

From the Regulation of Artificial Intelligence to the Governance of Artificial Cognitive Systems. Cognitive power, decision-making architectures and emerging regulatory challenges

Giovanni Giamminola, Conceptual Lead on Cognitive Governance and Decision Architectures, Senior Advisor on AI-Driven Cognitive Governance – Pollicino Advisory

Prof. Avv. Oreste Pollicino, Managing Partner Pollicino Aldvisory

Executive Summary

The current debate on Artificial Intelligence is still predominantly focused on regulating the technology and mitigating the risks associated with it. This policy brief proposes a change of perspective: contemporary AI, particularly in its generative and agentic forms, is no longer just a tool, but a real cognitive infrastructure that intervenes in individual and collective perception, interpretation and decision-making processes.

The result is an exercise of power that is both cognitive and de facto normative, often prior to and independent of formal legal intervention. In this context, AI governance cannot be limited to ex post compliance models, but must evolve towards a paradigm of integrated cognitive and regulatory governance, capable of intervening ex ante on the decision-making architectures and sensemaking mechanisms that AI helps to structure.

Over the last year, and especially due to the acceleration of generative AI, synthetic media, automated persuasion systems, and platform-mediated amplification, this cognitive power has acquired a direct impact on public discourse, electoral dynamics, and democratic opinion-formation. AI increasingly operates within the digital public sphere as a structuring force of visibility, relevance, agenda-setting, and political salience, thereby affecting the constitutional conditions under which democratic self-government unfolds.

At the same time, the rise of Foreign Information Manipulation and Interference (FIMI) in the broader context of hybrid conflict has revealed that AI is no longer only an economic or technological resource, but also a strategic instrument of cognitive power capable of affecting democratic sovereignty, the integrity of elections, and the resilience of constitutional democracies.

The brief puts forward operational recommendations for European and national policy makers, proposing tools for assessing cognitive impact, new accountability criteria and a strengthening of governance by design of artificial cognitive systems.

1. The paradigm shift: from AI as a tool to the integrated artificial cognitive system

The recent evolution of Artificial Intelligence marks a profound discontinuity with previous forms of digital automation. Generative and agentic AI systems are no longer limited to supporting the execution of tasks or optimizing predefined processes, but intervene directly in the mechanisms through which humans perceive, interpret and evaluate reality.

This transformation has been effectively framed by Chiriatti, Ganapini, Panai, Ubiali and Riva, who propose to read human-AI interaction as the emergence of a new cognitive modality, called **system 0 thinking** (The case for human-AI interaction as system 0 thinking, *Nature Human Behaviour*, 2024). In this perspective, AI operates neither as a simple extension of System 1 (intuitive) nor as a support to System 2 (reflective), but as a pre-decisional cognitive level that acts upstream of conscious processes, structuring the very context within which human thought takes shape.

The crucial point, from a governance perspective, is that these artificial cognitive systems do not just provide information or recommendations, but help define which options are visible, relevant, and plausible. In this sense, AI **participates in the construction of the cognitive frame within which the decision takes place**, exerting an influence that precedes the activation of individual will and legal responsibility.

The result is a form of distributed human-machine cognition, in which the boundary between technological support and cognitive agency becomes progressively more blurred. AI does not replace the human decision-maker, but co-determines the decision-making architecture, affecting attention, priorities, expectations and evaluation criteria. This phenomenon produces systemic effects that cannot be adequately understood or governed through normative categories traditionally oriented towards the act, output or individual unlawful behavior.

Recognizing AI as an integrated cognitive infrastructure therefore implies a paradigm shift for law and public policies: it is no longer just a matter of regulating a technology, but of governing systems that operate on the pre-normative level of the formation of judgment. In the absence of such recognition, the risk is that the power exercised by artificial cognitive systems will remain substantially invisible from a regulatory point of view, although it will produce significant effects on a social, economic and democratic level.

2. Cognitive governance: a new category of public interest

Cognitive governance emerges as an autonomous category of public interest when Artificial Intelligence systems cease to be limited to supporting human decisions and begin to systematically structure the decision-making environments within which these decisions are made. It is no longer just a matter of influencing individual choices, but of designing the cognitive context that makes some options visible, others marginal and others still unthinkable.

This dynamic has been analyzed particularly effectively by MIT Sloan Management Review's research on the concept of **Intelligent Choice Architectures (ICA)** (Schrage, Kiron, *Winning With Intelligent Choice Architectures*, MIT Sloan Management Review, 2025).

According to this approach, systems based on generative and predictive AI do not act as mere recommendation tools, but as architects of choices, capable of dynamically creating, refining and reorganizing the set of options available to human decision-makers.

