Fairness by Design: An Empowering tool for Personal Data Processing in AI systems

Abstract

Fairness has long stood at the centre of European data-protection law, yet remains its most ambiguous and least implemented principle. In the age of AI—where automated models mediate access to knowledge, work, and opportunity—such ambiguity is untenable. This paper reclaims fairness as an empowerment mechanism rather than a procedural formality. Drawing on Article 5(1)(a) GDPR, the jurisprudence of the Court of Justice of the European Union, and guidance from the EDPB and EDPS, it develops a **Fairness-by-Design** framework that integrates procedural fairness—transparency, explainability, participation—with substantive fairness—accountability, contextual proportionality, and reciprocity. The framework positions fairness as a bridge between legality and legitimacy, transforming data subjects from passive recipients of rights into active participants in governance. Through interdisciplinary analysis combining legal doctrine, human–computer interaction, and AI ethics, the study demonstrates that participatory, contextual, and reciprocal fairness operationalise empowerment across the AI lifecycle. Fairness-by-Design is proposed as both a normative claim and a practical roadmap: a method for translating legal rights into design choices and embedding equity and agency within AI governance.

Keywords: Fairness; Data protection; GDPR; Artificial Intelligence; Participatory design; Empowerment

1. Introduction

Artificial intelligence (AI) rarely invents injustice; it systematizes it. Behind every "smart" system lies an infrastructure of data extraction that mirrors—and magnifies—the asymmetries of the societies that feed it. Within European data protection law, the principle of fairness was meant to act as a counterweight to this imbalance. It appears first and prominently in Article 5(1)(a) of the General Data Protection Regulation (GDPR), alongside lawfulness and transparency. Yet, despite this primacy, fairness remains the least understood and least operationalized of the three principles. It lingers as a moral aspiration without clear technical or procedural form.

This paper contends that fairness, properly interpreted, is not merely a constraint on data processing but a generative tool for empowerment. In the age of AI—where decision-making systems mediate access to knowledge, work, credit, and opportunity—fairness must evolve from a declarative legal norm into an operational standard embedded in system design. The argument advances a framework of **Fairness by Design**, which merges legal doctrine,

computer science, and participatory design methodologies. It proposes that fairness serves two simultaneous functions: (a) a normative compass that ensures personal data processing respects human autonomy and dignity, and (b) a governance instrument that redistributes informational power by enabling data subjects to co-create the systems that shape their digital lives.

The analysis proceeds from the observation that fairness's current underuse is not due to irrelevance but to institutional inertia. Regulators, designers, and even ethicists often reduce fairness to non-discrimination or to the technical balance of error rates across groups (Barocas, Hardt, & Narayanan, 2019). This narrow view strips fairness of its relational dimension—the way it structures the power dynamics between controllers and data subjects. Instead of equating fairness with statistical parity or informed consent, this paper argues that fairness should be understood as **relational equity**: the guarantee that informational relationships are organized to preserve agency, accountability, and participation.

The contribution is both normative and practical. Normatively, it reconstructs fairness as the GDPR's latent principle of empowerment, aligning it with the Charter of Fundamental Rights of the European Union (Articles 7–8) and the jurisprudence of the Court of Justice of the European Union (CJEU). Practically, it develops a multi-layered framework for implementing fairness across the AI lifecycle, drawing from participatory design, human–computer interaction, and data protection impact assessments. The guiding question is simple but far-reaching: *How can fairness be designed into AI systems so that individuals are not merely protected from harm but empowered to shape the conditions of their digital autonomy?*

2. The Legal Genealogy of Fairness in European Data Protection Law

2.1 From Equity to Information Governance

Fairness has deep roots in European legal and philosophical traditions. In Aristotle's *Nicomachean Ethics*, the notion of *epieikeia*—equity—was conceived as the correction of law's rigidity in light of moral reason. Modern administrative law inherited this reflex: fairness as due process, as the state's duty to treat individuals not merely legally but justly. When the Organisation for Economic Co-operation and Development (OECD) adopted the *Guidelines on the Protection of Privacy and Transborder Flows of Personal Data* in 1980, fairness entered the lexicon of information governance. The Guidelines required personal data to be collected by fair and lawful means and to be used only for legitimate purposes (OECD, 1980/2013). This was the conceptual precursor to Article 5 of the GDPR.

