

Adversarial Intentionality: Extending Postphenomenology for Technology-Mediated Deception

Ryan Straight

Abstract

Postphenomenology provides powerful resources for analyzing how technologies mediate human experience, yet the framework maintains a foundational assumption that mediation operates through recognizable logics accessible to phenomenological examination. This paper identifies a boundary condition where that assumption fails, namely technologies engineered to conceal their operative logic while exploiting phenomenological vulnerabilities. I introduce *adversarial intentionality* as a theoretical extension, articulating three constitutive dimensions: intentional deception as mediation design, exploitative phenomenological engineering, and relational weaponization. Analysis of AI-enhanced social engineering (deepfake fraud, business email compromise, and romance scams) demonstrates how adversarial configurations systematically exploit vulnerabilities intrinsic to embodiment, hermeneutic, alterity, and background relations. This extension preserves postphenomenology’s relational ontology while revealing phenomenological vulnerability as a constitutive feature of human-technology relations. At stake is phenomenological integrity, the correspondence between mediated experience and the domains it presents, on which collective meaning-making depends.

Author’s manuscript. Accepted for publication in Techné: Research in Philosophy and Technology (forthcoming). This version is not the journal’s version of record; please cite the published article in Techné.

In early 2024, a finance worker at multinational engineering firm Arup made 15 transfers totaling US\$25.6 million after a video conference he believed included his CFO and colleagues. Every other participant was a deepfake (Chen and Magramo 2024; Magramo 2024; Kong 2024). The case exemplifies a crisis in which mediation, and the trust that ordinary experience places in it, has been destabilized. That same year, the United States Federal Bureau of Investigation (FBI)’s Internet Crime Complaint Center reported \$2.77 billion in business email compromise losses, where fabricated organizational communications enabled financial fraud at

institutional scale (Federal Bureau of Investigation, Internet Crime Complaint Center (IC3) 2025, 10). Romance scams, in which attackers cultivate false intimate relationships to extract money, generated \$1.14 billion in reported losses during 2023, showing how a fabricated persona met through a platform enables emotional exploitation across entire populations (Fair 2024). Beyond these quantified harms lies the threat of synthetic media becoming phenomenologically indistinguishable from authentic mediation, eroding the institutional trust and collective sense-making that rest on mediated evidence.

Existing frameworks for analyzing technology-mediated deception address symptoms, not phenomenological structures. They treat deception as deviation from proper technological operation rather than as constitutive mediation design. Technical security approaches focus on identifying spoofed domains, deepfake artifacts, and behavioral anomalies through algorithmic analysis. Yet detection cannot explain why legitimate-appearing mediations succeed phenomenologically despite technical indicators suggesting suspicion. A gap persists between what technical detection finds and what the target experiences. Psychological accounts (Ferreira and Teles 2019, 26–29) catalog the persuasion cues, chiefly authority and strong affect, that phishing messages carry, locating deception in message design while setting technological mediation aside. What this misses, I argue, is that mediation structures generate vulnerabilities of their own. Human-computer interaction work on online trust (Shneiderman 2000, 58–59) takes interface design seriously and places the remedy in disclosure, third-party certification, and guarantees, leaving unexamined how the fluency enabling skilled use also enables deceptive mediation.

Postphenomenology offers analytical resources these approaches lack. By examining how

technologies mediate rather than neutrally transmit experience, postphenomenological analysis illuminates why interfaces enable deception: hermeneutic relations presume interpretive correspondence between representations and world states, embodiment relations achieve transparency through pre-reflective habituation, alterity relations present the technology itself as a quasi-other, and background relations situate experience within infrastructural conditions that remain outside focal awareness (Ihde 1990; Verbeek 2005). Each constitutes a condition through which mediation enables world-access. Despite these resources, canonical postphenomenology maintains a limitation. The framework assumes technologies mediate in good faith or break down unintentionally, treating deception as contingent failure. Ihde's foundational typology analyzes relations enabling genuine world-access, however perspectival, presuming mediation operates toward epistemic fidelity even when transforming experience. Accounts of technological intentionality (Verbeek 2011; Mykhailov 2020, 2023) describe how artifacts shape action through affordances that arise in use relations, examining normative delegation without theorizing configurations designed to conceal their operative logic. Rosenberger's politically engaged work on hostile architecture addresses structural harm through unobtrusive material constraints (Rosenberger 2017, xii, 22), analyzing artifacts whose political inscriptions remain accessible to phenomenological examination. None theorize technologies engineered to fabricate rather than mediate reality, weaponizing mediation structures through configurations designed to resist the interrogation postphenomenology employs.

I introduce *adversarial intentionality* as a theoretical extension addressing an undertheorized domain in postphenomenology, namely deliberately antagonistic mediations injected to

exploit phenomenological vulnerabilities through deception. The concept describes socio-technical configurations (human-technology-world assemblages) designed to generate deceptive experiences resistant to reflective interrogation, succeeding phenomenologically by concealing their adversarial operative logic. I develop the concept through three constitutive dimensions that preserve postphenomenology's core commitments while expanding its analytical range to encompass configurations the tradition has excluded.

First, intentional deception as mediation design distinguishes adversarial configurations from ordinary mediation, which reduces through perspectival selectivity while preserving epistemic fidelity, and from complex systems, whose operative logics are difficult yet remain accessible to investigation. Adversarial mediations engineer discordance between appearance and function, designing interfaces that present themselves as mediating one domain while mediating fabricated alternatives.

Second, exploitative phenomenological engineering names the deliberate targeting of vulnerabilities that belong to mediation as such. Every relation succeeds through some structure that withdraws from scrutiny: the transparency of embodiment, the trust of hermeneutic interpretation, the felt presence of alterity, the invisibility of background infrastructure. These structures are not defects. They are what mediation functioning *as designed* looks like, and the engineering consists in turning each of them to deceptive ends. Third, relational weaponization holds that adversarial character emerges from configured assemblages rather than isolated components. By *weaponization* I mean the conversion of a mediation's enabling structure into the means of attack, so that the feature which lets a relation work is the feature through which it is exploited. Because the conversion happens at the level of the relation,

analysis must address whole socio-technical configurations.

The framework preserves relational ontology by maintaining human-technology assemblages as the analytical unit, pragmatist orientation by pursuing empirical investigation of concrete configurations, and analytical methods by adapting variational analysis for adversarial contexts. The concept's necessity becomes apparent through examining AI-enhanced social engineering, where adversarial configurations weaponize phenomenological structures, exploiting each type of human-technology relation. Attacks map onto specific phenomenological vulnerabilities rather than generic "human error," exploiting mediation structures. Postphenomenological examination exposes exploitation patterns invisible to technical security or psychological frameworks.

I contribute theoretically by extending postphenomenology to encompass adversarial mediations and methodologically by adapting variational analysis for weaponized configurations. Broader implications address AI governance and design, opening inquiry into what I call "defensive postphenomenology," the development of normative principles for technologies resistant to adversarial configuration while preserving mediation's world-disclosing function. Follows is a five-section argument. Section 1 establishes postphenomenological foundations, examining human-technology relations, multistability, and politically engaged analysis before identifying the shared accessibility assumption that adversarial cases challenge. Section 2 articulates adversarial intentionality through its three constitutive dimensions and develops a typology of phenomenological vulnerabilities across Ihde's four relation types. Section 3 demonstrates the framework's analytical necessity through three AI-enhanced social engineering cases. Section 4 addresses limitations and future directions, including defensive

postphenomenology. Section 5 concludes with implications for collective meaning-making under adversarial pressure.

1 Postphenomenological Foundations

Postphenomenology emerges from classical phenomenology’s concern with lived experience while reorienting its analytical focus toward concrete technologies. Where classical phenomenology sought universal structures of consciousness, postphenomenology investigates, following Ihde’s phenomenology of technics (1990, 72), how specific technologies shape particular experiences in concrete contexts. Merleau-Ponty’s insight that embodied perception structures world-access provided the groundwork. His analysis of how a cane extends tactile perception through pre-reflective integration into bodily schema anticipates postphenomenology’s recognition that technologies transform experience (2002).

