frontiers | Frontiers in Communication

TYPE Original Research
PUBLISHED 20 January 2026
DOI 10.3389/fcomm.2025.1620310

The right to game AI-systems: a speculative right for contestation

Freyja van den Boom[1,2]*

[1]Research Group Personal Rights & Property Rights, Centre National de la Recherche Scientifique, Paris, France, [2]Law Department, University of Antwerp, Antwerpen, Belgium

OPEN ACCESS

EDITED BY
Jesse Josua Benjamin,
Eindhoven University of Technology,
Netherlands

REVIEWED BY
Simon David Hirsbrunner,
University of Tübingen, Germany
Denisa Reshef Kera,
Tel Aviv University, Israel

*CORRESPONDENCE
Freyja van den Boom
✉ freyja.vandenboom@uantwerpen.be

RECEIVED 08 May 2025
ACCEPTED 22 September 2025
PUBLISHED 20 January 2026

CITATION
van den Boom F (2026) The right to game AI-systems: a speculative right for contestation.
Front. Commun. 10:1620310.
doi: 10.3389/fcomm.2025.1620310

This paper proposes the 'right to game AI-systems' as a speculative design artifact to challenge dominant narratives that position 'gaming' as a threat to algorithmic integrity. We argue that in high-stakes domains like insurance, health, and welfare, gaming the system should be recognized as a legitimate and necessary act of agency, resistance, and contestation. Rooted in a critical reading of the GDPR and the EU AI Act, and employing Causal Layered Analysis (CLA) and speculative design, we reframe the 'right to game' as a vital response to structural opacity and the unequal power dynamics inherent in AI governance. By connecting 'gaming' to established concepts of contestability, ethical hacking, and playful exploration, this paper argues for a radical shift in perspective that empowers individuals to become active participants in, rather than passive subjects of, algorithmic decision-making.

KEYWORDS

algorithmic governance, contestability, transparency, accountability, speculative design, futures studies, socio-legal studies, GDPR

1 Introduction: systems we're not allowed to game

Artificial Intelligence (AI) has rapidly evolved from a speculative technology into a ubiquitous, powerful force, fundamentally reshaping decision-making processes across critical sectors. From assessing creditworthiness and insurance premiums to allocating healthcare resources and determining eligibility for social welfare, AI-driven systems are now instrumental in governing people's lives (O'Neil, 2016). This pervasive integration means that automated decisions are routinely made about individuals, often with profound and life-altering consequences (O'Neil, 2016; Pasquale, 2015).

Numerous real-world examples underscore the critical need for transparency and contestability in algorithmic governance. In the German case of *Hesse v. Agentur für Arbeit*, a court questioned the transparency and accountability of a risk scoring algorithm (the Austrian AMS system) used to assess job seekers, highlighting concerns about individuals being disadvantaged by an inscrutable system.[1] The widely reported A-level grading scandal in the United Kingdom demonstrated how an opaque algorithm, initially presented as a neutral tool, disproportionately downgraded students from marginalised backgrounds based on factors seemingly unrelated to their individual performance, revealing a stark example of algorithmic bias and a lack of accountability for its discriminatory impact (UK Parliament, House of Commons Education Committee, 2020). In the Netherlands, the SyRI (System Risk Indication) system, designed to detect welfare fraud using a range of personal data, was ultimately ruled unlawful by a court due to its violation of human rights, particularly the right to privacy, but also raising significant concerns about its potential for discrimination against low-income and migrant communities (NJCM v The Dutch State, 2020; Van Bekkum and Borgesius, 2021).

[1] Hesse v Agentur für Arbeit Sozialgericht Gießen, Judgment of 20 February 2020, S 14 AS 101/19.

The opacity of the algorithm made it impossible for affected individuals to understand why they were flagged as potential risks, denying them the opportunity to meaningfully contest the basis of the state's suspicion.

These cases are not isolated incidents; they are symptomatic of a broader phenomenon where the opacity of algorithmic decision-making systems denies individuals the opportunity to understand, scrutinize, and effectively contest the basis upon which they are being judged. This lack of transparency is particularly problematic because, as critical scholars have demonstrated, algorithmic systems frequently reproduce and amplify existing societal biases and structural inequalities (Eubanks, 2018; Couldry and Mejias, 2019). Trained on historical data that reflects past and present discrimination, and designed with objectives and logics that may implicitly favor dominant groups, these systems can inadvertently or intentionally entrench disadvantage and reinforce existing power structures and social hierarchies under the guise of technical neutrality and efficiency (Couldry and Mejias, 2019).

The opacity of AI systems not only hinders individual understanding and contestation but also tends to obscure the political choices and value judgments embedded within algorithmic design and deployment. This can lead to outcomes that are not objective, data-driven inevitabilities but the result of deliberate design decisions, data selection, and power dynamics (Winner, 1980). This makes it harder to identify the mechanisms through which inequality is reproduced and limits opportunities for people to take action (Mittelstadt et al., 2016; Zarsky, 2013).

To make this more concrete, this is about when for example: your car insurance premium is determined by a telematics "black box" that monitors your every move; or when you need healthcare and your treatment options are ranked by a health algorithm; or when you are looking for a job and your opportunities are filtered by an automated hiring system These are systems, we are subject to, but not allowed to understand, much less challenge in a meaningful way.[2] In each of these cases, the system's logic may be protected as a corporate asset. This opacity denies individuals the opportunity to scrutinize and effectively contest the basis upon which they are being judged, often reproducing and amplifying existing societal biases and structural inequalities under a guise of technical neutrality (Mittelstadt et al., 2016).

