Artificial Intelligence and Public Policy: Disrupting Policy Sciences

Cambridge Elements on Public Policy

Cambridge University Press

Fernando Filgueiras

Associate professor at the School of Social Sciences at the Federal University of Goiás (UFG). Professor of the Doctorate in Public Policy at the National School of Public Administration (ENAP). Researcher at the National Council for Scientific and Technological Development (CNPq). Researcher at the National Institute of Science and Technology – Quality of Government. He is the Director of Innovation and Strategic Information at the Ministry of Education in Brazil.
ORCID: https://orcid.org/0000-0001-9570-8113
Email: fernandofilgueiras@ufg.br

**** This book is in proofreading and revision. Please, don't citate it****

Abstract

Based on the literature review, this element analyzes how artificial intelligence transforms and disrupts policy dynamics in three order issues: policy formulation and implementation, policy advisory and instrument constituencies. The impact of artificial intelligence comprises the entire policy process, changing various institutional dynamics, the behavior of actors, and the development of policy responses to societal problems. We start from the premise that AI is an epistemic instrument applied in general in the policy process. It alters the way knowledge is constructed and instills changes in the practical action of public policy. This element then discusses how the increasing use of artificial intelligence in public policy creates new challenges for governments to deal with the impacts and emerging issues. Scholars in public policy must consider elements of AI governance as essential in formulating and implementing public policies.

Keywords: artificial intelligence; governance; policy formulation; policy implementation; institutions; knowledge; learning

Contents

1. Introduction 4
2. Policy process, knowledge and AI 6
   2.1. Policy process and knowledge 6
   2.2. Epistemic changes and AI..9
3. AI as agents or epistemic instruments in the policy process? 13
   3.1. What is AI? 13
   3.2. Human-machine interactions and policy 19
   3.3. Risks, problems and catastrophic ambiguities 23
4. Modeling disruptive policy cycle with AI 27
   4.1. Instrument constituency and AI system design 35
   4.2. Policy advice in AI era..39
5. Governing AI in the policy cycle 43
   5.1. Ethical dilemmas and AI in policy cycle 43
   5.2. Humans in the loop and AI governance in the policy cycle 47
6. Concluding remarks..53
References 56

1. Introduction

The motivation of this book is to understand how and why AI provide a sociotechnical reengineering of public policy and changes the policy practice. In the same vein, this book analyzes the consequences and challenges of this disruptive changes infused by artificial intelligence in policy process and the requirements of new modes of governance. As a disruptive technology, artificial intelligence has a direct impact on the way public policy is implemented within government organizations. Although a controversial and poorly defined concept, artificial intelligence is an interdisciplinary field of research that involves computer science, engineering, mathematics, and a whole range of applied sciences. The objective is to analyze how artificial intelligence transforms policy science into a disruptive path that requires new standards of governance and control.

The pervasiveness of AI in the everyday life of contemporary society makes it disruptive in different sectors and areas. For instance, the way in which generative AI creates texts and images changes scientific practices, representing a contemporary challenge regarding how knowledge is created and under what rules. In many scientific fields, AI is a mechanism that increases the capacity to create materials, identify and edit genetic codes, produce projects, alter drug production systems, among many other fields. Likewise, the audiovisual industry is directly impacted by generative AI, changing production practices in cinema, music, and literature. The entire communication structure of society is being transformed through the personalization of information. Markets operate through various platforms and artificial intelligences that understand what consumers want when, where, and how. The pervasiveness of artificial intelligence makes it an instrument for reengineering society, including government practices. We assume here that artificial intelligence is a pervasive instrument for humanity to produce knowledge and transform the way we do things and perform actions in society. In other words, artificial intelligence is an instrument that transforms action in various fields of humanity.

Governments are engaged in this process of social reengineering, shifting all their sociotechnical instruments. The scope and speed with which artificial intelligence is becoming embedded in the public policy process is reforming the organizational and political structures by which government action is shaped to impact society. Governments are gradually adopting artificial intelligence into their organizational structures, producing silent, incremental, and effective reforms to produce changes in the structure of knowledge and government agency. Understanding how and why artificial intelligence produces changes at the roots of public policy becomes central to the constitution of policy science. The original imagery of the field of policy science is of a public policy based on knowledge and the constitution of a professional discipline within public organizations. At the present time, we are experiencing disruptive changes that imply new ways of doing, experimenting and reasoning to act in the public sphere. Making new sociotechnical instruments available, such as artificial intelligence, means producing new ways for humans to think and act to solve different public problems. We start from the premise that policy science is undergoing disruptive changes in the epistemic structure of the field, implying new patterns of action that organize new practices and new ways of understanding problems and solutions in public policy.

The book is divided into four more chapters. The following chapter produces the relationship between traditional knowledge and new emerging modes of knowledge with artificial intelligence. This chapter produces a general framework of change. The third chapter deals with the concept of artificial intelligence, its methodologies, and developments in policy cycle. We want to understand how AI modifies the policy process since it modifies decision-making and tasks. In the fourth chapter, we discuss how the introduction of AI into the policy process implies an overlapping of data and systems modeling across all stages of the policy cycle, creating new activities based on instrument constituencies and new practices of policy analysis and advisory. The fifth chapter summarizes the AI governance challenges that become central to the policy process. Finally, we conclude by pointing out how the changes introduced by AI in public policy transform the foundations of policy science and its dialogue with society.

2. Policy process, knowledge and AI

2.1. Policy process and knowledge

When Lasswell laid the foundations of policy science, he synthesized a scientific and professional perspective on this field of knowledge. The scientific perspective lies in the fact that public policy is problem-solving and that the connection between problems and solutions requires specific knowledge that mobilizes and encourages a particular pattern of decision-makers' agency in policy process. The action of policy actors is the empirical object of policy science so that the action dilemmas, perspectives, and outcomes can be understood and synthesized into theoretical knowledge. On the other hand, according to Lasswell, public policy involves a professional practice that mobilizes this knowledge in a practical way to build solutions. In summary, policy science is represented in the following statement by Lasswell: policy science is "knowledge of the policy process and the relevance of knowledge in the policy process" (Lasswell, 1970, p. 3).

Three characteristics of the field of policy science are essential. First, policy science deal with applied knowledge and it is oriented towards problem-solving. Secondly, this knowledge requires contextuality. Thirdly, the knowledge produced by policy science is interdisciplinary (De Leon, 1981). Understood as specialized, applied, and interdisciplinary knowledge, work with public policy involves the government's epistemic agency to solve problems. Placing the policy science as an epistemic agency means that policy decision-makers mobilize their action based on knowledge of problems and the connection of these problems with knowledge about solutions. Problems, represented through information, mean inputs for policymakers' actions, constituting a solution or outputs of the government agency in organizational contexts. Working with public policy involves agents understanding a problem and acting to build a solution and new knowledge about policy outcomes and dilemmas.

This perspective on policy science – knowledge and profession - like epistemic agency addresses different issues. We can suspect that the public policy cycle is a rational sequence of steps in which knowledge of the problems implies knowledge for and of the formulation, which in turn implies knowledge of and for the implementation and, finally, the generation of new knowledge through evaluation. Public policy embeds professional knowledge and constant learning to create solutions to public problems. A more skeptical conception of the public policy cycle might consider it a heuristic resource that organizes decision-making at different moments or stages of public policy. In all these stages, the policy cycle involves knowledge that motivates and shapes policymakers', bureaucrats', and societies' actions. The epistemic action of public policy professionals is shaped and infused by an interdisciplinary practical knowledge that enables them to reason and design solutions. Knowledge in policy science is essential because, in many situations, it defines how policymakers and managers select a purpose, create means of action, and diffuse policy objectives. The actors' actions lead to discussions about the theoretical framework of policy science. This theoretical framework leads to reflection on different models of understanding the construction of knowledge and its practical application by professionals from governments and interest groups.

Public policy is an applied, interdisciplinary, and professional-oriented knowledge field. Concerning professional aspects, it is necessary to consider that the public policy profession occurs in the organizational dimension. For example, the garbage can model criticizes the policy cycle's conception by pointing out that knowledge of solutions precedes knowledge of problems, shaping decision-making complex, infused by bounded rationality, and driven by fluid participation. Cohen, March, and Olsen (1972) recognize that decision-making, policy formulation, and implementation adhere to the phenomenon of organizations. They are the ones who fulfill the purpose of public policy and shape professional practice. Policymakers understand that, within organizations, public policy are artifacts to solve problems that involve professional practice within an organizational context and create public values (Barzelay, 2019).

The professional practice of public policy involves solving problems, which requires coordinated activity to mobilize several knowledge domains. Organizations make decisions, and policymakers design artifacts in their context to solve problems. Public policy is designed in organizational contexts, from which professionals engender a complex chain of interactions driven by analysis, prototyping, testing, evaluation, and decision-making (Howlett, 2019; Simon, 1970). In this dimension, public policy is a function of practical knowledge absorbed or produced within organizations, which shapes a professional discipline for policymakers.

As a function of knowledge, one of the essential tasks is how it is produced and influences the practical action of government professionals. Carol Weiss (1979) identifies seven ways in which knowledge produced by the social sciences influences public policy.

Firstly, *knowledge-drive* is how policies are treated as a research process. Secondly, *problem-solving* is when some scientific evidence is applied to an existing problem. Third, *interactive*, when scientific research is an informative resource for policymakers. Fourth, *political* is when evidence is used to support or certify political actors' preferences. Fifth, *tactical* is when evidence is used for political speeches against critics or to show that something should be done about a problem. Sixth, *enlightenment* is when evidence provides decision-makers with a means to make sense of world complexity. Finally, *social, intellectual enterprise* is when research responds to the moment's thoughts, fads, and fantasies. These are traditional ways in which knowledge influences the policy process, making it a contested political instrument, constituted to shape behaviors and practices throughout the policy process. This knowledge is organized into policy advice modes that disseminate this knowledge among policy decision-makers throughout the policy process.

These seven ways knowledge influences public policy, identified by Weiss (1979), pose an epistemic challenge for policy sciences. The epistemic challenge means understanding the public policy process. In other words, how knowledge of problems translates into knowledge of solutions to accomplish a purpose. For many analysts, the policy process is linear between problems and solutions, respecting rational policymaking and aimed at optimization (Braybrooke & Lindblom, 1963). On the other hand, there are those in which the policy process is non-linear, in which policies depend on political, ideational, or psychological factors since the idea of rationality is limited (Simon, 1947; Lindblom, 1959). Finally, there is an intermediate perspective of rationality. Knowledge is a factor that organize decisions rather than precisely makes rationality. Mixed scanning, for example, works with the idea that policy decisions are not linear, but that knowledge is a rationalizing factor for actors' strategies in different situations involving collective decision-making (Etzioni, 1968).

Knowledge mechanisms, their artifacts, and their usefulness in the policy process define many aspects of what we can call policy science. Conceived from a scientific and professional perspective, within organizations, public policy involves the interdisciplinary connection between knowledge and practice, thus shaping action within governments. Disruptions, however, make a new perspectives and changes that transform the knowledge production and practice.

2.2. Epistemic changes and AI

Regardless of types of policy knowledge and decision-making, the epistemology of policy science changes over time. Today, there is the idea that we live in a disruptive world in which the epistemic roots of policy science change radically due to new technologies (Hartley & Kuecker, 2022). This disruptive world is an interesting perspective to portray a moment in which the practice of public policy changes radically. Traditionally, policy advice and learning sources are consultancies, political parties, academia, and civil society organizations, which select, curate, and promote problems (Craft and Halligan, 2020; Craft and Howlett, 2013).

Currently, we are dealing with an epistemic change that occurs with the expansion of the volume of information available on different aspects of social life. The volume of information is expanding considerably due to the large volume of data collected, shared, and processed in everyday life. First, there is increased information and evidence about problems due to large volumes of data (Kitchin, 2014). Second, there is an expansion of knowledge sources that change the dynamics of policy advice (Safaei & Longo, 2024). In addition to consultancies, parties, academia, and civil society, there is an expansion in the volume of information, whether due to social media or expert systems that structure digital governments (Giest, 2017).

Although there are multiple perspectives on how knowledge influences policy decisions and tasks, the idea of having a greater volume of information has long been seen as a vital factor for policy rationalization and organization. Digital technologies have long played an essential role in the policy process. Since the foundations of policy science, computational systems have activated an imaginary of qualification and rationalization of the policy process through the dimension of professional technique. Lasswell's imagination, for example, is constructing a professional perspective of public policy mobilized by scientific knowledge, in which problems and evidence guide the work of policy analysts and can be instrumented by computational technologies (Lasswell, 1970).

In Lasswell's argument, the policy work, guided by "computerized" information, is essential in the scientific nature and in constructing a technical image of its professionals. The role that information plays in the policy process is part of a cybernetic conception of government action, in which actors understand the environment, react to information about the environment, make decisions, and establish value on the consequences of their decisions and actions implemented over time (Peters, 2012; Hood & Margetts, 2007). This cybernetic conception of government has been introduced previously and coincides with the emergence and construction of artificial intelligence (Deutsch, 1963; Samuel, 1962, Schwember, 1977).

This idea of strengthening the techniques and profession of policy analysts through digital technologies was reactivated by the possibility of using digital technologies throughout the public policy process. The use of blockchain to rearrange public organizations (Clifton & Pal, 2022), the Internet of Things to collect data and information in urban space (Chetfield and Reddick, 2019), the use of large volumes of data in the policy process (Giest, 2017), the use of robots in public health, industry, and transportation (Willems et. al., 2022) are examples of how digital technologies are becoming pervasive in the government domain. Digital technologies in government aim to strengthen policy science and create a new dynamic of more technical and evidence-oriented professional practices. This imagination became hype with the reinvention of artificial intelligence (Filgueiras, 2022).

Artificial intelligence promises a disruptive change in the practice of public policy. On the one hand, artificial intelligence reshapes the entire organization of the public sector, modifying the capacities of public agents, proposing a more vertical governance style of policy sectors and the horizontal allocation of power and functions between organizations through state integration, common capacity and needs-based joining-up of services (Dunleavy & Margetts, 2024). On the other hand, artificial intelligence provides a new epistemology and patterns of epistemic action (Coeckelbergh, 2023; Floridi, 2023), reflected in new analysis and new public policy practices by its professionals (Safaei & Longo, 2024). In other words, artificial intelligence allows policymakers to modify their knowledge and reasoning and think about government action in a complex and different world, with a new order of problems and new technologies to shape solutions.

For example, Open AI, an owner of ChatGPT, has a trained and customized policy advisor chat, which promises: (1) breaking down policy proposals, legislation, or regulations to assess their potential impacts, benefits, and drawbacks; (2) offer insights on how to approach policy challenges, engage stakeholders and navigate the policy-making process effectively; (3) identifying opportunities for innovation within public services and suggesting ways to implement new ideas; (4) helping to understand and interpret data relevant to policy issues, including economic indicators, public health statistics or educational outcomes; (5) providing examples of successful policies from other regions or sectors that could be adapted or learned.[1] This solution promises to replace policymakers' knowledge practice, constituting another advisory standard from which they learn about problems and solutions, create cognitive structures, and reason about their actions in the government organizational context. Furthermore, it changes the entire practice of policy advice and learning (Safaei & Longo, 2024; Henman, 2018).

Artificial intelligence that, in principle, can replace policy analysts, act as a mechanism for organizational rationalization in governments, or carry out tasks that enable policy implementation haunts the imagination. From a more realistic perspective, artificial intelligence is an instrument that assists policy analysts and policymakers. However, they are not just any instruments. AI represents a knowledge instrument that changes practices and interacts with humans to solve different problems. Thus, AI radically changes how humans learn from their experiences, sense context, reason their actions, or create cognitive structures. In public policy, AI goes beyond the conventional classification of policy instruments. AI is a pervasive instrument that can produce nodality, regulation, organization and define financial management mechanisms.

This book is dedicated to understanding the process of epistemic changes that result from the emergence of artificial intelligence and its application in policy process. AI has consequences in the professional practice of public policy and policy sciences, shifting the entire logic of policy knowledge. In many organizational domains, AI creates analytics that shape the actions of organizations with trust attributed by their members (Anthony, 2021). This generated knowledge changes the patterns of reason and action of policy decision-makers. We want to understand how these epistemic changes enable a new type of professional action for policymakers, bureaucrats, and analysts and how the changes introduced by AI change the way public policy is designed and implemented.