I take as foundational postphenomenology’s reconceptualization of the analytical unit from discrete subjects and objects to human-technology relations, shifting philosophical attention toward configured assemblages mediating world-access (Rosenberger and Verbeek 2015; Verbeek 2005). Technologies shape perception, action, and understanding, emerging as transformative mediators that co-constitute experience through perspectival selectivity and through the amplification-reduction structure Ihde establishes for embodiment relations (Ihde 1990, 75–78). Kudina (2021) models this co-constitution through the hermeneutic lemniscate, an infinity-loop structure in which humans appropriate technologies, technologies reshape the sociocultural world, and the reshaped world reconstitutes the interpreting subject. The lemniscate captures what linear models miss, that mediation is a recursive feedback

system in which all three nodes (human, technology, world) participate in meaning-making. This non-neutrality (Verbeek 2005) is key in analyzing *adversarial* systems, where actors configure mediations to produce deceptive experiences by poisoning the lemniscate’s feedback channels, injecting fabricated meanings while concealing the injection. Because human-technology relations are constitutive, they ground examination of how social engineering exploits structural conditions of mediated experience rather than individual cognitive failures.

1.1 Human-Technology Relations and Adversarial Multistability

Postphenomenology’s analytical power derives from its typology of human-technology-world relations, elaborated by Ihde (1990). Ihde positions human-technology relations as the fundamental analytical unit, revealing how artifacts mediate experience while remaining bound to human contexts. Four primary configurations (embodiment, hermeneutic, alterity, and background relations) provide vocabulary for analyzing how mediation takes diverse forms, and each matters for understanding adversarial intentionality. Table 1 details each type’s structural characteristics and corresponding phenomenological vulnerabilities.

Before turning to specific relations, I introduce a condition that cuts across all four. *Relegation* (Wellner 2020a, 181; 2024, 9) names a delegation imposed on the user, in which human intentionality is assigned a position inferior to the technology’s. Wellner develops the concept for augmented-reality and AI contexts, but the condition generalizes: it can attach to embodiment, hermeneutic, alterity, or background relations alike. I foreground it here because adversarial intentionality exploits relegation across the typology, and the analysis that follows turns on recognizing where users have been relegated to mediations they cannot

interrogate.

Multistability recognizes that technologies enable multiple uses and interpretations (Ihde 1990, 139, 144). Where Husserl’s variations sought essences, postphenomenology finds multistabilities (Rosenberger 2020, 2230–31). Artifacts afford diverse stable configurations limited by material and social contexts: eyeglasses correct vision while functioning as fashion accessories, benches enable sitting but become sleeping surfaces or skateboarding obstacles. Each represents a coherent human-technology relation, not repurposing.

Rosenberger’s (2020) variational cross-examination proposes a method for investigating multistability, first identifying possible stabilities through variation, then contrasting these to reveal dimensions otherwise difficult to perceive. Dominant stabilities appear natural because design, habit, and social context optimize them, and cross-examining alternatives can bring hidden phenomenological and sometimes political dimensions to light (2020, 2234).

Variational analysis of this kind always proceeds from a chosen anchor. Whyte (2015) distinguishes two senses of multistability, an imaginative sense (a first-person, testable hypothesis about how several stable patterns of one object can be perceived) and a practical sense (the entanglement of a technology within a lifeworld), and shows that practical multistability proceeds by *pivoting*, holding something constant so that a field of differences reads as variations of one thing. Whyte identifies a pivot in each of four authors, with Ihde pivoting on practice, Rosenberger on the artifact, Verbeek on standards, and Selinger on perspective. The pivot point matters for the adversarial case because it specifies what an attacker must hold stable. Adversarial configurations engineer a false practical stability by fixing the pivot the victim is most likely to trust, namely the familiar interface, the expected colleague, or

the routine authorization, so that the deceptive variation presents itself as the obvious one. Stability arises from the environment’s role as well as from material design (de Boer 2023). Drawing on Gibsonian affordance theory and on Rietveld and Kiverstein, de Boer conceptualizes the environment as a “rich landscape of affordances” where selective perception operates through participation in “forms of life” (2023, 2272), the stable normative practices characteristic of specific communities. Normativity embedded in these forms of life constrains which affordances solicit action, making specific uses appear natural while rendering alternatives transgressive. Adversarial exploitation succeeds because attackers engineer configurations aligning with victims’ forms of life, making adversarial stabilities appear legitimate within established patterns.

Adversarial multistability differs from ordinary repurposing. Skateboarding on benches represents alternative stability accessible through investigation as the practice remains visible and its operative logic recognizable through phenomenological inquiry or variational cross-examination. By contrast, the adversarial case involves deliberate concealment, presenting one stability while enacting another, engineering configurations that resist variational cross-examination by rendering alternatives phenomenologically inaccessible. The challenge emerges not from multistability’s existence, which postphenomenology treats as constitutive of technological mediation, but from weaponization: engineering false stabilities that exploit victims’ reasonable presumption that mediations correspond to their apparent operative logic, transforming mediation’s enabling conditions into vulnerabilities.

1.2 Politically Engaged Postphenomenology

Postphenomenology's critical turn examines how mundane artifacts are delegated the control of disenfranchised groups (Rosenberger 2014, 387–88). Work on hostile architecture, such as benches with armrests and ledges with spikes deterring rough sleeping, demonstrates how design features transform public infrastructure into social control mechanisms (Rosenberger 2017; though see Wellner 2020b, who complicates the count of stabilities). This approach refuses the presumption that exclusionary design is a neutral choice, and Rosenberger (2017, xiii, 30) locates responsibility for it across the many actors who specify, manufacture, and install such objects, not in designers alone. Verbeek (2020) sketches a postphenomenological politics of technology along three dimensions: the mediation of power through perceptual and interpretive shaping, the structuring of political interaction through technological arrangements, and the constitution of political issues. Together, these dimensions reveal how technologies participate in politics as active mediators configuring the conditions under which political interaction and issue formation occur.

Both Rosenberger's and Verbeek's politically engaged analyses, however, share an assumption that exclusionary or politically charged designs operate through constraints accessible to phenomenological examination. Hostile architecture's operative logic becomes apparent through critical investigation, its political character emerging through recognizable material configuration. Verbeek's political hermeneutics similarly presumes that sustained inquiry can reveal how technologies shape political perception and action. I speak of accessibility here rather than transparency, and the choice is deliberate. In postphenomenology, transparency names the withdrawal of an embodied technology from focal awareness, the eyeglasses one

sees through rather than at (Ihde 1990, 73; Rosenberger and Verbeek 2015), and the analysis below treats that withdrawal as a vulnerability rather than a virtue. What politically engaged analysis presumes is something else: that an artifact’s operative logic can be brought to reflection by an investigator who sets out to find it. Accessibility in this second sense distinguishes politically engaged analysis from the adversarial case. Where Rosenberger illuminates how artifacts exclude, often inconspicuously, and Verbeek reveals how technologies mediate political participation, adversarial cases demand analysis of artifacts that deceive.

1.3 STS Connections and Distinctions

Postphenomenology participates in Science and Technology Studies traditions challenging technological neutrality. Winner (1980) argued that technologies embody politics beyond their professed uses whether or not anyone intends it, Latour (1992, 232, 254) revealed how technical objects enforce behaviors through material constraints, and Feenberg (1991, 79–82, 86, 89) argued that the technical code embeds social interests in apparently neutral designs yet leaves a margin of maneuver for subverting it from within. These frameworks establish technology’s non-neutrality as foundational STS insight.

Yet postphenomenology and STS share a foundational assumption that technological mediation maintains operative logics accessible through phenomenological or critical examination. Even critical analyses addressing harmful technologies (Winner’s exclusionary bridges, Feenberg’s rationalized systems, Rosenberger’s hostile architecture) examine mediations whose operative logics are accessible through sustained investigation. This assumption is empirically warranted across most technological domains, where breakdown analysis, variational

examination, and cross-cultural comparison successfully reveal how technologies encode values and configure experience. It encounters a boundary condition, however, in adversarial cases. Adversarial technologies present false realities and exploit users while resisting the investigative methods postphenomenology employs.

2 Adversarial Intentionality: A Theoretical Extension

Adversarial intentionality, as I develop it here, extends postphenomenology's relational ontology to encompass socio-technical configurations purposefully engineered to exploit phenomenological vulnerabilities. These configurations succeed by concealing their adversarial operative logic beneath facades. The extension preserves postphenomenology's core commitments to the relational ontology of human-technology assemblages as the analytical unit and the amplification-reduction structure of mediation. What it adds is explicit theorization of malicious configuration as a distinct category.

A word on perspective, since the analysis moves among several parties. The experiencing subject of an adversarial relation is, throughout, the target: the person whose mediated experience is configured and who suffers the deception. The attacker occupies a position outside the relation, configuring it (Section 2.3 develops this as a fourth intentional position, shown in Figure 1), and a platform's designers occupy a third, supplying the affordances the attacker exploits. Where the designer's own aims are extractive, as Zuboff's (2019) analysis of surveillance capitalism argues they routinely are, the attacker's and the designer's positions can coincide; I return to that convergence in Section 2.3.