In this context, we ask how individuals can exercise meaningful control over their lives. This paper argues that addressing the harms of opaque algorithmic governance requires a fundamental shift in perspective. We propose a provocative concept: the right for individuals to 'game' AI systems. This concept is introduced not in the sense of malicious exploitation, but as a speculative design artifact that challenges dominant narratives. We argue that in high-stakes domains, understanding how these systems work and adjusting one's behavior in response should be recognized as a legitimate act of agency and participatory sense-making.

This paper will proceed as follows. Section 2 will analyze the current regulatory landscape and use Causal Layered Analysis (CLA) to deconstruct the fear of gaming that underpins algorithmic secrecy. Section 3 will reconceptualize gaming as a powerful form of contestation, linking it to academic literature on contestable AI, ethical hacking, and playful exploration. Section 4 will operationalize this concept through a speculative design scenario: the *"Fair Play Insurance Dashboard"* to illustrate its practical benefits. Section 5 will then discuss the ethical parameters and implications of such a right. Finally, we conclude by advocating for a future where individuals are empowered as active algorithmic citizens, not as passive data subjects.

2 The status quo: algorithmic secrecy and the fear of 'gaming'

The prevailing approach to AI governance is characterized by a fundamental tension: a stated desire for transparency on one hand, and a deep-seated institutional and economic structure that fiercely protects algorithmic secrecy on the other. This section first examines the legislative landscape that enables this conflict and then uses Causal Layered Analysis (CLA) to uncover the deeper narratives that fuel the fear of gaming the system.

2.1 The regulatory paradox: the right to access vs. trade secrecy

The European Union's General Data Protection Regulation (GDPR) and the AI Act represent landmark efforts to regulate algorithmic decision-making. Articles 15 and 22 of the GDPR grant individuals the right to meaningful information about the logic involved in automated decision-making (Regulation (EU), 2016; Malgieri and Comandé, 2017).

Article 15, the right of access by the data subject, is particularly relevant.[3] It grants individuals the right to obtain confirmation as to whether personal data concerning them is being processed, and if so, to access that data and specific information about the processing. Article 15(1)(h) states that this information includes *"meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject."* This provision seems to directly address the need for individuals to understand how automated decisions affecting them are made.

Furthermore, Article 22 grants data subjects the right "not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her." While this article provides a right against solely automated decisions in high-stakes scenarios, it implicitly underscores the need for transparency and human involvement where such decisions are permitted or used to inform human decisions.

However, these promises of transparency are significantly undermined by countervailing protections for corporate interests. Recital 63 of the GDPR explicitly states that the right of access "*should not adversely affect the rights or freedoms of others, including trade secrets or intellectual property*". [4]

[2] See *Dun & Bradstreet Austria*, Case C-203/22, Judgment of 27 February 2025, Court of Justice of the European Union.

[3] We argued for the need for a broad interpretation of access rights and trade secrets (Van den Boom, 2020).

[4] Recital 63 GDPR.

In practice, this exception is often used to block any meaningful access (Wachter et al., 2017). People trying to obtain greater algorithmic transparency under the GDPR have frequently encountered this trade secret defense from companies, highlighting the practical difficulties individuals face in exercising their rights when confronted with powerful corporate interests (Veale et al., 2021).

When a driver asks their telematics insurer why their premium has increased, they are likely to receive vague categories of risk factors (driving style, route choices) rather than the specific weights and data points that constitute the meaningful information they need (Bucher, 2017). The algorithms remain a black box. This creates a regulatory paradox where the right to transparency exists in theory but is often unenforceable in the face of corporate claims of confidentiality (Veale et al., 2021).

The European Union's Artificial Intelligence Act (AI Act) represents a first comprehensive legal framework.

A framework specifically designed to regulate AI systems.[5] Adopting a risk-based approach, the AI Act imposes varying levels of obligations on AI systems depending on their potential to cause harm. AI systems used in critical areas such as healthcare, education, employment, and law enforcement are classified as high-risk and are subject to stricter requirements.

The AI Act includes specific transparency obligations for high-risk AI systems under Article 13 and broader transparency obligations for certain AI systems under Article 52. For high-risk systems, providers must design and develop them to enable human oversight and provide accompanying information that is "clear and adequate." This information is intended to help users understand the system's capabilities and limitations.

Article 52 imposes specific transparency obligations for AI systems intended to interact with individuals or generate content, requiring users to be informed that they are interacting with or exposed to AI.

While the AI Act reinforces the importance of transparency and accountability for high-risk AI, it does not fully resolve the problems we face with implementing the GDPR in a way that protects trade secrets (Oxford Law Blogs, 2025). While it builds upon GDPR principles, the AI Act's focus is more on the safety and fundamental rights risks posed by the AI system itself, rather than the data protection rights of individuals concerning automated decisions.

The Act aims to balance innovation with safety and rights, but the specific details of how transparency will be enforced and how it will be weighed against claims of confidentiality remain subject to ongoing debate and the development of implementing standards and guidelines.

The burden often remains on affected individuals or civil society organizations to identify problematic AI systems and advocate for greater transparency and accountability.