An illustrative example of this process is how the Intergovernmental Panel on Climate Change (IPCC) has successfully laid a baseline of scientific facts for global discussions on climate action. Based on multiple layers and AI algorithms, the dashboard modulates complex scientific information about climate change, infusing policymakers with a new knowledge structure that modifies policy practice (Cowls et al., 2023). AI applied at the IPCC provides climate modeling, prediction and simulations, and the design of catastrophic scenarios, as well as chatbots to expand the communication of IPCC data and reports and monitor government actions. In addition to providing information about

[1] https://community.openai.com/t/policy-advisor-custom-gpt-for-public-policy-professionals/702969/1

climate data, AI applied in climate change supports studies on past environmental change around displacement hotspots and delivers future projections to inform adaptation measures and anticipatory action for integration in humanitarian programming. Similarly, AI can support disaster prevention, tracking pollution, and building climate-resilient agrifood systems that are more efficient, sustainable, and adaptable to climate change challenges. Within the practices of climate policy, carried out by humans, data and information are modeled to predict and simulate scenarios that guide government action in innovative areas and outside the regular pattern of action of policymakers and interest groups. Different climate policy challenges are shaped by AI, implying new recommendations for practical action taken by policymakers and bureaucrats (Muccione et al., 2024).

In all these situations, predictions, simulations, and AI modeling modify policymakers' epistemic action in the context of climate policies. First, climate change policy requires global negotiation and outreach, strategic partnerships, and implementation processes beyond issues internal to the nation-state. Secondly, climate change policies require the creation of networks to adapt global infrastructure, intensely shifting production and consumption chains. Finally, climate change requires policymakers to implement mitigation actions, which involve direct organizational changes in bureaucratic agencies and the infusion of behavioral changes in society. Mitigation actions require new formats, knowledge, and instruments for policy design (Braunerhjelm & Hepburn, 2023).

These changes in the epistemic action of policymakers, encouraged by an AI-based system, mean that artificial intelligence is providing new sociotechnical standards for public policy and have consequences in society. AI creates a new context for knowledge production and action that infuses policy changes in dimensions of government and its institutions. Understanding what AI is and the framework needed to think about its consequences is essential in the composition of artificial intelligence in the policy process.

3. AI as agents or epistemic instruments in the policy process?

3.1. What is AI?

Artificial intelligence is not a thing or a singular technology. AI does not have a precise concept in specialized literature (Wang, 2019). The conceptual bases of AI are in the machine intelligence to perform a task or solve a problem, which can vary according to the concept of intelligence (Russell, 2019). AI is a computational technology that imitate human intelligence but can vary depending on how the intelligence is conceived. The philosophical assumption behind the development of AI is imitation (Turing, 1950).

It is important to emphasize that AI can be deconstructed in two dimensions. The first has to do with the difficult concept of intelligence. If machines imitate human intelligence, the first step is to define what intelligence means. The concept of intelligence focuses on human agency and understanding the drivers of action in different contexts to solve problems. Humans are intelligent not because they have intellectual abilities but because they can intervene to achieve a particular purpose. Intelligence can be many different things. Intelligence can mean the idea of rationality, considering the ability of humans to realize their preferences in each context and maximize the greatest possible utility (Simon, 1957). The concept of preferences is based on the premise that humans know perfectly well what they want and that they will choose an optimal option in many situations. Intelligence can also be considered bounded rationality, which implies that human action is not based on perfect preferences but on heuristics that constitute an informational shortcut that leads humans to decide on the most satisfactory option and not the optimal option (Kahneman, 2003). Intelligence can also mean the human capacity to learn and form cognitive structures that allow them to perform in the world (Russell & Norvig, 2010). Intelligence can also mean the capacity for abstraction. Abstraction is a crucial function that enables humans to achieve the performance of different tasks (Minsky, 1985). Intelligence can be structured thinking, in which humans understand, plan, and reason to achieve specific goals and solve problems (Markram, 2006).

It is inappropriate here to adequately define intelligence, but this concept creates many initial difficulties. Intelligence is a human capacity to act based on abstraction, logic, understanding, self-awareness, learning, emotional knowledge, reasoning, planning, creativity, critical thinking, and problem-solving. In all these definitions, intelligence means some practical knowledge aimed directly at action in the world through social interactions and exchanges (Goldman, 2003). Intelligence involves how humans infer information and construct knowledge to adapt their behavior to the environment through exchanges that exist in social interactions (Simon, 1983; Goldman, 2003). Intelligence presupposes a capacity that is built in everyone. However, it also has a collective dimension established in meanings, common actions, and shared political knowledge. The collective dimension of intelligence is, by definition, a political dimension (Landemore, 2012).

In the other place of artificial intelligence concept, we have the issue of the artificial. According to Simon (1970), this dimension of the artificial entails the idea that systems interact with humans to achieve a particular purpose. An essential premise for understanding AI is launched here. What is different about AI systems is that they are in permanent interaction with humans through computational interfaces designed for human affairs (Cross & Ramsey, 2021). When interacting with humans, machines incorporate different ranges of human problems to solve them and perform tasks (Reid & Gibert, 2022). Systems only have a particular form and behavior because they adapt to their environment to accomplish objectives or purposes in interaction with humans. Thus, human artifacts, in terms of their behavior, are artificial. Simon (1970) characterizes an artificial system as an interface between internal and external environments. These environments belong to the domain of "natural science," but the interface that connects them is the domain of "science of artificial," which comprises an interdisciplinary dynamic that converges different forms of knowledge for constructing a given artifact. When an artificial system successfully adapts, its behavior mainly shows the shape of the external environment and reveals little of the structure or mechanisms of the internal environment (Simon, 1970).

We can construct a general concept of artificial intelligence using these two dimensions. *Artificial intelligence is a computer system capable of operating, in interaction with humans, different methodologies to calculate, predict, and simulate courses of action from large volumes of data to support humans in decision-making and carrying out tasks.* Five elements are embedded in this definition, especially if we consider this definition for applications focused on public policy.

- Firstly, artificial intelligence depends on large amounts of data to generate analysis, learning, reasoning, and action (Kitchin, 2014; DeSouza, 2017; Dunleavy, 2016).

- Second, different AI methodologies start from the premise of imitating human intelligence performed by algorithms (Russell & Norvig, 2010). As we have seen, intelligence has many definitions, each of which implies different ways of thinking and reasoning to solve a problem.
- Third, AI systems depend on interfaces through which machines interact with humans (Simon, 1995).
- Fourth, artificial intelligence systems are epistemic when processing large volumes of data to support decision-making and task accomplishment (McCarthy, 1981).
- Fifth, the constitution of an AI system has a purpose; therefore, this system intends to affect agency issues directly and can carry out some action (Tomasello, 2023; Barandiaran, Di Paolo, and Rohde, 2009).

Because AI is many different things and is not a singular technology, the label artificial intelligence must be understood in its epistemic and not ontological dimension. Understood this way, AI is shaped by different methodologies and can be understood as a field of research and knowledge. Among the best known, machine learning is an AI methodology in which computer systems calculate a specific output from learning vectors calculated in a structured database for training. In other words, they are algorithms that allow the machine to learn from the training database and produce an output from the learned vectors (Domingos, 2015; Samuel, 1959). Algorithms shape this learning process, and it can be supervised or unsupervised. It is supervised when the machine learns from a human-categorized database. This structured database serves as supervision of the knowledge constituted by machines and guides – or supervises - how outputs will be constructed. The other possibility is that the outputs are unsupervised. In this case, the machines calculate the learning vectors autonomously and use them to create the desired outputs.

Ensemble learning is another methodology involving aggregating two or more learners (e.g., regression models, neural networks) to produce better predictions. In other words, an ensemble model combines several individual models to produce more accurate predictions than a single model alone. At times, sources may refer to this technique as committee-based learning. Ensemble learning rests on the principle that a collectivity of learners yields greater overall accuracy than an individual learner (Zhou, 2012).

Deep learning is another offshoot and subfield of machine learning. Using premises from neuroscience, deep learning are AI methodologies in which multilayered neural networks are applied to simulate the complex decision-making power of the human brain (LeCun, Bengio, Hinton, 2015). Finally, reinforcement learning focuses on decision-making by autonomous agents. It is mainly applied in robotics, where an autonomous agent is any system that can make decisions and act in response to its environment independent of direct instruction by a human user. The environment allows machines to collect and react to information, creating learning vectors (Sutton and Barto, 2018).

Within machine learning methodologies, there are different applications like natural language processing; within NLP techniques, there are Large Language Models (LLMs). LLMs are embedded in systems such as ChatGPT and Gemini, enabling machine interaction with humans. LLM is a computational language processing system designed to generate sequences of words, codes, or other data (more recently, images) from an input sequence called "prompt" and processing by transformer algorithms. These are algorithms trained using tokenization and fine-tuning to calculate sequences of words and construct texts or images (Eisenstein, 2019). With less hype currently, there are AI-based expert systems, which are software that uses AI to simulate the judgment and behavior of an organization that has expertise and experience in a given field (Krishnamoorthy & Rajeev, 1996, Hurley and Wallace, 1986).

Finally, there are AI systems based on computer vision methodologies, which use images or videos from different sources to train systems that perform various tasks that depend on image recognition. Facial recognition applications in public safety (Chen, Surette, and Shah, 2021) or autonomous cars, for instance, use computer vision techniques to perform tasks and decision-making (Padmaja et al., 2023). In all these situations, designing AI systems involves creating large databases, choosing algorithm architectures, and producing interfaces that interact with humans by collecting, storing, sharing, and processing data.

Machine learning represents different methodologies supported by different techniques, which implies that the design of an AI system is based on the choice of different algorithmic architectures driven by the achievement of a purpose. Table 1 below lists the main methodologies and techniques embedded in the machine learning package:

Table 1. Types of AI methodologies and algorithms

Methodologies	Algorithms
Supervised learning	- Classification: KNN, Logistic Regression, Naive Bayes, Decision Trees, SVM, Random Forest, GBM, Neural Networks. - Regression: Linear, Polynomial, Ridge, Lasso, SVR, Decision Trees, Random Forest, Gradient Boosting.
Unsupervised Learning	- Clustering: K-Means, DBSCAN, Hierarchical Clustering, GMM. - Dimensionality Reduction: PCA, SVD, ICA, t-SNE, LDA. - Association: Apriori, ECLAT, FP-Growth.
Ensemble Learning	- Bagging: Random Forest, Bootstrap Aggregating. - Boosting: AdaBoost, GBM, XGBoost, LightGBM, CatBoost. - Stacking: Stacked Generalization, Blending
Deep Learning / Neural Networks	- Feedforward Networks: MLP, CNN. - Recurrent Networks: LSTM, GRU. - Generative Models: GAN, VAE. - Specialized Networks: Transformers, Autoencoders, RBFN.
Reinforcement Learning	- Value-Based: Q-Learning, DQN. - Policy-Based: REINFORCE, PPO. - Model-Based: AlphaZero, Dyna-Q. - Other Algorithms: A3C, DDPG, TD3, SAC

Source: own ellaboration.

In short, AI systems are designed and implemented with large volumes of data, different algorithms, and computational power embedded in interfaces that shape human-machine interactions. Data represents numbers, images, texts, and any element from which we construct information, which can be collected, stored, and shared in critical infrastructures (Kitchin, 2014). Algorithms represent the "rules of the game" (Turing, 1950), from which courses of action are calculated and infused into the interaction between humans and machines (Mendonça, Filgueiras, and Almeida, 2023). In many situations, algorithms are the elements that institutionalize rules and procedures, shaping human behavior and choices through diverse interfaces. Finally, AI systems are produced from large machines that amplify computational power in data storage and processing. Computer power is about the computing infrastructure where data is collected and processed, and algorithms perform complex calculations. This infrastructure is typically shared by cloud models that can assume private, public, or hybrid ownership. This infrastructure defines many elements of the geopolitics of AI (Lehdonvirta, 2024; Rikap, 2021).

We can establish different applications in policy processes by considering the two dimensions of AI – artificial and intelligence. For example, facial recognition systems applied to implement public security actions, systems that perform administrative tasks automatically, policy evaluation and monitoring processes, knowledge of citizens' preferences, learning, and cognition on the part of policymakers to understand problems and propose solutions (Mendonça, Filgueiras and Almeida, 2023). In a way, everything that involves agency can be implemented using AI. This premise of agency applies to the field of public policy, but there are also several other fields of human action. This definition, therefore, involves the fact that AI represents agents or multi-agents that perform tasks and make decisions (Wang, 2019).

This first frame implies thinking of AI as agents in the policy process. Thinking of AI as agents makes some sense if we consider that these agents perform actions to make decisions and carry out tasks autonomously. Autonomy is one of the assumptions of agency and is constituted controversially. The tendency to attribute autonomy to AI systems disregards the fact that these computational systems do not have consciousness (Floridi, 2023). Autonomy presupposes consciousness and intentionality, which obscures the idea of AI as agents or multi-agents. This discussion is foundational in the field of AI. Since the classic work of Joseph Weizenbaum (1976), the idea of autonomy and the anthropomorphization of artificial intelligence has been discussed. By anthropomorphization, we understand the attribution of human qualities to machines, especially emotional abilities and intentionality. Claims to avoid anthropomorphization are recurrent, especially concerning LLMs and chats (Bender et al., 2021; Coeckelbergh, 2021; Shardlow & Przybila, 2023). The fact is that within this anthropomorphization process is how AI represents systems that interact with humans to solve problems. In this process of interaction, the tendency of anthropomorphization occurs and solidifies in the human imagination. This first frame arises from the idea that users of these computer systems attribute human qualities to the machines.

The other possible frame is to conceive these computational systems from an instrumental reason (Weizenbaum, 1976). Instruments such as computational systems provide new possibilities for the imaginative reconstruction of the world. Or, more specifically, in this book, the imaginative reconstruction of policy science (Lasswell, 1970). However, according to Weizenbaum, the way computational systems designers re-create the world implies the prominence of an instrumental reason. This frame conceives AI as an instrument for human action in different fields, constituting knowledge and modifying the nature of human action. In this second frame, we can conceive of AI in the policy process as an essential instrument for policy actors (Filgueiras, 2022). In the policy process, artificial intelligence is an instrument to perform regulatory, nodal, organizational, and financial functions, altering the capacities and skills of policymakers and bureaucrats in different policy domains. AI is an epistemic instrument primarily designed, developed, and deployed for use in epistemic contexts such as policy analysis and research. AI is specifically designed, developed, and deployed to manipulate epistemic content such as data, and it is applied to do so, particularly through epistemic operations such as prediction and analysis (Alvarado, 2023).

Regardless of which frame we are using - agents or instruments - artificial intelligence applied in public policy is a type of knowledge technology with different epistemic effects on how we think, reflect, understand, or reason policy problems and solutions through interactions between humans and machines. The disruptive point of using this technology in public policy is that it produces epistemic changes in policy practice, creates new dilemmas and problems, and changes organizational dynamics from the moment that practitioners and AI-based computer systems interact to solve problems, produce analysis and apply this knowledge to the everyday practice of government interventions in society. Understanding these interactions between humans and AI systems is the next component of the analysis undertaken here.

3.2. Human-machine interactions and policy

One of the main characteristics of AI is that humans and machines interact to create knowledge and act to achieve a purpose. The determining factor in the conception of artificial intelligence is that humans and machines interact and constitute a new form of intelligence, shaped from data and institutionalized through algorithms (Mendonça, Filgueiras, and Almeida, 2023). The attribution of purpose to computer-based artifacts derives from the simple fact that each human action shapes an immediate machine reaction (Turkle, 1984). Interactions between humans and machines imply an exchange of meanings based on how users instill a need and how the machine responds. The human-machine interactions create a new cognitive structure based on normative meanings, preferences, or opinions interchangeable in interactions and situated in a context for action. The postulate is that the capacity for human agency is constantly reconfigured based on interaction dynamics between humans and machines (Suchman, 2007) and institutionalized in response to action situations, frames, and algorithmically defined action scripts (Mendonça, Filgueiras and Almeida, 2023).

As an epistemic agent or tool, the contemporary world lives with hybrid intelligence, driven by humans interacting with machines to solve problems (Jarrahi, Lutz, and

Newlands, 2022). This characteristic of hybrid intelligence is how humans interact with systems. Hybrid intelligence is essential for understanding the place of AI in public policy. At the current stage of evolution, it is not possible to attribute agency to artificial intelligence systems because the system itself cannot establish intentionality (Floridi, 2023b). Intentionality is established by humans interacting with these systems. In this sense, artificial intelligence is not an agent, but an instrument or tool with special characteristics. It is a general-purpose instrument that can be used in different policy domains and as a different means to replace other instruments. AI can perform actions to exercise nodality, authority, treasury, and organization. Furthermore, AI transforms the way policy actors will act in the policy process. The character of a hybrid intelligence derives from the way artificial intelligence, thought of as an instrument, changes the perceptions, reasoning, analysis, and work of policy actors. The literature on policy instruments shows how they are not neutral because they produce specific effects regardless of the objectives pursued (Lascoumes and Le Galès, 2007). As a pervasive and general policy instrument, artificial intelligence produces changes in the way policies are formulated, decided, implemented and evaluated.

