



15/10/2026

## Dual use, opacity and deterrence: artificial intelligence and the erosion of strategic signalling

1

## Introduction

Natural, demographic and technological resources are distributed unevenly among states. From a realist perspective, this asymmetry, together with the absence of an effective supranational authority, fuels competition for security. States seek to ensure their survival through military power, understood both as the capacity to project force externally and to defend against threats from others and protect their population, territory and institutions<sup>1</sup>. Relative superiority may facilitate coercion, but it does not make war inevitable: it can also sustain defence, negotiation and deterrence.

In this context, governments use signalling mechanisms to communicate capabilities and intentions and interpret those of their adversaries. The work of Thomas Schelling and Robert Jervis helps explain how threats, commitments and perceptions influence strategic calculations<sup>2</sup>. Exercises, arms procurement and technological tests provide information, although their meaning depends on the political context and the credibility of the actor conducting them.

The rise of artificial intelligence (AI) as a technology with applications across civilian and military sectors changes this architecture. A system's behaviour can change through adjustments to its models, data or software without its outward appearance revealing the transformation. This article argues that the gap between what is observable and actual capability erodes certain foundations of strategic signalling and calls for the adaptation of confidence-building and verification instruments.

This is a conceptual analysis drawing on strategic scholarship, institutional documents and public examples. It does not seek to measure the opacity of every AI system or demonstrate a uniform causal relationship between AI adoption and conflict. Its purpose is to identify mechanisms of uncertainty and the conditions under which they may affect the credibility of signals.

## The international system as an information dynamic

In offensive realism, the pursuit of relative power is central to explaining the behaviour of great powers<sup>3</sup>. Insecurity can, however, also lead to defensive strategies: a state may

---

<sup>1</sup> MEARSHEIMER, John J. *The Tragedy of Great Power Politics*. W. W. Norton, New York, 2001.

<sup>2</sup> SCHELLING, Thomas C. *The Strategy of Conflict*. Harvard University Press, Cambridge (MA), 1960. SCHELLING, Thomas C. *Arms and Influence*. Yale University Press, New Haven, 1966. JERVIS, Robert. *Perception and Misperception in International Politics*. Princeton University Press, Princeton, 1976.

<sup>3</sup> MEARSHEIMER, John J. *The Tragedy of Great Power Politics*, op. cit.

strengthen its capabilities to resist aggression, while its adversary interprets that reinforcement as preparation for an offensive. The problem concerns not only the material distribution of power but also the information available about its purpose.

States therefore face recurring questions: how strong is the adversary? How credible is its threat or its willingness to accept an agreement? The theory of strategic bargaining shows that communicating capabilities and resolve can alter incentives without the actual use of force<sup>4</sup>. A useful signal must convey relevant information and be sufficiently costly, verifiable or binding to distinguish it from a mere statement.

When signalling works, the parties adjust their expectations and expand the scope for negotiation. When it fails through deception, misinterpretation or difficulties with verification, the risk of miscalculation grows. Information alone does not explain the outbreak of war: incompatible interests, commitment problems and political decisions also matter. Its quality nevertheless shapes the ability to anticipate costs, recognise limits and avoid unintended escalation. Stability therefore depends both on credible defensive capabilities and on communication that allows others to interpret them.

It is also important to distinguish between providing information about a capability and committing to use it. An exercise may demonstrate a degree of readiness without revealing what a government will decide under pressure. Conversely, a political statement may express resolve without establishing the means available. Strategic interpretation combines both dimensions and checks them against observed behaviour. This distinction avoids assigning technical transparency a role that also depends on political relationships and the confidence built up between the parties.

### **Classical strategic signalling: from nuclear weapons to conventional deterrence**

Nuclear weapons offer a particularly relevant case of signalling. Their high cost, the facilities they require and certain physical signatures of their development provide external indications of capability<sup>5</sup>. Explosive tests can be detected using seismic, hydroacoustic, infrasound and radionuclide technologies<sup>6</sup>. Such observability does not

---

<sup>4</sup> SCHELLING, Thomas C. *The Strategy of Conflict and Arms and Influence*, op. cit.

