From Categorical to Contextual: Interpreting High-Risk Classification for Profiling-Based AI Recommender Systems in the EU AI Act

Luca Nannini[1, 2]

[1]Centro Singular de Investigación en Tecnoloxías Intelixentes da USC, Santiago de Compostela, Spain
[2]Piccadilly Labs, Barcelona, Spain
luca@piccadillylabs.co

Abstract

The EU's AI Act creates regulatory tension for recommender systems through its interaction with the GDPR's profiling provisions. This Article identifies a doctrinal challenge: Article 6(2)'s high-risk classification regime, when combined with GDPR Article 4(4), establishes a binary framework potentially inadequate for varied forms of algorithmic influence. Through doctrinal analysis, a "profiling paradox" is identified, manifesting in three dimensions: (1) conflation of data processing activities with behavioral influence mechanisms, (2) categorical treatment that undermines proportionate risk assessment, and (3) regulatory incentives that favor opaque non-profiling systems over transparent personalization. In comparison with the Digital Services Act's graduated approach, the AI Act's categorical treatment appears to overlook opportunities for more nuanced regulation. This Article proposes a "graduated influence" doctrine for interpreting Article 6(3)'s "significant risk" threshold through a multifactor test examining decisional weight, outcome reversibility, transparency, and user agency. This framework seeks to address regulatory challenges while aligning with fundamental rights objectives.

1 Introduction: The Regulatory Challenge

The European Union's Artificial Intelligence Act (European Parliament and Council 2024), which entered into force on August 1, 2024 and will be fully applicable by 2 August 2026 (European Parliament and Council 2024), represents the world's first comprehensive legal framework governing artificial intelligence. Yet beneath its risk-based taxonomy lies a potential doctrinal tension particularly apparent in the regulation of recommender systems.[1]—algorithmic frameworks that prioritize content or make personalized suggestions based on data analysis and pattern recognition (Zhang et al. 2019). These systems have become ubiquitous across digital platforms, from educational tools to employment services, influencing everything from the news articles users see to the job opportunities they discover (Ricci, Rokach, and Shapira 2022).

Recent scholarship has identified fundamental challenges in the AI Act's risk-based approach. Novelli et al. observe that while the Act "defines four risk categories for AI systems," it "lacks a clear methodology for the assessment of these risks in concrete situations" (Novelli et al. 2024). This methodological gap intersects with what the European Parliamentary Research Service identifies as "open questions" regarding the AI Act–GDPR "interplay" (European Parliamentary Research Service 2025). Ebers et al. further demonstrate that there remains "significant scope for applying a genuinely risk-based approach through careful interpretation and regulatory guidance" (Ebers et al. 2023), suggesting that courts need not await legislative amendment to address these doctrinal tensions. While these analyses recognize the need for more nuanced risk assessment methodologies, the specific regulatory tensions created by the Act's incorporation of the GDPR's profiling definition have not been examined in detail. This Article addresses that gap by identifying a "profiling paradox" and proposing concrete interpretive tools for judicial resolution.

AI recommender systems can be defined (Milano, Taddeo, and Floridi 2020; Zhu et al. 2023) as algorithmic systems that process user data to predict preferences and provide personalized suggestions, typically employing machine learning techniques to analyze behavioral patterns and contextual information. The widespread deployment of these systems has transformed digital interactions (Fleder and Hosanagar 2009), and their pervasive influence has prompted regulatory attention worldwide, with the EU positioning itself at the forefront through its comprehensive approach to AI governance (Veale and Borgesius 2021).

The AI Act's treatment of recommender systems carries global implications through the potential "Brussels Effect"—the EU's demonstrated ability to establish worldwide regulatory standards through market mechanisms (Bradford 2020). Recent analysis suggests that while the AI Act may achieve some extraterritorial influence, particularly for high-risk systems, its impact may be more limited than that of previous EU regulations such as the GDPR (Siegmann and Anderljung 2022; Pagallo 2023). Nevertheless, the Act's

[1]While this Article focuses on recommender systems, the analysis extends to search engines and other information retrieval systems that employ profiling. Though these systems differ in whether user queries are explicit (search) or implicit (recommendations), both raise similar profiling concerns under the AI Act. See Belkin and Croft (1992) on the convergence of information filtering and retrieval; Hanani, Shapira, and Shoval (2001) on distinctions between search and recommendation paradigms.

treatment of recommender systems could establish influential precedents for how democratic societies balance personalization benefits against manipulation risks (Helberger and Diakopoulos 2023). Therefore the regulation of recommender systems under the AI Act therefore warrants examination for several interconnected reasons. First, these systems increasingly mediate access to fundamental opportunities in education, employment, and essential services, making their regulatory treatment a matter of considerable social importance (Barocas and Selbst 2016; Eubanks 2018). Second, the interaction between the AI Act and existing regulations such as the GDPR creates novel compliance challenges that merit scholarly analysis, particularly given the limited judicial precedent in this emerging area (Kaminski 2023). Third, the AI Act's risk-based approach reflects broader trends in European digital regulation, yet implementing genuinely proportionate risk-based frameworks for algorithmic systems presents significant conceptual and practical challenges (Yeung and Lodge 2022). Recent judicial and regulatory developments, including the CJEU's SCHUFA decision (Court of Justice of the European Union 2023), and the European Commission's work on the Code of practice for general-purpose AI models (European Commission 2024) highlights the urgency of resolving AI Act–GDPR interpretive tensions and the larger regulatory landscape evolution.

This article investigates: *How might courts interpret the AI Act's risk classification provisions to address the regulatory challenges created by the mechanical incorporation of the GDPR's profiling definition, while remaining faithful to the Act's text and fundamental rights objectives?* This inquiry contributes to several strands of legal scholarship. It advances the literature on AI governance and the challenges of regulating emerging technologies (Selbst and Powles 2021), contributes to data protection scholarship examining how GDPR concepts translate to new contexts (Bygrave 2020), and engages with constitutional law scholarship on proportionality and fundamental rights in the digital sphere (De Gregorio and Dunn 2022).

Doctrinal analysis shows that the AI Act's incorporation of GDPR Article 4(4) creates a "profiling paradox" in which beneficial personalization in sensitive domains may trigger identical regulatory burdens to those imposed on potentially harmful manipulation. A "graduated influence" doctrine is proposed for interpreting Article 6(3)'s "significant risk" threshold, potentially introducing proportionality through a four-factor test that examines the nature and extent of algorithmic influence while remaining within the bounds of judicial interpretation.