In different situations, humans interact with machines to design and build AI, and AI systems incorporate policy design and institutional dynamics to achieve policy objectives and augment public value. The disruptive point of AI for public policy is that these interactions between humans and machines change the dynamics of policy process, which are now carried out by other means, with different consequences and new epistemics for policy science. Interactions are constituted from meanings, and the sociological premise is that individuals interact by how they interpret these meanings, which they attribute based on norms (Blumer, 1986). These meanings are institutionalized through responses to action situations, frames, and action scripts that define motivations for human action based on algorithmically created knowledge (Mendonça, Filgueiras, and Almeida, 2023).

Interactions between humans and machines in the policy process will manifest in two dimensions. According to Daugherty and Wilson (2018), these dimensions involve human dynamics in designing an AI, on the one hand, and the way the AI assists humans, on the other hand. This interactive dynamic between humans and machines depends on human inputs so that machines react to them and produce reciprocal feedback. Interactions between humans and machines are at the heart of AI and have sparked different discussions about machine intelligence (Searle, 1980; Simon & Eisenstadt, 2002). Reciprocal feedback shape new formats of knowledge, altering human action in society. Specifically in the field of public policy, interactions between humans and machines change the bases of the epistemic action of the actors involved in the policy process.

AI in public policy are systems that react to inputs posed by decision-makers and bureaucrats. The meanings attributed to this interaction between humans and machines in the policy process are still unknown. However, some clues already exist in emerging experimental research. In the work by Alon-Barkat and Busuioc (2023), the adoption of artificial intelligence in public policy is driven by the views and stereotypes of bureaucrats and decision-makers when they receive policy advice automatically. In other words, the adoption of artificial intelligence is motivated by the ability of these systems to augment data and information and constitute stereotypical policy advice based on the group's biases (Alon-Barkat & Busuioc, 2023). The meanings of interactions between humans and machines in the policy process are created from human interactions driven by ideas, norms, preferences, and perspectives that reinforce the biases of decision-makers and implementers.

In the policy process, these dimensions occur in the way developers and policy actors establish system design dynamics, in which humans assist machines so that they perform their functions. On the other hand, it matters how AI creates political and bureaucratic policy dynamics assisted by artificial intelligence in policy formulation, implementation, and evaluation. In both situations, these interactions create normative, functional, and symbolic meanings that produce consequences for humans in the policy process and social behavior construction.

When AI is designed to be adopted in the policy process, it incorporates knowledge from different sources to instrumentalize public action, whether augmented or faster. The policy design requires internal dynamics from bureaucrats and policymakers regarding the choice of the mix of instruments and their disposal for coherence and consistency to achieve objectives (Capano & Howlett, 2020; Howlett, 2019; Howlett, Mukherjee and Rayner, 2018; Siddiki, 2020). An AI that is applied in public policy involves defining the scope and characteristics of the system, in which developers, bureaucrats, and policymakers interact to choose algorithm architectures, create databases, train systems, explain these systems, and evaluate their capacity and accuracy to fulfill policy functions, sustain the system throughout the policy process.

Designing an AI means that humans constitute systems to perform a purpose. In many situations, systems design resembles policy design, depending on complex human interactions to choose algorithmic architectures and their instrumental requirements.

System design means that AI systems can embody the entire policy design, perform functions as an implementing agent, or embody specific tasks by interacting with humans to solve problems. In this sense, the rationale for interactions between humans and machines is instrumental rationality to create AI systems to augment capacities (Weizenbaum, 1976). In the policy cycle, AI is deployed to augment policy capacities (Filgueiras, 2022).

In another direction, artificial intelligence interacts with policymakers and bureaucrats in the policy process (Alon-Barkat and Busuioc, 2023). AI increases the dynamics of policy advice through recommendation algorithms, selection and hierarchy of preferences, clustering of social groups, and production of prediction and simulation on policy problems and solutions (Mendonça, Filgueiras & Almeida, 2023). AI is an instrument that produces organizational rationalization, new ways of thinking, and beliefs about objectives and future policy solutions. It is a pervasive technology driven by society's desire to build a decision system through agents that can control the future and interact with humans differently, defining new modes of social action (Nowotny, 2021; Weizenbaum, 1976). Social action is the essential point of interactions between humans and machines in the policy process: *AI transforms how policy actors understand and reason about problems and solutions through opaque technologies that support the entire construction of knowledge in different policy domains*.

According to Daugherty and Wilson (2018), the interaction between humans and AI in everyday life unfolds into three dynamics. First, AI *amplifies* analytical capacities and decision-making abilities. While policy decision makers have limited rationality because they cannot process all available information and then decide, AI can look at everything that can be datafied and, thus, increasing the human capacity to understand problems and solutions (Simon, 1995). Second, AI makes it possible to *interact* with different audiences. In the policy process, policymakers and bureaucrats can interact with citizens, companies, and other stakeholders more comprehensively and are mediated by interfaces that amplify collaboration (Campion et al., 2022). Third, AI enables *the embodiment* of a robot that augments a human worker. With their sophisticated sensors, motors, and actuators, AI-enabled machines can recognize people and objects and work safely alongside humans in factories, warehouses, and laboratories. In the public sector, robotics is a future field involving automation and human interactions with robots that perform repetitive or complex tasks (Dunleavy & Margetts, 2024).

Considering these interactive dynamics between AI and policymakers, analysts and bureaucrats, some questions emerge on the horizon. AI has a totalizing perspective on information with the expectation of increasing and creating collaborative structures and embodied robots throughout the policy process. This idea is compelling in producing policy change, new modalities of policy learning and advice, new modes of institutionalization and governance, and new perspectives on working with public policy in an augmented way. Although there are many promises about AI in the policy process, doubts still need to be made about the capacity for practical improvements through technological change or whether new risks are emerging for old problems.

3.3. Risks, problems and catastrophic ambiguities

The epistemic changes infused into policy science by artificial intelligence reinforce the technocratic thinking of policy actors (Hartley & Kuecker, 2021). This technocratic thinking derives from convergence defined by data patterns empirically, with greater or lesser receptivity to data and evidence. This technocratic bias is further reinforced when algorithms shape discourses, distribute resources, and increase or reduce the visibility of people and groups. Data and algorithms reproduce self-fulfilling prophecies in many situations through feedback loops, which forge new solutions to old policy problems (Mendonça, Filgueiras, & Almeida, 2023). In this sense, using AI in the policy process is challenging precisely because it transforms the structure of knowledge, modifying the framing of problems, the knowledge of solutions, and the ways of intervening, governing, and producing changes in society. Artificial intelligence reinforces a governance style based on an epistocratic dynamic, which distrusts democracy as a problem-solving method.

The epistemic changes spread by AI occur through how it classifies, clusters, hierarchizes, correlates, produces causality, and performs tasks based on data. Following the logic of imitation, AI uses these procedures based on the philosophical premise that humans use these procedures to learn about things, people, and nature and, thus, make decisions and act. The problem is that we do not have complete information about the policy process, and the decision-making process is formed by the biases of the actors who participate in it. The result is that the policy process is formed by ideas, values, and opinions that sustain decisions and actions on a factual representation of the world rather than precisely on a background rationality (Béland, 2019). Political, social, cultural, and economic biases are very important in the composition of the decision-making process and policy action (Banuri, Dercon & Gauri, 2019; Lindblom, 1959). The fact that AI sustain its decisions on data does not mean that these decisions in the policy process are based on evidence, much less that it is a technical decision. Artificial intelligence is designed through trial and error, decision-maker biases, and ideational heuristics that turn technology into a political artifact. In this way, the use of AI in the policy process is closer to dynamics based on muddling through, requiring incremental decisions and constant learning (Cox, 2019).

AI, as an epistemic instrument, incorporates, through data, all these biases and conceptions of the world, causing policy advice and learning to occur in a way that discursively and rationally reinforces the biases of decision-makers (Eubanks, 2018; Mendonça, Filgueiras, & Almeida, 2023; Noble, 2018). This incorporation of biases occurs from data, where bureaucrats and policymakers define the algorithmic architectures and databases that will inform the decision-making and carry out tasks automatically. Training and operation databases for AI systems in public policy express, through documents and public records, the choices and decisions of bureaucrats and policymakers, making these systems optimizer of algorithmically institutionalized decision biases.

This dynamic design and use of AI creates a series of risks in the policy process. These risks have been enumerated in the literature. However, considering the specificity of public policy, some risks become essential. Thus, there are risks related to algorithmic injustice, organizational risks, and the possibility of a rogue AI. Algorithmic unfairness arises from the algorithmically institutionalized biases of AI systems, which are related to issues of gender, race, ethnicity, sexuality, and social groups. When algorithms reproduce racial bias, for instance, a person can see how certain conceptions of the world are assimilated in seemingly technical ways (Buolamwini, 2023; Noble, 2018). In many situations, AI algorithms reproduce biases or discrimination from the lack of diversity in the technological design process (Benjamin, 2019). As a result, sexism, racism, and other forms of discrimination are built into the machine-learning algorithms that underlie the technology behind many ‘intelligent’ systems that shape how we are characterized and advertised.

In many instances, these forms of discrimination stem from a computer industry comprised of few women, black people, or people of different sexual orientations, for example (Crawford, 2021). This issue of inclusion in the design of technologies applied in public policy is fundamental to avoid bias and ensure co-creation and transparency processes that enable correct application (Yuwono et al., 2024; Noble, 2018). On the other hand, algorithmic injustice arises from data and the low visibility that data gives to topics related to social justice. Data incorporates society's biases, and the extent to which it makes certain social groups visible or invisible is fundamental to the performance of AI in the policy process (Gitelman, 2013).

Organizational risks refer to the possibility of systems malfunctioning due to organizational failures. Simple bugs in an AI's reward function could cause it to misbehave. For example, OpenAI researchers accidentally modified a language model to produce "maximally bad output." Gain-of-function research — where researchers intentionally train a harmful AI to assess its risks — could expand the frontier of dangerous AI capabilities and create new organizational hazards. For example, in the pharmaceutical field, AI has a dual potential to discover drugs, depending on human control and regulation to be directed towards human well-being (Urbina et al., 2022). Public organizations such as the Federal Audit Court in Brazil use ChatGPT to optimize processes and perform repetitive tasks. Organizational failures and the absence of measures to control risks and avoid human errors produce catastrophic results for organizations. Public policy requires public organizations to build security measures and address organizational dynamics such as segregation of duties, internal controls, and systems security measures so that operations and policy tasks are unaffected. From a risk perspective, organizations demand actions in line with the principles of trustworthy AI and taking accountability to mitigate the risks (Curtis, Gillespie, & Lockey, 2023).

Finally, we have the risks of rogue AI. Rogue AI are dangerous and powerful AIs that would execute harmful goals, irrespective of whether the outcomes are intended by humans (Bengio, 2023). Rogue AIs imply the idea of intentionally using powerful artificial intelligence to produce social harm. The assumption is that there is a misalignment between the use of technology and the purposes for which it was designed. For example, there is the possibility of using deepfakes created with AI to influence elections on social media (Diakopoulos & Johnson, 2021), explore markets (Lin, 2017), or generate political polarization (Jacobs, 2024). Likewise, AI can be used to create lethal autonomous weapons that intensely identify human targets to eliminate them (Russell, 2022) or AI that is used to create bioweapons (Urbina et al., 2022). AI threatens humanity with catastrophic or even existential consequences in all these situations.

AI is not a neutral instrument or an absolute technique in all these risk situations. Humans designed and deployed it in many policy domains to accomplish diverse objectives based on data-driven knowledge. In this sense, AI is an ambiguous, dual and opaque technology in the policy process, with diverse consequences for society. Conceived as an agent or as a policy instrument, the possibility of optimizing the policy cycle is great in the sense of increasing information and the possibilities of organizational rationalization of governments and policy capacity to achieve political objectives. The ambiguity lies in the fact that the design of AI systems can take on different facets and nurture the fantasies and beliefs of policymakers and bureaucrats (Filgueiras, 2022) and, in the same way, produce ambiguous results for society. Applying AI in the public policy cycle means taking risks and new possibilities to increase policy capacities.

4. Modeling disruptive policy cycle with AI

The traditional view of policy sciences understands the policy cycle as a central element for understanding the various dynamics through which a policy is formed and implemented with outcomes for society. The public policy cycle is an analytical device for explaining and prescribing a policy (Howlett & Ramesh, 2003). Fundamentally, the policy cycle is recognized as different phases or layers that follow one another but not in an orderly or sequential manner. The literature considers the policy cycle a heuristic resource formed by identifying the policy problem, formulation, decision-making, implementation, and evaluation.

Throughout the policy cycle, there is a flow of information that converts into decision and implementation processes. In problem definition, for example, agenda-setting issues involve knowing the problems and, based on this knowledge, deciding and acting in the formulation processes (Cháques-Bonafont, Palau, & Baumgartner, 2015). In implementation, for instance, information is central for governments to carry out their interventions and establish knowledge about the relationships between organizations and citizens to achieve effective outcomes that produce social change (Peeters, Rentería, & Cejudo, 2023; Pressman & Wildavsky, 1984). Information plays a central role in the policy cycle and enables the constitution of the cycle as a broader policy system and its effects on the practical action of governments (Howlett & Ramesh, 2003; Baumgartner & Jones, 2015).

The policy cycle strengthens a perspective that supports policy work from policy analysis. Understanding the policy cycle provides a broader framework of policy sciences as an interdisciplinary, analytical and practical work (De Leon, 1981). At the heart of policy science is analysis and how it informs policy practice. From the policy analysis perspective, information is central to constituting government action. As noted previously, public policy is an epistemic action that relies on information to shape different decisions at different stages. Public policy manuals define different steps by which a policy is formulated and implemented. For example, Bardach (2012) defines the policy analysis practice as an agency that defines the problem, assembles evidence, constructs alternatives, selects criteria, designs outcomes, confronts trade-offs, decides, and storytelling. Information and various micro or macro decisions are central to all these actions and define the practical dynamics surrounding government activities.

Throughout the dynamics of the policy cycle, information and advice are central, and the practical dynamics by which policy is formulated and implemented are defined (Wilson, 2009). The knowledge and recommendations that emerge in public policy are humanly created and involve interests, opinions, and perspectives on problems and solutions (Lowi, 1964). Because it is based on interests, opinions, and perspectives, public policy involves complex interactions between actors based on exchanging information and meanings for government action. In many situations, as Bardach states, "[t]he problem-solving process—being a process of trial and error—is iterative, so you usually must repeat each of these steps, sometimes more than once" (Bardach, 2012, p. xvii).

The policy cycle conceived as an analytical device, places at its center the idea that the success of a public policy depends on analytical capacities. Analytical capacities, in turn, depend on the central idea of policy analysis, which is the method for structuring information and providing opportunities to define alternatives for policymakers. Analytical capacities involve individual, organizational, and systemic approaches. Individual capacities involve technical skills and knowledge about the substance of policy. Organizational capacities concern budget and human resources for organizations to accumulate and disseminate knowledge. Finally, systemic capacities concern high-quality educational and training institutions and opportunities for knowledge generation, mobilization, and use (Howlett, 2015). Considering this idea of capacities and the role of information in the policy cycle, we can state that all public policy activities depend on knowledge and recommendations (Wilson, 2009).

Policy analysis depends on the capacities to frame solutions, create action situations, and define scripts for action (Ostrom, 2005). Artificial intelligence disrupts this perspective on public policy. AI changes the entire dynamics of policy analysis and changes regarding information flows and knowledge production. In principle, AI still has great potential, but it has already been implemented recurrently throughout the policy cycle. There are two reasons why AI disrupts the policy cycle. AI increases the speed of analysis and dissemination of information (Valle-Cruz et al., 2020) and augments analytical capacities in the individual, organizational, and systemic dimensions (Veale & Brass, 2019). In agenda-setting, for instance, problem identification combines existing administrative data with more granular or dynamic data collected in social media, platforms, or distributed systems for decision-making. AI enables policy analysts to shape unstructured and unconventional data through text mining and natural language processing (Allahyari et al. 2017). Furthermore, the possibility of using large language models for policy analysis is disruptive and fast to produce intelligence and advice (Safaei & Longo, 2024, Logan, 2024). In policy formulation, AI could provide some simulations about policy to assess viability. It could also help to improve previous decisions with machine learning algorithms and to expand or interrelate government decisions in several governmental layers (Valle-Cruz & Sandoval-Almanzán, 2022). In policy implementation, AI enables organizational rationalization and optimization and increases the organizational capacity to produce effective deliveries for society and connect citizens and governments by public service. Associated with the use of robotics, AI makes it possible to increase capacities and the speed of deliveries in a more holistic way of the State and public administration (Dunleavy & Margetts, 2024).