<sup>5</sup> BRODIE, Bernard. *Strategy in the Missile Age*. Princeton University Press, Princeton, 1959. FREEDMAN, Lawrence. *The Evolution of Nuclear Strategy*. 3rd ed., Palgrave Macmillan, Basingstoke, 2003.

<sup>6</sup> CTBTO. *Monitoring Technologies*. Available at: <https://www.ctbto.org/our-work/monitoring-technologies> (accessed: 03/10/2026).

amount to complete transparency: uncertainty persists over arsenals, doctrines, readiness and willingness to use these weapons.

Legal and technical verification instruments were built around some of these properties. The Treaty on the Non-Proliferation of Nuclear Weapons (NPT) establishes differentiated obligations, while the safeguards of the International Atomic Energy Agency (IAEA) verify compliance with commitments concerning the peaceful use of nuclear materials and technology<sup>7</sup>. This system should not be confused with a general inspection of all warheads or military delivery systems. Its effectiveness depends on the applicable agreements, available access and state cooperation.

The nuclear experience shows that deterrence can be sustained with incomplete information. Retaliatory capability and the prospect of unacceptable damage have helped restrain direct confrontation between nuclear powers, although they have not eliminated crises or proxy wars. Stability derives from a combination of capabilities, signals, institutions and political learning rather than from a perfectly transparent technology<sup>8</sup>.

The space race illustrates that dual use is not new either. The launch of Sputnik in 1957 was a scientific achievement and a signal of Soviet progress in developing long-range rockets<sup>9</sup>. Its orbit was observable by third parties, although it did not by itself demonstrate every performance characteristic of an operational arsenal. This distinction matters: physical evidence can establish a specific capability without resolving every uncertainty about its military use.

The problem posed by AI should be framed comparatively. Many technologies combine civilian and military uses; what matters is the relationship that can be established between external evidence and a specific capability. Concealment and deception also affect conventional systems. The difference emerges when an increasing share of capability depends on informational components that can be modified without equivalent material changes.

---

<sup>7</sup> UNITED NATIONS. Tratado sobre la No Proliferación de las Armas Nucleares. 1968, Article III. Available at: <https://www.un.org/en/conf/npt/2010/npttext.shtml> (accessed: 03/10/2026). OIEA. Nuclear safeguards explained. <https://www.iaea.org/topics/safeguards-explained> (accessed: 03/10/2026).

<sup>8</sup> FREEDMAN, Lawrence. The Evolution of Nuclear Strategy, op. cit.

<sup>9</sup> NASA. 65 Years Ago: Sputnik Ushers in the Space Age. 4 October 2022. Available at: <https://www.nasa.gov/history/65-years-ago-sputnik-ushers-in-the-space-age/> (accessed: 03/10/2026).

## Artificial intelligence as a dual-use technology: revisiting classical foundations

AI can contribute to intelligence analysis, logistics, decision support and the operation of platforms with varying degrees of autonomy. Its military impact depends on technical integration, organisation and training as well as algorithmic performance<sup>10</sup>. Nor should AI, automation and autonomy be treated as interchangeable: they are related concepts, but an uncrewed platform may be remotely controlled, and a support system can incorporate AI without making decisions independently<sup>11</sup>.

Industrial developments illustrate this diversity. In June 2025, Saab announced three flights with Helsing's Centaur agent integrated into a Gripen E; the first took place on 28 May<sup>12</sup>. In September 2025, Helsing presented the CA-1 Europa project through a full-scale design model, with entry into service still a development objective<sup>13</sup>. In March 2026, the US Army announced an enterprise contract with Anduril, worth up to twenty billion dollars, to consolidate procurement of its commercial solutions<sup>14</sup>. Tests, industrial announcements and contract ceilings constitute different signals; none alone establishes full deployment or combat performance.

These developments relate to multi-domain operations. NATO distinguishes the land, maritime, air, space and cyberspace domains, with effects that also extend into the cognitive dimension<sup>15</sup>. AI can facilitate information fusion and improve coordination across domains, but it is not an essential prerequisite: attacks and responses can be planned and coordinated through traditional procedures. The experience of Ukraine, examined in IEEE publications, shows the interaction of conventional, digital and

---

<sup>10</sup> HOROWITZ, Michael C. Artificial Intelligence, International Competition, and the Balance of Power. Texas National Security Review, vol. 1, no. 3, 2018, pp. 36–57. Available at: <https://doi.org/10.15781/T2639KP49> (accessed: 03/10/2026).

