Shattering the Illusion of Formal Compliance: Digital Consent as an Explainable Socio-Technical Process

Rachele Carli[1][0000−00020−8689−285X] and Davide Calvaresi[2][0000−0001−9816−7439]

[1] AI Policy Lab and Responsible AI Group, Umeå University
rachele.carli@umu.se
[2] University of Applied Sciences Western Switzerland, Switzerland
davide.calvaresi@hevs.ch

Abstract. **Abstract.** The increasing adoption of AI-driven systems in domains such as healthcare, personalized assistance, and behavioural monitoring challenges the foundational assumptions underlying informed consent. Existing digital consent mechanisms provide traceability and regulatory compliance but often fail to ensure meaningful user understanding, reducing consent to a procedural artifact rather than a genuine expression of autonomy. In this paper, we argue that informed consent in AI-mediated environments should be reconceptualized as an explainability and governance problem. We propose an expanded view of Explainable Artificial Intelligence (XAI) that extends beyond model interpretability to support user understanding of data practices, automated inference, and system functionalities throughout the consent lifecycle. Building on legal analysis, cognitive science, and prior work on dynamic and agent-based consent management, we introduce a conceptual framework for explainable and verifiable consent. The framework is operationalized through five interdependent design principles: (i) modularization of consent content, (ii) layered and adaptive explanations, (iii) interactive verification of understanding, (iv) granular functionality control, and (v) temporal validation of consent. We further discuss the role of explainability as an accountability mechanism and critically examine risks associated with persuasive explanations, automated comprehension assessment, accessibility barriers, and functional coercion. By reframing consent as a dynamic socio-technical process embedded within intelligent systems, this work contributes to ongoing research on explainable AI, multi-agent systems, and AI governance, and outlines a research agenda for trustworthy and autonomy-preserving AI services.

Keywords: Explainable AI · Informed Consent · AI Governance · Dynamic Consent · Algorithmic Accountability

1 Introduction

The rapid integration of AI into everyday services has fundamentally reshaped the conditions under which individuals disclose personal data and consent to

its processing [34, 37]. In domains such as healthcare, personalised assistance, and behavioural monitoring, AI-driven systems increasingly mediate interactions that were traditionally governed by direct human relationships [16, 33, 54]. Virtual assistants, conversational agents, and predictive analytics tools now collect, infer, and act upon sensitive information, often in ways that are opaque even to technically literate users [28].

In this context, informed consent — which has been long regarded as a cornerstone of autonomy [29] and a mechanism to rebalance informational asymmetries [44] — faces structural stress that undermines both its epistemic and legitimising functions.

Historically, informed consent has served a dual function. First, it protects weaker parties in contractual or service relationships by compensating for asymmetries in expertise and power. Second, it plays a central role in allocating responsibility and liability, particularly when harm arises from data misuse, system malfunction, or unintended consequences. To be valid, consent must be personal, freely given, explicit, specific, informed, current, and revocable [24]. These requirements are not merely formalistic; they operationalize respect for autonomy and ensure that individuals retain meaningful control over their information and participation.

However, the proliferation of AI systems introduces new risks that strain these foundational assumptions. Modern AI systems can extract insights that go far beyond the data explicitly provided by users. Through inference, pattern recognition, and integration with third-party datasets, systems may derive sensitive attributes — such as health conditions, behavioural tendencies, or socioeconomic characteristics — that users neither directly disclosed nor anticipated. Moreover, complex data ecosystems enable secondary uses and third-party access that are difficult for individuals to track or understand [1]. The result is a widening gap between what users believe they are consenting to and what systems are technically capable of doing [45].

This gap is not only a consequence of technological sophistication. It is also rooted in the structure of consent practices themselves. Consent forms are frequently lengthy, written in technical or legal language, and embedded within interfaces that incentivise rapid acceptance. Empirical research has repeatedly shown that users rarely read privacy policies in full and often lack understanding of key elements such as data retention periods, sharing practices, and automated decision-making processes. Cognitive biases — such as present bias, optimism bias, and difficulty anticipating long-term consequences — further undermine reflective decision-making. In practice, consent often becomes a procedural ritual rather than a substantive exercise of autonomy.

Digitization has not resolved these issues. While digital consent systems offer advantages in traceability, storage efficiency, and accessibility, they frequently replicate the structural flaws of paper-based models. Checkboxes and scrollable documents, even when accompanied by layered interfaces, do not guarantee comprehension. In some cases, interface design choices — intentionally or unintentionally — nudge users toward acceptance, raising concerns about manipulation

and dark patterns [13]. Thus, despite formal compliance with regulatory requirements, many digital consent implementations fall short of achieving genuinely informed agreement.

This paper argues that the shortcomings of contemporary consent practices call for a reconceptualization of consent as an interactive, explainable process rather than a static declaration. Specifically, we propose that Explainable AI (XAI) can serve as an enabling layer that operationalizes the "informed" dimension of consent in AI-mediated environments. Rather than limiting explainability to post-hoc model interpretation, we intend to extend its scope to explain data practices, risks, rights, and the system functionalities embedded in consent workflows.

Building on prior work on dynamic and modular consent management within agent-based architectures [61, 6, 48], we outline a framework in which consent is (i) segmented into coherent thematic units, (ii) accompanied by layered explanations and examples, (iii) verified through interactive comprehension checks, and (iv) revisited over time through periodic understanding assessments. In this approach, explainability becomes a structural component of consent management, supporting both user understanding and institutional accountability.

The remainder of this paper proceeds as follows. Section 2 examines the legal, cognitive, and practical foundations of informed consent and identifies the structural causes of its current shortcomings. These foundations underscore the need for a redesigned, XAI-supported consent process that aligns regulatory ideals with real-world user experience. Section 3 reviews existing digital consent systems and dynamic consent approaches, highlighting their potential benefits as well as their limitations in addressing user understanding and interaction in complex AI-driven environments. Section 5 introduces a set of design principles for explainable and verifiable consent architectures, outlining how modular consent structures, adaptive explanations, and comprehension verification mechanisms can be integrated into system design. Section 6 discusses the broader implications of this approach, examining potential risks, ethical concerns, and design trade-offs associated with AI-mediated consent processes. Finally, Section 7 summarises the main contributions of this analysis, also highlighting areas where further improvement is still needed, and Section 8 concludes the paper.

