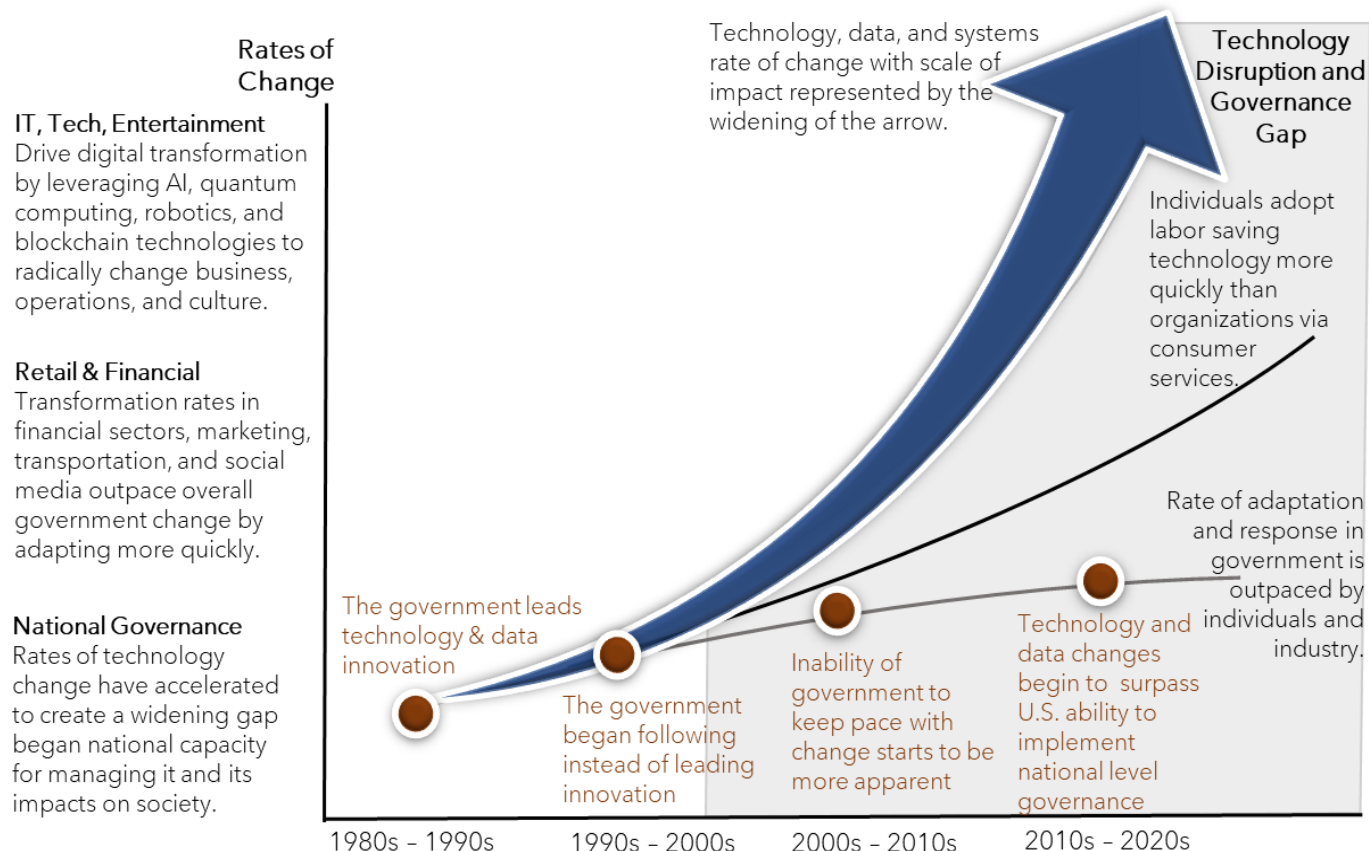
force development underground, offshore, or provide advantages in other countries. This may create considerable uncertainty at a national level regarding how to safely manage AGI's integration into society. That uncertainty can lead to an instinctive response towards techno-nationalist methods as previously described. If so, there may be some unexpected domestic results that include higher competition that results in rapid development and deployment without proper safeguards because government investment and protection incentivize entrants into AGI development. The ripple effect of this new, more competitive market could be premature deployment of platforms or tools, multiple platforms with differing standards, monopolistic behavior by the companies with the deepest pockets, and an even more unpredictable environment for creating governance policy. Conversely, uncertainty can create decision paralysis that leaves AGI development to corporations with less predictable cultural, economic, and social destabilizing outcomes that will require government policies and regulatory efforts to catch up.