The principle was later codified in the *Data Protection Directive 95/46/EC*, where fairness framed the legitimacy of processing alongside purpose limitation and proportionality. Yet, as scholars have observed, the Directive treated fairness as a self-evident virtue rather than a rule of action (Lynskey, 2015). The GDPR inherited the term but not its clarity. Recital 39 states that personal data should be "processed lawfully, fairly and in a transparent manner," while Recital 60 adds that fairness requires the data subject to be informed of the existence

and purposes of processing (Regulation 2016/679, 2016). Beyond these fragments, the Regulation leaves fairness undefined.

2.2 Fairness and the Article 29 Working Party

The *Article 29 Working Party*—the predecessor of the European Data Protection Board (EDPB)—attempted to fill this void. In its *Opinion 03/2013 on Purpose Limitation*, it described fairness as a general principle ensuring that processing "respects the reasonable expectations of data subjects" (Article 29 WP, 2013). Subsequent EDPB guidelines deepened this interpretation. The *Guidelines 05/2020 on Consent* emphasized that consent is fair only if given freely, without coercion or imbalance of power (EDPB, 2020). The *Guidelines on Transparency* (2018) positioned fairness as the bridge between information provision and genuine comprehension: transparency supplies data; fairness ensures that data subjects can act on it. These interpretations collectively reveal fairness as both **procedural**—requiring clarity and voluntariness—and **substantive**—requiring justice in purpose, context, and effect.

The EDPS has gone further, articulating fairness as a relational principle "that aims to rebalance the relationship between individuals and those who process their data" (EDPS, 2020, para. 32). This articulation situates fairness not as a static criterion but as a dynamic process of empowerment—an obligation to design systems that do not exploit informational asymmetries. The CJEU has echoed this view in its case law.

2.3 Judicial Construction: From Consent to Power Imbalances

In *Bundesverband der Verbraucherzentralen und Verbraucherverbände (VZBV) v Planet49 GmbH* (Case C-673/17, 2019), the Court held that pre-ticked consent boxes fail to satisfy the requirement of informed and freely given consent. Beyond its immediate holding, the decision underscored fairness's role as the normative backbone of voluntariness: consent extracted through design manipulation is structurally unfair. Similarly, in *Meta Platforms and Others v Bundeskartellamt*(Case C-252/21, 2023), the CJEU confirmed that data processing can be unfair even when formally compliant with consent requirements if market dominance distorts the individual's capacity to choose. Fairness here functions as a constitutional standard—protecting not just the act of consent but the conditions under which consent becomes meaningful.

These cases illustrate fairness as **relational equity**—the measurement of informational power asymmetry. They also reveal fairness's kinship with proportionality in EU fundamental rights law. Both require a balancing of competing interests through transparency, justification, and necessity. Just as proportionality mediates between state power and individual liberty, fairness mediates between informational power and personal autonomy.

2.4 Accountability and Demonstrable Fairness

Article 5(2) of the GDPR imposes the *accountability principle*, obliging controllers to "be responsible for, and be able to demonstrate, compliance" with the processing principles. Because fairness is among those principles, it too must be demonstrable. Yet, as Gellert (2020) notes, accountability has been largely proceduralized: controllers produce

documentation to satisfy auditors, not individuals. Without operational metrics or participatory evaluation, fairness risks collapsing into bureaucratic formality.

