https://doi.org/10.1016/j.eap.2021.05.001

WHAT EXPLAINS GOVERNMENTS INTEREST IN ARTIFICIAL INTELLIGENCE? A SIGNALING THEORY APPROACH

Running Header: What Explains Governments Interest in Artificial Intelligence? A Signaling Theory Approach

Samar Fatima
PhD Candidate
School of Management
QUT Business School
Queensland University of Technology
GPO Box 2434
Brisbane, QLD, Australia 4001
Email: s.fatima@qut.edu.au
(Corresponding Author)

Kevin C. Desouza
Professor of Business, Technology and Strategy
School of Management, QUT Business School, Queensland University of Technology
Centre for Future Enterprise, Queensland University of Technology
Email: kevin.desouza@qut.edu.au

James S. Denford
Associate Professor, Department of Management, Royal Military College of Canada,
PO Box 17000, Kingston, Ontario, Canada
Email: jamessdenford@gmail.com

Gregory S. Dawson
Clinical Professor
School of Accountancy, W.P. Carey School of Business
Arizona State University
Email: GregorySDawson@gmail.com

Declarations of interest: None

What Explains Governments Interest in Artificial Intelligence? A Signaling Theory Approach

Abstract

Since 2015, several countries have shown significant interest in artificial intelligence (AI) and have released national-level AI strategic plans. These plans reflect the country's rationale for embarking on AI. To identify what factors influence the AI approach of a country, this study employs the signaling theory to decode strategic national AI plans and understand each country's rationale. The study adapts the typology of signals and plots AI information given in national AI plans (AI-enabled public services, research, data, algorithmic ethics, governance) in a matrix of intentionality and veracity considering socio-economic and political conditions. Our findings indicate that countries with high democracy scores are more likely than less democratic countries to prioritize ethical and governance issues of AI, however, this is more pronounced in democratic countries with a lower technology base. The results also suggest that advanced research capability and data accessibility for AI is a precondition to developing a nationwide AI system.

Keywords

National AI plans, Signaling theory, Technology policy, Democracy, Intentionality

Classification Codes: F50, F68, H10, 020, P50

1. Introduction

Nations are conducting significant initiatives for technology supremacy, especially when it comes to leading the artificial intelligence (AI) race. For example, China announced $12 billion spending on artificial intelligence (AI) in 2017 and predicted spending of $20 billion by 2020 (Hao, 2019). Similarly, the US government has a budget of $2 billion in AI projects for the Department of Defense and quantum computing (Dwivedi et al., 2019). This race to AI is not limited to global superpowers. For example, the Government of Singapore has shown significant interest to pilot test autonomous vehicles for public transport (Trueman, 2019). As of January 2020, thirty- four countries have launched national strategic AI plans (Fatima et al., 2020; Future of Life, 2020).

These national plans provide details of a country's strategy to harness the potential of AI and also indicate their approach towards economic, social, and policy-making paradigms related to AI technologies. Additionally, these plans highlight the coordination and assessment of such technology initiatives among various stakeholders of the society, e.g. public agencies and industry partners (Fatima et al., 2020). In their study, Fatima et al., (2020) found that these plans discuss the potential of AI and propose a course of action for AI development and implementation.

With the belief that AI is critical for economic and military dominance, countries are racing in intense competition to develop AI technologies (Kapetas, 2020). In this battle enabled by algorithms, unique patterns are emerging in various countries. Such as AI research culture of countries differs based on numerous political orientations. However, recently a major shift is witnessed on how countries sought to do AI research. (O'Meara, 2019). For example, China, one of the least democratic nations, is using a partnership approach with more democratic countries to promote AI research (The Economist, 2018). For example, Xi'an Jiaotong University's Institute of Artificial Intelligence and Robotics as one of the most highly respected institutes in AI research has a strong collaboration with the USA, Germany, France and UK institutes. In 2017, the 22nd conference on Architectural Support for Programming Languages and

Operating Systems was held in Xi'an, China. As mentioned by conference organizing committee, computer science conferences have largely been held in US or Europe previously, however, due to increasing number of research collaborations, they were able to conduct the conference in China (O'Meara, 2019).

We would have expected great commonality amongst the plans as, worldwide, governments are grappling with the same issues. However, Fatima et al., (2020) found a significant variation among AI plans for inclusion or exclusion of an AI-related concept. For example, few plans emphasized the adoption of AI in the public sector more than in industry, similarly, some plans prioritized algorithmic ethics and AI governance while others did not. For example, France's AI plan explains the incorporation of ethics into the training of engineers and researchers studying AI (France AI Plan, 2018, p. 119). However, Russia's plan emphasized that the government would formulate ethical rules for human and AI interaction (Russia AI Plan, 2019, p. 17).

To study the differences between countries' approaches to AI planning, we draw on signaling theory, which postulates that a difference in information between two parties causes each of them to behave in different ways. The parties involved in signaling theory are the sender (has greater information) who choose whether or how to communicate (signal) the information that can impact or influence the behavior of the receiver (has lesser information) (Connelly et al., 2011). Such signals given in AI plans need to be explored to understand why AI plans differ in intent and veracity and how these differences can impact the future of government with AI and the future of AI in government. The reasons for such differences are not evident yet. The understanding of such differences has larger implications for AI development, AI policy-making, and strategic planning of economies.

However, it is not clear why the signals differ between countries and if the signals are intentional or inadvertent and whether they have high or low veracity. It is this question that motivates this research. As such, our research questions are:

RQ 1 – How do different types of governments signal using their AI plan?

RQ 2 – What are the veracity of these plans?

The paper is structured as follows, first, we present the background of national AI plans and signaling theory and use this to generate our research propositions. In the methodology section, we define the dataset and fsQCA. After methodology, we report the findings and decision criteria for configuration recipes. Lastly, with insight for future research, we conclude the paper.

2. Theory Development

2.1. National Strategic AI Plans

Strategic planning takes a future-oriented approach to develop organizational objectives and evaluating the performance against the objectives (Bryson et al., 2009). In the public sector, strategic planning provides a map of future direction and a course of action along with public agencies' capability to enhance public value (Poister, 2010). Studies on strategic planning in public agencies indicate that, despite budgetary, human and other resource constraints (Hatry, 2002), effective strategic planning in the public sector can bring meaningful change (Barzelay & Campbell, 2003; Hendrick, 2003).

Strategic planning is regarded as a black box until the content is operationalized into strategic plans (Bryson et al., 2009). According to Whittington et. al. (2006), to accelerate organizational change the tools between strategizing to organizing are the strategic plans that interlink desired outcomes with deliverables. The approach by Whittington et.al (2006) focuses on features of strategic plans as artifacts of strategic planning. Similarly, Giraudeau (2008) analyzed the literature on strategic plans and declared strategic plans as tools for practicing strategy and simulation tools to predict the future.

Strategic planning in science, technology, and innovation (STI) is different from general policymaking (Stine, 2009). The difference is mostly due to rapid advancements in STI as compared to other policies. Strategic planning for emerging technologies such as AI becomes more challenging for governments. Emerging technologies are largely discussed due to the uncertainty involved in their emergence (Wheatley

& Wilemon, 1999). Since the development and practical implications of emerging technologies are not fully developed, therefore, the adoption of emerging technologies entails an element of risk and uncertainty (Bonnín Roca et al., 2017). The potential impact of emerging technologies on the economy and society plays a key role in the decision of emerging technologies adoption (Porter et al., 2002).

To develop and diffuse new (emerging) technologies, Metcalfe (1995) argues that a national system of innovation is inevitable. Through such an innovation system, governments design and implement policies to launch technological change. The national system of innovation presents the idiosyncratic institutional environment. The features of a national system of innovation vary from country to country (Freeman, 1995).

As a part of technology diffusion, technology, industrial and economic policies have been used as functional tools since half-century (Clark & Guy, 1998). Clark and Guy (1998) defined technology policy as a set of policies that are intended to persuade firms to develop, commercialize and adopt new technologies. They also presented the framework for technological progress and showed that science and technology, industrial, education, and macroeconomic policies all are bilaterally related to the technical progress of a country. They also suggested that for sustainable technological change, the importance of contextual conditions is undeniable. The contextual conditions of a country largely determine the way a country sought to launch technological change.

Technology policies as a component of the national innovation system cannot work in isolation. The links between contextual conditions and technology policies might not be evident, however, both are strongly related. For example, Genus and Coles (2005) performed constructive technology assessment and found that governance structure and public participation in science and technology debates and decision making can impact technology design and shape the overall system of technology-enabled innovation. Interestingly, recently launched national AI plans present a comprehensive approach towards such innovation. The national AI plans cover a thorough outlook on technology adoption as issues ranging from technical capacity building from computational systems design to governance conundrums, from AI implementation in the industry to public agencies and AI ethics by design to AI ethics by regulations, variety of aspects have been identified and discussed.