The analysis proceeds as follows. Section 2 examines how Article 6's incorporation of GDPR Article 4(4) creates interpretive tensions. Section 3 articulates the resulting profiling paradox across technical, normative, and economic dimensions. Section 4 analyzes temporal boundaries, substantiality thresholds, and definitional ambiguities requiring judicial resolution. Section 5 proposes the graduated influence doctrine grounded in relational autonomy, choice architecture, and capabilities theories. Section 6 explores implementation through judicial interpretation, technical standards, and sectoral coordination. Section 7 concludes.

2 The Doctrinal Architecture of Algorithmic Risk

The analysis of Article 6's doctrinal architecture requires examining how its provisions interact to create potential interpretive challenges. This section dissects the article's structure to identify where judicial interpretation may be necessary, particularly regarding the incorporation of GDPR concepts into a fundamentally different regulatory framework. Understanding these structural tensions is essential for appreciating why courts may need to develop nuanced interpretive approaches.

Analyzing Article 6: A Taxonomy of Interpretive Challenges

The architecture of Article 6 reveals potential tensions in the AI Act's approach to risk classification. Article 6(2) establishes that "AI systems falling within any of the areas listed in Annex III shall be considered high-risk AI systems," creating what appears to be a straightforward categorical rule (European Parliament and Council 2024). Yet this apparent simplicity may mask interpretive challenges that emerge when examining the provision's interaction with its derogation clause.

Article 6(3) introduces a derogation that initially suggests proportionality: "By derogation from paragraph 2, an AI system referred to in Annex III shall not be considered to be high-risk where it does not pose a significant risk of harm to the health, safety or fundamental rights of natural persons, including by not materially influencing the outcome of decision making" (European Parliament and Council 2024). This provision suggests an approach that would allow courts and regulators to distinguish between AI systems based on their actual risk profiles rather than their mere presence in sensitive domains. The language "significant risk" and "materially influencing" introduces substantiality thresholds that invite contextual analysis. The importation of Article 4(4) into the AI Act exemplifies what Black terms "regulatory drift"—when legal concepts migrate between different regulatory contexts without adequate adaptation (Black 2002).

The GDPR's profiling definition emerged from concerns about notice, consent, and data subject rights in contexts where personal data processing was the primary regulatory target. The AI Act, by contrast, focuses on risks to fundamental rights arising from AI system deployment, not merely data processing. This shift in regulatory objective suggests that mechanical application of the GDPR definition may not serve the AI Act's purposes. A credit scoring algorithm that determines loan eligibility and a music recommendation system that suggests playlists could both "profile" under Article 4(4), yet their risk profiles diverge substantially—the former gates access to essential financial services while the latter enhances entertainment choice.

The concept of proportionality becomes more complex in the provision's third paragraph, which states that "[t]he first subparagraph shall not apply to AI systems that perform profiling of natural persons within the meaning of Article 4, point (4), of Regulation (EU) 2016/679" (European Parliament and Council 2024). This cross-reference to the GDPR

creates what could be termed as a "regulatory renvoi"—a referral that imports not merely a definition but an entire doctrinal apparatus developed in a fundamentally different regulatory context. While the operative text in Article 6(3)'s third paragraph refers simply to "profiling of natural persons," Recital 53 clarifies this means profiling "within the meaning of Article 4, point (4) of Regulation (EU) 2016/679." Article 4(4) defines profiling as "any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person"—a definition crafted for data protection notice and consent requirements rather than risk assessment (European Parliament and Council 2016). Indeed, the incorporation of the GDPR's profiling definition may transform Article 6(3)'s proportionality mechanism into a binary classification: if a system profiles, it may not benefit from the derogation regardless of its actual risk profile.

Table 1 illustrates the structural complexity of Article 6's risk classification framework.

Article 6(3) contains four exceptions in its second subparagraph. Each exception attempts to carve out categories of AI systems that, despite operating in Annex III domains, should not face high-risk classification. The first exception, found in Article 6(3)(a), exempts AI systems "intended to perform a narrow procedural task" (European Parliament and Council 2024). The notion of "narrow procedural task" lacks definition in the Act, and its boundaries remain uncertain. Drawing on teleological interpretation, this might encompass purely administrative functions with limited substantive decisional impact—perhaps systems that merely organize or format information without evaluating or ranking it. Yet recommender systems, by their nature, typically go beyond mere procedural tasks. Even basic recommendation involves some form of evaluation, ranking, or selection that shapes the information environment presented to users. The adjective "narrow" further constrains the exception, suggesting that systems with broad applicability or multiple functions would fall outside its scope.

Article 6(3)(b)'s exception for systems that "improve the result of a previously completed human activity" introduces questions of temporal causation and human-machine interaction (European Parliament and Council 2024). The requirement that human activity be "previously completed" suggests a clear temporal sequence: first human action reaches completion, then the AI system improves upon that completed action. But in the context of recommender systems, human activities may rarely reach such definitive completion. Consider a legal research platform that suggests relevant cases based on a lawyer's search history. Each search query might be viewed as a "completed" activity, but legal research is typically an iterative process where each search builds upon previous ones. The system's recommendations shape subsequent searches, creating feedback loops that may blur the distinction between human preferences and AI.

The GDPR's Profiling Definition: Contextual Considerations in Algorithmic Regulation

The complexity of these interpretive challenges reflects broader concerns about AI Act implementation. As institutional analysis recognizes, the AI Act-GDPR intersection creates "open questions" that require resolution through careful legal interpretation (European Parliamentary Research Service 2025). The incorporation of GDPR concepts into a fundamentally different regulatory framework exemplifies what regulatory scholars term "technology neutrality" challenges—ensuring that legal frameworks adapt appropriately to technological realities rather than creating artificial distinctions (Becker 2024).

The incorporation of GDPR Article 4(4)'s profiling definition into the AI Act's risk classification framework raises questions about cross-regulatory interpretation. Article 4(4) defines profiling as "any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements" (European Parliament and Council 2016). This definition emerged from a data protection context where the primary concern was notice, consent, and individual control over personal data processing. The breadth of the definition—"any form" of automated processing that evaluates "certain personal aspects"—may have been appropriate for establishing the scope of data protection rights. The low threshold helped ensure that data subjects would have rights regarding automated evaluations of their characteristics, regardless of how that evaluation was used.