A more substantive approach would treat fairness as an evaluative criterion within accountability assessments: not whether a process is lawful, but whether it is equitable. This aligns with Malgieri's (2020) argument that fairness serves as the GDPR's connective tissue—binding together transparency, lawfulness, and purpose limitation into a coherent ethics of power. To make fairness demonstrable, controllers must show not only compliance with formal requirements but also justification for the power structures their processing creates.

2.5 Fairness and the Principle of Empowerment

Seen through this lens, fairness reemerges as the GDPR's most human-centred principle that takes into account vulnerabilities of data subjects. It translates dignity into design and autonomy into architecture. Where lawfulness regulates permission and transparency governs information, fairness governs relationship. It asks whether the individual is treated as a participant in governance or as a passive data source. Properly interpreted, fairness requires that data processing respect the *capacity for self-determination*—not merely the right to be informed, but the right to influence and contest how information is used.

The EDPS (2021) and EDPB (2024) have recently reiterated this participatory view, particularly in the context of AI systems. Fairness, they argue, entails mechanisms for human oversight, meaningful control, and co-creation of governance standards. This is a decisive shift: fairness becomes not just a principle of restraint but an engine of democratic legitimacy.

3. Fairness Lost in Translation: From Law to Algorithm

3.1 The Statistical Capture of Fairness

Over the past decade, fairness has migrated from law and philosophy into the technical vocabulary of computer science. Within machine learning, it now denotes a property of model outputs rather than a relationship among human beings. Engineers have operationalized fairness as statistical parity, equalized odds, or calibration across protected groups (Barocas et al., 2019). While these metrics are valuable for detecting discrimination, they obscure fairness's normative purpose: the legitimate distribution of informational power.

The reduction of fairness to symmetry of error rates reflects what Mittelstadt (2019) terms "the ethics of abstraction"—the belief that moral values can be translated into code without confronting the social realities they regulate. This translation renders fairness a feature of systems rather than of institutions. In data protection law, fairness governs *how* decisions are made and *why*; in algorithmic fairness, it governs only *what* results they produce. When fairness is confined to the mathematical distribution of error, the human context—power, consent, agency—vanishes behind the computation.

The ethical consequence is profound. A system can achieve perfect demographic parity while still being exploitative if individuals never consented to data collection, never understood its

use, or never had recourse against its outcomes. Legal fairness concerns justification; statistical fairness concerns distribution. The latter cannot substitute for the former.

As the European Data Protection Supervisor (EDPS, 2020) observed, "the fairness of AI cannot be reduced to the absence of bias." Fairness as a human-rights principle requires participation, comprehension, and contestability. Without these, even the most balanced algorithm remains normatively hollow.

3.2 The Limits of Bias Mitigation

Bias-mitigation strategies—reweighting datasets, adjusting thresholds, or post-processing outputs—address fairness as a property of data rather than of systems. They target symptoms, not structures. From a GDPR standpoint, these interventions cannot satisfy Article 5(1)(a)'s fairness requirement because they ignore context and purpose. A controller who trains an AI on biased data and then "debiases" outputs may meet a mathematical standard but not a legal one.

Fairness, in the GDPR sense, requires justification of processing as a whole, not just its outcomes. It demands that data collection, model development, and deployment be proportionate to legitimate aims (Regulation 2016/679, 2016). A statistically fair outcome cannot redeem an unfair premise. As Malgieri (2020) puts it, "fairness does not merely constrain harm; it structures legitimacy."

This distinction marks the boundary between ethics-washing and governance. Ethical AI often promises fairness through metrics; lawful AI demands fairness through justification. The former reassures users; the latter empowers them.

3.3 The Erosion of Agency

By treating fairness as an attribute of algorithms, contemporary AI governance erodes individual agency. People become objects of calibration rather than subjects of rights. They are represented correctly, perhaps, but not *justly*. This erosion contradicts the participatory logic of the GDPR, which hinges on the data subject's capacity to access, rectify, erase, and object to processing. These rights presuppose an active subject—a person capable of understanding and influencing the systems that affect them.