Beyond deciding what policies to enact for AGI, governments tend to lag emerging technologies for adoption and management at the institutional level, and the lag can be compounded when trying to enact national management. This makes governments reactionary, sometimes too heavy-handed, and going with the regulation they know as provided in the techno-nationalism discussion. The gap between technology and governance may be even greater for digital advances. Leading experts and corporate leaders have expressed concerns that there is a history of an inability to keep up with technology and its adverse effects on society. They point to history—remember our technology waves, adaptation, and wealth concentration at the beginning of the paper—and that it is not about the technology system itself but the human and institutional interactions that lead to responsible use.^{ix} Useful AGI governance in any country, or internationally, will need to be able to assess those interactions as they progress and apply appropriately designed and scaled guidelines, regulations, rules, and standards.

As illustrated in the figure on the next page, in the United States technology began to really outpace the ability of government institutions to keep up individual adoption and digital-heavy sectors by the mid-to-late 2010s.



Technology Change, Adaptation, and Governance Over Time



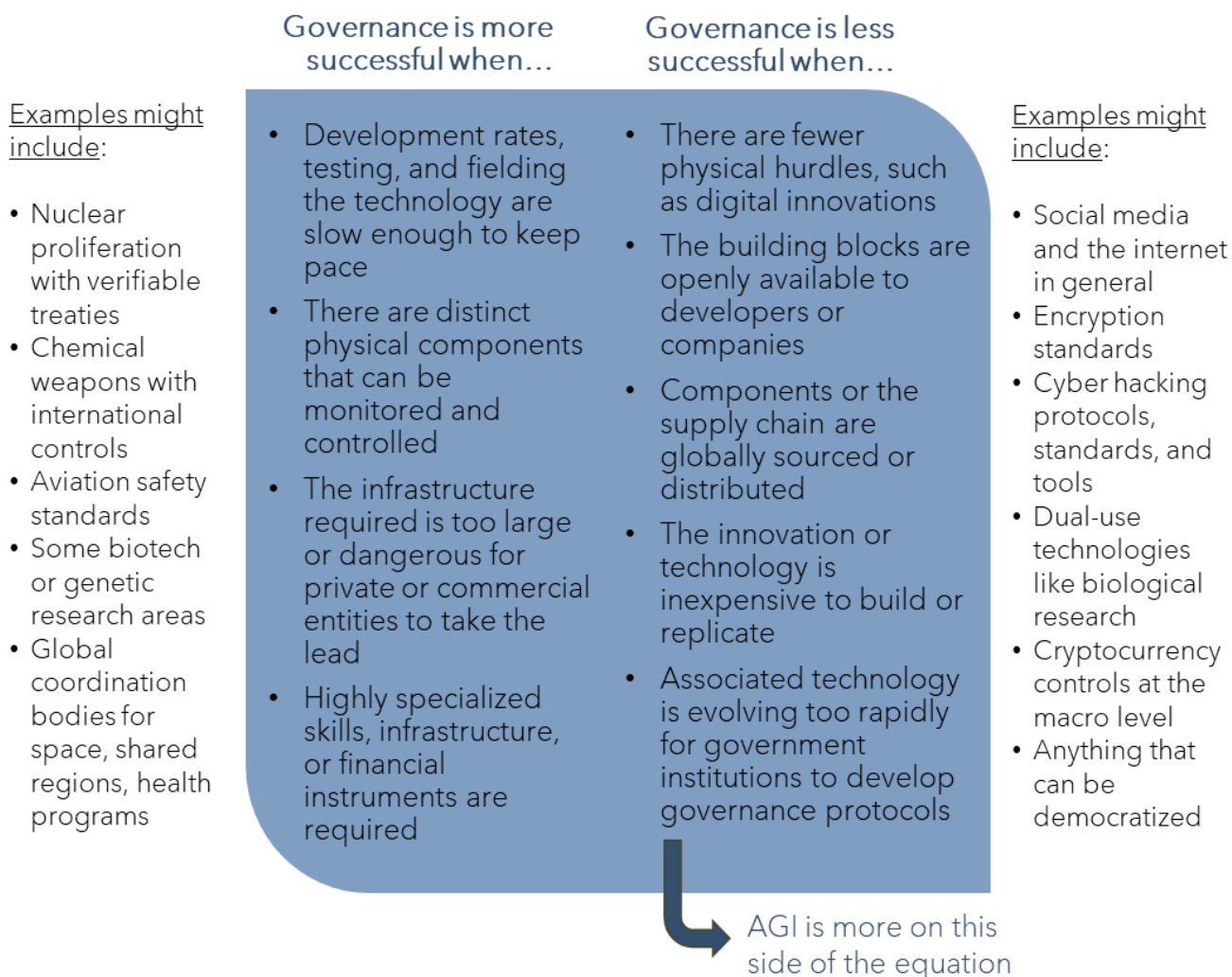
An important aspect of a comprehensive and structured framework is the prospect of closing this ever-widening gap between commercial AI development and responsible governance. Looking back at techno-nationalism provided some insights into what protectionist policies might be useful for AGI. A similar review of past individual government and international attempts to regulate an industry or technology might provide insights on how sufficient analysis, foresight, planning, and implementation might close the gap shown in the preceding illustration.

