



22 de octubre de 2026

Algocracy: The Threats of Automated Decision-Making

Introduction

Modern society has been characterised by unprecedented progress in its technological development. Since the Cold War and the competition between the blocs to demonstrate which was more advanced, society has developed through iteratively complex technological processes that have shaped contemporary reality. In this context, it was only a matter of time before the human mind's capacity to process data and solve problems was overtaken by the need for rapid, reliable and accurate results on an unimaginable scale. This led to the emergence of the first algorithm-based reasoning systems, which gradually evolved into modern artificial intelligence (AI) and neural networks. These systems have achieved such a degree of analytical refinement that they are able to mimic human cognitive behaviour with astonishing accuracy.

Consequently, algorithms, AI, neural networks and related technologies have, over recent decades, become increasingly integrated into every aspect of modern life. Beyond virtual assistants or conversation simulators – chatbots – to which we direct our everyday curiosity, we have discreetly integrated these technologies into internet search engines, content platforms, consumer forecasting, GPS navigation, industrial automation, smartphones... They are so deeply embedded in our lives that, in reality, they play a part in almost every significant aspect of our society.

There is even a trend for qualified professionals to carry out highly technical enquiries to make their work easier. Examples include: searches for case law by judges and lawyers; the use by politicians, public relations specialists and statisticians to analyse social patterns and public interests; or facial and fingerprint recognition programmes that streamline police work.

From an efficiency perspective, it is legitimate to use these mechanisms to speed up decision-making, but this raises the real question addressed in this article: where should we draw the line regarding the trust we place in these systems?

There is an increasingly clear tendency to rely excessively on algorithms, overlooking the fact that ultimate responsibility lies with human beings and not with the AI that processes the information. If we were to place our trust in algorithms to the extent that we no longer question their judgement, we could fall into a state of 'civic apathy' that would lead to one

of the central pillars of this article: a government governed directly or indirectly by the decisions of an AI. This form of government is what is known as ‘algocracy’.

Nothing of the sort has ever existed, nor does it seem feasible in the short term, but it serves as a speculative framework for analysing just how far our ever-increasing reliance on algorithms and AI might take us. It is not for nothing that entrusting certain aspects of life and governance to machines or *software* can entail significant strategic, ethical and social dangers, which are already being widely documented. This article therefore aims to analyse the risks we take on when we delegate part of our decision-making power to algorithms.

‘Algocracy’: a not-so-new trend

It is at this point that we enter fully into the realm of the purely speculative, as befits the hypothesis underlying the concept of ‘algocracy’. Essentially, it involves considering what would happen if, over time, our trust in AI were to shift from a purely advisory role to the delegation of sensitive, complex or far-reaching tasks.

In such circumstances, with the necessary technological development, we could reach a point where a sufficiently sophisticated artificial intelligence would assume decision-making powers which, *de facto*, would amount to control over all aspects governing society. We could say that this AI would be the one truly in charge. This would pose a significant risk if combined with the population’s lack of interest in its own governance (‘civic apathy’) — a phenomenon that already exists in society for other reasons. In that case, one might consider that the system of government would have become a new form of totalitarianism arising from social withdrawal: an ‘algocratic’ dictatorship¹.

Technically, ‘algocracy’ could be defined as a political phenomenon—as yet hypothetical—in which algorithms, particularly those based on artificial intelligence and machine learning, become active decision-makers regarding the governance of a territory, region or state. To this end, they would supplant or constrain national sovereignty, strategic autonomy and the

¹ VILLARROEL SOTO, Raúl. ‘Algocracy: Political Decisions and Citizen Representation in Algorithmic Governance’, **Otro siglo* Philosophy Journal*. 27 October 2023, pp. 193–212. ISSN 0719-921X. DOI: 10.5281/zenodo.10451886.

capacity for democratic decision-making in favour of automated systems managed by technocratic structures, which would likely be privately run or transnational.

From a philosophical perspective, this would represent a shift in the paradigm of authority—a conceptual evolutionary iteration that would entail a reconfiguration of the notions of power, responsibility and subjectivity. The challenge would arise of how to preserve human dignity when a logical machine is responsible for making decisions that may prove momentous for the individual. It would represent the ultimate paradigm of Max Weber's instrumental rationality, running the risk of leading society into an 'iron cage' of dehumanised efficiency².