Furthermore, the regulatory landscape for AI transparency is fragmented. Even with the GDPR and AI Act, individuals seeking to understand how algorithmic decisions are made about them must navigate complex interactions between data protection law, sector-specific regulations (e.g., in healthcare or finance), and intellectual property law (Mittelstadt, 2019).

The lack of a single, clear, and universally enforceable right to truly understand algorithmic logic significantly hinders the ability of individuals to exercise agency and contest decisions that impact their lives, contributing to the opacity that underpins the reproduction of inequality.[6]

2.2 Unpacking the fear of gaming: a causal layered analysis

To understand the persistent resistance to transparency, we must look deeper than legal texts. Causal Layered Analysis (CLA), a futures research method, helps deconstruct the narratives that sustain the status quo (Inayatullah, 1998). CLA is a poststructuralist futures research method that moves beyond conventional, surface-level analyses of issues to uncover the deeper causes, worldviews, and metaphors that shape our understanding and limit possibilities for change (Inayatullah, 1998). By applying CLA to the debate surrounding algorithmic transparency and the fear of providing access and transparency, we can peel back the layers and reveal the often-hidden power dynamics and ideological commitments that maintain the status quo of algorithmic opacity (Figure 1).

CLA operates on four distinct, yet interconnected layers:

(a) Litany: immediate concerns and the fear of gaming

At the surface level, the litany of the algorithmic transparency debate is dominated by immediate concerns about the potential negative consequences of granting individuals access to the inner workings of AI systems. The narrative frequently presented in media, policy discussions, and by corporate actors focuses on the risk that users will game the system if they understand its logic (Van den Boom, 2020). Examples cited include individuals manipulating credit scoring algorithms to improve their rating without genuine financial responsibility, drivers altering behavior only when telematics systems are active, or students finding loopholes in educational assessment AI.

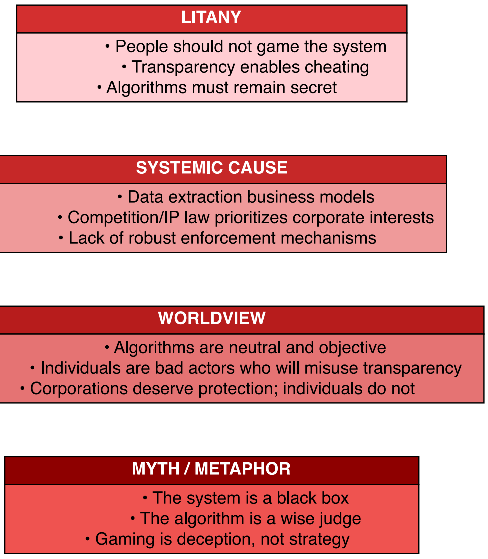
This layer emphasizes threats to algorithmic integrity, system security, and the potential for widespread manipulation or exploitation. The proposed solutions at this level often involve technical safeguards to prevent gaming, legal penalties for misuse, or simply maintaining secrecy to make gaming impossible. This litany, while containing elements of valid concern about system security, tends to frame the issue as a problem of malicious individual behavior that must be controlled.

(b) Systemic causes: structures of power and competition

Moving beneath the surface litany, the systemic layer reveals the underlying social, economic, and legal structures that give rise to and sustain the fear-of-gaming narrative. A primary systemic cause is the economic incentive for companies to protect their AI algorithms as

[5] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 (Artificial Intelligence Act).

[6] This is despite the CJEU, which recently confirmed the right to explanation of automated decision-making. Instead, the balancing of interests must be carried out on a case-by-case basis, *Dun & Bradstreet Austria*, Case C-203/22, para. 75 (CJEU, 2025). see https://curia.europa.eu/juris/document/document.jsf?text=&docid=295841&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=1555350.



LITANY
- People should not game the system
- Transparency enables cheating
- Algorithms must remain secret

SYSTEMIC CAUSE
- Data extraction business models
- Competition/IP law prioritizes corporate interests
- Lack of robust enforcement mechanisms

WORLDVIEW
- Algorithms are neutral and objective
- Individuals are bad actors who will misuse transparency
- Corporations deserve protection; individuals do not

MYTH / METAPHOR
- The system is a black box
- The algorithm is a wise judge
- Gaming is deception, not strategy

FIGURE 1
CLA applied.

valuable trade secrets and intellectual property. In a highly competitive market, the specific design, training data, and operational parameters of a performant algorithm can represent a significant competitive advantage. Transparency is perceived as a direct threat to this advantage, potentially allowing competitors to replicate successful models without incurring the same research and development costs (Van den Boom, 2020). This economic structure creates a powerful, vested interest in maintaining algorithmic opacity, and the legal frameworks surrounding intellectual property often provide robust mechanisms for doing so, creating a direct conflict with data protection and transparency rights.

Furthermore, the concentration of power within large technology companies and institutions that develop and deploy AI systems is a key systemic factor. This concentration of power allows these actors to shape the discourse around AI, influencing policy and public perception.

The fear of gaming can be strategically amplified by those in power to justify maintaining control and limiting external scrutiny. The complex technical nature of advanced AI also acts as a systemic barrier, creating an information asymmetry between those who build and deploy AI and those who are subjected to its decisions, reinforcing existing power inequality. The fragmented and often weakly enforced regulatory landscape for AI accountability also contributes to this layer by failing to create sufficient systemic pressure for meaningful transparency.