The possibility of new tools throughout the policy cycle has provided a new paradigm for policy analysis and new work practices in policy formulation and implementation. With augmented and faster information, the actions of policy actors change radically due to changes in how policy is implemented with data. Big data and the incorporation of AI systems into the policy cycle modify procedural and substantive instruments, in turn altering the actions of policymakers and bureaucrats in the policy process (Giest, 2017). From the perspective of policy instruments, AI can be implemented as a regulatory instrument, automating various activities related to the use of government authority (Yeung, 2018). Likewise, AI can be a nodal instrument, modifying the relationship between citizens and government (Margetts & Dorobantu, 2019). AI can also shape the financial management of the budget and public resources, being a treasury instrument (Valle Cruz, Fernandez-Cortez & Gil-Garcia, 2022). Finally, AI is an entire organizational instrument, changing the institutional framework of the public service (Dunleavy & Margetts, 2024).

Introducing AI in the policy cycle disrupts how policies are formulated and implemented. Policy sciences have a new layer of complexity for policy analysis and work. Over the public policy cycle, based on human interactive processes, there is an AI system modeling work based on interactions between humans and machines (Janssen & Helbig, 2018). For all phases of the public policy cycle, AI systems modeling is used to change the entire policy production chain. The disruptive element is AI creating and accelerating knowledge about various elements of the policy cycle, with humans and machines interacting to create solutions, set up and review government action, or predict and simulate outcomes and impacts. A new layer of policy work is to model databases and AI systems that will produce optimization across decision-making and task execution in the different phases of the policy cycle, innovating all activities related to analytical capacities and administrative execution.

The idea of simulating and predicting a policy is not exactly a new field of knowledge in policy process. Policy modeling is an important field of knowledge in public policy, working with predictions and simulations to guide the formulation. Policy modeling can be defined as an academic or empirical research work, that is supported using different theories as well as quantitative or qualitative models and techniques, to analytically evaluate the past (causes) and future (effects) of any policy on society, anywhere and anytime (Estrada, 2011). Policy modeling is related to the policy simulation through computational techniques and calculations. However, it does not support the idea of artificial intelligence. As we said before, AI assumes that humans and machines interact to achieve a purpose. The concept of policy modeling does not require this interactive dynamic. With artificial intelligence focusing on how decision makers, bureaucrats and citizens interact with machines, modeling extends to modeling policy work, in which predictions and simulations produced by modeling AI systems to produce public decisions and task execution.

In a particular public policy, each element of the policy cycle is superimposed on an element of the modeling policy work. We understand the modeling policy work as the entire cycle of analytical actions necessary and sufficient for humans to model a public policy by designing and deploying artificial intelligence. Policy work is overlapped by the activity of modeling AI-based systems that will be employed in the entire policy process. AI is applied to make decisions and perform tasks necessary to understand public problems, formulate alternatives, implement organizational actions, and evaluate outcomes and impacts. The disruption produced by AI is the fact that, beyond policy work, the application of AI throughout the policy process initiates a new work based on the modeling of AI systems that will perform different purposes in a singular public policy. Innovations with AI applied to public policy mean a different way of doing policy analysis and implying the knowledge generated throughout the policy cycle.

The modeling policy work is the sociotechnical reengineering of the public policy cycle through human interactions - traditional policy cycle - and human-machine interactions - disruptive policy. The modeling policy work is part of a broader concept of sociotechnical reengineering inscribed in the use of digital technologies (Frischmann &

Selinger, 2018; Filgueiras, 2022) and the reinstitutionalization of politics and society through algorithms (Mendonça, Filgueiras & Almeida, 2023). The modeling policy work means that human-machine interactions shift the policy dynamics in two ways: first, the human activities to design AI systems to sustain a policy; secondly, how humans and machines interacts and actors' agency shifts by new dynamics of policy advice. Table 2 below summarizes, in comparison, the elements of traditional policy work and modeling policy work.

Table 2. Policy and Modeling Work in Policy Analysis

	Traditional policy work	Modeling policy work
Instrument constituency	Interested choices of different actors for the instrument and its retention	Modeling of data and abstract system elements
Policy advice	Knowledge of the problems and recommendations that emerge from consultancies, actors, lobbying, parties or civil society	Recommendation, simulation and predictive systems
Policy dynamics	Human interactions	Human-machine interactions
Governance styles	Political	Epistocratic

Source: own elaboration

Incorporating AI into the policy cycle does not mean reducing its importance as an analytical device. It continues to exist but with an overlapping layer because the entire policy process is traversed by the modeling cycle, based on the design and use of AI systems. The effects of AI on the policy cycle arise from the fact that humans and machines interact in the dimensions of system design and in the dimension of application. In other words, on the one hand, we have the constitution of AI as an instrument or mix of instruments applied throughout the policy cycle. The design of AI in the public policy cycle should be understood as an instrument constituency, involving several actors to design and deploy systems applied throughout decision-making and task execution. On the other hand, AI changes the nature of policy advice. AI systems, especially those based on prediction and simulation, as well as recommendation systems, change the logic of policy advice, shaping the actions of policymakers and bureaucrats throughout the policy cycle.

The result of incorporating AI into the policy cycle is the disruption of the policymaking. This disruption shape different styles of policy governance. Traditional policy work comprises different styles, which can be based on state command-and-control or the composition of policy networks and more horizontal modes involving civil society, nonprofit organizations, markets, and social groups (Klijn & Koppenjan, 2004; Peters & Pierre, 2016; Salamon, 2002; Sorensen & Torfing, 2005; Stoker, 1998). By adding modeling policy work, we mean transforming policy governance, including another style based on knowledge. Adding a layer of modeling policy work makes the policy cycle governed in an epistocratic dynamic. Epistocracy is the idea that good political decisions should be based on knowledge instead of an aggregation of opinions (Estlund, 2008). The epistocratic style stems from the idea that democracies are incapable of delivering good solutions and advocates epistocratic procedures to protect political communities from the rule of ignorance. Technical solutions would overcome political disputes, thus enabling a more efficient government (Brennan, 2016). This epistocratic style resonates with the idea that public policy based on AI systems are neutral and that the instruments are technically designed.

Incorporating AI into the policy cycle adds a layer to traditional policy work through systems modeling. In problem definition, traditional policy work involves using opinion polls, monitoring media, identifying issues in social groups (Baumgartner & Jones, 2015), gathering national mood, or gathering ideas from visible and invisible actors (Kingdon, 1995). With the introduction of AI for problem identification, governments can now collect data from multiple sources, including social media, automated media monitoring, and data and evidence collected during implementation processes. For instance, text mining techniques and using LLMs provide rapid ways of building agendas (Gyódi et al., 2023). Collecting data to identify policy problems requires new policy work driven by database modeling, data governance, and skills to feed AI systems.

At the formulation stage, interactions between actors are fundamental in constructing policy alternatives, involving visible and invisible actors who interact to formulate a policy (Kingdom, 1995), usually through trial and error and incremental decisions (Lindblom, 1959). The introduction of AI in the formulation stage shifts the dynamics of defining alternatives, providing means to simulate and predict results reliably (Ramezani, 2023). In both dimensions, decisions are made with information constructed in diverse ways. In traditional policy work, information emanates from actors, studies, benchmarks, and interests. In the modeling policy cycle, information emanates from technical work with data and the construction of systems. This technical work involves the conceptualization and design dynamics of systems, in which human agents define the scope and purpose of systems. Considering problem identification and policy formulation, decision-making contrasts two distinct decision dynamics. While traditional policy work demands information exchange, conflicts or consensus among actors, modeling policy work depends on purely technical decisions such as building training databases or designing systems (Coeckelbergh & Saetra, 2023).

Implementation links government purpose and the world of actions and outcomes. Implementation is a function of government decision, government management and oversight, and resulting execution by bureaucracy (Hill & Hupe, 2009). In implementation, the traditional dynamic implies the existence of organizational structures that provide the action necessary to achieve outcomes. In organizational structures, the choice of institutional architectures is essential to shaping the actions of bureaucrats and society (Olsen, 2006). Furthermore, it involves incorporating organizations from the private and nonprofit sectors. The core implementation challenge is governance and crafting implementation structures that deliver services for society (Imperial, 2021).

The choice of institutional architectures implies top-down implementation structures following hierarchical guidelines or bottom-up models, which start with society and networks to policy implementation. Top-down models frame implementation in command-and-control relationships, where implementation is the ability to achieve predicted consequences after initial conditions have been met, like legislation and funds (Pressman & Wildavsky, 1984). Bottom-up structures, on the other hand, identify the actors network involved in service delivery and incorporate their goals and strategies as part of the policy-making process (Hjern & Porter, 1981). Bottom-up models recognize the importance of street-level bureaucrats and the discretionary nature of their actions, where small decisions can change the course of policy implementation (Lipsky, 2010). Implementation by bottom-up structures is based on the idea of policy effectiveness and how they depend on decentralized authority relationships based on formal and informal institutions, such as expertise, skill, and proximity to essential tasks that an organization performs (Elmore, 1979).

Modeling policy work understands that implementation implies the existence of AI systems that operate as mechanisms for rationalizing organizations implementing policy. This dynamic means that the deployment of AI in policy implementation is driven by the institutionalization of services through digital platforms that modify discretionary relationships with citizens. The discretion of government action is transferred to platforms with AI systems that perform organizational tasks and make decisions that directly affect the citizens' lives (Mendonça, Filgueiras, & Almeida, 2023). In modeling policy work, the work of policymakers is not defined by authority relationships within bureaucracies or policy networks or with citizens. Policy work is conditioned by the choice of algorithmic architectures, validation and verification processes, the creation of experiments and prototypes, and the scaling of service provision to humans – in this case citizens – interacting with machines. This choice implies the creation of services implemented through AI systems that provide automation, speed, accuracy, and a low possibility of deviations (Valle-Cruz et al., 2020; Veale & Brass, 2019). Implementation through AI reinforces the technocratic character of digital governments, in which implementation decisions are made in a technical rather than political manner. The implementation process is now mediated by computer systems that collect data and provide services digitally. However, implementation through AI and platforms implies that policy decision-makers, data analysts, and developers interact to construct information, facilitate framing, define action situations, and adjust action scripts to political preferences. In other words, data analysts and developers act in predefined policy contexts and are essential to defining the political frameworks for government action (Van der Voort et al., 2019).

Finally, policy evaluation in traditional policy work is carried out with monitoring structures and data that allow the assessment of outcomes, effectiveness, and impact of policies and programs on society (Vedung, 2009). The objective of evaluations is to generate knowledge and recommendations for policymakers and bureaucrats to review action or reinforce instruments to achieve policy objectives (Weiss, 1998). Although there are political and institutional constraints for evaluation (Bovens, t'Hart, & Kuipers, 2008), policymakers and bureaucrats use the knowledge generated by evaluations to establish value on the policies implemented and learn about possible corrections. Evaluation is crafted for different uses and depends on analytical capacities developed within organizations (Pattyn & Brans, 2015). The use of AI in evaluation makes it possible to create real-time monitoring structures and to automate evaluations depending on the quality of the data generated. AI can automate evaluation with text mining and consistently analyze policy outcomes, effectiveness, and impact (Sun & Medaglia, 2019). In modeling policy work, the choice of algorithmic architectures and coding and data structuring work is critical to delivering evaluation automation consistently. As with implementation, policymakers, data analysts, and developers interact and define the policy uses of AI-powered evaluation.

In disruptive public policy, traditional policy work iterates with modeling work for data and AI systems all over the policy cycle. The disruptive element is how policymaking depends on active interactions between bureaucrats, policymakers, and system developers.

On the other hand, citizens interact passively with interfaces - platforms - that have embedded artificial intelligence making decisions and performing tasks. Developers become central actors in public policy designed in the context of digital governments. This dynamic provides essential institutional changes in the way public policy is framed discursively, designed institutionally, becomes path-dependent on data, creates algorithmic regulation, and transforms procedural and substantive instruments to achieve objectives (Mendonça, Filgueiras & Almeida, 2023). In this framework, there is the possibility of different ethical problems and varied risks, for which there are still no adequate governance. The use of AI systems in public policy, although they have the potential to expand evidence and knowledge, is dependent on political frames offered in the system design dimension (Newmann & Mintron, 2023; Van der Voort et al., 2019).

In the policy cycle, artificial intelligence produces epistemic changes directly affecting public policy work incorporating human-machine interactions as essential to the policy context. Epistemic changes require new capacities from policymakers and bureaucrats due to the interactions between traditional policy work and modeling policy work. The disruptive point is that policymakers and bureaucrats must consider computational modeling a new skill for public policy. Policy cycle and computational modeling interact in a complex way, producing substantial changes in policy analysis and public policy practice (Süsser et al., 2021). This situation of interaction between modeling policy work and traditional policy work has implications for policy science in terms of knowledge and action, supporting new dynamics within the policy process.

4.1. Instrument constituency and AI system design

The debate over whether AI is an instrument, or an agent is substantial in the field of Computer Science. In terms of public policy, the definition of artificial intelligence is as an epistemic instrument (Alvarado, 2023) that transforms the way policymakers, decision-makers, bureaucrats, and networks make decisions and act based on knowledge generated with data. As a central epistemic instrument in the policy cycle, its constitution and design are essential. In the dynamics of interactions between humans and machines, modeling policy work implies a dynamic design of systems that will constitute actions so that AI can influence problem identification, formulation, decision-making, implementation, and evaluation. More specifically, in human-machine interactions, we address the way in which humans constitute AI-based instruments (Daugherty and Wilson, 2018).

Understanding how different agents act to create AI in public policy can be better understood in the dynamics of instrument constituencies. Instrument constituencies are networks of different types of institutions and organizations who share a common interest in promoting a specific policy instrument and related practices for their own benefit in material terms or sharing ideas (Simons & Voß, 2018; Voß & Simons, 2014). In general, instrument constituencies bring together networks of scientists, design experts, consultants, public administrators, and technicians that design and deploy instruments in policy process. This actors' network has an interest in the development, retention, and expansion of the instrument and they works to institutionalize it in policy practice.

In the case of AI conceived as an instrument, the constituencies are formed by networks of the actors listed above plus developers in private companies that control the global communications infrastructure - big techs. Instrument constituencies act as a network in a policy context defined by the digital transformation package, with the potential to change all types of policy instruments (Margetts & Dorobantu, 2019; Yeung, 2018; Dunleavy & Margetts, 2024). In this policy context, public policy has their instrumental dynamics modified by the system design applied to automate and rationalize public organizations.

The constitution of AI-based policy instruments is not a neutral activity, much less a technically constructed one. The AI instrument constituency involves networks of scientists, policymakers, bureaucrats, industry, design experts, consultants and developers, who interact politically and are dedicated to the articulation and promotion of kinds of solutions regardless of problem context, with the aim of producing technologies. Basically, this network provides the encounter between the solution and the problem, mobilizing articulations around solutions in search of problems (Béland & Howlett, 2016). Instrument constituencies deploying AI as a pervasive instrument across the policy cycle and changing all policy instruments - nodality, authority, treasury and organization. Instrument constituencies change the configuration of government, being a point of deep hype about the AI potentials and challenges and the visible and invisible interests that permeate political articulations. The design of technologies applied in governments is politically motivated, given the economic interests in retaining technological instruments (Mendonça, Filgueiras & Almeida, 2023).

Instrument constituencies involve design AI systems from policy modeling, defining the abstract objectives of systems, conceptualization and choice of algorithmic architectures, database modeling, methodological definition, coding and assessment of accuracy and validation of systems for making decisions and carrying out policy tasks. AI instruments have political implications as machine learning algorithms mean "a way of gathering and ordering society's knowledge that fundamentally transforms how the state and society come to understand each other" (Amoore, 2022, p. 21). Furthermore, AI instrument constituencies act by monopolizing knowledge and turning data and information into commodified resources (Rikap, 2021). As a result, we now have programmable and codified institutions that change the behavior of policymakers in a broader context (Mendonça, Filgueiras & Almeida, 2023). According to Louise Amoore (2014), working with artificial intelligence requires a different type of ability from designers, more imaginative and intuitive, applied to governing society in an innovative way. AI instrument constituencies act by defining the steering of public policy, embedding all policy cycle in systems that perform policy interventions and impact society.