The release of national strategic AI plans is the most modern initiative to adopt AI at the country level. The first formal national strategic AI plan was released by Canada in 2017. However, the United States of America and South Korea released AI plans in 2016 but did not declare them as national AI plan specifically. In 2017, five countries including China, Canada, Finland, Japan, and the United Arab Emirates released their national plans. Most of the European countries in 2018 and 16 countries around the world formalized AI in 2019 and released national plans (Fatima et al., 2020). Fatima et. al (2020) analyzed 34 national AI plans and found six common themes among them which are AI priorities for implementation (public sector and industry), capacity development for AI (research, education, public agencies' business model renovation) in data accessibility, algorithmic ethics and AI governance.

In the study by Fatima et. al (2020), six main themes i.e. Use of AI in public services 2) Use of AI in Industries 3) Data for AI 4) AI Research 5) Algorithmic Ethics and 6) AI Governance with thirty-seven subsidiary codes of AI were identified. They assigned a score of 1 (0 for absence) to countries where a subsidiary code was present. The results indicated that some countries scored higher than others due to the presence of codes. For example, European Union countries exhibited greater concern for data sharing among them, whereas, countries with an authoritarian form of government, such as Russia and China emphasized building capacity for in-house data availability and accessibility. Fatima et. al (2020) describe the commonalities and differences between national AI plans based on the presence and absence of codes, however, the underlying reasons for such differences were not pointed out. The understanding of underlying reasons is important to consider while investigating countries' future with AI because these reasons predict the future trajectories. Hence, we caught interest in exploring why a country approached AI in the way it did and what this approach signals about the future.

2.2. Signaling Theory

Signaling is defined as a process by which one entity attempts to convey important information that can induce the other party to make a favorable choice (Spence, 1978). The entity sending the information (signal) is termed signaler and the entity for whom the signal is sent is the receiver (Connelly et al., 2011).

In his seminal work, Spence (1973) defined signaling as the behavior demonstrated by a job applicant to support the selection decision by exhibiting their productive capacities that are not directly observable. The process of signaling occurs due to unequal information between two parties; the inequality of information is called information asymmetry. The core of signaling theory consists of the analysis of various types of signals and the situations in which they are used (Spence, 2002).

The signaling theory has been employed in vast areas of research. Such as in corporate governance studies to signal concern for society via financial statements to prospective investors (Zhang & Wiersema, 2009). Similarly, studies indicate the use of heterogeneous boards in recruitment to signal adherence to social values (Miller & Triana, 2009). In e-commerce studies, the use of signaling fits well because buyers have no access to the physical premise of the seller. The buyers rely on the signals given on the website and situations in which the signals are given to make a purchase decision (Mavlanova et al., 2012).

The typology of signals defines a 2 x 2 matrix of signals as shown in Table 1. This typology suggests that signal intention and signal veracity determine the properties of signals. The signal intention can be deliberate or inadvertent. Similarly, signal veracity varies from high to low. The typology formulates a matrix that compares signal intention and signals veracity with both types (Dawson et al., 2016).

Table 1. Typology of signals

		Signal Veracity	
		High	**Low**
Signal Intention	**Deliberate**	Traditional Signals	Opportunistic Signals
	Inadvertent	Inadvertent Disclosure Signals	Mixed Signals

Source: (Dawson et al., 2016)

Traditional signals are both deliberate and true. These signals are intended to reduce information asymmetry that is a core focus of signaling theory. *Inadvertent disclosure signals* transmit true information, but the sender does not send it deliberately. During inadvertent disclosure, true information is given that also reduces information asymmetry but not induced by the sender's intention. *Opportunistic signals* are not true and induced deliberately by the sender. Such false signals sabotage the objective of signaling theory and increase disadvantageous information asymmetry. *Mixed signals* transmit false information without the sender's deliberate intention. These signals can be taken anyway (true or false) by the receiver. Mixed signals can increase or decrease the disadvantageous information asymmetry (Dawson et al., 2016).

Signaling theory literature mentions use of signaling process in the public sector studies, however, it is a relatively new lens to study public sector interaction with citizens and other stakeholders (Raaphorst & Van de Walle, 2018). A pioneer work on policy reforms figured out that information asymmetry about the government's future intentions is the core reason for citizens and the private sector's insufficient credibility on government policy reforms. The study suggested that transmission of direct signals such as the speed of reforms can help the reform-minded government to gain the credibility of citizens and private sector partners (Rodrik, 1988).

Few studies, such as Goodsell (2000), referred to government agencies' magnificent building architecture and prime location as a signal to exhibit legitimacy and authority to citizens (Goodsell, 1977, 2000). Similarly, Raaphorst & Van (2018) drew on signaling theory to describe the communication between citizens and public officials and found how unobservable signals of trust can be translated into observable signals with both parties on the signaler and receiver sides. The findings of the study reinforced Spence's (2002) statement that the context in which signals are sent and received largely determines the interpretation (Raaphorst & Van de Walle, 2018).

However, the scope of these studies employing signaling theory in the public sector is limited to one signaler (government) and one or two receivers, i.e. citizens and private sector partners. In this study, we have taken a wide range of receivers including citizens, other countries, AI research centers, non-government regulatory entities e.g. OECD, EU.

2.3. Proposition Development

AI is the key to economic growth, national security, and strategic advantages, the competition between countries to dominate in AI is getting fierce. The development and implementation of AI technologies have become the national agenda. This national agenda is being propagated by governments through national AI plans. Countries like the USA and China are allocating billions of dollars to AI research and development of AI systems (Dwivedi et al., 2019). According to Castro et. al (2019), the USA leads the race for AI advancement despite China's enormous spending for AI development. USA's AI start-up ecosystem, production of computer chips, and high-quality AI research were declared some of the factors that help the USA is leading the competition.

However, in another report, China was declared to not only be the AI race-leading country but also surpassing the capabilities of the USA and European countries (Schmidt & Allison, 2020). China's command in both national and commercial security enabled by AI was revealed the reason to lead the race. The effective use of surveillance applications during the covid-19 pandemic has helped China to lead the AI trajectory (Schmidt & Allison, 2020). Similarly, the European Union has also shown a significant increase in AI investment deals from about 30 in 2011 to 350 in 2017 (OECD, 2018). Therefore, it is not possible to uncover the geopolitics of AI and determine what countries are in lead.

Also, to distinguish between technical capacity development and regulatory control development related to these emerging technologies is vital to understand the future of AI. For example, if a country fully deploys AI-enabled public services but public trust is damaged due to the government's tech-centric rather than citizen-centric approach, would such technological adoption likely sustain public value? A recent example of a public trust breach is where an autonomous system of the Dutch government wrongly accused more than 26,000 families of making fraudulent childcare benefits (BBC News, 2021; The Guardian, 2021). To answer such wicked questions, national AI plans are a useful tool to predict what the future of AI holds.

A policy does not exist in a vacuum rather it is influenced by the context in which it is designed, drafted, and implemented (Borrás, 2011). Innovation policy scholars have emphasized the inclusion of a mix of policy instruments to understand innovation at the national level (Branscomb & Florida, 1998). Policy instruments as intangible social constructs are defined in multiple ways. According to Lascoumes and Gales (2007), policy instruments are devices with technical and social dimensions that mediate between government and actors from policy design to policy implementation. The three types of policy instruments as defined by Borras (2011) are regulatory instruments e.g. (intellectual property rights) financial and economic instruments (tax exemptions) and soft instruments (public-private partnerships) (Borrás, 2011). Like general policies, AI policies do not have an "optimal" or fit for all list of policy instruments. The instruments used in AI policies vary according to their context (Borrás, 2011). Similarly, the objectives of AI plans are to inform internal and external stakeholders about AI initiatives. In terms of signaling theory, we define AI plans as signals, governments as signalers and internal and external stakeholders as receivers. However, the quality of signals varies depending on the contextual conditions of a country.

Not only each plan differs in showing intentions but also in some claims made are true while others are not. In this study, we are interested to figure out which claims are veracious with the existing information and which are not. The context of one country differs from the other, therefore no single set of policy instruments can be equally suitable for all countries. These contextual factors impact AI planning and the future of AI. The contextual factors could be several such as, the form of government (democratic or authoritarian), economic indicators, civil liberty, public participation in government decisions. Having mentioned that, we propose our model in figure 1. Next, we define why the information given in AI plans is important.