(c) Worldview: beliefs in control and market primacy

The systemic causes are, in turn, supported by deeper worldviews and paradigms. A dominant worldview underpinning the fear-of-gaming narrative is a strong belief in control and order imposed from the top down. This worldview views systems as entities to be managed and protected by experts and authorities, with users as passive recipients or potential disruptors who need to be controlled or contained. From this perspective, granting individuals the power that comes with understanding a system's inner workings is inherently risky and undesirable, as it might lead to unpredictable outcomes outside of intended control.

Another powerful worldview at play is the prioritization of market dynamics and corporate interests over individual rights and democratic accountability. This perspective holds that the pursuit of economic efficiency, innovation (often narrowly defined by market advantage), and corporate profitability are paramount. Within this worldview, the protection of trade secrets is seen as essential for market functioning and innovation, and concerns about individual transparency rights are secondary or viewed as obstacles to progress. This worldview often assumes that market competition will eventually lead to optimal outcomes, including trustworthy AI, without the need for extensive external regulation or mandatory transparency that might impede corporate strategies. This perspective often downplays or fails to adequately account for how market forces can exacerbate, rather than mitigate, inequality and social harms when left unchecked (Whittaker, 2021).

(d) Myth/Metaphor: Underlying Beliefs about Human Nature and AI

At the deepest layer, the worldviews are sustained by powerful, often unconscious, myths and metaphors. The fear of gaming taps into deep-seated cultural myths about human nature, often portraying individuals as fundamentally self-interested actors prone to cheating and exploiting systems for personal gain. This myth justifies the need for external control and surveillance, reinforcing the idea that individuals cannot be trusted with power or knowledge about the systems that govern them.[7]

At the same time, there are powerful myths surrounding AI itself. AI is often portrayed metaphorically as objective, a neutral judge capable of

[7] For an analysis of the arguments, see Busuioc et al. (2023) who found that *[.] the effectiveness of secrecy as an antidote for gaming is far from uncontested.* Busuioc M, Curtin D, Almada M. Reclaiming transparency: contesting the logics of secrecy within the AI Act. *European Law Open*. 2023;2(1):79–105; Rudin, C. (2019). Stop explaining black box machine learning models for high-stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*, 206.

making perfectly rational decisions free from human bias (Gillespie, 2014). We also see the narrative that AI systems are too hard to understand even for expert, so we just have to accept instead of scrutinize these systems.[8] These myths contribute to a sense of technological determinism or inevitability best left in the hands of a select few experts (Morozov, 2013; Broussard, 2019; De Liban, 2024). The combination of the myth of the untrustworthy individual and the myth of the inscrutable or infallible algorithm enables the narrative that denying transparency is necessary and appropriate (Schellmann, 2024; Kroll et al., 2017). The algorithm becomes a benevolent, necessary authority that must be protected from the potentially malicious actions of individuals seeking to exploit its secrets (Kak and West, 2023).

Using Causal Layered Analysis reveals that refusing transparency and access rights is not a simple response to security risks. It is a narrative deeply embedded in structures of power, competition, and worldviews that prioritise corporate control over individual agency (Van den Boom, 2022). Challenging this requires more than demanding transparency; it requires a fundamental reframing of what it means to interact with an algorithmic system (Rudin, 2019).

3 Reimagining the right to gaming systems as a right of contestation

We deliberately use the provocative term *gaming* to challenge its negative connotation and reframe it as a legitimate and necessary form of contestation in an algorithmic society. This section builds a theoretical foundation for this re-appropriation, connecting it to established legal and interdisciplinary debates on contestability (Wachter et al., 2020), civic resistance (Cohen, 2019), play as critique (Sicart, 2014), And ethical hacking (Bellaby, 2023).

3.1 From illegitimate cheating to a right of contestation

In everyday use, *gaming the system* implies deception or manipulation. We acknowledge this cultural framing. However, in the context of opaque, high-stakes AI systems, often shielded from scrutiny by trade secrets or proprietary protections, gaming becomes a rational and necessary response to structural power imbalances (Kak and West, 2023; Edwards and Veale, 2018). When official channels for redress or explanation fail, users are left with few options but to experiment, test, or subvert the system to understand or challenge it.

We argue that *gaming* in this context takes on multiple democratic functions:

- Agency: Reclaiming control in systems where individuals are typically positioned as passive subjects of computation, subject to automated decisions without recourse (Cohen, 2019).
- Resistance: Pushing back against dominant narratives that position algorithmic outcomes as objective or inevitable (Eubanks, 2018).
- Participatory sense-making: Engaging with algorithmic systems not just to interpret their outputs but to actively make sense of how they construct subjects and realities (Lindley et al., 2020).
- Behavioral self-modification: Using knowledge gained from gaming to adapt one's behavior and achieve fairer outcomes within algorithmic systems (Ananny and Crawford, 2018).

This reframing aligns with the growing literature on *contestable AI*, which seeks to provide procedural mechanisms for users to intervene in automated decision-making (Wachter et al., 2020). Our concept of the *right to game AI systems* is a speculative legal proposition that extends beyond explanation rights, advocating for user-driven practices of resistance, redress, and reappropriation.

