# Exploit Machina

*Andrea M. Matwyshyn*[*]

ABSTRACT

*Over half a century ago, Hannah Arendt cautioned us to "think what we are doing" when we build new technologies. Engaging with her counsel and a set of historical case studies, this Article frames what it calls exploit machina problems. Exploit machina refers to situations where broken technologies and broken governance combine to irreparably harm the public. In other words, exploit machina involves organizational choices to knowingly leverage technology as part of legally problematic conduct, including various forms of fraud. In the language of data science, exploit machina situations implicate strategic decisions in building and managing artificial intelligence (AI); they involve, for example, data quality, predictive and prescriptive analytics, and corporate governance. However, reframed in the language of computer security, exploit machina problems are functionally experienced as a form of insider attack. When broken technologies and broken governance converge, untrustworthy insiders can leverage superior information about a technology (and its flaws) to exploit the public and our democratic processes.*

*A portion of modern AI business models now reflect exploit machina dynamics. Using case studies of body-judging devices powered by predictive and prescriptive analytics, this article argues that some AI implementations*

 Professor of Law and Engineering Policy at Penn State Dickinson Law and a Professor in SEDI, Penn State College of Engineering. I wish to thank my repeat co-author Stephanie K. Pell with whom ideas for an earlier version of this article were incubated under a different title, as well as Chloe Anderson, Michael Antonino, Matt Blaze, Leisel Bogan, Shirley Chiu, Danielle Citron, Joshua Corman, Mary Helen Dupree, Sue Glueck, Mark Graber, Margaret Hu, Diana Slaughter-Kotzin, Christopher Marsden, Bethany Morgan, Miranda Mowbray, Martin Redish, Elizabeth Rowe, Desirae Satterlee, Alka Tandan, Marcia Tiersky, Michael Veale, Page Villarreal, and Jessica Wilkerson for their comments, critiques, and contributions to this project.

*threaten to repeat legally problematic historical patterns of insider attacks on confidentiality, integrity, and availability. Then, drawing inspiration from the technology theory of Arendt on cybernation, a form of destructive hyperautomation, this Article begins to reframe the legal and policy conversation around technology safety and exploit machina. It merges insights from data science and computer security theory with those from legal and policy scholarship around data analytics, data privacy, and AI governance. Specifically, this Article recasts technology safety in traditional legal terms — as a current problem where organizations knowingly or intentionally inflict irreparable harms on humans. As such, it rejects the dominant policy narrative of AI safety as a hypothetical future problem of machine supremacy. To combat exploit machina, this Article offers two concrete proposals. First, it introduces a set of (First Amendment sensitive) threat metamodeling techniques that expressly consider insider attacks and public safety. Second, after reviewing recent Supreme Court precedent, it proposes that Congress create a new technology regulator of last resort. The new agency would align existing governmental efforts in technology safety, fill regulatory and enforcement gaps in existing agencies' enabling statutes and practical capabilities, and facilitate international cooperation. The new agency might be called the Bureau of Technology Safety.*

TABLE OF CONTENTS

***

## INTRODUCTION

> "These machines are keeping us alive while others are coming to kill us. Interesting isn't it? The power to give life and the power to end it."
>
> — Councilor Hamann[1]

> "Law never is, but is always about to be . . . . No doubt the ideal system, if it were attainable, would be a code at once so flexible and so minute, as to supply in advance for every conceivable situation the just and fitting rule. But life is too complex . . . ."
>
> — J. Benjamin N. Cardozo[2]

Imagine standing at the top of a staircase and having your artificial vision fail.[3] You would know immediately that your retinal implant had malfunctioned.[4] But what you may not realize is that your visual prosthetics provider might refuse to fix the broken bionic device in your eyes.[5] For business reasons, the company that manufactured your implanted device might inform you that it has discontinued technical support for your "last generation" body device.[6] At that point, you will face a no-win choice: you can have surgery to remove the dead technology from your eyes,[7] or you can allow it to age in place,[8]

---

[1] THE MATRIX RELOADED, at 35:53 (Warner Bros. Pictures 2003).

[2] BENJAMIN N. CARDOZO, THE NATURE OF THE JUDICIAL PROCESS 126, 143 (1921).

[3] *See, e.g.*, Eliza Strickland & Mark Harris, *Their Bionic Eyes are Now Obsolete and Unsupported*, IEEE SPECTRUM (Feb. 15, 2022), https://spectrum.ieee.org/bionic-eye-obsolete [https://perma.cc/XPK7-3393].

[4] *Id.*

[5] *Id.*

[6] *Id.*

[7] *Id.*

[8] *See, e.g.*, Liam Drew, *Abandoned: The Human Cost of Neurotechnology Failure*, NATURE (Dec. 6, 2022), https://www.nature.com/immersive/d41586-022-03810-5/index.html [https://perma.cc/5UUQ-PNNL].

accepting the health risks that will accompany its deterioration.[9] In both cases, your "vision" [10] will be "down"[11] and unavailable[12] — the result of a hardware failure and a corporate product deprecation choice.[13] But, perhaps the company will turn your vision back on if you agree to "upgrade" your body — to purchase and implant the company's newer device.[14]

Stories of broken, remotely deactivated or corrupted bionic body parts may sound like the plot of a frightening dystopian science-fiction novel or film. Perhaps you are reminded of the biometric eye replacement surgery depicted in *Minority Report*,[15] the body part markets

[9] For example, certain other medical procedures are impacted due to the presence of the implant. *Id.*; *see generally* Andrea M. Matwyshyn, *The Internet of Bodies*, 61 WM. & MARY L. REV. 77, 128 (2019) ("You can get your lenses removed, risking physical harm and absorbing the cost. . . . Alternatively, you can buy "upgraded" lenses, absorbing those associated risks and costs. In all cases, the IoB manufacturer has contractually and technically forced an 'upgrade' onto the body of the consumer.") [hereinafter Matwyshyn, *Internet of Bodies*].

[10] For a discussion of the engineering of retinal prosthetics, see, for example, Kevin Y. Wu, Mina Mina, Jean-Yves Sahyoun, Ananda Kalevar & Simon D. Tran, *Retinal Prostheses: Engineering and Clinical Perspectives for Vision Restoration*, 23 SENSORS, no. 5782, at 1, 13, explaining "recent advancements in retinal prosthesis technology . . . , emphasizing progress in engineering and the outlook of retinal prostheses."

[11] For a discussion of downtime in the context of data, see, for example, Barr Moses, *What Is Data Downtime*, MONTE CARLO (Feb. 4, 2024), https://www.montecarlodata.com/blog-the-rise-of-data-downtime/ [https://perma.cc/F98L-8D53].

[12] Unavailability refers to a situation where a resource is not available on demand and on an as-needed basis. *See Availability*, NAT'L INST. STANDARDS & TECH., https://csrc.nist.gov/glossary/term/availability (last visited Sept. 9, 2025) [https://perma.cc/92MP-5K4G] [hereinafter NIST]. For a discussion of unavailability and body-implanted devices, see, for example, Andrea M. Matwyshyn, *Unavailable*, 81 U. PITT. L. REV. 349, 371 (2020) [hereinafter Matwyshyn, *Unavailable*]: "In addition to the public-safety concerns arising from unavailability of Internet access needed for emergency services, a second set of concerns threatens to damage physical safety: when human bodies rely on body-attached and body-embedded Internet of Things devices, unavailability will mean physical injury to human bodies."

[13] Matwyshyn, *Unavailable*, *supra* note 12, at 371-74.

[14] *See infra* text accompanying notes 23–27.

[15] *See* MINORITY REPORT, at 1:09:00 (20th Century Fox & DreamWorks Pictures 2002) (In *Minority Report*, the protagonist receives an eye transplant but keeps his original eyes to use as a biometric identifier).

in *Altered Carbon*[16] or *Repo Men*,[17] or the brain implanted devices in *Johnny Mnemonic*[18] and *Severance*.[19] But this story of deprecated eyes is no longer the stuff of science fiction: it happened to a woman named Barbara in 2022.[20] Despite discontinuing support for her failed retinal implant, Barbara's[21] bionic eye provider later allegedly proceeded to raise[22] millions from the public, stating that it intended to develop[23] a

[16] *See Altered Carbon* (Mythology Entertainment & Skydance Television 2018–2020).

[17] *See* REPO MEN (Universal Pictures, Relativity Media, & Stuber Pictures 2010).

[18] *See* WILLIAM GIBSON, *Johnny Mnemonic*, *in* BURNING CHROME 5, 22 (1995) ("The stored data are fed in through a modified series of microsurgical contraautism prostheses.").

[19] *See* Henry St Leger, *Apple TV's Severance Explained: Is the Personality-Splitting Procedure Actually Possible?*, LIVESCIENCE (Sept. 28, 2022), https://www.livescience.com/severance-tv-show-explained [https://perma.cc/Z55L-BFB5].

[20] *See* Strickland & Harris, *supra* note 3.

[21] Barbara was not the only customer of the company to experience physical consequences of the product deprecation. Eliza Strickland & Mark Harris, *Their Bionic Eyes Are Now Obsolete and Unsupported*, IEEE SPECTRUM (Feb. 15, 2022), https://spectrum.ieee.org/bionic-eye-obsolete [https://perma.cc/XPK7-3393].

[22] *See, e.g.*, Ravikash Bakolia, *Second Sight Raises Additional $7.5M From Underwriters of Stock Offering*, SEEKING ALPHA (June 25, 2021, 3:10 PM), https://seekingalpha.com/news/3710249-second-sight-raises-additional-75m-from-underwriters-of-stock-offering [https://perma.cc/FA8B-FUN4]; *see Second Sight Medical Products Announces Proposed Public Offering of Common Stock*, BUS. WIRE (June 22, 2021), https://www.businesswire.com/news/home/20210622006070/en/Second-Sight-Medical-Products-Announces-Proposed-Public-Offering-of-Common-Stock [https://perma.cc/6UJS-44DS] ("Second Sight Medical Products, Inc. (Nasdaq: EYES) develops implantable visual prosthetics that are intended to deliver useful artificial vision to blind individuals. A recognized global leader in neuromodulation devices for blindness, the Company is committed to developing new technologies to treat the broadest population of sight-impaired individuals."). For a discussion of shelf registrations, see, for example, Roberta S. Karmel, *Assessment of Shelf Registration: How Much Diligence is Due Investors?*, 3 YALE J. ON REG. 401, 401-02 (1986), explaining that "[t]he Shelf Registration Rule is but one component of the SEC's integrated disclosure system, a deregulatory initiative begun during the Carter Administration and put into final form under the leadership of SEC Chairman John S. R. Shad during the Reagan Administration."

[23] *See Second Sight Medical Products*, *supra* note 22 (The S-3 shelf offering documents stated that the company intends to focus its work on visual prosthetic devices.).

new artificial vision[24] product. However, the next generation of the visual prosthetic device would not be a retinal implant:[25] the new device would be brain implanted and powered by artificial intelligence.[26]

It would be comforting for us to dismiss the story of Barbara's[27] dead "cybereyes" as a one-off, as a problem with legal gray areas that jurists and policymakers can postpone for another day. But, unfortunately, Barbara's situation is not exceptional.[28] Similar body "downtime" issues have impacted people with bodies reliant on neural implants, spinal cord stimulators, and various other implanted devices.[29] A number of these people have even resorted to desperate acts of self-help, including hacking the devices implanted in their bodies.[30]

Often caught by surprise when their devices become unavailable, the impacted people (and their doctors) express dismay and shock. [31] They use language of unfair surprise and helplessness, expressing feelings of betrayal and regret over (mis)placing trust in people and organizations that they now view as untrustworthy.[32] They use language that mirrors

[24] Second Sight Medical Products, Inc., Annual Report (Form 10-K) 4 (Mar. 16, 2021) ("[L]everag[ing] proven . . . technology to develop the . . . visual cortical prosthesis and significantly expand our addressable market.").

[25] *Id.*

[26] *See Orion*, CORTIGENT, https://www.cortigent.com/orion (last visited Feb. 7, 2026) [https://perma.cc/89PB-8MVQ].

[27] *See id.* (The company stated in its 2020 10-K that it discontinued the product because of the small potential market: "Given the limited addressable market of Argus II, we no longer market the Argus II and have focused all of our resources on the development of Orion.").

[28] *See, e.g.*, Drew, *supra* note 8 (In complaints filed with the FDA and disclosures to journalists, some customers allege they were promised software upgrades that failed to happen, while other customers allege to have experienced external hardware failures that the company refused to address.).

[29] *Id.*

[30] *Id.*

[31] In the words of one doctor, "You expect them to receive essentially lifelong care from the device manufacturer." *Id.* In the words of another, "Making [patients] the victims of bad business practices or a bankruptcy is horrible for patients, horrible for the field and grossly unethical." *Id.*

[32] *See* Strickland & Harris, *supra* note 3 ("Had I known three years ago what I know now, I probably wouldn't have signed up for it . . . .").

the statements computer security professionals regularly hear from the targets of insider attacks. [33]

Again, it might be comforting to believe that these kinds of irreparable harms are limited to the context of body-implanted devices, but they are not. Existing judicial precedent and agency enforcement reveal a troubling broader pattern: Broken technologies and broken governance are combining in legally problematic ways with growing frequency, and they are imposing irreparable harms on the public and on social systems at scale.[34] Indeed, the stakes are now literally life and death.[35] At least two people have already died following a corporate technology deprecation choice, [36] despite the foreseeability of death as a likely outcome.[37] As (the various technologies we call)[38] artificial intelligence (AI)[39] are incorporated into the systems we trust with our safety, the public has grown more concerned.[40] So too have legislators; according

[33] *Id.*

[34] *See infra* Part I.

[35] *See, e.g.*, Andrea Matwyshyn, *It's Morning Again in Pennsylvania: Rebooting Computer Security Through a Bureau of Technology Safety*, LAWFARE (Jan. 30, 2024, 2:52 PM), https://www.lawfaremedia.org/article/it-s-morning-again-in-pennsylvania-rebooting-computer-security-through-a-bureau-of-technology-safety [https://perma.cc/8N4V-XKJB] [hereinafter Matwyshyn, *It's Morning Again in Pennsylvania*] ("As the Cybersecurity and Infrastructure Security Agency (CISA) explained recently, deaths tied to health care ransomware are mounting, and hospitals are now closing due to ransomware.").

[36] *See* Yasemin Craggs Mersinoglu, *Two Died After UK Shift from Analogue to Digital Phone Lines*, FIN. TIMES (Apr. 26, 2024), https://www.ft.com/content/accd4c72-a7fa-4f82-9c48-4a20cf75124a [https://perma.cc/QXE8-M2QH].

[37] *See, e.g.*, Matwyshyn, *Unavailable*, *supra* note 12.