Extending postphenomenology to encompass configurations defying existing categories represents a productive trajectory already underway. Besmer (2012) analyzes cochlear implants as “embodied translation technologies” (2012, 310) generating “cyborg intentionality” (Verbeek 2008), hybrid relations where hermeneutic translation becomes embodied perception that enables world-access otherwise impossible. I extend this trajectory toward configurations that exploit rather than enable, weaponizing phenomenological structures.

2.1 Dimension 1: Intentional Deception as Mediation Design

Intentional deception as mediation design distinguishes adversarial intentionality through its focus on deception operating at the mediation level, weaponizing the interpretive and constitutive functions through which technologies enable world-access. Verbeek’s (2011) development of technological intentionality, a term he takes from Ihde, describes how artifacts embody directionality shaping action through invitations and inhibitions that arise in use relations. Yet this directionality operates through genuine mediation of world states, however selectively or perspectively. Adversarial intentionality describes a qualitatively different condition, where mediation itself constitutes deception and the mediating technology becomes complicit in constructing false realities rather than shaping access to genuine phenomena.

Ordinary mediation involves inevitable reduction. For example, thermal imaging reveals heat signatures while concealing visible-light information, constituting mediation’s inherent selectivity through amplification-reduction dynamics. Such reduction represents mediation’s perspectival character, not deception. It preserves what I term *epistemic fidelity*, the degree to which a mediation maintains systematic correspondence between its phenomenological

presentation and the domain it purports to represent, including the source it claims to relay. Adversarial mediation differs by engineering systematic discordance between appearance and operative logic, exploiting users' reasonable presumption that mediations correspond to their apparent function. This performative dimension reveals how deception operates through configured relations. Dark patterns literature (Gray et al. 2018, 8–9) infers designer motives from interface examples practitioners share and grants that not all are equally dark in intent, suggesting that adversarial intentionality operates across a spectrum from subtle nudging through interface design to overt deception through systematic falsification.

Composite intentionality (Verbeek 2008) provides a foundation for this distinction. Technologies and humans combine their intentionalities in coupling, generating composite intentionality. Adversarial configurations exploit this coupling by engineering components that misrepresent their contribution. Where legitimate composite intentionality involves mediating world-access through perspectival selectivity, adversarial composite intentionality mediates deception through systematic falsification. The extension encompasses antagonistic configurations while preserving Verbeek's core insight about composite intentionality. I develop the phenomenological mechanism for this engineering below, where I argue that adversarial configurations hijack the unifying and hiding activities Mykhailov (2020) identifies as constitutive of any technological intentionality.

2.2 Dimension 2: Exploitative Phenomenological Engineering

Exploitative phenomenological engineering targets *phenomenological vulnerabilities*, structural features that make mediation exploitable. These vulnerabilities arise from mediation's

design, not from cognitive weaknesses or user deficiencies. Ihde (1990, 73, 75) shows how artifacts achieve a quasi-transparency through habituation: once learned, eyeglasses withdraw from focal awareness while enabling direct visual engagement, barely noticed within pre-reflective bodily schema. Yet this same pre-reflective incorporation creates exploitable pathways. Embodiment relations operating transparently resist conscious interrogation when the compromise preserves fluency rather than producing obvious breakdown. Subtly altered interfaces pass unnoticed where cracked lenses would be immediately foregrounded. The vulnerability emerges from mediation structure, for the transparency enabling skilled perception simultaneously generates exploitability through resistance to reflective examination. Merleau-Ponty's (2002, 126–27, 165–67) analysis of motor intentionality and habit shows how the body schema incorporates tools through artifactual coupling that operates beneath thematic awareness. This incorporation enables skilled action through pre-reflective integration, yet makes bodily habits resistant to cognitive override because they operate through motor rather than cognitive processes. Technologies incorporated into habitual comportment shape perception through operative intentionality, constituting world-access through embodied habits that resist reflective interrogation. Adversarial configurations exploit this by targeting incorporated technologies whose pre-reflective operation resists conscious interrogation. The vulnerability is constitutive, with habitual fluency necessarily involving pre-reflective integration that resists immediate conscious access, generating exploitability as structural feature rather than contingent failure.

In Mol's account the objects of medical practice are multiple rather than singular, done in practices and distributed across sites (Mol 2002, 5–6, 88, 115), and I take embodiment to be

multiple in the same way. I draw on Mol's framework to argue that adversarial configurations exploit this multiplicity through what I call *synthetic coherence*, the fabrication of a unified embodied presence (a face, a voice, a gestural repertoire) from independently generated signals. Victims experience that unity despite the ontological absence of any singular body producing the cues. Put plainly, ordinary perception trusts a hidden remainder to complete the manifest face, much as an apple, in Mykhailov and Liberati's reading of Husserl, has a backside we have no reason to doubt (2023, 845). Adversarial mediation supplies no such remainder, yet our perceptual habits assemble one anyway.

Different relations generate distinct vulnerabilities through their characteristic mediating structures. Embodiment relations are vulnerable through pre-reflective habit exploitation. Transparency resists foregrounding when compromised, so incorporated technologies operate beneath conscious interrogation at the moment they are most thoroughly weaponized. Hermeneutic relations generate vulnerability through semiotic manipulation as interpretation presumes that representations correspond to represented domains, operating through trust in interface signifiers that adversarial configurations falsify. When correspondence breaks, hermeneutic relations continue mediating deception, interpreting fraudulent signifiers with the same fluency applied to legitimate mediations. The structure of alterity relations cultivates vulnerability through fabricated presence. What elicits interpersonal engagement in an alterity relation is relational structure rather than the ontological status of the other, as it is the responsiveness, turn-taking, address, and apparent regard that make an encounter feel like an encounter with someone. Because those structural features can be produced without anyone behind them, fabricated presence generates quasi-personal connections with

absent entities. From the target’s side, such a connection is a relation through the medium to an apparent person, Ihde’s (I-telephone)-you (1990, 78). The attack configures the alterity structure behind that relation. Each vulnerability illustrates the same underlying breakdown: a phenomenological presentation without the corresponding domain it claims to represent, a systematic failure of epistemic fidelity engineered by the adversarial configuration.

Background relations present the most theoretically challenging case because they remain the least systematically theorized within canonical postphenomenology (though see Braga (2023) for recent attention to social dimensions). Ihde’s (1990) original treatment addressed analog infrastructures (climate control, lighting, electrical grids) that shape environmental conditions while remaining outside focal awareness, and extending this analysis to digital infrastructures is a necessary theoretical development. Aydin et al. (2019) introduce *technological environmentality* to analyze smart environments “intentionally directed at humans” (2019, 337) while merging with the world. They extend Verbeek’s immersion relation, which they distinguish from Ihde’s passive background relations by its interactivity (2019, 336). Yet their positioning addresses environments designed to respond beneficially to users, not infrastructures that remain phenomenologically inert while constituting attack surfaces.

Kranc (2017) characterizes the backgrounded artifact, in Ihde’s phrase, as a *present absence*, “just as real as foregrounded entities, but somehow resistant to conscious attention” (2017, 20), and argues that the background harbors the community’s socio-political imaginaries, the “strategies, goals, and choices” within which “the built world arises” (2017, 22–23). This widening matters for the adversarial case. If the background carries ideology as well as artifacts, then compromising it corrupts not merely a technical substrate but the taken-for-granted

conditions through which users take the world to be as it appears.

Here I take up a distinct phenomenon: digital infrastructures that maintain Ihde’s original passive character for users while remaining exploitable by adversarial actors, generating vulnerabilities through their withdrawal from attention rather than their responsiveness to it. Unlike embodiment relations, which breakdown can foreground, and unlike foreground relations generally, which variational cross-examination can open to inspection, these infrastructures resist phenomenological access even when compromised. I draw on Mykhailov (2020) for the phenomenological grounding of this resistance. Every technological intentionality, he argues, performs two structural activities: a *unifying activity* binding artifacts into functional meaning, and a *hiding activity* concealing the network of dependencies that makes the binding possible. Hiding activity, on this account, is ordinary and constitutive rather than deceptive. It makes experience “smooth and fluent” by keeping infrastructural dependencies outside focal awareness (Mykhailov 2020, 631–32). Background relations concentrate this structural hiding most thoroughly, since infrastructures function by withdrawing from attention. I argue that adversarial configurations weaponize the duality, engineering exploitative operative logic to tuck itself within the user’s already-expected opacity.