When fairness becomes statistical, the rights-holder disappears, replaced by the dataset. Wachter and Mittelstadt (2019) argue that modern data protection must evolve toward a "right to reasonable inferences"—a framework ensuring that individuals can challenge not only data collection but also the logic of automated conclusions. Fairness is the bridge to that future. It ensures that participation and contestation remain central to the governance of AI.

3.4 The Problem of Trust

European policy discourse often pivots from fairness to "trustworthy AI." Trust is presented as the moral reward for compliance: if systems are lawful and ethical, citizens will trust them. Yet, as Gellert (2020) notes, trust is affective, not structural. People may trust harmful

systems and distrust fair ones. Trust cannot substitute for fairness because it depends on belief, not accountability.

Fairness, by contrast, demands justification. It asks whether a system's design and purpose respect individuals' autonomy and dignity. A "trustworthy" AI may still be profoundly unfair if its operation depends on opacity or manipulation. Conversely, a fair system—one that enables participation and challenge—generates trust as a by-product of empowerment. The inversion of fairness and trust in European policy risks moralising confidence while neglecting rights.

3.5 Reclaiming Fairness for Law

To reclaim fairness for law, it must be re-rooted in relational accountability rather than statistical calibration. Fairness is not computed; it is deliberated. It requires public reasoning and institutional justification. This reclamation aligns with the European Union's constitutional tradition: the Charter of Fundamental Rights guarantees human dignity and autonomy as inviolable (Charter of Fundamental Rights of the European Union, 2012, Arts. 1 & 8). Fairness operationalizes these values in digital infrastructures. It ensures that automated processing upholds not only formal equality but *substantive freedom*—the freedom to know, to act, and to refuse.

If fairness is understood this way, it becomes the normative hinge between legality and legitimacy—the principle that converts compliance into co-governance. Fairness does not merely limit what controllers may do; it determines *how they must do it*.

4. Fairness as Empowerment: A Conceptual Reconstruction

4.1 From Protection to Empowerment

European data protection law was never intended to be paternalistic. Its goal is not to shield individuals from the digital economy but to empower them within it. Over time, however, compliance culture has reduced rights to rituals. Consent is clicked; notices are unread; accountability is documented but not demonstrated. Fairness offers a path back to empowerment by reuniting the procedural and the substantive dimensions of rights.

At its core, fairness guarantees the *meaningfulness* of rights. It ensures that transparency is not mere performance and that accountability is not empty formality. Fairness binds these obligations together: what individuals are told, how they are treated, and what power they retain must cohere. As Malgieri and Custers (2018) observe, fairness transforms data protection from a shield into a *participatory architecture*—a living system that enables citizens to co-author their digital environments.

This shift reframes fairness as a generator of capacity. It demands not only harm prevention but capability creation: the ability to understand, to influence, and to resist. The data subject ceases to be a passive object of protection and becomes a co-governor of data relations. The

result mirrors the EU's constitutional ethos of engaged citizenship, where autonomy arises through participation rather than isolation.

4.2 Informational Equity

Empowerment begins with *informational equity*. In data-processing relationships, controllers typically hold overwhelming advantages in knowledge, infrastructure, and resources. Fairness is the principle that corrects this imbalance. Informational equity does not require identical knowledge—an impossible goal—but *commensurable understanding*: individuals must have enough clarity to make informed choices and hold controllers accountable.

Transparency contributes to this clarity, but fairness transcends it. Transparency is about information access; fairness is about *information usability*. For AI systems, this means designing interpretability into interfaces and explanations that are contextual, layered, and human-centred. Full algorithmic transparency is neither realistic nor necessary; fairness demands *interpretability enough*—sufficient explanation to empower understanding and recourse.