At first glance, it may seem merely a new form of technocracy or digital democracy, or even a new form of virtual bureaucracy, but 'algocracy' differs from these concepts in its entirely automatic nature, its algorithmic logic (apparently) devoid of human bias, and its impact on sovereignty.

Various types of 'algocracy' can be defined: based on their aims, who controls the *software*, whether there is social collusion in its establishment, or whether it has been a gradual and imperceptible process. However, some authors consider that the most relevant classification would be based on the implicit consequences of the decision-making processes, that is, whether its recommendations become directly binding or whether ratification by a human who takes responsibility for them is required³.

Nevertheless, the thought experiment proposed in this article is not entirely speculative. Decision-making mechanisms are already being implemented today, utilising algorithms, AI and neural networks to make decisions of governmental significance. Some examples include the automation of the review of legal documents in the EU, the prediction of failures in bridges and roads in the US, automated customer service via *chatbots* in many local authorities and public bodies around the world, the optimisation of credit approvals, the automatic receipt of complaints against the Mexican police, and the assessment of human resources in some UK agencies⁴.

² GIL VILLEGAS, Francisco. 'The Concept of Rationality in the Work of Max Weber', *Revista Mexicana de Ciencias Políticas y Sociales*, 31(119). 26 November 2019. DOI: 10.22201/fcpys.2448492xe.1985.119.72143.

³ CELESTE DANESI, Cecilia. 'The danger of automating decisions with AI without democratic oversight', *OnIA, Ethical AI*. Buenos Aires, 15 June 2025. Available at: www.elnacional.cat/oneconomia/es/on-ia/peligro-automatizar-decisiones-con-ia-sin-control-democratico_1434472_102.html (accessed 15 September 2025).

⁴ WWW.GOBIERTO.ES. 'Artificial Intelligence in public administration: use cases'. Available at: <https://www.gobierno.es/blog/20190920-inteligencia-artificial-administracion->

The automation of the public sector: benefits or illusions

When discussing the benefits that the use of AI brings to public administration and governance, we must not limit our analysis to the perspective of operational efficiency alone. The use of data analysis tools brings about a substantial improvement in decision-making, as it facilitates the identification of patterns and trends, thereby enabling public policies to be guided by logical evidence.

Furthermore, AI has no economic interests of its own — although this statement will be qualified later on — so its use allows for greater transparency and traceability in decision-making. This minimises the opacity of public administration and facilitates accountability in the event of erroneous or underhand decisions. In this regard, the use of AI would help reduce the risk of fraud through mechanisms for the early detection of irregularities, automatically monitoring transactions and expenditure records in a constant search for anomalies that indicate possible breaches⁵.

Furthermore, the use of AI to address citizens' requests and needs would enable a more professional, efficient and immediate service, improving the citizen's experience and even allowing for the intelligent targeting of benefits and resources. This would increase trust in the public administration since, on the one hand, it is perceived as a sign of modernisation and professionalisation, whilst, on the other, it would reduce human error and the risk of administrative arbitrariness⁶. In short, it would be perceived as a more professional, equitable, automated and efficient public administration, with the capacity to respond to the needs of the society it serves with a speed that is often lacking in today's bureaucracy.

However, the use of AI in administration and government is not without its drawbacks. The main problem facing a society when implementing these systems is the misconception that algorithms, because they are based on mathematical and statistical logic, are free from human bias and are completely objective and impartial in their

[publica?gad_source=1&gad_campaignid=22714986660&gclid=Cj0KCQjw8p7GBhCjARIsAEhghZ1p5-7ST3sXHWIu7FDRwnwHDhWJkm583icbOYkuLEL6iw90QovEWIaAgcPEALw_wcB](https://www.mdpi.com/2271-4986/6/04/0044/publica?gad_source=1&gad_campaignid=22714986660&gclid=Cj0KCQjw8p7GBhCjARIsAEhghZ1p5-7ST3sXHWIu7FDRwnwHDhWJkm583icbOYkuLEL6iw90QovEWIaAgcPEALw_wcB) (accessed 15 September 2025).