In other words, sociotechnical systems based on artificial intelligence are produced on an industrial scale based on algorithms embedded in data processing platforms, shared or marketed in the cloud, with ready-made architectures that can be customized for different problems, both in the public and private sectors. Scientists and developers promote the choice of algorithmic architectures based on trial and error concerning the problem, choosing the solution that presents the best accuracy and optimization (Amoore, 2022). The design dynamics of AI systems, in many ways, emulate the dynamics of policy design: elements of rationalization and optimization are sought to implement a politically shaped idea. If policy instruments are politically shaped, AI instruments are also politically constructed and guided by interests, perspectives, and opinions. Digital transformation is infused by an idea of austerity shaped in economic crises (Mendonça, Filgueiras & Almeida, 2023) and makes governments dependent on private cloud-based data infrastructures that give big techs an intellectual monopoly on political ideas (Rikap, 2021).

Machine learning algorithms are solutions looking for problems within broader organizational processes (Filgueiras, 2022). Although there are technical measurements of the accuracy of an algorithm, it does not allow checking the degree of adherence, coherence, and consistency when applied in the policy process. The technical accuracy of algorithms does not consider policy objectives, public values, justice criteria, or the solution's effectiveness because systems designers consider the accuracy of knowledge, not its epistemological status, aroused in confronting reality. AI-based policy modeling means that the network of instrument constituencies interacts politically with a view to achieving a goal. Industry, developers, and scientists benefit materially from the industrial development of AI, defining actions aimed at the permanence and retention of AI-based instruments and their use in public policy.

The permanence and retention of AI-based policy instruments means a struggle to adapt systems to policy contexts infused with broader political and institutional orientations based on interests and values. Ideas, therefore, count in the broader framework of technology design (Mendonça, Filgueiras & Almeida, 2023). In other words, AI design is a political struggle to realize an idea framed in discursively expressed political interests of instrument constituencies. In this case, AI design in public policy fits into techno-solutionism (Sætra, 2023; Paul, 2022). Techno-solutionism is the metaphysical power of advanced technology to transmute the universe's complex, indeterminate nature into obedient, mechanical certainty to be manipulated by the fantasies and fads of policymakers. The AI hype in public policy stems from techno-solutionist ideas that enable constituents to achieve retention of policy instruments.

For example, the digital welfare state is based on the premise of redesigning all welfare policies through digital instruments. Machine learning algorithms can perform different tasks related to the welfare state, promoting greater access for society (Coles-Kemp et al., 2020). However, they are often developed in line with a logic of control and dispositions around surveillance and efficiency which challenge careful engagements (Zakharova, Jarke & Kaun, 2024). For example, the Danish government built a surveillance behemoth, dedicated to increasing surveillance against citizens receiving welfare state benefits (Kayser-Bril, 2020). The case of Danish reforms between 2002 and 2019, public sector has entailed the transfer of responsibility for key infrastructure to private actors through digitalization. As Collington (2021) points out, the main objective of public sector digitalisation has rather been the growth of Denmark's nascent digital technology industries as part of the state's wider export-led growth strategy, adopted in response to functional pressures on the welfare state model. Reforms in Denmark were driven by fiscal stability and have produced a retrenchment of critical assets and capacities (Collington, 2021).

Instrument constituencies, therefore, are central to defining the frames of AI applied in public policy. The networks of actors that participate in the instrument constituency strive for the realization of ideas that create the frames from which AI technologies will be deployed to realize policy ideas. The network of instrument constituencies strives to define the objectives, data modeling, choice of algorithmic architectures, coding, and framing of system outcomes, to compose all or part of the policy design through AI. This incorporation of AI into policy design is done through trial and error, so that the policy analysis resulting from the use of AI confirms and disseminates the results expected by constituents and their partners in an opaque and unaccountable manner. In many situations, AI applied to public policy confirms the bias of policymakers (Alon-Barkat and Busuoic, 2023), in a network of business and control of central infrastructures for the development of opaque and unaccountable technologies.

4.2. Policy advice in AI era

Policy advice is the other side of the AI coin when applied to public policy. In the instrument constituency dynamic, we deal with how public policy actors and industry interact to build and feed AI systems. On the policy advice side, we deal with how policy actors receive AI outputs and recommendations like advice to act in policy practice. Policy advice is the set of activities that support policymakers' decisions by analyzing problems and connecting them with solutions and recommendations (Hallingan, 1995). Policy advice activities start from problems and define courses of action for decision-makers (Althaus, 2013). Policy advice encompasses the epistemic nature of policy science and applies to the entire policy cycle (Wilson, 2009). Thus, policy advice is a special type of policy work that connects problems and solutions. Furthermore, policy advice is a complex system that provides decision-makers with a political perspective through values and beliefs, on the one hand, and knowledge, through evidence and information (Veselý, 2017). Disruption in public policy arises from the fact that AI provides policy advice, both in a macro dimension of policy as a whole, and in a micro dimension, which involves small decisions and tasks.

Sharing information and knowledge about problems and formulating policy recommendations for action is the heart of policy advice and is disseminated throughout the policy cycle. Typically, policy advice is provided by multiple actors and multiple levels. Actors provide information, knowledge and recommendations for action to policymakers. These actors include both individuals and organizations. Among several actors, policy advice encompasses the actions of consultants, academics, scientists, third sector organizations, philanthropic organizations or international organizations. There are levels of policy advice that include the influence of the knowledge generated on the action of policymakers and bureaucrats, on the one hand, and the broader organizational action of the government (Veselý, 2017). Furthermore, policy advice relates more directly to policy capacities, especially those related to analytical capacities (Craft, Head & Howlett, 2024).

The practice of policy advice is fundamental in the construction of knowledge and in the conversion of this knowledge into policy action, both at the individual and organizational levels (Wilson, 2009; Craft, Head & Howlett, 2024). Policy advice builds knowledge through the analysis of outcomes, production of evidence, use of data in business intelligence, studies focused on a policy topic or recommendations that emanate from experts and scientists in a given field of knowledge. Recommendations, therefore, are central to policy advice and they connect problems with changes in policy work (Veselý, 2017). Modeling AI systems for the entire public policy cycle transforms the dynamics of policy advice. The heart of the use of AI is to transform the entire epistemic basis of public policy through systems that perform predictions and simulations quickly and reliably. In particular, the use of recommendation algorithms is fundamental to generate and disseminate content and change the course of action of policy actors. Algorithmic recommenders are systems aimed at generating meaningful recommendations for content or products that might interest a given set of users. The main function of algorithmic recommendation systems is to estimate a utility function that automatically and mathematically predicts, ranks, and presents the user's top preferences for a specific content or product (Schrage, 2020).

Recommendation algorithms are embedded in government platforms and produce content for policy analysts. Another alternative is create LLM that uses different techniques that enable the generation of knowledge and influence on action. For example, the Federal Court of Auditors in Brazil developed the ChatTCU tool, a chatbot based on the use of GPT-4 and the retrieval-augmented generation (RAG) technique to integrate all the specialized knowledge produced by auditors and system users in Brazil to create a chatbot that recommends or generate audit and policy evaluation content. This content recommendation instills a standardization of knowledge among auditors and a radical change in the auditing and monitoring of public policies implemented by the Brazilian federal government (Silva et. al, 2024). The use of artificial intelligence means a new interactive dynamic of knowledge production and infusion of action. Policy advice dynamics go beyond human interactions and incorporate relationships between humans and machines to shape policy content. The example of ChatTCU in Brazil means that knowledge emanating from public policy monitoring and accountability processes disseminates a new type of public action driven by interaction with chatbots.

The use of LLM has the potential to produce automatic policy advisers, increasing the capacity of science to support the practice of public policy. Artificial intelligence has the potential to produce evidence syntheses in all fields of knowledge such as medicine and health (Nowak, 2022), environment (Wani et. al, 2024), education (Ifenthaler et al., 2024), for example. Artificial intelligence has interacted with scientists and experts and radically changed the way science operates, creating difficulties and disruption in the construction of knowledge. Large language models and other artificial intelligence systems could be excellent at synthesizing scientific evidence for policymakers. However, this use requires appropriate safeguards and humans in the loop (Tyler et al., 2023).

The use of AI shifts the logic of policy advice, increasing and accelerating knowledge and modifying the practice of policy work. This generated knowledge modifies both individual actions and the organizational framework and institutional dynamics that are transformed by AI. AI will not replace policymakers, but it can enable a comprehensive, faster, and more efficient approach to policymaking in the short run and different way. The premise is that in the dynamics of the policy process, humans and machines will interact to generate new forms of policy advice for recommending and shaping government action. Assuming the inherent risks, policy makers, bureaucrats, lobbyists, consultants, members of civil society organizations, and citizens will interact with artificial intelligence to synthesize evidence, understand the problems, propose solutions, evaluate alternatives, and model the entire institutional and organizational architecture to implement a policy.

This dynamic occurs in a broader dimension of policy advice. But it also occurs in a micro dimension, in which AI creates evidence, information, and knowledge that influences the policy practice among street-level bureaucrats. For example, the use of predictive policing in security policies radically changes the way security agents act in society. Predictive policing involves the collection of a broad variety of data to estimate, through several correlations, when and where crime is likely to occur, thereby more efficiently employing the existing resources to avoid it (O'Neil, 2016; Meijer & Wessels, 2019). Usually, predictions are made via machine-learning advances and the construction of artificial neural networks that shape police action. In public security, AI advice micro actions fundamental to the policy implementation, which change the relationships between governments and citizens.

AI transforms the entire logic of policy advice, both in a macro sense, in the dynamics of formulation, and in a micro sense, shifts the implementation actions. Considering the way instrument constituencies act in the design of AI instruments, their deployment as an essential knowledge structure in implementation changes the entire structure of political decision-making and the construction of public action in society. Transforming the structure of policy advice means having AI instruments that change the knowledge dynamics of formulators and implementers, changing, in turn, the outcomes of interventions and the impact on society.

5. Governing AI in the policy cycle

5.1. Ethical dilemmas and AI in policy cycle

The advancement of digital technologies tends to create an imaginary of greater political neutrality and accuracy of public decisions based on data (Esko & Koulu, 2023). The use of artificial intelligence in public policy is surrounded by an imaginary of policies implemented in a faster, more effective, safer and more neutral way, capable to augment productivity of public service and benefits for society, which would be exempt from "naturally" human flaws. This imaginary creates a tech-solutionist frame for AI in public policy, without paying attention to the social and political dilemmas that emerge. For example, in 2019 the Dutch government's tax authority used an AI based on a machine learning algorithm to create risk profiles to spot fraud among people applying for childcare benefits. In practice, this AI would decide - not autonomously - which families would be eligible or not to receive the benefit. Authorities penalized families over a mere suspicion of fraud based on the system's risk indicators. Tens of thousands of families were pushed into poverty because of exorbitant debts to the tax agency. Reports indicate that some victims committed suicide. More than a thousand children were taken into foster care due to the scandal (Newmann & Mintrom, 2023).

This case of the Dutch government demonstrates how AI can incur ethical dilemmas from the perspective of its use and application in the policy process. This techno-solutionist imaginary provides a frame of opaque technologies that infuse human actions by defining individual and collective choices. Within human-machine interactions, these frames are discursively constructed and are based on ideas and values that make AI-based instruments ambiguous and powerful in defining action. The Dutch government’s goal was to reinforce austerity policies and improve the provision of public benefits with the moral duty of fairness and honesty. It is not possible to discuss here whether the Dutch government acted well or badly. The fact is that the use of technologies is driven by discursive frames that instill values and norms into the use of technologies (Mendonça, Filgueiras & Almeida, 2023). The sociotechnical reengineering of public policy has a series of implications for society and the political system, bringing to the center of the debate on the use of AI a complex set of social dilemmas, normally discussed from a risk perspective. A social dilemma is a situation of interdependence between people in which there is conflict between doing what is best for oneself and doing what is best for the group. The social dilemmas of AI in public policy emerges from human-machine interactions and the mode of how algorithms institutionalize policy practices and knowledge. By delegating the task of solving problems and making decisions to an AI, a false image is created that social dilemmas in public policies have been overcome. Sociotechnical reengineering of public policy with AI produces social dilemmas due to ethical issues.

The first dilemma is epistemological. Artificial intelligence introduces a dilemma related to its predictive power and how it resonates with human action. The deployment of AI in policy cycle has the objective to predict and simulate all decision making and tasks. According to Nowotny, we use artificial intelligence to increase our control over the future and uncertainty, while the performativity of AI, the power it has to make us act in the way it predicts, reduces our agency over the future (Nowotny, 2021). The way policymakers delegate to artificial intelligence the power to decide and perform tasks reduces the margin of human control over technology. AI creates the sense that public policy governance is data-driven and neutral. However, algorithmic governance does not supplant democracy because of an epistemic impossibility. Even if AI were to exercise algorithmic governance, it would not supplant democracy because humans continue to construct data and information that feed decision-making (Innerarity, 2024).

The spectrum of uncertainties in public policy and complexity means that defining problems and constructing solutions depends on a new type of knowledge, which is comprehensive and uncertain, yet granular and focused on individuals. On the other hand, the modeling policy work based on AI imposes many challenges for governments in designing and implementing effective policies to govern AI. We do not fully understand the problems posed by AI, which makes the technology itself unpredictable, intractable, and nonlinear, making it very difficult for governments to create an institutional framework and correct objectives for their policies (Gasser & Almeida, 2017).

The second dilemma is control technology. Not knowing the problems and uncertainties related to AI makes the paths forward ambiguous, in which we expect greater knowledge with the deployment of AI in public policy, while at the same time we do not know its dynamics for producing knowledge. Social dilemmas arise in the way in which the framing of AI in public policy implies interventions in human life, without us having control over it (Russell, 2019). AI significantly reduces human control over their decisions, creating new challenges for ascribing responsibility and legal liability for the harms imposed by AI on others. Artificial intelligence is a technology that learns and adapts to the environment by following the rules set out in algorithms, without humans being able to control the results and impose responsibility on systems. The unpredictability of machine learning based decisions implies that many erroneous decisions made by AI are beyond the control of and cannot be anticipated by society (Lim & Taeihagh, 2019).

The ethical challenge of AI in public policy concerns the extent to which governments can construct interventions in society, embracing people's lives, using technologies whose decisions are unpredictable and uncertain. These decisions are subject to diverse forms of algorithmic injustice (Eubanks, 2018), invisibility of identities or visibility of prejudice (Noble, 2018), or even policy failures that cause various harms in society. The Dutch case has been presented as an example in which the harms caused by AI create ethical dilemmas and require innovations in governance processes that go beyond traditional patterns in policy theory. The ethical dilemma of the Dutch case lies in the fact that wanting to do the right thing with AI - regardless of what we think is right or wrong - produces unforeseen effects on society, making it necessary to have a governance framework that is also uncertain and experimental.

The epistemological dilemma posed by Nowotny (2021) extends into uncertainties and the way in which delegating decisions to an AI system produces consequences for society. According to Floridi, the world of artificial intelligence produces a divorce between agency and intelligence. Epistemologically, AI as a new form of agency can be harnessed ethically and unethically (Floridi, 2023b). When AI is applied throughout the policy process, this divorce is amplified. On the one hand, it produces epistemic changes in policy work, in which policymakers, bureaucrats, citizens and corporations delegate decision-making and task performance to AI systems, changing the entire policy advisory system and forging actors' actions shaped without a properly human intelligence, but rather artificial. Within the interactions between humans and machines in the policy process, AI as a new form of agency can be modeled and deployed ethically and unethically. In this sense, faced with the ethical dilemmas of AI, we need governance frameworks that make it possible to instill practices and procedures that ensure the possibility of ethical AI development.

The AI governance in public policy starts from principles that frame AI development and deployment to governments and companies, creating a discursive and normative frame. AI ethics require principles that guide innovation and technological deployment. AI ethics do not have a consensus about the principles. However, international organizations like

OECD discloses and disseminates principles. The AI principles from OECD include: (1) inclusive growth, sustainable development and well-being; (2) human rights and democratic values, including fairness and privacy; (3) transparency and explainability; (4) robustness, security and safety; (5) accountability. AI principles is not "one size fits all" approach, but an action guideline that frame institutional arrangements to governance. Similarly, organizations such as UNESCO advocate an ethical perspective on AI that is grounded in strengthening democracy. The cornerstone of UNESCO's Recommendation on the Ethics of Artificial Intelligence is the advancement of fundamental principles such as transparency and fairness, while always remembering the importance of human oversight of AI systems. The principles formulated by UNESCO comprise a framework for AI development aimed at:

- Proportionality and do no harm: the use of AI systems should not go beyond what is necessary to achieve a legitimate objective. Risk assessment should be used to avoid harm that may result from such uses.
- Safety and security: unintended harm (security risks) as well as vulnerabilities to attacks should be prevented and addressed by AI actors.
- Right to privacy and data protection: privacy should be protected and promoted throughout the AI lifecycle. Adequate data protection frameworks should also be established.
- Collaborative, adaptive and multi-stakeholder governance: international law and national sovereignty should be respected in the use of data. Furthermore, multi-stakeholder engagement is necessary for inclusive approaches to AI governance.
- Responsibility and accountability: AI systems should be auditable and traceable. There should be oversight, impact assessment, auditing and due diligence mechanisms in place to avoid conflicts with human rights standards and threats to environmental well-being.
- Transparency and explainability: The ethical deployment of AI systems depends on their transparency and explainability (T&E). The level of T&E should be appropriate to the context, as there may be tensions between T&E and other principles such as privacy, safety and security.
- Human oversight and determination: Member States should ensure that AI systems do not displace ultimate human responsibility and accountability.