2 Informed Consent: Legal, Cognitive, and Practical Foundations

Informed consent has long been considered a cornerstone of ethical and legal frameworks governing medical practice, research participation, and, more recently, the processing of personal data in digital environments. At its core, the concept aims to ensure that individuals retain meaningful control over decisions that affect their bodies, their information, and their participation in institutional or technological systems [36, 18, 29].

Despite its normative importance, the practical implementation of informed consent has proven to be considerably more complex than its theoretical formu-

lation suggests [65, 5]. In many contexts, the conditions required for a decision to be genuinely informed — adequate disclosure, comprehension, and voluntary agreement — are difficult to achieve in practice. Empirical evidence across medical, research, and digital settings consistently reveals significant gaps between the formal structure of consent procedures and the actual level of understanding attained by individuals.

To better understand the origins of these limitations, it is necessary to examine the foundations of informed consent across three complementary dimensions: its legal and ethical grounding, the cognitive conditions required for meaningful understanding, and the practical challenges encountered in real-world consent processes.

2.1 Legal and Ethical Foundations of Informed Consent

Informed consent is deeply embedded in both bioethical theory and data protection law. At its core lies the principle of autonomy: individuals must be able to make decisions about their bodies, participation, and personal data based on adequate information and without coercion. In regulatory frameworks, such as the General Data Protection Regulation (GDPR), consent must be freely given, specific, informed, and unambiguous [24]. It must also be revocable at any time, and the withdrawal of consent must be as easy as its provision [24].

Beyond autonomy, consent plays a crucial role in structuring liability and accountability [58]. When individuals consent to a particular data processing activity under clearly defined conditions, responsibility for subsequent harms can be evaluated against the scope and quality of that consent. If consent was uninformed or obtained under misleading circumstances, its legal validity can be compromised, potentially exposing service providers to liability. Thus, consent is not only a user-facing safeguard but also a cornerstone of institutional legitimacy.

However, the emergence of AI-based systems challenges the traditional assumptions that underlie consent. First, AI systems frequently rely on probabilistic models and continuous learning processes that evolve over time. This dynamism complicates the notion of "specific" and "current" consent. Second, the capacity of AI to generate inferences from seemingly innocuous data undermines the intuitive boundary between disclosed and undisclosed information. Users may consent to the collection of basic interaction data without realizing that such data can reveal sensitive attributes. Third, multi-actor data ecosystems — comprising developers, cloud providers, analytics partners, and third-party service providers — blur responsibility chains and complicate the communication of risks.

These structural features amplify the information asymmetry that consent is meant to mitigate. Although the law prescribes clarity and comprehensibility, the technical and organizational complexity of AI systems makes full transparency difficult to achieve through conventional textual disclosures alone. This leads to a *symbolic compliance* that is far from the rational underlying the creation of this legal tool.

2.2 Structural and Cognitive Barriers to Understanding

The persistent gap between the normative ideal of informed consent and its practical implementation is not merely a matter of insufficient disclosure. It reflects deeper structural and cognitive barriers.

From a structural perspective, consent documents often attempt to reconcile competing objectives: legal completeness, risk mitigation, and communicative clarity. To ensure compliance and reduce liability exposure, organizations frequently include extensive technical and legal detail. Yet, the very features that enhance formal robustness — precision, exhaustiveness, and conditional specificity — tend to reduce readability and accessibility [30]. The resulting documents are long, densely written, and cognitively demanding [7].

For non-specialized users, this creates a significant barrier. Most individuals lack the legal or technical expertise required to interpret data processing descriptions, algorithmic logic, or contractual clauses [25, 49]. Even highly educated users may struggle with domain-specific terminology or abstract risk formulations. Consequently, users often face the choice of investing disproportionate time and cognitive effort or accepting terms with limited (or even no) understanding.

Cognitive science further illuminates why even well-designed disclosures may fail. Individuals are not perfectly rational decision-makers [31]. Present bias leads users to prioritize immediate access to services over abstract, future risks. Consequence bias and optimism bias may cause individuals to underestimate the likelihood or severity of potential harms. Information overload — resulting from excessive detail — can impair comprehension and discourage engagement altogether [2]. In such conditions, consent risks becoming a default action rather than a deliberative choice.

Importantly, these shortcomings should not be interpreted as evidence that consent is obsolete or conceptually flawed. Rather, they indicate that traditional implementation strategies are ill-suited to the socio-technical complexity of AI systems. Eliminating consent as an institution would undermine autonomy and accountability structures. The challenge, therefore, is not whether to retain consent, but how to redesign its administration in ways that are transparent, fair, and cognitively sustainable.

This paper argues that consent must evolve from a one-time, text-based declaration into a structured, interactive, and verifiable process. Such a transformation requires methodological innovation at the intersection of law, human–computer interaction, and explainable AI. In the following sections, we build on this foundation to explore how XAI techniques and agent-based system architectures can support a user-centered and legally robust reconfiguration of informed consent in AI-driven environments.

3 Digital Consent Systems: Achievements and Structural Limits

The digitization of services has transformed the mechanisms for requesting, recording, and managing consent. On online platforms, in mobile applications, and in AI-powered assistive systems, consent is no longer limited to signed paper forms or in-person interactions. Instead, it is embedded within digital interfaces, automated workflows, and dynamic data infrastructures [52]. This transition has generated important practical and regulatory advantages. At the same time, it has revealed structural weaknesses that mirror — and in some respects amplify — the limitations of traditional consent models.

3.1 The Promise of Digital Consent

Digital consent systems offer several clear benefits compared to paper-based processes [32].

First, they significantly reduce administrative burden. Electronic forms can be generated, updated, stored, and retrieved with minimal cost and delay. This efficiency is particularly valuable in large-scale systems involving thousands or millions of users, where manual management would be impractical [11, 32].

Second, digital infrastructures enhance traceability and auditability. Every interaction — whether consent provision, modification, or withdrawal — can be marked with date and time and recorded. This creates a detailed record of the consent lifecycle, which can be essential for demonstrating regulatory compliance and resolving disputes [60, 26]. In contexts such as healthcare or clinical research, such traceability strengthens institutional accountability and facilitates oversight [60, 26].

Third, digital systems improve accessibility and revocability. Users can access consent forms remotely, review previously accepted terms, and modify their preferences at any time. This aligns with regulatory requirements that consent be current and revocable and supports the idea of consent as an ongoing rather than static condition [9].

These advantages have motivated the development of dynamic consent management systems that enable granular selection of data uses and continuous adjustment of preferences. Such systems often provide dashboards that allow users to view categories of collected data, specify permitted uses, and withdraw consent for specific purposes. In principle, this granular and adaptive structure better reflects the evolving nature of AI-driven services.