⁵ CAIZA, Gustavo *et al.* 'Navigating Governmental Choices: A Comprehensive Review of Artificial Intelligence's Impact on Decision-Making', **Informatics**, 11, 64, MDPI, 4 September 2024. DOI 10.3390/informatics11030064.

⁶ BABŠEK, Matej *et al.* 'Artificial Intelligence Adoption in Public Administration: An Overview of Top-Cited Articles and Practical Applications', *AI Journal*, 6, 44, MDPI, 21 February 2025. DOI: 10.3390/ai6030044.

conclusions. This idea is known as the ‘illusion of algorithmic neutrality’⁷ and has been extensively analysed by the academic world, which has empirically demonstrated that algorithms can reproduce, amplify or even create new forms of discrimination, depending solely on how they are designed, trained and applied. At this point, it is necessary to clarify that AI systems are developed by private entities with their own economic interests and institutional policies, which are not necessarily aligned with those of the government or the society using these systems.

One must also consider the possibility of operating in an environment already skewed by social inequalities, ethnic differences or political polarisation. A notable example is that of facial recognition systems, where it has been shown that models trained predominantly on white people have a higher error rate when identifying Black people — a finding that can be extrapolated to any other ethnic group. This error highlights a bias arising from the limitations or segmentation of the available data, but it conceptually illustrates how even a correctly executed analysis can make errors due to the source data⁸.

However, we must not assume that this ‘mirage’ arises solely from human errors—whether unintentional or negligent—when feeding the system. Let us not forget that Google was fined in 2017 and again in 2025 for favouring its own products in search engines, demonstrating that biases can be introduced into algorithms to favour their creator and/or owner. As noted in the introduction to this section, it is true that algorithms have no interests of their own, but they can have an external bias introduced by their owners or designers⁹.

The dangers of algocracy

Let us return to the initial hypothesis: ‘algocracy’, a form of government entirely governed by algorithms, whether through the explicit acceptance of citizens or through an implicit dependence on human decision-makers.

⁷ PHILLIPS-BROWN, Milo. *Algorithmic neutrality*. University of Oxford, v4, 22 July 2025. DOI: 10.48550/arXiv.2303.05103.

⁸ STINSON, Catherine. *Algorithms are not neutral: Bias in collaborative filtering*. AI Ethics 2, 763–770, 31 January 2022. DOI: 10.1007/s43681-022-00136-w.

⁹ XAVIER, Bibin. ‘Biases within AI: challenging the illusion of neutrality’, *AI & Society*, Volume 40, 13 June 2024, pp. 1545–1546. DOI: 10.1007/s00146-024-01985-1.

In this hypothetical form of government, the first issue we would encounter would be one of 'algorithmic opacity': a complete lack of understanding on the part of people of the intricate reasoning processes of the governing artificial intelligence. This problem could lead to a phenomenon of political ineptitude, providing those ultimately responsible for decisions — if such people were still to exist — with arguments to shirk their responsibility when the results were negative¹⁰.

As we have seen, it is possible to introduce specific biases into the decisions of these systems simply by providing incomplete or pre-selected statistical data, whether through indolence, incompetence or corruption. This could lead to decisions that run counter to what is considered fair, ethical or lawful, generating undesirable effects such as the automated discrimination against specific groups¹¹.

Furthermore, attention must be paid to the reliance on algorithms and AI in decision-making that is critical to the survival of the state. Consequently, these systems become priority targets for potential enemies. If human decision-makers lose the ability to make decisions autonomously, technological dependence would have become a strategic vulnerability of the utmost criticality¹².

Finally, we must not overlook the digital divide as a social divide. This phenomenon exists between those who have the capacity to understand, manage or control algorithms and AI, and those who would simply accept the decisions derived from their analyses, even if they do not understand them. This would run the risk of creating a new privileged class, based on technological skills, which would shift the social balance towards increasingly pronounced inequality¹³.

¹⁰ SZLAJEN, Fishel. 'Algorithmic opacity and bioethical implications', *InfoBAE*. 16 November 2024. Available at: <https://www.infobae.com/opinion/2024/11/16/opacidad-algoritmica-e-implicancias-bioeticas/> (accessed 15 September 2025).

