KSTARR Enterprises, LLC

"Negligent Digital Access™" is an expert opinion introducing a new legal theory: how platform design choices enable foreseeable harm before legal protections apply.

Negligent Digital Access™ Before the Shield: Section 230's Legal Blindspot

By Katherine Starr

Trust Isn't the Problem. Design Is.

In the physical world, we install locks. In the digital world, we post disclaimers. Platforms shape the outcome long before harm becomes visible. The moment a user logs in, creates a profile, or gains access to another person's trust, the platform could have made a series of decisions to determine whether harm would be foreseeable or preventable, but chose not to in that moment. These early-stage choices, made long before a post is published or a message is sent, are where negligence quietly takes root. And yet, when the damage surfaces, platforms retreat behind the legal shield of Section 230 as though their only role was to host content not to design the path that led directly to harm.

We must begin with a basic contrast between the protections we expect in the physical world and those absent from our digital lives. In the real world, we do not grant strangers unverified access to secure buildings, children's camps, or professional environments without some form of screening, reference, or credential. Yet in the digital world, users are handed access to private communication, vulnerable populations, and intimate spaces with no meaningful vetting or control. The inconsistency is not just a policy failure it is a design failure.

As someone who has worked on more than thirty cases over the past thirteen years involving institutional negligence from youth sports to university hiring practices, I have seen one pattern repeat with alarming consistency. The harm does not begin at the moment of abuse. It begins earlier, during the phase no one scrutinizes: the access point. The moment someone is allowed in. The systems we are taught to rely on platforms, schools, hiring bodies are often built not to prevent harm but to respond to it later, after the damage has already been done.

© 2025 KATHERINE STARR | WWW.KATHERINESTARR.COM | ALL RIGHTS RESERVED

Section 230 of the Communications Decency Act is frequently cited as the reason victims cannot hold platforms accountable. But Section 230 protects platforms from liability after content is created not for what they do beforehand. This distinction is not just legal; it's moral. Because when platforms allow bad actors in without meaningful screening, tracking, or intervention, they are not passive hosts. They are enabling a foreseeable harm. And the law does not protect that. Nor should it.

To address the current crisis of trust in digital and institutional environments, we must shift our lens. The core problem is not that people trust too much it's that systems are designed to appear safe while remaining structurally negligent. Trust isn't the weakness. It's the target. And that changes everything about where accountability belongs.

Defining Negligent Digital Access

Negligent Digital Access refers to the failure of a digital platform to act on foreseeable risk at the point of access before any content is generated, before harm is done, and before legal immunity through Section 230 is triggered. This form of negligence is rooted in design choices, not content moderation. It begins when platforms choose not to implement safeguards, they already have access to: identity verification tools, behavioral pattern recognition, IP and device flagging systems, complaint-driven pattern tracking, or trust indicators that reflect real accountability.

This theory is not based on wishful enforcement or idealistic standards. The digital infrastructure to detect bad actors already exists. AI-driven fraud detection, facial verification, real-time risk scoring, metadata analysis, and other protective mechanisms are not aspirational they are active across the financial, insurance, and cybersecurity industries.

Some may argue that implementing full verification at scale is unrealistic or privacy-invasive. But risk-reduction need not mean surveillance it means using already-available, ethically deployed tools to prevent repeated harm.

Platforms are not failing because they lack the tools. They are failing because they have chosen not to apply them at the most critical moment: when trust is first extended.

Negligent Digital Access is therefore not just about who enters, but about what the platform does or refuses to do when granting access. The digital environment is not a neutral street corner. It is a constructed architecture where emotional vulnerability, private information, and relational trust are part of the transaction. And where platforms simulate safety without structurally supporting it, they are no longer passive intermediaries. They are facilitators of foreseeable harm.