- Sustainability: AI technologies should be assessed against their impacts on sustainability, understood as a set of evolving goals, including those set out in the UN Sustainable Development Goals.
- Awareness and literacy: Public understanding of AI and data should be promoted through open and accessible education, civic engagement, digital skills training and AI ethics, and media and information literacy.
- Fairness and non-discrimination: AI actors should promote social justice, impartiality, and non-discrimination, while adopting an inclusive approach to ensure that the benefits of AI are accessible to all.

From this frame, AI ethics guide the reflection of human-machine interactions, in which the discussion on moral status of AI is essential. The AI ethics framework addresses the moral implications of AI to society, assessing the moral status of AI instruments and its social, economic, cultural and political effects. The principles guide human action to design AI systems and the working of this systems in human-machine interactions. The main claim of AI ethics is that humans create mechanisms and instruments to control AI. This claim is appropriate to public policy, in which the technological deployment is challenging and disruptive. The main claim is translating these principles in practical action to create governance instruments to AI in public policy. In the same way, create mechanisms that assign humans in control, specially in public policy. Maintain humans in the loop of AI deployment in policy cycle is the main challenge in public policy, concerning the mode of knowledge created by AI shapes human behavior and government action.

5.2. Humans in the loop and AI governance in the policy cycle

Essentially, humans in the loop approach reframes an automation problem from human-machines interactions (Amershi et al., 2014). Reframe an automation problem means that humans create controls on AI to calibrate outcomes and to involve humans in system design. The application of the humans in the loop approach perspective in public policy is that humans calibrate AI instruments by exercising supervision and controlling the flow of input data and evaluate the outcomes and impacts of the output data. The heart of human in the loop approach is human-machine interactions. As interactions is dependent from means construction and sharing by individuals, the social relationships is infused by means exchange and action (Blumer, 1986). The means exchange is sustained by reciprocal feedback that organize action in social and political landscapes. Humans in the loop is an approach that instrumentalize by governance the calibrations of outcomes and design and redesign of the AI systems. In AI-shaped public policy, the human in the loop approach means that humans can calibrate both the design of systems and the constituency of AI-based policy instruments, on the one hand, and the outcomes of systems that reorganize policy advice. This means that, on the one hand, human in the loop approach meet the instrument constituencies, requiring compliance with procedures like transparency, open training databases or privacy and data protection. In the other hand, human in the loop approach meet the systems outcomes, providing procedures that require algorithmic audits, system validation or regulatory sandboxes.

Humans in the loop approach demands governance instruments which are experimental and emerging in policy landscape. Calibrations from human feedback are essential to develop AI in public policy, as AI systems are based in machine learning algorithms which adapt to their users and environment. Guided by the AI principles, humans can calibrate outcomes and review the entire dynamics of AI systems in policy landscape. Human in the loop and AI governance emerges from AI regulations delivered recently by governments. The European Union AI Act, for example, require that AI designers to allow human control to achieve effective human oversight. Under article 14, AI systems should be designed in a path that they can be overseen by people in the AI lifecycle. The objective is formulating policies to compel AI designers to integrate human control function as part of safeguard against AI risks and malfunctions.

AI governance is framed by human in the loop approach. Create human control on AI is essential to align technologies and human objectives. Three challenges are foundational to AI governance particularly applied in public policy cycle. First, information asymmetries between people and instrument constituencies. Instrument constituencies are involved in AI development and deployment in a policy context shaped by information asymmetries. Second, the lack of normative consensus increases the complexity of technology control and obscures the AI potential. Thirdly, the government mismatch to the design of effective, efficient, and legitimate means (strategies, approaches, tools, and so forth) to resolve the substantive issues, concerning the conditions of uncertainty and complexity in the AI ecosystem. Governments has failed to design policies to AI because limits on traditional approaches to law and policymaking in the digital age (Gasser & Almeida, 2017).

AI governance is structured in layers, with different problems and approaches. A technical, ethical, and social and legal layers put in AI landscape the challenges to institutionalize governance for emerging technologies deployed in the policy cycle. The main respect of AI governance is two directions: in the one hand, AI is a substantial instrument in policy cycle applied in policy formulation, implementation and evaluation; in the other hand, AI governance require a set of procedural instruments that steering the mode of constituencies, adaptation to institutional framework, alignment with ethical principles, accountability, transparency, and technical requirements. Procedural instruments, typical in AI governance, is used to indirectly but significantly affect policy processes and outcomes (Bali et al., 2021). In the different layers that set-up AI governance, all instruments are procedurals, with the objective to steering the actions and organizations to governing emerging technologies. The table 3, below, produce a synthesis of procedures concerning all layers of AI governance.

Table 3. Layers and procedural instruments of AI governance

Layer of AI governance	Topic	Procedural instruments
Ethical layer	Principles	Code of ethics
	People	Courses and training for public servants Guides for developing inclusive AI National AI Strategy
	Constituencies	Centers of excellence and innovation in AI Common development framework Cooperation with universities Coordination Political support Transparency
Technical layer	Data governance	Data assessment and exclusion in public databases Data collection Data qualification Data sharing Interoperability Open data policy Opening system training databases
	Algorithmic accountability	Responsibility ecosystems AI oversight and control agency Algorithmic audits Algorithmic risk allocations Public algorithm registries
	Standards	AI development toolkits Cybersecurity infrastructures Standardization of algorithm selection. Standardization of technical documentation System validation Training database standards
Social & legal layer	Norms	Principles Rights
	Regulation	Periodic assessment of AI systems Regulatory sandboxes Privacy and data protection Intellectual property

Source: Gasser & Almeida and own elaboration.

AI governance in public policy is about procedures concerning ethical, technical and regulatory requirements to assign humans in the loop with technologies that make decisions and carry out tasks in policy process. In other words, AI governance in public policy is about define who can do where, how and when with requirements to action to align artificial intelligence with government objectives, democracy, and public values (Innerarity, 2024; Korinek & Balwit, 2023). The integration of main layers of AI governance is essential to align technology development and social good (Wirtz, Weyerer & Sturm, 2020). The procedural logic of AI governance is pushed by necessity to maintain humans in the loop with AI applied in policy process. The main objective is designing an institutional framework that integrate the ethical, technical and social and regulatory layers

(Gasser & Almeida, 2017) delivering procedural instruments that create an logic of appropriateness for AI development (Filgueiras, 2022). Between the procedural instruments for AI governance, table 3 enumerates the main instruments applied to AI development and deployment. Data governance procedural instruments, like interoperability, data sharing, data collection and qualification are essential to AI in public policy. In the same vein, regulatory procedural instruments like sandboxes or privacy and protection are essential to align AI with social goods. Finally, define ethical principles and codes of conduct for AI designers and developers are essential to create a public value perspective.

The challenge of AI governance is that the procedural instruments are beyond the government. AI governance landscape require approaches that are beyond the power of governments to regulate and define procedures. Another alignment in AI governance is a global digital ecosystem with local needs and institutions. For example, diverse artificial intelligence is designed and deployed in cloud systems, without government capacities to regulate and impose control (Filgueiras & Almeida, 2021). In the AI governance, the institutional framework to define rules and procedures to develop and deploy AI systems is labelled in a multistakeholder perspective. From this perspective, the objective is to couple global and local institutions with emerging practices to govern AI. The challenge is that humans supervise decisions and tasks performed by AI and create a perspective that it evolute and create knowledge from human feedback framed in data, controls and regulations. Humans in the loop, yet, involve the micro and macro institutions like Table 4 presents:

Table 4. Layers and levels of AI governance

Level	Layers		
	Ethical	**Technical**	**Social & regulatory**
Micro	Principles that shape AI development	Reciprocal feedback between humans and machines	Humans supervise
Macro	Frameworks to governance	Standards and rules	Norms and rights

Source: own elaboration

The macro level concerns more global and systemic norms and practices, which require multi-stakeholder constructions to define technical standards, internationally shared norms and rights, and frameworks for governance. At the micro level, AI governance deals with the institutionalization of principles and how they translate into practical action. Similarly, at the micro level, it deals with how reciprocal feedback between humans and machines will be processed and how humans supervise the work of systems. The micro level deals with the institutionalization of practices related to compliance with fundamental procedures that need to be performed by developers, companies, governments and other organizations.

In the human-machine interactions, AI governance is essential to policy process. The challenge is how to govern disruptive technologies with large uncertainty and ambiguities. AI governance require a policy perspective based on procedural instruments to align technological development in policy process and the outcomes in terms of policy knowledge, practices, ideas, analyses, and work. This challenge is strange to create a policy environment adapted to disruptive AI technologies applied in policy process. Policymakers and bureaucrats have limited knowledge concerning how it works and why, and what are the possible applications and consequences of its deployment. Furthermore, policymakers and bureaucrats have a policy environment shaped by uncertainty an structural power dynamics framed by big techs (Taeihagh, Ramesh & Howlett, 2021). AI in policy cycle and structural power dynamics create a governance context based in uncertainty and difficult to governments to design policies and regulatory perspectives.

6. Concluding remarks

The development and deployment of AI in public policy is quite broad and encompasses the entire policy cycle and can be a general-purpose instrument. The condition for AI to make decisions or perform some policy tasks is to have large volumes of data, computing power, human developers and interfaces that enable constant interaction between humans and machines to perform actions throughout the policy cycle. In the implementation of health policy, for example, the use of AI to treat diseases can represent an organizational optimization in the availability of instruments to serve society. In 2018, IBM launched Watson for Oncology, a revolutionary tool for AI-enabled personalized cancer treatment. Watson for Oncology worked with limited data from real-world patients and was dependent on synthetic data. The result was the recommendation of unsafe and inaccurate treatments that threatened human health. This led IBM to discontinue the solution.

This example demonstrates that the development and use of AI in public policy is not a simple matter of pressing a button, nor that there are solutions that follow a "one size fits all" pattern. AI, especially in the field of public policy, demands a broad governance process to frame technological development in procedures and rules that aim at effectiveness, safety and realization of public value. As an instrument, it has the potential to be a factor in organizational rationalization and produce profound epistemic changes in the policy process. However, this is no guarantee that we will have a rational, neutral, and effective policy process. The social and political dilemmas of public policy remain from the moment that AI in public policy follows a logic of instrument constituency, and its development is permeated by interests, opinions, and perspectives on which solution is best at a given moment. Throughout the policy cycle, AI is a solution looking for problems, with a view to the production of systems on an industrial scale, framed in a techno-solutionist perspective that changes the ways of making decisions and performing tasks in an unpredictable way.

The increasing pervasiveness of AI throughout the policy cycle means that system design actions are intertwined with typical actions for policy formulation, implementation, and evaluation. This intertwining creates a higher layer, in which the design of artificial intelligence systems becomes actions for designing policy in the context of uncertainty, ambiguity, and opacity that are framed within the broader context of the political system. Designing an AI system with the goal of automating and augmenting a public policy represents a political activity at the time when the objectives will be outlined, resources will be allocated, the instruments to achieve the objectives will be defined, and the deployment of the system will put the policy into implementation.

Artificial intelligence does not change the dilemmas for decision-making and policy design. The political and social dilemmas remain the same, but with different practices, involving a data-driven language, algorithmic architectures, logical and abstract system design, operation, and implementation. Overall, policy work is radically transformed as actors begin to interact with AI to perform various actions based on a radical and abrupt change in the structure of knowledge in society. Working with large volumes of data, systems that deliver information in real time and that automate repetitive tasks cause the focus of policy work to change. Humans interact with systems applied in public policy by offering data and information, abstract logic of systems and objectives. On the other hand, artificial intelligence systems return information and actions that optimize various elements of the policy process. Furthermore, it modifies the entire structure of policy advisory, creating new patterns of action, new ways of understanding problems and shaping solutions that aim to achieve a political objective.

Artificial intelligence does not overcome the set of political activities that are circumscribed in the frames that surround its defense and hype, as well as its criticisms and identification of threats that technological advances pose to governments. In the field of public policy, understanding the construction and use of AI to carry out government activities must be thought of in a more realistic and effective framework. AI is not a solution to all problems nor is it a neutral technology that will automatically provide effective changes. The adoption of artificial intelligence in the policy process requires legitimacy, given the constitutive dilemmas of any public policy, such as effectiveness, efficiency, security and achievement of public purposes. Likewise, AI is not an existential threat to humanity. The framing of the discussion is very important, and AI should be thought of in a framework focused on its instrumental character, associated with its possibilities of use and challenges that involve its development and construction.

Conceived as an instrument of knowledge within the policy process, artificial intelligence, however, produces a series of changes in policy work. Science and evidence are commonly seen as the epistemic building blocks of rationalist policymaking. Accordingly, the Lasswellian imaginary of policy science, democracy embodies the romanticized image of a professional moving between scientific and political realms. The fact is that AI does not mean the most “scientific”, evidence-based policymaking. AI systems are created by trial and error in uncertain contexts and with opaque technologies, driven by a metric of accuracy to solve a problem or perform a task. AI is an important tool for producing organizational optimization. However, it produces a series of new problems that require human oversight, governance frameworks, and new modes of policy work to deal with emerging issues, such as algorithmic unfairness, security issues, or threats.

The professional field of policy science must absorb a new capacity related to the development of systems. And it must also absorb new ways of dealing with the knowledge created by AI systems applied in various areas of the policy cycle. AI systems are knowledge systems that transform the practical action of policy analysts and practitioners. The epistemic changes that emerge with artificial intelligence transform all policy work, underpinning new rationalities, frameworks and action scripts that transform the way policies are formulated, decided, implemented and evaluated. On the one hand, policy work must be attentive to system design processes, considering the dynamics of instrument constituency. On the other hand, AI is an instrument to accelerate and increase policy analysis and to influence new patterns of action through prediction and simulation.

In the context of these epistemic changes, the disruptive changes in policy science impact the way governments understand their problems, formulate solutions and evaluate them. The risk is that policymakers and implementers delegate analytical capabilities to AI systems to predict the future, while at the same time these policymakers and implementers lose control over the future. These risks require governments to rethink problems and solutions and create initiatives that are capable of adapting AI to the complex reality of policy sciences, both in terms of knowledge construction and analysis, and from a professional perspective. At the same time, we must face a totalizing and epistocratic perspective of public policy formulated and implemented with AI. This is the main challenge in view, which calls the policy scientist to think about sociotechnical reengineering in a broader framework of defense of democracy and its virtues, in a broader scenario of defense of freedom and human autonomy.

References

Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). A brief survey of text mining: Classification, clustering and extraction techniques. *Arxiv*. Doi: https://doi.org/10.48550/arXiv.1707.02919

Alon-Barkat, S., Busuioc, M. (2023). Human–AI interactions in public sector decision making: "Automation bias" and "selective adherence" to algorithmic advice, *Journal of Public Administration Research and Theory*, 33 (1), p. 153–169. Doi: https://doi.org/10.1093/jopart/muac007

Althaus, M. (2013). Reflections on advisory practice in politics. *PSCA—Political Science Applied.* 2, p. 5–15.