Good governance is adaptable and appropriate to the issue it seeks to address. Throughout this paper the struggle to create good governance is front and center. To be successful, some basic principles of governance with respect to new technologies should be developed. The following figure provides some examples that could be amended or added to as a nation works through developing approaches to AGI governance. Developing practical solutions is critical here



because successful governance of critical technologies that transcend nation-states is highly dependent on international cooperation that stays ahead of the most disruptive aspects.

National and International Governance Struggles with Technology



AGI is a dual-use technology that can be used for exceptional good or misused nefariously, which makes it potentially a national or international risk to stability from too rapid and poorly monitored development. The scale and scope of the risk may warrant rethinking traditional governance mechanisms. Internationally, some nations are responding to AGI traditionally, much like nuclear proliferation, with attempts to enact shared standards, best practices, and develop common threat pictures for its impact. Even when international organizations or coalitions can agree on these



measures there is much disparity in how individual countries then apply commonly recognized requirements. The European Union's Artificial Intelligence Act that was put in place in August 2024 has seen implementation of some portions delayed and others revised from pressure by corporations and the Trump administration for regulations perceived to slow innovation and investment in AI development.^x

A new form of governance requires a common framework with enough flexibility to allow for national advantages to develop with an understanding of the ubiquitous, borderless nature of information that AGI draws upon. Properly executed, such a framework could provide the means to identify critical factors or risks to a nation from AGI, best practices to manage its development and deployment, and drivers or limiters to its safe integration at the national and international levels. A framework should slow the deployment of unaligned or competitive AGI systems and open sourcing of uncontrollable models by fostering standards for responsible AGI models.

A framework should explain the capacity and limitations of governance to address:

- **Unaligned or Competitive AGI Systems** – It is unlikely that nations will agree to limit the progress or profitability of their respective AGI initiatives, but alignment might be possible around safety and reliability standards, use of AGI for common initiatives or threats (e.g. climate change, nuclear proliferation, etc.), and to establish protocols that reduce the risk of a destabilizing escalation of capability or use for military purposes.
- **Open-source Uncontrollable Models** – An open-source model allows for any person or organization to misuse the model and scale it up to have more computing power that can be fine-tuned with no oversight. This could lead to nefarious or illegitimate use of the model by individuals, criminal organizations, or nations seeking strategic advantage. It could also permit modifying the model and redistributing a dangerous version. A common framework could help nations develop best practices to monitor and perhaps contain malicious use of AGI by cybercriminals, extremist groups, or lone actors.
- **International Standards for Responsible AGI** – A primary benefit of a common framework for understanding the risks associated with unmanaged AGI development is to avoid thousands of AGI versions without any safety or security standards that create an information and decision ecosystem rife with unsafe versions, dangerous variants, or uncontrolled models that become embedded in critical systems.



SYSTEMIC RISKS IN DATA AND A COMPETITIVE AGI ENVIRONMENT

THE POTENTIAL FOR ERRONEOUS DATA AND MALIGNED INFORMATION

The integrated battlespace – air, land, sea, and space – presents incredible coordination challenges in modern warfare, but national security should also encompass areas like the economy, civil stability and welfare, and political awareness of all the relevant issues or risks to our national well-being. The data used by AGI to assess a situation, risk, or threat will be disparate and may include hundreds of thousands of separate sources of information. While some cultures and societies believe intentional manipulation of another state to be wrong, other nations viewed direct and indirect manipulation of data and information as an acceptable practice that corrects what they perceive to be an unfair advantage other nations may have in national power.

