Human Dignity and Meaningful Control Under the EU AI Act

11 SEPTEMBER 2026

ELIF SU COMERT

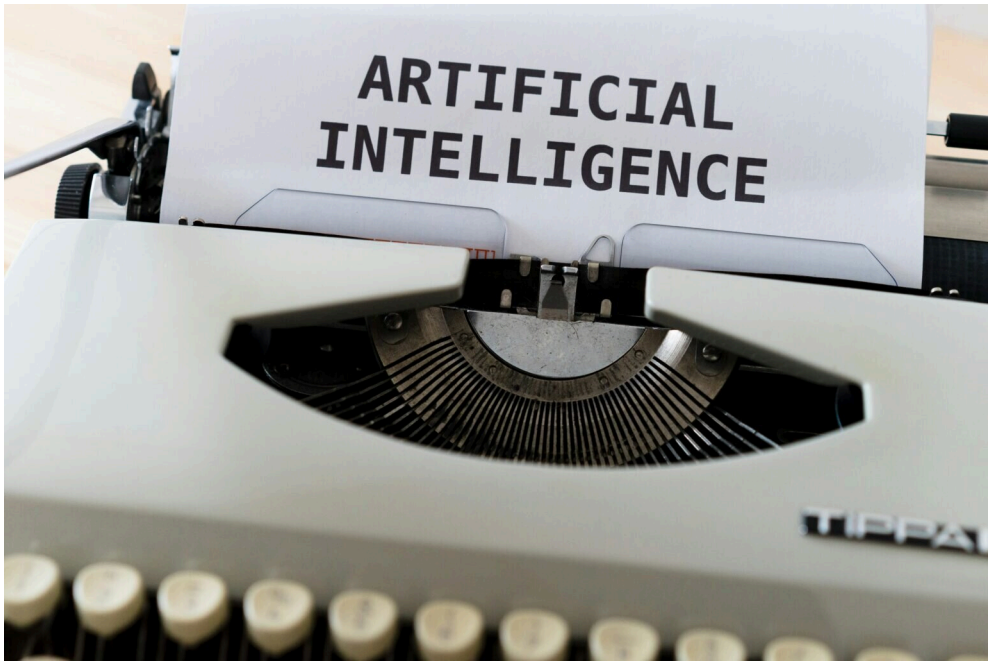

Photo by Markus Winkler on Unsplash

From Formal Oversight to Meaningful Human Control: Human Dignity under Article 14 of the EU AI Act

I. Introduction

Artificial intelligence (AI) is no longer confined to routine administrative or commercial tasks. AI systems increasingly influence decisions determining individuals' access to employment, education, healthcare, financial services, migration procedures, and law enforcement. Part of the decision-making process shifts from human judgement toward algorithmic processes capable of handling vast quantities of data at a speed and complexity no human reviewer can fully replicate. Within the EU's constitutional order, this shift engages more than technological or regulatory concerns. Article 1 of the Charter of Fundamental Rights of the European Union (CFR) establishes human dignity as inviolable and requires that it be respected and protected. Dignity is closely connected to personal autonomy and the recognition of individuals as active subjects rather than

passive recipients of decisions. Where algorithmic outputs predominate and human involvement becomes merely formal, the protection afforded by Article 1 CFR may be undermined.

The EU AI Act (Regulation (EU) 2024/1689) is the Union's central legislative instrument for regulating artificial intelligence. It adopts a risk-based approach: certain practices are prohibited outright, systems classified as "high-risk" face extensive obligations, and lower-risk systems face mainly transparency duties. Human oversight, data governance, and transparency obligations for high-risk systems form the core of this regime, and it is against this backdrop that the present article situates Article 14.

The EU legal framework governing artificial intelligence has generated a rapidly expanding body of legal scholarship, much of it approaching the AI Act primarily as an exercise in risk governance and internal-market harmonisation. Veale and Zuiderveen Borgesius (2021) offered an early systematic assessment of the draft Regulation, situating it within EU product-safety law and identifying provisions whose practical effectiveness they considered doubtful. Smuha (2021) cautioned that a human-rights-based framing of AI governance risks remaining aspirational without concrete mechanisms of enforceability and attention to societal, rather than purely individual, dimensions of algorithmic harm. Hildebrandt (2015) argued that meaningful legal protection against automated decision-making must be built into the architecture of AI systems through "Legal Protection by Design." Mantelero (2022) developed a framework for assessing AI's human-rights impact that moves beyond data-protection law, emphasising collective and social effects alongside individual harms.