Alvarado, R. (2023). AI as an epistemic technology. *Science Engineering Ethics*, 29, 32. Doi: https://doi.org/10.1007/s11948-023-00451-3

Amershi, S., Cakmak, M., Knox, W.B., Kulesza, T. (2014). Power to the people: The role of humans in interactive machine learning. *AI Magazine*, 35 (4), p.105–120. Doi: https://doi.org/10.1609/aimag.v35i4.2513

Amoore, L. (2022). Machine learning political orders. *Review of International Studies*, 49 (1), p. 20-36. Doi: https://doi.org/10.1017/S0260210522000031

Amoore, L. (2014). Security and the incalculable. *Security Dialogue*, 45 (5), p. 423–439. Doi: http://dx.doi.org/10.1177/0967010614539719

Anthony, C. (2021). When knowledge work and analytical technologies collide: The practices and consequences of black boxing algorithmic technologies. *Administrative Science Quarterly*, 66 (4), p. 1173-1212. Doi: https://doi.org/10.1177/00018392211016755

Bali, A.S., Howlett, M., Lewis, J.M., Ramesh, M. (2021). Procedural policy tools in theory and practice. *Policy and Society*, 40 (3), p. 295–311. Doi: https://doi.org/10.1080/14494035.2021.1965379

Banuri, S., Dercon, S., Gauri, V. (2019). Biased policy professionals. *The World Bank Economic Review*, 33 (2), p. 310–327. Doi: https://doi,org/10.1093/wber/lhy033

Barandiaran, X., Di Paolo, E., Rohde, M. (2009). Defining agency: individuality, normativity, asymmetry and spatial temporality in action. *Adaptive Behaviour*, 17 (5), 367-386.

Bardach, E. (2012). *A practical guide for policy analysis.* The eightfold path to more effective problem solving. Los Angeles: Sage.

Barzelay, M. (2019). Public management as a design-oriented professional discipline. Cheltenham: Edward-Elgar.

Baumgartner, F.R., Jones, B.D. (2015). The politics of information: Problem definition and the course of public policy in America. Chicago. The University of Chicago Press.

Béland, D. (2019). *How ideas and institutions shape politics and policy*. Cambridge: Cambridge University Press.

Béland, D., Howlett, M. (2016). Instrument constituencies in the policy process. *Governance*, 29 (3), p. 393-409. Doi: https://doi.org/10.1111/gove.12179

Bender, E.M., Gebru, T., McMillan-Major, A., Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21).* New York: Association for Computing Machinery, 610–623. Doi: https://doi.org/10.1145/3442188.3445922

Bengio, Y. (2023). AI and catastrophic risk. *Journal of Democracy,* 34 (4), p. 111-121.

Benjamin, R. (2019). *Race after technology*: *Abolitionist tools for the new Jim Code*. New York: Polity Press.

Blumer, H. (1986). *Symbolic interactionism. Perspective and method*. Berkeley: University of California Press.

Bovens, M., 't Hart, P., Kuipers, S. (2008). The politics of policy evaluation, in Robert Goodin, Michael Moran, and Martin Rein (eds), *The Oxford Handbook of Public Policy* Oxford: Oxford University Press.

Braunerhjelm, P., Hepburn, C. (2023). Climate change, complexity, and policy design. *Oxford Review of Economic Policy*, 39 (4), p. 667–679. Doi: https://doi.org/10.1093/oxrep/grad047

Braybrooke, D., Lindblom, C.E. (1963). *A strategy of decision*. New York: Free Press.

Brennan, J. (2016). *Against democracy*. Princeton: Princeton University Press.

Buolamwini, Joy (2023). Unmasking AI: My Mission to Protect What Is Human in a World of Machines. New York: Random House.

Campion, A., Gasco-Hernandez, M., Jankin Mikhaylov, S., Esteve, M. (2022). Overcoming the challenges of collaboratively adopting artificial intelligence in the public sector. *Social Science Computer Review*, 40 (2), p. 462-477. Doi: https://doi.org/10.1177/0894439320979953

Capano, G., Howlett, M.P. (2020). The knowns and unknowns of policy instrument analysis: Policy tools and the current research agenda on policy mixes. *SAGE Open* 10 (1), p. 1–13. Doi: https://doi.org/10.1177/2158244019900568

Chaqués-Bonafont, L., Palau, A.M., Baumgartner, F. (2015). *Agenda dynamics in Spain*. London: Palgrave.

Chatfield, A.T., Reddick, C.G. (2019). A framework for Internet of Things-enabled smart government: A case of IoT cybersecurity policies and use cases in U.S. federal government. Government Information Quarterly, 36 (2), p. 346-357. Doi: https://doi.org/10.1016/j.giq.2018.09.007

Chen, C., Surette, R., and Shah, M. (2021). Automated monitoring for security camera networks: Promise from Computer Vision Labs. *Security Journal*, 34, p. 389–409. Doi: https://doi.org/10.1057/s41284-020-00230-w

Clifton, J., Pal, L.A. (2022). The policy dilemmas of blockchain, *Policy and Society*, 41 (3), p. 321–327. Doi: https://doi.org/10.1093/polsoc/puac025

Coeckelbergh, M. (2021). Narrative responsibility and artificial intelligence: How AI changes human responsibility and sense making. *AI & Society*, 38 (4), 2437-2450. Doi: https://doi.org/10.1007/s00146-021-01375-x

Coeckelbergh, M. (2023). Democracy, epistemic agency, and AI: political epistemology in times of artificial intelligence. *AI Ethics*, 3 (1), 1341–1350. Doi: https://doi.org/10.1007/s43681-022-00239-4

Coeckelbergh, M., Sætra, H.S. (2023). Climate change and the political pathways of AI: The technocracy-democracy dilemma in light of artificial intelligence and human agency. Technology in Society, 75, 102406. Doi: https://doi.org/10.1016/j.techsoc.2023.102406

Cohen, M.D., March, J.G., Olsen, J.P. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly,* 17 (1), p. 1-25.

Coles-Kemp, L.; Ashenden, D.; Morris, A.; Yuille, J. (2022). Digital welfare: Designing for more nuanced forms of access. *Policy Design and Practice***,** 3 (2), p. 177-188. Doi: https://doi.org/10.1080/25741292.2020.1760414

Collington, R. (2021). Disrupting the welfare state? Digitalisation and the retrenchment of public sector capacity. *New Political Economy*, 27 (2), p. 312–328. Doi: https://doi.org/10.1080/13563467.2021.1952559

Connolly, R. (2020). Why computing belongs within social science. *Communications of the ACM*, 63 (8), p. 54-59. Doi: https://doi.org/10.1145/3383444

Cowls, J., Tsamados, A., Taddeo, M. *et al.* (2023). The AI gambit: leveraging artificial intelligence to combat climate change—opportunities, challenges, and recommendations. *AI & Society,* 38 (2), p. 283–307. Doi: https://doi.org/10.1007/s00146-021-01294-x

Cox, T. (2019). Muddling-through and deep learning for managing large-scale uncertain risks. *Journal of Benefit-Cost Analysis,* 10 (2), p. 226-250. Doi: https://doi.org/10.1017/bca.2019.17

Craft, J., Head, B., Howlett, M.P. (2024). Expertise, policy advice, and policy advisory systems in an open, participatory, and populist era: New challenges to research and practice. *Australian Journal of Public Administration*, 83, p. 143–155. Doi: https://doi.org/10.1111/1467-8500.12630

Craft, J., Halligan, J. (2020). Advising governments in the Westminster tradition: Policy advisory systems in Australia. Cambridge: Cambridge University Press.

Craft, J., Howlett, M.P. (2013). The dual dynamics of policy advisory systems: The impact of externalization and politicization on policy advice. *Policy and Society*, 32 (3), 187-197. Doi: https://doi.org/10.1016/j.polsoc.2013.07.001

Crawford, K. (2021). *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. New Haven: Yale University Press.

Cross, E., Ramsey, R. (2021). Minds meets machines: Towards a cognitive science of human-machine interactions. *Trends in Cognitive Sciences*, 25 (3), p. 200-2013. Doi: https://doi.org/10.1016/j.tics.2020.11.009

Curtis, C., Gillespie, N., Lockey, S. (2023). AI-deploying organizations are key to addressing 'perfect storm' of AI risks. *AI Ethics*, 3, p. 145–153. Doi: https://doi.org/10.1007/s43681-022-00163-7

Daugherty, P., Wilson, H. (2018). *Human + Machine: Reimagining work in the age of AI*. Boston: Harvard Business Press.

De Leon, P. (1981). Policy sciences: Discipline and profession. *Policy Sciences*, 13 (1), 1-7.

DeSouza, K.C., Jacob, B. (2017). Big data in the public sector: Lessons for practitioners and scholars. *Administration & Society*, 49 (7), p. 1043-1064. Doi: https://doi.org/10.1177/0095399714555751

Deutsch, K.W. (1963). *The nerves of government. Models of political communication and control.* New York: Free Press.

Diakopoulos, N., Johnson, D. (2021). Anticipating and addressing the ethical implications of deepfakes in the context of elections. *New Media & Society*, 23 (7), p. 2072-2098. Doi: https://doi.org/10.1177/1461444820925811

Domingos, P. (2015). The master algorithm. How the quest for the ultimate learning machine will remake our world. New York: Basic Books.

Dunleavy, P. (2016). Big data and policy learning. In: *Evidence-Based policy making in the social sciences. Methods that matter*. Edited by Stoker, Gerry, Evans, Mark. Chicago: Policy Press.

Dunleavy, P., Margetts, H. (2024). Data science, artificial intelligence and the third wave of digital era governance. *Public Policy and Administration*, early view. Doi: https://doi.org/10.1177/09520767231198737

Eisenstein, J. (2019). Introduction to natural language processing. Cambridge: MIT Press.

Elmore, R.F. (1979). Backward mapping: Implementation research and policy decisions. *Political Science Quarterly*, 94 (4), p. 601-616.

Esko, T., Koulu, R. (2023). Imaginaries of better administration: Renegotiating the relationship between citizens and digital public power. *Big Data & Society*, 10 (1), . Doi: https://doi.org/10.1177/20539517231164113

Estlund, D. (2008). *Democratic authority: A philosophical framework*. Princeton: Princeton University Press.

Estrada, M.A.R. (2011). Policy modeling: Definition, classification and evaluation. Journal of Policy Modeling, 33 (4), p. 523-536. Doi: https://doi.org/10.1016/j.jpolmod.2011.02.003

Etzioni, A. (1968). *The active society*. A theory of societal and political process. New York: Free Press.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York: St. Martin's Press.

Filgueiras, F. (2022). Big data, artificial intelligence and the future of regulatory tools. In: Howlett, M.P. (ed.). *Routledge Handbook on Policy Tools*. London: Routledge, p. 529-539.

Filgueiras, F. (2022). New Pythias of public administration: ambiguity and choice in AI systems as challenges for governance. *AI & Society*, 37 (4), p. 1473–1486. Doi: https://doi.org/10.1007/s00146-021-01201-4

Filgueiras, F., Almeida, V. (2021). *Governance for the digital world: Neither more state nor more market*. London: Palgrave.

Floridi, L. (2023). AI as agency without intelligence: On ChatGPT, Large Language Models, and other generative models. *Philosophy & Technology*, 36 (1), 15.

Floridi, L. (2023b). *The ethics of artificial intelligence. Principles, challenges and opportunities*. Oxford: Oxford University Press.

Frischmann, B., Selinger, E. (2018). *Re-engineering humanity*. Cambridge: Cambridge University Press.

Gasser, U., Almeida, V. (2017). A layered model of AI goverance. *IEEE Internet Computing*, 21 (6), p. 58-62. Doi: http://dx.doi.org/10.1109/MIC.2017.4180835

Giest, S. (2017). Big data for policymaking: fad or fasttrack? *Policy Analysis*, 50 (3), p. 367-382. Doi: https://doi.org/10.1007/s11077-017-9293-1

Gitelman, S. (2013). *Raw data is an oxymoron*. Cambridge: MIT Press.

Goldman, A.I. (2003). "The Technology and Economics of Communication", *Knowledge in a Social World*. Oxford: Oxford University Press.

Gyódi, K., Nawaro, Ł., Paliński, M. *et al.* (2023). Informing policy with text mining: technological change and social challenges. *Quality & Quantity* **57**, p. 933–954. Doi: https://doi.org/10.1007/s11135-022-01378-w

Halligan, J. (1995). Policy advice and the public service. In B. G. Peters & D. J. Savoie (Eds.), *Governance in a changing environment* (pp. 138–172). Montreal: McGill-Queens' University Press and Canadian Centre for Management Development.

Hartley, K, Kuecker, G.D. (2022). *Disrupted governance*: Towards a new policy science. Cambridge: Cambridge University Press.

Henman, P. (2018). Of algorithms, Apps and advice: digital social policy and service delivery. *Journal of Asian Public Policy*, 12 (1), p. 71–89. Doi: https://doi.org/10.1080/17516234.2018.1495885

Hill, M., Hupe, P. (2009). *Implementing public policy*. Thousand Oaks: Sage.

Hjern, B., Porter, D. (1981). Implementation structures: A new unit of administrative analysis. *Organizational Studies,* 2 (3), p. 211-227. Doi: https://doi.org/10.1177/017084068100200301

Hood, C.C., Margetts, H. (2007). The tools of government in digital age. London: Palgrave MacMillan.

Howlett, M.P. (2019). *Policy design primer: Choosing the right tools for the job*. London: Routledge.

Howlett, M.P. (2019). Procedural policy tools and the temporal dimensions of policy design. *International Review of Public Policy,* 1 (1), p. 27–45. Doi: https://doi.org/10.4000/irpp.310

Howlett, M.P., Mukherjee, I., Rayner, J. (2018). Understanding policy designs over time. In *Routledge Handbook of Policy Design*, edited by M. Howlett. London: Routledge.

Howlett, M.P. (2015). Policy analytical capacity: The supply and demand for policy analysis in government. *Policy and Society,* 34 (3-4), p. 173-182. Doi: http://dx.doi.org/10.1016/j.polsoc.2015.09.002

Howlett, M.P., Ramesh, M. (2003). *Studying public policy. Policy cycles and policy subsystems.* Oxford: Oxford University Press.

Hurley, M.W., Wallace, W.A. (1986). Expert systems as decision aids for public managers: An assessment of the technology and prototyping as a design strategy. *Public Administration Review*, 46 (4), p. 563–571.

Ifenthaler, D., Majumdar, R., Gorissen, P. *et al.* (2024). Artificial intelligence in education: Implications for policymakers, researchers, and practitioners. *Technology, Knowledge and Learning,* early view, p. 1-18. Doi: https://doi.org/10.1007/s10758-024-09747-0

Imperial, M. (2021). Implementation structures: The use of top-down and bottom-up approaches to policy implementation, in: *Oxford research encyclopedia – Politics*. Doi: https://doi.org/10.1093/acrefore/9780190228637.013.1750

Innerarity, D. (2024). The epistemic impossibility of an artificial intelligence take-over of democracy. *AI & Society*, 39(4), p. 1667–1671. Doi: https://doi.org/10.1007/s00146-023-01632-1

Jacobs, J. (2024). The artificial intelligence shock and socio-political polarization. *Technological Forecasting and Social Change*, 199, 123006, Doi: https://doi.org/10.1016/j.techfore.2023.123006

Janssen, M., Helbig, N. (2018). Innovating and changing the policy-cycle: Policy-makers be prepared! *Government Information Quarterly*, 35 (1), p. 99-105. Doi: https://doi.org/10.1016/j.giq.2015.11.009

Jarrahi, M.H., Lutz, C., Newlands, G. (2022). Artificial intelligence, human intelligence and hybrid intelligence based on mutual augmentation. *Big Data & Society*, 9 (2), p. 1-6. Doi: https://doi.org/10.1177/20539517221142824

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, 93 (5), p. 1449–1475**.** Doi: https://doi.org/10.1257/000282803322655392

Kayser-Bril, N. (2020). In a quest to optimize welfare management, Denmark built a surveillance behemoth. *Algorithmic Watch.* Retrieved from: In a quest to optimize welfare management, Denmark built a surveillance behemoth - AlgorithmWatch

Kingdon, J.W. (1995). *Agendas, alternatives, and public policies*. New York: Harper & Collins.

Kitchin, R. (2014). The data revolution. Big data, open data, data infrastructures and their consequences. Thousand Oaks: Sage.

Klijn, E. H., & Koppenjan, J. (2012). Governance network theory: Past, present, and future. *Policy and Politics*, 40 (4), p. 587–606. Doi: https://doi.org/10.1332/030557312x655431

Korinek, A., Balwit, A. (2023). Aligned with whom? Direct and social goals and AI systems, in: J.B. Bullock, Y.C. Chen, J. Himmelreich, V.M. Hudson, A. Korinek, M.M. Young, B. Zhang (eds.). *The Oxford handbook of AI governance*. Oxford: Oxford University Press.

Koga, N.M., Palotti, P.L.M, Pontes, P.A.M.M., Couto, B.G., Soares, M.L.V. (2023). Analytical capacity as a critical condition for responding to COVID-19 in Brazil. *Policy and Society*, 42 (1), p. 117–130. Doi: https://doi.org/10.1093/polsoc/puac028

Krishnamoorthy, C.S., Rajeev, S. (1996). *Artificial intelligence and expert systems for engineers*. Boca Raton: CRC Press.