[38] As explained by one author: "The term AI is now ubiquitous in the technology industry. Its use bears little relation to the things with which [it was originally] concerned, and it's often used to sex-up stuff that has little to do with what practitioners of artificial intelligence care about." Tiernan Ray, *Separating AI from the Nonsense*, FORBES (Sept. 2, 2019, 1:30 PM), https://www.forbes.com/sites/tiernanray/2019/09/02/separating-ai-from-the-nonsense/ [https://perma.cc/22D8-96UK].

[39] The term "artificial intelligence" was first coined as a marketing term by computer science professor John McCarthy to describe his work on nonmonotonic reasoning. *Professor John McCarthy, General Information*, STAN. COMPUT. SCI., http://jmc.stanford.edu/general/index.html [https://perma.cc/YJT9-6ZNX].

[40] Fifty percent of Americans are more concerned than excited about AI according to some recent research. Brian Kennedy, Eileen Yam, Emma Kikuchi, Isabelle Pula &

to the National Conference of State Legislatures, "[i]n the 2025 legislative session, all 50 states, Puerto Rico, the Virgin Islands, and Washington, D.C., have introduced legislation on [AI] . . . [and t]hirty-eight states adopted or enacted around 100 [AI] measures."[41]

This public and legislative interest is understandable. Much like Barbara's eyes, our individual safety and that of our society have become increasingly contingent upon two interlocking safety assumptions. First, we assume the absence of serious technical flaws in the increasingly complex technologies that we trust with our lives and our critical infrastructures. Second, we assume the salutary intent of technology creators and operators. We trust that they will not knowingly choose to inflict harm on us and on our country through their governance and product design choices. But caselaw and enforcement activity now caution us to question these assumptions; and, when these assumptions are flawed, a magnified set of harms results.[42] In other words, insiders are sometimes leveraging their superior information about a particular technology (and its technical flaws) to "game" their legal duties at the expense of safety.

As the safety stakes continue to increase, so too does the severity of potentially irreparable harm in situations where knowingly flawed technologies and knowingly flawed oversight processes combine.[43] This two-pronged form of technological public exploitation deserves its own name — the problem of exploit machina. *Thus, exploit machina refers to situations where broken technologies and broken governance merge to potentially irreparably harm the public and its safety.*

---

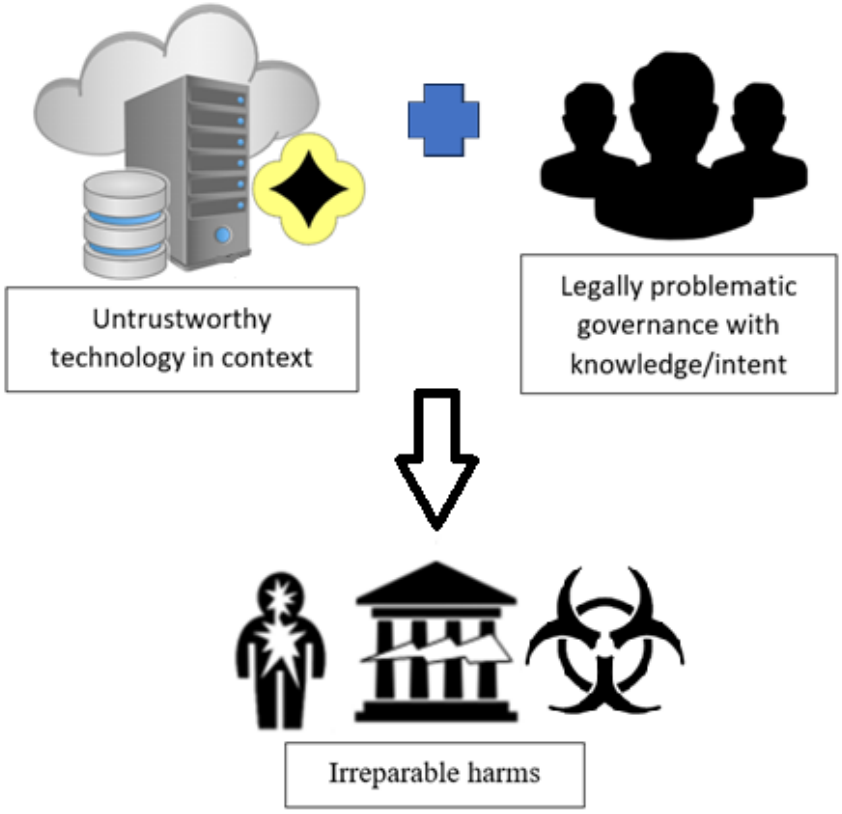
Javier Fuentes, *How Americans View AI and its Impact on People and Society*, PEW RSCH. CTR. (Sept. 17, 2025), https://www.pewresearch.org/science/2025/09/17/how-americans-view-ai-and-its-impact-on-people-and-society/#ai-awareness-and-attitudes [https://perma.cc/B76A-WY3D].

[41] *Artificial Intelligence 2025 Legislation*, NAT'L CONF. OF STATE LEGISLATURES, https://www.ncsl.org/technology-and-communication/artificial-intelligence-2025-legislation (last updated July 10, 2025). These measures have included legislation on AI intellectual property, AI intimate images and harassment, AI whistleblowing, AI misrepresentations of medical licensing, and other related topics. *Id.*

[42] *See infra* Part I.

[43] Subsequent recourse cannot always make harmed parties and society whole. National security and democratic harms in particular are not adequately compensable through retrospective judicial means.

Figure 1: Exploit Machina




Stated another way, through exploit machina organizations can knowingly leverage unsafe[44] technology as part of legally problematic conduct, including in furtherance of various traditional forms of fraud.

For example, recent exploit machina incidents have included criminal scenarios where companies created purpose-built technologies to

[44] Unsafe as used herein means:

> reflecting a deficit of technical and governance substantiation sufficient to reasonably believe that irreparable physical, psychological, financial, national security, infrastructure or other public safety harms will not result from the selected design, operation, and maintenance of the technology, in light of the current state of the art of safety and computer security knowledge in light of the context of deployment and use.

intentionally game regulatory oversight of public safety.[45] They have included intentional misrepresentations of technological capabilities that defrauded the public,[46] and unsafe design choices in shipped[47] technology, where death foreseeably resulted.[48] They have also included emergent technology failures where the people who controlled a particular technology ignored known risks in their operations or refused

[45] *See* Press Release, U.S. Dep't of Just., Volkswagen AG Agrees to Plead Guilty and Pay $4.3 Billion in Criminal and Civil Penalties; Six Volkswagen Executives and Employees are Indicted in Connection with Conspiracy to Cheat U.S. Emissions Tests (Jan. 11, 2017), https://www.justice.gov/opa/pr/volkswagen-ag-agrees-plead-guilty-and-pay-43-billion-criminal-and-civil-penalties-six [https://perma.cc/699D-S2CZ] (for example, Volkswagen AG plead guilty and paid $4.3 billion in criminal and civil penalties for purpose-built defeat devices and six executives and employees were indicted for attempting to game the EPA. VW had created the defeat device software to trick consumers, the market, and regulators into believing that cars met environmental safety specifications); *see also, e.g.*, Press Release, EPA, Hino Motors, a Toyota Subsidiary, Agrees to Plead Guilty and Pay Over $1.6B to Resolve Emissions Fraud Scheme (Jan. 15, 2025), https://www.epa.gov/newsreleases/hino-motors-toyota-subsidiary-agrees-plead-guilty-and-pay-over-16b-resolve-emissions [https://perma.cc/M5BG-6YRE] ("Hino Motors, Ltd. engineers also failed to disclose software functions that could adversely affect engines' emission control systems. As a result of the fraud, Hino Motors, Ltd. imported and sold over 105,000 non-conforming engines between 2010 and 2022.").

[46] For example, some of Theranos' officers were convicted for intentionally making misrepresentations to investors, the public, and business partners about the functionality of the company's flawed technology that (mis)diagnosed terminal illnesses. *See infra* Part I.

[47] Shipping technology is a term of art meaning commercial release. *See, e.g.*, Jocelyn Goldfein, *The Right Way to Ship Software*, FIRST ROUND, https://review.firstround.com/the-right-way-to-ship-software/ (last visited Nov. 18, 2025) [https://perma.cc/K7BM-VD3M] (discussing experiences in building and releasing software into the commercial marketplace, one former executive writes, "I've discovered and rediscovered the 'right' way to build and ship software many times").

[48] For example, courts appear to be growing comfortable with the idea of assessing the degree of direct product design control a company possessed, product use and behavior, and knowledge of likely harm. *See, e.g.*, Lemmon v. Snap, Inc., 995 F.3d 1085, 1092-94 (9th Cir. 2021) (explaining that, "In short, Snap, Inc. was sued for the predictable consequences of designing Snapchat in such a way that it allegedly encouraged dangerous behavior" and "It is thus apparent that the Parents' amended complaint does not seek to hold Snap liable for its conduct as a publisher or speaker. Their negligent design lawsuit treats Snap as a products manufacturer, accusing it of negligently designing a product (Snapchat) with a defect (the interplay between Snapchat's reward system and the Speed Filter)").

to evolve products[49] to mitigate ongoing irreparable harms.[50] In some cases, the failure to correct promptly[51] occurred despite public outcry and despite express demands from public safety and consumer protection agencies.[52] Indeed, sometimes companies have "decided not to resolve . . . vulnerabilities" despite direct notice of serious, ongoing safety risks,[53] or they have allegedly intentionally destroyed or withheld technical evidence to foil forensic inquiry when their products are

[49] Meanwhile, Meta shareholders alleged that officers repeatedly violated Federal Trade Commission consent decrees knowingly, allegations that a Delaware court recently deemed potentially credible and a basis for suit — a signal of a potentially "captured" board of directors. *See* Verified Shareholder Derivative Complaint, CONF ORD, Emp.'s Ret. Sys. v. Zuckerberg, No. 2023-0304-JTL (Del. Ch. Mar. 10, 2023).

[50] For example, whistleblowers and litigants claim that Meta's officers knowingly turned a blind eye to use of the company's technology potentially being used in furtherance of war crimes. *E.g.*, Dan Milmo, *Rohingya Sue Facebook for £150bn over Myanmar Genocide*, GUARDIAN (Dec. 6, 2021, 12:03 PM), https://www.theguardian.com/technology/2021/dec/06/rohingya-sue-facebook-myanmar-genocide-us-uk-legal-action-social-media-violence [https://perma.cc/S9L9-7UCS].

[51] An example involves the potential sensor and design issues on cars, *see, e.g.*, *Update Vehicle Firmware to Correct Brake Fluid Level Sensor*, TESLA, https://www.tesla.com/support/recall-brake-fluid-level-sensor (last visited Sept. 10, 2025) [https://perma.cc/7JUQ-PARV]. *See, e.g.*, Vanessa Romo, *Judge Says Evidence Shows Tesla and Elon Musk Knew About Flawed Autopilot System*, NPR (Nov. 23, 2023, 6:00 AM), https://www.npr.org/2023/11/23/1214966530/judge-says-evidence-shows-tesla-and-elon-musk-knew-about-flawed-autopilot-system [https://perma.cc/49RV-X67H] (safety advocates and plaintiffs have alleged that known sensor failures on cars have already resulted in death); *Update Vehicle Firmware to Correct Brake Fluid Level Sensor*, TESLA, https://www.tesla.com/support/recall-brake-fluid-level-sensor (last visited Sept. 10, 2025) [https://perma.cc/7JUQ-PARV].

[52] *See, e.g.*, Sean O'Kane, *Tesla Ignored Safety Board's Autopilot Recommendations, Chairman Says*, VERGE (Feb. 25, 2020, 11:32 AM), https://www.theverge.com/2020/2/25/21152984/tesla-autopilot-safety-recommendations-ignored-ntsb-crash-hearing [https://perma.cc/PQJ3-68P4] ("Tesla ignored safety recommendations made by the National Transportation Safety Board (NTSB) about its Autopilot driver assistance system, the board's chairman Robert Sumwalt said on Tuesday.").

[53] *See, e.g.*, *ICS Advisory: Festo CECX-X-(C1/M1) Controller Vulnerabilities*, CYBERSECURITY & INFRASTRUCTURE SEC. AGENCY, https://www.cisa.gov/news-events/ics-advisories/icsa-14-084-01 (last revised Sept. 6, 2018) [https://perma.cc/3T3W-XYJT].

involved in the death of a member of the public.[54] In other cases, companies have chosen to move too slowly in correcting dangerously flawed technology design, again resulting in preventable loss of life.[55]

Exploit machina dynamics may sound counterintuitive. You might ask: "Why would an organization harm its own users and the public? Surely that cannot be profitable or efficient." Yet current financial and legal incentives can lead technology builders to push out a "minimum viable product,"[56] rather than a minimum *safe* product that is fit for purpose.[57] A faster, less safe product is sometimes perceived by technology builders as the logical default because of three reasons — the time value of money, the tying of individual officer compensation to short term outcomes, and the low risk of swift regulatory enforcement. In the limited cases where enforcement or litigation does follow, then the incentives shift toward protracting the legal process to minimize short term financial impact.[58] Without timely oversight and

[54] *See, e.g.*, Fred Lambert, *Tesla Withheld Data, Lied, and Misdirected Police and Plaintiffs to Avoid Blame in Autopilot Crash*, ELECTREK (Aug. 4, 2025, 8:54 AM), https://electrek.co/2025/08/04/tesla-withheld-data-lied-misdirected-police-plaintiffs-avoid-blame-autopilot-crash/ [https://perma.cc/T7CP-5DHD] (alleging that based on a review of trial transcripts in *Benavides v. Tesla, Inc.*, 804 F. Supp. 3d 1242 (S.D. Fla.) (No. 1:21-cv-21940) "Tesla withheld the crash-snapshot data that its own server received within minutes of the collision" and "[w]hen the plaintiffs asked for the data, Tesla said that it didn't exist").

[55] *See, e.g.*, STAFF OF S. COMM. ON COM., SCI., & TRANSP., 117TH CONG., AVIATION SAFETY WHISTLEBLOWER REP. (Comm. Print 2021).

[56] *See* Maksym Babych, *A Review of the Minimum Viable Product Approach*, FORBES, https://www.forbes.com/councils/theyec/2021/12/08/a-review-of-the-minimum-viable-product-approach/ (last updated Apr. 21, 2022, 8:20 AM) [https://perma.cc/QA5M-VBBJ].