Taken together, this scholarship has advanced understanding of the AI Act's regulatory architecture, enforcement gaps, and relationship to human rights. It has been comparatively less concerned with the specific constitutional relationship between human oversight and human dignity. Discussions of Article 14 have tended to treat human oversight as one governance safeguard among several, rather than as a provision whose legitimacy depends on a distinct constitutional value: the recognition of individuals as autonomous subjects rather than objects of algorithmic determination. The analysis therefore addresses this doctrinal gap.

Accordingly, this article asks: To what extent do the human oversight obligations applicable to high-risk AI systems under Article 14 of the AI Act effectively protect human dignity under Article 1 CFR?

The article does not attempt a comprehensive assessment of every fundamental-rights implication of artificial intelligence. Data protection, non-discrimination, and transparency are considered only insofar as they clarify the distinct function that human oversight performs in protecting human dignity. Nor does the article assess AI systems outside the high-risk category, or the criminal law and national security provisions of the AI Act, which raise distinct legitimacy questions beyond its scope.

It argues that Article 14 provides an essential legal basis for human control over high-risk AI systems, but that formal compliance with its oversight requirements does not by itself guarantee meaningful human judgement or, consequently, effective protection of human dignity under Article 1 CFR.

Methodologically, the article adopts a doctrinal and normative approach, examining the relevant provisions of the AI Act and the Charter alongside Court of Justice case law and academic literature on dignity, autonomy, human agency, and automation bias, while drawing on empirical public-administration research on human oversight. Because the research question turns on how effectively Article 14 protects dignity, the analysis also evaluates whether the provision can achieve its stated aim in practice.

The article proceeds in five further sections. Section II establishes human dignity, autonomy, and human agency as the normative basis against which the AI Act's safeguards are assessed. Section III examines how privacy, non-discrimination, and transparency obligations under the GDPR and the AI Act protect necessary but insufficient preconditions of a dignity-respecting decision. Section IV evaluates whether Article 14's human oversight requirements can ensure meaningful control in practice, given automation bias and organisational limits. Section V considers additional safeguards that might close the gap between formal and substantive oversight, before Section VI concludes.

II. Human Dignity as the Normative Anchor

Article 1 CFR is notably brief: human dignity is inviolable and must be respected and protected. Yet this brevity conceals considerable doctrinal weight. The official Explanations accompanying the Charter make clear that dignity is not one entitlement among the rights the instrument enumerates. It constitutes the substantive foundation from which the others derive their normative force, such that no Charter right may be invoked or construed in a manner that compromises the dignity of another. Its placement at the head of Title I, structurally distinct from values such as equality and solidarity, reflects what Dupré (2015) characterises as an organising principle of the Union's constitutional architecture: a meta-value that underwrites and circumscribes the exercise of every other right, rather than sitting alongside them as a peer.

This foundational character recommends dignity as the analytical vantage point for the present inquiry, in preference to any single substantive right taken in isolation. Privacy, non-discrimination, and transparency each possess an autonomous doctrinal architecture and capture different modalities of AI-related harm. What this fragmentation obscures is a structural feature common to all three: the displacement of the individual from the position of legal subject to that of computational object. The jurisprudence of the Court of Justice of the European Union (CJEU) situates dignity precisely at this fault line, tying it to autonomy and self-determi-

nation in a manner consonant with Halbertal's account of dignitary harm as instrumentalisation, the reduction of a person to a means, or to a status of helplessness before processes over which they exercise no meaningful will (as discussed in Aizenberg & van den Hoven, 2020). In Omega (Case C-36/02, 2004), the Court restricted an otherwise protected economic freedom on precisely this ground, holding that treating the human person as an object rather than a subject fell foul of dignity notwithstanding its commercial legitimacy, confirming that dignity in the EU's constitutional order is an operative threshold capable of constraining market and technological conduct alike, not a rhetorical flourish.