Yet incorporating this definition into the AI Act's risk-based framework creates analytical challenges that reveal fundamental tensions between data protection and risk regulation objectives. The AI Act purports to regulate based on risk to fundamental rights, not merely based on the existence of data processing. A music streaming service that analyzes listening history to recommend songs engages in profiling no less than a credit scoring algorithm that determines loan eligibility. Both systems process personal data to evaluate personal aspects (musical preferences versus creditworthiness), yet their risk profiles may differ substantially. The music recommender may enhance user autonomy by expanding musical horizons; the credit scorer gates access to essential financial services.

Figure 1 demonstrates the broad scope of GDPR Article 4(4)'s profiling definition. The open-ended nature of "personal aspects" and the low threshold for "automated processing" help explain why this definition, while appropriate for data protection purposes, may create regulatory tensions when imported into risk-based AI regulation.

The enumerated personal aspects in Article 4(4) employ the phrase "in particular," indicating that the list is illustrative rather than exhaustive (European Parliament and Council 2016). This open-ended quality, while appropriate for data protection's broad preventive approach, creates regulatory uncertainty when imported into the AI Act's risk-based framework—a tension that becomes especially apparent when examining how courts might apply this definition to recommender systems. Courts interpreting the AI Act must consider whether any evaluation of any personal characteristic triggers high-risk classification, or whether some

Provision	Classification Rule	Key Terms	Interpretive Considerations
Art. 6(1)	High-risk if safety component + third-party assessment	"Safety component" (Art. 3(14)); "Third-party conformity assessment"	Scope of Annex I products; Definition of safety criticality
Art. 6(2)	High-risk if listed in Annex III	Eight enumerated domains	Breadth of domain definitions; Sectoral overlaps
Art. 6(3)	Exception if no significant risk	"Significant risk"; "Materially influencing"	Substantiality threshold; Causation standards
Art. 6(3) ¶2	No exception if profiling	"Profiling" per GDPR Art. 4(4)	Categorical exclusion; No proportionality

Table 1: Doctrinal Structure of AI Act Article 6 Risk Classification Framework. This table illustrates how Article 6's provisions interact to create potential interpretive challenges, particularly regarding the categorical exclusion for profiling systems in paragraph 3(b).

implicit threshold of significance applies.

Recent scholarship on the relationship between data protection and algorithmic decision-making has highlighted the complexity of determining when profiling activities warrant regulatory intervention (Malgieri 2019; Bygrave 2020). Moreover, the GDPR's profiling definition predates the widespread deployment of recommender systems and reflects concerns primarily about discriminatory credit scoring, employment screening, and government surveillance rather than content curation and personalization. The definition's focus on "evaluation" may not adequately distinguish between different types and purposes of evaluation. A recommender system evaluating reading preferences to suggest books operates differently from one that evaluates the same preferences to determine insurance premiums, yet both fall within Article 4(4)'s scope. This breadth creates challenges when the definition is imported into different regulatory contexts with varying objectives (Gellert 2020; Black 2002).

The DSA's Alternative Approach: Graduated Transparency and Proportionate Regulation

The Digital Services Act (European Parliament and Council 2022), developed contemporaneously with the AI Act, embodies a different regulatory philosophy that offers instructive contrasts. Rather than imposing binary classification based on technical characteristics, the DSA employs graduated obligations that scale with platform size, reach, and systemic impact. This approach recognizes that algorithmic influence exists on a spectrum and that regulatory responses should be proportionate to actual rather than theoretical risks.

DSA Article 27 establishes baseline transparency requirements for all online platforms that use recommender systems, requiring them to "set out in their terms and conditions, in plain and intelligible language, the main parameters used in their recommender systems, as well as any options for the recipients of the service to modify or influence those main parameters" (European Parliament and Council 2022). This provision applies regardless of whether the system performs profiling, focusing instead on ensuring users understand how recommendations are generated and maintaining some degree of user control.

The sophistication of the DSA's approach becomes apparent in Article 38, which applies only to Very Large Online Platforms (VLOPs) and Very Large Online Search Engines (VLOSEs)—those with more than 45 million monthly active users in the EU. Article 38 requires these platforms to "provide at least one option for each of their recommender systems which is not based on profiling as defined in Article 4, point (4), of Regulation (EU) 2016/679" (European Parliament and Council 2022). This provision recognizes that profiling-based recommendations may pose risks when deployed at scale, but rather than prohibiting such systems or subjecting them to potentially disproportionate compliance requirements, it preserves user choice.

3 The Profiling Paradox: A Critical Analysis

The Paradox Defined and Illustrated

The profiling paradox emerges where the AI Act's categorical risk classification collides with the technical realities of recommender system design, creating three interconnected regulatory tensions. The paradox may manifest in three interconnected dimensions that reveal possible challenges in the regulatory architecture.

Modern recommender architectures—whether collaborative filtering, content-based, or hybrid approaches—require analyzing user characteristics to generate relevant suggestions (Aggarwal 2016). Collaborative filtering systems build user preference profiles by analyzing behavioral similarities across user populations. Content-based systems evaluate user interests against item characteristics. Matrix factorization techniques, widely used in production systems, explicitly model latent user factors that correspond to personal preferences and behavioral patterns. Each of these approaches necessarily "evaluates certain personal aspects relating to a natural person" as Article 4(4) requires.

The regulatory consequence is that sophisticated personalization—which may enhance user experience and expand opportunity access—becomes indistinguishable from potentially manipulative profiling under the current framework. A job recommendation system that successfully matches candidates with opportunities based on skills, interests, and career goals faces identical classification as one that reinforces