<sup>11</sup> BOULANIN, Vincent and VERBRUGGEN, Maaïke. Mapping the Development of Autonomy in Weapon Systems. SIPRI, Stockholm, 2017. Available at: <https://www.sipri.org/publications/2017/policy-reports/mapping-development-autonomy-weapon-systems> (accessed: 03/10/2026). SCHARRE, Paul. Army of None: Autonomous Weapons and the Future of War. W. W. Norton, New York, 2018.

<sup>12</sup> SAAB. Saab achieves AI milestone with Gripen E. 11 June 2025. Available at: <https://www.saab.com/newsroom/press-releases/2025/saab-achieves-ai-milestone-with-gripen-e> (accessed: 03/10/2026).

<sup>13</sup> HELSING. Helsing unveils CA-1 Europa, an autonomous fighter jet. 25 September 2025. Available at: <https://helsing.ai/newsroom/helsing-unveils-ca-1-europa-an-autonomous-fighter-jet> (accessed: 03/10/2026).

<sup>14</sup> U.S. ARMY PUBLIC AFFAIRS. U.S. Army awards enterprise contract for IT commercial solutions. 13 March 2026. Available at: [https://www.army.mil/article/291074/u\\_s\\_army\\_awards\\_enterprise\\_contract\\_for\\_it\\_commercial\\_solutions](https://www.army.mil/article/291074/u_s_army_awards_enterprise_contract_for_it_commercial_solutions) (accessed: 03/10/2026).

<sup>15</sup> NATO ALLIED COMMAND TRANSFORMATION. Multi-Domain Operations. Available at: <https://www.act.nato.int/activities/multi-domain-operations/> (accessed: 03/10/2026).

uncrewed capabilities, whose outcomes also depend on command and organisational adaptation<sup>16</sup>.

The information dimension adds another difficulty. Generative models can lower the cost of producing synthetic content and facilitate misleading representations of capabilities or events<sup>17</sup>. This does not remove the possibility of checking an image or a statement against sensors, independent sources and contextual analysis. It does increase the authentication effort required for public signals to be credible. Deception predates AI; its scale, speed and accessibility are changing.

Infrastructure is a second source of ambiguity. A data centre may host civilian services and military applications, although not every facility has the specifications needed to train advanced models. Chips, power supplies and facilities remain physical and potentially observable; it is harder to infer which model is running, with what data and for what purpose<sup>18</sup>. Computing capacity provides a partial indicator, not a direct measure of military effectiveness.

The distributed nature of the value chain complicates this inference. Models, data and applications can move between organisations, while integration with sensors, communications and operational procedures determines their practical utility. Counting equipment or inspecting facilities therefore remains informative, but is insufficient to verify every capability. Table 1 summarises these differences as analytical tendencies that vary according to the system under consideration.

---

<sup>16</sup> SAURÍN MARTÍNEZ, Daniel. Origen y fundamentos teóricos de la guerra multidominio y su aplicación en la guerra de Ucrania. Documento de Opinión IEEE 100/2025. Available at: [https://www.defensa.gob.es/ceseden/-/ieeee/guerra\\_multidominio\\_2025\\_dieeeo100](https://www.defensa.gob.es/ceseden/-/ieeee/guerra_multidominio_2025_dieeeo100) (accessed: 03/10/2026). ABAD SOTO, Jorge. Cambio de paradigma de la seguridad y la defensa: capacidades futuras, ya presentes. Documento de Opinión IEEE 04/2025. [https://www.defensa.gob.es/ceseden/-/ieeee/cambio\\_de\\_paradigma\\_de\\_la\\_seguridad\\_y\\_la\\_defensa\\_2025\\_dieeeo04](https://www.defensa.gob.es/ceseden/-/ieeee/cambio_de_paradigma_de_la_seguridad_y_la_defensa_2025_dieeeo04) (accessed: 03/10/2026).