[57] This minimum viable product approach might then sometimes be strategically coupled with subsequent intentional delay in remediation of unsafe conditions or in responsiveness to legal inquiries. *Breach Claims from Delayed Product Modifications or Fixes*, AARON HALL, https://aaronhall.com/breach-claims-from-delayed-product-modifications-or-fixes/#How_Can_Delays_in_Product_Updates_Lead_to_Legal_Liability (last visited Feb. 8, 2026) [https://perma.cc/F335-AVGY] ("Contracts involving software development, manufacturing supply, and service levels are particularly vulnerable to issues from delayed product modifications.").

[58] These dynamics are particularly concerning in a circumstance where the CEO works for a Board of Directors that is perceived as feeble or whose members have been captured. Press accounts have alleged that such dynamics may exist at a number of large

consequences, officers and directors who knowingly make unsafe, legally problematic governance choices can reframe them as "cost-effective" decisions, choices aimed at "driving shareholder value" in the short term, even if it leads to self-harm in the long term[59] and externalizing irreparable harms[60] onto the public.[61] Troublingly, even technology insiders now describe this dynamic as culturally endemic — as a short-termist ethos of disposability common in Silicon Valley, and one that now limits meaningful progress of science and useful arts.[62]

But exploit machina problems extend beyond the private sector, beyond consumer products, and beyond for-profit entities: sometimes nonprofit organizations and governmental organizations are also part of the problem.[63] For example, government programs have sometimes

technology companies. *See, e.g.*, Kirsten Grind, Emily Glazer, Rebecca Elliot & Coulter Jones, *The Money and Drugs That Tie Elon Musk to Some Tesla Directors*, WALL ST. J. (Feb. 3, 2024, 9:00 PM), https://www.wsj.com/tech/elon-musk-tesla-money-drugs-board-61af9ac4 [https://perma.cc/P3EZ-3ZGW]; Steven Musil, *Facebook CEO Mark Zuckerberg Reportedly Overhauled Board to Assert More Control*, CNET (Apr. 28, 2020, 6:30 PM), https://www.cnet.com/tech/services-and-software/facebook-ceo-mark-zuckerberg-reportedly-overhauled-board-to-assert-more-control/ [https://perma.cc/JVC6-4T9G]. *But see, e.g.*, Felix von Meyerinck, Jonas Romer & Markus Schmid, *CEO Turnover and Director Reputation*, HARV. L. SCH. F. ON CORP. GOVERNANCE (Jan. 7, 2025)), https://corpgov.law.harvard.edu/2025/01/07/ceo-turnover-and-director-reputation/ [https://perma.cc/G6RA-5KTU].

[59] The reframing, however, often comes at the expense of corporate self-harm in the long run. However, current fiduciary duty obligations have not been interpreted adequately robustly to create a check on this short-termist corporate self-harm. *See* Andrea M. Matwyshyn, *Imagining the Intangible*, 34 DEL. J. CORP. L. 966, 984 (2020).

[60] *See Id.* at 966, 997-98.

[61] For a discussion of Silicon Valley's focus on short term consumer product wins at the expense of national security and national social cohesion interests, see, for example, ALEXANDER KARP & NICHOLAS ZAMISKA, THE TECHNOLOGICAL REPUBLIC, 103-12 (2025). For a discussion of the limitation of the entire project of artificial general intelligence as it is currently framed in Silicon Valley, see generally ARVIND NARAYANAN & SAYASH KAPOOR, AI SNAKE OIL (2024).

[62] *Id.* As explained by AI pioneer Sir Demis Hassabis, juxtaposing the approaches of the US and UK technology entrepreneurship ecosystems, "In Silicon Valley, everybody is having a new company every year and if it doesn't work you chuck it and then you start something new. That is not conducive to a long-term research challenge." THE THINKING GAME (Cityspeak Films & Reel as Dirt 2024).

[63] In other words, these dynamics include potentially irreparable harms at scale to infrastructure, national security, and society. For example, manufacturers of industrial

misplaced their reliance on untrustworthy predictive and prescriptive technologies that were not fit for purpose (and they have sometimes chosen to ignore actual knowledge of ongoing harms from these technologies).[64] Tragically, exploit machina in government contexts has already resulted in hundreds of wrongful criminal convictions,[65] tens of thousands of wrongful criminal[66] accusations,[67] and irreparable psychological harms, including potentially leading to multiple suicides.[68]

control systems whose systems contain critical and exploitable security vulnerabilities have sometimes refused to correct the issue even when repeatedly advised to do so by agencies. *ICS Advisory*, *supra* note 53.

[64] *See, e.g.*, Ben King, *Post Office: Governments of Last 20 Years Should 'Regret' Horizon Scandal, Says David Cameron*, BBC (Feb. 22, 2024) https://www.bbc.com/news/business-68362392 [https://perma.cc/L2H4-GFKU] ("Lord Cameron has expressed regret over the Post Office scandal after revelations emerged suggesting his government knew about an axed probe that may have cleared sub-postmasters. . . . Lord Cameron described 'the appalling way' post office branch managers had been treated. 'I've said very clearly already many times, I think anyone who has been in government for the last 15, 20 years or perhaps more, should deeply regret what's happened.'").

[65] *See, e.g.*, Stephen Castle, *How a TV Show Forced Britain's Devastating Post Office Scandal into the Light*, N.Y. TIMES (Jan. 10, 2024), https://www.nytimes.com/2024/01/10/world/europe/uk-itv-mr-bates-vs-post-office.html [https://perma.cc/2TD2-82NU] ("More than 700 people convicted of a crime they didn't commit. At least four suicides. A woman sent to jail while pregnant. Bankruptcies. Marriages broken, lives ruined.").

[66] In other words, the antidemocratic, pre-crime world from films such as *Minority Report* is arguably arriving; but, instead of relying on human Precogs, ours relies on sometimes flawed prognostications from predictive analytics and sensor data. *See Agatha*, FANDOM: MINORITY REP. WIKI, https://minorityreport.fandom.com/wiki/Agatha [https://perma.cc/92BF-WRVW].

[67] *See* David Eggert, *State Apologizes for Fraud Fiasco, Wants to Reduce Penalties*, ASSOCIATED PRESS, https://apnews.com/united-states-congress-c0e2346e85854a5b827ca42653c1fb40 (last updated Jan. 28, 2017, 2:35 PM) [https://perma.cc/JJ6T-WYLX].

[68] *See* Castle, *supra* note 65; *see also, e.g.*, Frances Mao, *Robodebt: Illegal Australian Welfare Hunt Drove People to Despair*, BBC (July 7, 2023), https://www.bbc.com/news/world-australia-66130105 [https://perma.cc/3XKN-M38A] ("A landmark inquiry in Australia has found an illegal welfare hunt by the previous government made victims feel like criminals and caused suicides."). *See also The Post Office Scandal*, YOUTUBE, at 2:37 (Dec. 23, 2025), https://www.youtube.com/watch?v=HxjLFgX-eP0.

In each of the above exploit machina scenarios, untrustworthy technologies, on the one hand, and legally problematic organizational governance choices, on the other, combined to exploit the trust of both individual humans and society at scale. In other words, exploit machina is an inherently public-private safety problem,[69] just as computer security always has been public-private due to the nature of vulnerability.[70] In the language of data science, exploit machina situations implicate strategic choices in AI and the data lifecycle involving, for example, choices with respect to data quality,[71] predictive[72] and prescriptive analytics,[73] data fabric,[74] data management,[75] and data governance.[76] However, reframed in the language of computer security, *exploit machina problems are experienced*

[69] In other words, the harms of exploit machina are both individual and social simultaneously.

[70] S*ee, e.g.*, Andrea M. Matwyshyn, *CYBER!*, 2017 BYU L. REV. 1109, 1109 (2017) (explaining that dominant "cybersecurity" paradigms . . . fail to recognize that corporate information security and national "cybersecurity" concerns are inextricable, which might be called the "problem of 'reciprocal security vulnerability'") [hereinafter Matwyshyn, *CYBER!*]; *see also, e.g.*, Kristen E. Eichensehr, *Public-Private Cybersecurity*, 95 TEX. L. REV. 467, 478-79 (2017).

[71] Data quality refers to an assessment of a data set's condition based on factors including completeness, consistency, reliability, and validity. *See* Scott Robinson, Robert Sheldon & Craig Stedman, *Data Quality*, TECHTARGET (Aug. 13, 2025), https://www.techtarget.com/searchdatamanagement/definition/data-quality [https://perma.cc/HY7W-KBKP].

[72] *See What Is Predictive Analytics?*, IBM (Aug. 8, 2022), https://www.ibm.com/think/topics/predictive-analytics [https://perma.cc/MLV7-WFQU].

[73] *See* Cole Stryker, *What Is Prescriptive Analytics?*, IBM (May 2, 2024), https://www.ibm.com/think/topics/prescriptive-analytics [https://perma.cc/H4ML-GECC].

[74] *See* Alexandra Jonker & Tom Krantz, *What Is a Data Fabric?*, IBM (Aug. 5, 2025), https://www.ibm.com/think/topics/data-fabric [https://perma.cc/XQP7-TQ33].

[75] *See* Craig Stedman, *What Is Data Management and Why Is It Important?*, TECHTARGET (May 28, 2024), https://www.techtarget.com/searchdatamanagement/definition/data-management [https://perma.cc/GJG2-QTSB].

[76] *See* Craig Stedman, *What Is Data Governance and Why Does It Matter?*, TECHTARGET (Feb. 23, 2024), https://www.techtarget.com/searchdatamanagement/definition/data-governance [https://perma.cc/XBP6-RZCB].

*as a form of insider attack*.[77] When these destructive dynamics of technology and governance converge, insiders can leverage superior information about a technology (and its flaws) to exploit the public and to harm its safety, including harming constitutional interests.

And so we arrive at our current technology policy precipice. Data-intensive technologies such as AI potentially offer transformational life improvements for individuals and for society in suitable applications. However, this success is deeply contingent: As ever-greater portions of our daily lives come to rely on information technology systems and AI, our safety and thriving[78] depend not only on the trustworthiness of those systems and on the trustworthiness of the organizations that create, manage, and monitor them. We also rely on the trustworthiness of the technology marketplace as a whole. Transformational technologies that are built safely[79] cannot succeed when unsafe

[77] *See Defining Insider Threats*, CYBERSECURITY & INFRASTRUCTURE SEC. AGENCY, https://www.cisa.gov/topics/physical-security/insider-threat-mitigation/defining-insider-threats (last visited Nov. 18, 2025) [https://perma.cc/4GJF-VAQ3] ("The Cybersecurity and Infrastructure Security Agency (CISA) defines insider threat as the threat that an insider will use their authorized access, intentionally or unintentionally, to do harm to the department's mission, resources, personnel, facilities, information, equipment, networks, or systems. Insider threats manifest in various ways: violence, espionage, sabotage, theft, and cyber acts.").

[78] Thriving is a term of art in developmental psychology. *See, e.g.*, Daniel J. Brown, Rachel Arnold, David Fletcher & Martyn Standage, *Human Thriving: A Conceptual Debate and Literature Review*, 22 EUR. PSYCH. 167, 179 (2017), https://econtent.hogrefe.com/doi/full/10.1027/1016-9040/a000294 [https://perma.cc/FWK8-M9MW] (explaining that "variety of indicators suggests that thriving is multifaceted and may appear qualitatively different across individuals" and "can be broadly defined as the joint experience of development and success"). For an introduction to developmental psychology, see, for example, *Developmental Psychology Studies Humans Across the Lifespan*, AM. PSYCH. ASS'N (2014), https://www.apa.org/education-career/guide/subfields/developmental [https://perma.cc/6578-CNWJ]: "Developmental psychologists study human growth and development over the lifespan, including physical, cognitive, social, intellectual, perceptual, personality and emotional growth."

[79] As the Supreme Court has explained, risk of future harm does not offer standing for an individual plaintiff to pursue a claim, even when a statutory violation has occurred. *See* TransUnion LLC v. Ramirez, 594 U.S. 413, 417 (2021). However, the ability of government enforcement to defend the individual and collective interests of the public is not similarly constrained. Even false speech which falls into categories of "advocacy intended, and likely, to incite imminent lawless action; obscenity;

technologies saturate the marketplace; the signal of high-quality technologies cannot be heard in a marketplace where the noise of fraudulent and abusive technologies has become deafening. Market trust and national security are both undercut when exploit machina becomes pervasive. Thus, exploit machina arguably presents the lodestar technology governance, policy, and legal issue of the next decade.[80]

To wit, this Article argues a straightforward legal and policy premise: to address the growing threat of exploit machina to public safety, the stability of our markets, national security, and democracy, Congress should act to create a new technology regulator and enforcer of last resort — a Bureau of Technology Safety (BoTS). As the burgeoning list of exploit machina scenarios demonstrates, we exist in what is known in computer security as a race condition[81] and what legal scholars have

defamation; speech integral to criminal conduct; so-called 'fighting words'; child pornography; fraud; true threats; and speech presenting some grave and imminent threat the government has the power to prevent" can be the basis for government regulatory or criminal sanction. *See* United States v. Alvarez, 567 U.S. 709, 717 (2012) (citations omitted). While definitions of safety may vary across statutory regimes, they generally involve protecting the public from unreasonable risks of serious harm, injury or death. *See, e.g.*, *Consumer Product Safety Commission (CPSC)*, USAGOV, https://www.usa.gov/agencies/consumer-product-safety-commission (last visited Nov. 18, 2025) [https://perma.cc/8Q7P-SEWD]. As explained by Professor Ido Kilovaty in the context of recourse for psychological harms arising from security compromises, "[c]onsumers suffering [psychological] harms are unlikely to pursue litigation and, even if consumers do pursue litigation, are unlikely to prevail because of both standing and cause of action reasons." Ido Kilovaty, *Psychological Data Breach Harms*, 23 N.C. J.L. & TECH. 1 (2021).

[80] At the moment, the list of exploit machina examples ripped from the headlines is uncomfortably long. *See As Nationwide Fraud Losses Top $10 Billion in 2023, FTC Steps Up Efforts to Protect the Public*, FTC (Feb. 9, 2024), https://www.ftc.gov/news-events/news/press-releases/2024/02/nationwide-fraud-losses-top-10-billion-2023-ftc-steps-efforts-protect-public [https://perma.cc/5HL4-CE3E] (Technology-related fraud complaints topped the list of consumer submissions at the Federal Trade Commission in 2023; losses exceeded $10 billion, reflecting a fourteen percent increase over reported losses in 2022).