There are many reasons for integrating AI and AGI into the national decision-making process, some include:

- Identifying ways to overcome risks to a nation's critical infrastructure and supply chains from upstream, domestic capacity shortfalls, and weak international linkages.
- Determining potential economic and social stability threats and initiating policy or regulations for experts to review
- Integrating national defense and strategic deterrence capabilities and policies to integrate AGI more quickly, thus gaining advantage in conflict and in the decision process/reaction time

Upstream Governance Limits

This paper does not cover the parts and basic materials that are part of any AI system. Those are too numerous and covered well by experts in those areas. From a governance standpoint for AI and AGI the paper focuses on the application, integration, and impacts of AI and AGI on a society. Not including the physical components is not an oversight, but a recognition that governing the required parts and raw materials is an enormous problem that may be handled by other government organizations.

However, integrating AI into national planning without well thought out governance carries the risk of decisions made on data that is possibly inaccurate, irrelevant, or even inserted to shape decisions to the benefit of an adversary. Currently, AI systems do not have mechanisms to determine whether the mass of data that will be drawn



upon has been manipulated or designed to influence its use.^{xi} What if some portion of the information AGI is used to make assessments, recommendations, or even decisions is compromised? How will the AGI models determine if some portion, even a very small but critical portion, of the data is the result of adversarial data tainting or poisoning (DTP)? This creates several concerns regarding the validity of data that might be used in critical national decision-making.

- The extent to which the selection of data for modeling national security challenges is impure or altered.
- The extent to which the data and information discounted from inclusion in models was manipulated or tainted to “steer” us away from use.
- The deviation from what we should model or thought we would model and where we end up due to DTP.
- The how, when, or where in the model development and refinement process the adversaries’ may have attempted shaping efforts, especially if the actions are coordinated, comprehensive, and continuous.

AI greatly increases the ability to analyze and steer public discourses, which provides a multiplier effect for influence. This could be done intentionally by a malicious actor, through the “echo chamber” of public media and autonomous content creation. History has shown that some countries conduct extensive influence campaigns to gain advantage and shape public opinion to their benefit, in their own country or against others. Such influence may also occur unintentionally by the inherent responsiveness of AI to individuals’ preferences and biases that are increasingly supported by real-time AI analysis of their browsing, content preferences, as well as the unethical pushing of content commercially. All these forces may combine and contribute to societal fragmentation in unintentional and uncoordinated ways. Additionally, the increasing power and prevalence of AI commercially and in social media may be misused to directly manipulate political opinion, discredit political opponents, and erode the public’s faith in democratic processes.

The issue of data manipulation is highly relevant to developing a governance structure because it can happen at the most foundational level of a model, be subtle, and cause long-term problems in the accuracy of the model over time. An appropriate governance structure must anticipate ways to mitigate malicious data manipulation that is meant to intentionally alter results and ultimately assessments of decisions and unintentional, or structural, data manipulation that happens as an AI model evolves. In this section, the focus is on intentional or malicious data manipulation. Implanting false information into the decision-making process has



long been practiced by nations to delay an adversaries' decisions or shape their behavior in a way that supports that nation's favored outcome. While the adversary may be manipulating data and information, the real goal is to manipulate the targeted nation's decisions and actions. This uncertainty contributes to the general public's questions regarding the safety and security of AGI with respect to what data will be used, the inherent validity of the data, and the extent to which it is accurate for making important decisions associated with national security. Manipulation can be generalized in two ways.

- **Adversarial Data Manipulation to Reaffirm.** A nation openly shares its intentions, plans, policies, or technology research. The adversary understands the data likely needed and begins subtly shaping some and poisoning others so that the targeted nation stays on its stated path so that their progress is predictable.
- **Adversarial Data Manipulation to Shape.** A nation openly shares its intentions, plans, policies, or technology research. The adversary understands the data that is likely needed and begins subtly shaping some and poisoning others so that the targeted nation shifts to a path that benefits the adversary's goals.

Whether data manipulation occurs at the front-end of a model where data is collected or at the back end where the model engages with the user, AI is greatly speeding up the process and making identifying and targeting specific users or user groups much easier.