<sup>17</sup> GARCÍA HUETE, Estíbaliz et al. IA generativa y sistemas serverless en desinformación bélica con LUTs: impacto y ética. Documento de Opinión IEEE 45/2025. Available at: [https://www.defensa.gob.es/ceseden/-/ieeee/ia\\_generativa\\_y\\_sistemas\\_serverless\\_en\\_desinformacion\\_belica\\_con\\_luts\\_impacto\\_y\\_etica](https://www.defensa.gob.es/ceseden/-/ieeee/ia_generativa_y_sistemas_serverless_en_desinformacion_belica_con_luts_impacto_y_etica) (accessed: 03/10/2026).

<sup>18</sup> GARCÍA RETUERTA, Oscar and GARCÍA RETUERTA, David. La cadena de valor de la inteligencia artificial: estrategias de autonomía para España. Documento de Opinión IEEE 03/2026. Available at: [https://www.defensa.gob.es/ceseden/-/ieeee/cadena\\_de\\_valor\\_ia\\_2026\\_dieeeo03](https://www.defensa.gob.es/ceseden/-/ieeee/cadena_de_valor_ia_2026_dieeeo03) (accessed: 03/10/2026).

Table 1. Classical signalling and AI-based capabilities

| Dimension            | Classical capabilities                     | AI-based capabilities                            |
|----------------------|--------------------------------------------|--------------------------------------------------|
| External evidence    | Physical signatures; conditional access    | Observable infrastructure; less visible function |
| Location             | Identifiable facilities and platforms      | Physical computing sites; distributed services   |
| Dual use             | Present in components and delivery systems | Widespread in data, models and infrastructure    |
| Attribution          | Varies by system and context               | Complex in cyberspace; requires multiple sources |
| Verification         | Sector-specific regimes with limitations   | Specific evaluations and partial agreements      |
| Entry costs          | High for strategic systems                 | Variable; high for advanced capabilities         |
| Public demonstration | Establishes specific performance           | Depends on version and test conditions           |

Source: own elaboration based on the works cited. AI: artificial intelligence. Conceptual comparison, not an exhaustive classification.

The comparison must also account for time. A demonstration provides information about a particular system version under specific conditions; its value declines if the data, connections or operating environment change. Interpreting a signal requires more than knowing the outcome of a test: it matters which aspect of capability the test establishes and how long it remains representative. This dynamic character complicates agreements based exclusively on inventories and shifts some attention towards evaluation procedures. The difference concerns degree and the object of assessment, rather than the disappearance of any material basis.

## Implications: opacity, attribution and the erosion of deterrence

Opacity affects deterrence calculations. Deterrence may rest on denying an adversary its objectives or threatening to impose costs. Both approaches require capabilities and the willingness to use them to be sufficiently credible. If an update changes performance characteristics that are difficult to observe, the scope for underestimating or overestimating a rival increases. This uncertainty may encourage miscalculation, but it can also induce caution; its effect is not automatic and depends on the strategic situation<sup>19</sup>.

Cyber attribution presents a related problem. Intermediaries, the reuse of tools and the manipulation of evidence already made it difficult to identify those responsible before AI became widespread<sup>20</sup>. AI may amplify some risks, but it can also support defensive analysis<sup>21</sup>. Technical attribution, intelligence assessment and the political decision to make an accusation public should be distinguished. Uncertain attribution may weaken the threat of retaliation without negating other forms of deterrence based on resilience, defence and cooperation.

The object of international verification is also changing. In early-warning, analytical support or nuclear command-and-control systems, algorithmic components may affect the interpretation of critical information<sup>22</sup>. Verifying a physical facility does not necessarily reveal how a model will behave beyond test conditions. Evaluations require functions, operating environments and reliability criteria to be specified. Access to sensitive information also entails political and security constraints that make it difficult to transfer traditional inspection models directly.

Meaningful human control is a relevant response, but it must translate into an effective capacity to understand and intervene. The formal presence of an operator is insufficient if that person lacks the time, information or authority to change a decision. System design must allow limitations to be recognised, failures to be managed and clear responsibilities

---

<sup>19</sup> JERVIS, Robert. Perception and Misperception in International Politics, op. cit.