3.2 Playing with the trouble

This re-appropriation finds further grounding in the legal recognition of *ethical hacking, red teaming, and adversarial testing* as legitimate modes of system critique (European Union., 2024; Veale, 2020). However, these practices are usually reserved for technical experts under institutional oversight. Our speculative intervention envisions a future in which such practices are *democratized.* Here, the individuals most impacted by algorithmic systems (welfare recipients, insured drivers) are empowered to act as *civic auditors,* drawing on their lived experience to test, probe, and challenge the logic and effects of these systems (Van den Boom, 2023).

In this framing, gaming shifts from a self-interested or deceptive practice into a distributed method of algorithmic accountability. It functions as a form of *grassroots red teaming*, allowing users to stress-test decision systems and demand more robust, just, and transparent design. Unpacking these socio-technical infrastructures requires interventions from multiple, distributed vantage points (Crawford and Joler, 2018).

3.3 Why gaming and not just contesting?

While the notion of *contestability* has gained traction in AI governance discourse, it often remains tethered to formal legal procedures or institutional processes (Wachter et al., 2020). We retain the term *gaming* because of its speculative, insurgent potential. *Contesting* suggests recourse within an existing system; *gaming* implies tactics deployed precisely when those systems are inaccessible, incomplete, or untrustworthy.[9]

The *Right to Game AI Systems* is thus presented not as a formal legal right in the traditional sense, but as a speculative legal artifact, a tool to provoke discussion about what rights might be needed when we are governed by inscrutable, non-negotiable infrastructures. It reflects the critical legal insight that law itself is often structured to deny access or recognition to certain subjects, and that resistance must

[8] This understanding that systems are too complex to be understood by ordinary people was also refuted by the CJEU in *Dun & Bradstreet Austria*, Case C-203/22 (CJEU 2025).

[9] Cohen (2019); Eubanks.[13]

often come from outside its formal channels (Delacroix and Wagner, 2021).

4 A speculative artifact: the right to game AI-systems in practice

Causal Layered Analysis (CLA) helps us break down the deeper stories and systems that support the lack of transparency in algorithmic technologies. Speculative design builds on this by offering a way to go beyond critique (Van den Boom, 2023). It allows us to imagine and explore futures where people are not simply governed by algorithms but have agency and control over them. This section introduces a speculative scenario that puts into practice the idea of a right to game AI systems.

Speculative design is not about solving current issues or creating market products. Rather, it provides a way to ask "what if?" about technological and societal direction, allowing us to envision futures beyond existing legal, social, or technological frameworks. It challenges dominant worldviews and opens doors to imagining alternatives (Dunne and Raby, 2013; Lindley and Green, 2021). Within AI governance, it enables us to question the passive roles often assigned to users and imagine futures where power is redistributed toward individuals (Lindley et al., 2020).

The *Right to Game AI-Systems* is a speculative legal artifact, a fictional, provocative tool. It is not a legally enforceable right, but a means to rethink how people might contest or engage with algorithmic systems. By imagining individuals who can test, deceive, or resist algorithmic decisions, the artifact aims to surface assumptions embedded in current governance frameworks and invite alternative models of fairness, accountability, and agency (Lindley et al., 2020).

Speculative design turns abstract values like transparency and resistance into tangible experiences that can be imagined, lived through, and discussed. It concretizes intangible concepts in narrative or material form.[10] By envisioning a future where individuals have the right to game AI systems, we can explore both positive outcomes and unintended consequences and anticipate ethical or legal challenges (Pschetz et al., 2017; Tallyn et al., 2018).

More broadly, this speculative approach reshapes our view of human–AI relations. Rather than focusing solely on protection from harm, it invites consideration of how people might actively shape, resist, or subvert AI systems. The *Right to Game* framework prompts reflection on who holds power in algorithmic systems and under what conditions such power can be contested (Lindley et al., 2020).

Consider the example of a driver whose insurance premium is set by a telematics black box. Instead of passive acceptance, we imagine a speculative tool called *The Fair Play Insurance Dashboard*. Although fictional, this interface makes algorithmic decisions visible, contestable, and even gameable by the user.[11]

Our driver logs into her insurance portal and is greeted by the Fair Play Dashboard. This innovative tool goes beyond simply displaying her premium; it empowers her with four key features: the module, the explorer, the simulator, and alerts.

The fair play insurance dashboard

Radical transparency module: This section lists every data point the insurer's algorithm uses: every trip, start/end times, speed, acceleration/braking patterns, routes taken, and even contextual data like weather and traffic density. It also lists non-driving data points that might be used, such as the car's model, age, and color.	**Logic and weights explorer:** This is the core of the dashboard. It displays the key factors influencing the Driver risk score and, crucially, their relative weights. For example: *Hard Braking Events: 35%* *Driving Between 11 p.m. - 5 a.m.: 25%* *Exceeding Speed Limit: 20%* *Driving in High-Risk Zones: 15%* *Total Mileage: 5%*
The gaming simulator: This is an interactive tool where the driver can play with the model. They can use sliders to adjust variables and see the immediate impact on simulated premium. For example, "*What if I had made 50% fewer hard-braking maneuvers last month?"* The simulator shows a premium drop of 15%. *"What if I avoided all driving after 11 p.m.?"* The simulator shows a premium drop of 20%.	**Bias and fairness alerts:** *"Our model flags the 'North Industrial' zone as high-risk, which increases your premium. We recognize this may disproportionately affect residents or workers in this area.* *You have the right to request a review of this factor."* This feature turns the driver from a passive subject into an active participant in ensuring the system's fairness.