Understood in these terms, privacy, non-discrimination, transparency, and human oversight cease to appear as four discrete regulatory silos. They can instead be understood as four doctrinal expressions of a common underlying concern: whether the individual subjected to an AI-assisted decision retains meaningful agency in decisions affecting their legal and personal position, or is instead relegated to a data point processed according to a logic they can neither access nor contest. Privacy secures control over the informational substrate of the decision. Non-discrimination guards against being classified and judged according to criteria beyond the individual's control. Transparency preserves the capacity to understand and contest the decision once made. Human oversight, the safeguard on which this article ultimately turns, preserves the more basic claim that responsibility for the outcome remains attributable to a human decision-maker rather than to the system alone. It is this last guarantee, and its capacity to render the others meaningful rather than merely formal, that occupies the remainder of this article.

III. Fragmented Safeguards: Privacy, Non-Discrimination, and Transparency

The three safeguards examined in this section protect necessary, but individually insufficient, preconditions of a dignity-respecting decision. Table 1 summarises the function and structural limitation of each, alongside human oversight, before Sections III.A–C and Section IV examine them in turn.

Table 1. AI Act and GDPR safeguards and the preconditions they protect.

A. Data Protection as Informational Self-Determination

The GDPR supplies the first layer of protection against algorithmic objectification. It operates on the premise that dignity is compromised when an individual loses control over the informational substrate from which decisions about them are constructed. Principles such as purpose limitation and data minimisation, together with the obligation to build data protection into system architecture by design, preserve a residual sphere of informational self-determination even where processing is otherwise lawful (GDPR, arts. 5, 25). Article 10 of the AI Act supplements this regime

with AI-specific data-governance obligations. It requires that training, validation, and testing datasets be relevant, sufficiently representative, and, to the extent possible, free of errors, irrespective of whether the underlying data is personal in the GDPR sense (AI Act, Art. 10(1)–(3)).

The limitation of this layer is structural. Both regimes regulate the inputs into automated decision-making, not the exercise of judgement that follows once lawfully processed data has been transformed into an algorithmic output. A dataset may be impeccably minimised, purpose-limited, and GDPR-compliant, while the individual may nonetheless be treated as a mere object of computation at the point of decision.

B. Categorical Prohibitions and the Limits of Statistical Debiasing

The second layer operates on two distinct registers, one categorical, one statistical, both of which bear on dignity though neither is confined to non-discrimination in the narrow doctrinal sense. Article 5 of the AI Act prohibits outright certain practices: social scoring, the exploitation of vulnerabilities arising from age, disability, or socio-economic circumstance, and specified biometric applications (AI Act, Art. 5(1)). The Union legislature considered these practices incompatible with human dignity by their very nature, regardless of any procedural safeguard that might accompany them. These prohibitions are better understood as dignity-protective, in the sense identified in Section II, than as instances of anti-discrimination law proper. Several of the practices they capture, such as social scoring or manipulative exploitation of vulnerability, would cause dignitary harm even to a demographically homogeneous population.

Short of this categorical prohibition, Article 10(2)(f)–(g) imposes a narrower, genuinely statistical duty: to examine datasets for biases likely to affect health, safety, or fundamental rights, and to adopt appropriate mitigation measures (AI Act, Art. 10(2)(f)–(g)). This is a meaningful constraint, but a statistical one. It addresses the risk that a class of persons will be systematically disadvantaged, not whether the individual before the decision-maker in a particular case has been treated as a subject entitled to genuine consideration. Bias mitigation may equalise outcomes in the aggregate while leaving open whether any single decision reflects an exercise of judgement, as opposed to the mechanical application of a debiased but still unreviewed output.

C. Transparency as a Precondition, Not a Guarantee, of Understanding

The third layer, transparency, is addressed by Article 13, which requires that high-risk systems be accompanied by instructions enabling deployers to interpret the system's output and use it appropriately (AI Act, Art. 13(1)–(3)). This obligation is reinforced, from the perspective of the affected individual rather than the deployer, by Article 86. That provision grants a person subject to an adverse, legally or similarly significant decision the right to obtain from the deployer a clear and meaningful

explanation of the AI system's role and the main elements of the decision taken (AI Act, Art. 86(1)).