[81] *See* Ben Lutkevich, *Race Condition*, TECHTARGET (June 16, 2021), https://www.techtarget.com/searchstorage/definition/race-condition [https://perma.cc/H9PX-9CLA].

called a technology law Red Queen Effect.[82] We are running as fast as we can, yet we are losing legal ground. Court judgments and agency enforcement during the first quarter of the twenty-first century have proven inadequately persuasive to discourage organizations from engaging in exploit machina, and we are now out of time — traceable deaths and other irreparable harms have begun. Significantly more nimble and efficient legal oversight of technology safety is necessary. This Article argues that the reason for urgency in addressing exploit machina rests in what Hannah Arendt called "cybernation," a form of socially destructive hyperautomation.[83]

Specifically, Part I continues from Barbara's story. Using body sensing technologies as a case study, Part I further introduces exploit machina problems. It explains that the exploit machina problems visible today are not entirely historically unprecedented.[84] In past eras, untrustworthy technologies and insider decisions have leveraged flawed predictive and prescriptive data analytics to exploit legal coordination, oversight, and enforcement gaps.[85] Part I then maps the historical exploit machina conversation to modern AI and data science concerns

[82] Andrea M. Matwyshyn, *Penetrating the Zombie Collective: Spam as an International Security Issue*, 3 SCRIPT-ED 371, 385 n.85 (2006), https://script-ed.org/wp-content/uploads/2016/07/3-4-Matwyshyn.pdf [https://perma.cc/9N2W-JU24] ("A Red Queen Effect generally refers to a situation where individuals must adjust quickly to changing threats to survive from generation to generation, derived from Lewis Carroll's *Through the Looking Glass*, where Alice complains to the character of the Red Queen that it is necessary to run simply to stay in the same place and advancing means running twice as fast.").

[83] *See* Part II.A.

[84] For a discussion of repeating patterns in the context of computer security, corporate conduct, and harms, see, for example, *infra* Part I.

[85] For a discussion of this exploitation, see, for example, Matwyshyn, *It's Morning Again in Pennsylvania*, *supra* note 35: "Sometimes they exploit coverage gaps across agencies' enabling statutes, and sometimes they game suboptimal coordination when a single technology safety problem cuts across multiple agencies' authority."

such as data quality[86] and downtime.[87] It also intersects this analysis with the characteristics of an insider attack as defined in the field of information security — an unexpected failure of confidentiality, integrity, or availability of a system due to an untrustworthy insider.[88] Today's sensor-reliant Internet of Bodies (IoB)[89] technologies often include AI with predictive and prescriptive analytics as part of device functionality.[90] Yet, particularly in circumstances where use of these devices is mandatory, outside a user's control or weakly governed, a user may face a (legally problematic) exploit machina whipsaw. One set of harms may arise from flaws in the devices themselves, and a second set may arise from (the legal, economic, and physical consequences of) flawed secondary judgments of AI — from the predictive and prescriptive analytics that rely on those devices' outputs. Part I then argues that these flawed devices and judgments may present a patina of infallibility despite the underlying exploit machina problems. More reliable sources of information and judgment may face devaluation. Part I concludes by explaining why "multiple function" [91] IoB devices enabled

[86] For an overview of data quality issues, see, for example, Jingran Wang, Yi Liu, Peigong Li, Zhenxing Lin, Stavros Sindakis & Sakshi Aggarwal, *Overview of Data Quality: Examining the Dimensions, Antecedents, and Impacts of Data Quality*, 15 J. KNOWLEDGE ECON. 1159 (2023), developing "the notion of 'Data Analytics Competency' . . . as a five-dimensional formative measure (i.e., data quality, the bigness of data, analytical skills, domain knowledge, and tool sophistication)"

[87] *See, e.g.*, Moses, *supra* note 11 (explaining the meaning of downtime).

[88] *See Defining Insider Threats*, *supra* note 77.

[89] *See* Maria Gomez De Sicart & Cristy Garratt, *The Next Generation of the 'Internet of Bodies' Could Meld Tech and Human Bodies Together*, CNBC (June 1, 2024, 9:40 AM), https://www.cnbc.com/2024/06/01/internet-of-bodies-could-meld-tech-and-human-bodies-together.html [https://perma.cc/Q52F-5S24]. For a legal analysis of the Internet of Bodies framework, see Matwyshyn, *Internet of Bodies*, *supra* note 9, at 129-56.

[90] These technologies also often merge existing databases of information tethered to identifiable humans with live data, seeking to bolster predictive power through "big data." Big data has multiple definitions that variously seek to encompass the idea of machine learning assisted processing of databases with large numbers of data points. For a discussion of "big data", see, for example, UW Online Collaboratives, *What Is Big Data?*, UNIV. OF WIS. DATA SCI., (May 18, 2015), https://datasciencedegree.wisconsin.edu/data-science/what-is-big-data/ [https://perma.cc/29JQ-PPFR].

[91] For a discussion of legal status of so-called multiple function medical devices, see *infra* Part I.C.

by the 21st Century Cures Act,[92] in particular, introduce heightened risk of exploit machina.

Part II calls for explicitly engaging with the Arendtian shibboleth of "thinking what we are doing."[93] Exploit machina problems impact not only individual safety; they also menace various Constitutional values. They potentially lead us down a self-destructive and antidemocratic societal path, a path that ends in a form of self-harm that Hannah Arendt called "cybernation." [94] Part II engages with Arendt's work on "cybernation" alongside her consonant insights on the philosophy of Immanuel Kant on imagination. Part II then pairs Arendt's work with scholarship from developmental psychology theory and the writings of Founder and Framer Benjamin Franklin, highlighting the developmental importance of self-narrated identity formation. It also argues that today's exploit machina problems endanger not only public safety but also tomorrow's democratic process, public mental health, and the kernel of our creative innovation economy. Without new trustworthiness backstops in law, exploit machina brings a form of self-harm that quietly undercuts baseline assumptions of U.S. law — the default assumptions that undergird antitrust, the First Amendment, intellectual property, and contract law. Yet these baselines are critical to both promotion of the "progress of science and useful arts" and to democratic governance.

Building on Arendt and Franklin, Part III then argues, first, that the path forward lies in creating technology-neutral trustworthiness backstops in governance — technical, organizational, and legal. It advocates adopting a policy framework centering "thinking what we are doing," a framework whose goal is preventing exploit machina. Part III reviews existing threat modeling frameworks and proposes the need for threat "metamodeling." It argues in favor of explicitly considering human and social impact of technologies and the role of insider attacks in organizational governance. It then proposes one such metamodel, the TROL metamodel. TROL is grounded in three First Amendment-

[92] For background on the 21st Century Cures Act, see, for example, 21st Century Cures Act, Pub. L. No. 114-255, 130 Stat. 1033 (2016).

[93] *See infra* Part II (explaining the political philosophy of Arendt on cybernation).

[94] *See infra* discussion in Part II (explaining that Arendt's cybernation is a state of self-destructive hyper-automation).

sensitive variables that are explicitly technology neutral and easily scale across international boundaries: context sensitivity and control, harm severity, and intent and knowledge. Part III then reviews the current dynamics of software liability. Inspired by design principles from Benjamin Franklin, Part III proposes the (re)centering of progress — not innovation — and the creation of a robust but self-restrained technology regulator of last resort to address exploit machina: BoTS. BoTS would act as a second trustworthiness backstop, an efficient, coordinating technology safety agency with rulemaking, injunctive, fining, and other authority. In other words, BoTS would align technology safety work across government with private sector technology evolution. It would expedite resolution of exploit machina situations and ensure that government conduct aligns with building the public trust and public-private progress.[95]

## I. MOVING (TOO) FAST AND BREAKING THINGS: INSIDER ATTACKS

> "The Matrix is older than you know"
>
> — The Architect[96]

> "Thankfully its [sic] not the 16 hundreds, or else we'd be burned at the stake!"
>
> — Elizabeth Holmes, @eholmes2003[97]

In the second film of the *Matrix* series, the protagonist Neo and the people of Zion realize that they have been misled.[98] Their most trusted technology, the Oracle, has lied to them, and a corrupt insider among

[95] Stated another way, BoTS would ensure public-private alignment on avoiding irreparable technology harms and on more effectively defending other technology-related public safety interests.

[96] THE MATRIX RELOADED, *supra* note 1, at 1:51:42.

[97] Theranos (@theranos), TWITTER (Oct. 26, 2015), https://twitter.com/theranos/status/658802936013811712 [https://perma.cc/V4W3-S9TR] ("Thankfully its not the 16 hundreds, or else we'd be burned at the stake!").

[98] *See generally* THE MATRIX RELOADED, *supra* note 1.

them has sabotaged their technology defenses without detection.[99] The combination of these two betrayals proves devastating for their efforts at preserving their freedom; yet, they persist.[100] This same toxic combination — untrustworthy technology, on the one hand, and untrustworthy governance processes, on the other — presents potentially the most formidable challenge we face in today's technology ecosystem. This is the dynamic that this Article calls exploit machina.

Consider the fraud of Theranos, perhaps the most (over)hyped medical technology company of the last decade. In January 2022, a Northern California jury convicted Elizabeth Holmes, the founder of Theranos, on wire fraud charges.[101] In its indictment, the Department of Justice[102] asserted that the scientific claims about her "revolutionary" technology that promised on-site blood analysis through software-reliant machines[103] were, in fact, false[104] — aspirational fictions about

[99] *Id.*

[100] *Id.*

[101] U.S. Att'y's Off., N. Dist. of Cal., *U.S. v. Elizabeth Holmes, et al.*, U.S. DEP'T OF JUST. (Dec. 12, 2022), https://www.justice.gov/usao-ndca/us-v-elizabeth-holmes-et-al [https://perma.cc/XJ7S-Y8RU]; Miles Cohen, *Theranos Founder Elizabeth Holmes Convicted on 4 Counts of Fraud*, ABC NEWS (Jan. 3, 2022, 6:00 PM), https://abcnews.go.com/US/elizabeth-holmes-trial-jury-unable-unanimous-verdict-counts/story?id=82055043 [https://perma.cc/49PP-MLHR].

[102] Indictment at 8, United States v. Holmes, No. 18-cr-00258 (N.D. Cal. 2018), 2018 U.S. Dist. LEXIS 176120.

[103] Thomas Claburn, *Theranos Blood-Test Machine Demos for VIPs Rigged to Hide any Failures, Court Told*, REGISTER (Oct. 21, 2021, 1:15), https://www.theregister.com/2021/10/21/theranos_machine_trial/ [https://perma.cc/76U9-QW6B]; Joel Rosenblatt, *Theranos Devices Were Built to Impress, with Error Code Hidden*, BLOOMBERG.COM (Oct. 19, 2021, 14:27), https://www.bloomberg.com/news/articles/2021-10-19/theranos-devices-were-built-to-impress-with-error-code-hidden [https://perma.cc/9J4Y-NJ8D]; Thomas Claburn, *Theranos Blood-Test Demo Machines Hid Errors, Court Told*, REGISTER (Oct. 21, 2021), https://www.theregister.com/2021/10/21/theranos_machine_trial/ [https://perma.cc/76U9-QW6B].

[104] *See* Indictment, *supra* note 102, at 6 ("Holmes and Balwani defrauded doctors and patients (1) by making false claims concerning Theranos's ability to provide accurate, fast, reliable, and cheap blood tests and test results, and (2) by omitting information concerning the limits of and problems with Theranos's technologies. The defendants knew Theranos was not capable of consistently producing accurate and reliable results for certain blood tests . . . .").

health tech AI vaporware.[105] As set forth in the indictment, Theranos was merely a scheme to defraud.[106] Evidence presented to the jury demonstrated that, the technology never worked as described.[107] Yet Holmes made express claims to the contrary to Theranos's investors, doctors, and patients,[108] and the device's untrustworthy results placed patient bodies (and their mental states) at risk.[109] The Securities and Exchange Commission similarly brought an enforcement action against Holmes on the basis of defrauding investors,[110] seeking to address the

[105] *Vaporware*, MERRIAM-WEBSTER, https://www.merriam-webster.com/dictionary/vaporware [https://perma.cc/W9X5-GVD7] (last visited Mar. 7, 2025) (vaporware refers to the development of a technology "product that appears to be on its way but that never actually materializes").

[106] Despite actual knowledge of these risks to safety, Holmes promoted her medical misinformation product through various means. *See* Press Release, U.S. Dep't of Just., Theranos Founder and Former Chief Operating Officer Charged In Alleged Wire Fraud Schemes (June 15, 2018), https://www.justice.gov/usao-ndca/pr/theranos-founder-and-former-chief-operating-officer-charged-alleged-wire-fraud-schemes [https://perma.cc/3BXT-U5YM] ("The indictment alleges that the defendants used a combination of direct communications, marketing materials, statements to the media, financial statements, models, and other information to defraud potential investors.").

[107] Indictment, *supra* note 102, at 6.

[108] See, generally, JOHN CARREYROU, BAD BLOOD (2018). Theranos products were a form of IoB vaporware: they were widely advertised but did not exist as promised. *See* Julia Shapero, *Elizabeth Holmes Sentenced to over 11 Years in Prison Following Fraud Conviction*, HILL (Nov. 18, 2022), https://thehill.com/regulation/court-battles/3742503-elizabeth-holmes-sentenced-to-over-11-years-in-prison-following-fraud-conviction/ [https://perma.cc/Y828-AZHH].

[109] *See* Yasmin Khorram, *Former Theranos Patient Testifies Blood Test Gave Her False Miscarriage Diagnosis*, CNBC (Sept. 21, 2021, 7:05 PM), https://www.cnbc.com/2021/09/21/theranos-test-gave-false-miscarriage-diagnosis-witness-testifies.html [https://perma.cc/WFG4-WMJB] ("'She told me your numbers are falling, unfortunately, and that I was miscarrying,' Gould said, getting emotional on the stand."); Yasmin Khorram, *Former Theranos Patient Testifies That a Blood Test at Walgreens Came Back with False Positive for HIV*, CNBC (Nov. 17, 2021, 8:08 PM), https://www.cnbc.com/2021/11/17/theranos-patient-says-blood-test-came-back-with-false-positive-for-hiv.html [https://perma.cc/ZU27-MAUF].

[110] Press Release, SEC, Theranos, CEO Holmes, and Former President Balwani Charged with Massive Fraud (Mar. 14, 2018), https://www.sec.gov/news/press-release/2018-41 [https://perma.cc/N22P-Z5V6] (Holmes and Theranos settled these charges).

harm to market integrity. Facing a maximum sentence of twenty years in prison,[111] she was ultimately sentenced to over eleven years.[112]

But a larger question remains unadjudicated and without satisfactory resolution: Why did Theranos's experienced board of directors,[113] the company's sophisticated business partners,[114] and some of the most prominent investors in Silicon Valley[115] all fail to identify (or why did they choose to ignore) the massive health tech fraud of Theranos?[116]

---

[111] Erin Griffith, *Elizabeth Holmes Denied New Trial and Is Set to Be Sentenced*, N.Y. TIMES (Nov. 8, 2022), https://www.nytimes.com/2022/11/08/technology/elizabeth-holmes-denied-new-trial.html [https://perma.cc/P4RL-JV3L].