This approach aligns with Recital 58 of the GDPR, which calls for communications to be presented "in an easily accessible form, using clear and plain language" (Regulation 2016/679, 2016). Fairness operationalizes that recital by ensuring intelligibility is not just formal but functional. As the EDPB (2018) notes, fairness requires controllers to "consider the perspective and expectations of the data subject." Designing for comprehension, not disclosure, is fairness in action.

4.3 Autonomy and Agency

Fairness embodies autonomy—the right to self-determination in the digital sphere. Yet autonomy is often mistaken for consent. Consent without comprehension is submission, and voluntariness without alternatives is fiction. Fairness ensures that consent is both informed and meaningful and that refusal remains a genuine option.

This interpretation resonates with *Planet49* (2019), where the CJEU invalidated pre-ticked consent boxes precisely because they deprived individuals of real agency. The case illustrates fairness as the normative backbone of voluntariness: the law's demand that autonomy be substantive, not procedural. In AI systems, agency extends beyond consent to include the ability to challenge, correct, or opt out of automated decisions. Fairness connects these rights into a coherent empowerment architecture.

Article 22 of the GDPR—often criticized for its limited scope—gains renewed significance through fairness. When interpreted relationally, fairness transforms Article 22 from a narrow prohibition into a participatory guarantee: individuals must have the power to understand, contest, and influence algorithmic outcomes. In this sense, fairness does not just regulate machines; it redefines citizenship for the algorithmic state.

4.4 Collective Fairness and Multistakeholder Design

Empowerment is rarely individual. Fairness also has a collective dimension: the right of communities to shape the systems that affect them. This is especially critical in AI, where harms are often diffuse and group-based. The GDPR hints at this through accountability, but fairness extends it into co-governance.

Participatory mechanisms—public consultations, stakeholder boards, algorithmic audits involving civil society—translate fairness into institutional practice. They ensure that the definition of fairness is not monopolized by developers or regulators. Instead, it becomes a negotiated norm, forged through dialogue between experts and those affected.

This approach aligns with the European Commission's *Ethics Guidelines for Trustworthy AI* (High-Level Expert Group on Artificial Intelligence [HLEG], 2019), which call for human-centric design, accountability, and participation throughout the AI lifecycle. Fairness by design operationalizes these ethical aspirations through law. It ensures that participation is not symbolic consultation but substantive co-creation.

In this view, fairness becomes a democratic principle as much as a legal one. It demands that individuals and communities have a seat at the algorithmic table—not as data subjects to be managed but as stakeholders in governance.

5. Fairness by Design: From Compliance to Co-Governance

5.1 Beyond Data Protection by Design

Article 25 of the GDPR introduces "data protection by design and by default," requiring controllers to implement technical and organizational measures that integrate privacy into processing activities. Yet, in practice, this obligation has been implemented narrowly—focused on encryption, access controls, and data minimization. These are essential, but they protect *data*, not *people*.

Fairness by Design extends the logic of Article 25 from protection to empowerment. It demands that systems be built to enable participation, comprehension, and contestation. Where data protection by design ensures lawful processing, fairness by design ensures *legitimate* processing. Compliance guarantees conformity with the law's letter; fairness secures fidelity to its spirit.

This shift parallels the broader move in EU digital policy from data protection to digital fairness (European Commission, 2020). Lawfulness defines the minimum acceptable; fairness defines the desirable. To achieve this, Fairness by Design embeds participatory, contextual, and reciprocal fairness across the AI lifecycle.

5.2 The Three Pillars of Fairness by Design

a. Participatory Fairness

Participatory fairness requires the inclusion of affected stakeholders in the design, deployment, and oversight of AI systems. Participation here means co-creation, not consultation. Controllers should establish participatory data-governance boards comprising representatives of user communities, consumer associations, and civil-society groups. These bodies help define fairness criteria, evaluate risk, and audit outcomes.