¹¹ ARAYA PAZ, Carlos. 'Algorithmic transparency: a regulatory or technological problem?', *CUHSO* (Temuco). ISSN-e 2452-610X, vol. 31, no. 2. December 2021, pp. 306–334. DOI: 10.7770/cuhso-v31n2-art2196.

¹² ROCA RIVERO, Francisco Javier. 'The Age of Artificial Intelligence Is Already Here', *Revista Española de Defensa RED*, no. 430, vol. 38. September 2025. Available at: <https://publicaciones.defensa.gob.es/media/downloadable/files/links/r/e/red-430.pdf>.

¹³ SARDEGNA, Paula Costanza and PELIZA, Eleonora. 'Legal challenges regarding data governance and algorithmic opacity in employment', *International and Comparative Journal of Labour Relations and Employment Law*, vol. 12, no. 4. October–December 2024. ISSN 2282-2313.

Civic apathy: a new risk

‘Civic apathy’, which leads to an ‘algocracy’, is not a well-established idea but rather an inferred and nascent notion. These are simply two concepts that interrelate to form a single idea: a cause-and-effect relationship which, in the specific situation analysed in this article, provides a microcosm of a problem whose depth could fill entire encyclopaedic volumes.

However, based on the information we have already presented in this text, we can deduce that ‘algocracy’ arising from ‘civic apathy’ would be the consequence of the public’s indifference, lack of interest or demoralisation regarding public affairs. In this case, it would stem from the convenience of delegating management and governance decisions to algorithmic systems, which are supposed to be infallible but which, in practice, would be perceived as opaque, uncontrollable and unchallengeable¹⁴.

The causes could be very diverse, and cases would need to be studied individually should they arise. However, we can anticipate that educational and cultural factors would lead to a lack of understanding of how algorithms work, which would ultimately result in either blind acceptance or scepticism. Initially, the average citizen would come to delegate any aspect of governance to automated decision-making processes, thereby delegitimising traditional politics and replacing it with a new technocracy based on technical, impersonal and inhuman processes (in the sense of being artificial and logical). It would be such a convenient and unquestionable system that it would risk progressively eroding citizen participation¹⁵.

It is at this point that disinformation or information overload would provide fertile ground for magnifying the imbalances or controversies generated by the system, thereby legitimising any school of thought opposed to the prevailing regime (even if radical). Gradually, any citizen who perceived a decision by the ‘algocracy’ as unfair would be prone to feelings of disaffection towards the system, succumbing to the fatalism of attributing all their ills to technology and becoming demotivated with regard to civic

¹⁴ ALNEMR, Nardine. ‘Democratic self-government and the algocratic shortcut: the democratic harms in algorithmic governance of society’, *Contemporary Political Theory*, vol. 23. 2023, pp. 205–227. DOI: 10.1057/s41296-023-00656-y.

¹⁵ GRIMMELIKHUIJSEN, Stephan and MEIJER, Albert. ‘Legitimacy of Algorithmic Decision-Making: Six Threats and the Need for a Calibrated Institutional Response’, *Perspectives on Public Management and Governance*, Volume 5, Issue 3. 28 March 2022, pp. 232–242. DOI: 10.1093/ppmgov/gvac008.

participation in political life. This process would provide a dangerous breeding ground for the erosion of the social contract and the gradual delegitimisation of the system of government¹⁶.

Implications for security and defence

Quite apart from the hypothesis put forward, we can examine this field from countless perspectives. However, none entails the risks and moral dilemmas associated with delegating decisions relating to military and national security contexts¹⁷.