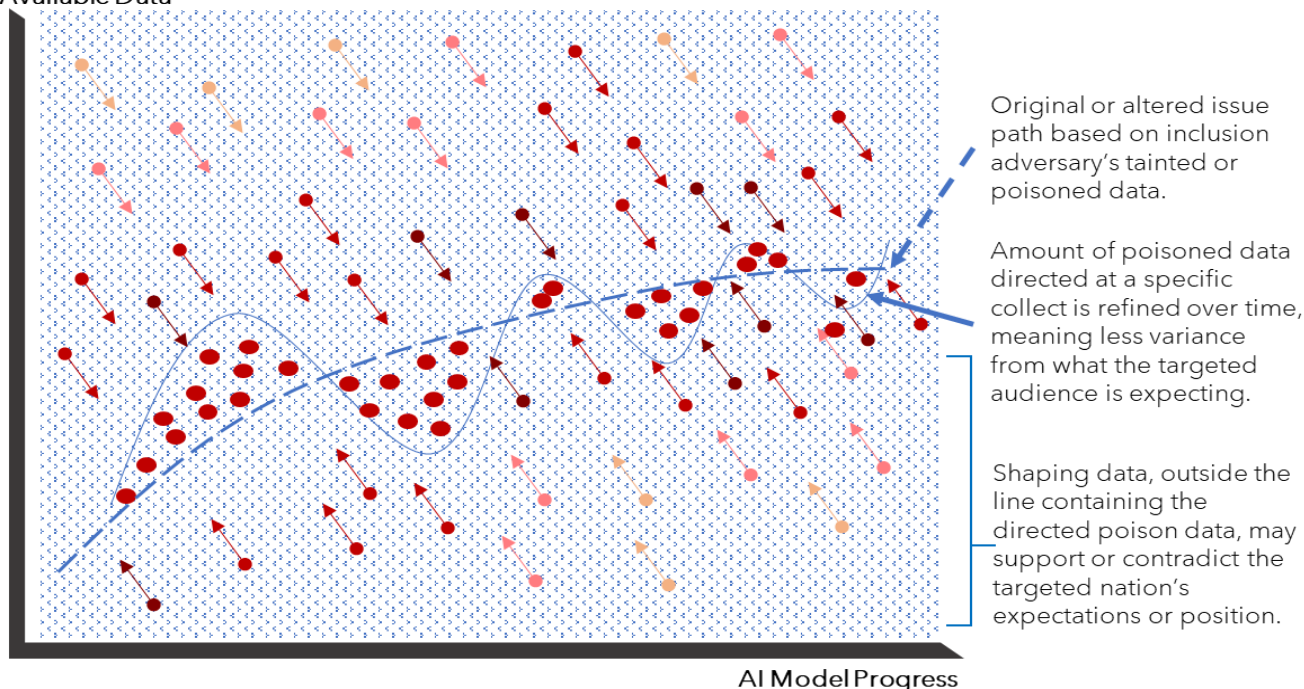
The role of governance may not be to identify data manipulation, as this more often falls to specific organizations in the intelligence or law enforcement communities of a nation. Understanding the extent to which AI models are at risk in a nation will require close formal and informal cooperation between a nation's policy community and the skilled government workers in the organizations responsible for identifying data manipulation. In some nations, this is considered a counterintelligence issue but that can also place too narrow an optic on the data review. Influence campaigns and deceptive data may be incremental, subtle, and not associated with intelligence collection or protection. That suggests that it may be a unique mix of capabilities that understand adversarial intent, capabilities, and domestic efforts to use information in decision-making. If such organizations do not currently exist, then setting up such functions may be an additional objective of the AGI strategy for governance. The following illustration expands on some critical considerations for nations determining governance and use of data for national planning and security.



Model Data Manipulations and Intent Shaping

The world of data is represented by innumerable blue dots in the graphic, with some portion of that data potentially misinformation or disinformation (represented by the directed arrows). Manipulated data may target likely information to be queried, research to be referenced, or the validity of sources. Adversaries can manipulate data to misrepresent the truth about the progress of their national programs, resources, and security. It allows the adversary to support or contradict another nation's decisional process to their benefit.

Available Data



Legend

Tainted Data Purpose

- Subtle to heavy data manipulation to shape the adversary's thought process, research, and solution sets.

- Available data for model development
- Tainted data
- Poisoned data
- Predicted model path
- Data used to refine the model

Directed Data Flow

The data selected for the model's learning may contain an undetermined amount of data directed at likely algorithms, data sets, tools, or users to force specific behavior or shape desired outcomes for the adversary.

Adversary identifies national efforts of interest, often through open-source statements of policy intent on defense, economics, or social equities.

Feedback and adjustment for adversary data manipulation may be nearly constant and supports prioritizing or planning means or type of data manipulation.

The adversary refines and repeats data manipulation as needed with a hidden hand or showing their hand to burn data sources the targeted nations may use.



IS IT TOO LATE TO PREVENT A COMPETITIVE NATURE IN AGI?