Both provisions are indispensable, yet both share the same structural limitation: they furnish the precondition for informed human judgement without guaranteeing that such judgement is actually exercised. Instructions for use do not compel an operator to read them attentively. A right to ex-post explanation does not prevent the decision from having been made, in substance, by the system rather than the deployer. Transparency addresses what a human being is capable of understanding, not what a human being, under the pressures of caseload, time, and institutional incentive, actually does with that capability. It is precisely this residual gap between the capacity for judgement that these safeguards help secure and the exercise of judgement itself that Article 14's human oversight requirement purports to close.

IV. Human Oversight as the Decisive Safeguard: Article 14 of the AI Act

If Articles 10, 5, and 13 address the informational, statistical, and cognitive preconditions of a dignity-respecting decision, Article 14 addresses the exercise of decision-making authority itself, requiring that high-risk AI systems be designed so that they can be effectively overseen by natural persons throughout the period they are in use (AI Act, Art. 14(1)–(2)). It is the only provision within the high-risk regime that speaks directly to the moment algorithmic output is converted into a decision affecting a person's rights.

Article 14(4) is unusually specific for a framework instrument. It requires that the measures built into the system enable the natural person to whom oversight is assigned to fully understand the system's capacities and limitations and correctly interpret its output; to remain aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system, a phenomenon the Regulation names explicitly as "automation bias"; and to decide not to use the system, disregard, override, or reverse its output, or intervene in or interrupt its operation through a "stop" function or equivalent mechanism (AI Act, Art. 14(4)(a)–(e)). These obligations are reinforced upstream by Article 16, which requires providers to ensure that high-risk systems undergo the conformity procedures necessary to satisfy Article 14 before being placed on the market, and downstream by Article 26, which obliges deployers to assign human oversight only to natural persons who have the necessary competence, training, authority, and support to exercise it (AI Act, arts. 16(1), 26(2)).

The significance of Article 14(4)(b) should not be understated. The Union legislature did not merely hope that human overseers would exercise independent judgement; it expressly anticipated that they might not, and required systems to be designed to counteract that tendency. This legislative foresight is corroborated by empirical literature on human–

AI interaction in decisional settings. Alon-Barkat and Busuioc (2023) studied public sector decision makers and found that officials exposed to algorithmic recommendations exhibited both automation bias, a general tendency to align with algorithmic advice, and "selective adherence": disproportionately following recommendations that confirmed pre-existing organisational or social stereotypes, while more readily overriding those that did not. This finding exposes a significant limitation in the architecture examined in Section III. An overseer operating within a system that is fully GDPR-compliant, statistically debiased under Article 10, and accompanied by exemplary Article 13 instructions may nonetheless default to algorithmic output because the formal conditions for independent judgement, while present, are psychologically and organisationally costly to exercise.

This is the central doctrinal issue of this article. Article 14 formally answers the question left open by privacy, non-discrimination, and transparency safeguards: who is ultimately responsible for converting algorithmic output into a decision affecting a legal subject? But the formal assignment of oversight responsibility to a natural person does not, without more, guarantee that the responsibility is substantively discharged. Three distinct failure modes recur across the literature and merit distinction.

The first is epistemic. Contemporary high-risk systems, particularly those relying on machine-learning techniques, may resist meaningful interpretation even by a trained overseer, such that the "correct interpretation" Article 14(4)(a) demands is, in practice, unattainable within the time available for the decision. The second is psychological. Automation bias and selective adherence persist even where the legal authority to override exists and is understood, because deviation from algorithmic output carries a cognitive and reputational cost: a wrong decision that departs from the system's recommendation is more readily attributed to the human overseer than one that follows it. The third is organisational. Article 26 requires that oversight personnel possess "necessary competence, training, authority and support," yet says comparatively little about the caseload, time pressure, or performance metrics under which that authority is exercised. This leaves deployers considerable latitude to design workflows in which meaningful override is theoretically available but practically discouraged.