Strategic decisions are already being made on the basis of recommendations from algorithms or AI. Added to this is their growing presence as intrinsic components of weapons systems, guided munitions, tactical intelligence processing and cyber-defence systems¹⁸. However, we must be aware that any action in such scenarios carries a high probability of harm to people, whether they are legitimate targets or collateral damage. Consequently, the issue of accountability for erroneous actions takes on greater importance: if the decision-making process relies increasingly on highly complex systems, to the extent that it hinders the traceability of the decisions taken, how can gaps in accountability be avoided?^{19,20}

It may seem a trivial matter; indeed, it is tempting to suggest that ultimate responsibility will always lie with the human decision-maker, but it is not quite that simple. In a scenario where an erroneous, unfortunate, negligent or punishable decision is made, an auditor would have to determine not only who accepted a particular outcome of the analysis and acted accordingly, but also how the process of reaching that decision unfolded. The very nature of the analyses carried out — using *software*, algorithms or AI — means that, in order to ensure proper traceability of the processes, it is necessary to rely on analogous

¹⁶ VIEIRA MAGALHÃES, João. 'Algorithmic Resistance as political disengagement', *Media International Australia*, vol. 183. 1 May 2022, pp. 77–89. DOI: 10.1177/1329878X221086045.

¹⁷ OLIER, Eduardo and CORCHADI, Juan Manuel. *Artificial Intelligence: Applications to Defence*. Research paper 01/2022. Spanish Institute for Strategic Studies. Available at: https://www.ieeee.es/Galerias/fichero/docs_investig/2022/DIEEEINV01_2022_EDUOLI_Inteligencia.pdf (accessed 10 September 2025).

¹⁸ MEZA, Milton. *Fully Autonomous Weapon Systems: A Challenge for the International Community*. Opinion Paper 85/2016, IEEE, 18 August 2016. Available at: https://www.ieeee.es/Galerias/fichero/docs_opinion/2016/DIEEO85-2016_SistemasArmas_UNU_MiltonMeza.pdf (accessed 16 September 2025).

¹⁹ NATO Official Texts. *Summary of NATO's revised Artificial Intelligence (AI) strategy*. Washington DC, 10 July 2024. Available at: https://www.nato.int/cps/en/natohq/official_texts_227237.htm (accessed 16 September 2025).

²⁰ SPANISH INSTITUTE FOR STRATEGIC STUDIES. 'Artificial intelligence in geopolitics and conflicts', *Strategy Paper* 226. IEEE, 2024. Available at: https://www.ieeee.es/Galerias/fichero/cuadernos/CE_226/CE_226.pdf (accessed 20 September 2025).

systems, as these are designed to perform calculations in a matter of seconds that would be impossible for human beings.

In cases where a decision is based on incorrect input data or premises, this traceability becomes critical. It is then imperative to determine whether the damage is a consequence of that data or those premises, of the algorithms' design itself, or whether it is due to an unintended external bias. It could even be intentional, given that technological dependence makes these tools prime targets for cyber-attacks. After all, strategic dominance in cyberspace is characterised by a succession of operations aimed at limiting, controlling or neutralising the enemy's data infrastructure, and that includes decision-making algorithms²¹.

Implementing systems that we might define as 'algocratic' in security and defence would entail a gradual and significant loss of human control in favour of a substantial improvement in efficiency, immediacy and results. This raises significant ethical and strategic dilemmas, eroding the classical doctrine of command and control and replacing it with a series of automated logical systems that are not constrained by moral judgements.

As we have seen, delegating such critical activities as surveillance, cyber defence or target identification to automated systems entails specific risks. The automation of decision-making in these areas is particularly complex due to the impact it may have. When the life, liberty or rights of any person – even a declared enemy – are at stake because of the logic of an algorithm, it is imperative to establish effective mechanisms of human control, ensuring that the 'mirage of neutrality' does not lead us to regard them as superfluous or acceptable²².

Responsible algorithmic governance

It must be emphasised that a system of governance based wholly or partly on 'algocracy' remains purely hypothetical at present; therefore, the proposals for mitigating it set out below should be interpreted as theoretical. Nevertheless, the intention of this article is for

²¹ NATO Standardisation Office. *Allied Joint Doctrine for Cyberspace Operations*. NATO STANDARD AJP-3.20, Edition A, Version 1, 29 January 2020.

²² U.S. Department of Defense. *DoD Directive 3000.09: Autonomy in Weapon Systems*. Washington, D.C., 25 January 2023 (first edition 21 November 2012).

them to serve as examples of how to cushion the side effects of delegating too many decisions to automated logical systems – a reality that is already upon us.

It is also necessary to clarify that the guiding principles by which any general decision-making process should be managed are transparency and accountability. This also applies to the use of algorithms or AI for this purpose²³.