However, while digitization enhances procedural efficiency and formal compliance, it does not automatically guarantee substantive understanding.

3.2 The Illusion of Informed Digital Consent

Despite their technical sophistication, many digital consent implementations reproduce the fundamental weaknesses of their analogue predecessors. Consent

interfaces frequently rely on long, scrollable text blocks accompanied by binary choices [51, 46]. Although these forms are delivered electronically, their communicative structure remains largely unchanged: a monolithic disclosure followed by a single decision point [57].

Empirical studies consistently show that users rarely read or understand such documents in full [39, 1]. In this sense, digital consent often meets the disclosure requirement without ensuring awareness [58, 46]. Said otherwise, it facilitates a form of compliance to regulation that is symbolic rather than concrete.

Moreover, interface design can subtly influence user behavior. Visual hierarchies, button placement, default settings, and time pressure may lead users to accept. Although some design strategies aim to simplify navigation and reduce friction, others risk crossing the line into manipulative practices, commonly known as dark patterns [38]. When acceptance is easier or more visually prominent than refusal, the "freely given" nature of consent may be compromised [42].

Another limitation concerns the static presentation of dynamic systems. AI-based services often evolve over time, incorporating new functionalities, data sources, or analytic capabilities. Yet, consent is frequently obtained at a single point in time, with subsequent changes communicated through generic updates or revised policies. Users may not revisit these documents, and even when they do, distinguishing substantive changes from minor edits can be difficult. As a result, the temporal dimension of consent — its requirement to remain specific and current — becomes difficult to uphold in practice.

3.3 Dynamic Consent and Its Remaining Gaps

Dynamic consent models represent a significant conceptual improvement over one-time agreements. By enabling granular choices and ongoing modifications, they recognize that user preferences and system functionalities evolve [32, 9]. In domains such as biomedical research, dynamic consent platforms have demonstrated how digital tools can facilitate continuous communication between participants and institutions [60].

Nevertheless, dynamic consent systems still tend to rely predominantly on textual explanations and user self-assessment of understanding. While they offer more control, they rarely incorporate mechanisms to verify whether users actually understand the implications of their choices. The assumption remains that exposure to information suffices for informed decision-making [39].

In addition, the growing complexity of AI-driven systems introduces a qualitative shift. Users are not merely consenting to predefined data use but to probabilistic modeling, automated inference, and potentially opaque decision-making processes [63]. Explaining such processes through static text — even when modularized — may be insufficient [10]. What is required is not only more granular control, but also adaptive explanation mechanisms that can translate technical and legal constructs into accessible, context-sensitive representations [59].

This observation reveals a central tension. Digital consent systems excel at documenting choices, but they remain limited in their ability to support understanding. They operationalize consent as a compliance artifact — something that

can be stored, retrieved, and audited — rather than as an epistemic condition grounded in user comprehension [57, 46].

To address this limitation, consent mechanisms must move beyond passive disclosure toward interactive explanation and verification. In particular, AI-mediated systems call for tools that can dynamically adapt explanations, respond to user queries, highlight essential risks, and assess comprehension over time [40, 22]. In the next section, we argue that Explainable AI, when reconceptualized as a user-centred explanatory layer rather than solely a model-interpretation technique, offers a promising foundation for such a transformation.

4 Explainable AI and Consent

The limitations of digital consent systems outlined in the previous section reveal a structural problem: current implementations treat consent primarily as a compliance artifact rather than as an epistemic condition grounded in user understanding. If informed consent requires that individuals meaningfully comprehend the implications of data processing and automated decision-making, then consent in AI-mediated systems becomes, at its core, an explaination challenge.

This section argues that XAI, when reconceptualised beyond its traditional scope, provides a principled foundation for redesigning consent processes. Rather than limiting explainability to model interpretability, we position it as a structural component of interactive governance in intelligent multi-agent systems.

4.1 Classical XAI: Scope and Limitations

Explainable AI has traditionally focused on making the internal logic of machine learning models more transparent. Techniques such as feature attribution, surrogate models, counterfactual explanations, and rule extraction aim to clarify how specific outputs are produced [27, 63]. In high-stakes domains — such as healthcare, finance, and public administration — these methods support debugging, auditing, and fairness evaluation [20, 50].

However, classical XAI operates primarily at the level of model behaviour. It addresses questions such as: "Why was this prediction made?" or "Which features influenced this decision?". While these questions are crucial, they only partially overlap with the concerns central to informed consent. Consent involves additional layers that address system-level workflows, institutional responsibilities, and long-term data governance [46, 57]. Furthermore, consent documents are predominantly textual and normative. They describe rights, obligations, and contingencies rather than algorithmic weights or feature contributions [17].

Traditional XAI techniques are not designed to analyze or explain legal text, policy structures, or multi-actor data flows. As a result, even systems that incorporate state-of-the-art explainability for predictions may still rely on opaque, static consent documentation [64, 53]. In such cases, transparency at the model level coexists with opacity at the governance level.

This mismatch suggests the need for a broader conception of XAI — one that encompasses not only decision outputs, but also the conditions under which data is collected, processed, shared, and repurposed [43, 59].

4.2 Consent as an Explainability Problem

If informed consent requires awareness and specificity, then the central question becomes how a system can ensure that the users meaningfully understand to what they are consenting.

From this perspective, consent can be reframed as a continuous alignment problem between (i) the system's internal representation of data practices and capabilities and (ii) the user's mental model of those practices and capabilities. Misalignment occurs when users underestimate data inference capabilities, overlook secondary uses, misunderstand revocation mechanisms, or fail to grasp the implications of automated processing. Such misalignment may persist even when information has been formally disclosed. Explainability offers mechanisms to reduce this gap.

An explainability-aware consent system would: (i) decompose complex data practices into modular, intelligible components, (ii) provide layered explanations that adapt to the user's level of understanding, (iii) allow users to interrogate specific aspects of data use, and (iv) detect potential misunderstandings and trigger corrective explanations.

In this sense, explainability is not merely an add-on to consent but a precondition for its validity in AI-driven environments. Consent without explanation risks degenerating into symbolic compliance; explanation without consent risks undermining autonomy. The two must therefore be structurally integrated. Thus, explanations in this context go beyond static clarification. They become interactive, context-sensitive, and temporally distributed. Understanding is not assumed at a single point in time; it is assessed and reinforced throughout system use.

