

Digital Trauma in the Age of Artificial Intelligence

The rapid expansion of artificial intelligence has transformed everyday human interactions, communication, and decision-making processes, this is simultaneously giving rise to new forms of digital and psychological trauma. Trauma has been identified as not merely a personal state, but a social state as mentioned by Erikson, K. (1976). This paper examines how AI-driven technologies such as surveillance systems, social media algorithms, and algorithmic profiling contribute to experiences of anxiety, emotional distress, and digital vulnerability. Zuboff, S. (2019) have clearly stated that technological systems do not merely process data; they actively shape human vulnerability and lived experience. This study relies of trauma theory and critical technology studies, to analyse the ways in which constant monitoring, feedback and data extraction intensify power imbalances. The paper explores the research in the field and proposes the need for trauma-informed approaches to the development and implementation of AI technologies, emphasizing empathy, transparency, user autonomy, and ethical accountability.

Keywords: Digital Trauma, Psychological Trauma, Surveillance system, Artificial Intelligence, Critical technology studies

The expansion of artificial intelligence has become an ordinary feature of contemporary life. Algorithmic systems now mediate communication, learning, work, governance, and social interaction in ways that are often invisible yet have far reaching consequences. Although Artificial Intelligence has received great scholarly attention, most of the focus is on the efficiency, innovation, and ethical governance. It is clear based on exploring available literature that less attention has been paid to the psychological and social costs that accompany sustained exposure to automated decision-making and digital surveillance that has become a widespread use of AI in the digital age. These problems may not be immediately identified as dramatic or catastrophic events, instead, they are accumulated slowly and quietly, shaping everyday experience in multiple ways. These lead to a constant fear that leads to

anxiety, uncertainty, self-censorship, and a persistent sense of being watched and evaluated. The idea of digital trauma is identified in this ordinary, ongoing condition of anxiety, uncertainty and self-censorship.

Trauma is an area that has been extensively researched and for long is understood not merely as an individual psychological response to extreme events, but as a condition that can emerge through social structures and collective experiences. Erikson's (1976) sociological account situates trauma within damaged social relations and broken trust. Caruth's (1996) psychoanalytic work emphasizes the temporal disjunction between experience and understanding. The analysis of complex trauma by Herman (1992) further extends this framework by emphasising the conditions of prolonged exposure, power imbalance, and the impossibility of escape. These three theories are considered to provide the lens through which contemporary digital life can be examined, where systems that use Artificial Intelligence have the power to monitor, evaluate and predict continuously and the fact that makes it more dangerous is that all this often happens without explanation.

This paper approaches digital trauma as a qualitative, socio-technical condition that could be explored and understood using the trauma theories proposed by Caruth, Erikson and Herman. Drawing on trauma theory, and digital humanities, it argues that artificial intelligence restructures everyday life in ways that mirror the dynamics of traumatic experience: opacity, loss of agency, chronic vigilance, and erosion of trust. The four factors: opacity, loss of agency, chronic vigilance, and erosion of trust together constitute the experiential core of digital trauma. Opacity in the age of AI, refers not only to technical unreadability but the impossibility to understand how decisions or choices about one's life are made, leaving individuals unable to explain, contest, or narrate their own experiences of harm (Caruth, 1996; Pasquale, 2015; Fricker, 2007). This opacity directly reduces the capacity and the power of humans. In this digital context, humans are positioned as objects of evaluation rather than as subjects or participants in decision-making processes. This state of humans emerging as the object in the digital age who is being observed and analysed leads to a sense of powerlessness similar to that observed in traumatic contexts where control is structurally denied (Herman, 1992; Maier & Seligman, 2016). Digital trauma is not explored as an exceptional, unexpected or pathological response, it is rather seen as an ordinary feature of life in the digital age. The common conditions of continuous surveillance and prediction using artificial intelligence are seen as the backdrop in which digital trauma is studied.

The state of continuous surveillance and predictive analytics that are a part of the AI induced digital age intensify the effects by shattering the boundary between present action and anticipated future behaviour. It can be seen that AI systems do not merely observe what individuals do; they infer intentions, forecast risks, and pre-emptively shape opportunities through scores and classifications that remain largely incontestable (Eubanks, 2018; O'Neil, 2016). This situation is seen everyday in the social media platforms, AI systems go beyond showing users what they clearly choose to see. The systems track likes, pauses, shares, and scrolling speed to infer interests and intentions. Based on these inferences, the systems predict what the user prefers and pre-emptively curates the feed. It is common that over time, certain posts, viewpoints, or opportunities are amplified, while others quietly disappear. At times the user becomes the object where he has to choose from what is curated and offered to him. The user is also aware of why certain pages are suggested while others are ignored, this lack of clarity is considered as the opaque nature of the digital age. The user is mostly not aware of why a particular content is shown or hidden, and sometime the user or the master is left helpless with no clear way to challenge or override these decisions. Thus, it can be observed that slowly, the system shapes attention, choices, and behaviour, not by direct compulsion, but by structuring what appears possible or visible at the background. This process indirectly results in making the user feel helpless in the digital interactions. This shift of the digital age from observation to prediction transforms everyday life into a space of permanent exposure, where individuals are not only required to manage what they are currently doing but they are also forced to manage what they might be assumed to do next (Zuboff, 2019; Cheney-Lippold, 2017). Thus, such environments which have become part of our day-to-day life in the digital age have changed surveillance and monitoring from episodes to a constant state, producing a sustained psychological condition rather than a momentary intrusion (Andrejevic, 2007; van Dijck, 2014).