As established above, relegation is one of several phenomenological preconditions adversarial intentionality exploits. What distinguishes the adversarial case is that concealment becomes an explicit design goal rather than a structural byproduct. Users are relegated to dependence on infrastructures engineered to resist interrogation.

These infrastructures constitute the conditions enabling all other relation types to function.

DNS resolution, certificate validation, authentication protocols, and update mechanisms¹ operate outside focal awareness while shaping the possibility conditions for hermeneutic, embodiment, and alterity engagement. When adversarial actors compromise these background systems, they alter the environmental context within which all subsequent relations operate without ever becoming phenomenologically present to users. Compromised certificate authorities or poisoned DNS resolvers redirect authentication flows and web interactions without demanding the interpretive attention that might reveal discordance.

This constitutive opacity generates what I term *cascading vulnerability*, where background exploitation corrupts the entire relational ecology, thus enabling adversarial configurations to operate across foreground relation types simultaneously while remaining infrastructurally invisible. A single compromised certificate authority enables hermeneutic exploitation through phishing sites displaying ostensibly valid trust indicators. Embodiment exploitation occurs through deepfake platforms appearing legitimately secured. Alterity's relational structure is exploited through fraudulent dating services presenting authenticated identities. Further background exploitation propagates as compromised authentication corrupts update mechanisms and adjacent infrastructural systems.

Phenomenological vulnerability, then, describes structural features intrinsic to successful mediation, not critical thinking failures addressable through individual intervention. Mitigating them requires reconfiguring mediations to resist adversarial appropriation, not merely training users to overcome structural conditions enabling technological world-access.

¹DNS (Domain Name System) maps domain names to network addresses; certificate authorities issue digital certificates verifying site authenticity; trust indicators in browsers (lock icons, certificate badges) signal successful verification. These infrastructures function transparently in ordinary use, which is what makes their compromise consequential for the analysis.

2.3 Dimension 3: Relational Weaponization

By *relational weaponization*, I mean that adversarial intentionality emerges from configured human-technology relations. The term names the coupling of technological affordances, interface designs, user populations, and deployment contexts into assemblages that generate exploitation. This preserves postphenomenology’s ontological commitment that human-technology relations constitute the analytical unit, since weaponization occurs at the relational level. The configured assemblage constitutes the attack instrument. Locating the attack there rules out two reductions: to actors alone, as though the technologies they wield were neutral instruments, a position postphenomenology has always rejected, and to artifacts alone, as though harm resided in a technology’s properties independent of the relations in which it is deployed.

Computer intentionality, as Mykhailov (2023) describes it, forms a triad in which the designer’s intentional states transfer into the computer system and shape user experience, with autonomous learning eroding the causal closure of that transfer. I argue that adversarial intentionality extends this account by introducing a fourth intentional position (Figure 1): the attacker who configures the assemblage from outside the designer-computer-user triad, engineering the conditions under which transfer produces exploitation. This fourth position cannot be reduced to rogue designer intent, corrupted computer autonomy, or naive user misuse. It occupies a distinct ontological station that Mykhailov’s descriptive framework does not require. The extension also operates on a different axis. Where Mykhailov analyzes how computer systems promote or demote values, paradigmatically fairness, adversarial intentionality analyzes engineered discordance between a mediation’s appearance and its

operative function. Where Wellner (2024) describes relegation as imposed delegation users cannot resist, relational weaponization goes further by deliberately engineering the conditions of relegation toward deceptive ends, configuring assemblages that subordinate users while misleading them about the nature of their subordination.

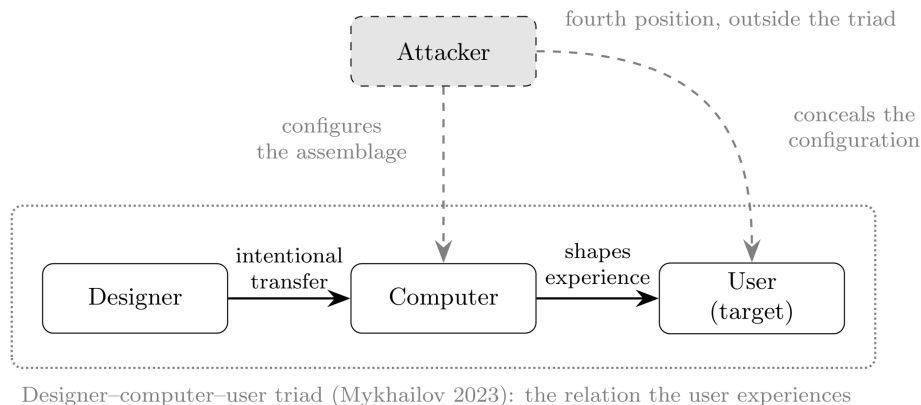


Figure 1: The four intentional positions. Mykhailov’s designer-computer-user triad (solid) is the relation the user experiences: the designer’s intentional states transfer into the computer system, which shapes the user’s experience. The attacker (dashed) occupies a fourth position outside the triad, configuring the assemblage and concealing the configuration from the user, who is the experiencing subject throughout.

This fourth-position account sits within a longer tradition of showing that technologies carry politics, and it extends that tradition. Relational weaponization inherits STS’s foundational insight that technologies embed politics, but radicalizes it. Winner (1980) argued that technologies embody values and settle political questions through design, from bridges built, on Caro’s account, to exclude buses to agricultural technologies that restructured labor relations without any such plot. Latour (1992, 229–32, 243–44) revealed how technical mediators enforce behaviors through material constraints, such as speed bumps and door closers translating moral programs into non-human actants operating persistently across time. Adversarial intentionality extends these insights by theorizing deliberately antagonistic

configurations. Where Winner examined political inscriptions accessible to critical examination and Latour analyzed delegation through material constraints, adversarial intentionality addresses performative mediations engineered to present false appearances while enabling exploitation, weaponizing mediation’s concealment rather than its political character alone. This emphasis on concealment connects adversarial intentionality to Zuboff’s (2019) analysis of surveillance capitalism, in which everyday platforms extract behavioral surplus while concealing their operative aims behind ordinary-seeming interfaces. Surveillance capitalism weaponizes concealment at population scale, mediating social connection while hiding the logic that turns that connection into a resource. Relationality remains the fundamental commitment. The configured assemblage, not actors or artifacts in isolation, constitutes the analytical unit, and this requires investigation of human-technology couplings.

Composite intentionality (Verbeek 2008) provides an ontological foundation for this relational approach in that technologies carry an intentionality of their own, and human and technological intentionalities combine in the coupling, so that the intentionality of an assemblage belongs to neither on its own. Mykhailov and Liberati (2023) extend technological intentionality to generative adversarial networks, whose permanently interacting generator and discriminator components produce an internal complexity, an “inner horizon,” exceeding what programmer, designer, or user can perceive.² Adversarial configurations exploit composite intentionality by engineering assemblages that mislead about their own operation. The result is a *deceptive*

²The term “adversarial” here shares etymology but not conceptual apparatus with “adversarial” in machine learning contexts (e.g., adversarial training, generative adversarial networks). Mykhailov and Liberati (2023) analyze GANs as instances of technological intentionality, using the permanent interaction of generator and discriminator components to illustrate the system’s phenomenological “inner horizon.” My usage of “adversarial” denotes the against-ness of the configured human-technology relation, the fact that the assemblage works against the user’s epistemic interests while appearing to serve them. The terminological overlap is coincidental rather than conceptual, though both usages share the structural insight that adversarial dynamics emerge from relational configurations rather than isolated components.

composite intentionality of coupled components producing false phenomenological presentations that conceal exploitative operative logic. Weaponization occurs through configuration, assembling components with particular affordances, designing interfaces presenting false signifiers, and establishing contexts shaping interpretive presumptions. The resulting coupling generates deception resistant to reflective interrogation.

2.4 Typology of Phenomenological Vulnerabilities

Each human-technology relation type generates characteristic phenomenological vulnerabilities through its distinctive mediating structure, arising from the mechanisms that enable each relation to mediate. Following variational cross-examination (Rosenberger 2020), Table 1 systematizes how structural characteristics of Ihde’s four relation types produce exploitable vulnerabilities.