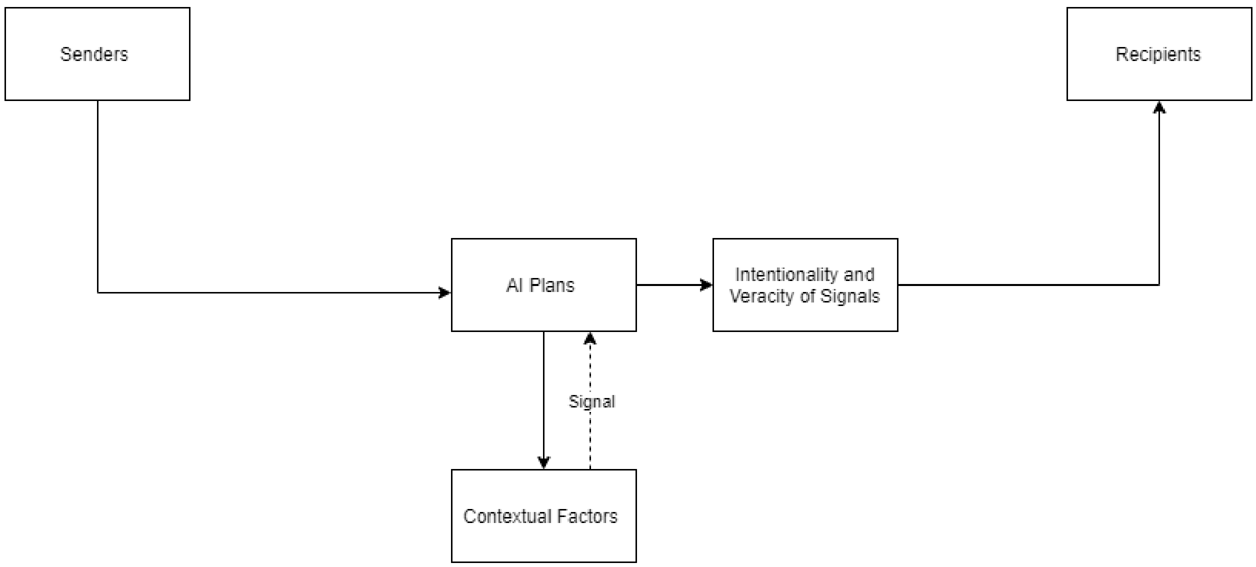

Fig. 1. Signaling theory and AI plans

We suggest that intentionality and veracity of the information on the AI initiatives are important for both; countries releasing AI plans (senders) and recipients of the information, i.e. internal and external stakeholders. AI plan releasing countries convey the information to reduce the information asymmetry and invite opportunities for collaboration of AI research, attract foreign AI experts, build regional data network, etc. By soliciting such information, AI plan releasing countries can highlight the information they deem favorable for AI development, implementation, or governance.

Similarly, internal and external stakeholders (termed as buyers in signaling theory) are those entities that can influence or be influenced by a country's approach to AI. For example, AI research centers find opportunities with similar research interests. Non-government regulatory entities such as Organisation for Economic Co-operation and Development (OECD) or World Economic Forum (WEF) observe the missing AI governance components and highlight governance-related issues. Therefore, reducing information asymmetry between countries and recipients of the information is a cornerstone for AI development, implementation, and governance.

However, reducing information asymmetry depends on the intentionality/deliberateness and veracity of the signals, and not all AI plans exhibit similar intentionality and veracity of signals. Further, the intentionality and veracity of signals are not straightforward to determine. We use three dimensions to ascertain the intentionality and veracity of AI signals to develop our typology table for AI plans (Table 1).

1. **Signal fit -** The extent to which the signal is correlated with unobservable quality (Busenitz et al., 2005; Zhang & Wiersema, 2009). For example, the tendency of a country to lower the spread of covid-19 can be signaled through their international border closure or open status. There is a logical connection between the number of international travelers entering a country with the number of covid-19 cases. We define unobservable quality as the contextual conditions of a country that directly or indirectly impact the policies. Further details on contextual conditions are given in the methods sections.
2. **Signal consistency -** The extent to which there is an agreement between signals from different sources (Fischer & Reuber, 2007). For example, if the democracy score of a country is higher, the tendency of having free and fair elections at the stipulated time is also present. Both these factors democracy score and conduct of elections indicate a consistency among two political factors. However, the two sources of information can be related or unrelated. The source of signals and conditions in this study are countries. Thus, signal consistency in this regard is how aligned the claims made in AI plans are with contextual conditions of the country.
3. **Signal reliability -** The combination of a signal's fit and consistency (Arthurs et al., 2009). If there is a logical explanation and consistency between what signals indicate and contextual conditions, we refer to it as signal reliability. For example, the use of AI in public services signals citizen-centric AI. However, if the use of AI is abiding by rules of data protection and other governance-related issues, only then the signal would be considered reliable.

We refer to four types of signals as discussed in table 1. For example, in AI plans, the traditional signals with deliberate intention and high veracity can be a description of AI projects already initiated or information of budget amount allocated for AI research. The inadvertent signals with high veracity are the ranking of a country in digital literacy or the number of AI research publications per year. These signals even if not included in the plan are already available.

Likewise, if a plan claims that citizens will be included in the process of AI policy design while the contextual factor of citizen engagement rate in policy design is very low, such contradiction is regarded as an opportunistic signal. Mixed signals on the other hand are difficult to catch as the intention of the sender is not clear (whether deliberate or inadvertent). An example of mixed-signal in AI plans is the declaration of using anonymized public data in AI systems, while analysis of contextual conditions fails to depict the use of anonymized data. However, the reason for such inconsistency is not clear. Based on these statements, our propositions are:

- ***Proposition 1:*** *National AI plans signal contextual factors of countries*
- ***Proposition 2A:*** *The intentionality of AI plans (signals) is influenced by national contextual factors*
- ***Proposition 2B:*** *The veracity of AI plans (signals) is influenced by national contextual factors*

We propose that national AI plans (signals) as artifacts of strategic planning of countries (senders) transmit information to a wide variety of receivers (internal and external stakeholders). Employing signaling theory, we decode these signals to predict their intent and to judge their veracity. We consider socio-political and economic factors of a country and explore the link between these contextual factors and signals using Fuzzy Set Qualitative Comparative Analysis (fsQCA). Having done that, we identify the configuration models among the contextual factors and AI plans (signals)

3. Methodology

3.1. Approach

The approach used to study the national system of innovation must enable theoretical multiplicity where multifaceted phenomena can be explored (Park et al., 2020). To conduct systems perspective research, a configurational analytic approach such as Qualitative Comparative Analysis (QCA) is suggested as it uses both theory and method to identify the casual recipes for the occurrence of an outcome (Fiss et al., 2013; Levallet et al., 2020). QCA works with configurational approaches to find out which parts of the system (called conditions) are necessary or sufficient for the occurrence of an outcome (Rihoux & Ragin, 2009; Thiem et al., 2016). QCA was launched as an innovative research approach in Information Systems (IS) to explore complex causal relationships (Fichman, 2004).

Today, QCA is being used in several IS research as the main methodological approach (Park et al., 2020; Tsolakis & Tsekouras, 2016). We sought QCA as a suitable approach to explore the relationship between technology policies (AI plans) and socio-economic contextual factors. By employing QCA, we identify the causal complexities between AI priorities and contextual factors.

Among the types of QCA, we chose fuzzy-set qualitative comparative analysis (fsQCA). In fsQCA calibration of conditions and outcome ranges from 0 (non-membership) to 1 (full membership). Fuzzy sets offer both qualitative and quantitative insights. The calibration of data from 0 to 1 provides features of interval and ratio scales, while such calibration is performed using theoretical and substantive knowledge thus depicting vital qualitative features (Ragin, 2008).

The three steps involved in performing fsQCA are 1) preparation of dataset, 2) construction of truth table and 3) logical reduction of outcomes (Park et al., 2020). In the first step, data is calibrated on a scale of 0 to 1, the calibration standardizes all variables on fully in to fully out in membership. The two important statistical measures considered in fsQCA are coverage and consistency of solutions (Denford et al., 2019). Consistency is the degree to which a relation of necessity or sufficiency between a combination of conditions and an outcome is met within a given set of data, whereas coverage is a measure of empirical relevance that captures the degree of overlap between sets or between a set and the overall solution space, again ranging from values of 0 and 1 (Ragin, 2008). Analysis in fsQCA produces three sets of solutions 1) complex solution 2) intermediate solution and 3) parsimonious solution. As a result of the analysis, core and peripheral conditions are identified. Conditions appearing in both parsimonious and intermediate solutions are considered core while those only in intermediate are considered peripheral. The positive dimension of a condition in a solution is deemed presence (core or peripheral) and the negative dimension of a condition in a solution is taken as absence (core or peripheral) (Ragin, 2008). Detailed information on data preparation is given in the next section.

3.2. Calibration and Principal Component Analysis (PCA)

The dataset used in the study has two components: conditions and outcomes.