Disclaimers Do Not Displace Duty

Digital platforms frequently rely on broad disclaimers embedded in their terms of service to preempt responsibility. They may warn users that "harm can occur," that "verification is not

© 2025 KATHERINE STARR | WWW.KATHERINESTARR.COM | ALL RIGHTS RESERVED

guaranteed," or that interactions are "at your own risk." These statements are designed to insulate platforms from liability, suggesting that their only obligation is to host content not manage the consequences of the connections they enable.

But this defense collapses under the weight of their own design. Platforms do not operate as neutral bulletin boards; they structure interactions, curate access, and create systems that invite emotional, financial, and personal trust. When platforms apply labels such as “verified,” use trust-building language like “safe community,” or enable repeat interactions through design features that prioritize engagement, they are no longer acting as passive hosts. They are active participants in the environment they have engineered.

You cannot disclaim responsibility while simultaneously constructing a system that appears secure. A warning buried in fine print does not negate the fact that platforms profit from a user’s belief in safety especially when that belief is based on engineered cues, not reality. In this way, the digital environment begins to resemble the physical: just as a hotel cannot advertise a secure resort and then blame a guest for injury in an unlit hallway, a platform cannot project trust while doing nothing to maintain it. Legal disclaimers cannot erase foreseeability, and they cannot substitute for the duty of care owed when systems are designed to facilitate vulnerable interactions.

Evolution from Negligent Dating and the Roots in Institutional Negligence

Negligent Digital Access builds upon the foundation laid by the Negligent Dating theory, which identified how platforms that invite emotional proximity without structural safeguards create conditions for foreseeable harm. But its legal roots stretch deeper—into doctrines long established in physical-world liability: *negligent hiring and negligent supervision.*

In the offline world, when an institution fails to vet someone before placing them in a position of trust—like a teacher, coach, or caregiver they can be held liable for *negligent hiring*. When that institution ignores red flags or fails to act on warning signs once harm is foreseeable, they can be held accountable for *negligent supervision*. These doctrines exist because the law recognizes a simple truth: institutions that create access to others especially vulnerable populations have a duty to assess risk before granting entry, and a duty to respond to patterns of harm once they emerge.

The same is true in digital spaces. Platforms aren’t simply neutral hosts; they are architects of access. Whether it's a dating app that allows known scammers to create new profiles, or a freelance marketplace that reinstates previously reported users without review, the failure is not just reactive it is structural. The digital architecture is designed to allow anyone in, maintain frictionless onboarding, and scale trust without accountability. And just like in a workplace, when that access results in foreseeable harm, the liability should follow.

© 2025 KATHERINE STARR | WWW.KATHERINESTARR.COM | ALL RIGHTS RESERVED

Negligent Digital Access reframes platform responsibility not through the lens of speech, but through the lens of institutional conduct. It asks not what the user posted or said, but what the platform knew or should have known when it granted access, and what it failed to act on afterward. The theory mirrors the logic of employment law, where enabling access to others without due care is a breach of duty, regardless of the employee's words. This is not a reinvention of liability it is a digital translation of standards that have long governed schools, churches, employers, and care facilities.

This evolution matters because it positions digital negligence within a lineage of legal reasoning, rather than outside it. It anchors new harm in familiar frameworks and gives courts, attorneys, and policymakers a roadmap forward one that connects emotional betrayal in dating platforms to broader patterns of system-enabled harm, and ultimately, to the duty of access control that every institution, digital or physical, must bear.

Use Cases and Institutional Parallels

Negligent Digital Access is not a theory reserved for dating platforms alone. It applies across a range of digital environments where trust is extended without verification, where access is granted without scrutiny, and where harm is both foreseeable and preventable but not prevented. From freelance marketplaces to social media platforms to virtual event environments, the problem is not only what users do. The problem is what platforms allow and continue to allow despite patterns, reports, and available safeguards.

Take gig economy platforms as one example. A user hires a web developer or virtual assistant through a platform that promises safety, verified reviews, and curated professionals. The platform, however, does not verify identity beyond a working email. It does not flag when a previously reported individual returns under a new account. It does not screen for prior abuse reports or escalate based on repeated complaint patterns. If that contractor gains admin-level access to the client’s website, injects malicious code, or compromises customer data, the harm is not just in the act. it’s in the platform’s architecture that gave that person access and trust, despite warning signs it chose not to use.