[112] Shapero, *supra* note 108.

[113] *See* Lydia Ramsey Pflanzer, *Controversial Health Startup Theranos Has Barely Any Medical Experts on Its Board of Directors*, BUS. INSIDER (Oct. 16, 2015), https://www.businessinsider.com/theranos-board-of-directors-2015-10 [https://perma.cc/9J7F-D3RX] ("To make sure we got all that: that's six former government officials, two former military leaders, two corporation leaders, two members of Theranos' leadership, and two men who graduated from medical school."). For a discussion of what the role of the board could have been, see, for example, Brent T. Wilson, *Theranos and the Tale of the Disappearing Board of Directors*, IDAHO STATE BAR (Mar. 11, 2020), https://isb.idaho.gov/blog/theranos-and-the-tale-of-the-disappearing-board-of-directors/ [https://perma.cc/2TRF-MBJA].

[114] *See, e.g.*, Nick Statt, *Walgreens Brought Theranos to Its Stores Without Even Testing the Technology*, VERGE (May 25, 2016, 3:13 PM), https://www.theverge.com/2016/5/25/11776018/theranos-walgreens-blood-testing-partnership-validation [https://perma.cc/BDG8-ZPRJ] ("Throughout the process, officials shrugged off concerns . . . . The news is yet another unsettling revelation in the troubled partnership between the well-established Walgreens and a startup whose core tech may be more a product of hype than groundbreaking science.").

[115] *See, e.g.*, Sophia Kunthara, *A Closer Look at Theranos' Big-Name Investors, Partneand Board as Elizabeth Holmes' Criminal Trial Begins*, CRUNCHBASE NEWS (Sept. 14, 2021), https://news.crunchbase.com/health-wellness-biotech/theranos-elizabeth-holmes-trial-investors-board/ [https://perma.cc/M8HH-PBRV] ("Some of the most high-profile investors in the company include: Media mogul Rupert Murdoch, who led a $5.8 million Series A in February 2005; Venture capitalist and Draper Fisher Jurvetson partner Tim Draper, who remained an outspoken defender of Theranos at least until 2018; Oracle Executive Chairman and founder Larry Ellison; and National pharmacy and retail chain Walgreens.").

[116] For one theory, see, for example, Tom Relihan, *4 Red Flags that Signaled Theranos' Downfall*, MIT SLOAN (Oct. 29, 2018), https://mitsloan.mit.edu/ideas-made-to-matter/4-red-flags-signaled-theranos-downfall [https://perma.cc/2BTM-5H7Q].

Despite ample warnings that the technology was fatally flawed,[117] why did the board look away from the fraud[118]

The problem of Theranos is a prime example of the dynamic that this Article calls exploit machina — legally problematic scenarios where a broken technology and a broken governance structure combine to cause potentially irreparable harm to the public. But exploit machina problems are not limited to internet-enabled health technologies; Theranos is merely a recent example where flawed "state of the art" technologies[119] and AI and data analytics[120] have combined with flawed organizational governance.[121] Yet, as the Theranos fraud illustrates, exploit machina presents what is arguably the key safety, security, legal, and economic challenge for our future.

---

[117] *See* Erin Griffith, *Theranos Whistle-Blower Testifies She Was Alarmed by Company's Blood Tests*, N.Y. TIMES, https://www.nytimes.com/2021/09/14/technology/elizabeth-holmes-trial-theranos.html (last updated Oct. 6, 2021) [https://perma.cc/36N9-JPL7] ("John Bostic, a prosecutor and an assistant U.S. attorney, argued that documents showing Theranos's internal issues were relevant to the case . . . .").

[118] *See* My Say, *The Theranos Crisis: Where Was the Board?*, FORBES (Apr. 27, 2016, 3:57 PM), https://www.forbes.com/sites/groupthink/2016/04/27/the-theranos-crisis-where-was-the-board/?sh=6e8735f6c58e [https://perma.cc/3UAK-NJRE].

[119] *See* Arvind Narayanan, *How to Recognize AI Snake Oil*, PRINCETON UNIV., https://www.cs.princeton.edu/~arvindn/talks/MIT-STS-AI-snakeoil.pdf [https://perma.cc/5BTM-2JTN] (this problem is framed differently by Professor Arvind Narayanan as one of AI snake oil).

[120] Much excellent scholarship exists on issues of training data bias and other representativeness concerns in AI/ML and data analytics policy. *see, e.g.*, THOMAS S. MULLANEY, BENJAMIN PETERS, MAR HICKS & KAVITA PHILIP, YOUR COMPUTER IS ON FIRE (2021) (collecting essays on various aspects of AI and data equity); *see* SAFIYA UMOJA NOBLE, ALGORITHMS OF OPPRESSION (2018); *see also* RUHA BENJAMIN, RACE AFTER TECHNOLOGY, *passim* (2019).

[121] Another example might be seen in the conduct of fintech lender LendUp, a startup that was subject to multiple enforcement actions by CFPB and a judgment under the Military Lending Act, before ultimately being shut down under a stipulated final judgment. As explained in the CFPB press release, "'LendUp was backed by some of the biggest names in venture capital,' said CFPB Director Rohit Chopra. 'We are shuttering the lending operations of this fintech for repeatedly lying and illegally cheating its customers.'" Press Release, CFPB, CFPB Shutters Lending by VC-Backed Fintech for Violating Agency Order (Dec. 20, 2021), https://www.consumerfinance.gov/about-us/newsroom/cfpb-shutters-lending-by-vc-backed-fintech-for-violating-agency-order/ [https://perma.cc/9DDJ-MTSM].

At its core, exploit machina requires that we *distinguish between trustworthy technologies that advance human thriving,*[122] *on the one hand, and those technologies that are merely variants of fraud and abuse* perpetrated through insider attacks, on the other. Exploit machina scenarios, particularly those in connection with AI safety,[123] may seem a novel byproduct of our modern economy. However, perhaps counterintuitively, they are less new than we might assume. Examples of exploit machina dynamics predate computing.

The next section channels the advice of two intellectual giants, Justice Benjamin Cardozo[124] and developmental and cognitive psychologist Jerome Bruner:[125] it engages the explanatory power of historical examples. The section that follows presents case studies of body sensing devices through the lens of exploit machina and the language of AI, data science, and computer security. It elaborates on the three categories of

[122] *See supra* note 78.

[123] The term "AI safety" is used in this work in a traditional legal sense — the safety of currently existing, deployed AI systems in context for humans today. However, in contrast, the term "AI safety" is used by industry groups to refer to the futuristic possibility of artificial general intelligence seeking to destroy humanity. *See, e.g.*, Samantha Kelly, *Sam Altman Warns AI Could Kill Us All. But He Still Wants the World to Use It*, CNN (Oct. 31, 2023, 6:00 AM), https://www.cnn.com/2023/10/31/tech/sam-altman-ai-risk-taker/index.html [https://perma.cc/XR6Z-62SW] ("Two weeks after the hearing, Altman joined hundreds of top AI scientists, researchers and business leaders in signing a letter stating: 'Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war.'").

[124] *See* BENJAMIN N. CARDOZO, *Law and Literature*, *in* LAW AND LITERATURE AND OTHER ESSAYS AND ADDRESSES 3, 9 (1931) (As Justice Cardozo advised, "only blindness or indifference will fail to turn in all humility, for guidance or for warning, to the study of examples.").

[125] Jerome Bruner wrote in his foundational work *Two Modes of Thought* that:

> A good story and a well-formed argument are different natural [ways of ordering experience]. Both can be used as means for convincing another. Yet what they convince *of* is fundamentally different: arguments convince one of their truth, stories of their lifelikeness. The one verifies by eventual appeal to procedures for establishing formal and empirical proof. The other establishes not truth but verisimilitude.

JEROME BRUNER, ACTUAL MINDS, POSSIBLE WORLDS 11 (1987) (introducing constructivist psychology and "two modes of thought," or "distinctive ways of ordering experience, of constructing reality," — "a good story and a well-formed argument," what Bruner calls "the narrative mode" and "the paradigmatic or logico-scientific").

attacks visible in both historical and modern exploit machina scenarios: attacks on confidentiality, attacks on integrity, and attacks on availability.

### *A. Attacks on Confidentiality: Sensors and Self-Pwns*[126]

The first category of exploit machina involves compromises of confidentiality. In computer security, an attack on the confidentiality of a system refers to the situation where an attacker seeks to trick a system into revealing information in excess of what the creator or user of that system intended.[127] While attacks are usually perpetrated by a third party, perhaps unexpectedly, sometimes that "attacker" is the ultimate insider — the person who designed or installed the system. This phenomenon is known in computer security as a "self-pwn."[128] An (in)famous example might be found in the repurposing of the Watergate tapes, tapes which existed only because of an on-body automatic recording device that President Nixon wore by choice and with knowledge of its operation.[129] In other words, the Watergate tapes might be characterized as an insider attack on system confidentiality, an attack by President Nixon on himself.[130]

Today, members of the public describe their experiences with body sensing devices and data collection in terms reminiscent of insider attacks on confidentiality. For example, patients reliant on CPAP

[126] *What Does 'Pwn' Mean?*, MERRIAM-WEBSTER, https://www.merriam-webster.com/wordplay/pwn-what-it-means-and-how-you-say-it (last visited Sept. 11, 2025) [https://perma.cc/2ZP8-HGAB] (defining pwn as "to have power or mastery over (someone)").

[127] For the NIST definition of confidentiality, see, for example, *Confidentiality*, NAT'L INST. OF STANDARDS & TECH., https://csrc.nist.gov/glossary/term/confidentiality (last visited Sept. 11, 2025) [https://perma.cc/43V3-VTHB], defining confidentiality as "[p]reserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information."

[128] *See, e.g.*, *What Does 'PWN' Mean?*, *supra* note 126 (defining pwn).

[129] *See, e.g.*, Lee M. Smithyman, *John Dean Presents Watergate II*, 82 J. KAN. BAR ASS'N 6 (2013) (explaining that President Nixon oversaw installation of the recording system and wore an on-body device to automatically record Oval Office conversations).

[130] *See id.*; *see also* Russell G. Donaldson, Annotation, *Propriety and Scope of Protective Order Against Disclosure of Material Already Entered into Evidence in Federal Court Trial*, 138 A.L.R. FED. 153 (1997).

machines for breathing during sleep sometimes feel unfairly surprised when they find out that their medical device can repurpose body-derived information in ways potentially adverse to their interests.[131] Patients report that their CPAP machines sometimes misreport use data to remote system or otherwise malfunction; [132] they object that these technical errors[133] have sometimes been construed against them as "proof" of these patients' disuse of the device,[134] fueling erroneous insurance denials.[135] Patients have also raised concerns over unexpectedly gamified business models[136] and risk of financial abuse through rental agreements.[137] In some markets, these practices have also potentially fueled exploration of new kinds of adjacent "wellness" financial products.[138] Thus, despite in theory agreeing to use these medically necessary devices and accepting the privacy policy (and receiving a doctor's explanation), patients still express feeling betrayed by their device's unexpected operations.

But unexpected repurposing of body data by insiders is not limited to medical safety contexts. Consider one company's reported reuse of users' live body location data as a form of entertainment at corporate events.[139] For some users, this unexpected reuse of real-time data from

---

[131] As explained by one patient:

> You view it as a device that is yours and is serving you . . . . And suddenly you realize it is a surveillance device being used by your health insurance company to limit your access to health care. . . . I wish they would spend as much time providing me actual care as they do monitoring whether I'm "compliant."

Marshall Allen, *You Snooze, You Lose: Insurers Make the Old Adage Literally True*, PROPUBLICA (Nov. 21, 2018, 5:00 AM), https://www.propublica.org/article/you-snooze-you-lose-insurers-make-the-old-adage-literally-true [https://perma.cc/R3RT-S92D].

[132] *See* Allen, *supra* note 131.

[133] *Id.*

[134] *Id.*

[135] *Id.*

[136] *Id.* ("'Congratulations! You've earned yourself a badge!' the email said.").

[137] *Id.*

[138] Mithun Dasgupta, *Insurers Seek Clarity on Other Financial Products*, FIN. EXPRESS (Dec. 7, 2022, 06:15 IST), https://www.financialexpress.com/money/insurance-insurers-seek-clarity-on-other-financial-products-2903084/ [https://perma.cc/C5F3-LF9U].

[139] *See* Ben Smith, *Uber Executive Suggests Digging Up Dirt on Journalists*, BUZZFEED NEWS (Nov. 17, 2014, 8:57 PM),

(on-body) mobile phones was unnerving; some users perceived it as a risk to their personal safety.[140] Problematic data reuse also became part of the company's settlements with the New York Attorney General[141] and the FTC.[142] Or consider the case of an Ohio man who was prosecuted for arson and insurance fraud in reliance in part on his own pacemaker data.[143] Similarly, the recent South Carolina murder trial and conviction of an attorney included location data from the defendant's phone provider and car, which were central to establishing his movements.[144] But, even if these situations might seem unproblematic due to the nature of the alleged criminality, consider the recent press accounts of social media companies allegedly sharing user location and other information with law enforcement organizations that are potentially seeking to prosecute defendants[145] who had an abortion,[146]

https://www.buzzfeednews.com/article/bensmith/uber-executive-suggests-digging-up-dirt-on-journalists [https://perma.cc/WGS2-2UXU].

[140] *Id.*

[141] Johana Bhuiyan, *Uber Settles with New York Attorney General over 'God View' Tracking Program*, BUZZFEED NEWS (Jan. 6, 2016), https://www.buzzfeednews.com/article/johanabhuiyan/uber-settles-godview [https://perma.cc/X46F-3QJX].

[142] Press Release, FTC, Federal Trade Commission Gives Final Approval to Settlement with Uber (Oct. 26, 2018), https://www.ftc.gov/news-events/news/press-releases/2018/10/federal-trade-commission-gives-final-approval-settlement-uber [https://perma.cc/J2H7-V68W]; Brian Fung, *Uber Settles with FTC over 'God View' and Some Other Privacy Issues*, L.A. TIMES (Aug. 15, 2017, 3:40 PM), https://www.latimes.com/business/technology/la-fi-tn-uber-ftc-20170815-story.html [https://perma.cc/5H74-FBTQ].

[143] *Man's Pacemaker Data Used Against Him in Arson Case*, CBS NEWS, https://www.cbsnews.com/news/mans-cardiac-pacemaker-data-led-to-arson-charges/ (last updated Feb. 11, 2017, 1:15 PM) [https://perma.cc/SYB5-PQEG].