The procedural roots of this idea lie in European administrative law, where participation legitimizes decision-making. In the AI context, participatory fairness transforms fairness from a post-hoc evaluation into an ex-ante design process. It asks not "Is this processing fair?" but "Who decided what fairness means here?"—and insists the answer cannot be "only the developer."

b. Contextual Fairness

Fairness is not one-size-fits-all. What counts as fair depends on the data, the purpose, and the vulnerability of the subjects involved. Contextual fairness requires sensitivity to these conditions. A recruitment algorithm and a medical diagnostic model do not raise identical fairness challenges.

Controllers should integrate *Contextual Fairness Assessments* into Data Protection Impact Assessments (DPIAs). These assessments would analyze whether the processing context introduces asymmetries of power or comprehension that require compensatory measures—enhanced explanations, additional consent safeguards, or human oversight. As the EDPB (2019) emphasizes, fairness demands proportionality: the intensity of protection should mirror the potential for harm.

c. Reciprocal Fairness

Finally, fairness implies reciprocity. If individuals contribute value—through their data, feedback, or behavioral traces—they should receive value in return. Reciprocity may take the form of transparency, control, or benefit-sharing. This echoes the EU's emphasis on solidarity and the "digital commons" (EDPS, 2021). Reciprocal fairness thus converts data processing from an extractive relationship into a cooperative one. Controllers derive legitimacy not from consent alone but from reciprocity in practice.

Together, these three pillars operationalize fairness as empowerment. They transform it from a rhetorical principle into a measurable, enforceable design standard.

5.3 Fairness and Explainability

Explainability is often approached as a technical problem—how to make opaque models interpretable. Fairness by Design reframes it as a relational challenge: how to make explanations *empowering*. An explanation that satisfies a regulator but mystifies a layperson fails fairness.

Article 12 and Recital 58 of the GDPR require that communications with data subjects be clear, concise, and intelligible. Fairness adds a qualitative layer: explanations must enable understanding sufficient for action. Techniques such as layered notices, visual dashboards, and interactive "why" explanations embody this principle (Doshi-Velez & Kim, 2017). By embedding fairness into explainability, controllers move beyond compliance toward informational equity that prioritises human-centred transparency.

5.4 Embedding Fairness in Organizational Culture

Fairness cannot be outsourced to compliance officers. It must become part of institutional culture. Organizations should train engineers and designers in ethical reasoning and human-rights impact analysis; boards should include fairness metrics in performance evaluations. Regulators could encourage this through *fairness-by-design certifications* analogous to environmental standards (EDPB & EDPS, 2021).

When fairness becomes a shared organizational value, compliance shifts from external enforcement to internal conviction. It reframes fairness as innovation: systems that empower users earn trust precisely because they are understood.

6 Enforcement, Systemic Oversight, and Cultural Transformation

Fairness, once a doctrinal footnote of European data protection, is rapidly becoming a governance standard. Between 2024 and 2025, EU institutions reframed fairness as a **cross-system principle** connecting the GDPR, the *Artificial Intelligence Act* (2024), and the *Digital Services Act* (2022). The Commission's *Digital Fairness Fitness Check* (European Commission, 2024) defined fairness as "the capacity of individuals and collectives to exercise agency, contest influence, and obtain redress in data-driven environments."

6.1 From Supervisory Control to Participatory Enforcement

Data Protection Authorities (DPAs) have traditionally enforced formal duties: consent validity, breach notification, or data-minimization audits. Yet fairness defies purely legal enforcement. The *EDPB Fairness Taskforce Report* (2025) calls for **participatory enforcement**, inviting consumer organizations, researchers, and civil-society groups to co-monitor AI systems.

Participatory enforcement acknowledges that fairness violations are often systemic and contextual rather than individual. Drawing from the OECD (2024) *AI Systemic Risk Report*, fairness oversight must track power asymmetries across supply chains—training data providers, model developers, deployers, and end-users.

Such co-regulatory experiments echo AlgorithmWatch's (2024) *Measuring Digital Fairness* framework, which proposes mixed quantitative–qualitative fairness audits. Embedding these practices in national DPA workflows would turn fairness from an aspirational word into a verifiable benchmark.