3.2.1. Country Conditions

To determine the intention and veracity of signals (AI plans), we gathered information on the characteristics of each country. As fsQCA allows the use of numerical data (when standardized and calibrated), we used country characteristics from the Global Competitiveness Index of the World Economic Forum (WEF) (World Economic Forum, 2017). The most recently available values of variables (yearly, monthly, quarterly, etc.) are used in the study. Initially, we collected a total of 53 variables that define the socio-political and economic characteristics of the countries e.g. democracy score, diversion of public funds, government support for R&D etc. The initial dataset of characteristics with time, scale and value is shown in appendix A.1.

3.2.2. Plan Outcomes

The second component of the dataset is outcomes that have been taken from the appendix of the paper "National strategic artificial intelligence plans: A multi-dimensional analysis" (Fatima et al., 2020).

This paper analyzed national AI plans of countries and assigned a value to various components found in the plan. A total of five outcomes has been used in the study (dataset of outcomes is shown in appendix A.2). Fatima et. al(2020) analyzed thirty four national AI plans and found common themes across them using content analysis. They coded data following Dey's (1993) guidelines and identified six themes with 37 subsidiary codes in them. The common theme among national AI plans are 1) Implementation of AI in public sector functions 2) Implementation of AI in industry sector 3) Data for AI 4) Algorithms 5) Capacity development for AI and 6) AI Governance. To ascertain the coverage of various codes within a plan, they assigned "1" for a code present in the plan and "0" if a code was not present in the plan. Based on the scoring of 0 and 1, they calculated the composite score for themes and countries and declared some plans more detailed (in terms of coverage of concepts) than others. Drawing on the findings of Fatima et. al (2020), we prepared our dataset and used five themes from their analysis that are 1) Use of AI in public services 2) Data for AI 3) AI Research 4) Algorithmic Ethics and 5) AI Governance. We did not use the sixth theme i.e. use of AI in industry, since the scope of our study is limited to the use of AI in the public sector. The second component of data; outcomes data extracted from a secondary data source i.e. findings by Fatima et. al (2020).

In the first step, we calibrate the data set of all country conditions. We use 0 and 1 as fully-out and fully-in values. We use logical reasoning based on the original formulation of each condition as suggested by Ragin (2008) to calibrate the data. We did not use means, minima, and maxima for calibration as these forms the weakest type of calibration. For example, the democracy score is 1 to 10 with 8-10 being defined as a full democracy. This suggests that a logical argument can be made for setting the fully-in point at 8 out of 10 as all cases above this are, by definition, fully democratic. Similarly, the scale defines hybrid as being between 4 and 6, making 5 the midpoint and cross-over between the two. Finally, while authoritarian (i.e. not democratic) is defined as 1 to 4, a case can be made for either setting fully-out as 4 (based on the scale) or 2 (based on parallelism with partial democracy) to calibrate the data.

To create composite variables (data reduction to capture the variance), we ran principal component analysis (PCA) using the calibrated data of country conditions. We found 17 variables that were grouped in 5 groups with all factors loading over 0.800, cross-loadings under 0.250, and 88.0% variance explained, making these highly consistent factors with strong explanatory power (Nunally, 1967). We named these groups according to their common features. Table 2 below presents the five groups and variables in each group with the factor-loaded value.

Table 2. Principal component analysis of country conditions

Conditions	Sub-conditions	1	2	3	4	5
Democracy	Democracy	0.980	0.059	0.024	-0.033	-0.009
	Voice and Accountability	0.969	0.142	0.175	-0.025	-0.019
	Electoral Democracy	0.939	-0.170	-0.074	-0.086	-0.070
	Freedom of Elections	0.910	-0.140	-0.009	0.103	0.038
	Freedom of Internet	0.855	-0.051	0.103	-0.222	0.072
Effective Government	Trust in Politicians	-0.140	0.955	0.167	-0.020	-0.049
	Government Political Stability	-0.008	0.941	-0.009	0.010	0.135
	Diversion of Funds	0.062	0.932	0.222	0.090	-0.014
	Government Future Orientation	-0.195	0.908	-0.024	-0.004	0.104
	Judicial Independence	0.106	0.856	0.231	0.023	0.066
Reform Orientation	Reforms Social	0.020	0.187	0.919	0.048	-0.140
	Reform Society	0.096	0.080	0.887	-0.192	0.109
	Reform Health and Education	0.049	0.174	0.859	0.185	0.007
Political Participation	Public Participation (Local)	-0.020	0.031	0.066	0.974	0.044
	Public Participation (National)	-0.151	0.029	-0.029	0.968	0.045
Technical Environment	Technical Environment for Firms	0.052	0.088	-0.224	-0.078	0.816
	Public Authorities Support to R and D	-0.042	0.078	0.210	0.174	0.805

Next, we standardized (0-1) the calibrated values using PCA values and created composite values. The standardized composite scores of each country are shown in appendix B.1. The values shown in appendix B.1 are fsQCA ready-to-use conditions data. To prepare a dataset of plan outcomes, we performed calibration and used theoretical reasons rather than taking minima, maxima, mean or median. The outcomes data have single values since there was no need for factor analysis. The fsQCA prepared plan outcomes data is shown in appendix B.2.

4. Results

4.1. Correlational Analysis

To identify if there were any dominant conditions, we first examined the correlations between country conditions and plan outcomes as shown in Table 3.

Table 3. Correlation matrix

Country Condition / Plan Outcomes	Democracy	Effective Government	Reform Orientation	Political Participation	Technical Environment		Public Services	Research	Data	Algorithmic Ethics	Governance
Democracy	1										
Effective Government	-.066	1									
Reform Orientation	.183	.590	1								
Political Participation	-.247	.101	.100	1							
Technical Environment	-.003	.468	.266	.218	1						
Public Services	-.231	.069	-.176	-.067	.180		1				
Research	-.019	.023	-.095	-.261	.143		.257	1			
Data	.121	-.130	-.066	-.157	-.190		.289	.555	1		
Algorithmic Ethics	.205	-.043	.007	-.081	-.284		.110	.267	.530	1	
Governance	.213	-.308	.006	-.275	-.253		.199	.416	.482	.547	1

The interesting insights from the correlation matrix (among country conditions) indicate that effective government was significantly correlated with reform orientation (.590) and technical environment

(.468). However, reform orientation and technical environment were not correlated with each other (.266). Such results indicate the orthogonal connection between these three conditions. So while either reform-oriented governments are effective or governments in technical environments are effective, technical environment and reform orientation are generally unrelated to each other. Therefore, a country can be advanced in technical capabilities but still can avoid reformative initiatives. Another interesting and relatively less expected connection was found between democracy and effective government (-.066). The (negative) low value of the correlation score indicates that not all democratic countries are working effectively nor are only democratic countries perceived as the most effective.

Next, we discuss the correlation scores among outcomes. As expected, a strong correlation was found between data and research (.555) indicating that countries with high accessibility to data to be used for AI have greater concern for AI research. Similarly, a strong positive correlation was found between data and algorithmic ethics (.530). As expected, governance and algorithmic ethics also showed a positive correlation (.547) indicating that countries with high concern for algorithmic ethics signal formulation of AI governance mechanism.

No strong correlations were found between conditions and outcomes. This validates the choice of fsQCA as the suitable methodology since fsQCA works with causal recipes among conditions and outcomes and more relevant in situations like this where no clear one-to-one connection can be made and interpreted. In summary, the correlation matrix identifies no dominant conditions and proves the use of fsQCA as the right choice for investigating such phenomenon. Next, we present the results of fsQCA performed both for an aggregate AI plan outcome independently for each sub-plan outcome.

4.2. Configuration Analyses

To present the configuration analysis, we first create the indicators and their description to be used in configuration tables. The indicators and their descriptions are given in Table 4.

Table 4. Indicators for Configuration Tables

Indicator	**Description**
⬤	Necessary presence of a core condition
●	Necessary presence of a peripheral condition
⨂	Necessary absence of a core condition
⊗	Necessary absence of a peripheral condition
Blank	The presence or absence of the condition does not impact the outcome
High	High outcome configuration
Low	Low outcome configuration
Note: Solutions that have the same core conditions are grouped by those conditions (i.e. High 1, High 2 or Low 1, Low 2) with configurations with the same core conditions but different peripheral conditions labeled with letters (i.e. 1A, 1B), while configurations that include two core conditions are labeled with both (i.e. 1A/2A).	

4.2.1. Composite

Table 5 shows the configurational analysis for all of the country conditions and all of the components of the AI plan.