Cathy Caruth's work is particularly useful in understanding how harm and trauma operate in this system. Algorithmic systems are not disconnected from social psychology as they are technologies that use computer programs to collect information about people, analyse it, and make judgments or predictions that influence real-life decisions. These systems are not just bits of code that work separately; they are part of larger systems that include institutions, policies, and economic interests. In everyday life, it can be observed that algorithmic systems decide majority of our digital interactions including the videos or posts seen online, assess students' performance, screen job applications, track workplace productivity, and determine

97 access to services. Although the information and details are mostly presented as neutral or
98 objective, in most situations, systems reflect human choices, priorities, and power
99 structures. This clearly shows how the society or users are evaluated and treated mostly
100 without their knowledge or consent. For Caruth (1996), trauma is not defined by the
101 magnitude of an event but by the subject's inability to fully grasp the experience as it occurs.
102 Trauma returns belatedly, disrupting memory, language, and narrative coherence. Thus, what
103 wounds is not only what happens, but what cannot be immediately known or articulated. This
104 emphasis on belatedness resonates strongly with algorithmic decision-making, where
105 individuals are frequently subjected to outcomes whose logic remains inaccessible. The
106 monitoring and future predictions of AI has evolved to a level where the users become
107 puppets and are made to act based on the selected options that are provided by AI analysis.

108 When a student is flagged by an automated proctoring system, when a job applicant is
109 rejected through algorithmic screening, or when welfare eligibility is denied by an automated
110 system, the experience of harm is often immediate but unintelligible. The decision might
111 appear to be objective but yet opaque and comes from a system that does not provide a proper
112 explanation and there are no avenues to challenge or question the decision. Individuals are
113 left to infer reasons, often internalizing blame or anxiety without knowing what triggered the
114 outcome. As Pasquale (2015) argues, such black-boxed systems institutionalize opacity,
115 producing forms of power that are difficult to challenge precisely because they resist
116 understanding.

117 This gap between experience and comprehension is not accidental. It is structurally
118 embedded within algorithmic design, reinforced by proprietary secrecy, technical complexity,
119 and institutional distancing. In this sense, algorithmic harm reproduces the temporal structure
120 of trauma identified by Caruth: injury without immediate meaning, followed by anxiety,
121 repetition, and a fractured sense of agency. The individual knows something has gone wrong
122 but cannot locate or narrate the cause. This condition is not episodic but persistent, replayed
123 each time one interacts with systems that evaluate without explanation.

124 While Caruth illuminates the psychic dimensions of this experience, Erikson's sociological
125 account draws attention to its collective implications. In *Everything in Its Path*, Erikson
126 (1976) describes trauma as damage to the "social tissue" that binds communities together.
127 Trauma, in this sense, is not reducible to individual suffering but manifests as a breakdown of

trust, shared meaning, and institutional reliability. When the structures that organize social life become sources of harm, entire communities experience a form of collective disruption.

Digital environments have evolved from being a tool or assistant to critical social infrastructures. The digital environment can be considered as critical social infrastructures because they organize many of the basic activities through which people participate in everyday life. The influence of the digital environment can be seen in every walk of life including education, employment, healthcare, communication, governance, and social relationships. All the major social infrastructures increasingly depend on digital platforms and systems. It can be seen that learning takes place through online classrooms and assessment tools, work is managed through productivity software and algorithmic monitoring, public services are accessed through digital portals, and social connections are maintained through platform-mediated interaction. These environments are no longer optional or supplementary; they shape how individuals encounter institutions, exercise rights, and form social identities. They have become essential components like the physical infrastructure such as roads or electricity. It has been observed that digital systems structure access, distribute resources, and define possibilities for participation, making them central to contemporary social life rather than neutral technological add-ons (van Dijck, 2014; Couldry & Mejias, 2019). Artificial Intelligence mediates education, employment, healthcare, and civic participation, shaping how individuals relate to institutions and to one another. The use of technology in every aspect of life is being normalized within these environments. The idea of privacy, autonomy, and accountability are changing and getting transformed very fast. Apart from the social infrastructures, even individual are connected and monitored through the digital environment.