The cumulative effect is that human oversight, as currently codified, can operate in either of two registers. It may function substantively, where the overseer genuinely retains and exercises the capacity to understand, question, and, where appropriate, depart from algorithmic output. On this register, the individual subject to the decision remains, in the terms established in Section II, a legal subject whose circumstances are determined through accountable human decision-making rather than being reduced to an object of automated assessment. Or oversight may function merely

procedurally, where the formal architecture of Article 14 is fully satisfied, a named natural person is assigned, a stop function exists, and instructions have been issued, while the actual decision-making dynamic reduces to ratification of the system's output. The AI Act's textual apparatus cannot, by itself, distinguish between these two registers. That determination depends on facts external to Article 14: organisational design, incentive structures, and the psychological dynamics the Regulation itself acknowledges in Article 14(4)(b) but does not neutralise. It is this indeterminacy that Section V addresses.

V. Toward Meaningful AI Governance

The preceding analysis suggests that the AI Act, read together with the GDPR, constructs a legal architecture that is formally comprehensive but substantively contingent. It is comprehensive in addressing the informational, statistical, cognitive, and decisional dimensions of algorithmic harm through Articles 10, 5, 13, and 14, respectively. It is contingent in that its dignity-protective function depends on implementation choices the Regulation does not fully determine. The Fundamental Rights Impact Assessment introduced by Article 27, a structured ex-ante process requiring certain deployers, currently confined to bodies governed by public law and private entities providing certain public services, to assess the risks a specified high-risk system poses to individuals' rights and identify oversight measures, represents a partial acknowledgment of this contingency (AI Act, Art. 27(1)). Its scope, however, is narrower than the problem identified in Section IV: it addresses oversight arrangements ex ante, but contains no equivalent mechanism for verifying ex post that oversight is exercised substantively rather than procedurally once the system is operational.

A sceptic might object that this proposal asks too much of a single provision. No oversight mechanism can fully eliminate the epistemic and psychological pressures identified in Section IV. Demanding evidentiary proof of substantive judgement in every high-risk decision risks reintroducing the administrative burden the AI Act's risk-based approach was designed to avoid. This objection has force, but it misstates the claim. The argument is not that Article 14 must guarantee substantive oversight in every instance, an unattainable standard. It is that the Regulation provides no mechanism for distinguishing, even approximately, systems and deployers where oversight functions substantively from those where it has collapsed into ratification. A proportionate audit obligation, confined to the sensitive-domain subset of high-risk systems already subject to the Article 27 impact assessment, would not eliminate automation bias. It would make its prevalence visible to regulators, courts, and affected individuals, creating a precondition for corrective intervention.

Three refinements would narrow this gap. First, conformity assessment and post-market monitoring obligations could be extended, for a defined

subset of high-risk systems operating in sensitive domains, to include an audit of oversight practice rather than system design alone: examining override rates, recorded reasons for departures from algorithmic output, and whether such departures correlate with caseload or time pressure, thereby distinguishing substantive from nominal oversight. Second, Article 26's requirement that deployers assign oversight to persons with "necessary competence, training, authority and support" would benefit from harmonised, sector-specific minimum standards, analogous to those developed for data protection officers under the GDPR, rather than being left to deployers' discretion. Laux (2023) reaches a comparable conclusion from a different angle, proposing that institutional design draw on democratic-theory principles of "institutionalised distrust" which includes periodic review of overseers' mandates and collective decision-making for higher-stakes determinations precisely because the AI Act leaves the competence and authority of Article 26 overseers largely undefined. Third, the individual right to explanation under Article 86, while a valuable ex-post safeguard, could be strengthened by a clearer evidentiary link to Article 14. A deployer's explanation should specifically address whether and how human judgement was exercised in the individual's case, rather than describing the system's general logic in the abstract. None of these refinements would displace the AI Act's risk-based architecture. Each would sharpen its capacity to distinguish substantive human oversight, the condition on which, this article has argued, the protection of human dignity under Article 1 CFR ultimately depends, from its merely formal appearance.