Table 5. Composite Configurations

	Comp High 1A	Comp High 1B/2A	Comp High 1C	Comp High 1D	Comp High 2B	Comp Low 1
Democracy	⬤	⬤	⬤	⬤		⨂
Effective Government	●	●		⊗	●	●
Reform Orientation	●		⊗	⊗		●
Political Participation			●		●	●
Technical Environment		⬤		⊗	⬤	⨂
Raw coverage	0.615	0.587	0.358	0.291	0.583	0.177
Unique coverage	0.069	0.025	0.019	0.040	0.104	0.177
Consistency	0.840	0.822	0.878	0.963	.831	0.955
Solution Coverage	0.906					0.177
Solution Consistency	0.816					0.955

As shown, equifinality, which refers to multiple paths for achieving the same outcome, and causal complexity, that many different "recipes" exist, are present with our high configurations. While democracy is seen in most of the AI plans, it is not seen in all of them and the variety of other factors indicates high casual complexity. Not surprisingly, all our countries except UAE (the single, authoritarian and low-technology country in the low solution) are found in one of our configurations and that indicates that there are multiple ways to develop an AI plan. Since the composite AI plan index does not differentiate to a great extent the various countries, and so we delve deeper by looking at the five different components of the AI plans.

4.2.2. Public Services

Table 6 shows the configurational analysis for Public Services.

Table 6. Public Services Configurations

	PS Low 1	PS Low 2A	PS Low 2B	PS Low 3A	PS Low 3B
Democracy	●		●	●	
Effective Government	●	●	⊗		●
Reform Orientation	●	●	⊗	⊗	⊗
Political Participation		●		●	●
Technical Environment		⊗	⊗		●
Raw coverage	0.625	0.346	0.272	0.355	0.315
Unique coverage	0.243	0.015	0.042	0.017	0.010
Consistency	0.847	0.976	0.893	0.863	0.873
Solution coverage	0.830				
Solution consistency	0.810				

For high public services (Public Services), no solution was generated. However, low public services had all but two configurations load. The findings suggest that no countries in the sample had deployed AI for public services. Looking specifically at the output, it is Low1 that dominates with the greatest raw and unique coverage. The very small unique coverage in 2A/2B and 3A/3B shows that these paired configurations share a great deal of commonality with each other; it is also worth noting that they had low unique coverage in the parsimonious solution too. Essentially, almost all the solution space is covered by nations that lack AI for public services.

4.2.3. Research

Interestingly, no configurations – high or low – were generated (and thus no table is shown). This is a very interesting result as it means that the countries' conditions do not differentiate in the area of research. These findings suggest that the population of all countries with AI Policies are very strong in AI research/research strategy & policy. Also, it implies that strong research and research policy is a precondition to developing a national AI plan. Since no parsimonious and intermediate solutions were generated for research and ~ research, therefore, it is worth noting that AI research orientation is common across all countries, irrespective of their characteristics.

4.2.4. Data

The configurations for Data are shown in table 7.

Table 7. Data Configurations

	Data High 1A	Data High 1C	Data High 1D	Data High 1B/2A	Data High 2B
Democracy	⬤	⬤	⬤	⬤	
Effective Government	●		⊗	●	●
Reform Orientation	●	⊗	⊗		
Political Participation		●			●
Technical Environment			⊗	⬤	⬤
Raw coverage	0.515	0.300	0.240	0.491	0.489
Unique coverage	0.070	0.017	0.032	0.024	0.089
Consistency	0.868	0.912	0.985	0.850	0.860
Solution coverage	0.780				
Solution consistency	0.867				

Equifinality and causal complexity were present in our high data configurations. The output of the fsQCA standardized test showed that the greatest raw coverage is High 1A with democracy as core and effective and reformative government as peripheral conditions for data. The smallest raw coverage is High 1D where democracy is suggested as core presence and participative, reformative and effective government as a peripheral absence for the outcome (1) of data, while democracy was the most frequently seen factor, it was not present in all configurations.

The results of the data had close similarities to research configurations. However, the only point of difference is that all data did not show a high outcome (1). The exception was the UAE. which is a non-democratic and lower-tech environment. One probable reason for this is UAE's recent heavy investments in AI are significant, however, it has not yet caught up to countries that started earlier. Again, this is a single country outlier and there is almost uniform adoption of AI data policy in the population of nations with AI policies. In the second part of the analysis ~Data had no configurations to show the low outcome (0). Like research, data also indicates that availability and accessibility of data for AI is a feature common across AI plan releasing countries. This finding also signals the importance of data to build AI capabilities.

4.2.5. Algorithmic Ethics

The configurations for Algorithmic Ethics are shown in Table 8.

Table 8. Algorithmic Ethics Configurations

	Ethics High 1A	Ethics High 1B	Ethics High 1C	Ethics Low 1	Ethics Low 2
Democracy	⬤	⬤	⬤	⨂	⨂
Effective Government	●	●	⊗	●	●
Reform Orientation		●	⊗	⨂	●
Political Participation	●			●	●
Technical Environment	⨂	⨂	⨂	●	⨂
Raw coverage	0.320	0.315	0.262	0.160	0.128
Unique coverage	0.016	0.032	0.086	0.055	0.022
Consistency	0.877	0.826	0.776	0.840	0.825
Solution Coverage	0.434			0.183	
Solution Consistency	0.814			0.857	

Algorithmic ethics found the most diverse and therefore most interesting results. Non-democracies are all right at the bottom of the truth table (truth tables are given in appendix C) but, interestingly, those democratic paragons in the public services truth table are not represented here. There is visibly a combination of democracy and a low technical environment that is important – New Zealand, India, Lithuania, Spain, Serbia, Czech Republic, Mexico, Italy, Uruguay. The results underpin several reasons for low tech and high democratic countries.

Firstly, the absence of a strong national technical base indicates that the country is focused on ethical issues to prevent external actors from applying AI tools and techniques to the nation. Secondly, such a focus on algorithmic ethics lays the foundation of ethical innovation because AI as the industry has not developed yet. Thirdly, countries with average and high democracy scores are highly likely to anticipate and mitigate the risks associated with the use of AI in wake of AI malfunctioning incidents. To gain and retain public trust in AI system deployment, these countries signal higher concern for algorithmic ethics.

In the high outcome of algorithmic ethics (1), some of the high democratic countries such as Australia, Belgium, Canada, Denmark, Estonia, Korea, Luxembourg, Malta, Netherlands, Norway, Portugal, and the United Kingdom indicated the presence of all conditions but still did not result in the high outcome of algorithmic ethics. This finding signal that despite having high democracy, effective, participative, and reformative government, concern for algorithmic ethics is largely determined by the technical environment. Countries with high technical environments seem less prudent about algorithmic ethics as compared to those with low technical environments. This phenomenon is quite evident in the solution. The greatest raw coverage is High 1A where democracy is shown as core presence and technical environment is core absence.

In the low outcome of algorithmic ethics, it was non-democratic countries at the top of the list of those without algorithmic ethics policies, which would tie into their authoritarian use in public services – Russia and UAE are the two in the low algorithmic ethics. The overall assessment of algorithmic ethics shows that democracy and the technical environment play a significant role. Surprisingly, a low technical base reinforces the cautious intentions to inculcate algorithmic ethics in national AI plans.

4.2.6. Governance

The configurations for AI governance are shown in Table 9.

Table 9. Governance Configurations

	AI Gov High 1	AI Gov High 2	AI Gov High 3A	AI Gov High 3B/4A	AI Gov High 4B	AI Gov High 5	AI Gov Low 1
Democracy	⬤	●	⬤	⬤	●	⨂	⨂
Effective Government	●	●	●	⨂	⨂	●	●
Reform Orientation	⬤			⊗	⊗	⨂	●
Political Participation		⨂	●		●	●	●
Technical Environment		●	⨂	⨂		●	⨂
Raw coverage	0.577	0.296	0.273	0.298	0.257	0.106	0.168
Unique coverage	0.177	0.023	0.034	0.012	0.004	0.054	0.168
Consistency	0.821	0.884	0.942	0.952	0.917	0.874	0.842
Solution Coverage	0.781						0.168
Solution Consistency	0.829						0.842

Equifinality and causal complexity are again high in AI governance and this suggests that many paths are possible for achieving this condition but challenge interpretation. Referring to the outcome table, the highest raw coverage is found in solution High 1 where democracy and reform orientation are core presence for AI Governance. It was mostly non-democratic countries clustered at the bottom of the Governance truth table and the top of the low Governance one, but with only one showing up in the low Governance solution. Again, this suggests that authoritarian governments are less concerned about unrestricted AI developments than democratic governments in a similar way that they are more likely to use public services for control purposes. On the low side of the Governance outcome, only one country, UAE was present again.