We see the same in social media. Platforms routinely allow users who have been banned for harassment or predatory behavior to return with new accounts sometimes with the same name or photo. These platforms are not unaware. They choose not to track IPs, cross-reference behavioral flags, or structure onboarding around safety. Instead, they prioritize reengagement over restriction. This is not the oversight of a neutral host it is a systemic design flaw with real-world consequences.

Even in large digital events and conferences, platforms often promise curated networking and safety protocols while doing little to nothing to vet participants. As with physical events, the failure is not only in the moment of harm it is in the failure to control the conditions under which trust is created. In the real world, we would never allow unscreened individuals into a school

© 2025 KATHERINE STARR | WWW.KATHERINESTARR.COM | ALL RIGHTS RESERVED

overnight camp. Yet digitally, platforms do exactly that—placing vulnerable users in spaces with strangers under the guise of connection, all without a single meaningful safeguard.

What ties all of these examples together is the structure of access. Harm arises not simply from what users say or do, but from what they are allowed to do, and where, and with whom—by design. Negligent Digital Access theory focuses the legal and moral lens not on individual bad actors, but on the systems that repeatedly fail to stop them from walking through the same open door.

The Legal Gap: Where Section 230 Ends

Section 230 of the Communications Decency Act has long been cited as a foundational protection for online platforms. It was written to prevent platforms from being treated as the publisher or speaker of user-generated content. In its original intent, it helped the internet grow without holding every site liable for what individuals posted. But it was never meant to shield platforms from liability for their own conduct and especially not for the design of access systems that facilitate foreseeable harm.

The key distinction is this: Section 230 protects platforms from what users say. It does not protect them from what platforms do. And the act of granting access knowingly, repeatedly, and without using available tools to reduce risk is a design decision. It's not content. It's not speech. It's conduct. And that is where the statute ends.

Courts have already begun to acknowledge this boundary. In *Doe v. Internet Brands*[1], the Ninth Circuit allowed a claim to proceed against a platform not for what it failed to take down, but for what it failed to do in the face of known risk. The plaintiff argued that the site knew predators were targeting users for offline exploitation and yet failed to warn or intervene. The court agreed: the duty to warn was based on platform knowledge and conduct, not user content. Section 230 did not apply.

This is the foothold for Negligent Digital Access. It builds on the principle that when a platform designs access points, collects user data, receives complaints, or enables reentry by bad actors it is no longer neutral. It is exercising judgment, making structural decisions, and shaping outcomes. That conduct is not shielded by 230. It never was.

The legal system has yet to fully recognize the digital equivalent of negligent hiring or supervision, but the logic is already there. No school could defend a hiring decision by saying, "We didn't know what the teacher would say to the students." The liability isn't about speech it's about access. In the same way, platforms should not be allowed to grant digital access to intimate or vulnerable spaces and then plead neutrality once harm occurs.

[1] *Doe v. Internet Brands, Inc.*, 824 F.3d 846 (9th Cir. 2016). The court held that Section 230 does not shield a platform from liability when the claim is based on failure to warn users of known offline risks, not on user-generated content.

© 2025 KATHERINE STARR | WWW.KATHERINESTARR.COM | ALL RIGHTS RESERVED

This is not an attack on free speech. It's a call to clarify the limits of a law written in the 1990s, long before platforms became engineered systems of trust. Section 230 should not and legally does not extend to the structural decisions that enable foreseeable harm before speech even begins.

The Physical vs Digital Standard

In the physical world, access is managed, regulated, and often enforced by law. We require background checks for teachers and camp counselors. We expect security protocols for public events, vetting for medical professionals, and credentialing for those placed in positions of authority or intimacy. Institutions that ignore these responsibilities are held liable—not because they caused the harm directly, but because they created the conditions that allowed it.