[144] Maddie Saines, *Phone Location Data Is Center Stage at Murdaugh Trial*, GPS WORLD (Feb. 23, 2023), https://www.gpsworld.com/phone-location-data-is-center-stage-at-murdaugh-trial/ [https://perma.cc/T9DX-NRDP]; Jeffrey Collins, *Both Sides Use Trove of Cell Data at Alex Murdaugh Trial*, ASSOCIATED PRESS NEWS (Feb. 1, 2023, 3:34 PM), https://apnews.com/article/south-carolina-homicide-crime-3da14ec557407f0a253b460bee9573f1 [https://perma.cc/E5P4-9JC7].

[145] CHRIS D. LINEBAUGH, CONG. RSCH. SERV., LSB10786, ABORTION, DATA PRIVACY, AND LAW ENFORCEMENT ACCESS: A LEGAL OVERVIEW 1 (2022), https://crsreports.congress.gov/product/pdf/LSB/LSB10786 [https://perma.cc/M2LR-FJ78].

[146] Jason Koebler & Anna Merlan, *This Is the Data Facebook Gave Police to Prosecute a Teenager for Abortion*, VICE (Aug. 9, 2022, 2:45 PM),

conduct that is legal in many states.[147] Or consider the potentially problematic governance practices of a dating safety application that may have resulted in disclosure of users' private messages regarding abortion services, phone numbers, and other sensitive matters, as well as their verification images and personal IDs potentially being posted on 4chan.[148] Again, users perceive these dynamics as insider attacks on confidentiality[149] and as a threat to their physical safety; some of these unexpected data reuses may remind users of a form of pre-crime dragnet, one conducted with the help of an organization that they unwisely trusted.[150]

https://www.vice.com/en/article/n7zevd/this-is-the-data-facebook-gave-police-to-prosecute-a-teenager-for-abortion [https://perma.cc/8YRY-UGMP].

[147] The right to travel is also protected. For a discussion of the Constitutional right to travel, see, for example, *Saenz v. Roe*, 526 U.S. 489 (1999). Justice Stevens, writing for the majority, explained that:

> The word "travel" is not found in the text of the Constitution. Yet the "constitutional right to travel from one State to another" is firmly embedded in our jurisprudence. United States v. Guest, 383 U.S. 745, 757 (1966). Indeed, as Justice Stewart reminded us in *Shapiro v. Thompson*, 394 U.S. 618 (1969), the right is so important that it is "assertable against private interference as well as governmental action . . . a virtually unconditional personal right, guaranteed by the Constitution to us all."

*Id.* at 643 (concurring opinion) (citations abbreviated).

[148] *See* Chase DiBenedetto, *New Hack of Women-Only App Tea Exposes Personal Chats, Phone Numbers*, MASHABLE (July 28, 2025), https://mashable.com/article/tea-app-data-break-chats-abortion [https://perma.cc/Z5TJ-HCFH] ("In addition to exposing thousands of user verification images and personal IDs, which were later abused by users on platforms like 4Chan, the app's recently discovered security flaws make it possible for hackers to access private messages between users includ[ing] sensitive information like shared phone numbers, conversations about intimate relationships, and discussions of abortion.").

[149] *See* Alfred Ng, '*A Uniquely Dangerous Tool': How Google's Data Can Help States Track Abortions*, POLITICO (July 18, 2022, 4:30 AM), https://www.politico.com/news/2022/07/18/google-data-states-track-abortions-00045906 [https://perma.cc/U8R4-K869].

[150] Although the constitutionality of such pre-crime dragnet lists is in doubt, particularly when not accompanied by warrants or when constructed on the basis of gender or race, some companies choose to share user information to the detriment of their users as the path of least legal resistance. *See, e.g.*, David Gray, *Bertillonage in an Age of Surveillance: Fourth Amendment Regulation of Facial Recognition Technologies*, 24 SMU SCI. & TECH. L. REV. 3, 9-10 (2021) (arguing that "Carpenter instructs us to focus on the

The organizations managing these technologies, on the other hand, would likely allege that, like President Nixon, users had consented to their self-pwn through end user license agreements or terms of use ("EULAs").[151] However, in *Carpenter v. United States*, the Supreme Court signaled that such arguments have limits;, in particular, the Court noted that the third party doctrine is stretched too thin in some modern on-body device contexts, even despite ostensible user consent through a EULA.[152] The Court reiterated its position that consent to a EULA has limits in *Van Buren v. United States*, a case involving consent to a government employer's permitted use policies and a prosecution under the Computer Fraud and Abuse Act.[153] Legal scholars have also amply

technology at issue, to ask about the extent to which the information it gathers might reveal intimate details about our lives, including our associations and activities, the 'retrospective quality of the data,' whether the data can be 'store[d] and efficiently mine[d] for information years into the future,' whether the technology can be scaled up easily, facilitating 'dragnet-type law enforcement practices' such as 'twenty-four hour surveillance of any citizen of this country,' whether the technology 'by design, proceeds surreptitiously,' and whether the deployment and use of the technology 'evades the ordinary checks that constrain abusive law enforcement practices: limited police resources and community hostility.'").

[151] For example, a defendant in a wrongful death suit recently argued that the terms of use that accompanied a television service extended liability protection to the company for harms in connection with the death of a restaurant patron at an affiliated restaurant. *See* Minyvonne Burke, *Disney Says Man Can't Sue Over Wife's Death Because He Agreed to Disney+ Terms of Service*, NBC NEWS (Aug. 14, 2024, 1:54 PM), https://www.nbcnews.com/news/us-news/disney-says-man-cant-sue-wifes-death-agreed-disney-terms-service-rcna166594.

[152] *See* Carpenter v. United States, 585 US 296, 305-09 (2018). ("As technology has enhanced the Government's capacity to encroach upon areas normally guarded from inquisitive eyes, this Court has sought to 'assure [ ] preservation of that degree of privacy against government that existed when the Fourth Amendment was adopted.' Kyllo v. United States, 533 U. S. 27, 34 (2001). . . . This sort of digital data — personal location information maintained by a third party — does not fit neatly under existing precedents. . . . Given the unique nature of cell phone location records, the fact that the information is held by a third party does not by itself overcome the user's claim to Fourth Amendment protection.").

[153] Van Buren v. United States, 593 U.S. 374 (2021) ("If the 'exceeds authorized access' clause criminalizes every violation of a computer-use policy, then millions of otherwise law-abiding citizens are criminals.").

considered questions of the limits of consent[154] in criminal,[155] civil,[156] and other[157] contexts. But considerations of technology safety and

[154] Professors Evan Selinger and Woodrow Hartzog argue that "valid consent cannot be given for face surveillance" and advocate the enactment of moratoria "to prevent entrenchment of and dependence on facial recognition systems before they can be properly considered by lawmakers and society." Evan Selinger & Woodrow Hartzog, *The Inconsentability of Facial Surveillance*, 66 LOY. L. REV. 101, 101, 105 (2019). Meanwhile, Professor Elizabeth Rowe cautions against regulating facial recognition and related technologies as an "all or nothing or one-size-fits-all endeavor." Elizabeth A. Rowe, *Regulating Facial Recognition Technology in the Private Sector*, 24 STAN. TECH. L. REV. 1, 48 (2020) (hyphens omitted); *see also* Nila Bala, *The Danger of Facial Recognition in Our Children's Classrooms*, 18 DUKE L. & TECH. REV. 249, 260 (2020); Lindsey Barrett, *Ban Facial Recognition Technologies for Children — and for Everyone Else*, 26 B.U. J. SCI. & TECH. L. 223, 223-24 (2020).

[155] Professor Monika Zalnieriute explains that "[l]ive automated facial recognition technology, rolled out in public spaces and cities across the world, is transforming the nature of modern policing." Monika Zalnieriute, *Burning Bridges: The Automated Facial Recognition Technology and Public Space Surveillance in the Modern State*, 22 COLUM. SCI. & TECH. L. REV. 284, 284 (2021). Professor Andrew Guthrie Ferguson proposes a framework embodying Fourth Amendment limits on law enforcement use of facial recognition. Andrew Guthrie Ferguson, *Facial Recognition and the Fourth Amendment*, 105 MINN. L. REV. 1105, 1109 (2021); Andrew Guthrie Ferguson, *Surveillance and the Tyrant Test*, 110 GEO. L.J. 205, 205 (2021); *cf.* Andrew Guthrie Ferguson, *Structural Sensor Surveillance*, 106 IOWA L. REV. 47, 47 (2020). Professor David Gray advocates striking a "balance, allowing for the reasonable use of facial recognition while guarding against its use to facilitate broad, indiscriminate, and intrusive searches." Gray, *supra* note 150, at 9-10. Professor Marc Blitz engages the limitations of the Fourth Amendment, arguing that its current constraints are not well-suited for analyzing emerging video surveillance systems such as facial recognition. *See* Marc Jonathan Blitz, *Video Surveillance and the Constitution of Public Space: Fitting the Fourth Amendment to Aa World That Tracks Image and Identity*, 82 TEX. L. REV. 1349, 1365 (2004). Professor Kimberly Brown argues that "Fourth and First Amendment doctrine should be reconciled to address the manipulation — versus acquisition — of [facial recognition] data to derive new information about individuals which is exceedingly intimate and otherwise out of the government's reach." Kimberly N. Brown, *Anonymity, Faceprints, and the Constitution*, 21 GEO. MASON L. REV. 409, 411 (2014). Professor Laura Moy offers a model equity impact assessment for proposed police technologies. Laura M. Moy, *A Taxonomy of Police Technology's Racial Inequity Problems*, 2021 U. ILL. L. REV. 139, 139-40 (2021). Professor Hannah Bloch-Wehba explains that "[i]n the absence of meaningful Fourth Amendment safeguards, transparency litigation makes policing practices increasingly visible to the public and democratic institutions in areas where constitutional criminal procedure today has minimal reach." Hannah Bloch-Wehba, *Visible Policing: Technology, Transparency, and Democratic Control*, 109 CALIF. L. REV. 917, 918 (2021).

insider attacks are not limited to issues of confidentiality: the most complex insider attacks often involve manipulations of integrity and availability of systems and the information within them.[158]

### *B. Attacks on Integrity: Data Lakes[159] and Drowning Witches*

The second category of exploit machina attacks involves compromises of integrity. In computer security, an insider attack

[156] Professor Tiffany Li cautions that downstream harms can include the sharing and reselling of data or the inclusion of personal data in machine learning systems, including those that are used in facial recognition technologies. Tiffany C. Li, *Post-Pandemic Privacy Law*, 70 AM. U. L. REV. 1681, 1717 (2021).

[157] Professor Catherine Crump points out the role of procurement in surveillance policy. *See* Catherine Crump, *Surveillance Policy Making by Procurement*, 91 WASH. L. REV. 1595, 1595 (2016). A robust legal literature also exists around the limits of contractual consent formation and end user license agreements in data privacy and security contexts. Andrea M. Matwyshyn, *Privacy, the Hacker Way*, 87 S. CAL. L. REV. 1 (2013) (arguing in favor of a "reasonable data stewardship" approach that relies on a set of implied promises — nonwaivable contract warranties and remedies in EULAs) [hereinafter Matwyshyn, *Privacy, the Hacker Way*]; Woodrow Hartzog, *Website Design As Contract*, 60 AM. U. L. REV. 1635, 1635 (2011) (explaining that when courts seek to determine a website user's privacy expectations and the website's promises to that user, they almost invariably look to the terms of use agreement or to the privacy policy); Daniel J. Solove, *Murky Consent: An Approach to the Fictions of Consent in Privacy Law*, 104 B.U. L. REV. 593, 593 (2024) (arguing that most of the time, privacy consent is fictitious); Mark Lemley, *Protecting Consumers in a Post-Consent World*, 77 STAN. L. REV. ONLINE 247, 248 (2025) (explaining that notice and consent offer neither notice nor consent); Lilian Edwards, Edina Harbinja, *Protecting Post-Mortem Privacy: Reconsidering the Privacy Interests of the Deceased in A Digital World*, 32 CARDOZO ARTS & ENT. L.J. 83, 86 (2013) (explaining how EULAs contribute to poor protection for the postmortem privacy of the deceased).

[158] *See, e.g.*, Joseph Cox, *Hacker Plants Computer 'Wiping' Commands in Amazon's AI Coding Agent*, 404 MEDIA (July 23, 2025, 9:48 AM), https://www.404media.co/hacker-plants-computer-wiping-commands-in-amazons-ai-coding-agent/ [https://perma.cc/H4QV-QKFC] ("A hacker compromised a version of Amazon's popular AI coding assistant 'Q', added commands that told the software to wipe users' computers, and then Amazon included the unauthorized update in a public release of the assistant").

[159] *See generally,* Matthew Kosinski, *What Is a Data Lake?*, IBM (Jan. 16, 2025), https://www.ibm.com/think/topics/data-lake [https://perma.cc/29AX-UYUN] (defining a data lake as "a low-cost data storage environment designed to handle massive amounts of raw data in any format, including structured, semi-structured and unstructured data.").

against integrity refers to the manipulation or corruption of information and processes contained within a system by an insider.[160] Perhaps surprisingly, when viewed through the lens of exploit machina, the seventeenth- and eighteenth-century witchcraft trials in England and the British colonies offer a historical example of insider attacks on (legal) system integrity. Specifically, these trials leveraged body sensing "technology" and early forms of prescriptive analytics.

For example, the witchcraft Tryal at Bury St. Edmunds, presided over by Lord Matthew Hale, relied heavily on pseudoscientific "body sensor" technologies and fraudulent prescriptive data analytics[161] as the primary evidence against the defendants.[162] As scholars have explained, at a time

[160] *Integrity*, NIST, https://csrc.nist.gov/glossary/term/integrity (last visited Sept. 11, 2025) [https://perma.cc/96RU-8BRX].

[161] Experiments included invasive physical exams of the defendant, searching for any indication of a mole or skin abnormality — allegedly a body-sensed mark of consort with the devil — as well as spectral evidence generated in court, where accusers allegedly writhed due to sensing the presence of the body of the defendant or would appear to regurgitate pins upon sensing the presence of the body of the defendant. For example:

> Poor Rose Cullender, moreover, was made to furnish evidence against herself. Mary Chandler, Susan's mother, being appointed with five other women by Sir Edmond Bacon, the Magistrate who issued the Warrant on the complaint of Mr. Pacey, 'to search the Bodies of the Prisoners,' they found in the abdominal region of Rose, 'something like a teat about an inch long,' and then a smaller one. Of course, these were simple hernias and were so explained by Rose — but in vain, they were clearly the identifying marks of a favorite of Satan.

William Renwick Riddell, *Sir Matthew Hale and Witchcraft*, 17 J. CRIM. L. & CRIMINOLOGY 5, 10-11 (1926).