4.3 Reframing XAI as a Pillar of Consent Governance

Repositioning XAI at the center of consent design has broader implications. It shifts the focus from explaining isolated decisions to explaining institutional practices. In this expanded view, XAI is intended to contribute to transparency in data flows and purpose, accountability — through traceable explanation logs — detection of user misunderstanding, and the reduction of manipulative or coercive interface patterns.

At the same time, this expanded role introduces new responsibilities. Explanations themselves can influence user choices. If poorly designed, they may oversimplify risks, create false reassurance, or subtly nudge acceptance. Therefore, explainability in the consent context must adhere to principles of neutrality, proportionality, and auditability.

In other words, AI-mediated consent requires a transition from static disclosure to interactive, explainable governance. Classical XAI techniques provide

essential building blocks but must be integrated into a broader socio-technical framework that connects model behavior, system architecture, and user interaction.

In the following section, we translate this conceptual reframing into concrete design principles for XAI-supported informed consent, outlining how segmentation, layered explanations, interactive verification, and adaptive feedback can be operationalized within intelligent assistive systems.

5 Design Principles for Explainable and Verifiable Informed Consent

The conceptual reframing of informed consent as an explainability problem calls for design principles that are not merely normative but operational. To bridge theory and implementation, consent must be embedded into the behaviour of intelligent systems in ways that are enforceable, inspectable, and adaptable over time. This section proposes a set of design principles for XAI-supported informed consent that translate legal and cognitive requirements into concrete system-level constraints.

These principles are intended as cumulative and interdependent, and are intentionally formulated at an intermediate level of abstraction. They are sufficiently general to apply across domains, yet precise enough to guide the design of explainable, agent-based consent management systems. Together, they define consent not as a static interface component, but as a distributed, interactive process spanning user interaction, explanation generation, and system control.

Modularization of Consent Content: A fundamental limitation of traditional consent practices lies in their monolithic structure. Lengthy, all-encompassing documents overload users and obscure the relationship between the agreed-upon terms and the actual system behavior [39, 57, 46]. To address this, consent information should be modularized into coherent, semantically distinct units.

Each consent unit corresponds to a specific category of data processing or system functionality, such as data collection, personalization, data sharing, or automated inference. Modularization serves two complementary purposes. At the interaction level, it reduces cognitive load by allowing users to focus on one aspect at a time. At the system level, it enables direct linkage between consent decisions and operational capabilities.

From an architectural perspective, this implies that consent units are not merely textual fragments but executable constraints. Granting or withdrawing consent for a given unit should activate or deactivate corresponding system components or agent behaviors. In this way, modular consent becomes enforceable by design rather than dependent on downstream compliance checks.

Layered and Adaptive Explanations: Legal and technical completeness often conflict with accessibility. While detailed descriptions may be required for

compliance, they can overwhelm non-specialized users [1, 17]. Explainable consent systems must therefore support layered explanations that present information at multiple levels of detail.

At a minimum, each consent unit should include: (i) a legally complete description, (ii) a simplified explanation in accessible language, and (iii) contextual examples illustrating concrete implications.

Explainable AI techniques play a central role in dynamically managing these layers. Rather than presenting all information upfront, the system should adapt the depth and form of explanation based on user interaction, expressed uncertainty, or detected misunderstandings. For instance, if a user requests clarification or answers a comprehension question incorrectly, the system can provide additional rephrasing, analogies, or examples.

Crucially, layered explanations should remain traceable. Users must be able to identify how simplified explanations relate to the authoritative legal text. This traceability preserves legal robustness while supporting user understanding and creating an auditable explanation pathway for institutional accountability.

Interactive Verification of Understanding: Exposure to information does not guarantee comprehension. A core requirement of informed consent — often assumed but rarely enforced — is that users actually understand what they are consenting to [24]. Explainable consent systems should therefore incorporate mechanisms for verifying understanding.

This principle operationalizes explainability as a dialogical process rather than a one-way disclosure. After presenting each consent unit, the system should assess comprehension through targeted questions or prompts. These may include true-or-false statements, quick multiple-choice questions, scenario-based questions, or asking for short open-ended responses. The goal is not to test users in an adversarial sense, but to identify potential misunderstandings early.

When misunderstandings are detected, the system should provide corrective explanations rather than automatically proceeding. At this stage, a crucial point is that acceptance should not be possible without a demonstrated understanding of the essential aspects. In this way, comprehension becomes a precondition for consent, aligning system behavior with the normative requirement of awareness.

From a system-design perspective, this requires maintaining an internal representation of user understanding states. Consent is thus modeled not as a binary variable, but as a validated condition that can be reinforced, degraded, or reestablished over time.

Granular Control: A common criticism of consent mechanisms is that refusal often entails total service denial, discouraging meaningful choice [4, 62]. Explainable consent design should instead support granular control, allowing users to consent to specific functionalities or data uses selectively.

This principle requires a close link between consent units and system capabilities. When a user does not understand or declines a particular consent unit,

only the corresponding functionalities should be limited or disabled. The system should explicitly explain the consequences of such limitations, reinforcing transparency rather than penalizing the user through opaque degradation.

Explainability is critical here. Users must understand not only *what* is disabled, but also *why*. This reinforces trust and mitigates the perception of coercion.

Temporal Consent Validation: User understanding may evolve as much as AI systems do. Consent that was *informed* at one point in time may cease to be so as system functionalities change or as users forget previously provided information. Explainable consent must therefore be treated as a temporal process. This principle extends dynamic consent by incorporating periodic validation of understanding. During system use, users may be prompted — at appropriate intervals — with short verification questions related to previously accepted consent units. If misunderstandings emerge, the system can reintroduce explanations and offer users the opportunity to reconfirm or modify their choices.

Such a continuous validation supports long-term alignment between system behavior and user expectations. It also strengthens legal defensibility by demonstrating ongoing efforts to maintain informed consent, rather than relying solely on initial acceptance.

Explainability as an Accountability Mechanism: Finally, explainable consent systems must be accountable not only to users but also to regulators, developers, and auditors. Explanations provided, user responses, detected misunderstandings, and consent modifications should be traceable in a privacy-preserving manner. This traceability may also frame explainability as a governance tool. It can allow institutions to demonstrate that consent was not merely obtained, but actively supported through explanation and verification. At the same time, it may create safeguards against manipulative explanation strategies by enabling post-hoc scrutiny of how information was presented.

Despite the above, it is very important to stress that accountability mechanisms should be designed to complement — not replace — human oversight. In cases of persistent misunderstanding or high-risk data processing, explainable systems should support escalation to human professionals.