According to the truth table of high Governance, democratic countries dominate with a range of other combinations of factors. These findings suggest that democratic governments recognize the benefits of good governance in managing technology with far-ranging societal and ethical implications such as AI. This finding pairs well with the algorithmic ethics one. The interesting insight of the Governance truth table is that countries with high democracy and low technical environment cover the first few rows of the truth table; these are Lithuania, New Zealand, Czech Republic, Serbia, Spain, Uruguay, India, Italy, and Mexico.

The underlying reason to prioritize governance among technically low capability countries is also because of the deliberate effort to develop AI capabilities with a strong governance mechanism in the first place.

5. Discussion

Our research questions were to decode the information (signals) given in AI plans considering various contextual conditions and predict the pattern of AI in countries. To discuss the fsQCA results, we refer to the signaling theory and recall the four types of signals discussed in earlier sections of the study. The four outcomes used in the study are 1) Public services 2) Research 3) Data 4) Algorithmic ethics and 5) Governance.

5.1. Criteria to determine intentionality and veracity

The criteria to determine the intentionality and veracity of outcomes is selected as follows:

Table 10. Criteria to determine intentionality and veracity

<table>
<tr><th colspan="2" rowspan="2"></th><th colspan="2">Signal Veracity</th></tr>
<tr><th>High</th><th>Low</th></tr>
<tr><td rowspan="2">Signal Intention</td><td>Deliberate</td><td>1. Expressed in Plans (Outcomes)
2. Established in Contextual Factors (Conditions)</td><td>1. Not Expressed in Plans (Outcomes)
2. Established in Contextual Factors (Conditions)</td></tr>
<tr><td>Inadvertent</td><td>1. Expressed in Plans (Outcomes)
2. Not Established in Contextual Factors (Conditions)</td><td>1. Not Expressed in Plans (Outcomes)
2. Not Established in Contextual Factors (Conditions)</td></tr>
</table>

As discussed in the theory development section, signal fit, signal consistency and signal reliability help in determining the intention and veracity of signals. We found signal fit among signals and country conditions when there existed a logical explanation in terms of established knowledge. For example, democratic countries are more likely to exhibit concern for AI governance. Due to the increased involvement of the public in governmental decision-making, chances of concern for AI governance e.g.

data privacy and fair and equitable treatment by autonomous systems are more likely to occur. Thus, we used a logical explanation of the outcome in deciding signal fit.

To determine signal consistency, we used an empirical approach and found out which of the signals agreed with contextual data. The determination criteria for signal consistency was based on evidence found from empirical data i.e. signals data and conditions data. Meanwhile, signal reliability was determined using a combination of both signal fit and signal consistency.

Using findings of table 10, we placed signal fit, consistency, and reliability in table 11. A higher value of fit, consistency and reliability are shown by plus sigh (+) and a lower value of the three concepts is shown as a minus sign (-).

Table 11. Signal Assessment

		Signal Veracity	
		High	**Low**
Signal Intention	**Deliberate**	Signal Fit (+) Signal Consistency (+)	Signal Fit (-) Signal Reliability (-) Signal Consistency (-)
	Inadvertent	Signal Consistency (-) Signal Fit (+)	Signal Reliability (-) Signal Consistency (-) Signal Fit (+ or -)

Next, we present how national AI plans (outcomes) are categorized in terms of signal fit, signal consistency, and signal reliability about contextual conditions. The study has five outcome variables, defined as signals i.e. AI Research. AI Data, Algorithmic Ethics, AI Governance and Use of AI in Public Services. In this section, we categorize these signals in the template of table 11.

The first signal is AI research. The analysis showed that the contextual conditions of all countries are in full agreement with the claims made in the national AI plans. The population of AI plans releasing countries has shown that building research capabilities are a precondition to developing AI. Bolstering AI research is vital to develop the national AI landscape. It was one of the most dominating themes found in almost all AI plans (Fatima et al., 2020). Thus, AI research counts as a high signal fit (+). Also, the analysis showed that there is consistency between information given in signal and contextual conditions, therefore, the signal consistency for AI research is also high (+). Based on the presence of high signal fit and high signal consistency, we categorize AI research as a traditional signal with high intention and high veracity as shown in quadrant 1 of table 12. Traditional signals reduce information asymmetry and fulfils the objective of signaling theory (Dawson et al., 2016). By highlighting AI research capabilities, countries showcase their research priorities to relevant stakeholders such as citizens, technology companies, sponsors and academicians. Moreover, as a traditional signal, AI research also indicates its role as a prerequisite to developing AI at the national level.

The second signal is AI data or data required to build AI systems. Data works as fuel for AI systems, thus shows a high signal fit (+). The role of data for AI system development has been well recognized by AI plans (Fatima et al., 2020). Upon investigation of the information about data in AI plans, we noticed that data signals are largely validated by contextual conditions, indicating a high signal consistency. co. One exception was found among authoritarian countries – UAE. Results of data suggest that the signal has high signal fit (+) and high signal consistency (+). Like AI research, data is a primary factor for countries to develop and deploy AI systems. Meanwhile, data accessibility was found highly prevalent among democratic countries. We suggest data signals as one of the traditional signals with high intention and high veracity. Data is a resource required for building AI capability and AI plans indicated use and accessibility of data and contextual conditions are found in agreement with such claims. The data signal is also categorized in the first quadrant as signal fit and signal consistency are both on the higher end.

The third signal is Algorithmic Ethics. The results indicated that democratic countries have shown greater concern for algorithmic ethics that shows a high signal fit (+). However, among democratic countries, countries with low technical capability topped the list. This indicates that countries, with high technical capabilities, were not in full agreement with contextual conditions indicating a low signal consistency (-). The results also indicate that countries with low technical capabilities are proactive in building ethics by design since they have to lay the foundation for technology (AI specifically), while those with a strong AI foundation might assume it a challenge to re-build systems with ethics by design, this indicates lack of consistency between both sources of data. Thus, democratic countries indicate the emission of inadvertent signals for concern for AI ethics.

Similarly, the results indicated authoritarian countries have shown less concern for algorithmic ethics that depicts a high signal fit with contextual information (authoritarian countries have less involvement of the public in decision making and eventually lesser concern for ethical implications of AI). Meanwhile, the signal consistency is low as plans claim to make higher concern for algorithmic ethics but are not validated by the contextual information. Therefore, authoritarian countries also indicate the emission of inadvertent signals. The algorithmic ethics signals are placed in the second quadrant of the intention and veracity matrix i.e. table 12.

The fourth signal is AI governance. There are considerable similarities between results of algorithmic ethics and governance, and both are placed in the second quadrant of the matrix (Table 12). Democratic and authoritarian countries showed similar patterns for AI governance as shown for algorithmic ethics. Democratic countries with low technical capability have shown greater concern for AI governance (high signal fit and low signal consistency). Meanwhile, authoritarian countries depicted lesser concern for AI governance (high signal fit and low consistency). Algorithmic ethics and governance place in the category of inadvertent signals. Inadvertent signals as discussed in the theory section, reduce information asymmetry and fulfill the objective of signaling theory but the disclosure of this information is not induced by the sender. For algorithmic ethics and governance, the true information is disclosed by analysis of contextual conditions. However, AI plan releasing countries didn't intend to disclose because there is no consistency among the two sources of data. Meanwhile, inadvertent signals do not indicate manipulation by the sender that might occur in opportunistic signals.

The last signal is Use of AI in public services. This signal indicates various interesting insights. As the use of AI in Public services is extensively discussed in national AI plans (Fatima et al., 2020). It was expected that contextual conditions of countries will validate such signals, however, the results showed that countries have not yet prioritized AI-enabled public services for citizen support. For authoritarian countries, the use of AI in public services was not expected (low signal fit), however, two authoritarian countries indicated the use of AI in negation set of fsQCA tests, indicating low consistency and low reliability of signals. According to signal fit, consistency and reliability, the veracity of these signals was very low with a high intention. Such a combination of the typology of signals is considered opportunistic signals. Thus, the use of AI in public services in authoritarian countries is placed in the third quadrant of the matrix that indicates the opportunistic nature of signals (table 12). This trend indicates the use of AI for citizen control and surveillance more than citizen support in authoritarian countries, and the signaling is likely involuntary and hence an opportunistic signal. We placed authoritarian countries' use of AI in public services in the third quadrant with low signal fit, low signal reliability and low consistency.

Opportunistic signals falsely sabotage the objective of signaling theory and increase disadvantageous information asymmetry. For example, the information given in AI plans that do not turn to be true can increase information asymmetry among stakeholders who are users of such information. The information asymmetry has not only the tendency to slow the process of AI deployment but can also misperceive the AI priorities.

The results of democratic countries indicated low use of AI in public services. However, the plans made claims about the use of AI in public services that indicate a high signal fit. The results did not indicate the use of AI with contextual conditions. Also, the intention of democratic countries falls in the inadvertent quadrant as claims were about but not verified by the contextual conditions. Therefore, the use of AI in public services for democratic countries is placed in the fourth quadrant of the matrix and these signals are labeled as mixed signals as intention and veracity of signals are difficult to ascertain.