Landemore, H. (2012). *Democratic reason.* Politics, collective intelligence, and the rule of many. Princeton: Princeton University Press.

Lascoumes, P., Le Galès, P. (2007). Understanding public policy thorough its instruments – From the nature of instruments to the sociology of public policy instrumentation. *Governance*, 20 (1), p. 1-21. Doi: https://doi.org/10.1111/j.1468-0491.2007.00342.x

Lasswell, H.D. (1970). The emerging conception of policy sciences. *Policy Sciences*, 1 (1), 3-14.

LeCun, Y., Bengio, Y, Hinton, G. (2015). Deep learning. *Nature*, 521, 436-444. Doi: https://doi.org/10.1038/nature14539

Lehdonvirta, V. (2024). *Cloud empires: How digital platforms are overtaking the State and how we can regain control*. Cambridge: MIT Press.

Lim H.S.M., Taeihagh, A. (2019). Algorithmic decision-making in AVs: Understanding ethical and technical concerns for smart cities. *Sustainability*, 11 (20), 5791. Doi: https://doi.org/10.3390/su11205791

Lin, T.C.W. (2017). The new market manipulation. *Emory Law Journal*, 66 (6), 1253-1314.

Lindblom, C. E. (1959). The science of muddling through. *Public Administration Review*, 19 (2), p. 79-88.

Lipsky, M. (2010). *Street-level bureaucracy: Dilemmas of the individual in public service*. New York: Russell Sage Foundation.

Logan, S. (2024). Tell me what you don't know: large language models and the pathologies of intelligence analysis. *Australian Journal of International Affairs*, 78 (2), p. 220–228. Doi: https://doi.org/10.1080/10357718.2024.2331733

Lowi, T. (1964). American business, public policy, case-studies, and political theory. *World Politics*, 16 (4), p. 677–715. Doi: https://doi.org/10.2307/2009452

Margetts, H., Dorobantu, C. (2019). Rethink government with AI. *Nature*, 568 (7751), p. 163–165.

Markram, H. (2006). The Blue Brain Project. *Nature Reviews Neuroscience*, 17 (2), 153–160. Doi: https://doi.org/10.1038/nrn1848

McCarthy, J. (1981). Epistemological problems of artificial intelligence. *In*: Webber, B.L.; Nilsson, N.J. (eds.) *Readings in artificial intelligence*. Berlin: Springer, p. 459-465.

Meijer, A., Wessels, M. (2019). Predictive policing: Review of benefits and drawbacks. *International Journal of Public Administration*, 42 (12), 1031–1039. Doi: https://doi.org/10.1080/01900692.2019.1575664

Mendonça, R.F., Filgueiras, F., Almeida, V.A. (2023). *Algorithmic institutionalism: The changing rules of social + political life.* Oxford: Oxford University Press.

Minsky, M. (1985). *The society of mind*. New York: Simon and Schuster.

Muccione, V., Vaghefi, S.A., Bingler, J. *et al.* (2024). Integrating artificial intelligence with expert knowledge in global environmental assessments: opportunities, challenges and the way ahead. *Regional Environmental Change*, 24 (121), 1-8. Doi: https://doi.org/10.1007/s10113-024-02283-8

Newman, J., Mintrom, M. (2023). Mapping the discourse on evidence-based policy, artificial intelligence, and the ethical practice of policy analysis. *Journal of European Public Policy*, 30 (9), p. 1839–1859. Doi: https://doi.org/10.1080/13501763.2023.2193223

Nobel, S.U. (2018). *Algorithms of opression: How search engines reinforce racism*. New York: NYU Press.

Nowak, A.J. (2022). Artificial intelligence in evidence-based medicine. In: Lidströmer, N., Ashrafian, H. (eds) *Artificial intelligence in medicine*. Berlin: Springer. Doi: https://doi.org/10.1007/978-3-030-64573-1_43

Nowotny, H. (2021). *In AI we trust: Power, illusion, and control of predictive algorithms.* New York: Polity Books.

Nowotny, H. (2021). *In AI we trust: Power, illusion, and control of predictive algorithms.* New York: Polity Books.

Olsen, J.P. (2006). Maybe it is time to rediscover bureaucracy. *Journal of Public Administration Research and Theory*, 16 (1), p. 1–24. Doi: https://doi.org/10.1093/jopart/mui027

O'Neil, C. (2016). *The weapons of math destruction: How big data increases inequality and threatens democracy*. New York: Crown Publishing.

Ostrom, E. (2005). *Understanding institutional diversity*. Princeton: Princeton University Press.

Padmaja, B., Moorthy, C.V.K.N.S.N., Venkateswarulu, N. *et al.* (2023). Exploration of issues, challenges and latest developments in autonomous cars. *Journal of Big Data*, 10, p. 1-24. https://doi.org/10.1186/s40537-023-00701-y

Pattyn, V., Brans, M. (2015). Organisational analytical capacity: Policy evaluation in Belgium. *Policy and Society*, 34 (3-4), p. 183-196. Doi: https://doi.org/10.1016/j.polsoc.2015.09.009

Paul, R. (2022). Can *critical policy studies* outsmart AI? Research agenda on artificial intelligence technologies and public policy. *Critical Policy Studies*, 16 (4), p. 497–509. Doi: https://doi.org/10.1080/19460171.2022.2123018

Peters, B. G., Pierre, J. (2016). *Comparative governance*. Rediscovering the functional dimension of governing. Cambridge University Press.

Peters, B.G. (2012). Information and governing: Cybernetic models of governance, in David Levi-Faur (ed.), *The Oxford Handbook of Governance.* Oxford: Oxford University Press.

Peeters, R., Rentería, C., Cejudo, G.M. (2023). How information capacity shapes policy implementation: A comparison of administrative burdens in COVID-19 vaccination programs in the United States, Mexico, and the Netherlands. *Government Information Quarterly*, 40 (4), 101871. Doi: https://doi.org/10.1016/j.giq.2023.101871

Pressman, J.L., Wildavsky, A. (1984). *Implementation.* Berkeley: University of California Press.

Ramezani, M., Takian, A., Bakhtiari, A. *et al.* (2023). The application of artificial intelligence in health policy: a scoping review. *BMC Health Services Research*, 23, 1416. Doi: https://doi.org/10.1186/s12913-023-10462-2

Reid, T., Gibert, J. (2022). Inclusion in human-machines interactions. *Science*, 375 (6577), p. 149-150. Doi: https://doi.org/10.1126/science.abf2618

Rikap, C. (2021). Capitalism, power and innovation. London: Routledge.

Russell, S. (2022). 'Why We Need to Regulate Non-State Use of Arms', 18 May. https://www.weforum.org/agenda/2022/05/regulate-non-state-use-arms/#:~:text=Stuart%20Russell&text=An%20emerging%20arms%20race%20between,while%20the%20technology%20rapidly%20advances .

Russell, S.J. (2019). *Human compatible*. Artificial intelligence and the problem of control. New York: Viking Books.

Russell, S.J.; Norvig, P. (2010). *Artificial intelligence*. A modern approach. Englewood Cliffs, NJ: Prentice-Hall.

Sabatier, P.A. (1999). The need for better theories. In P.A. Sabatier (ed.), *Theories of the policy process*, pp. 3–17. Boulder: Westview.

Sætra, H.S. (2023). "The promise and pitfalls of techno-solutionism", in: *Technology and sustainable development. The promises and pitfalls of techno-solutionism.* London: Routledge.

Safaei, M.; Longo, J. (2024). The end of policy analyst? Testing the capability of artificial intelligence to generate plausible, persuasive and useful policy analysis. *Digital Government: Research and Practice*, 5 (1), p. 4-35. Doi: https://doi.org/10.1145/3604570

Salamon, L. (2002). *The tools of government. A guide to the new governance*. Oxford University Press

Samuel, A. L. (1962). Artificial intelligence: A frontier of automation. *The ANNALS of the American Academy of Political and Social Science*, 340 (1), p. 10-20. Doi: https://doi.org/10.1177/000271626234000103

Samuel, A. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 44 (1.2), 210-229. Doi: https://doi.org/10.1147/rd.33.0210

Schrage, M. (2020). *Recommendation engines*. Cambridge, MA: MIT Press.

Schwember, H. (1977). Cybernetics in government: Experience with new tools for management in Chile 1971–1973. In *Concepts and tools of computer-assisted policy analysis: Vol. 1: Basic Concepts*, Hartmut Bossel (Ed.). Birkhäuser: Basel, p. 79–138.

Searle, J.R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences,* 3 (3), p. 417-424. Doi: https://doi.org/10.1017/S0140525X00005756

Shardlow, M., Przybila, P. (2023). Deanthropomorphising NLP: Can a Language Model Be Conscious? ArXiv, p. 1-20. Doi: https://doi.org/10.48550/arXiv.2211.11483

Schuett , J. (2023). Risk Management in the Artificial Intelligence Act. European Journal of Risk Regulation. Doi:10.1017/err.2023.1, 1-19.

Siddiki, S. (2020). *Understanding and analyzing public policy design*. Cambridge: Cambridge University Press.

Silva, E.H.M., Santos, E.M.F., Miranda, M.L.B.M., Bezerra, S.L., Miranda, S.C. (2024). ChatTCU: inteligência artificial como assistente do auditor. *Revista do TCU,* 153, p. 19-45.

Simons, A., Voß, J.P. (2018). The concept of instrument constituencies: Accounting for dynamics and practices of knowing governance. *Policy and Society*, 37 (1), p. 14–35. Doi: https://doi.org/10.1080/14494035.2017.1375248

Simon, H.A. (1947). *Administrative behavior.* New York: Free Press.

Simon, H.A. (1970). *The sciences of artificial.* Cambridge: MIT Press.

Simon, H.A. (1983). *Reason in human affairs.* Stanford: Stanford University Press.

Simon, H.A. (1995). Artificial intelligence: an empirical science. *Artificial Intelligence*, 77 (1) 95-127.

Simon, H.A., Eisenstadt, S.A. (2002). A Chinese room that understands. In: J.M. Bishop, J. Preston (eds.), *Views into the Chinese room: New essays on Searle and artificial intelligence*. Oxford: Oxford University Press.

Sorensen, E., & Torfing, J. (2005). The democratic anchorage of governance networks. *Scandinavian Political Studies*, 28 (3), p. 195–218. https://doi.org/10.1111/j.1467-9477.2005.00129.x

Stoker, G. (1998). Governance as theory: Five propositions. *International Social Science Journal*, 50 (155), p. 17–28. Doi: https://doi.org/10.1111/1468-2451.00106

Suchman, L. (2007). *Human-machine reconfigurations. Plans and situated actions*. Cambridge: Cambridge University Press.

Sun, T.Q., Medaglia, R. (2019). Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. Government Information Quarterly, 36 (2), p. 368-383. Doi: https://doi.org/10.1016/j.giq.2018.09.008

Süsser, D., Ceglarz, A., Gaschnig, H., Stavrakas, V., Flamos, A., Giannakidis, G., Lilliestam, J. (2021). Model-based policymaking or policy-based modelling? How energy models and energy policy interact. *Energy Research & Social Science*, 75, 101984. Doi: https://doi.org/10.1016/j.erss.2021.101984

Sutton, R., Barto, A. (2018). *Introduction to reinforcement learning*. Cambridge: MIT Press.

Taeihagh, A., Ramesh, M., Howlett, M. (2021). Assessing the regulatory challenges of emerging disruptive technologies. *Regulation & Governance*, 15 (4), p. 1009-1019. Doi: https://doi.org/10.1111/rego.12392

Tomasello, M. (2022). The evolution of agency: Behavioral organizations from lizards to humans. Cambridge: MIT Press.

Turing, A. (1950). Computing machiney and intelligence. *Mind – A Quarterly Review of Psychology and Philosophy,* 59 (236), 433-460.

Turkle, S. (1984). *The second self: Computers and the human spirit.* New York: Simon & Schuster.

Tyler, C., Akerlof, K.L., Allegra, A., Arnold, Z., Canino, H., Doornenbal, Goldstein, J.A., Pedersen, D.B., Sutherland, W.J. (2023). AI tools as science policy advisers? The potential and the pitfalls. *Nature,* 622, p. 27-30. Doi: https://doi.org/10.1038/d41586-023-02999-3

Urbina, F., Lentzos, F., Invernizzi, C. *et al.* (2022). Dual use of artificial-intelligence-powered drug discovery. *Nature Machine Intelligence,* 4, p. 189–191. Doi: https://doi.org/10.1038/s42256-022-00465-9

Valle-Cruz, D., Fernandez-Cortez, V., Gil-Garcia, J.R. (2022). From E-budgeting to smart budgeting: Exploring the potential of artificial intelligence in government decision-making for resource allocation. Government Information Quarterly, 39 (2), 101644. Doi: https://doi.org/10.1016/j.giq.2021.101644

Valle-Cruz, D. Sandoval-Almazán, R. (2022). Role and governance of artificial intelligence in the public policy cycle, in: Justin B. Bullock et al. (eds.). *The Oxford handbook of AI governance*. Oxford: Oxford University Press.

Valle-Cruz, D., Criado, J. I., Sandoval-Almazán, R., Ruvalcaba-Gomez, E. A. (2020). Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation. *Government Information Quarterly* ,37 (4), 101509. Doi: https://doi.org/10.1016/j.giq.2020.101509

Van der Voort, H.G., Klievink, A.J., Arnaboldi, M., Meijer, A.J. (2019). Rationality and politics of algorithms. Will the promise of big data survive the dynamics of public decision making? *Government Information Quarterly*, 36 (1), p. 27-38. Doi: https://doi.org/10.1016/j.giq.2018.10.011

Veale, M., Brass, I. (2019). Administration by algorithm? Public management meets public sector machine learning, in: Karen Yeung, and Martin Lodge (eds), *Algorithmic regulation.* Oxford: Oxford University Press. Doi: https://doi.org/10+1093/oso/9780198838494.003.0006

Vedung, E. (2009). *Public policy and program evaluation*. New Brunswick: Transaction.

Veselý, A. (2017); Policy advice as policy work: a conceptual framework for multi-level analysis. *Policy Sciences,* 50, p. 139–154. Doi: https://doi.org/10.1007/s11077-016-9255-z

Voß, J.-P., Simons, A. (2014). Instrument constituencies and the supply side of policy innovation: The social life of emissions trading. *Environmental Politics*, 23 (5), p. 735–754. Doi: https://doi.org/10.1080/09644016.2014.923625

Wang, P. (2019). On defining artificial intelligence. *Journal of General Artificial Intelligence,* 10 (2), p. 1-37. Doi: https://doi.org/10.2478/jagi-2019-0002

Wani, A.K., Rahayu, F., Ben Amor, I. *et al.* (2024). Environmental resilience through artificial intelligence: innovations in monitoring and management. *Environmental Science Pollution Research,* 31, p. 18379–18395. Doi: https://doi.org/10.1007/s11356-024-32404-z

Weiss, C.H. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19 (1), p. 21-33. Doi: https://doi.org/10.1016/S1098-2140(99)80178-7

Weiss, C.H. (1979). *Using social research in public policy making*. London: Lexington Books.

Weizenbaum, J. (1976). *Computer power and human reason: From judgment to calculation.* San Francisco: W. H. Freeman.

White House. (2023) Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence | The White House.

Willems, Jurgen et. al. (2022) Ethics of robotized public services: The role of robot design and its actions. Government Information Quarterly Volume 39, Issue 2 , April, https://doi.org/10.1016/j.giq.2022.101683

Wilson, R. (2009). Policy analysis as policy advice. In R. E Goodin, M. Moran and M. Rein (ed.), *The Oxford handbook of public policy*. Oxford University Press.

Wirtz, B.W., Weyerer, J.C., & Sturm, B.J. (2020). The dark sides of artificial intelligence: An integrated AI governance framework for public administration. *International Journal of Public Administration*, 43 (9), p. 818–829. Doi: https://doi.org/10.1080/01900692.2020.1749851

Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance*, 12 (4), p. 505-523. Doi: https://doi.org/10.1111/rego.12158

Yuwono, E.I., Tjondronegoro, D., Riverola, C., Loy, J. (2024). Co-creation in action: Bridging the knowledge gap in artificial intelligence among innovation champions. *Computers and Education: Artificial Intelligence*, 7, 100272. Doi: https://doi.org/10.1016/j.caeai.2024.100272

Zakharova, I., Jarke, J., & Kaun, A. (2024). Tensions in digital welfare states: Three perspectives on care and control. *Journal of Sociology*, *60*(3), 540-559. Doi: https://doi.org/10.1177/14407833241228312

Zhou, Z.H. (2012). *Ensemble methods: Foundations and algorithms.* New York: CRC Press.