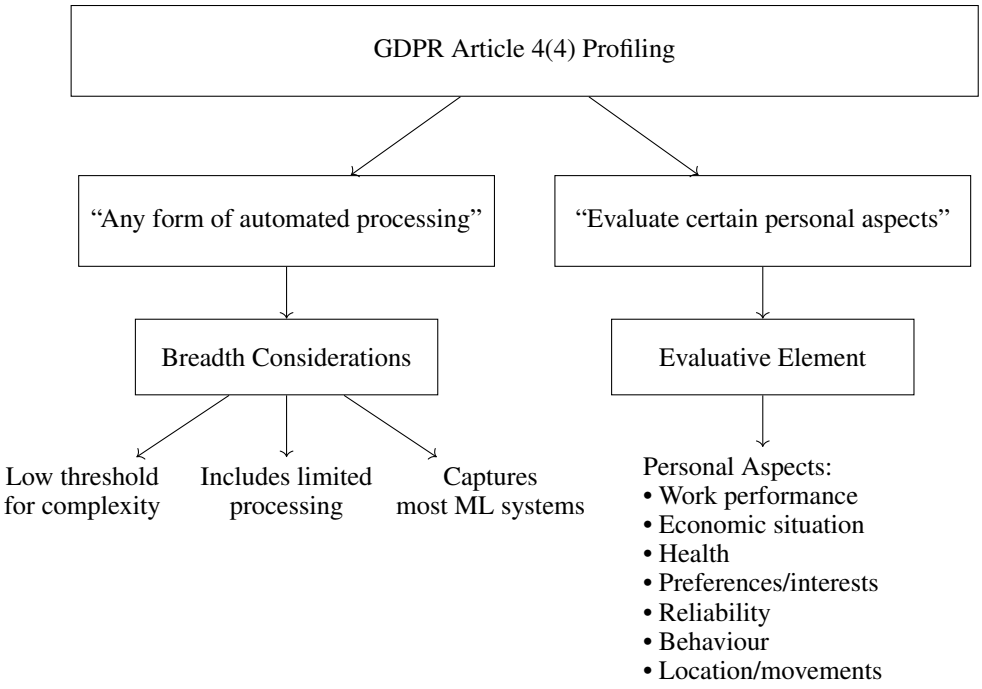

Figure 1: Structural Analysis of GDPR Article 4(4) Profiling Definition. The definition's broad scope and low thresholds, appropriate for data protection contexts, may create regulatory tensions when incorporated into the AI Act's risk-based framework.

discriminatory hiring patterns, despite their vastly different social impacts. A system stripped of profiling capabilities may devolve into a broadcasting mechanism that shows all users identical content, undermining the purpose of recommendation. Consider an educational platform operating in an Annex III domain: without profiling, it may not identify that a student struggles with mathematics and would benefit from additional resources, nor recognize that another student excels and needs advanced challenges. The regulation thus may force a choice between sophisticated personalization that serves individual needs and less targeted approaches that might not address diverse educational requirements.

The normative dimension reveals a potential contradiction. The AI Act's stated objective, articulated in Article 1 and elaborated in Recitals 1 through 13, is to ensure AI systems respect fundamental rights while favoring beneficial innovation (European Parliament and Council 2024). Yet applying high-risk classification to all profiling recommender systems in sensitive domains directly contradicts these objectives by treating beneficial personalization identically to manipulative targeting. A job recommendation platform that uses profiling to match candidates with opportunities based on skills and interests might help reduce employment barriers by expanding awareness beyond traditional networks. A non-profiling alternative that shows all users the same job listings might not achieve this goal as effectively.

The economic dimension of the paradox may create market effects that favor incumbent platforms while potentially discouraging innovation in socially beneficial applications. The compliance costs associated with high-risk classification—including conformity assessments, quality management systems, technical documentation, and ongoing monitoring—can be substantial. Large technology companies may absorb these costs more easily, while SMEs seeking to develop innovative recommender systems for education, employment, or essential services may face significant barriers to entry. The regulation thus may risk concentrating market power while potentially preventing the emergence of specialized alternatives.

Analytical Challenges in the Regulatory Scheme

The profiling paradox reveals potential analytical challenges in the EU's approach to algorithmic regulation, manifesting in three areas that may warrant scholarly attention.

1. *The relationship between data processing and behavioral influence*. The GDPR's profiling definition emerged from concerns about data protection—ensuring individuals knew when their data was being processed to evaluate their characteristics and maintaining their control over such processing. The definition intentionally employed a broad scope because its purpose was to establish data protection rights, not to assess risks to fundamental rights. When the AI Act incorporates this definition as a trigger for high-risk classification, it may conflate the fact of data processing with the potential for harmful influence. A book recommendation system that processes reading history to suggest similar titles engages in the same technical profiling as a system that processes reading history to infer political affiliations. The regulation may not adequately distinguish between these different use cases.

2. *Proportionality in what purports to be a risk-based framework.* The Charter of Fundamental Rights (European Union 2012), particularly Article 52(1), establishes that limitations on rights should be proportionate to their objectives. The Court of Justice has developed proportionality analysis that weighs the intensity of interference against the importance of the objective and considers whether less restrictive means could achieve the same goals. As de Búrca observes, this entails a three-part test examining whether measures are suitable, necessary, and proportionate in their effects on affected interests (de Búrca 1993). Yet the AI Act's treatment of profiling recommender systems may depart from proportionality. Once a system profiles in an Annex III domain, it may face identical compliance obligations regardless of whether it poses minimal or substantial risks to fundamental rights. This approach could create situations where beneficial systems face the same regulatory burden as potentially problematic ones.
3. *The treatment of systemic and individual risks*. The AI Act's preamble recognizes that some AI systems pose risks primarily through their aggregate effects on society—such as filter bubbles, polarization, or manipulation at scale—while others pose direct risks to individuals through specific decisions. Recommender systems typically fall into the former category, with their primary risks emerging from collective effects rather than individual harms. Yet the regulation's approach may treat all high-risk systems similarly, potentially imposing requirements designed for individual decision-making systems onto systems whose risks are primarily systemic.

These challenges reflect what Novelli et al. identify as a broader problem in AI risk assessment: the lack of clear methodology for the assessment of these risks in concrete situations (Novelli et al. 2024). The profiling paradox represents a specific manifestation of this general methodological deficit. Where existing scholarship recognizes the need for proportionality in AI regulation, the profiling context reveals how categorical approaches may work against the Act's fundamental rights objectives. The European Parliamentary Research Service's recognition of AI Act-GDPR "interplay" uncertainty (European Parliamentary Research Service 2025) validates our analysis that these cross-regulatory tensions require specific attention rather than mechanical rule application.

4 Interpretive Challenges and Judicial Considerations

The Temporality Question in Article 6(3)(b)

The exception in Article 6(3)(b) for systems that "improve the result of a previously completed human activity" raises questions about the temporal boundaries of human-machine interaction that courts may need to address (European Parliament and Council 2024). The provision's apparent clarity may mask uncertainties about when human activities can be considered "completed" in iterative digital environments.

Consider a legal research recommender system used by a law firm. When a lawyer searches for cases on corporate liability, the system analyzes the query and suggests relevant precedents. The lawyer reviews these suggestions, refines the search, and receives new recommendations. This process continues through multiple iterations until the lawyer has gathered sufficient authorities for a brief. Under Article 6(3)(b), we must determine which, if any, human activities in this sequence are "previously completed" such that the AI system merely "improves" their results.

One interpretation would treat each individual search query as a completed human activity, with the system's suggestions improving the results of that specific search. This reading might potentially bring many recommender systems within the exception. Yet this interpretation faces several challenges.