One reason for less validating signals (mixed signals) for AI in public agencies is the inflexibility in public agencies' business models to design and deliver AI-enabled public services. The intentions to use AI in public services are deliberate and abundantly discussed in the AI plans, however, the current state of contextual conditions shows that public agencies are not yet ready to fulfill such plans. This highlights the need to renovate the business models of public agencies to develop readiness for AI. Based on the analysis, we suggest that signals of AI in public services in democratic countries are likely opportunistic.

Table 12. AI Plans Intention and Veracity Matrix

		Signal Veracity	
		High	**Low**
Signal Intention	**Deliberate**	*Research Data*	*Public Services (Authoritarian)*
	Inadvertent	*Algorithmic Ethics Governance*	*Public Services (Democratic)*

Adapted from (Dawson et al., 2016)

Summarizing the AI Plans Intention and Veracity Matrix, countries want to share information on their research and data initiatives for AI. The reason is that they seek potential collaboration opportunities about research or data sharing. Further research with highly relevant conditions for data (e.g. number of data generated in a day, government share in the generated data) and research indicators (e.g. number of international conferences, the proportion of foreign speakers, and topics covered in conferences) can better predict what kind of technologies are sought to be implied in countries.

Algorithmic ethics and governance issues related to AI are plotted in the inadvertent signals category. This finding is most exciting to be explored further. Democratic countries led authoritarian countries about algorithmic ethics and governance. However, among these countries, those with a low technical environment topped the list. We suspect that this may be due to a low-technology country having to grapple with such issues for the first time. However, further research can further investigate why the technical environment hinders (does not support) algorithmic ethics.

This finding is very useful for AI capability developing countries since the truth table listed countries such as New Zealand, Lithuania, Italy, and India, etc. at top of those that intend to enable algorithmic ethics and their contextual conditions also are in the right direction. As mentioned earlier, the geopolitics of AI is launching new trends, this is another influential one. Where technology advanced countries US and China are building AI capabilities (advanced research and system design), countries who are lagging, such as New Zealand, India and Italy, are building capability along with adherence to governance issues. It is important to reiterate that merely building AI capabilities would not help in leading the race rather ensuring a sustainable technology adoption process.

Another thought-provoking finding is the country's ambiguous status on the use of AI in public services. All countries irrespective of their democracy scores have given mixed/opportunistic signals about AI-enabled public services. One of the core objectives to adopt AI at the national level is to improve the quality of life of citizens. However, this objective has not been witnessed through the findings. Authoritarian countries rather showed a negative connection with the negation of AI in public services, which indicates that the use of AI seems to control the citizens instead of facilitating them. However, a further investigation with more relevant indicators of AI in public services can support or deny the proposition that emerged out of this study.

6. Limitations and Conclusion

As with all studies, this one has certain limitations. First, we had no way to validate that the plans were developed using any kind of consensus in the various countries but have no reason to believe that the plans do not represent the intentions of that country. Second, our dataset was limited to those countries that have actually produced a plan and cannot infer why a country might not have already developed one.

This study employs signaling theory to explore how national AI plan releasing countries view AI and to what extent their contextual conditions are related/unrelated to these plans. To do so, we used a fuzzy set qualitative comparative analysis to decode the national AI plans according to a set of socio-political and economic conditions. The study uses conditions data from WEF and outcomes data from the analysis of 34 national AI plans (Fatima et al., 2020). After performing statistical processes to ensure rigor (exploratory factor analysis, principal component analysis and calibration of values), we prepared the data for fsQCA. Using fsQCA software, we generated the truth tables and conducted logical reduction to discuss the presence and absence of core and peripheral conditions in the causal recipes.

According to our results, strong technological and data capabilities underpin the ability of nations to deploy AI capabilities. For those nations that have fielded AI policies, AI is used by governments to reinforce their underlying tendencies. Democratic countries signal to use AI for further transparency and ensure they are well-governed and used effectively. Conversely, authoritarian countries signal to use AI to control and eschew governance and ethics. This study uses five independent outcomes (public services, research, data, algorithmic ethics, and governance) and general socio-economic conditions chosen for contextual factors.

The current study gives a solid base to further examine these contextual factors in the development of AI policy across nations. The study takes initiative in reading the between the lines messages as emitted from AI plans. These indirect signals can not only inform stakeholders about agenda of countries towards AI but also a guiding tool for countries who have not released national AI plans. By incepting the debate on reading signals from national AI plans, the study opens vast areas to be explored in future research.

In future research, each outcome can be investigated separately with a set of context-specific conditions. Also, since the study considers the AI initiatives as mentioned in national plans, it is important to investigate the feasibility of such initiatives. For example, how well the countries are capable to develop, and diffuse AI-enabled systems. For such exploration, computational capabilities of countries such as the availability of supercomputers can be an indicator to investigate in future research. Similarly, other indicators to gauge feasibility of AI initiatives can be exploration of AI workforce composition.

We conclude the paper with the reflection that on the surface the plans depict all manner of honorable goals for racing to AI implementation but our deeper examination into the intentionality and veracity of the plans reflects a far more complex reality.

7. References

Arthurs, J. D., Busenitz, L. W., Hoskisson, R. E., & Johnson, R. A. (2009). Signaling and initial public offerings: The use and impact of the lockup period. *Journal of Business Venturing*, *24*(4), 360–372. https://doi.org/10.1016/j.jbusvent.2008.02.004

Barzelay, M., & Campbell, C. (2003). *Preparing for the Future: Strategic Planning in the U.S. Air Force*. Brookings Institution Press.
BBC News. (2021, January 15). Dutch Rutte government resigns over child welfare fraud scandal. *BBC News*. https://www.bbc.com/news/world-europe-55674146
Bonnín Roca, J., Vaishnav, P., Morgan, M. G., Mendonça, J., & Fuchs, E. (2017). When risks cannot be seen: Regulating uncertainty in emerging technologies. *Research Policy*, *46*(7), 1215–1233. https://doi.org/10.1016/j.respol.2017.05.010
Borrás, S. (2011). Policy learning and organizational capacities in innovation policies. *Science and Public Policy*, *38*(9), 725–734. https://doi.org/10.3152/030234211X13070021633323
Branscomb, L. M., & Florida, R. (1998). Challenges to technology policy in a changing world. *American Chemical Society*, *28*(6), 13–22.
Bryson, J. M., Crosby, B. C., & Bryson, J. K. (2009). Understanding Strategic Planning and the Formulation and Implementation of Strategic Plans as a Way of Knowing: The Contributions of Actor-Network Theory. *International Public Management Journal*, *12*(2), 172–207. https://doi.org/10.1080/10967490902873473
Busenitz, L. W., Fiet, J. O., & Moesel, D. D. (2005). Signaling in Venture Capitalist—New Venture Team Funding Decisions: Does it Indicate Long–Term Venture Outcomes? *Entrepreneurship Theory and Practice*, *29*(1), 1–12. https://doi.org/10.1111/j.1540-6520.2005.00066.x
Castro, D., McLaughlin, M., & Chivot, E. (2019). *Who Is Winning the AI Race: China, the EU or the United States?* Center for Data Innovation. https://datainnovation.org/2019/08/who-is-winning-the-ai-race-china-the-eu-or-the-united-states/
Clark, J., & Guy, K. (1998). Innovation and competitiveness: A review: Practitioners' forum. *Technology Analysis & Strategic Management*, *10*(3), 363–395. https://doi.org/10.1080/09537329808524322
Connelly, B., Certo, T., Ireland, R., & Reutzel, C. (2011). Signaling Theory: A Review and Assessment. *Journal of Management - J MANAGE*, *37*, 39–67. https://doi.org/10.1177/0149206310388419
Dawson, G., Watson, R., Boudreau, M.-C., & Pitt, L. (2016). A Knowledge-centric Examination of Signaling and Screening Activities in the Negotiation for Information Systems Consulting Services. *Journal of the Association for Information Systems*, *17*(2), 77–106. https://doi.org/10.17705/1jais.00422
Denford, J. S., Dawson, G. S., & Desouza, K. C. (2019). Performance impacts of structure and volition in implementing policy through IT-enabled government-to-citizen and government-to-employee interactions. *Economic Analysis and Policy*, *64*, 116–129. https://doi.org/10.1016/j.eap.2019.08.005
Dey, I. (1993). *Qualitative Data Analysis: A User-Friendly Guide for Social Scientists*. Routledge.
Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., … Williams, M. D. (2019). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002
Fatima, S., Desouza, K. C., & Dawson, G. S. (2020). National strategic artificial intelligence plans: A multi-dimensional analysis. *Economic Analysis and Policy*, *67*, 178–194. https://doi.org/10.1016/j.eap.2020.07.008
Fichman, R. (2004). Going Beyond the Dominant Paradigm for Information Technology Innovation Research: Emerging Concepts and Methods. *J. AIS*, *5*. https://doi.org/10.17705/1jais.00054
Fischer, E., & Reuber, R. (2007). The Good, the Bad, and the Unfamiliar: The Challenges of Reputation Formation Facing New Firms. *Entrepreneurship Theory and Practice*, *31*(1), 53–75. https://doi.org/10.1111/j.1540-6520.2007.00163.x
Fiss, P. C., Marx, A., & Cambré, B. (2013). Chapter 1 Configurational Theory and Methods in Organizational Research: Introduction. In P. C. Fiss, B. Cambré, & A. Marx (Eds.), *Research in*

the Sociology of Organizations (Vol. 38, pp. 1–22). Emerald Group Publishing Limited. https://doi.org/10.1108/S0733-558X(2013)0000038005