Table 1: Phenomenological Vulnerabilities by Relation Type

		Adversarial		
Relation		Structural	Phenomenological	Exploitation
Type	Notation	Characteristic	Vulnerability	Mechanism
Embodi- ment	(human-	Pre-reflective	Constitutive opacity:	Exploitation of
	technology)	incorporation	incorporated	habitual fluency,
	→ world	achieving	technologies resist	attacks embedded in
		transparency	conscious	familiar interfaces
		through	interrogation even	operate beneath
		habituation	when compromised	reflective awareness

				Adversarial
Relation		Structural	Phenomenological	Exploitation
Type	Notation	Characteristic	Vulnerability	Mechanism
Hermeneu- tic	human $\rightarrow$ (technology- world)	Interpretive	Semiotic trust:	Fabricated
		engagement with	interpretation	information
		technological	assumes	presented through
		representations	representations	legitimate interfaces,
		presuming	genuinely	false displays exploit
		systematic	correspond to	interpretive
Alterity	human $\rightarrow$ technology- (-world)	correspondence	represented domains	presumption
		Quasi-	Relational	Fabricated presence,
		interpersonal	attribution:	technological
		engagement	interaction patterns	configurations elicit
		elicited by	emerge from	interpersonal
		relational	structural features,	engagement from
		structure rather	not genuine agency	ontologically absent
than ontological		entities		
		status		

				Adversarial
Relation		Structural	Phenomenological	Exploitation
Type	Notation	Characteristic	Vulnerability	Mechanism
Back-ground	human (-	Contextual	Infrastructural	Environmental
	technology-) world	operation outside focal awareness while enabling conditions for action	invisibility and constitutive dependence: background systems shape possibility conditions while resisting foregrounding even during active compromise. Users depend on infrastructures they cannot directly interrogate	manipulation through infrastructure compromise, including DNS poisoning, certificate spoofing, update mechanism hijacking, and algorithmic recommendation manipulation, alters mediation contexts for all subsequent relations while remaining phenomenologically absent

Likewise, Figure 2 visualizes the core structural insight. The diagram contrasts the standard postphenomenological relation, where feedback between human, technology, and world remains open, with the adversarial relation, where the assemblage substitutes a counterfeit world for the one the user expects to encounter. Table 2 specifies how this doubling manifests across each relation type.

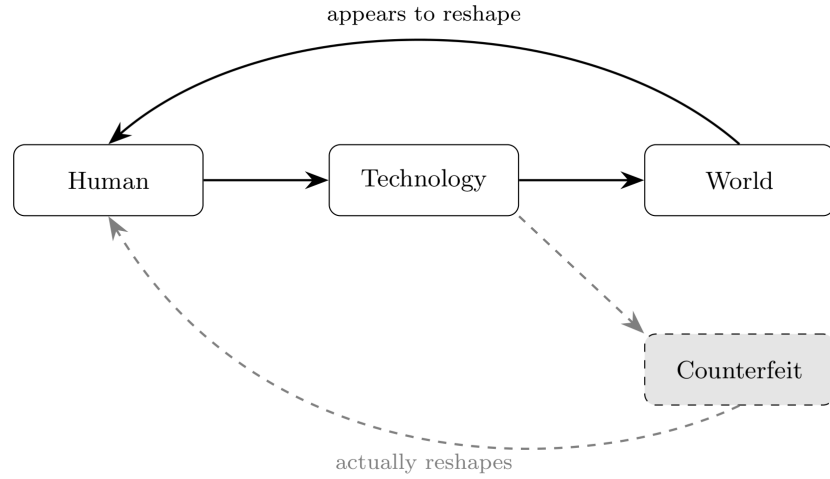


Figure 2: Adversarial intentionality as doubled relation. Solid lines represent the expected relation accessible to the user. Dashed lines represent the operative relation concealed by the assemblage.

Table 2: Expected and operative relations across Ihde’s (1990) four relation types, with “counterfeit” replacing the element through which deception operates

Relation	Expected	Operative
Embodiment	(human-technology) → world	(human-technology) → counterfeit
Hermeneutic	human → (technology-world)	human → (technology-counterfeit)
Alterity	human → technology-(-world)	human → counterfeit-(-world)
Background	human (-technology-) world	human (-technology-) counterfeit

2.5 STS Positioning and Critical Extensions

Adversarial intentionality extends STS insights while revealing their limits for analyzing deliberately deceptive configurations. Where Winner and Latour analyze political inscriptions and delegations whose operative logics remain accessible to investigation, adversarial configurations encode deception as constitutive logic rather than merely biasing outcomes.

Mediation theory argues that technologies shape moral perception and action (Verbeek 2011). Adversarial configurations invert this insight by corrupting moral participation through manufactured false perceptions, deceiving users about their mediated experience. Feenberg's (1991, 79–82, 88) technical code lets domination pass for technical necessity. Adversarial configurations instead engineer a false phenomenological situation, leaving no reliable margin of maneuver for the tactics through which Feenberg locates resistance (pp. 86, 89).

2.6 Preserving Relational Ontology

Adversarial intentionality might appear to violate postphenomenology's commitment to relational ontology by attributing malicious properties to technologies. The framework avoids both technological determinism and instrumentalism by maintaining relationality while extending it to encompass deliberately antagonistic configurations. The analytical unit remains the configured assemblage, and technologies possess no adversarial character in isolation from relational configurations. Deepfake algorithms operating in film production constitute different relations than the same algorithms deployed in romance fraud. The technology's material properties remain constant. Relational configuration determines adversarial character.

This account follows postphenomenology's relational ontology, in which human and technology

constitute one another and neither is reduced to the other (Rosenberger and Verbeek 2015, 19–20; Verbeek 2005, 112–13). Adversarial intentionality reduces neither to social intentions, as if malicious actors simply misused neutral technologies, nor to technological properties, as if certain artifacts inherently enabled deception independent of configuration. Adversarial character emerges from configured relations through what I term *exploitative co-constitution*: attacker, technology, and target jointly constitute the deceptive relation, and the target’s part in it is unknowing, supplied through ordinary skilled use rather than through any participation in the exploitation. Technological affordances shape possible attacks by enabling specific phenomenological manipulations, attack strategies drive technological development by creating demand for sophisticated deceptive capabilities, and human-technology couplings generate composite intentionality that misleads about its own operation. Neither intentions nor properties alone determine adversarial outcomes. The configured relation is analytically primitive, requiring investigation of entire socio-technical assemblages rather than isolated components.

Maintaining such relational ontology preserves postphenomenology’s pragmatist orientation toward empirical investigation. Locating adversarial intentionality in configured relations demands empirical investigation of concrete human-technology couplings, examining how particular affordances, interface designs, and deployment contexts couple to generate exploitation. This orientation distinguishes postphenomenological analysis from essentialist accounts that attribute harm to artifact properties and from constructivist accounts that reduce mediation to social interpretation.

2.7 Societal Implications of Adversarial Intentionality

Phenomenological integrity, as I use the term, names the capacity of technologically mediated experience to maintain recognizable correspondence between its presentations and the domains it purports to represent, the experiential ground on which judgments about what is the case depend. This integrity is a prerequisite for collective deliberation, the shared reasoning through which people who must decide together settle what is true, since deliberation requires that participants can distinguish authentic from fabricated mediation (Rini 2020, 2, 5). Deliberation is itself technologically mediated, as Verbeek argues of political interaction and issue formation (2020, 148, 151). When adversarial configurations undermine this correspondence, they threaten the foundations of shared public life. The challenge extends beyond occasional misinformation to AI-enabled fabrication of entire phenomenological realities indistinguishable from authentic mediation. The shared foundations of collective deliberation are this integrity held in common, the presumption each participant extends to the others' mediated evidence. When synthetic video becomes phenomenologically equivalent to documentary evidence, that presumption fails for the deceived and the undeceived alike (Rini 2020, 8).

Beyond individual exploitation, adversarial intentionality bears on the distributed trust systems that enable social coordination and on the uneven distribution of phenomenological vulnerability across populations. How adversarial mediation threatens the trust sustaining institutional and interpersonal life, and how its harms concentrate in already-marginalized communities, warrants dedicated treatment that I pursue in subsequent work.

3 AI-Enhanced Social Engineering as Paradigmatic Case

AI-enhanced social engineering demonstrates adversarial intentionality’s analytical necessity by embodying the framework’s core insights across multiple phenomenological dimensions. Unlike technical vulnerabilities targeting software implementations or psychological manipulation emphasizing cognitive biases, AI-enhanced attacks systematically weaponize technological mediation. Two characteristics make these attacks philosophically significant.