Educational spaces offer a clear illustration of this process. The increasing use of AI-driven surveillance tools in online learning environments has altered how students experience education. Platforms that track attention, monitor eye movement, and flag “suspicious behaviour” transform classrooms into sites of continuous evaluation. Learning becomes less about intellectual exploration and more about compliance with algorithmic norms (Swauger, 2020; Dutta, 2021). The trust between students and teachers, willingness to take intellectual risks, and collaborative learning which are the core pillars of education begins to fall under the modern digital system. What emerges is not a single traumatic incident but a gradual normalization of surveillance that damages collective educational life.

Judith Herman's work on complex trauma offers further insight into the dynamics of digital environments. Unlike models that focus on discrete traumatic events, Herman (1992) examines trauma that results from prolonged exposure to harm in situations where escape is limited or impossible. Complex trauma is marked by chronic anxiety, hypervigilance, emotional numbing, and erosion of self-trust. Complex trauma develops when stress or harm is not a single event but something a person is exposed to again and again over time. It often shows up as constant anxiety, where the person feels uneasy or on edge even when there is no immediate danger. The need to constantly check messages and updates is a simple reference to constant anxiety. Hypervigilance means always being alert, watching oneself and the surroundings carefully, as if something might go wrong at any moment. Smartwatches and self-monitoring devices that track steps, heart rate, sleep, and activity can also produce forms of hypervigilance. Instead of moving naturally, individuals begin to constantly check their step count, worry about whether they have met daily targets, or feel uneasy when the numbers are low. A slow day, illness, or rest can start to feel like failure, even when the body needs recovery. Over time, people may trust the device more than their own physical sensations thereby pushing themselves to walk more despite fatigue or guilt. This shows how continuous self-monitoring can turn everyday bodily awareness into constant self-surveillance and anxiety. Emotional numbing happens when feelings become muted or disconnected, not because the person does not care, but because shuffling down feels safer than being overwhelmed. Over time, this leads to an erosion of self-trust, where individuals begin to doubt their own judgment, reactions, and feelings, relying instead on external signals or authorities to tell them whether they are safe, acceptable, or doing "well enough." Emotional numbing can be seen when individuals stop reacting strongly to experiences that would normally evoke feeling. For instance, a person who is constantly evaluated by performance metrics or online feedback may no longer feel excitement at praise or disappointment at criticism; both begin to feel flat and distant. A student repeatedly marked low by learning analytics systems may stop feeling anxious or hopeful about outcomes and instead become emotionally detached, going through tasks mechanically. This numbing is not a lack of care, but a protective response when emotional investment repeatedly leads to stress or disappointment, withdrawing feelings becomes a way to cope and continue functioning.

Algorithmic surveillance closely resembles these conditions. Monitoring is continuous rather than episodic, embedded within platforms that are difficult or impossible to avoid without

withdrawing from essential aspects of social and institutional life. Learning management systems record every click, workplace software tracks productivity minute by minute, and social media algorithms monitor and shape interaction continuously (Andrejevic, 2007; Lupton & Michael, 2017). Over time, this produces a state of heightened vigilance: individuals learn to anticipate evaluation, censor themselves, and adapt behaviour to invisible standards. Invisible standards refer to the hidden rules and expectations by which people are judged, even though those rules are never clearly stated or explained. In digital and algorithmic systems, individuals are evaluated according to criteria they cannot see, understand, or question. Posting a picture on social media and tracking its likes or reach illustrates how invisible standards operate in digital spaces. Users are evaluated by opaque algorithmic criteria they cannot see or understand, yet they often internalise these outcomes as measures of personal worth. Over time, repeated exposure to unexplained judgment fosters self-monitoring, anxiety, and emotional vulnerability. Because these standards are invisible, people are left guessing what counts as “good,” “normal,” or “acceptable,” leading them to constantly adjust their behaviour without knowing whether it will actually meet the system’s expectations. Over time, living under invisible standards creates anxiety, self-doubt, and chronic self-monitoring, as individuals try to align themselves with measures that remain out of reach and beyond their control. Herman emphasizes that complex trauma is enabled by social structures that normalize harm. This insight is particularly relevant to digital contexts, where corporations and institutions wield disproportionate power over users who have limited ability to refuse participation or challenge outcomes. The harm is not only psychological but relational, reshaping how individuals understand authority, agency, and accountability.

Drawing on these theoretical strands, digital trauma can be understood as a socio-psychological condition produced by sustained exposure to algorithmic systems characterized by opacity, surveillance, and unequal evaluative power. It manifests as chronic overwhelm, erosion of agency, heightened vigilance, and disruption of social trust. Unlike event-based trauma, digital trauma is cumulative and ambient. It unfolds through everyday interactions rather than extraordinary moments.