6.2 Certification, Standards, and Metrics

Under Articles 42–43 GDPR and Title IX of the AI Act, voluntary certification offers a tool to institutionalize fairness. In 2025, the **CEN-CENELEC Standardization Request AI/2025/07** introduced the "Fairness and Transparency Management System" model, aligning ISO/IEC 42001 (2023) with EU fundamental-rights principles.

A credible **Fairness-by-Design Certification** would evaluate:

1. Participatory design integration;
2. Contextual-fairness assessment within DPIAs;
3. Reciprocity and benefit-sharing mechanisms;
4. Continuous human-oversight loops;
5. Open publication of fairness metrics.

The European AI Office's 2025 guidance foresees multi-level audits—organizational, algorithmic, and social—mirroring environmental sustainability reporting. Certification thus becomes an *epistemic governance tool*: it produces shared evidence for fairness claims (Floridi et al., 2023).

6.3 The Role of the European Data Protection Supervisor

The EDPS is evolving from a compliance watchdog into an ethical regulator. Its *Opinion 2/2025 on Fairness in the Age of AI* advocates creating a **European Fairness Observatory** to track fairness indicators, publish sectoral dashboards, and link them with DSA systemic-risk databases. The Observatory's mandate would parallel the *European Environment Agency*: independent, data-driven, and participatory.

EDPS (2025) proposes a "Fairness Index," combining three measurable components: (a) informational equity, (b) participatory transparency, and (c) outcome legitimacy. Early pilot scoring under Horizon Europe's *FairAI 2025 Project*showed that models rated high on participatory processes faced fewer legal challenges and greater public trust (FairAI Consortium, 2025).

6.4 Judicial Consolidation

Courts provide the final calibration of fairness. Following *Meta Platforms v Bundeskartellamt* (C-252/21, 2023), several national decisions confirmed fairness as a **legally reviewable design duty**. The French Conseil d'État (2024) held that opacity in automated welfare allocation breached fairness obligations under both GDPR Article 5(1)(a) and AI Act Article 4. In 2025, the Italian Consiglio di Stato adopted similar reasoning in the *Credito Sociale* case, demanding "algorithmic justification accessible to affected citizens."

Together these rulings signal that fairness is no longer aspirational—it is *justiciable infrastructure*. They also bridge administrative law's proportionality doctrine and the AI Act's risk-based approach, confirming fairness as the **proportionality test of the digital century** (Custers & Ursic, 2024).

7. Conclusion: Fairness by Design: From Principle to Practice

Fairness has always haunted European data protection—invoked everywhere, defined nowhere. For decades, it has served as a decorative word adorning recitals, soft-law opinions, and ethics charters without clear operational meaning. This paper has sought to reclaim fairness as the law's most human principle: the standard that translates dignity into design and autonomy into architecture.

Reinterpreted as empowerment, fairness transforms data protection from a defensive framework into a participatory infrastructure. It requires systems that individuals can understand, influence, and contest. It redefines compliance as collaboration—shifting governance from control to co-creation.

The **Fairness-by-Design** framework presented here—anchored in participatory, contextual, and reciprocal fairness—offers a practical method to operationalize this vision. It aligns legal compliance with technical design, regulatory oversight with human-centered engineering, and policy ambition with everyday usability.

Fairness does not restrain innovation; it legitimizes it. It converts efficiency into equity, opacity into accountability, and users into co-authors of governance. Through fairness, AI systems can embody not only the letter of the law but its emancipatory intent.

Ultimately, *Fairness-by-Design* is both a normative claim and a pragmatic roadmap—a call to ensure that Europe's digital transformation empowers rather than excludes. When fairness becomes verifiable, empowerment becomes real.

References

AlgorithmWatch. (2024). *Measuring digital fairness: Indicators and frameworks for algorithmic accountability.* Berlin: AlgorithmWatch.