First, cascading exploitation across relation types reveals phenomenological targeting. Deep-fake manipulations target embodiment relations through synthetic presence, business email compromise exploits hermeneutic relations through interface fabrication, and romance scams weaponize the relational structure of alterity through fabricated intimacy. Each maps onto specific phenomenological vulnerabilities identified in the typology, demonstrating that adversarial intentionality operates through configured relations targeting mediation structure. The vignettes are organized by the relation the attack exploits, which is not always the relation the target experiences. In the romance scam, the target relates through the platform to an apparent person, while the attack works on the structure of alterity.

Second, AI scales adversarial mediation beyond what technical enhancement alone would allow. Voice synthesis fabricates embodied presence indistinguishable from authentic mediation, large language models generate contextually sophisticated phishing at industrial scale, and chatbots presented as people sustain durational relations across extended temporal horizons. AI does not merely enhance existing attacks, instead enabling new adversarial

configurations by automating the construction of deceptive mediation, generating attacks that map comprehensively onto phenomenological relation types. Postphenomenological analysis reveals exploitation patterns invisible to technical and psychological frameworks alike, as the following three vignettes demonstrate.

3.1 Vignette 1: Deepfake Video Calls

Returning to the Arup case, the employee initially suspected email phishing but abandoned that skepticism after seeing and hearing familiar colleagues in what presented itself as a live video conference. The attackers kept up contact by message, email, and one-on-one video calls until, about a week later, he checked with the company’s head office and discovered the deception (Chen and Magramo 2024; Kong 2024). The case reveals systematic exploitation of embodied presence mediation through adversarial intentionality’s three dimensions. Video conferencing platforms mediate remote presence through embodiment relations: users experience quasi-bodily extension, perceiving through technology to remote others rather than at mediated representations. Deepfake synthesis weaponizes this mediating function by fabricating audiovisual coherence that produces identical phenomenological transparency, so that the interface mediates false presence indistinguishable from authentic embodiment.

This exploitation succeeds because embodiment relations achieve transparency through habitual fluency operating beneath conscious awareness. Seeing colleagues’ faces and hearing their vocal patterns generates immediate experiential certainty without interpretive mediation. The pre-reflective trust that “seeing is believing” constitutes embodiment’s operational success, yet it simultaneously generates structural vulnerability. Deepfake attacks exploit exactly this

pre-reflective level. At Arup the fabrication needed no real-time responsiveness. Recognizable faces and voices reading from a script, interacting only to ask him to introduce himself, were enough to produce embodied confirmation resistant to deliberative interrogation (Kong 2024). Neither deepfake technology alone nor video platforms themselves constitute the attack. The weaponized relation emerges from purposeful assemblage coupling synthetic audiovisual generation, video conferencing affordances, and users' habituated embodied trust.

Security awareness training emphasizing verification of unusual requests assumes deliberation can override embodied certainty. Yet embodiment operates pre-reflectively: users experience presence phenomenologically even when reflectively suspicious. Technical detection addresses pixel artifacts and lip-sync anomalies yet misses the phenomenological vulnerability. The framework demonstrates why "seeing" no longer guarantees presence. Technological mediation constructing genuine embodied experience from ontologically absent entities exploits the gap between experiential immediacy and reflective verification.

At broader scale, deepfake exploitation threatens institutional systems depending on audiovisual evidence. This erosion feeds what Chesney and Citron term the "Liar's Dividend": declining confidence in audiovisual authenticity lets individuals evade accountability by claiming evidence is deepfaked, threatening what Horvitz calls a "post-epistemic world, where fact is indistinguishable from fiction" (as cited in Lundberg and Mozeliuss 2025, 2160, 2164). News media, legal systems, and political discourse face parallel credibility crises as synthetic footage becomes indistinguishable from authentic reporting. Adversarial intentionality transforms embodied mediation's transparency from enabling condition into systematic vulnerability, forcing choice between zero-trust verification and acceptance of synthetic presence as undetectable

reality.

3.2 Vignette 2: Business Email Compromise

Business email compromise (BEC), responsible for the billions in losses noted earlier, exemplifies adversarial intentionality through hermeneutic relation exploitation. Typical attacks involve spoofed executive emails authorizing fraudulent wire transfers. An employee receives apparently authentic correspondence from their CFO requesting urgent payment, displaying correct sender details and organizational branding. Attacks succeed not through technical penetration but through phenomenological exploitation of how email interfaces mediate perceived organizational authority.

Adversarial intentionality illuminates how attackers systematically target hermeneutic relations through interlocking mechanisms. Attackers fabricate interface signifiers (sender addresses, visual formatting, organizational branding), weaponizing the interface’s hermeneutic function so that the technology itself becomes complicit in constructing false organizational reality. This exploitation succeeds because hermeneutic relations achieve fluency through interpretive withdrawal, and legitimate-appearing emails produce identical withdrawal, generating false certainty resistant to interrogation. The adversarial relation emerges from purposeful coupling of fabricated interface elements, organizational authority structures, and users’ habituated practices, configuring human-email-organization assemblages toward exploitation.

Hermeneutic relations operate through nested interpretation, as users simultaneously interpret sender identity, organizational authority, message urgency, and action legitimacy. Each

interpretive layer reinforces the others when adversarially configured, and fabricated interface signifiers generate genuine hermeneutic relations mediating false organizational realities.

Beyond individual victims, BEC threatens the institutional infrastructure enabling economic coordination. When employees cannot distinguish genuine from fraudulent executive communications, organizational hierarchies become operationally compromised and supply chain systems face vulnerability as partner communications are indistinguishable from legitimate processes. BEC exploitation reveals how organizational trust systems presume mediation’s hermeneutic reliability, that interface signifiers correspond to organizational realities. Adversarial intentionality transforms this reasonable assumption into systematic vulnerability, demanding either prohibitively inefficient verification or acceptance of structural fraud as routine reality.

3.3 Vignette 3: Romance Scams

Romance scams operate through durational cultivation spanning weeks or months, and I take this extended temporal dimension to mark a feature of adversarial intentionality that the deepfake and BEC cases show only briefly by comparison. Victims frequently cannot recognize deception even after revelation, which suggests how adversarial intentionality exploits an intimate other constructed over time as an apparent person. The technology mediates the ontology of the intimate other itself. The “person” exists solely through platform-mediated construction. What Loewith (2025, 3) calls *generative likeness*, where personal data becomes a mimetic computational proxy simulating social presence beyond the modeled individual’s control, illuminates this transformation, though Loewith develops it for conversational AI.

These scams weaponize this generative capacity with stolen profiles and scraped images, producing interactive entities that engage across extended temporal scales while eliciting genuine emotional attachment. The attack is constituted through sustained assemblage rather than isolated incident.

Variational cross-examination (Rosenberger 2020) exposes what I take to be adversarial intentionality's deepest philosophical challenge. Consider two parallel stabilities the same dating platform mediates. In one, relationship-seekers reach actual others through the platform. In the other, adversarial actors fabricate optimized personas, exploiting identical platform affordances to construct phenomenologically indistinguishable intimate others serving exploitation rather than connection. Cross-examining these stabilities exposes an insight, as within the platform's own mediation no internal criteria distinguish an authentic other from a fabricated one. Both generate genuine emotional attachment through identical relational structures and progressive habituation, producing phenomenologically real intimate presence regardless of ontological status. Detection requires pluralizing mediation across channels the platform does not support, like video calls that introduce embodiment relations, social network triangulation through mutual acquaintances, or physical meetings that test embodied consistency, though the Arup case shows that video can be faked too. None of these operate within the platform's singular mediational frame.

Psychological accounts (Whitty 2013, 678, 682) frame romance scams as persuasion and cognitive exploitation, crediting communication technologies with building trust, but do not ask how technological mediation constitutes the intimate other. The "person" trusted never existed beyond platform affordances. It was pure mediation weaponized across time. I argue

that victims live a genuine relation to an apparent person who is a technological construction, which renders the distinction between authentic and fabricated intimacy phenomenologically impossible from within the mediated relation. Romance scams also sever *social authorship* (Loewith 2025), the individual’s control over self-presentation across social contexts. Stolen photographs and fabricated narratives appropriate modeled individuals’ social identity without consent, cutting them off from control over how their likeness mediates intimate connection. Privacy is the least of what is lost here. The harm to the modeled individual is not the disclosure of information about them but the conscription of their likeness into a relation they never joined, which is why I frame it as a loss of social authorship rather than as a privacy violation. Scam targets simultaneously lose capacity for authentic social authorship, as their vulnerability expressions and emotional disclosures become raw material for adversarial configurations rather than genuine relational exchanges.