Digital trauma is structural rather than individual, emerging from socio-technical systems rather than personal vulnerability. It is opacity-driven, rooted in the inability to understand or contest algorithmic decisions. It is power-asymmetric, reflecting the imbalance between those who design and deploy AI systems and those subjected to their judgments. Finally, it is

225 socially mediated, damaging institutional trust and collective participation alongside
226 individual well-being.

227 Digital humanities scholarship strengthens this analysis by foregrounding technology as a
228 cultural and interpretive system rather than a neutral tool. Scholars such as Berry (2012) and
229 Drucker (2014) emphasize that digital systems shape how knowledge is produced,
230 represented, and understood. Interfaces, visualizations, and infrastructures carry assumptions
231 about authority, objectivity, and legitimacy. From this perspective, algorithmic systems do not
232 merely process information; they reorganize perception, language, and affect.

233 The influence of AI on language itself echoes Orwellian anxieties articulated in *1984*, where
234 control over language functions as a mechanism of power. Contemporary digital life
235 increasingly relies on compressed expressions, predictive text, automated feedback, and
236 platform-specific vocabularies. These shifts subtly reshape how individuals articulate
237 uncertainty, dissent, and emotion. Language becomes optimized for platforms rather than for
238 reflection, reinforcing the affective flattening associated with digital trauma (Chun, 2016).

239 Understanding digital trauma through a socio-technical lens also highlights its uneven
240 distribution. Marginalized communities are disproportionately subjected to intrusive
241 surveillance and algorithmic judgment, particularly within welfare systems, policing, and
242 education (Eubanks, 2018; O'Neil, 2016). In the Global South, where digital infrastructure is
243 often adopted rapidly without robust regulatory frameworks, these effects are intensified. In
244 India, for instance, the integration of AI into education and governance has raised concerns
245 about surveillance, exclusion, and psychological distress, particularly among students and
246 economically vulnerable populations (Dutta, 2021).

247 Seen through this lens, digital trauma is not a failure of individual resilience, but a condition
248 produced by contemporary socio-technical arrangements. Addressing it requires more than
249 ethical guidelines or technical fixes. It demands trauma-informed approaches to AI design
250 and implementation that prioritize transparency, user autonomy, and accountability. Such
251 approaches must also recognize the role of institutions in mediating technological harm and
252 the importance of restoring trust within social systems.

253 In conclusion, digital trauma names a condition that has become ordinary rather than
254 exceptional. By bringing together trauma theory, critical technology studies, and digital
255 humanities, this paper argues for a reorientation of how harm is understood in the age of
256 artificial intelligence. Digital systems shape not only what the users do, but how the users

feel, relate, and make sense of our lives. Recognizing the traumatic dimensions of these systems is a necessary step toward imagining more humane technological futures.

Reference

Andrejevic, M. (2007). *iSpy: Surveillance and power in the interactive era*. University Press of Kansas.

Berry, D. M. (2012). The computational turn: Thinking about the digital humanities. *Culture Machine*, 12, 1–22.

Buell, L. (1995). *The environmental imagination: Thoreau, nature writing, and the formation of American culture*. Harvard University Press.

Caruth, C. (1996). *Unclaimed experience: Trauma, narrative, and history*. Johns Hopkins University Press.

Cheney-Lippold, J. (2017). *We are data: Algorithms and the making of our digital selves*. New York University Press.

Chun, W. H. K. (2016). *Updating to remain the same: Habitual new media*. MIT Press.

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

Drucker, J. (2014). *Graphesis: Visual forms of knowledge production*. Harvard University Press.

Dutta, S. (2021). Surveillance, students, and suspicion: Proctoring software and the future of education. *Postdigital Science and Education*, 3(3), 958–975. <https://doi.org/10.1007/s42438-021-00285-1>

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

Erikson, K. (1976). *Everything in its path: Destruction of community in the Buffalo Creek flood*. Simon & Schuster.

Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.

284 Herman, J. L. (1992). *Trauma and recovery: The aftermath of violence—from domestic abuse*
285 *to political terror*. Basic Books.

286 Lupton, D., & Michael, M. (2017). *Digital sociology*. Routledge.

287 Maier, S. F., & Seligman, M. E. P. (2016). Learned helplessness at fifty: Insights from
288 neuroscience. *Psychological Review*, 123(4), 349–367. <https://doi.org/10.1037/rev0000033>

289 O’Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and*
290 *threatens democracy*. Crown.

291 Orwell, G. (1949). *Nineteen eighty-four*. Secker & Warburg.

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

294 vanDijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific
295 paradigm and ideology. *Surveillance & Society*, 12(2), 197–208.
296 <https://doi.org/10.24908/ss.v12i2.4776>

297 Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new*
298 *frontier of power*. PublicAffairs.

299