1. It may artificially segment what is actually a continuous, iterative process of legal research.
2. It may not account for how recommendations shape subsequent queries, creating feedback loops that blur the distinction between human activity and machine influence.
3. It may conflict with the provision's apparent intent to exclude systems that actively influence ongoing human decision-making.

An alternative interpretation would view the entire research process holistically, considering it "completed" only when the lawyer finishes researching and begins drafting. Under this reading, recommendations provided during the research process would not qualify for the exception because they influence an ongoing activity rather than improving a completed one. This interpretation better captures the reality of human-machine interaction but might exclude most recommender systems from the exception.

The Substantiality Threshold in Article 6(3)

Article 6(3)'s requirement that systems "not pose a significant risk of harm to the health, safety or fundamental rights of natural persons" introduces a substantiality threshold that may require judicial interpretation (European Parliament and Council 2024). The qualifier "significant" distinguishes between AI systems that pose theoretical or minimal risks and those warranting high-risk classification.

Recital 53 offers some interpretive guidance, noting that AI systems referred to in Annex III "may not necessarily always be considered to be high-risk" and that classification should depend on whether they pose "little to no impact on fundamental rights of natural persons" (European Parliament and Council 2024). The recital specifically mentions systems performing "narrow procedural tasks" or improving "the result of a previously completed human activity" as examples of potentially low-risk systems. However, this circular reference to the statutory exceptions provides limited additional clarity on the meaning of "significant risk."

Drawing on proportionality analysis from fundamental rights jurisprudence, courts might develop a multi-factor assessment for evaluating significance. The intensity of potential interference with fundamental rights would be a primary consideration—does the system merely influence preferences or does it determine access to essential services?

The probability of harm occurring would also matter—is adverse impact speculative or demonstrable? The scale of affected persons could factor into the analysis—does the system affect individuals occasionally or systematically shape outcomes for entire populations?

Table 2 summarizes the graduated influence framework's four-factor assessment, providing concrete criteria for evaluating the "significant risk" threshold in Article 6(3).

The Profiling Boundary Question

Perhaps the most fundamental interpretive challenge concerns the boundaries of "profiling" itself when imported from the GDPR context into the AI Act framework. While GDPR Article 4(4) provides a definition, its application to recommender systems may raise edge cases that courts must resolve.

Consider a news aggregation platform that allows users to select topic preferences (technology, sports, politics) and then shows articles from those categories in reverse chronological order. Does this constitute profiling? The system processes personal data (topic preferences) to deliver personalized content, which could be seen as evaluating "interests." Yet the processing is minimal—merely filtering content by explicitly stated preferences without inference or prediction. Courts must determine whether such minimal processing falls within Article 4(4)'s scope or whether profiling requires some degree of inference beyond explicit user choices.

The question becomes more complex with hybrid systems that combine explicit preferences with behavioral analysis. A job board that asks users to specify desired roles and locations, then refines recommendations based on which listings users click, straddles the line between preference-based filtering and behavioral profiling. The explicit preferences alone might not constitute profiling, but the click-based refinement likely does. Such scenarios create interpretive ambiguities about when Article 22 GDPR obligations are triggered, particularly in systems involving "multiple stages and/or effects with different levels of legal or similar significance" (Veale and Binns 2021). Must the entire system be classified as high-risk, or can components be assessed separately?

5 Toward a Graduated Influence Doctrine

Theoretical Foundations

The challenges identified in the current regulatory framework suggest the need for a new interpretive approach that can introduce proportionality while remaining within the bounds of judicial interpretation. We propose a "graduated influence" doctrine that courts could consider when interpreting Article 6(3)'s "significant risk" threshold. This doctrine draws on three established theoretical frameworks that provide principled bases for distinguishing between beneficial and potentially harmful algorithmic influence.

The first foundation comes from relational autonomy theory, developed by philosophers like Jennifer Nedelsky and Catriona MacKenzie (MacKenzie and Stoljar 2000; Nedelsky 2011). This theory rejects the notion of autonomy as complete independence, instead understanding it as constituted through relationships and social contexts. Recommender systems, from this perspective, do not inherently threaten autonomy merely by shaping choices. Rather, the question becomes whether they support or undermine the relational conditions necessary for autonomous action. A job recommender that expands awareness of opportunities and helps users understand their options may enhance relational autonomy. One that manipulates preferences or constrains choices undermines it. This theoretical lens allows courts to move beyond binary thinking about algorithmic influence.

The second foundation draws on Thaler and Sunstein's work on libertarian paternalism and choice architecture (Thaler and Sunstein 2008). Their insight that all choice environments are constructed, never neutral, applies directly to recommender systems. The regulatory question should not be whether influence occurs—it always does—but whether the influence preserves meaningful choice while improving decision-making contexts.

The third foundation comes from Amartya Sen's capabilities approach (Sen 1999), which evaluates systems based on whether they expand or constrain human capabilities—the genuine opportunities people have to lead lives they value. Recommender systems can enhance capabilities by surfacing previously unknown opportunities, reducing search costs, and enabling better-informed decisions. They can also constrain capabilities by creating filter bubbles, perpetuating biases, or limiting perceived options. This framework provides courts with a normative basis for distinguishing beneficial from harmful systems based on their effects on human flourishing rather than their technical characteristics. Recent scholarship on recommender systems has highlighted the importance of considering these broader social effects when evaluating algorithmic accountability frameworks (Milano, Taddeo, and Floridi 2020).

The Four-Factor Assessment

Building on these theoretical foundations, we propose that courts interpret Article 6(3)'s "significant risk" threshold through a four-factor analysis that examines the nature and extent of a recommender system's influence on human decision-making. This assessment would allow proportionate evaluation while providing sufficient structure to ensure consistency across jurisdictions.

Factor 1: Decisional Weight examines the role recommendations play in final decisions. Courts should analyze whether recommendations merely provide information that users can freely accept or reject, or whether they effectively determine outcomes. Key considerations include the availability of alternative information sources, the framing of recommendations, and whether the system creates artificial constraints. A university course recommender that presents options alongside comprehensive course catalogs exhibits lower decisional weight. An employment screening system that determines which applications recruiters see has higher decisional weight because it gates access to opportunities.

