Artificial Intelligence in Brazil still lacks a strategy

Report by the Center for Technology and Society at FGV Law School
Walter B. Gaspar and Yasmin Curzi de Mendonça

Artificial Intelligence in Brazil still lacks a strategy

Report by the Center for Technology and Society at FGV Law School

Walter B. Gaspar and Yasmin Curzi de Mendonça

Anyone who takes the time to read the recently published Brazilian Artificial Intelligence Strategy (EBIA) will not be able to get a very concrete idea of what the strategy really is. The document describes, in about fifty pages, some general considerations about the implementation of AI in several sectors, but without ever delving deeply into planning issues that would be basic to a successful strategy. Many questions remain unanswered, making the document take on the appearance more of a letter of intent than a pragmatic planning effort.

We will address below some of these issues, dealing with how EBIA i) does not identify the actors responsible for governance, failing to follow the example of other strategic documents already produced by the Executive; ii) does not specify measurable benchmarks; iii) is too generic in character; iv) does not sufficiently incorporate the expertise of the contributions offered during the public consultation; v) does not delve into the methods available to provide transparency and explicability to AI systems; and vi) uncritically incorporates research on the use of AI in Public Security.

Governance uncertainty

A first essential point that remains undefined is that of the governance structures responsible for its management. Many contributions made during the public consultation on EBIA suggested the creation of regulatory bodies, specific authorities or the use of existing structures. For example, we at FGV Direito Rio's Center for Technology and Society (CTS) suggested, in our contribution (p. 18), that a possible solution to some problems inherent to the implementation of AI systems would be

> "[The] creation of a specialized and independent regulatory body, capable of reviewing and licensing algorithmic decision systems. This authority would have the function of defining what types of audits can be carried out; what technical and/or legal requirements must be met for each case; determine possible types of decisions or contexts in which the use of machine learning algorithms should be

prohibited, due to their 'intrinsic opacity'; state any types of decision or contexts that require a more accurate explanation of the decision or the possibility of human review; define the technical requirements to be followed by organizations both in the development and in the use of AI systems ".

Likewise, the Rio de Janeiro Institute of Technology and Society (ITS Rio) highlighted in its contribution (p. 16) that many jurisdictions around the world discussed the creation of a specific authority to address the issue, in some cases creating a subdivision of their national data protection authorities dedicated exclusively to Artificial Intelligence. The creation of a specific authority is also identified as a guideline shared by several letters of principles in the Principled Artificial Intelligence study, by the Berkman Klein center at Harvard University, in which multiple approaches on artificial intelligence regulation were considered for an overall balance of the most frequent recommendations.

Many of the actions in the first axis of the strategy - "Legislation, Regulations and Ethical Use" - as well as the others, would benefit from a better definition of who will be their active subject. When, for example, there is talk of "Creating and implementing best practices or codes of conduct regarding the collection, implementation and use of data, encouraging organizations to improve their traceability, safeguarding legal rights" or "Promoting innovative approaches to regulatory supervision", it is very important to know who will be creating, implementing, encouraging or promoting, as this defines the extent of what can actually be accomplished. To repeat the point, a definition of the governance structure - responsible actors and their respective capacities - would bring clarity on the implementation of these actions.

EBIA is silent on this point. Although at times it mentions existing governance structures, none of the "strategic actions" listed at the end of each axis of the document is decisive in relation to a governance body or bodies responsible for monitoring the execution of the strategy as a whole. *Portaria* (ordinance) No. 4,617/21 of the Ministry of Science, Technology and Innovations (MCTI), which creates the strategy, establishes on this point only that it will be up to the Ministry "to create governance instances and practices to prioritize, implement, monitor and update strategic actions established in the Brazilian Artificial Intelligence Strategy".

Given this, it is impossible to know the form that AI governance in Brazil will take in the future. If we look at other neighboring documents produced by the Federal Executive itself, the insufficiency of EBIA is evident:

- Decree nº 9.319/18, which institutes the National System for Digital Transformation, establishes a specific governance structure formed by the Interministerial Committee for Digital Transformation (CITDigital, composed of representatives of the Federal sphere), by the Consultative Council for Digital Transformation (composed of specialists and representatives of the scientific community, civil society and the productive sector) and bodies, entities and institutions linked to digital transformation policies.
- Decree nº 9.854/19, which institutes the National Internet of Things Plan, creates the Management and Monitoring Chamber for the Development of Machine-to-Machine and Internet of Things Communication Systems (IoT Chamber), formed by representatives of the MCTI, Ministry of Health, Economy, Agriculture, Livestock and Supply and Regional Development.

On the subject, it is worth noting that the broad participation of civil society, academia and the productive sector is essential, given the complex nature of the topic addressed. As repeatedly commented on in the contributions to the public consultation process of the strategy, different applications of artificial intelligence in different sectors have radically different potential risks and benefits, so that the composition of groups that are too homogeneous and univocal for the monitoring of the strategy can lead to blind spots that jeopardize the achievement of the stated objectives.