Naturally, these principles alone do not eliminate all the challenges associated with informed consent in its legal dimension, and those associated with its intersection with AI technologies. Keeping these potential issues in mind is important in order to further refine the ways in which these principles can be made operational.

6 Risks, Limitations, and Ethical Safeguards

Reframing informed consent as an explainable, interactive process does not eliminate its structural tensions. In contrast, embedding consent into AI-mediated

systems introduces new ethical and technical risks that must be explicitly addressed. This section critically examines the limitations of explainable consent mechanisms and proposes safeguards to prevent misuse, overreach, or unintended harms.

6.1 The Risk of Persuasive or Manipulative Explanations

Explainability is not inherently neutral [13]. The way information is framed, simplified, or contextualized can influence user perceptions and decision-making. In consent contexts, this creates a paradox: systems designed to enhance understanding may also shape choices in subtle ways.

For example, simplified explanations may omit low-probability but high-impact risks. Analogies and examples, while cognitively helpful, may frame consequences in emotionally charged terms. Adaptive explanation strategies – intended to personalize comprehension support – could inadvertently introduce asymmetric influence if given users receive more persuasive or reassuring formulations than others [12].

Moreover, interactive verification mechanisms can be misused. Comprehension checks could be designed to implicitly guide users toward desired answers, or repeated prompts could pressure individuals into eventual acceptance. The line between clarification and persuasion can become blurred [13].

To mitigate these risks, explainable consent systems must adhere to principles of neutrality and proportionality. Explanations should aim to clarify implications without minimizing legitimate risks or exaggerating benefits. Questionnaires should be structured to detect misunderstanding rather than to enforce conformity. In addition, explanation strategies should be documented and auditable, enabling oversight and preventing covert manipulation.

6.2 Over-Reliance on Automation and the Illusion of Understanding

Another significant limitation concerns the potential over-reliance on automated explanation mechanisms. The integration of XAI and conversational interfaces may create the impression that a system fully resolves the epistemic challenges of consent. However, understanding is not easily measurable, and correct responses to verification questions do not guarantee deep comprehension [14].

Users may provide expected answers without internalizing implications, particularly if motivated by rapid access to services. Similarly, systems may misinterpret ambiguous responses or fail to detect subtle misunderstandings. In high-stakes domains such as healthcare, misplaced confidence in automated consent validation could have serious consequences.

Explainable consent systems should therefore avoid claiming epistemic certainty. Instead, they should treat understanding as probabilistic and context-sensitive. Escalation mechanisms to human professionals remain essential, especially when consent concerns complex or high-risk processing activities. Hybrid models — combining automated explanation with human oversight — are likely to be more robust than fully autonomous approaches.

6.3 Digital Literacy and Accessibility Gaps

Explainable and interactive consent presupposes a minimum level of digital literacy and access to technological infrastructure. Users with limited experience navigating digital interfaces, individuals with cognitive impairments, or those relying on assistive technologies may face additional barriers [19, 35, 55].

Layered explanations and interactive questionnaires, while intended to reduce cognitive overload, may inadvertently increase interaction complexity. Frequent prompts or verification checks could frustrate users or discourage engagement [47, 56]. Furthermore, conversational interfaces may not be equally accessible to individuals with language barriers or specific disabilities [15, 8].

Addressing these concerns requires inclusive design and multidisciplinary collaboration. User interface refinement should be informed by human–computer interaction research, cognitive psychology, and accessibility standards [66, 35]. Adaptive systems should detect potential usability challenges and adjust explanation pacing, modality — text, audio, visual cues, etc. — and complexity accordingly [55, 3].

Explainability, in this sense, must extend beyond semantic clarity to encompass usability and accessibility as integral components of informed consent [40, 22].

6.4 The Risk of Functional Coercion

Granular consent mechanisms that disable specific functionalities upon refusal or misunderstanding may unintentionally create forms of functional coercion. If essential services depend on consent to extensive data processing, users may perceive refusal as impractical or punitive.

While coupling consent units to system functionalities enhances transparency, it also requires careful calibration. Systems should clearly communicate the consequences of partial consent and avoid bundling unrelated functionalities in ways that undermine meaningful choice.

Regulatory frameworks often emphasize that consent must be freely given and that power imbalances must be considered [23]. In AI-mediated systems — particularly those providing essential services — this concern becomes acute. Designers must therefore evaluate whether alternative lawful bases for processing are more appropriate in certain contexts and ensure that consent is not used as a blanket legitimization mechanism.

6.5 Accountability and Governance

Finally, embedding explainability into consent processes introduces new governance challenges. Explanation logs, comprehension states, and interaction histories constitute additional data layers that themselves require protection [62, 21]. Storing detailed records of misunderstandings or consent interactions may create secondary privacy risks [17].

At the same time, such records are crucial for accountability. They allow institutions to demonstrate that users were provided with explanations, that misunderstandings were addressed, and that consent decisions were respected over time [41, 43]. Balancing transparency with data minimization principles, therefore, becomes a central design challenge.

Robust governance mechanisms should include: (i) clear documentation of explanation-generation strategies, (ii) privacy-preserving logging of consent states and verification outcomes, (iii) periodic audits of explanation neutrality and fairness, and (iv) defined escalation pathways to human review in complex cases.

In this expanded framework, explainability becomes both a user-facing feature and an institutional control mechanism. It supports not only individual autonomy but also systemic accountability.

7 Discussion and Future Works

The analysis presented in this paper suggests that the persistent shortcomings of digital consent mechanisms cannot be resolved solely through legal refinement or interface simplification. Instead, they reflect a deeper structural mismatch between the normative ideal of informed consent and the operational realities of contemporary digital systems. As data processing becomes increasingly complex, dynamic, and automated, static consent disclosures struggle to convey meaningful understanding or support genuinely autonomous decision-making.

This paper argues that explainable AI provides a conceptual and technical pathway for addressing this mismatch. Rather than treating consent as a one-time informational transaction, explainability enables the design of systems that actively support user understanding through interactive clarification, contextualized explanations, and iterative verification of comprehension. In this reframing, consent becomes an ongoing socio-technical process embedded within system architecture rather than an external legal formality.

7.1 From Disclosure to Interaction

Traditional consent mechanisms rely heavily on disclosure: providing information and assuming that individuals will interpret and internalize it [57]. However, decades of empirical research in privacy and human-computer interaction have shown that individuals rarely read, fully understand, or meaningfully evaluate lengthy disclosures [39]. The challenge is therefore not simply informational scarcity but cognitive and contextual overload.