Factor 2: Reversibility of Outcomes assesses whether decisions influenced by recommendations can be easily changed or create lasting consequences. Systems influ-

Factor	Assessment Criteria	Risk Level Indicators	Example Applications
Decisional Weight	Role of recommendations in final outcomes; availability of alternatives; constraints created	Low: Advisory suggestions High: Gatekeeping function	Education platform (Low) Employment screening (High)
Reversibility	Ease of changing influenced decisions; temporal persistence of effects	Low: Easily changed High: Lasting consequences	Music recommender (Low) Credit scoring (High)
Transparency	Explainability of recommendation logic; user understanding of preference model	Low: Opaque algorithms High: Clear, accessible explanations	Black-box system (Low) Interpretable model (High)
User Agency	Control over preference formation; access to alternatives; override capabilities	Low: Constrained choice High: Meaningful alternatives	Forced recommendations (Low) Multiple feed options (High)

Table 2: Graduated Influence Assessment Framework

encing reversible decisions with low switching costs may pose less risk than those affecting choices with significant path dependencies. A music recommender suggesting songs poses minimal risk because users can easily skip suggestions and switching costs are negligible. An educational pathway recommender that influences course selection may have lasting effects, though many educational decisions remain modifiable. A loan approval recommender affecting access to credit creates potentially lasting impacts on financial futures.

Factor 3: Transparency of Preference Formation evaluates whether users can understand how recommendations are generated and whether the system's model of their preferences is accessible and correctable. Transparency encompasses both technical explainability and practical comprehensibility. Systems that clearly communicate their recommendation logic and allow users to view and modify their preference profiles may pose less risk than opaque systems with hidden factors. This factor recognizes that transparency enables user agency by allowing people to understand and potentially resist algorithmic influence.

Factor 4: User Agency Preservation examines whether the system maintains meaningful user choice and control. This includes assessing whether users can customize recommendation parameters, access non-personalized alternatives, and easily override or ignore suggestions. Systems that offer multiple configuration options and make it easy to access unfiltered information preserve agency. Those that make it difficult to escape recommendation filters or that use design patterns to enforce compliance undermine agency.

Courts applying this framework should consider several practical guidelines for factor integration.

- First, *threshold analysis*: systems scoring "high risk" on multiple factors warrant presumptive high-risk classification, while those showing "low risk" across most factors may qualify for Article 6(3)'s derogation.
- Second, *compensatory relationships*: strong performance on some factors may offset weaker performance on others—high transparency may compensate for moderate decisional weight, while strong user agency may mitigate lower reversibility.
- Third, *contextual weighting*: factor importance may vary by domain—decisional weight carries more significance in employment contexts than entertainment, while reversibility matters more for educational than informational applications.
- Fourth, *temporal considerations*: risk assessments should account for changing circumstances—systems with initially high transparency may become problematic if algorithmic updates reduce explainability without corresponding user notification.

Application to Paradigm Cases

Our approach builds upon existing scholarship while addressing identified gaps. Where Novelli et al. propose scenario-based proportionality testing for general AI risk assessment (Novelli et al. 2024), we develop profiling-specific factors that courts can apply within Article 6(3)'s existing framework. While Ebers et al. demonstrate that courts have significant scope for applying a genuinely risk-based approach through careful interpretation (Ebers et al. 2023), our four-factor test provides concrete judicial tools for resolving regulatory tensions in the specific context of recommender systems. This approach respects principles of technology neutrality (Becker 2024) by focusing on actual impacts rather than technical characteristics alone.

Applying this framework to concrete cases illustrates how it might operate in practice while addressing the profiling paradox's most challenging effects.

- **Case Example 1#**: Consider first an educational course recommender system operating in a university setting. Under current law, such a system may face automatic high-risk classification if it performs profiling, regardless of its actual risk profile. Applying the graduated influence assessment yields a more nuanced evaluation. The system exhibits moderate decisional weight—while it influences course selection, students retain ultimate enrollment authority and typically consult advisors, course catalogs, and peers. Reversibility remains high during add/drop periods and through subsequent semester adjustments. If the system clearly explains its recommen-

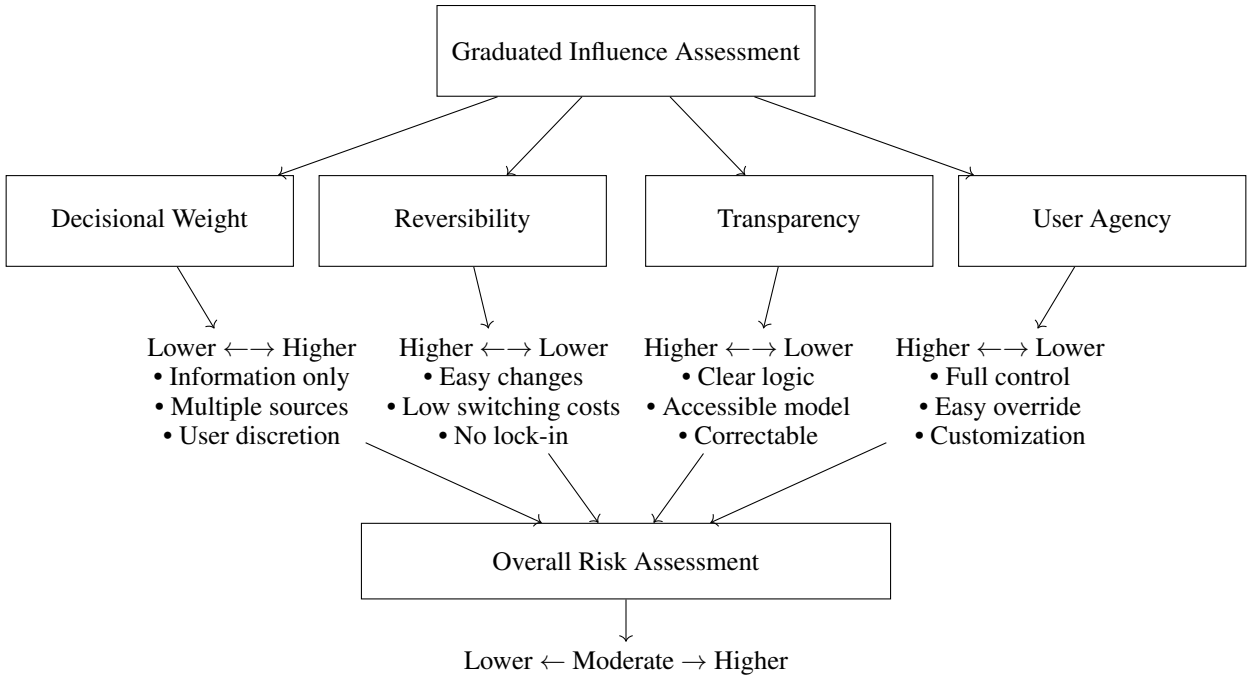

Figure 2: The Graduated Influence Framework

dation logic (matching courses to declared majors, prerequisites, and past performance) and allows students to adjust their profiles, transparency is maintained. User agency is preserved through easy access to full course catalogs and the ability to ignore recommendations entirely. Under this analysis, the system might not pose "significant risk" despite performing profiling, avoiding the potentially disproportionate result of treating beneficial educational technology as high-risk.