Measuring progress

The issue of governance related to the artificial intelligence strategy is a topic that will need to be defined by MCTI in the future to make EBIA more practical and concrete. This is crucial, since it gives rise to several other questions that the document left open. What will be the frequency of review and control of the actions? What are the completion indicators for each one? At which point will the strategy need to be reformulated - an important topic, given the speed with which the technological landscape changes?

Looking at, for example, the Brazilian Strategy for Digital Transformation, we see that each of its nine axes brings not only strategic actions - which in some cases reach a level of specificity greater than that presented in EBIA -, but also measurable benchmarks to verify the successful implementation of the strategy. This is an essential element in order to verify if all the effort involved in designing the strategy has been effective.

It would be important, therefore, to develop indicators and a schedule of periodic reviews, with publication of targets for their monitoring. This would contribute to accountability regarding the objectives outlined in EBIA and would serve as a stepping stone for its execution, facilitating the work of the MCTI and other government agencies involved.

Inaccuracy of actions

In addition, it is interesting to note how some of EBIA's strategic actions are not as much actions as they are general objectives. For example, remaining in the first axis, the action of "Stimulating efforts of transparency and responsible disclosure regarding the use of AI systems, and promoting the observance, by such systems, of human rights, democratic values and diversity", in in its current form, it sounds more like an introduction to a letter of principles than a concrete action. In fact, it seems more like a reorganization of points already elaborated in the prompts of the public consultation phase (emphasis added):

> "In addition, it is often stated that systems must be designed in a way that **respects human rights, democratic values and diversity**, imposing the inclusion of appropriate safeguards that enable human intervention, whenever necessary, to guarantee a just society.
>
> Another point widely discussed refers to **transparency and responsible disclosure about Artificial Intelligence systems**, stressing the need to adopt measures to ensure the understanding of the processes associated with automated decision-making, making it possible to identify biases involved in the decision-making process and challenge those decisions, when appropriate".

The public consultation page that contains this topic brings relevant contributions that could make the statement more concrete. BRASSCOM, for example, highlights the Singapore AI Framework, which lists principles for the application of AI systems and extracts good practices from them:

> "Keeping these principles in mind, the AI Framework of Singapore suggests some good practices to assist those organizations that choose to follow the principles presented by the PDPC, divided into categories: 1) Internal Governance Measures and Structures ● Establishment of clear roles and responsibilities within the organization; ● Personnel involved in data protection practices and policies must monitor and manage risks; and ● Internal training. 2) Determining the level of human involvement in decision making for AI systems ● Appropriate level of human involvement, taking into account the entire context; and ● Take measures to minimize damage to individuals. 3) Operations Management ● Minimize bias in data and models; and ● Risk-based approach to ensure robustness and regular tuning. 4) Communication and interaction with all stakeholders ● Share with users the adopted AI policy(s) ● Allow users to provide feedback, if possible; and ● Make communications easy to understand".

Then, focusing on the use of AI solutions by the State, the Data Privacy Brasil Association highlights the terms of the 2018 Toronto Declaration, according to which "the main ethical principles are: (i) Identify risks, (ii) Ensure transparency and accountability and (iii) Put in place supervisory mechanisms". Specifically dealing with point ii, which relates more closely to the strategic action discussed above, they continue:

> "The principle of guaranteeing transparency means that 'States must ensure and require accountability and maximum possible transparency around public sector use of machine learning systems. This must include explainability and intelligibility in the use of these technologies so that the impact on affected individuals and groups can be effectively scrutinised by independent entities, responsibilities established, and actors held to account'. This principle unfolds in three active obligations on the part of the State: (i) publicly disclosing where machine learning systems are used in the public sphere, providing information that explains in clear and accessible terms how automated decision-making or machine learning processes are carried out, and document actions taken to identify, document and mitigate discriminatory impacts or against other rights, (ii) allow independent analysis and supervision through systems that are auditable, and (iii) avoid the use of 'black box' systems that may not be subjected to significant parameters of accountability and transparency, and refrain from using these systems under any circumstances in high-risk contexts".

Despite being cited in the body of the EBIA text, the Toronto Declaration is only mentioned at a generic level, without going into the minutiae of how to implement accountability in the use of AI systems by the government. Strategic actions, as already exposed, also only touch the subject superficially. This seems like a missed opportunity, given the level of detail that can be found in some of the contributions to EBIA's public consultation process.

Both contributions previously mentioned bring practical mechanisms for achieving the objective of "transparency and responsible disclosure regarding the use of AI systems".

CTS itself did so in its contribution, by describing several technical and organizational options to promote the transparency and explainability of AI systems - such as, for example, the use of model cards throughout the process of developing algorithms. Remaining on the point of responsible disclosure of the use of AI, it is interesting to resort to international references. The AI Roadmap developed by the independent advisory board of artificial intelligence experts to the UK government, for example, indicates the following action:

> "Commit to achieving AI and data literacy for everyone. The public needs to understand the risks and rewards of AI so they can be confident and informed users. An Online Academy for understanding AI, with trusted materials and initiatives would support teachers, school students and lifelong learning".