These features allow her to actively 'play' with her premium and driving behavior. These features allow her to interactively explore her premium and driving behavior.

The dashboard is intended to spark discussion rather than provide direct assistance. By enabling drivers to experiment with different scenarios without altering their actual behavior, they can "win" by lowering their premiums.

This approach highlights the fact that insurers may not fully understand or have access to the algorithms they use. Research shows that insurers set premiums using algorithms that lack transparency regarding the factors influencing risk scores. By outsourcing risk scoring algorithms, insurers no longer know whether there is a clear link between driving behavior, risk scores, and the premiums drivers pay. This ambiguity can lead to potential unfair discrimination, as drivers may be unaware of the underlying data and logic that determine their rates (Van Bekkum et al., 2025). The Fair Play Dashboard aims to provoke people to challenge how we allow algorithmic decision-making in insurance and other aspects of our lives that have a serious impact.

5 The parameters and implications of a right to game

With this example of a beneficial vision of gaming the AI system, we now turn to the ethical framework needed to guide its implementation. Having the right to game AI systems should not be absolute. There are serious concerns and negative consequences when people are allowed to

[10] Dunne & Raby[12].

[11] See for other examples, Bitbarista, which was designed to provoke reflection on autonomy and data norms (Tallyn et al., 2018).

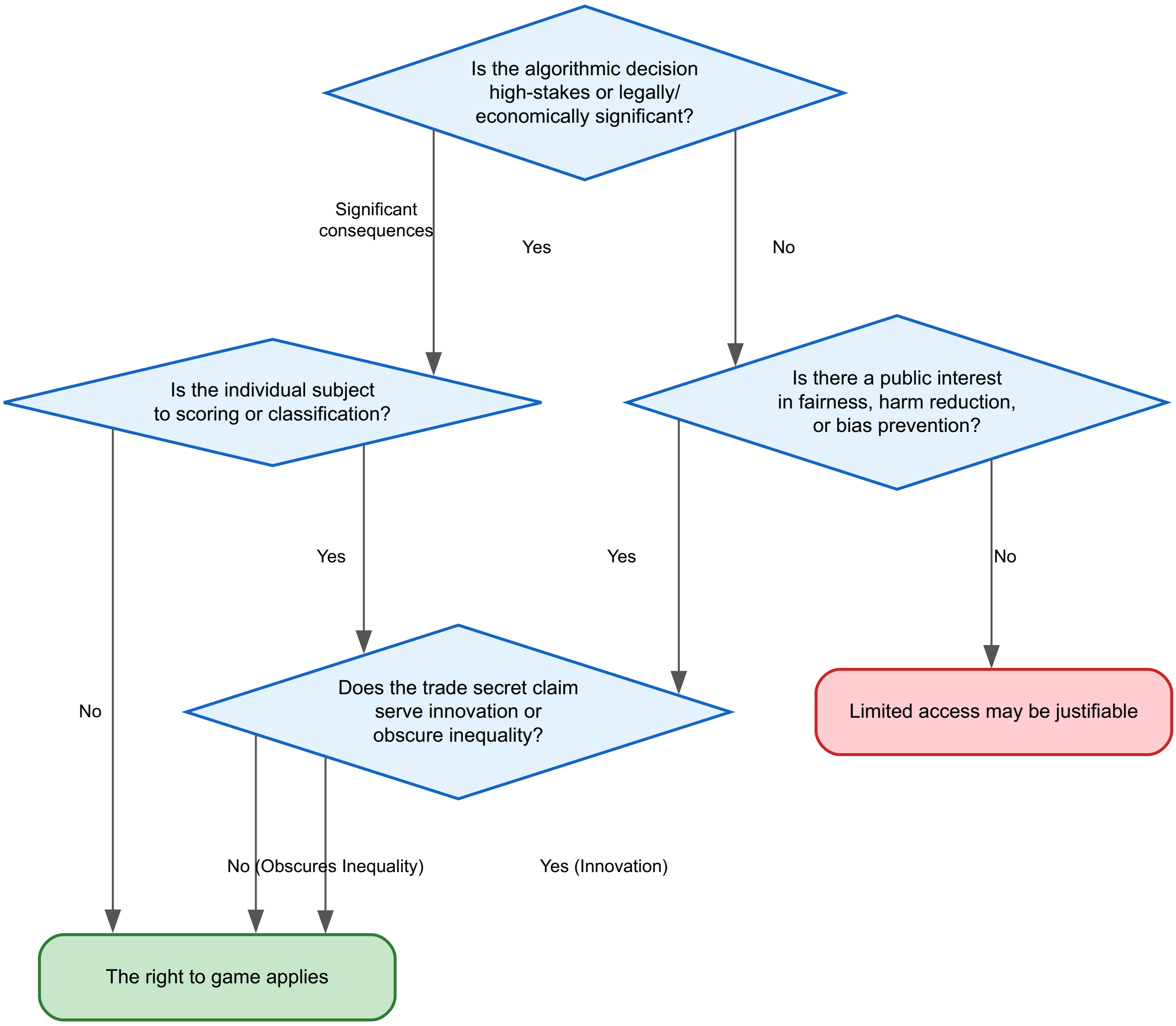

Is the algorithmic decision high-stakes or legally/ economically significant?
Significant consequences
Yes
No
Is the individual subject to scoring or classification?
Is there a public interest in fairness, harm reduction, or bias prevention?
Yes
Yes
No
No
Does the trade secret claim serve innovation or obscure inequality?
Limited access may be justifiable
No (Obscures Inequality)
Yes (Innovation)
The right to game applies

FIGURE 2
Decision flow for applying the right to game.

challenge systems through creative practices. In the context of insurance, for example, when people know the company will have a threshold before they will investigate, people may abuse this knowledge and make sure their claims are below the threshold instead of filing for the actual (lower) amount. Therefore, the right we propose must be balanced against societal interests and be the least likely to cause harm.