- **Case Example 2#**: Contrast this with an employment screening system that uses profiling to filter job applications before human review. Such a system exhibits high decisional weight because it determines which applications receive human consideration (i.e., gatekeeping access to employment opportunities). Reversibility is limited because missed opportunities cannot be recaptured and early career impacts may compound over time. Transparency is often minimal due to proprietary algorithms and competitive concerns. User agency is severely constrained because applicants cannot override filtering decisions or access unfiltered consideration. This system properly warrants high-risk classification based on its actual impacts rather than merely its use of profiling.
- **Case Example 3#**: An edge case involves news recommendation systems on major platforms. These systems exhibit moderate decisional weight—they significantly influence information consumption but users can seek alternative sources. Reversibility varies—individual reading choices are easily changed but cumulative effects on worldview may persist. Transparency depends on implementation—some platforms clearly signal recommendation factors while others do not. User agency varies considerably based on whether platforms offer chronological feeds, topic controls, and easy access to diverse sources. Under the graduated influence assessment, identical technical systems might receive different risk assessments based on their implementation choices, potentially incentivizing transparency and user control rather than technical workarounds.

6 Implications for Regulatory Interpretation

Immediate Interpretive Strategies

Legislative amendment would resolve the profiling paradox definitively, yet courts possess interpretive authority to address its most problematic effects within existing statutory boundaries. As Ebers et al. (2023) demonstrates, there is significant scope for applying a risk-based approach through careful interpretation and regulatory guidance.

This interpretive authority allows courts to address regulatory tensions within existing statutory frameworks, particularly where mechanical application would frustrate legislative purposes. Several interpretive strategies remain available that could help mitigate the paradox while respecting the regulation's text and purpose. The European AI Board, established under Article 65 (European Parliament and Council 2024), has authority to issue guidelines that could operationalize Article 6(3)'s "significant risk" threshold through substantive rather than formal analysis. Such guidance should clarify that the presence of profiling does not automatically equate to significant risk, encouraging case-by-case assessment based on actual rather than theoretical impacts.

First, courts should consider adopting a contextual interpretation of what constitutes "profiling" when imported into the AI Act context. While GDPR Article 4(4) defines profiling broadly for data protection purposes, its incorporation into risk classification may warrant contextual interpretation. Courts could distinguish between minimal preference processing and substantive behavioral evaluation, potentially excluding from high-risk classification systems that merely implement explicit user choices without inference or prediction. This approach finds support in the principle that identical terms may bear different meanings in different regulatory contexts, particularly when mechanical application would frustrate legislative purposes.

Second, regulators should develop guidance that operationalizes Article 6(3)'s "significant risk" threshold through substantive rather than formal analysis. The European AI Board, established under Article 65 (European Parliament and Council 2024), has authority to issue guidelines on the Act's implementation. Such guidance should clarify that the presence of profiling does not automatically equate to significant risk, encouraging case-by-case assessment based on actual rather than theoretical impacts. This interpretation aligns with Recital 53's recognition that Annex III systems may have "little to no impact on fundamental rights" (European Parliament and Council 2024). As Dunn and de Gregorio argue in their analysis of risk-based regulation in digital constitutionalism, the challenge lies in ensuring that risk assessments genuinely reflect the constitutional values the regulation seeks to protect (Dunn and De Gregorio 2022). Recent Commission guidelines published in February 2025 provide initial interpretive guidance, though as Shrishak observes, these guidelines indicate that "opaque recommender systems could be within the scope of the prohibitions" while "AI system using personalised recommendations based on transparent algorithms and user preferences and controls engages in persuasion [and] is not prohibited" (Shrishak 2025). This distinction between opaque and transparent systems aligns with our transparency factor in the graduated influence framework, suggesting emerging regulatory recognition of the need for nuanced assessment.

Third, courts should apply the principle of proportionality, fundamental to EU law, when interpreting the Act's classification provisions. Article 52(1) of the Charter requires that limitations on rights be proportionate to their objectives (European Union 2012). Interpreting the Act to impose identical compliance burdens on beneficial educational recommenders and discriminatory employment screeners may violate proportionality. Courts could introduce proportionality through careful interpretation of undefined terms like "significant risk" and "materially influencing," creating space for contextual assessment within statutory boundaries.

The Role of Technical Standards

The Act's framework for technical standards, outlined in Articles 40 and 41, provides another avenue for addressing the profiling paradox (European Parliament and Council 2024). The European standardization organizations tasked with developing harmonized standards could create differentiated requirements for recommender systems based on their risk profiles rather than their technical characteristics. Standards could establish presumptions of compliance for systems meeting certain transparency, user control, and reversibility criteria, effectively operationalizing the graduated influence approach through technical specifications rather than legal interpretation.

Coordination with Sectoral Regulators

The AI Act's relationship with sector-specific regulation offers additional interpretive flexibility. Article 2(3) provides that the Act applies "without prejudice to" existing Union law, while Recital 10 emphasizes coordination with sectoral requirements (European Parliament and Council 2024). Sectoral regulators in education, employment, and financial services could issue guidance on how recommender systems in their domains should be assessed, potentially establishing safe harbors for systems meeting sector-specific criteria. This approach would allow nuanced assessment based on domain expertise while maintaining consistency with the Act's framework.

7 Conclusion

This Article identified a "profiling paradox" created by the AI Act's incorporation of GDPR Article 4(4) into its risk classification framework. The mechanical importation of a broad data protection definition into risk-based regulation treats beneficial personalization and potentially harmful manipulation as equivalent risks, undermining the Act's proportionality objectives.

The graduated influence doctrine offers courts an interpretive framework to address these tensions within existing statutory bounds. By examining decisional weight, reversibility, transparency, and user agency, judges can introduce proportionality into Article 6(3)'s "significant risk" threshold while preserving legal certainty. This approach aligns with established constitutional principles and recognizes the complex realities of human–machine interaction.

The framework contributes to ongoing scholarly debates by providing specific tools to address profiling-related tensions where Novelli et al. identify general methodological deficits (Novelli et al. 2024) and institutional analyses highlight cross-regulatory uncertainties (European Parliamentary Research Service 2025). It operationalizes Ebers et al.'s demonstration of judicial interpretive authority (Ebers et al. 2023) while respecting technology neutrality principles (Becker 2024). The profiling paradox exemplifies challenges when regulatory concepts migrate across legal domains without adequate translation. Through principled interpretation grounded in proportionality and fundamental rights, courts can establish foundations for algorithmic accountability that serve both innovation and human agency in an algorithmic society.