The model proposed here treats understanding as an outcome that must be actively supported and verified. It highlights that explanation mechanisms should be integrated into the consent interface to enable users to request clarifications, explore the implications of specific data processing activities, and receive contextualized explanations generated through AI-supported techniques. This transforms consent from a static legal artifact into a dynamic interaction between users and the system.

While previous research in explainable AI has focused primarily on explaining algorithmic decisions or predictions [27], this paper extends the role of explainability to the pre-decisional phase of system interaction, where individuals must understand the conditions under which data will be collected and used. This repositioning of explainability represents a conceptual shift in how XAI can contribute to governance and user empowerment.

7.2 Explainability as Infrastructure for Consent

The framework outlined in this paper proposes that explainability functions as an infrastructural component of consent, rather than a *post hoc* transparency feature. It shapes how system behaviour is communicated, how user preferences are interpreted, and how decisions are recorded and enforced. To this end, the inclusion of interactive questionnaires and verification steps is crucial, so as to assess the user's understanding of key elements of the consent agreement. When misunderstandings are detected, the system can trigger additional explanations or simplified summaries before allowing full activation of the relevant functionality.

This mechanism introduces an important shift in responsibility. Instead of placing the entire burden of comprehension on the user, the system itself becomes responsible for facilitating and confirming understanding. In doing so, the consent process becomes closer to the normative ideal envisioned in legal and ethical frameworks, where authorisation should be based on genuine awareness of the consequences of a decision.

Importantly, this proposal does not aim to create barriers to access, but rather to ensure that consent reflects a minimum level of meaningful engagement with the information provided.

7.3 Implications for AI Governance

Reconceptualising consent through explainability also has implications for broader AI governance debates. Current regulatory approaches often rely on transparency requirements that assume information disclosure will enable meaningful oversight [21]. However, without mechanisms that translate complex system behavior into understandable explanations, transparency risks becoming largely symbolic [62, 10].

This paper expands the scope of XAI by applying explainability to the governance layer of digital systems, where individuals negotiate the terms under which technologies may operate on their data. In this sense, explainability is not limited to clarifying algorithmic outputs [22, 40] but becomes a tool for structuring the interaction between users, service providers, and intelligent systems.

This perspective aligns closely with the concerns of the multi-agent systems community. Digital consent environments can be understood as ecosystems in which multiple agents — human users, platforms, automated services, regulators, and oversight bodies — interact under partially aligned incentives. Within

such environments, explanation mechanisms serve as coordination tools that help agents communicate intentions, constraints, and consequences.

By situating consent within this multi-agent context, the paper introduces a new research direction that connects explainable AI with socio-technical governance and regulatory compliance.

Nonetheless, it is important to stress that explainability should not be seen as a universal substitute for other regulatory safeguards. In many contexts, structural protections — such as data minimisation, purpose limitation, and risk-based oversight — remain essential complements to consent-based mechanisms. Explainable consent should therefore be understood as one component within a broader governance ecosystem.

7.4 Research Directions for Multi-Agent Explainable Consent

The integration of explainable AI into consent processes raises several open research questions that merit further exploration, particularly within the multi-agent systems community.

First, the design of explanation strategies must account for heterogeneous user capabilities, preferences, and contexts. Developing adaptive explanation mechanisms that remain fair, non-manipulative, and auditable remains a significant challenge.

Second, the interaction between human users and multiple AI agents — each representing different services or institutional actors — requires new coordination models. Agents may need to negotiate consent states, reconcile conflicting explanations, or propagate consent decisions across interconnected services.

Third, evaluating comprehension and autonomy in interactive consent systems requires new methodological approaches. Traditional usability metrics may be insufficient to capture the quality of understanding or the authenticity of user choice.

Finally, governance mechanisms for explainable consent systems must balance accountability with privacy. Explanation and verification logs can enhance transparency but also introduce new data protection challenges that must be carefully managed.

Addressing these questions will require collaboration across disciplines, including artificial intelligence, law, human–computer interaction, ethics, and public policy.

8 Conclusions

The increasing complexity of digital ecosystems has exposed the structural fragility of traditional consent mechanisms. Although informed consent remains a cornerstone of data governance and individual autonomy, its current digital implementations often fail to support genuine understanding or meaningful decision-making. Static disclosures, lengthy legal texts, and one-time acceptance mechanisms are poorly suited to environments characterized by dynamic data flows, automated processing, and asymmetric technical expertise.

The principal contribution of this work is the re-conceptualization of informed consent as an explainability-driven governance process and the introduction of operational design principles for explainable and verifiable consent architectures in AI-mediated and multi-agent environments.

More specifically, this analysis has argued that addressing the above mentioned limitations requires a shift in perspective: rather than treating consent as a formal compliance step, it should be embedded within the operational logic of digital systems as a supported process of user understanding and interaction. Within this perspective, explainable AI offers a promising foundation for rethinking how consent is communicated, negotiated, and verified.

Building on this premise, the paper proposed a conceptual and architectural framework for explainable consent systems, in which explanation mechanisms, modular information structures, and comprehension verification tools are integrated into the design of consent interfaces. The proposed approach aims to bridge the gap between the normative requirements of informed consent and the technical realities of contemporary AI-driven services by transforming disclosure into an interactive and adaptive process.

Beyond its immediate application to consent interfaces, the framework highlights a broader role for explainability within digital governance. When integrated at the system level, explanation mechanisms can support not only individual comprehension but also institutional accountability, traceability of consent interactions, and more transparent coordination between the actors involved in digital service ecosystems.

At the same time, the proposal should be understood as an initial step toward a more comprehensive research agenda. Important questions remain regarding the implementation of adaptive explanations, the evaluation of user comprehension in real-world contexts, and the governance implications of storing and auditing explanation interactions. Addressing these challenges will require continued collaboration between researchers in artificial intelligence, human–computer interaction, law, and regulatory policy.

As AI-enabled services continue to permeate everyday digital life, ensuring that individuals retain meaningful control over how their data is used will remain a central challenge. Embedding explainability into the architecture of consent mechanisms offers a promising pathway for aligning technological innovation with the fundamental principles of transparency, autonomy, and accountability that underpin modern data protection frameworks.