The following conditions, presented here as a decision flow, can serve as a guide for determining when this right should apply (Figure 2).

This framework helps to set the boundaries of the right. It is not a call against the protection of trade secrets, but to contest that when an algorithm functions as a gatekeeper to important interests of individuals, the decision must be in favor of radical transparency and individual agency.[12]

The Fair Play Dashboard scenario meets all these criteria: insurance is high-stakes, Drivers are being scored, there is a public interest in fair pricing, and the insurer's claim of secrecy risks obscuring biases against certain drivers or neighborhoods.

6 Conclusion: toward algorithmic citizenship

This paper has argued for a fundamental reframing of our interaction with AI systems. We have challenged the narrative that positions gaming the system as a threat, re-imagining it as a legitimate and necessary right of contestation. By connecting this provocative concept to established academic literature and illustrating its potential through a speculative design scenario, we have shown how empowering individuals to understand and engage with algorithmic logic can lead to fairer, more equitable, and more effective outcomes.

The fear of individuals gaming AI systems they are subject to is a narrative that serves to protect existing power structures. By reframing this, we advocate for a future where individuals transition from being passive data subjects to becoming active

[12] On whether the public interest should outweigh secrecy. Oxford Law Blogs. (2025, July 23). *Secrecy without oversight: How trade secrets could potentially undermine the AI Act's transparency mandate*.

stakeholders, enabled to make well-informed decisions (Gillespie et al., 2014). This requires a fundamental shift in how we develop and regulate AI, prioritizing transparency, agency, and accountability. Embracing this openness is essential for fostering trust and ensuring that the future shaped by AI is one where power is more equitably distributed and fundamental human rights are protected.

In other words, what we have experienced is that using speculative scenarios helps make the benefits of the right to game tangible. It transforms the relationship between the driver and their insurer from one of opaque judgment to one of transparent negotiation. Furthermore, it can help companies to shift from merely avoiding punishment to actively pursuing improvement because there are clear rules they and others can follow. Instead of accepting given narratives, using speculative design can raise awareness and improve stakeholder engagement to lead to better outcomes for both the individual (lower premiums, safer driving) and society (fewer accidents, fairer pricing; Gillespie et al., 2014).

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

FB: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that Gen AI was used in the creation of this manuscript. For conceptualisation and redraft.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Amerirad, B., Cattaneo, M., Kenett, R. S., and Luciano, E. (2023). Adversarial Artificial Intelligence in Insurance: From an Example to Some Potential Remedies. *Risks*, 11, 20. doi: 10.3390/risks11010020

Ananny, M., and Crawford, K. (2018). Seeing without knowing: limitations of the transparency ideal. *New Media Soc.* 20, 973–989. doi: 10.1177/1461444816676645

Bellaby, R. W. (2023). *Hacks, hackers, and political hacking*. Bristol: Bristol University Press.

Broussard, M. (2019). *Artificial unintelligence: How computers misunderstand the world*. Cambridge, MA: MIT Press.

Bucher, T. (2017). The algorithmic imaginary: exploring the ordinary effects of Facebook algorithms. *Inf. Commun. Soc.* 20, 30–44. doi: 10.1080/1369118X.2016.1154086

Busuioc, M., Curtin, D., and Almada, M. (2023). Reclaiming transparency: contesting the logics of secrecy within the AI act. *Eur. Law Open* 2, 79–105. doi: 10.1017/elo.2022.47

Cohen, J. E. (2019). *Between truth and power: The legal constructions of informational capitalism*. Oxford, UK: Oxford University Press.

Couldry, N., and Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Redwood City, CA: Stanford University Press.

Crawford, K., and Joler, V. (2018). *Anatomy of an AI system*.

Delacroix, S., and Wagner, B. (2021). Constructing a mutually supportive interface between AI and human values. *Nat. Mach. Intell.* 3, 103–105. doi: 10.2139/ssrn.3404179

De Liban, K. (2024). *Inescapable AI: The ways AI decides how low-income people work, live, learn, and survive*. Techtonic Justice.

Dunne, A., and Raby, F. (2013). *Speculative everything: Design, fiction, and social dreaming*. Cambridge, MA: MIT Press.

Edwards, L., and Veale, M. (2018). Enslaving the algorithm. *IEEE Secur. Privacy* 16, 46–54. doi: 10.1109/MSP.2018.2701152

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, exclude, and punish the poor*. New York: St. Martin's Press.

European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending certain Union acts and repealing Commission Decision 2022/C 227/04. Official Journal of the European Union, L 2024/1689

Gillespie, T., Boczkowski, P. J., and Foot, K. A. (2014). *Media technologies*. Cambridge, MA: MIT Press.

Gillespie, T. (2014). "The relevance of algorithms" in *Media technologies: Essays on communication, materiality, and society*. eds. T. Gillespie, P. J. Boczkowski and K. A. Foot (Cambridge, MA: MIT Press), 167–194.