Ethics Statement

This paper did not benefit from any grant from funding agencies in the public, commercial, or not-for-profit sectors. The author declares no conflicts of interest related to this work.

References

Aggarwal, C. C. 2016. *Recommender Systems: The Textbook*. Cham, Switzerland: Springer. ISBN 978-3-319-29659-3.

Barocas, S.; and Selbst, A. D. 2016. Big Data's Disparate Impact. *California Law Review*, 104: 671–732.

Becker, F. 2024. Regulating the Risks of AI: Challenges in the Intersection of the GDPR and AI Act. *SSRN Electronic Journal*. Available at SSRN: https://ssrn.com/abstract=4870387.

Belkin, N. J.; and Croft, W. B. 1992. Information filtering and information retrieval: Two sides of the same coin? *Communications of the ACM*, 35(12): 29–38.

Black, J. 2002. Regulatory Conversations. *Journal of Law and Society*, 29(1): 163–196.

Bradford, A. 2020. *The Brussels Effect: How the European Union Rules the World*. Oxford University Press.

Bygrave, L. A. 2020. Machine Learning, Automated Profiling, and the GDPR. *European Data Protection Law Review*, 6(1): 24–35.

Court of Justice of the European Union. 2023. Opinion of Advocate General Pikamäe in Case C-634/21, SCHUFA Holding and Others. ECLI:EU:C:2023:220.

de Búrca, G. 1993. The Principle of Proportionality and its Application in EC Law. *Yearbook of European Law*, 13(1): 105–150.

De Gregorio, G.; and Dunn, P. 2022. The European Risk-Based Approaches: Connecting Constitutional Dots in the Digital Age. *Common Market Law Review*, 59(2): 473–500.

Dunn, P.; and De Gregorio, G. 2022. Risk-Based Regulation in Digital Constitutional Democracy. *Computer Law & Security Review*, 45: 105681.

Ebers, M.; Hoch, V.; Rosenkranz, F.; Ruschemeier, H.; and Steinrötter, B. 2023. Regulating Explainable AI in the European Union: An Overview of the Current Legal Framework(s). *European Journal of Risk Regulation*, 14(2): 234–255.

Eubanks, V. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press.

European Commission. 2024. Public Consultation on a Code of Practice for providers of general-purpose AI models. Available at: https://digital-strategy.ec.europa.eu/en/consultations.

European Parliament and Council. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). OJ L 119, 4.5.2016.

European Parliament and Council. 2022. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services (Digital Services Act). OJ L 277, 27.10.2022.

European Parliament and Council. 2024. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). OJ L 2024/1689, 12.7.2024.

European Parliamentary Research Service. 2025. Algorithmic discrimination under the AI Act and the GDPR. Authors: De Luca, Stefano and Federico, Marina. PE 769.509, February 2025.

European Union. 2012. Charter of Fundamental Rights of the European Union. OJ C 326, 26.10.2012.

Fleder, D.; and Hosanagar, K. 2009. Blockbuster Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity. *Management Science*, 55(5): 697–712.

Gellert, R. 2020. Understanding the Notion of Risk in the General Data Protection Regulation. *Computer Law & Security Review*, 34(2): 279–288.

Hanani, U.; Shapira, B.; and Shoval, P. 2001. Information filtering: Overview of issues, research and systems. *User Modeling and User-Adapted Interaction*, 11: 203–259.

Helberger, N.; and Diakopoulos, N. 2023. The European AI Liability Directives – Critique of a Half-hearted Approach and Lessons for the Future. *Digital Journalism*, 11(9): 1645–1649.

Kaminski, M. E. 2023. Regulating the Risks of AI. *Boston University Law Review*, 103: 1347–1411.

MacKenzie, C.; and Stoljar, N. 2000. *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*. Oxford University Press.

Malgieri, G. 2019. Automated Decision-Making in the EU Member States: The Right to Meaningful Information and the 'Suitable Safeguards' for Profiling. *Computer Law & Security Review*, 35(5): 105327.

Milano, S.; Taddeo, M.; and Floridi, L. 2020. Recommender Systems and their Ethical Challenges. *AI & Society*, 35(4): 957–967.

Nedelsky, J. 2011. *Law's Relations: A Relational Theory of Self, Autonomy, and Law*. Oxford University Press.

Novelli, C.; Casolari, F.; Rotolo, A.; Taddeo, M.; and Floridi, L. 2024. AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act. *Digital Society*, 3(13): 1–29.

Pagallo, U. 2023. Why the AI Act Won't Trigger a Brussels Effect. *AI Approaches to the Complexity of Legal Systems*, forthcoming.

Ricci, F.; Rokach, L.; and Shapira, B. 2022. *Recommender Systems Handbook*. Springer, 3rd edition.

Selbst, A. D.; and Powles, J. 2021. Meaningful Information and the Right to Explanation. *International Data Privacy Law*, 7(4): 233–242.

Sen, A. 1999. *Development as Freedom*. Oxford University Press.

Shrishak, K. 2025. EU's AI Act: Tread the Guidelines Lightly.

Siegmann, C.; and Anderljung, M. 2022. The Brussels Effect and Artificial Intelligence: How EU regulation will impact the global AI market. *arXiv preprint arXiv:2208.12645*.

Thaler, R. H.; and Sunstein, C. R. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Yale University Press.

Veale, M.; and Binns, R. 2021. Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR. *International Data Privacy Law*, 11(4): 319–335.

Veale, M.; and Borgesius, F. Z. 2021. Demystifying the Draft EU Artificial Intelligence Act. *Computer Law Review International*, 22(4): 97–112.

Yeung, K.; and Lodge, M. 2022. Demystifying AI Governance: A Conceptual Framework for Understanding and Regulating Artificial Intelligence. *The Oxford Handbook of AI Governance*, forthcoming.

Zhang, S.; Yao, L.; Sun, A.; and Tay, Y. 2019. Deep Learning based Recommender System: A Survey and New Perspectives. *ACM Computing Surveys*, 52: 1–38.

Zhu, Y.; Wu, L.; Guo, Q.; Hong, L.; and Li, J. 2023. How Can Recommender Systems Benefit from Large Language Models: A Survey. *ACM Transactions on Information Systems*, 42(2): 1–30.