[162] As explained by Professor Andrew Amos in 1856, citing to the official account of the trial:

> The celebrated trial before Sir. M. Hale of two old women for witchcraft, took place at Bury, in the year 1665. Rose Callender and Amy Duny, widows, were tried before him for bewitching seven people, and assuming for the purpose various shapes, as those of a bee, and a mouse; but collateral evidence was given of bewitching two carts, a chimney, and a firkin of fish. The provocation alleged for the malice of the supposed witches, was the refusal of some herrings . . . . Among other absurdities at the trial, an experiment was tried of the supposed witches being made to touch the persons bewitched, whereupon they would suddenly shriek out and open their hands, which were before clenched, that would not happen on the touch of any other person . . . . Upon an ingenious person suggesting that there might be a fallacy in the experiment,

when witchcraft trials were already viewed by many jurists[163] and members of the public as inherently fraudulent,[164] a fraction of judges, including Lord Hale, nevertheless chose to ignore these concerns and still willingly admitted evidence from problematic "experiments," even over the objections of law enforcement experts in court.[165] Scholars note that the experiments' structures and metrics lacked scientific substantiation and were designed by committees of wealthy locals, often with vested financial interests in the outcomes.[166] Scholars have similarly highlighted accusers' personal interests in assisting with the elimination of people considered challenging or undesirable,[167] often for

> Lord Cornwallis, Mr. Serjt. Keeling, and some other gentlemen in court were deputed to examine one of the distempered women together with one of the supposed witches in the further part of the hall; an apron was put before the eyes of the woman said to be bewitched, and then another person, not of the accused, touched the woman's hand, "which produced the same effects as the touch of the witch did in court." Whereupon the reporter . . . adds, "the gentlemen returned, openly protesting that they did believe the whole transaction of this business was a mere imposture."

Amos points out that Lord Hale ignored the opinions of these third parties, one of whom would shortly thereafter become chief judge, and instead sua sponte solicited contradictory testimony from a known witchcraft believer to contradict the controlled test result. ANDREW AMOS, RUINS OF TIME EXEMPLIFIED 240 (1856); *see also* A TRYAL OF WITCHES AT THE ASSIZES HELD AT BURY ST. EDMONDS 17 (London, John Russell Smith 1838).

[163] As explained by Professor Amos, Hale's "contemporaries on the bench were not all equally disposed." AMOS, *supra* note 162, at 241.

[164] *See* IVAN BUNN & GILBERT GEIS, A TRIAL OF WITCHES: A SEVENTEENTH CENTURY WITCHCRAFT PROSECUTION Preface, at xii (1997) ("It is vital to appreciate that by 1662, English opinion no longer was widely consensual regarding the reality of witchcraft").

[165] *See id.*

[166] *See infra* text accompanying note 170.

[167] *Witchcraft*, UK PARLIAMENT, https://www.parliament.uk/about/living-heritage/transformingsociety/private-lives/religion/overview/witchcraft/ (last visited Sept. 11, 2025) [https://perma.cc/65XF-89BW]. In fact, a cottage industry of for-profit witch finders arose that would happen to identify precisely the local deemed most problematic by locals while creating a patina of impartiality. *See, e.g.*, https://www.thebritishacademy.ac.uk/podcasts/the-rise-and-fall-of-matthew-hopkins-witchfinder-general/ (describing the role of Matthew Hopkins in false accusations of witchcraft for personal profit).

economic revenge[168] or to advance anticompetitive tactics in, for example, beermaking.[169]

In a perhaps similar manner, today's body sensing devices increasingly rely on sensors of variable quality[170] and machine learning processes with demonstrated technological limitations; consequently, these devices may leave their users vulnerable to attacks on integrity, either by external attackers[171] or by insiders' choices.[172] In a medical device context, consider the long-known integrity shortcomings of device sensors.[173] Yet these known problems have gone largely uncorrected in,

[168] A dispute over a refused herring sale may have in part caused the Trial at Bury St. Edmunds. *See* Riddell, *supra* note 161, at 9.

[169] *See* Laken Brooks, *Why Did Women Stop Dominating the Beer Industry?*, SMITHSONIAN MAG. (Mar. 8, 2021), https://www.smithsonianmag.com/history/women-used-dominate-beer-industry-until-witch-accusations-started-pouring-180977171/ [https://perma.cc/VH3X-EJXM]; Christina Wade, *Witchcraft, Alewives, and Economics*, BRACIATRIX (Aug. 7, 2017), https://braciatrix.com/2017/08/07/witchcraft-alewives-and-economics/ [https://perma.cc/KGH2-MMPB]. *But see* Christina Wade, *Nope, Medieval Alewives Aren't the Archetype for the Modern Pop Culture Witch.*, BRACIATRIX (Oct. 27, 2017), https://braciatrix.com/2017/10/27/nope-medieval-alewives-arent-the-archetype-for-the-modern-pop-culture-witch/ [https://perma.cc/8NCB-WE9Y].

[170] Inherent hardware limitations on accuracy of sensor perception also generate imperfect data that feeds automated judgments. *See, e.g.*, Timothy B. Lee, *NTSB: Autopilot Steered Tesla Car Toward Traffic Barrier Before Deadly Crash*, ARSTECHNICA (June 7, 2018, 8:50 AM), https://arstechnica.com/cars/2018/06/ntsb-autopilot-steered-tesla-car-toward-traffic-barrier-before-deadly-crash/ [https://perma.cc/HCC6-PP5N] ("At four seconds before the crash, the Tesla vehicle was no longer following the car ahead of it. The car's cruise control was set to 75mph, so it began to accelerate, reaching a speed of 70.8mph just before the crash. There was 'no precrash braking or evasive steering movement detected,' the NTSB says.").

[171] *See* Gerald Lynch, *Driverless Cars Can Be Tricked with Simple Stickers on Road Signs*, TECHRADAR (Aug. 7, 2017), https://www.techradar.com/news/driverless-cars-can-be-tricked-with-simple-stickers-on-road-signs [https://perma.cc/4W46-ZX9A].

[172] For example, a driver might expect that a car in self-driving mode would by default obey state law regarding stop signs. Yet that may not be the car's default programming. *See* Matthew Sparkes, *Tesla Recalls 50,000 Cars that Disobey Stop Signs in Self-Driving Mode*, NEW SCIENTIST (Feb. 3, 2022), https://www.newscientist.com/article/2307147-tesla-recalls-50000-cars-that-disobey-stop-signs-in-self-driving-mode/ [https://perma.cc/JGD6-HBRZ].

[173] "It [has] been known for decades that skin pigmentation and melanin [levels] can affect a pulse oximeter's ability to accurately measure oxygen saturation." Haley Bridger, *Skin Tone and Pulse Oximetry*, HARVARD MED. SCH. (July 14, 2022),

for example, pulse oximeters, shortfalls that could in theory be corrected either in the devices themselves through improvements in sensor technology, or through compensating human controls in practice.[174] As a result, these sensor shortcomings may have contributed to the higher rate of COVID complications and deaths in some segments of the population.[175] Compensating controls may have been inadequate.[176] In other sensor contexts, trusting that human insiders will prevent and correct sensor malfunction has a recent history of deadly failures.[177] But some of the most potentially troubling exploit

https://hms.harvard.edu/news/skin-tone-pulse-oximetry [https://perma.cc/WS3L-LVS2].

[174] *See id.*

[175] Ana M. Cabanas, Macarena Fuentes-Guajardo, Katina Latorre, Dayneri León & Pilar Martín-Escudero, *Skin Pigmentation Influence on Pulse Oximetry Accuracy: A Systematic Review and Bibliometric Analysis*, 22 SENSORS, April 29, 2022, at 1.

[176] Particularly when these technology shortcomings exacerbate preexisting inequalities in the healthcare system, human thriving becomes impacted. For a discussion of preexisting inequalities in the health care system, see, for example, DAYNA BOWEN MATTHEW, JUST HEALTH (2022), presenting evidence of discrimination in housing, education, employment, and the criminal justice system to argue that racial inequality cumulatively undermines the health of minority populations in the United States.

[177] *See, e.g.*, Rene Marsh & Gregory Wallace, *Whistleblower Testifies that Boeing Ignored Pleas to Shut Down 737 MAX Production*, CNN (Dec. 11, 2019, 6:10 PM), https://www.cnn.com/2019/12/11/politics/fatally-flawed-737-max-had-significantly-higher-crash-risk-faa-concluded/index.html (last updated Dec. 11, 2019, 6:10 PM) [https://perma.cc/BKH7-493A] (explaining that a whistleblower highlighted that faulty sensor data should have been a red flag of a problem during testing, "because the sensor in question is a 'historically reliable part'"). For other potential connections between real-time tracking and physical, psychological, and national security harms, see, for example, Brian Fung, *Execs Ignored the Damage Instagram Does to Teens, Meta Whistleblower Tells Congress*, CNN BUS., https://www.cnn.com/2023/11/07/tech/meta-ignored-warnings-instagrams-harm/index.html (last updated Nov. 7, 2023, 3:44 PM) [https://perma.cc/5UZQ-CVWT] (Senator Josh Hawley "accused Meta of 'cooking the books' on data related to mental health harms."); and Joseph Menn, Elizabeth Dwoskin & Cat Zakrzewski, *Former Security Chief Claims Twitter Buried 'Egregious Deficiencies,'* WASH. POST, https://www.washingtonpost.com/technology/interactive/2022/twitter-whistleblower-sec-spam/ (last updated Aug. 23, 2022, 12:27 PM) [https://perma.cc/G8L9-QP9G]:

> The complaint — filed last month with the Securities and Exchange Commission and the Department of Justice, as well as the FTC — says

machina integrity issues may arise in mandatory body monitoring scenarios, particularly those that involve children. Consider a potentially problematic device currently in use by school districts in multiple states[178] with kindergarten aged children, a so-called on-body "talk pedometer" that alleges to measure a "critical factor . . . in early brain development,"[179] a metric that a school administrator might then use to prejudge (potentially in error and therefore artificially limit) children's (alleged) future learning prospects.

Particularly if the use of a body-sensing technology is unavoidable or mandatory, integrity failures ignored by insiders can quickly lead to irreparable harms. In these contexts, as in the witchcraft trials, a slippery slope frequently exists: monitoring human bodies with sensors often results in not only a(n alleged) measurement but also a prejudgment about whether the measured body is "good" or "undesirable." Legal scholars have ably elaborated on the privacy[180] and

> thousands of employees still had wide-ranging and poorly tracked internal access to core company software, a situation that for years had led to embarrassing hacks, including the commandeering of accounts held by such high-profile users as Elon Musk and former presidents Barack Obama and Donald Trump. . . . For example, [current and former employees] said the company implied that it had destroyed all data on users who asked, but the material had spread so widely inside Twitter's networks, it was impossible to know for sure.

[178] *See School Districts Use LENA Programs to Improve Kindergarten Readiness, Child Outcomes*, LENA (May 12, 2020) https://www.lena.org/school-districts/ [https://perma.cc/83WM-FHV9].

[179] *See* K.C. Compton, *In Babies' Brains, White Matter Is Crucial — and Conversational Turns Make It Grow*, EARLY LEARNING NATION (Mar. 14, 2023), https://earlylearningnation.com/2023/03/in-babies-brains-white-matter-is-crucial-and-conversational-turns-make-it-grow/?mc_cid=9bb96543f9&mc_eid=bc3fb7d497 [https://perma.cc/V5UB-JT4G].

[180] *See, e.g.*, Margot E. Kaminski, Matthew Rueben, William D. Smart & Cindy M. Grimm, *Averting Robot Eyes*, 76 MD. L. REV. 983 (2017) (arguing that in-home sensing robots "will raise privacy problems of three basic types: (1) data privacy problems; (2) boundary management problems; and (3) social/relational problems."); Jonathan Turley, *Anonymity, Obscurity, and Technology: Reconsidering Privacy in the Age of Biometrics*, 100 B.U. L. REV. 2179, 2179 (2020) (recommending "a comprehensive approach to biometric technology that would obscure increasingly available images and data while recasting privacy protections to fit a new and unfolding biometric reality"). *See generally* Mike Hintze, *Science and Privacy: Data Protection Laws and Their Impact on Research*, 14

procedural fairness[181] considerations raised by sensing and scoring technologies.[182] Exploit machina considerations expand upon the important cautionary notes raised by these scholars, focusing attention on the negative self-reinforcement when broken technologies exist in the context of broken governance.

### *C. Attacks on Availability: Access and X-Ray Specs*[183]

The final category of exploit machina involves availability. In computer security, an attack on availability refers to an external or insider attacker's interference with readiness of information and systems, preventing them from being accessible on an as-needed basis.[184] Availability issues and problematic insider conduct also predate modern technology. For example, in the 1970s New York City faced allegations that mobile tuberculosis monitoring trucks, a form of body-sensing technology, were selectively unavailable.[185] As tuberculosis

WASH. J.L. TECH. & ARTS 103, 111-12 (2019) (addressing "how privacy laws can and should allow for scientific research while still providing meaningful protections for personal information" and providing specific recommendations).

[181] Professor Frank Pasquale and Professor Danielle Citron argue in favor of procedural regularity to compensate for the opacity and deficits of oversight in artificially intelligent scoring systems. They argue that "a regime of total surveillance undermines the free development of personality upon which free expression depends" and, citing Professor Daniel Solove, explain that "privacy is not just a problem of concealing isolated facts." Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1413, 1417-19 (2014).

[182] Professor Margaret Hu includes first generation law enforcement IoB devices — body cameras and smart glasses — and cautions against "[t]he potential of ubiquitous and continuous data collection and analysis." Margaret Hu, *Bulk Biometric Metadata Collection*, 96 N.C. L. REV. 1425, 1425-26 (2018); *see also* Chaz Arnett, *Race, Surveillance, Resistance*, 81 OHIO ST. L.J. 1103, 1103 (2020).

[183] Like many of the technologies discussed in this article, this section heading is dual use, intended both as a summary of the section regarding X-ray truck specifications that follows and as a quiet acknowledgement of Poly Styrene, the leader of the iconic 1970s punk band X-Ray Spex and her innovation in music, despite facing access challenges from insiders. *See Oh Bondage up Yours!*, X-RAY SPEX, https://www.x-rayspex.com/ (last visited Sept. 11, 2025) [https://perma.cc/YQS7-T5MC].

[184] NIST, *supra* note 12.