These patterns carry systemic consequences extending beyond individual victims. Dating platform ecosystems become structurally vulnerable when the affordances enabling genuine intimacy are equally effective for exploitation, confronting users with verification demands the platform gives them no means to meet. Self-disclosure, the openness that intimacy requires, becomes a competitive disadvantage against scalable fabricated personas, reduced to what Loewith (2025, 1) calls “raw material for synthetic optimization,” a displacement Loewith describes for modeled individuals and that I extend here to the platform as a whole. Those with the least recourse when deceived bear the heaviest losses, and this dynamic transforms digital intimacy platforms into infrastructures that redistribute harm toward those least able to absorb it.

3.4 Cross-Case Patterns in Adversarial Intentionality

The three vignettes demonstrate that adversarial intentionality operates systematically across distinct phenomenological relations and temporal scales. Three patterns emerge from their comparison, namely cascading exploitation across relation types, temporal variability, and vulnerability located at mediation interfaces.

Cascading exploitation across relation types, the first of these patterns, recurs across all three cases. While each vignette foregrounds particular relations (deepfakes through embodied presence, BEC through hermeneutic fabrication, romance scams through a constructed other), all three depend upon background infrastructures and touch multiple relation types. Romance scams exploit alterity's relational structure through fabricated intimacy while also weaponizing hermeneutic relations through platform signifiers and background relations through authentication systems validating fraudulent profiles. The layering itself is ordinary. An established technology such as the car already rests on background infrastructures of roads, fuel stations, and traffic systems while engaging embodiment relations in driving and hermeneutic relations in reading signs and instruments, so that a compromise of the infrastructure would reach every foreground relation at once. This many-to-many pattern demonstrates that adversarial intentionality targets mediation architecture holistically rather than exploiting isolated vulnerabilities.

Temporal variability shows that attackers are bound to no single tempo. Attacks operate across vastly different timescales, with deepfake fraud concentrating its decisive moment in one call (at Arup, within an episode of about a week), BEC exploiting organizational urgency (minutes to days), and romance scams cultivating habituation (weeks to months). This

diversity reveals that adversarial intentionality constitutes adaptable configuration rather than fixed attack pattern.

Vulnerability emerges at mediation interfaces, not from individual deficiencies. Each case reveals exploitation where technologies mediate specific experiential domains, such as video conferencing mediating embodied presence, email mediating organizational authority, dating platforms mediating intimate connection. Background relations provide infrastructural foundation, operating outside focal awareness yet constituting conditions enabling foreground exploits. Traditional security frameworks cannot explain why legitimate-appearing mediations succeed phenomenologically, and psychological accounts miss how technological mediation constitutes the experiential conditions enabling exploitation. Adversarial intentionality reveals how assemblages weaponize mediation structures, generating authentic embodiment relations with synthesized presence, genuine hermeneutic relations mediating false realities, genuine relations to apparent persons, and background relations operating within compromised infrastructure.

4 Limitations, Future Directions, and Defensive Design

I have prioritized postphenomenological investigation while bracketing adjacent approaches (political economy, psychological deception research, computer science authentication), each of which offers complementary insights. The concept offers analytical resources for examining exploitation patterns rather than prescriptive solutions. The view of the practitioners interviewed by Pharris and Perez-Mira (2023, 20–21), that social engineering attacks can be reduced but never eliminated, resonates here. Namely, adversarial intentionality exploits

fundamental structures of human-technology relations that cannot be eliminated without destroying the relations themselves.

The question remains of whether the adversarial inaccessibility I describe is a matter of principle or merely of fact, such that better detection tools or more sustained phenomenological inquiry could eventually penetrate adversarial concealment. Adversarial inaccessibility is neither a metaphysical property of technologies (which would render the claim unfalsifiable) nor a merely contingent limitation (which would reduce it to a temporary engineering problem). It is a designed condition. Attackers engineer adversarial configurations to resist phenomenological interrogation from within the configured relation, and they do so by exploiting affordances that the platform’s designers built for other ends. This inaccessibility is practically irreversible for users operating within the relation as configured, even though it is not irreversible in principle. Out-of-band verification, forensic analysis, or reconfiguration of the mediating assemblage can reveal the deception. The lemniscate model (Kudina 2021) clarifies the stakes with its co-shaping loop, which on my reading assumes that feedback between human, technology, and world remains in principle accessible to reflective examination. Adversarial configurations challenge this assumption by engineering conditions under which reflective access fails systematically within the operative relation. This succeeds *because* the user has no phenomenological reason to initiate the inquiry that would reveal the deception. The question of defensive intervention deserves dedicated treatment. Building on Rosenberger’s (2017) hostile architecture work and Verbeek’s (2011) moral mediation analysis, a defensive postphenomenology might reconfigure amplification-reduction structures, introducing friction where attackers depend on pre-reflective engagement and pluralizing mediations

where exploitation requires singular configurations. Such investigation must address inevitable trade-offs like security friction that diminishes operational fluency or verification requirements that attenuate communicative spontaneity. The distributional justice implications of who bears these costs remain unresolved, as does the relation between adversarial mediation and privacy, which I have bracketed here. Future research could systematically apply variational cross-examination to proposed defensive technologies, interrogating whose interests they serve and how they distribute protective burdens.

For AI governance, adversarial intentionality suggests incorporating phenomenological impact assessment as a governance tool, evaluating how AI systems shape conditions of experience. Such assessment would examine how proposed systems exploit cognitive and temporal pressures and emotional vulnerabilities. Another direction is *phenomenological literacy* as a form of critical technological education, helping users recognize experiential manipulation while maintaining openness to legitimate mediation.

5 Conclusion

Adversarial intentionality extends postphenomenology by revealing an analytical boundary in its conceptual architecture. The tradition's foundational assumption, that technological mediation maintains operative logics accessible through phenomenological examination, encounters a limit case in configurations designed to conceal rather than disclose their mediating operations. These configurations weaponize the phenomenological vulnerability that makes human engagement with technology possible, exploiting the constitutive features enabling mediation to function.

The concept provides vocabulary for understanding how deceptive technologies exploit the structural features of mediation through which humans encounter their technological environment, transforming mediation's enabling conditions into systematic vulnerabilities. This extension preserves postphenomenology's commitment to relational ontology while expanding analytical scope to encompass mediations engineered to present false appearances that conceal exploitative operative logic. Recognizing phenomenological vulnerability as a constitutive feature reveals the depth of human embeddedness in technological relations. Mediation's enabling structures simultaneously generate systematic exploitability.

Beyond postphenomenological theory, adversarial intentionality addresses an analytical gap in technology studies approaches to artificial intelligence. Current frameworks struggle to theorize the phenomenological specificity of AI-mediated deception, often defaulting to discussions of bias or algorithmic accountability that fail to capture the experiential dimensions of adversarial mediation. By theorizing the exploitation of phenomenological vulnerability, the framework I have developed provides conceptual resources for understanding AI systems operating through deliberate phenomenological manipulation. Future work might develop the temporal phenomenology of adversarial mediation, examining how deceptive configurations unfold as trust accumulates and vulnerability deepens through progressive habituation.

At stake are questions about the conditions necessary for shared public life and collective meaning-making under the ubiquity of AI-enhanced deception. Phenomenological integrity, the correspondence between mediated presentations and what they purport to present, emerges as a prerequisite for the deliberation and judgment shared public life requires. The challenge, I contend, is to develop philosophical vigilance adequate to the systematic weaponization of

phenomenological vulnerability without abandoning the structural features of mediation that make meaningful technological engagement possible.

Acknowledgments

During the preparation of this work, the author used Claude (Anthropic) for editing and readability. After using this tool, the author reviewed, edited, made the content their own, and validated the outcome as needed. The author takes full responsibility for the content of the publication.

References

- Aydin, Ciano, Margoth González Woge, and Peter-Paul Verbeek. 2019. “Technological Environmentalism: Conceptualizing Technology as a Mediating Milieu.” *Philosophy & Technology* 32 (2): 321–38. <https://doi.org/10.1007/s13347-018-0309-3>.
- Besmer, Kirk. 2012. “Embodying a Translation Technology: The Cochlear Implant and Cyborg Intentionality.” *Techné: Research in Philosophy and Technology* 16 (3): 296–316. <https://doi.org/http://dx.doi.org/10.5840/techne201216319>.
- Braga, Joaquim. 2023. “On Don Ihde’s Concept of Technological Background Relations.” In *Thinking Togetherness: Phenomenology and Sociality*, edited by Andrej Božič. Institute Nova Revija for the Humanities.