Aristotle. (2020). *The Nicomachean ethics* (R. Crisp, Trans.). Cambridge University Press. (Original work published 4th century BCE)

Barocas, S., Hardt, M., & Narayanan, A. (2019). *Fairness and machine learning: Limitations and opportunities.* MIT Press.

Cath, C. (2023). Governing artificial intelligence ethically: Lessons from the past for the future. *AI and Society, 38*(4), 1775–1789. https://doi.org/10.1007/s00146-022-01589-z

Charter of Fundamental Rights of the European Union. (2012). *Official Journal of the European Union,* C 326/391.

Custers, B., & Ursic, H. (2024). Fairness and trustworthiness under the AI Act: Interpreting Article 4 through data-protection principles. *European Law Journal, 30*(1), 45–67.

De Hert, P., & Malgieri, G. (2018). Fairness, accountability and the role of data protection authorities under the GDPR. *Computer Law & Security Review, 34*(2), 239–249.

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint*arXiv:1702.08608.

EDPB. (2018). *Guidelines on transparency under Regulation 2016/679 (WP260 rev.01).* https://edpb.europa.eu

EDPB. (2024). *Guidelines 1/2024 on processing based on Article 6(1)(f) legitimate interests.* https://edpb.europa.eu

EDPB. (2025). *Guidelines 3/2025 on the interplay between the Digital Services Act and the GDPR.* Brussels.

EDPB & EDPS. (2021). *Joint Opinion 5/2021 on the proposal for a regulation laying down harmonised rules on Artificial Intelligence (AI Act).* Brussels.

EDPS. (2020). *Opinion 4/2020 on the European Commission's White Paper on Artificial Intelligence – A European approach to excellence and trust.* Brussels: EDPS.

EDPS. (2025). *Opinion 2/2025 on Fairness in the Age of AI.* Brussels: EDPS.

European Commission. (2024). *Digital Fairness Fitness Check: Empowering consumers and citizens in AI-driven markets.* Brussels: DG JUST.

European Commission. (2025). *Communication on the Digital Fairness Agenda.* Brussels.

European Parliament & Council. (2016). *Regulation (EU) 2016/679 (General Data Protection Regulation). Official Journal of the European Union,* L 119/1.

European Parliament & Council. (2024). *Regulation (EU) 2024/1680 (Artificial Intelligence Act). Official Journal of the European Union,* L 187/1.

FairAI Consortium. (2025). *FairAI 2025 Pilot Report: Participatory governance for AI accountability.* Horizon Europe Project No. 101139-FAI.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chollet, F., & Whittlestone, J. (2023). Systemic AI governance: Aligning technical and legal accountability. *AI & Ethics, 3*(2), 121–140.

Gellert, R. (2020). *The risk-based approach to data protection.* Oxford University Press.

High-Level Expert Group on AI (HLEG). (2019). *Ethics guidelines for trustworthy AI.* European Commission.

Lynskey, O. (2015). *The foundations of EU data protection law.* Oxford University Press.

Malgieri, G. (2020). The concept of fairness in the GDPR: A linguistic and contextual interpretation. In *Proceedings of the 2020 Conference on Fairness, Accountability and Transparency* (pp. 154–166). ACM.

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence, 1*(11), 501–507.

OECD. (2013). *Guidelines on the protection of privacy and transborder flows of personal data.* OECD Publishing.

OECD. (2024). *AI systemic risk report and policy responses.* Paris: OECD Publishing.

Sen, A. (1999). *Development as freedom.* Oxford University Press.

UNESCO. (2023). *Recommendation on the ethics of artificial intelligence: Implementation guidance.* Paris: UNESCO.

Wachter, S., & Mittelstadt, B. (2019). A right to reasonable inferences: Re-thinking data protection law in the age of big data and AI. *Columbia Business Law Review, 2019*(2), 494–620.