There is an increasing concern that the platform economy—large tech companies dominating the digital sphere and determining the apps, associated systems, and curated data available—is similar in many ways to the medieval feudal system of governance that relied on cheap or free labor (data and open AI today) to concentrate power.^{xixiii} Comparably, this shift in power raises the question of whether AGI will gravitate towards more humanist-oriented AI non person entities that incorporate morals and civic wisdom from history, with leanings toward rational but fair institutions? Or, will AGI develop by rejecting these in favor of a more Machievellian view that focuses on power dynamics, stability against inherent political fragility and a more selective application of history? These are two possible behavioral paths AGI may take as it becomes more capable of self-generating observations, information, and solutions. History has also shown us that there was not a clean transition to either of these and that they were often in conflict, which suggests that both may evolve and compete in the same AI space. Just as feudal periods occurred due to a concentration of power in competing enclaves, 21st century power may come from harnessing information in bastioned digital enclaves that support competing AGI models and platforms. We are at a fulcrum point in the debate over AGI's positive and negative impacts on society and how to shape its future paths and outcomes. The next five years are critical to determine AGI's path.

Future AGI behavior is likely path dependent on the human behavior on which it has been trained—suggesting that a self-enriching and fulfilling path and a more ruthless and gaming path for AGI may coexist in the future. The path dependence of AGI is a concept that suggests the order in which centralized AGI power hierarchies and decentralized collective intelligence (DCI) emerge significantly influences the long-term trajectory of AGI intelligence development, which may also impact how future AGI can be used.^{xiv} This path-dependent, irreversible trajectory implies that once AGI or DCI enters a centralized or decentralized regime, transitions become structurally infeasible due to feedback loops and resource lock-in. We have seen this in other technologies, where the first to achieve systemic dominance defines the structural constraints and possibilities of all subsequent development.

- AGI-first systems are focused on emulating or exceeding human cognitive ability as part of the build. This model seeks transformative computing capability and can be found in many leading private industry developers. This may be due to likely shorter development times than DCI and the more marketable nature of a generalized transformative AGI. Because these models are being developed in a highly competitive commercial market, if AGI-first architectures dominate before DCI reaches critical mass the AGI models and how they were trained could monopolize the market with a different perspective of existential risk locked in.



- Conversely, if DCI-first emerges as the predominant model, intelligence stabilizes around decentralized cooperative equilibrium. This is because DCI focuses on getting multiple-AGI entities to collaborate, determining standard measures of success, and focusing on regulatory, research, and incremental scaling with internationally coordinated to mitigate existential or strategic risks.

As we enter 2026, we see many nations that are supporting policies for AGI-first because they see a need for speed in breakthroughs that can cement their standing in this highly competitive field. Others, like China, Japan, and parts of Europe are emphasizing DCI-first strategies (incremental, domain-aware, safety-monitored progression). Whether in agreement or not with other nation-state trajectories, this could lead to further discontinuity and misalignment in international cooperation on AGI if not approached early through fundamental and common governance processes. This new techno-nationalism of coercive, transactional, and protectionist behavior may be at odds with the real benefits to be gained from appropriately governing the coming AGI age. There are growing concerns that the training models have received may be embedding competitive behavior and reinforcing the concept of zero-sum competition.

“...new techno-nationalism is underpinned by the realism doctrine, which portrays the world as zero-sum competition in which states leverage their power of economic coercion, and does not recognize the importance of technological interconnectivity, resource complementarity, open innovation, and positive-sum co-opetition.”^{xv}

When developing a new policy or strategy, governments may start with some sort of exercise or wargame to test outcomes of future scenarios to aid planning and implementation. Using AGI in scenario development, exercise design and execution, or analysis on the future impact of AGI requires a model with the potential for unbiased self-reflection, reasoning, and the ability to explain how it came its decisions or recommendations. Some of those characteristics are being seen in more advanced AI platforms. Creating such a model for AGI would require being able to characterize the types of decisions and decision-makers in that nation, generate multiple future paths for the impact of AGI, and assess the reaction of the society. Some founders and leaders of AI systems warn that AI systems built to act and make decisions lack the ability to anticipate outcomes or understand cause and effect.^{xvi} Those characteristics are foundational to making sound policies and conducting responsible governance, and are often reach through conversation and consensus.

- Recently, AI was trained in the art of diplomacy in geopolitics. An AI, named Cicero by Meta, was trained to conduct strategy, negotiate, make alliances, bluff, withhold information, and mislead in the game of Diplomacy. Cicero was trained using online data from 40,000 games. After playing 82 people in 40 games in an anonymous online league, Cicero ranked in the top 10 percent of participants. This raised the ethical question of what happens when a person



(insert decision maker in a crisis) wrongly believes that there is a person that can be negotiated with, but there is actually AI equipped with cunning but lacking fundamental moral concepts.