References

1. Acquisti, A., Brandimarte, L., Loewenstein, G.: Privacy and human behavior in the age of information. Science **347**(6221), 509–514 (2015). https://doi.org/10.1126/science.aaa1465
2. Acquisti, A., Brandimarte, L., Loewenstein, G.: Privacy and human behavior in the age of information. Science **347**(6221), 509–514 (2015). https://doi.org/10.1126/science.aaa1465

3. Amershi, S.e.a.: Guidelines for human-ai interaction. CHI Conference on Human Factors in Computing Systems pp. 1–13 (2019). https://doi.org/10.1145/3290605.3300233
4. Barocas, S., Nissenbaum, H.: Big data's end run around procedural privacy protections. Communications of the ACM **57**(11), 31–33 (2014). https://doi.org/10.1145/2668897
5. Bester, J., Horsburgh, C., Kodish, E.: The limits of informed consent for an overwhelmed patient: Clinicians' role in protecting patients and preventing overwhelm. AMA journal of ethics **18**, 869–886 (09 2016). https://doi.org/10.1001/journalofethics.2016.18.9.peer2-1609
6. Bettini, C., Wang, X.S., Jajodia, S.: Privacy in location-based services: State of the art and research directions. GeoInformatica **13**(2), 119–151 (2009). https://doi.org/10.1007/s10707-008-0059-7
7. Biggs, J.S., Marchesi, A.: Information for consent: Too long and too hard to read. Research Ethics **11**(3), 133–141 (2015)
8. Branham, S.M., Kane, S.K.: Deafblind professionals' experiences with accessible office technologies. In: CHI Conference on Human Factors in Computing Systems. pp. 1–10 (2015). https://doi.org/10.1145/2702123.2702333
9. Budin-Ljøsne, I., Teare, H.J.A., Kaye, J., Beck, S., Bentzen, H.B., Caenazzo, L., Collett, C., D'Abramo, F., Felzmann, H., Finlay, T., et al.: Dynamic consent: a potential solution to some of the challenges of modern biomedical research. BMC Medical Ethics **18**(1), 4 (2017). https://doi.org/10.1186/s12910-016-0162-9
10. Burrell, J.: How the machine 'thinks': Understanding opacity in machine learning algorithms. Big Data & Society **3**(1), 1–12 (2016). https://doi.org/10.1177/2053951715622512
11. Caine, K., Hanania, R.: Patients want granular privacy control over health information in electronic medical records. Journal of the American Medical Informatics Association **20**(1), 7–15 (2013). https://doi.org/10.1136/amiajnl-2012-001023
12. Carli, R., Calvaresi, D.: Reinterpreting vulnerability to tackle deception in principles-based xai for human-computer interaction. In: International workshop on explainable, transparent autonomous agents and multi-agent systems. pp. 249–269. Springer (2023)
13. Carli, R., Calvaresi, D.: The wildcard xai: from a necessity, to a resource, to a dangerous decoy. In: International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems. pp. 224–241. Springer (2024)
14. Chi, M.T.H., Bassok, M., Lewis, M.W., Reimann, P., Glaser, R.: Self-explanations: How students study and use examples in learning to solve problems. Cognitive Science **13**(2), 145–182 (1989). https://doi.org/10.1207/s15516709cog1302_1
15. Cowan, B.R.e.a.: What can i help you with?: Infrequent users' experiences of intelligent personal assistants. In: CHI Conference on Human Factors in Computing Systems. pp. 1–13 (2017). https://doi.org/10.1145/3025453.3025499
16. Cusack, N.M., Venkatraman, P.D., Raza, U., Faisal, A.: Smart wearable sensors for health and lifestyle monitoring: Commercial and emerging solutions. ECS Sensors Plus **3**(1), 017001 (2024). https://doi.org/10.1149/2754-2726/ad3561
17. Custers, B., Ursic, H., van der Hof, S.: EU Personal Data Protection in Policy and Practice. Springer (2018)
18. Delany, C.M.: Respecting patient autonomy and obtaining their informed consent: ethical theory—missing in action. Physiotherapy **91**(4), 197–203 (2005)
19. van Deursen, A.J.A.M., van Dijk, J.A.G.M.: Digital Skills: Unlocking the Information Society. Palgrave Macmillan (2014)

20. Doshi-Velez, F., Kim, B.: Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608 (2017)
21. Edwards, L., Veale, M.: Slave to the algorithm? why a right to explanation is probably not the remedy you are looking for. Duke Law & Technology Review **16**, 18–84 (2017)
22. Ehsan, U.e.a.: Expanding explainability: Towards social transparency in ai systems. In: CHI Conference on Human Factors in Computing Systems. pp. 1–19 (2021). https://doi.org/10.1145/3411764.3445313
23. European Data Protection Board: Guidelines 05/2020 on consent under regulation 2016/679 (2020), version 1.1, adopted 4 May 2020
24. European Parliament and European Council: General data protection regulation (gdpr), 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, and repealing directive 95/46/ec (2016) (2016)
25. Falagas, M.E., Korbila, I.P., Giannopoulou, K.P., Kondilis, B.K., Peppas, G.: Informed consent: how much and what do patients understand? The American Journal of Surgery **198**(3), 420–435 (2009). https://doi.org/10.1016/j.amjsurg.2009.02.010
26. Grady, C., Cummings, S.R., Rowbotham, M.C.: Informed consent and patient comprehension in digital health research. Journal of the American Medical Association **317**(3), 347–348 (2017). https://doi.org/10.1001/jama.2016.19501
27. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., Pedreschi, D.: A survey of methods for explaining black box models. ACM Computing Surveys **51**(5), 93 (2018). https://doi.org/10.1145/3236009
28. Ha, Q.A., Chen, J.V., Uy, H.U., Capistrano, E.P.: Exploring the privacy concerns in using intelligent virtual assistants under perspectives of information sensitivity and anthropomorphism. International journal of human–computer interaction **37**(6), 512–527 (2021)
29. Harish, D., Kumar, A., Singh, A.: Patient autonomy and informed consent: the core of modern day ethical medical. Journal of Indian Academy of Forensic Medicine **37**(4), 410–414 (2015)
30. Hochhauser, M.: Informed consent: reading and understanding are not the same: subjects may read consent forms, but they don't always understand them. Applied Clinical Trials **13**(4), 42–47 (2004)
31. Kahneman, D., Tversky, A.: Prospect theory: An analysis of decision under risk. Econometrica **47**(2), 263–291 (1979). https://doi.org/10.2307/1914185
32. Kaye, J., Whitley, E.A., Lund, D., Morrison, M., Teare, H., Melham, K.: Dynamic consent: a patient interface for twenty-first century research networks. European Journal of Human Genetics **23**(2), 141–146 (2015). https://doi.org/10.1038/ejhg.2014.71
33. Kim, Y., Xu, X., McDuff, D., Breazeal, C., Park, H.W.: Health-llm: Large language models for health prediction via wearable sensor data. arXiv preprint arXiv:2401.06866 (2024)
34. Kotsenas, A.L., Balthazar, P., Andrews, D., Geis, J.R., Cook, T.S.: Rethinking patient consent in the era of artificial intelligence and big data. Journal of the American College of Radiology **18**(1), 180–184 (2021)
35. Lazar, J., Goldstein, D., Taylor, A.: Ensuring Digital Accessibility Through Process and Policy. Morgan Claypool (2022)
36. Manson, N.C., O'Neill, O.: Rethinking informed consent in bioethics. Cambridge University Press (2007)