In addition to a generic intention - to inform everyone about AI -, the action indicates an effectively concrete and achievable path, in order to configure an objective action. This is a problem that permeates many of EBIA's actions, which speak of "stimulating", "structuring", "encouraging" and "defining" without saying who, when or how.

In addition, some key terms would need better definition to be operationalized. For example, when talking about "Facilitating access to open government data" ("AI Governance" axis), it would be important to specify what is intended by "facilitating", since for AI applications not only access to open data, but the quality and structure of these data are crucial factors. When enunciating the action of "Stimulating the retention of specialized ICT talent in Brazil" (axis "Workforce and training"), it would be important to list the ways in which this objective can be achieved, under the risk of merely resorting to an obvious objective without indicating a real path forward.

The problem of public security

In addition to these more general considerations, it is important to note at least one substantial and specific problem that EBIA poses when dealing with the use of AI in Public Security.

EBIA introduces survey statistics from the Carnegie Endowment for International Peace uncritically. However, the aim of such study was to establish concerns about the advancement of the active and indiscriminate use of AI technologies by governments for surveillance purposes. The use of research data to affirm a favorable position for the use of AI in this sector is deeply problematic, given that the study aimed to bring attention and alarm to the use of AI by public authorities in the area of public security. As the conclusion of the study attests, the potential problems of the use of AI, even those arising from apparently benevolent technologies, evoke the need for absolute caution by public authorities in the implementation of AI systems in the public security sector, especially in relation to facial recognition technologies (SRFs in the Portuguese acronym).

> "The spread of AI surveillance continues unabated. Its use by repressive regimes to engineer crackdowns against targeted populations has already sounded alarm bells. But even in countries with strong rule of law traditions, AI gives rise to troublesome ethical questions. Experts express concerns about facial recognition error rates and heightened false positives for minority populations. The public is increasingly aware of algorithmic bias in AI training datasets and their prejudicial impact on predictive policing algorithms and other analytic tools used by law

> enforcement. Even benign IOT applications—smart speakers, remote keyless entry locks, automotive intelligent dash displays—may open troubling pathways for surveillance. Pilot technologies that states are testing on their borders—such as iBorderCtrl's affective recognition system—are expanding despite criticisms that they are based on faulty science and unsubstantiated research. The cumulative impact gives pause. Disquieting questions are surfacing regarding the accuracy, fairness, methodological consistency, and prejudicial impact of advanced surveillance technologies. Governments have an obligation to provide better answers and fuller transparency about how they will use these new intrusive tools." (p. 24)

With regard specifically to the dissemination of facial recognition systems in Brazil, EBIA cites a study published by the Igarapé Institute on the implementation of facial recognition systems in Brazil to denote that "since 2011, SRFs have been used in Brazil for different purposes. Of 47 reported cases, 13 were intended for use in the context of public security. For example, in the city of Rio de Janeiro, between July and October 2019, 10% of the arrests of the 19th Military Police Battalion - BPM were due to the use of SRF". It is a fundamental conclusion of the aforementioned study, however, that the adoption of these systems implied the collection of detailed data on civil and criminal backgrounds of individuals, facial and biometric records, even before the elaboration and adoption of the general data protection law (LGPD).

Despite also mentioning the use of SRFs in conjunction with closed circuit television monitoring systems (CCTVs) and also referring to Igarapé Institute's research on the topic in a footnote, EBIA does not portray the conclusion of the study, which, in turn, states that:

> "Several studies point out that video surveillance has a limited effect on crime reduction. Surveillance cameras tend to be more effective in reducing crimes against property, to the detriment of crimes against life. Factors such as place of installation and type of system affect the effectiveness of video surveillance. The presence of surveillance cameras does not necessarily lead to an increase in the perception of security. The improvement in urban lighting may contribute more to the reduction of certain types of crime than the installation of cameras".

EBIA criticizes the potential perverse effects of the implementation of SRFs, such as the possibility of algorithmic discrimination and inefficiency of applications. However, in its strategic actions, it outlines few effective ways of approaching and addressing these problems and corroborates their implementation, delegating the standardization initiatives and possible planning necessary for the structuring of safe systems to other "regulatory bodies", without even pointing out appropriately what these would be (pp. 49-50).

Conclusion

The generality with which themes are addressed in EBIA's strategic actions, combined with the lack of a clear outline of the intended governance structure, deadlines and goals, give the document the appearance of a timid first step on the path of AI regulation.

A series of questions remains open: what are the technical and organizational guidelines for facing problems linked to the implementation of AI systems? Who will be able to

define or revise these guidelines, with what frequency, through which regulatory instruments? How will civil society participate in decision-making processes on the topic? What incentive instruments will be applied for the development of AI applications in Brazil? What are the priority sectors for this development?

These are questions that EBIA hints at, but does not answer. However, given the long process of public consultation, the volume of existing knowledge and the experiences of other countries in similar undertakings, a more significant advance was expected. The Executive will face the hard work of drawing firm contours for AI in Brazil. Listening to the voices that indicate concrete and possible paths and watching the progress of similar experiments in other countries can give more strength and effectiveness to any further steps from now on.