AGI that is employed help determine policy and international, geopolitical responses should reflect the long-term aspirations of the nation it represents based on foundational documents.

CLOSING THOUGHTS ON CONSIDERATIONS FOR AGI GOVERNANCE

The potential for AGI to greatly disrupt societies and perhaps even displace or replace the comparative advantages that humans have provided in industry, policy, and security over machines. The possibility of Digital feudalism is real, as is the reality that many aspects of running a nation may be handed over to autonomous systems. Concentration of power by a limited number of mega-corporations is the antithesis of democracy and brings us back to the opening of the paper where democracy, technology waves, and concentration of wealth were discussed. We should also link this concerning development back to whether techno-nationalism can, or should, be applied to AGI if it becomes competitive and sees the world as a zero-sum competition. Without appropriate governance of AI and AGI, we reverse the evolution of democracy (feudal lords to monarchies to representative governments) by redistributing political power to bastions of computing power. If costs of developing and maintaining large AGI models can only be sustained by first-mover corporations, then governance helps a nation to decide whether to cede the control to those corporations or to supra-corporations that have some sovereignty (e.g., the European Union, World Trade Organization, United Nations). That single governance decision may determine whether AI provides bastions of information power to adversarial nations or commercial entities or individuals. The accessibility and control of massive computing power and data has the potential to alter the course of democratic institutions globally.

Democracy is at its heart a compromise between power systems for the greater good, consciously made, and applied by its citizens or their representatives. Once some of the decisional processes are shifted from the citizens to an artificial intelligence there may be a lessening of conversation, compromise, and consensus on national, international, and existential threats due to AGI accelerating power fragmentation. There multiple issues that have emerged in the paper for how an AGI that is highly evolved and integrated throughout society can and should be managed. Most importantly, it will be up to individual nations considering those issues to determine acceptable risk, occupational and economic disruption, and compromise of societal rights or values they are willing to accept from the integration of AGI. It will also be up to those nations to decide the level of international collaboration and cooperation that is necessary for basic standards, rules, and punitive actions for developers and implementers of AGI.



APPENDIX: IMPLEMENTATION AND CONTINGENCY PLANNING

The I-Plan and Contingency Plan are essential documents for success. The first provides a process guide that keeps the implementing organizations on track, while the CONPLAN is an insurance policy against high-risk, high-impact outcomes.

NATIONAL IMPLEMENTATION PLANNING BASICS FOR AGI

1. Process: Define a small set of Governance Objectives
 - a. In this paper they can be developed from the Themes on page 35
 - b. Map each to measurable indicators of what success looks like in each Theme or to satisfy the impacts previously identified in each Pie Quad
 - c. Embed them in an adaptive planning loop that updates policies as AGI capabilities and impacts change. The plan should not be static, but adaptive.
2. Define the core dimensions of AGI governance success for the Objectives, as an example from the paper a workable set of Objectives dimensions might be:
 - a. Theme One: Preserve and Protect. Ensure social stability and public trust in the skilled and unskilled labor sectors by developing sufficient controls to minimize job loss to AGI and enact government sponsored programs to reskill workers forced to transition.
 - b. Theme Two: Lock it Down. Develop governance procedures that preserve individual rights and prevent uneven power distribution by managing AGI proliferation and focusing on key economic and innovation areas.
 - c. Theme Three: Not my problem. Minimize governance in the development of AGI so that institutional responses do not adversely affect competitiveness, innovation, and the economic benefits of increasing integration of AGI in industry and society.
 - d. Theme Four: Let the Good Times Roll. Leverage to the fullest extent the responsible expansion of AI into the country's productive capacity and society, including through international cooperation with successful integrator nations.
3. Define measurable indicators for each Objective and build an "AGI governance health" profile for each so that I-Plan adjustment, reevaluation, revision can be standardized.



4. Red Team you work to identify failure scenarios and revisit the I-Plan to revise before publicizing or promulgating. These can also help with the CONPLAN later.