To this end, we can apply the same decision-making control mechanisms that should be applied to any form of government: auditing and public participation. The first would consist of systematic processes aimed at evaluating the functioning of decision-making algorithms, verifying their fairness, transparency, traceability and absence of bias. As indicated in the previous section, this includes an additional problem arising from the volume of data and complex processes they handle: they can only be analysed by another algorithm or artificial intelligence. This would lead us into a vicious circle regarding who audits the auditing algorithm. The other method would be that of public participation, which could consist of panels or working groups of non-experts who verify whether the decisions taken comply with the ethical and moral criteria of the society being governed. The problem, in this case, is deciding who should make up these working groups: should they be professionals dedicated to this task, or volunteer citizens? Both models present difficulties when it comes to ensuring genuine independence or a sincere commitment to the intended objective. These mechanisms would be analogous to those of lay juries or polling station committees²⁴.

On the other hand, the speed at which these systems are being integrated into all aspects of daily life is overwhelming. However, if the integration process were gradual enough to be accompanied by policies on digital education and algorithmic literacy, we could avoid falling into 'civic apathy'. If a significant proportion of the population had a basic understanding of how algorithms, neural networks and AI work, this would enable a more responsible use of these technologies, preventing a social divide that could give rise to a new form of discrimination. This would ensure that those who 'suffer' the

²³ LAMAS LÓPEZ, Francisco and PERALTA GUITERREZ, Alfonso. 'Notes on the international and EU legal context for the military use of artificial intelligence'. *Defence Technology Observation Bulletin (BOTD)*, No. 78, 3rd quarter of 2023. Ministry of Defence. Available at:

https://publicaciones.defensa.gob.es/media/downloadable/files/links/b/o/bot_78.pdf

²⁴ KOENE, Andreas; CLIFTON, Chris; HATADA, Yuko *et al.* *A governance framework for algorithmic accountability and transparency*. European Parliamentary Research Service (STOA), 2019. Available at:

https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624262/EPRS_STU%282019%29624262%28ANN1%29_EN.pdf (accessed 17 September 2025).

consequences of decisions derived from algorithmic analysis can interpret them critically and constructively, thereby reducing the cognitive asymmetry between those in authority and those under their authority²⁵.

Finally, we must not overlook the importance of determining the chain of responsibility. A legal and regulatory framework must be put in place to facilitate the identification of roles, traceability, 'explainability' and fundamental rights in decisions made by automated systems. To this end, the concept of 'meaningful human oversight' – the human decision-maker acting as a supervisor with veto power – would ensure that there is always someone responsible for managing decisions, thereby preventing the dilution of accountability when undesirable outcomes are punishable^{26,27}.

A new reality

Whilst this article was being written, a news story emerged suggesting that the paradigm outlined here is closer to becoming a reality than it might seem. The report states that the Albanian Government has implemented an AI called Diella ('Sun' in Albanian) as a minister dedicated to the fight against corruption, responsible for all public procurement within the Albanian public administration²⁸.

Although this is a symbolic act, it highlights the issue of algorithms' growing decision-making power, entrusting them with tasks as important as overseeing a country's expenditure²⁹. It is not the aim of this article to assess the democratic quality of this political experiment, nor the interests of the companies behind its development (under their own government contract), nor whether members of the government have reserved

²⁵ GAGRČIN, Elma; NAAB, Teresa K. and GRUB, Marvin F. *Algorithmic media use and algorithm literacy: An integrative literature review*. New Media & Society, 2024. Available at: <https://madoc.bib.uni-mannheim.de/68149/1/gagrcin-et-al-2024-algorithmic-media-use-and-algorithm-literacy-an-integrative-literature-review.pdf> (accessed 17 September 2025).

²⁶ CHEONG, Boon Chong. 'Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making', *Frontiers in Human Dynamics*, vol. 6, 2024. DOI: 10.3389/fhumd.2024.1421273.