This analysis is doctrinal and normative, and it is worth being explicit about what that approach cannot show. It cannot establish empirically how human oversight operates across different sectors, institutions, or professional cultures, nor quantify how often formal compliance collapses into ratification in practice. The effectiveness of Article 14 is likely to vary with the type of AI system, institutional setting, and professional context in which oversight is exercised. The claim advanced here is accordingly conditional: Article 14, as currently drafted, is insufficient to guarantee, though it is not incapable of supporting, the substantive human judgement that Article 1 CFR requires.

VI. Conclusion

This article has argued that the human oversight obligations established by Article 14 of the AI Act constitute the decisive, though not self-executing, mechanism through which the human dignity guaranteed by Article 1 CFR is given practical effect in AI-assisted decision-making. Privacy, non-discrimination, and transparency safeguards under the GDPR and Articles 5, 10, and 13 of the AI Act each protect a necessary precondition of a dignity-respecting decision: control over informational inputs, protection against categorical and statistical harm, and the capacity to understand

the system's operation. None of them, individually or cumulatively, guarantees that a human being actually exercises judgement at the point of decision. That guarantee is the specific function Article 14 is designed to perform, reinforced by Articles 16 and 26 and supplemented ex post by Article 86.

Yet the very language of Article 14(4)(b), which names automation bias as a risk the system must be designed to counteract, reveals the limits of the provision's capacity to secure substantive human judgement through formal oversight architecture alone. Empirical findings on selective adherence in comparable decisional settings bear this out (Alon-Barkat & Busuioc, 2023). The central claim advanced here is accordingly that the AI Act's protection of human dignity is not self-executing. It depends on regulatory attention shifting from the existence of oversight mechanisms to the organisational, evidentiary, and psychological conditions under which those mechanisms are exercised. Absent that shift, human oversight risks becoming, in practice, a formal condition satisfied on paper while the underlying decision is made, in substance, by the AI system.

What remains, accordingly, is an empirical question this article's doctrinal method cannot resolve: whether the formal powers Article 14 vests in human overseers are, across the diversity of high-risk sectors, actually exercised as genuine intervention rather than merely held in reserve.

Legal Sources

Charter of Fundamental Rights of the European Union, 2012 O.J. (C 326) 391, Art. 1.

Explanations Relating to the Charter of Fundamental Rights, 2007 O.J. (C 303) 17.

Case C-36/02, Omega Spielhallen- und Automatenaufstellungs-GmbH v. Oberbürgermeisterin der Bundesstadt Bonn, 2004 E.C.R. I-9609.

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation), 2016 O.J. (L 119) 1, arts. 5, 25.

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 (Artificial Intelligence Act), 2024 O.J. (L 1689), Recital 27, arts. 5(1), 10(1)–(3), 10(2)(f)–(g), 13(1)–(3), 14(1)–(2), 14(4)(a)–(e), 16(1), 26(2), 27(1), 86(1).

References

Aizenberg, E., & van den Hoven, J. (2020). Designing for human rights in AI. Big Data & Society, 7(2). https://doi.org/10.1177/2053951720949566

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

Dupré, C. (2015). The age of dignity: Human rights and constitutionalism in Europe. Hart Publishing.
Hildebrandt, M. (2015). Smart technologies and the end(s) of law: Novel entanglements of law and technology. Edward Elgar Publishing.
Mantelero, A. (2022). Beyond data: Human rights, ethical and social impact assessment in AI (Information Technology and Law Series, Vol. 36). T.M.C. Asser Press/Springer. https://doi.org/10.1007/978-94-6265-531-7
Smuha, N. A. (2021). Beyond a human rights-based approach to AI governance: Promise, pitfalls, plea. Philosophy & Technology, 34, 91–104. https://doi.org/10.1007/s13347-020-00403-w
Veale, M., & Zuiderveen Borgesius, F. (2021). Demystifying the draft EU Artificial Intelligence Act: Analysing the good, the bad, and the unclear elements of the proposed approach. Computer Law Review International, 22(4), 97–112. https://doi.org/10.9785/cri-2021-220402
Laux, J. (2023). Institutionalised distrust and human oversight of artificial intelligence: Towards a democratic design of AI governance under the European Union AI Act. AI & Society, 39(6), 2853–2866. https://doi.org/10.1007/s00146-023-01777-z-