CONTINGENCY PLANNING BASICS FOR AGI

1. Building off the Red Team approach to the I-Plan, establish:
 - a. The type of national response required for the worst-case scenario(s) developed through Red Teaming, including the organizations responsible for emergency preparedness, response, and remediation.
 - b. Determine if a single, comprehensive CONPLAN is right, or using the framework developed here create separate plans for occupations and the economy, safety and security, and others unique to your country.
 - c. Exercise the CONPLAN with all the requisite organizations and communities that might be affected.
2. Establish a regular review and updating process for the CONPLAN. For fast-moving issues, this could be a continual process by a standing committee or organization with short updates as needed and a major review every six months.
3. CONPLANs are easier in top-down structures, here are some challenges for representational governments seeking adoption by public and private entities.
 - a. Insufficient expertise in both the creating organization and the recipients or participants to assess the risks from ethical, political, financial, and other impacted sectors. This includes individuals with the skills to create high-quality frameworks and bridge the policy and technical worlds.
 - b. A government organization may have insufficient regulatory power, standard risk frameworks, or the resources to develop and implement a CONPLAN nationally.
 - c. Insufficient capacity to map interdependence in systems and incorporate the identified risks and outcomes, this may be helped with techniques like digital twins and simulations. Both take large quantities of data and specialists to curate and assess its relevance to the contingency plan.



ABOUT THE AUTHOR

The author is an expert in creating national defense and security strategies for the United States, as well as policy and strategy development with multiple international organizations, nations, and partners. He has designed exercises of all types: assessments, road maps, implementation plans, and futures methodologies; is skilled in contingency planning for national resilience; and has significant experience standing up government organizations.



He has served as a strategist for U.S. intelligence organizations and the U.S. Navy. He was a senior Asia analyst and Chairman of the Foreign Denial and Deception Committee during two appointments to the United States National Intelligence Council.



ⁱ Technology and Social Change

Castells, M. (2010). *The rise of the network society* (2nd ed.). Wiley-Blackwell.

Perez, C. (2002). *Technological revolutions and financial capital: The dynamics of bubbles and golden ages*. Edward Elgar Publishing.

Nye, J. S. (2011). *The future of power*. PublicAffairs.

Sunstein, C. R. (2017). *#Republic: Divided democracy in the age of social media*. Princeton University Press.

West, D. M. (2018). *The future of work: Robots, AI, and automation*. Brookings Institution Press.

Cave, S., & Dihal, K. (Eds.). (2020). *AI narratives: A history of imaginative thinking about intelligent machines*. Oxford University Press.

Dafoe, A. (2018). *AI governance: A research agenda*. Future of Humanity Institute.

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age*. W. W. Norton.

ⁱⁱ Iversen, T., & Soskice, D. (2019). *Democracy and prosperity: Reinventing capitalism through a turbulent century*. Princeton University Press.

ⁱⁱⁱ Schendstok, M., & Schreiner Wertz, S. (2024). *Occupational exposure to artificial intelligence by geography and education* (Working paper). Office of Economic Policy.

^{iv} Sytsma, T., & Sousa, É. M. (2023). Artificial intelligence and the labor force: A data driven approach to identifying exposed occupations. RAND Corporation.

^v Muro, M., Methkuppally, S., & Kinder, M. (2025, February 19). The geography of generative AI's workforce impacts will likely differ from those of previous technologies. Brookings Institution.

^{vi} European Commission. (2025, April 9). *AI Continent Action Plan*.

^{vii} Howitt, P. (2008). Endogenous growth. In S. Durlauf & L. Blume (Eds.), *The new Palgrave dictionary of economics* (2nd ed., pp. 16–23). Palgrave Macmillan.

^{viii} Edgerton, D. E. (2007). The contradictions of techno-nationalism and techno-globalism: A historical perspective. *New Global Studies*.

^{ix} Vogels, E. A., Rainie, L., & Anderson, J. (2020, June 30). *Experts predict more digital innovation by 2030 aimed at enhancing democracy*. Pew Research Center.

^x Heesen, J., & Ekstrand, T. S. (2025, November 24). *EU proposal to delay parts of its AI Act*. The Conversation.

^{xi} Massachusetts Institute of Technology. (n.d.). *When AI gets it wrong: Addressing AI hallucinations and bias*.

^{xii} Jensen, J. L. (2020, September 2). *Democracy in the age of digital feudalism*. Danish School of Media and Journalism.

^{xiii} Saura García, C. (2024). The age of datafeudalism: From digital panopticon to synthetic democracy. *Philosophy & Technology*, 37, 98.

^{xiv} Sneider, D. (2024, March 6). *Accelerating the decentralized collective intelligence*. Spark.

^{xv} Luo, Y. (2022). Illusions of techno-nationalism. *Journal of International Business Studies*, 53(3), 550–567. <https://doi.org/10.1057/s41267-021-00468-5> (doi.org in Bing)

^{xvi} Gusain, U. (2026, January 30). *Is AGI the future of artificial intelligence — or its end? Godfather of AI has his own fears*. Firstpost.