<sup>20</sup> BUCHANAN, Ben. The Hacker and the State: Cyber Attacks and the New Normal of Geopolitics. Harvard University Press, Cambridge (MA), 2020.

<sup>21</sup> BRUNDAGE, Miles et al. The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. 2018, arXiv:1802.07228. Available at: <https://arxiv.org/abs/1802.07228> (accessed: 03/10/2026).

<sup>22</sup> RODRÍGUEZ LA BARRERA, Luciana Alejandra. Armas nucleares e inteligencia artificial: un equilibrio entre la automatización y el factor humano. Documento de Opinión IEEE 10/2026. Available at: [https://www.defensa.gob.es/ceseden/-/iee/armas\\_nucleares\\_2026\\_dieeo10](https://www.defensa.gob.es/ceseden/-/iee/armas_nucleares_2026_dieeo10) (accessed: 03/10/2026).

to be maintained<sup>23</sup>. Human oversight helps reduce operational risks; it does not, on its own, replace external verification of international commitments.

Asymmetries in the value chain add a structural dimension. Access to semiconductors, infrastructure, data and talent shapes states' autonomy, while competition between great powers relies on incomplete indicators<sup>24</sup>. The spread of certain tools lowers entry barriers, but developing and sustaining advanced capabilities may require substantial resources. It is therefore misleading to characterise all AI as inexpensive or to assume that access to a model amounts to an integrated military capability.

In the information environment, synthetic content can erode trust in public messages, evidence and commitments. Signalling requires distinctions between announced capability, demonstrated capability and capability available under operational conditions. A reasonable response combines source authentication, independent corroboration and communication of uncertainty. The aim is not to eliminate all opacity, an objective difficult to reconcile with national security, but to prevent ambiguity from making it impossible to recognise limits and commitments during a crisis.

Uncertainty about an adversary should be distinguished from uncertainty about one's own system. An organisation may not know how a tool will respond outside familiar conditions, even when it has internal access to it. If it communicates a capability on the basis of limited testing, it may convey confidence that is not justified even for those responsible for the system. Signalling therefore requires attention both to external credibility and to the soundness of an organisation's own knowledge. The risk of miscalculation arises not only from deliberate concealment but also from excessive expectations, unrepresentative results or incompatible interpretations of the same demonstration. This distinction allows the problem to be framed without assuming that every instance of opacity reflects a strategy of deception.

Doctrinal adaptation should also consider defensive signals. Demonstrating the ability to maintain essential services, restore functions or retain control over a decision may strengthen deterrence without revealing details of offensive capabilities. The aim is not to turn these demonstrations into universal guarantees, but to make clear which commitment they support and which uncertainties remain. During a crisis, cautious

---

<sup>23</sup> BOULANIN, Vincent and VERBRUGGEN, Maaïke. Mapping the Development of Autonomy in Weapon Systems; SCHARRE, Paul. *Army of None*, op. cit.

<sup>24</sup> SCHARRE, Paul. *Four Battlegrounds: Power in the Age of Artificial Intelligence*. W. W. Norton, New York, 2023.

statements about limitations may be more useful than promises of autonomy or general superiority. Credibility depends on consistency between the message, the available evidence and subsequent behaviour. AI introduces new difficulties into this relationship, but does not remove the role of political judgement or the need for stable channels of communication.

## The regulatory landscape and Europe's position

Regulation (EU) 2024/1689 establishes obligations differentiated by risk and a phased application timetable. Its strategic scope has an essential limit: Article 2(3) excludes systems insofar as they are intended exclusively for military, defence or national security purposes. Civilian or mixed uses require an assessment of the relevant scope of application<sup>25</sup>. The Regulation therefore does not constitute a general arms-control regime for AI-based systems.

Spain has the Spanish Agency for the Supervision of Artificial Intelligence (AESIA), whose statute was approved by Royal Decree 729/2023<sup>26</sup>. The 2024 Artificial Intelligence Strategy updates the effort initiated with the 2020 national strategy and links technological capacity, AI adoption and responsibility<sup>27</sup>. These instruments can provide experience in oversight and evaluation, but their existence does not guarantee military verification capability or international acceptance of European standards.