Inayatullah, S. (1998). Causal layered analysis: Poststructuralism as method. *Futures* 30, 815–829. doi: 10.1016/S0016-3287(98)00086-X

Kak, A., and West, S. M. (2023). *Landscape: Confronting tech power*. AI Now Institute.

Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., et al (2017). Accountable Algorithms. *University of Pennsylvania Law Review*, 165, 633–705.

Lindley, J., Akmal, H., and Coulton, P. (2020). Design research and object-oriented ontology. *Open Philos.* 3, 11–41. doi: 10.1515/opphil-2020-0002

Lindley, J., and Green, D. P. (2021). The ultimate measure of success for speculative design is to disappear completely. *Interact. Design Architect.* 51, 32–51. doi: 10.55612/s-5002-051-002

Malgieri, G., and Comandé, G. (2017). Why a right to legibility of automated decision-making exists in the general data protection regulation. *Int. Data Privacy Law* 7, 243–265. doi: 10.1093/idpl/ipx019

Mittelstadt, B., Allo, P., Taddeo, M., Wachter, S., and Floridi, L. (2016). The ethics of algorithms: mapping the debate. *Big Data Soc.* 3, 1–21. doi: 10.1177/2053951716679679

Mittelstadt, B. D. (2019). *Principles for regulating medical AI*. Life Sciences, Society and Policy, No. 15.

Morozov, E. (2013). *To save everything, click here: The folly of technological solutionism*. New York: Public Affairs.

NJCM v The Dutch State. (2020). The Hague District Court, ECLI:NL:RBDHA, No. 2020, pp. 1878.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. New York: Crown Publishing Group.

Oxford Law Blog. (2025). July 23. *Secrecy without oversight: How trade secrets could potentially undermine the AI Act's transparency mandate*. Oxford Law Blog

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Cambridge, MA: Harvard University Press.

Pschetz, L., Gianni, R., Tallyn, E., and Speed, C. (2017). *Bitbarista: exploring perceptions of data transactions in the internet of things*. Proceedings of the 2017 CHI conference on human factors in computing systems, pp. 2964–2975.

Regulation (EU). (2016). *679 (general data protection regulation) [2016] OJ L 119/1*. Regulation (EU).

Rudin, C. (2019). Stop explaining black box machine learning models for high-stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* 1, 206–215. doi: 10.1038/s42256-019-0048-x

Schellmann, H. (2024). The algorithm. Hachette books.; Accountable algorithms. University of Pennsylvania Law Review, No. 165, pp. 633–705.

Sicart, M. (2014). *Play matters*. Cambridge, MA: MIT Press.

Tallyn, E., Pschetz, L., Gianni, R., and Speed, C. (2018). Exploring machine autonomy and provenance data in coffee consumption: a field study of Bitbarista. *Proc. ACM Hum. Comput. Interact.* 2, 1–25. doi: 10.1145/3274439

UK Parliament, House of Commons Education Committee. (2020). The impact of COVID-19 on education and children's services. UK Parliament, House of Commons Education Committee.

Van Bekkum, M., and Borgesius, F. Z. (2021). Digital welfare fraud detection and the Dutch SyRI judgment. *Eur. J. Soc. Secur.* 23, 323–340. doi: 10.1177/13882627211031257

Van Bekkum, M., Zuiderveen Borgesius, F., and Heskes, T. (2025). AI, insurance, discrimination and unfair differentiation: an overview and research agenda. *Law Innov. Technol.* 17, 177–204. doi: 10.1080/17579961.2025.2469348

Van den Boom, F. (2020). Vehicle data controls, balancing interests under the trade secrets directive. *Int. J. Technol. Policy Law* 3:11.

Van den Boom, F. (2023). *The state of glitch, a speculative design provocation for inclusive AI futures*, in Morals & Machines Journal published by nomos publishing house.

Van den Boom, F. (2022). "Driven by digital innovations: regulating in-vehicle data access and use" in *Informational rights and informational wrongs: A Tapestry for Our Times*. eds. M. Borghi and R. Brownsword (Abingdon: Routledge).

Veale, M. (2020). A critical take on the policy recommendations of the EU high-level expert group on artificial intelligence. *Eur. J. Risk Regul.* 11, 1–14. doi: 10.1017/err.2019.65

Veale, M., Binns, R., and Edwards, L. (2021). Algorithms that remember: model inversion attacks and data protection law. *Phil. Trans. R. Soc. A* 379:83. doi: 10.1098/rsta.2018.0083

Wachter, S., Mittelstadt, B., and Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the general data protection regulation. *Int. Data Privacy Law* 7, 76–99. doi: 10.1093/idpl/ipx005

Wachter, S., Mittelstadt, B., and Russell, C. (2020). Why fairness cannot be automated: bridging the gap between EU non-discrimination law and AI. *Comput. Law Secur. Rev.* 36:105373. doi: 10.2139/ssrn.3547922

Whittaker, M. (2021). The steep cost of capture. *Interactions* 28, 50–55. doi: 10.1145/3488666

Winner, L. (1980). Do artifacts have politics? *Daedalus* 109, 121–136.

Zarsky, T. (2013). The trouble with algorithms. *Univ. California Hastings Law J.* 61, 57–116. doi: 10.1177/0162243915605575