The crucial step, relevant for law and public policy, consists in the fact that ICAs do not limit themselves to improving the accuracy of decisions, but learn how to improve the decision-making environment itself, intervening on framing, priorities, trade-offs and evaluation criteria.

In other words, AI does not decide *for the human*, but helps decide **what counts as a relevant decision**.

This ability to model the architecture of choices shifts the center of gravity of cognitive power upstream of the conscious decision-making moment. As MIT points out, the real shift of power in the age of AI is not about who makes the final decision, but who designs and governs the cognitive environment within which that decision becomes likely, legitimate, or even inevitable.

In this perspective, cognitive governance cannot be reduced to a question of applied ethics or the protection of the informed consumer. It directly invests:

- individual decision-making autonomy,
- the distribution of power in organizational and institutional processes,
- the quality of the democratic process,
- the transparency of high-impact public and private choices.

A further critical element lies in the fact that intelligent choice architectures are not neutral. They incorporate assumptions, values, objectives, and performance

metrics that reflect priorities that are often implicit or opaque. MIT emphasizes how ICAs can generate cognitive empowerment, but also reinforce systemic biases or produce forms of heterodirection that are invisible if not properly governed.

From a legal point of view, this implies that the traditional focus on the responsibility of the act or on the output of the system is insufficient. The central node becomes responsibility for the design of decision-making architectures, i.e. for those meta-decisions that determine which choices are made accessible, which trade-offs are highlighted and which rationality criteria are privileged.

Cognitive governance is therefore configured as a new frontier of public interest, which requires regulatory tools capable of intervening not only on the use of AI, but on the structure of artificial cognitive contexts. In the absence of such intervention, the architectures of choice risk being transformed into forms of de facto regulation, exercised by technical systems that silently assume para-normative functions.

3. Regulatory governance: the limits of the current legal approach, and its cognitive blind spot

Positive law has begun to respond to the challenges posed by AI through highly relevant tools, such as the European Regulation on Artificial Intelligence, the Digital Services Act and the strengthening of personal data protection protections. However, these regulatory instruments remain largely anchored to a **risk-and-compliance paradigm**, whereas the most impactful effects of AI increasingly operate at the **cognitive and pre-decisional level**. However, these instruments share a structural limitation: they mainly intervene on outputs, identifiable risks and ex post responsibilities.

AI systems, on the other hand, exert their main impact upstream, in the configuration of decision-making environments and in the structuring of available options. The result is a misalignment between the time of law and the time of technology, aggravated by the proprietary opacity of algorithmic architectures and their rapid evolution.

This gap risks making regulatory governance always reactive and never really preventive. The core problem is that **regulatory AI governance and cognitive governance have so far evolved on parallel but insufficiently integrated tracks**: law focuses on outputs, harms, and liabilities, while AI systems exercise influence on framing, attention, prioritization, and decision architectures upstream of formal accountability.

4. From the compliance model to integrated cognitive-regulatory governance by design

To bridge this gap, a conceptual shift is needed: from regulating the use of AI to the **integrated governance of artificial cognitive architectures**. This means explicitly linking **regulatory AI governance** (obligations, supervision, enforcement) with **cognitive AI governance** (control over framing, choice architectures, sensemaking, and human decision autonomy).

This implies the introduction of governance by design principles that take into account not only the technical performance of the system, but also its impact on human cognitive processes. In this perspective, transparency cannot be limited to the explainability of the output, but must include the traceability of decision-making logics and information framing mechanisms.

Similarly, accountability must extend from the responsibility of individual use to **systemic responsibility for the decision-making architectures embedded in the models**. In this integrated perspective, compliance is no longer limited to legal risk mitigation, but becomes a **mechanism for preserving cognitive autonomy, democratic integrity, and constitutional proportionality**.

5. Policy recommendations: aligning cognitive and regulatory AI governance

In light of the above considerations, the policy brief proposes some operational recommendations:

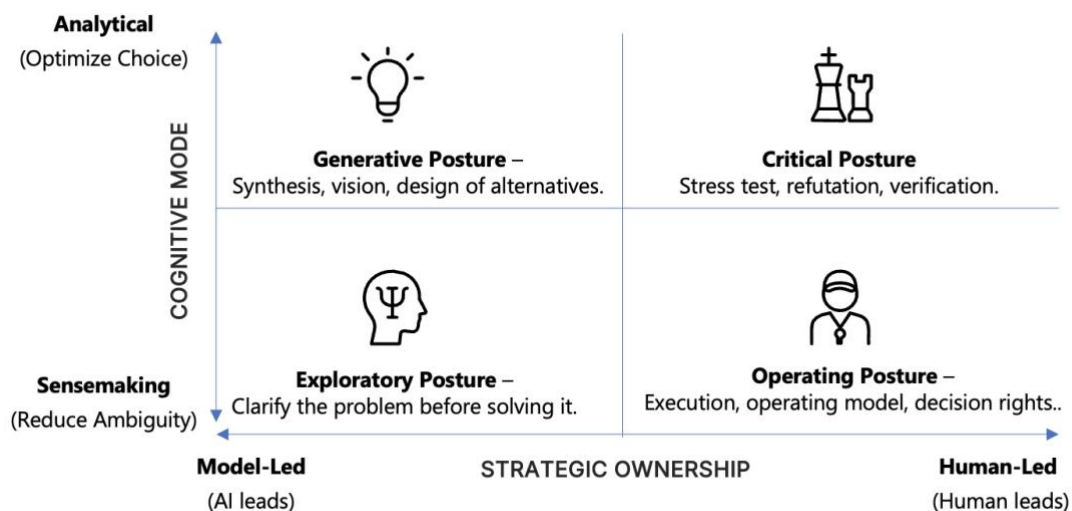
- Introduction of a **Cognitive Impact Assessment (CogIA)**, complementary to the risk assessment provided for by the AI Act, to assess the impact of AI systems on human decision-making processes. CogIA should be treated as a **regulatory obligation**, embedding cognitive-impact criteria directly into AI Act compliance and enforcement mechanisms. The CogIA should employ structured methodologies such as the Thinking Quadrants framework to evaluate whether systems preserve or compromise cognitive autonomy across all decision-making phases. High-risk AI systems should be required to pass CogIA certification before deployment.
- Regulatory classification of AI systems based on their cognitive role (informational, decision-making, agentic), with differentiated regimes of obligations and responsibilities.
- Strengthening of human oversight mechanisms, understood not only as functional control, but as an effective possibility of cognitive intervention and critical review of the decisions suggested by the system.
- Regulatory experimentation (sandboxes) oriented towards cognitive governance, involving independent authorities, companies and the scientific community.

- Expansion of the powers of supervisory authorities, including the evaluation of decision-making architectures and not just outputs or individual applications.

Methodology

Thinking Quadrants as the foundation of Cognitive Impact Assessment

To operationalize the assessment of the cognitive impact of high-impact AI systems, this policy brief proposes the adoption of the **Thinking Quadrants framework**, developed as part of the Systemic Zero research.



Thinking Quadrants (Decision Postures). The vertical axis captures the cognitive mode—Analytical (optimise choice) versus Sensemaking (reduce ambiguity). The horizontal axis captures Strategic Ownership—Model-Led (AI leads the reasoning process) versus Human-Led (humans lead, AI supports). The four postures describe recurring ways executives engage AI in strategic work. **Source: Systemic Zero (2026).** A managerial publication further developing the framework is currently in preparation

The Thinking Quadrants are an analysis methodology designed to assess how an artificial cognitive system intervenes on the architecture of human and collective thought, influencing the framing of decisions even before conscious choice. The framework stems from the assumption that the main risk in the adoption of AI is not technological lock-in, but **cognitive lock-out**: the progressive shift in the level of structuring of thought and decision-making priorities towards external and opaque systems.

The method articulates the Cognitive Impact Assessment through four decision postures, enabling assessment not only of outputs, but of how the system shapes framing, option-generation, verification, and execution governance:

- **Exploratory Posture (Sensemaking | Model-Led):** Does the system expand and clarify the framing, or does it anchor users to a premature interpretation?
- **Generative Posture (Analytical | Model-Led):** Does it generate genuinely diverse alternatives, or does it converge too quickly toward a single narrative?
- **Critical Posture (Analytical | Human-Led):** Does it enable structured refutation and verification, or does it create an illusion of robustness through fluent argumentation?
- **Operating Posture (Sensemaking | Human-Led):** Does it preserve accountability (decision rights, traceability, escalation paths), or does it obscure decision chains and weaken governance?

The diagonal bypass risk identifies conditions in which AI short-circuits human sensemaking, producing **cognitive lock-out**.

6. Conclusions

Artificial Intelligence is profoundly reshaping the way individuals and institutions think, decide and act. In the absence of adequate cognitive governance, regulatory governance risks being structurally lagging behind, intervening when systemic effects are already consolidated.

Governing AI today means governing the **cognitive power** it wields. Doing so in a timely and informed manner is a necessary condition for preserving decision-making autonomy, the quality of democracy and trust in digital technologies.

Giovanni Giamminola is an expert in cognitive governance and AI-driven decision architectures. He is the originator of the Systemic Zero framework and of the Thinking Quadrants methodology, which provide the conceptual foundation for Cognitive Impact Assessment and AI governance by design. He works with boards, public institutions and senior leaders on the governance of artificial cognitive systems and advanced decision-making environments. He has taught in executive education programmes at the UCL School of Management (University College London) and is the author of *The Augmented Manager* and *The Augmented Brand*.