At the Allied level, NATO's AI Strategy, revised in July 2024, sets out six principles of responsible use: lawfulness, responsibility and accountability, explainability and traceability, reliability, governability and bias mitigation<sup>28</sup>. For Spain, the challenge is to connect military interoperability, the industrial base and evaluation practices while respecting the distinction between civilian legal obligations and doctrinal commitments. Governability includes the ability to intervene when unwanted behaviour occurs; it should not be reduced to a general statement about good management.

---

<sup>25</sup> EUROPEAN PARLIAMENT AND COUNCIL. Regulation (EU) 2024/1689 of 13 June 2024. Official Journal of the European Union, 12 July 2024, Article 2(3), recital 24 and Article 113. Available at: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj?locale=en> (accessed: 03/10/2026).

<sup>26</sup> GOVERNMENT OF SPAIN. Royal Decree 729/2023 of 22 August approving the Statute of AESIA. BOE no. 210, 2 September 2023. Available at: <https://www.boe.es/eli/es/rd/2023/08/22/729> (accessed: 03/10/2026).

<sup>27</sup> GOVERNMENT OF SPAIN. IA: ¿qué es la Estrategia de Inteligencia Artificial 2024? La Moncloa, 5 June 2024. Available at: <https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/transformacion-digital-y-funcion-publica/Paginas/2024/ia-inteligencia-artificial-estrategia-espana.aspx> (accessed: 03/10/2026).

<sup>28</sup> NATO. Summary of NATO's revised Artificial Intelligence (AI) strategy. 10 July 2024. Available at: <https://www.nato.int/en/about-us/official-texts-and-resources/official-texts/2024/07/10/summary-of-natos-revised-artificial-intelligence-ai-strategy> (accessed: 03/10/2026).

This distinction suggests a cooperation agenda that warrants examination. Sharing evaluation criteria or communicating relevant incidents may be less intrusive than demanding full access to models, data and command systems. None of these measures, however, offers automatic guarantees: participants must agree on what information allows a commitment to be checked and how to prevent its exchange from exposing vulnerabilities. Reciprocity, the protection of sensitive information and the ability to revise mechanisms as technology changes are conditions for feasibility. Their strategic value should be judged by their contribution to better-informed decisions, particularly during periods of tension.

## Conclusions

Strategic signalling contributes to stability when it allows capabilities and intentions to be interpreted with sufficient credibility. Defence against external threats forms part of this logic, alongside the ability to impose costs. Nuclear and conventional experience offers useful instruments, but also shows that deterrence has always operated with incomplete information and risks of misperception.

AI changes some foundations of this exchange by weakening the correspondence between visible infrastructure and actual capability. Its dual-use character, software updates and dependence on the context of use complicate external inference. Physical signatures persist, and verification remains possible; what changes is the combination of evidence needed to sustain a reasonable assessment.

Europe's response should recognise the limits of civilian regulation and develop its connections with defence frameworks. Discussions in the Group of Governmental Experts on lethal autonomous weapons systems, under the Convention on Certain Conventional Weapons, provide a relevant forum, although their scope does not cover every military application of AI<sup>29</sup>. Translating principles into verifiable commitments requires agreement on scope, admissible evidence and the protection of sensitive information.

As a doctrinal direction, this article proposes gradual transparency about functions and limitations, agreed evaluations and incident communication mechanisms. These

---

<sup>29</sup> UNITED NATIONS. Convention on Certain Conventional Weapons: Group of Governmental Experts on Lethal Autonomous Weapons Systems (2026). Available at: <https://meetings.unoda.org/meeting/79329> (accessed: 03/10/2026).

measures should be assessed in terms of technical feasibility, reciprocity and usefulness in crises. Spain can contribute through evaluation capabilities, Allied cooperation and multilateral engagement. The aim is to rebuild sufficient conditions for confidence so that technological innovation does not increase the risk of conflict through miscalculation without countervailing safeguards.

*Pau Muñoz Pairet \**

PhDs in Artificial Intelligence and Strategic Studies  
Lecturer in International Relations, Universidad Pontificia Comillas