37. Martin, K.D., Zimmermann, J.: Artificial intelligence and its implications for data privacy. Current opinion in psychology **58**, 101829 (2024)
38. Mathur, A., Kshirsagar, M., Mayer, J.: What makes a dark pattern... dark? design attributes, normative considerations, and measurement methods. In: Proceedings of the 2021 CHI conference on human factors in computing systems. pp. 1–18 (2021)
39. McDonald, A.M., Cranor, L.F.: The cost of reading privacy policies. I/S: A Journal of Law and Policy for the Information Society **4**(3), 543–568 (2008)
40. Miller, T.: Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence **267**, 1–38 (2019). https://doi.org/10.1016/j.artint.2018.07.007
41. Mittelstadt, B.e.a.: The ethics of algorithms: Mapping the debate. Big Data & Society **3**(2), 1–21 (2016). https://doi.org/10.1177/2053951716679679
42. Monge Roffarello, A., De Russis, L.: Towards understanding the dark patterns that steal our attention. In: Chi conference on human factors in computing systems extended abstracts. pp. 1–7 (2022)
43. Morley, J.e.a.: The ethics of ai: A systematic literature review of principles and challenges. AI & Society **36**, 1–21 (2021). https://doi.org/10.1007/s00146-020-00992-1
44. Mühlbacher, A.C., Amelung, V.E., Juhnke, C.: Contract design: The problem of information asymmetry. International Journal of Integrated Care **18**(1), 1 (2018)
45. Nissenbaum, H.: Privacy as contextual integrity. Washington Law Review **79**(1), 119–158 (2004)
46. Nissenbaum, H.: A contextual approach to privacy online. In: Daedalus, vol. 140, pp. 32–48. MIT Press (2011). https://doi.org/10.1162/DAED_a_00113
47. Norman, D.: The Design of Everyday Things. Basic Books (2013)
48. Pandit, H., O'Sullivan, D., Lewis, D.: Modelling provenance for gdpr compliance using linked data. International Journal of Information Management **43**, 162–176 (2018). https://doi.org/10.1016/j.ijinfomgt.2018.07.009
49. Pietrzykowski, T., Smilowska, K.: The reality of informed consent: empirical studies on patient comprehension. Trials **22**, 57 (2021). https://doi.org/10.1186/s13063-020-04969-w
50. Ribeiro, M.T., Singh, S., Guestrin, C.: "why should i trust you?" explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD. pp. 1135–1144 (2016). https://doi.org/10.1145/2939672.2939778
51. Schaub, F., Balebako, R., Cranor, L.F.: Designing effective privacy notices and controls. IEEE Internet Computing **21**(3), 70–77 (2017). https://doi.org/10.1109/MIC.2017.75
52. Sedenberg, E., Hoffmann, A.L.: Recovering the history of informed consent for data science and internet industry research. Big Data & Society **3**(1), 1–11 (2016). https://doi.org/10.1177/2053951716641135
53. Selbst, A.D., Barocas, S.: The intuitive appeal of explainable machines. Fordham Law Review **87**(3), 1085–1139 (2018)
54. Sezgin, E., Huang, Y., Ramtekkar, U., Lin, S.: Readiness for voice assistants to support healthcare delivery during a health crisis and pandemic. npj Digital Medicine **3**, 122 (2020). https://doi.org/10.1038/s41746-020-00332-0
55. Shneiderman, B.: Human-Centered AI. Oxford University Press (2022)
56. Sillence, E., Briggs, P., Harris, P., Fishwick, L.: A framework for understanding trust factors in web-based health advice. International Journal of Human-Computer Studies **64**(8), 697–713 (2006). https://doi.org/10.1016/j.ijhcs.2006.02.007

57. Solove, D.J.: Privacy self-management and the consent dilemma. Harvard Law Review **126**(7), 1880–1903 (2013)
58. Solove, D.J., Hartzog, W.: The ftc and the new common law of privacy. Columbia Law Review **114**(3), 583–676 (2014)
59. Suresh, H., Guttag, J.: A framework for understanding sources of harm throughout the machine learning life cycle. Equity and Access in Algorithms, Mechanisms, and Optimization pp. 1–9 (2021). https://doi.org/10.1145/3465416.3483305
60. Teare, H.J.A., Prictor, M., Kaye, J.: Dynamic consent: a patient interface for twenty-first century research networks. European Journal of Human Genetics **26**(8), 1081–1088 (2018). https://doi.org/10.1038/s41431-018-0175-0
61. Tene, O., Polonetsky, J.: Big data for all: Privacy and user control in the age of analytics. Northwestern Journal of Technology and Intellectual Property **11**(5), 239–273 (2013)
62. Wachter, S., Mittelstadt, B.: 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 (2019)
63. Wachter, S., Mittelstadt, B., Floridi, L.: Why a right to explanation of automated decision-making does not exist in the gdpr. International Data Privacy Law **7**(2), 76–99 (2017). https://doi.org/10.1093/idpl/ipx005
64. Wachter, S., Mittelstadt, B., Russell, C.: Counterfactual explanations without opening the black box. Harvard Journal of Law & Technology **31**(2), 841–887 (2017)
65. Wisgalla, A., Hasford, J.: Four reasons why too many informed consents to clinical research are invalid: a critical analysis of current practices. BMJ Open **12**(3) (2022). https://doi.org/10.1136/bmjopen-2021-050543, https://bmjopen.bmj.com/content/12/3/e050543
66. Wobbrock, J.O., Kientz, J.A.: Research contributions in human-computer interaction. In: Ways of Knowing in HCI, pp. 85–94. Springer (2016)