Freeman, C. (1995). The 'National System of Innovation' in historical perspective. *Cambridge Journal of Economics*, *19*(1), 5–24. https://doi.org/10.1093/oxfordjournals.cje.a035309

Future of Life. (2020, February). *National and International AI Strategies*. Future of Life Institute. https://futureoflife.org/national-international-ai-strategies/

Genus, A., & Coles, A. (2005). On Constructive Technology Assessment and Limitations on Public Participation in Technology Assessment. *Technology Analysis & Strategic Management*, *17*(4), 433–443. https://doi.org/10.1080/09537320500357251

Giraudeau, M. (2008). The Drafts of Strategy: Opening up Plans and their Uses. *Long Range Planning*, *41*(3), 291–308. https://doi.org/10.1016/j.lrp.2008.03.001

Goodsell, C. T. (1977). Bureaucratic Manipulation of Physical Symbols: An Empirical Study. *American Journal of Political Science*, *21*(1), 79–91. https://doi.org/10.2307/2110449

Goodsell, C. T. (2000). *American Statehouse: Interpreting Democracy's Temples* (Illustrated edition). University Press of Kansas.

Hao, K. (2019). *Yes, China is probably outspending the US in AI—but not on defense*. MIT Technology Review. https://www.technologyreview.com/2019/12/05/65019/china-us-ai-military-spending/

Hatry, H. P. (2002). Performance Measurement. *Public Performance & Management Review*, *25*(4), 352–358. https://doi.org/10.1080/15309576.2002.11643671

Hendrick, R. (2003). Strategic Planning Environment, Process, and Performance in Public Agencies: A Comparative Study of Departments in Milwaukee. *Journal of Public Administration Research and Theory*, *13*(4), 491–519. https://doi.org/10.1093/jpart/mug031

Kapetas, A. (2020, December 23). *The geopolitics of artificial intelligence*. The Strategist. https://www.aspistrategist.org.au/the-geopolitics-of-artificial-intelligence/

Lascoumes, P., & Gales, P. L. (2007). Introduction: Understanding Public Policy through Its Instruments—From the Nature of Instruments to the Sociology of Public Policy Instrumentation. *Governance*, *20*(1), 1–21. https://doi.org/10.1111/j.1468-0491.2007.00342.x

Mavlanova, T., Benbunan-Fich, R., & Koufaris, M. (2012). Signaling theory and information asymmetry in online commerce. *Information & Management*, *49*(5), 240–247. https://doi.org/10.1016/j.im.2012.05.004

Metcalfe, J. S. (1995). The economic foundations of technology policy / equilibrium and evolutionary perspectives. *Handbook of the Economics of Innovation and Technological Change*.

Miller, T., & Triana, M. D. C. (2009). Demographic Diversity in the Boardroom: Mediators of the Board Diversity–Firm Performance Relationship. *Journal of Management Studies*, *46*(5), 755–786. https://doi.org/10.1111/j.1467-6486.2009.00839.x

OECD. (2018). *A portrait of innovative start-ups across countries* (OECD Science, Technology and Industry Working Papers No. 2018/02; OECD Science, Technology and Industry Working Papers, Vol. 2018/02). https://doi.org/10.1787/f9ff02f4-en

O'Meara, S. (2019). AI researchers in China want to keep the global-sharing culture alive. *Nature*, *569*(7758), S33–S35. https://doi.org/10.1038/d41586-019-01681-x

Park, Y., Fiss, P., & Sawy, O. E. (2020). Theorizing the Multiplicity of Digital Phenomena: The Ecology of Configurations, Causal Recipes, and Guidelines for Applying QCA. *Management Information Systems Quarterly*, *44*(4), 1493–1520.

Poister, T. H. (2010). The Future of Strategic Planning in the Public Sector: Linking Strategic Management and Performance. *Public Administration Review*, *70*, s246–s254. https://doi.org/10.1111/j.1540-6210.2010.02284.x

Porter, A. L., Roessner, J. D., Jin, X.-Y., & Newman, N. C. (2002). Measuring national 'emerging technology' capabilities. *Science and Public Policy*, *29*(3), 189–200. https://doi.org/10.3152/147154302781781001

Raaphorst, N., & Van de Walle, S. (2018). *Trust, Fairness and Signalling: Studying the interaction between officials and citizens*.

Ragin, C. C. (2008). *Redesigning Social Inquiry: Fuzzy Sets and Beyond* (1st edition). *University of Chicago Press.
Rihoux, B., & Ragin, C. C. (2009). *Configurational Comparative Methods*. SAGE Publications Inc. https://us.sagepub.com/en-us/nam/configurational-comparative-methods/book230411
Rodrik, D. (1988). *Promises, Promises: Credible Policy Reform via Signaling* (No. w2600; p. w2600). National Bureau of Economic Research. https://doi.org/10.3386/w2600
Schmidt, E., & Allison, G. (2020, August 5). *Is China winning the AI race?* Australian Financial Review. https://www.afr.com/technology/is-china-winning-the-ai-race-20200805-p55imu
Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374. https://doi.org/10.2307/1882010
Spence, M. (1978). Job Market Signaling. In P. Diamond & M. Rothschild (Eds.), *Uncertainty in Economics* (pp. 281–306). Academic Press. https://doi.org/10.1016/B978-0-12-214850-7.50025-5
Spence, M. (2002). Signaling in Retrospect and the Informational Structure of Markets. *The American Economic Review*, *92*(3), 434–459.
Stine, D. D. (2009). *Science and Technology Policymaking: A Primer* (p. 42). Congressional Research Service.
The Economist. (2018, March 1). How the West got China wrong. *The Economist*. https://www.economist.com/leaders/2018/03/01/how-the-west-got-china-wrong
The Guardian. (2021, January 14). *Dutch government faces collapse over child benefits scandal*. The Guardian. http://www.theguardian.com/world/2021/jan/14/dutch-government-faces-collapse-over-child-benefits-scandal
Thiem, A., Baumgartner, M., & Bol, D. (2016). Still Lost in Translation! A Correction of Three Misunderstandings Between Configurational Comparativists and Regressional Analysts. *Comparative Political Studies*, *49*(6), 742–774. https://doi.org/10.1177/0010414014565892
Trueman, C. L. and C. (2019, March 12). *How Singapore is using AI*. CIO. https://www.cio.com/article/3292616/how-singapore-is-using-artificial-intelligence.html
Tsolakis, D. M., & Tsekouras, G. E. (2016). A Fuzzy-Soft Competitive Learning Approach for Grayscale Image Compression. In M. E. Celebi & K. Aydin (Eds.), *Unsupervised Learning Algorithms* (pp. 385–404). Springer International Publishing. https://doi.org/10.1007/978-3-319-24211-8_14
Wheatley, K. K., & Wilemon, D. (1999). From emerging technology to competitive advantage. *PICMET '99: Portland International Conference on Management of Engineering and Technology. Proceedings Vol-1: Book of Summaries (IEEE Cat. No.99CH36310)*, 35–41 vol.2. https://doi.org/10.1109/PICMET.1999.787785
Whittington, R., Molloy, E., Mayer, M., & Smith, A. (2006). Practices of Strategising/Organising: Broadening Strategy Work and Skills. *Long Range Planning*, *39*(6), 615–629. https://doi.org/10.1016/j.lrp.2006.10.004
World Economic Forum. (2017). Global Competitiveness Index 2017- 2018. *Global Competitiveness Index 2017-2018*. https://wef.ch/2L4l4G9
Zhang, Y., & Wiersema, M. F. (2009). Stock market reaction to CEO certification: The signaling role of CEO background. *Strategic Management Journal*, *30*(7), 693–710. https://doi.org/10.1002/smj.772