- Chen, Heather, and Kathleen Magramo. 2024. “Finance Worker Pays Out \$25 Million After Video Call with Deepfake ‘Chief Financial Officer.’” CNN, February. <https://www.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html>.
- de Boer, Bas. 2023. “Explaining Multistability: Postphenomenology and Affordances of Technologies.” *AI & SOCIETY* 38 (6): 2267–77. <https://doi.org/10.1007/s00146-021-01272-3>.
- Fair, Lesley. 2024. “‘Love Stinks’ – When a Scammer Is Involved.” Federal Trade Commission, February. <https://www.ftc.gov/business-guidance/blog/2024/02/love-stinks-when-scammer-involved>.
- Federal Bureau of Investigation, Internet Crime Complaint Center (IC3). 2025. *2024 IC3 Annual Report*. Federal Bureau of Investigation. https://www.ic3.gov/AnnualReport/Reports/2024_IC3Report.pdf.
- Feenberg, Andrew. 1991. *Critical Theory of Technology*. Oxford University Press.
- Ferreira, Ana, and Soraia Teles. 2019. “Persuasion: How Phishing Emails Can Influence Users and Bypass Security Measures.” *International Journal of Human-Computer Studies* 125 (May): 19–31. <https://doi.org/10.1016/j.ijhcs.2018.12.004>.
- Gray, Colin M., Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. 2018.

- “The Dark (Patterns) Side of UX Design.” *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC Canada), April, 1–14. <https://doi.org/10.1145/3173574.3174108>.
- Ihde, Don. 1990. *Technology and the Lifeworld: From Garden to Earth*. Indiana University Press.
- Kong, Harvey. 2024. “HK\$200 Million Lost in Deepfake Conference Call Scam in Hong Kong First.” South China Morning Post, February. <https://www.scmp.com/news/hong-kong/law-and-crime/article/3250851/everyone-looked-real-multinational-firms-hong-kong-office-loses-hk200-million-after-scammers-stage>.
- Kranc, Stanley C. 2017. “Picturing the Technologized Background.” *Kunstlicht* 38 (4): 20–28. <https://tijdschriftkunstlicht.nl/vol-38-2017-no-4-mediated-imaginations-technologies-touching-upon-art/>.
- Kudina, Olya. 2021. “‘Alexa, Who Am I?’: Voice Assistants and Hermeneutic Lemniscate as the Technologically Mediated Sense-Making.” *Human Studies* 44 (2): 233–53. <https://doi.org/10.1007/s10746-021-09572-9>.
- Latour, Bruno. 1992. “Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts.” In *Shaping Technology/Building Society*, edited by Wiebe E Bijker and John Law. MIT Press.

- Loewith, Chloe. 2025. “Shhh! Your Proxy Is Speaking: Real Persona Social AI and the Appropriation of Likeness.” *AI & SOCIETY*, ahead of print, November. <https://doi.org/10.1007/s00146-025-02735-7>.
- Lundberg, Ebba, and Peter Mozelius. 2025. “The Potential Effects of Deepfakes on News Media and Entertainment.” *AI & SOCIETY* 40 (4): 2159–70. <https://doi.org/10.1007/s00146-024-02072-1>.
- Magramo, Kathleen. 2024. “British Engineering Giant Arup Revealed as \$25 Million Deepfake Scam Victim.” CNN, May. <https://www.cnn.com/2024/05/16/tech/arup-deepfake-scam-loss-hong-kong-intl-hnk/index.html>.
- Merleau-Ponty, Maurice. 2002. *Phenomenology of Perception: An Introduction*. Routledge.
- Mol, Annemarie. 2002. *The Body Multiple: Ontology in Medical Practice*. Duke University Press. <https://doi.org/10.1215/9780822384151>.
- Mykhailov, Dmytro. 2020. “The Phenomenological Roots of Technological Intentionality: A Postphenomenological Perspective.” *Frontiers of Philosophy in China* 15 (4): 612–35. <https://doi.org/10.3868/s030-009-020-0035-6>.
- Mykhailov, Dmytro. 2023. “Philosophical Inquiry into Computer Intentionality: Machine Learning and Value Sensitive Design.” *Human Affairs* 33 (1): 115–27. <https://doi.org/10.1007/s11266-023-0035-6>.

[1515/humaff-2022-2035](#).

Mykhailov, Dmytro, and Nicola Liberati. 2023. “A Study of Technological Intentionality in C++ and Generative Adversarial Model: Phenomenological and Postphenomenological Perspectives.” *Foundations of Science* 28 (3): 841–57. <https://doi.org/10.1007/s10699-022-09833-5>.

Pharris, Lily, and Begona Perez-Mira. 2023. “Preventing Social Engineering: A Phenomenological Inquiry.” *Information & Computer Security* 31 (1): 1–31. <https://doi.org/10.1108/ICS-09-2021-0137>.

Rini, Regina. 2020. “Deepfakes and the Epistemic Backstop.” *Philosophers’ Imprint* 20 (24): 1–16. <https://philarchive.org/rec/RINDAT>.

Rosenberger, Robert. 2014. “Multistability and the Agency of Mundane Artifacts: From Speed Bumps to Subway Benches.” *Human Studies* 37 (3): 369–92. <https://doi.org/10.1007/s10746-014-9317-1>.

Rosenberger, Robert. 2017. *Callous Objects: Designs Against the Homeless*. University of Minnesota Press.

Rosenberger, Robert. 2020. “On Variational Cross-Examination: A Method for Postphenomenological Multistability.” *AI & SOCIETY* 38 (6): 2229–42. <https://doi.org/10.1007/>

s00146-020-01050-7.

Rosenberger, Robert, and Peter-Paul Verbeek. 2015. “A Field Guide to Postphenomenology.” In *Postphenomenological Investigations: Essays on Human-Technology Relations*, edited by Robert Rosenberger and Peter-Paul Verbeek. Postphenomenology and the Philosophy of Technology. Lexington Books. <https://research.utwente.nl/en/publications/a-field-guide-to-postphenomenology>.

Shneiderman, Ben. 2000. “Designing Trust into Online Experiences.” *Commun. ACM* 43 (12): 57–59. <https://doi.org/10.1145/355112.355124>.

Verbeek, Peter-Paul. 2005. *What Things Do: Philosophical Reflections on Technology, Agency, and Design*. Penn State University Press.

Verbeek, Peter-Paul. 2008. “Cyborg Intentionality: Rethinking the Phenomenology of Human-Technology Relations.” *Phenomenology and the Cognitive Sciences* 7 (3): 387–95. <https://doi.org/10.1007/s11097-008-9099-x>.

Verbeek, Peter-Paul. 2011. *Moralizing Technology: Understanding and Designing the Morality of Things*. University of Chicago Press.

Verbeek, Peter-Paul. 2020. “Politicizing Postphenomenology.” In *Reimagining Philosophy and Technology, Reinventing Ihde*, edited by Glen Miller and Ashley Shew. Springer

International Publishing. https://doi.org/10.1007/978-3-030-35967-6_9.

Wellner, Galit. 2020a. “Postphenomenology of Augmented Reality.” In *Relating to Things: Design, Technology and the Artificial*, edited by Heather Wiltse. Bloomsbury Publishing Plc. <https://doi.org/10.5040/9781350124288>.

Wellner, Galit. 2020b. “The Multiplicity of Multistabilities: Turning Multistability into a Multistable Concept.” In *Reimagining Philosophy and Technology, Reinventing Ihde*, edited by Glen Miller and Ashley Shew. Springer International Publishing. https://doi.org/10.1007/978-3-030-35967-6_7.

Wellner, Galit. 2024. “A Postphenomenological Guide to AI Regulation.” *Journal of Human-Technology Relations* 2 (October). <https://doi.org/10.59490/jhtr.2024.2.7397>.

Whitty, Monica T. 2013. “The Scammers Persuasive Techniques Model: Development of a Stage Model to Explain the Online Dating Romance Scam.” *The British Journal of Criminology* 53 (4): 665–84. <https://doi.org/10.1093/bjc/azt009>.

Whyte, Kyle Powys. 2015. “What Is Multistability? A Theory of the Keystone Concept of Postphenomenological Research.” In *Technoscience and Postphenomenology: The Manhattan Papers*. Postphenomenology and the Philosophy of Technology. Lexington Books.

Winner, Langdon. 1980. “Do Artifacts Have Politics?” *Daedalus* 109 (1): 121–36. <https://www.jstor.org/stable/20024652>.

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. First edition. PublicAffairs.