²⁷ CAVALCANTE, Paulo. 'Artificial Intelligence and Algorithm Accountability: Potential Risks and Proposed Governance Measures', *Blavatnik School of Government Working Paper*. University of Oxford, 2024. Available at: https://www.bsg.ox.ac.uk/sites/default/files/2024-09/BSG-WP-2024-60_EN_WorkingPaper_ArtificialIntelligence.pdf (accessed 17 September 2025).

²⁸ DAVIES, Pascale. 'Albania appoints the world's first AI-generated "minister" to root out corruption', *Euronews online*, Business & Technology section. 13 September 2025. Available at: <https://es.euronews.com/next/2025/09/13/albania-nombra-a-la-primera-ministra-de-inteligencia-artificial-del-mundo-para-erradicar-l> (accessed 18 September 2025).

²⁹ ELECONOMISTA.ES. 'Albania appoints an artificial intelligence system as anti-corruption minister: Meet Diella', *Eleconomista.es*, Technology section. 16 September 2025. Available at: <https://www.eleconomista.es/tecnologia/noticias/13548088/09/25/albania-nombra-a-una-inteligencia-artificial-ministra-anticorrupcion-asi-es-diella.html> (accessed 18 September 2025).

any kind of loophole to introduce bias. What does need to be analysed in a country that does not stand out for its transparency — ranked 80th out of 180 in terms of perceived corruption in 2024³⁰ — is whether there is a hierarchy of responsibility for any errors Diella might make or whether, on the contrary, policy-makers will simply take refuge in the ‘mirage of algorithmic neutrality’, pointing to its supposed infallibility. In time, we shall see the results, how this idea evolves and the new problems that will need to be addressed.

Conclusions

This text has presented the delegation of decision-making to algorithms as a hypothetical yet plausible scenario; however, significant steps are already being taken that bring us a little closer to this situation every day. This raises a series of ethical, social and strategic risks that it would be negligent to ignore, especially given the breakneck speed at which this technological leap is taking place.

Lack of transparency, loss of traceability, biases (whether unintentional or not), technological elites and digital illiteracy can create a series of legal loopholes that allow responsibilities to be diluted in the face of negative consequences. This would lead to a gradual deterioration of the system, triggering the much-feared process of ‘civic apathy’ whereby citizens lose interest in decision-making processes, thereby increasing mistrust and eroding the legitimacy of the established social contract. These problems are common to any form of government that ignores the risks it faces, and ‘algocracy’ appears to be no exception.

Developing efficient regulatory frameworks, establishing independent audit systems and educating the public to understand these new technologies would be the fundamental pillars for ensuring transparency. All of this must be underpinned by the principle of ‘meaningful human oversight’, so that ethical and moral criteria are always upheld and accountability is ensured. As we have seen, this is essential for decisions that may cause harm to individuals, as is the case in security and defence.

Documentation automation, infrastructure forecasting, automated customer service, complaint handling, facial recognition, candidate assessment...; ‘algocratic’ systems are

³⁰ TRANSPARENCY INTERNATIONAL. *Corruption Perceptions Index 2024* (CPI-2024). Available at: <https://transparencia.org.es/que-hacemos/investigacion/indices/indice-de-percepcion-de-la-corrupcion-2024/> (accessed 18 September 2025).

already being effectively implemented within public administrations. Algorithms, neural networks, Big Data and AI are concepts that are already part of administrative vocabulary and, increasingly, of governments' analytical and decision-making processes. It can therefore be said that a process of gradual transformation is already underway, giving rise to a 'partial algocracy' in which all our decisions will be based on information obtained through these technologies. The problem lies in determining where to draw the line between purely advisory or analytical roles and where to allow algorithms to make complex decisions.

At both extremes, human control must be non-negotiable in order to prevent these systems from achieving complete autonomy in their decision-making ('total algocracy'), by establishing a hierarchy of decision-makers with the power to halt any automated action, whilst relying solely on the operator's subjective judgement.

However, before reaching that point, we must determine whether it is feasible to aim to govern ourselves with the ever-increasing support of automated decision-making systems, or whether this would entail sacrificing too much of our sovereignty or our fundamental rights. After all, let us remember that the challenge is no longer solely technological development, but how to assimilate this new technology without losing our essence as a society, our ethical perspective and human moral judgement.

*Alfredo Gil Laso**
Capitán CIP del ET, INTA