The digital world, by contrast, has been allowed to function in a vacuum. Platforms claim the authority to organize human interaction at scale—socially, romantically, financially—without any of the responsibilities that would be demanded in an offline context. They build systems that invite vulnerability while declining to implement the most basic safeguards. A platform that encourages emotional connection or grants backend access to a user's business infrastructure is not offering a neutral product it is creating a proximity environment. And in any other domain, proximity without safeguards equals liability.

The result is a fractured accountability structure: in person, we expect screening. Online, we're told to read the fine print. But the risks are not different they're simply harder to see. When digital spaces mirror the intimacy and exposure of real-world interactions, the standard of care must rise to meet the same threshold. If we expect it of a school, a hospital, or a hotel, we must expect it of a platform that inserts itself into lives at equal or greater depth.

The Platform's Role in Shaping Harm

Platforms do not simply host users they shape behavior. They guide who we see, what we engage with, who gets visibility, and who gets silenced. They deploy behavioral nudges, curated trust signals, and algorithmic reinforcement to steer attention. These are not neutral choices. They are design decisions. And with those decisions comes a duty of care.

When a platform promotes a profile with a "*verified*" badge but offers no meaningful verification, it is projecting safety while providing none. When it reinstates a banned user without screening or tracking, it is enabling harm in a predictable cycle. The platform's role is not passive it is participatory. It builds the environment, decides who gets in, and reinforces the illusion that trust is earned when in fact it is merely implied.

In architecture, form follows function. In law, liability follows foreseeability. The structure of digital access creates the function of digital harm. And platforms that design for scale,

© 2025 KATHERINE STARR | WWW.KATHERINESTARR.COM | ALL RIGHTS RESERVED

engagement, and vulnerability without safeguards are not innovating they are exposing. Their choices shape the harm before it happens. Their systems decide how easily it spreads.

Trust as the Target

Trust is not the flaw in these systems. It is the target. Platforms are built to gain user trust quickly, scale it widely, and convert it into engagement, data, and revenue. But when that trust is misplaced when users are harmed by others who never should have been allowed through the gate the platform shrugs and points to a statute that was never meant to protect those choices.

Negligent Digital Access is not a theory of content. It is a theory of structure. It says that when platforms allow access without care, patterns without intervention, and proximity without protection, they are no longer just carriers of speech. They are enablers of harm.

Section 230 was never meant to shield negligence. Trust, when systemically exploited, is not a design flaw it's a liability strategy. It's time the law recognized the architecture of access as the architecture of harm.

For law firms, policymakers, or organizations exploring liability, prevention, or institutional response models for online platforms, I provide strategic guidance and expert analysis grounded in lived experience and legal relevance. If you're thinking about this space even quietly I invite you to start the conversation now.

Citation and Attribution

Negligent Digital Access: When Platform Design Enables Harm
By Katherine Starr (2025)

To cite this work in academic, legal, or professional contexts, please use one of the following formats:

APA (7th Edition)

Starr, K. (2025). Negligent digital access: When platform design enables harm. KStarr Enterprises, LLC. https://www.katherinestarr.com/negligent-digital-access/

© 2025 KATHERINE STARR | WWW.KATHERINESTARR.COM | ALL RIGHTS RESERVED

MLA (9th Edition)

Starr, Katherine. Negligent Digital Access: When Platform Design Enables Harm. KStarr Enterprises, LLC, 2025. www.katherinestarr.com/negligent-digital-access/.

Bluebook (Legal)

Katherine Starr, Negligent Digital Access: When Platform Design Enables Harm, KStarr Enterprises, LLC (2025), https://www.katherinestarr.com/negligent-digital-access/.

© 2025 Katherine Starr | All rights reserved.
For permissions or media inquiries, contact: info@katherinestarr.com

© 2025 KATHERINE STARR | WWW.KATHERINESTARR.COM | ALL RIGHTS RESERVED