[185] *See* Alfonso A. Narvaez, *The Young Lords Seize X-Ray Unit*, N.Y. TIMES (June 18, 1970), https://www.nytimes.com/1970/06/18/archives/the-young-lords-seize-xray-unit-take-it-to-area-where-they-say-it.html [https://perma.cc/TG9B-5VLT] ("[A] Health

infections exploded in the city, residents engaged in their own data collection. Armed with this data, they alleged that access to the trucks did not map to the neighborhoods that were most impacted by the tuberculosis outbreak; they feared that insider manipulation of truck availability was afoot.[186] Ultimately, when they failed to persuade officials with their data and entreaties, desperate residents engaged in an illegal form of self-help, hijacking a truck.[187]

In the twenty-first century, strategic unavailability with the help of AI is potentially in use to exclude certain "undesirable" people from access to opportunities.[188] Consider the recent case of a New Jersey attorney who was repeatedly[189] denied admittance to events at Madison Square Garden, allegedly because facial recognition connected her with an "undesirable" associate: a member of her law firm who was involved in litigation against the operators of the venue.[190] Or consider the recent shortcomings of mandatory body sensing devices used in bar exam administration systems, systems that prevented some test takers from

Department spokesman said the tuberculosis rate in East Harlem was 68 cases for every 100,000 people — about double the rate outside the city's poorer neighbor hoods. The national average, he said, is 21 for every 100,000. . . . Mr. Guzman said the truck had been operating only from noon to 6 P.M. every other day, in various parts of the city. He said that in the poorer neighborhoods, such as East Harlem and the South Bronx, most persons who were working were unable to make use of the service during those hours.").

[186] *See id.*

[187] An activist group, The Young Lords, hijacked a tuberculosis monitoring truck and redirected it to the most impacted areas. *See id.* ("'On Monday,' Mr. Guzman said, 'we went to the Tuberculosis Association and asked them for a truck. They said our request was ridiculous. So, at noon, we walked in and we took it.'").

[188] *See* Sarah Wallace, *Face Recognition Tech Gets Girl Scout Mom Booted from Rockettes Show — Due to Where She Works*, NBC, https://www.nbcnewyork.com/investigations/face-recognition-tech-gets-girl-scout-mom-booted-from-rockettes-show-due-to-her-employer/4004677/ (last updated Dec. 20, 2022, 3:27 PM) [https://perma.cc/FS5K-MFWL].

[189] *See generally* Kashmir Hill, *Which Stores Are Scanning Your Face? No One Knows*, N.Y. TIMES, https://www.nytimes.com/2023/03/10/technology/facial-recognition-stores.html (last updated June 5, 2023) [https://perma.cc/PL8E-RVGL] ("The City Council convened a hearing last month to discuss how Madison Square Garden and other local businesses were using the technology. . . . Madison Square Garden hadn't sent a representative, as requested. . . . And no one at the hearing knew which other businesses were using the technology.").

[190] Wallace, *supra* note 188.

having access to their exam.[191] The testing company alleged that the failure was due to a "sophisticated attack."[192] However, some users alleged that representatives of the test administration company had provided contrary information to them: that it was, in reality, a governance shortfall — allegedly a systems management issue in connection with a software upgrade.[193] Legal scholars have also extended the discussion of selective availability into adjacent

[191] Monica Chin, *ExamSoft's Proctoring Software Has a Face-Detection Problem*, VERGE (Jan. 5, 2021, 6:21 PM), https://www.theverge.com/2021/1/5/22215727/examsoft-online-exams-testing-facial-recognition-report [https://perma.cc/C25Q-CD2Y].

[192] ExamSoft (@ExamSoft), X (July 28, 2020, 3:34 PM), https://twitter.com/examsoft/status/1288241360039010307 [https://perma.cc/S78U-TD3R].

[193] *See The Biggest Bar Exam Disaster Ever? ExamSoft Makes Everyone's Life Hard*, ABOVE L. (July 29, 2014), https://abovethelaw.com/2014/07/bar-exam-disaster-examsoft-makes-everyones-life-hard/ ("And then ExamSoft isn't even prepared for the surge despite having X years of notice. And their customer service is just a busy signal"); *When Software Fails a Bunch of Future Lawyers*, NETWORKWORLD (Aug. 12, 2014), https://www.networkworld.com/article/928068/software-when-software-fails-a-bunch-of-future-lawyers.html ("a recent software upgrade caused the snafu"). In other words, allegedly the problems arose from an intentional infrastructure management choice. *See* Joe Patrice, *ExamSoft Tells Senators That Facial Recognition Problems Are Everyone's Fault But Theirs*, ABOVE L. (Feb. 1, 2021), https://abovethelaw.com/2021/02/examsoft-tells-senators-that-facial-recognition-problems-are-everyones-fault-but-theirs/ [https://perma.cc/UEA4-DAGQ]. Historically, availability failures have sometimes been a symptom of common corporate cost cutting errors, unwisely assuming risk of infrastructure failure. Because a DDoS attack and inadequate server capacity are experienced in a parallel manner by users, sometimes self-sabotage through suboptimal governance is not always immediately apparent during an incident. *See, e.g.*, Moira Alexander, *5 Ways Poor Capacity Planning Can Sabotage a Project*, TECHREPUBLIC (Jan. 4, 2019), https://www.techrepublic.com/article/5-ways-poor-capacity-planning-can-sabotage-a-project/ [https://perma.cc/PZ3S-VJ69] ("Poor capacity planning leads to resource shortages and, eventually, exhausted resources."). *But see Break the Internet*, MERRIAM-WEBSTER, https://www.merriam-webster.com/wordplay/break-the-internet (last visited Sept. 11, 2025) [https://perma.cc/SX3Y-CSLA] (One "meaning has to do with overwhelming the Internet system by having too many excited users direct too much traffic to a particular website").

technology concerns, such as the impact of data-intensive technologies[194] on the availability of future financial opportunities.[195]

But, returning to potentially deadly situations, exploit machina availability threats are already visible in various safety contexts, such as those involving aspects of critical infrastructure.[196] Consider the

[194] *See* Mark A. Lemley & Bryan Casey, *Fair Learning*, 99 TEX. L. REV. 743, 748 (2021) (Professors Mark Lemley and Bryan Casey encourage attention to training data quality in machine learning systems, regardless of underlying copyright restrictions); Amanda Levendowski, *How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem*, 93 WASH. L. REV. 579, 579 (2018) (Professor Amanda Levandowski explores the intersection of artificial intelligence bias and copyright).

[195] *See* Citron & Pasquale, *supra* note 181, at 533-34 ("While scorers often characterize their work as an oasis of opportunity for the hardworking, the following are examples of ranking systems that are used to individuals' detriment."); *see also, e.g.*, Frank A. Pasquale, *Reforming the Law of Reputation*, 47 LOY. U. CHI. L.J. 515, 534 (2015) (arguing that "new threats to reputation have seriously undermined the efficacy of health privacy law, credit reporting, and expungement."); Larry Catá Backer, *Next Generation Law: Data-Driven Governance and Accountability-Based Regulatory Systems in the West, and Social Credit Regimes in China*, 28 S. CAL. INTERDISC. L.J. 123 (2018) (suggesting "that social credit represents the expression of new forms of governance that are possible only through the correct utilization of big data management" and that "[t]he shift in regulatory forms also point to significant shifts in the relationship between law, the state and government."); Kristin Johnson, Frank Pasquale & Jennifer Chapman, *Artificial Intelligence, Machine Learning, and Bias in Finance: Toward Responsible Innovation*, 88 FORDHAM L. REV. 499, 504 (2019) (arguing that "learning algorithms may help regulators and lenders fulfill an altruistic promise of inclusion, compensating for decades of discrimination and exclusion in financial markets" but that "should learning algorithms fail to fulfill this promise, fintech firms may hardwire predatory inclusion, existing inequities, and unconscious biases into financial markets for the next several generations, compounding wealth gaps and undermining the welfare of the most vulnerable communities."); Nizan Geslevich Packin & Yafit Lev-Aretz, *On Social Credit and the Right to Be Unnetworked*, 2016 COLUM. BUS. L. REV. 339, 343 (2016) ("introducing a limited "right to be unnetworked," to accommodate the welcomed aspects of social credit systems while mitigating many of their undesired consequences"); Julia Simon-Kerr, *Credibility in an Age of Algorithms*, 74 RUTGERS U. L. REV. 111, 134 (2021) (using analogies to U.S. financial credit scores and China's experiment with a "social credit" scoring system to argue that "a predictive approach to credibility is structurally distinct from a worthiness-centered one"); Monika Zalnieriute, Lyria Bennett Moses & George Williams, *The Rule of Law "By Design"?*, 95 TUL. L. REV. 1063, 1091-92 (2021) (using social credit as an example to ask "whether technological solutions that embed rule of law values do in fact promote the rule of law").

[196] Consider an incident involving the throttling of emergency communications from a fire department because the account allegedly was coded for a lower tier of service.

software design and governance failures in Boeing's MCAS system that resulted in the deaths of 346 people.[197] As explained in congressional testimony, the planes in question (imprudently) relied on the availability of a single sensor to prevent planes from unexpectedly plunging to the ground; the company allegedly made the design choice with knowledge of its safety risks and failed to adequately protect the public from it,[198] both before and after the first of two planes crashed.[199] Or consider the alleged (potentially legally problematic) choice of a train builder to build in a technology capability into the control systems of their trains to stop them from functioning if a GPS tracker indicated that the train was stopped for several days at an independent repair shop.[200] But imagine a repurposing of this capability to stop a train while in use, a design choice that could potentially lead to death or bodily injury if the train is remotely disabled while in service.[201] Or consider the recent software update with a "defect"[202] that a cybersecurity

Rob Marvin, *Verizon Throttled Fire Department's Data During California Wildfire*, PCMAG (Aug. 22, 2018), https://www.pcmag.com/news/verizon-throttled-fire-departments-data-during-california-wildfire [https://perma.cc/D3M5-3ZAA].

[197] *See* David Gelles, *Boeing 737 Max: What's Happened After the 2 Deadly Crashes*, N.Y. TIMES, https://www.nytimes.com/interactive/2019/business/boeing-737-crashes.html (last updated Oct. 28, 2019) [https://perma.cc/9U3W-3A2J].

[198] *See* STAFF OF S. COMM. ON COM., SCI. & TRANSP., *supra* note 55; David Schaper, *Congressional Inquiry Faults Boeing and FAA Failures for Deadly 737 Max Plane Crashes*, NPR (Sept. 16, 2020, 5:46 AM), https://www.npr.org/2020/09/16/913426448/congressional-inquiry-faults-boeing-and-faa-failures-for-deadly-737-max-plane-cr [https://perma.cc/57QL-BM9A].

[199] STAFF OF S. COMM. ON COM., SCI. & TRANSP., *supra* note 55.

[200] Ashley Belanger, *Trains Were Designed to Break Down After Third-Party Repairs, Hackers Find*, ARSTECHNICA (Dec. 13, 2023, 2:14 PM), https://arstechnica.com/tech-policy/2023/12/manufacturer-deliberately-bricked-trains-repaired-by-competitors-hackers-find/ [https://perma.cc/A4J9-FY9M].

[201] This type of disabling while in use has happened in connection with creditors and cars. Aimee Picchi, *Why the Repo Man Can Remotely Shut off Your Car Engine*, CBS NEWS (Sept. 25, 2014), https://www.cbsnews.com/news/why-the-repo-man-can-remotely-shut-off-your-car-engine/ [https://perma.cc/H62X-RYXE].

[202] George Kurtz (@George_Kurtz), X (July 19, 2024), https://x.com/George_Kurtz/status/1814235001745027317 [https://perma.cc/KTB3-6Q5Z].

company pushed to hospitals, disrupting care[203] — a disruption that some medical researchers and other experts believe placed the physical safety of patients at risk.[204] These governance choices in the design of products and organizational processes directly impact availability, potentially leading to irreparable harms to members of the public.

In particular, in healthcare scenarios physical safety of bodies is directly contingent on availability of (correctly functioning) sensor-reliant technologies. Yet, the future is rife with potential for exploit machina. The 21st Century Cures Act[205] contemplates that nonmedical functionality may share the hardware initially installed in bodies for medical purposes. These dual use scenarios where medical and nonmedical technologies coexist in a single hardware component may open the door to new categories of selective unavailability scenarios, ones fraught with safety concerns and potential for insider abuse. Service interruption and unavailability scenarios could start to involve not only reasons of alleged customer nonpayment, but also other concerns that could impact both medical and nonmedical capabilities. For example, consider the possibility of exploit machina in a scenario where AI analytics (incorrectly) allege a contractually impermissible or copyright infringing use of an eye implant with recording capability, the "ultimate AR-VR display."[206] Consider a scenario where the patient is playing a video game and encounters content that causes an AI process to disable the implant automatically, a scenario that is perhaps foretold by recent allegedly recurring AI-powered errors leading to service

[203] *See* Jeffrey L. Tully, Sumanth Rao, Isabel Straw, Rodney A. Gabriel, Christopher A. Longhurst, Stefan Savage, Geoffrey M. Voelker & Christian J. Dameff, *Patient Care Technology Disruptions Associated with the CrowdStrike Outage*, JAMA NETWORK, https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2836824?resultClick=3 (last updated Aug. 13, 2025).

[204] Andy Greenberg, *At Least 750 US Hospitals Faced Disruptions During Last Year's CrowdStrike Outage, Study Finds*, WIRED (July 19, 2025, 11:54 AM), https://www.wired.com/story/at-least-750-us-hospitals-faced-disruptions-during-last-years-crowdstrike-outage-study-finds/ [https://perma.cc/R9NZ-YQ88].

[205] 21st Century Cures Act, Pub. L. No. 114-255, § 3060, 130 Stat. 1033 (2016).

[206] *See* Noor Al-Sibai, *Neuralink Cofounder Says New Company's Eye Implant Could Be "Ultimate" VR Tech*, BYTE (Jan. 3, 2023, 4:42 PM), https://futurism.com/the-byte/max-hodak-ultimate-vr-tech [https://perma.cc/32SK-V4AY].

interruptions in other platform contexts.[207] As bodies become increasingly reliant on real time services provided by third parties, those third parties are likely to assess the use of these IoB devices in real time. Or consider the safety risks of selective unavailability that might arise due to forced advertisement displays in smart eyewear, AI-powered contacts or eye implants.[208] Services providers are likely to seek to influence the behaviors of the users. This influence may occur through the threat of selective unavailability of their IoB devices, through changes in user interface quality such as increasing involuntarily pushed advertising and through use of discretionary access limitation and termination rights that they may allege to arise from EULAs and other contract rights (that they may also allege to be unilaterally amendable).[209]

---

[207] *See* Michael Gwilliam, *YouTube AI Mistakes Streamer's Microphone for a Firearm and Shuts Down Broadcast*, DEXERTO (Dec. 15, 2025) https://www.dexerto.com/youtube/youtube-ai-mistakes-streamers-microphone-for-a-firearm-and-shuts-down-broadcast-3294635/ (explaining that a user "had his video restricted after his laugh was marked as "graphic content" [by AI] and it wasn't until he removed his laugh from the footage that his video could be fully monetized again," and another user's livestream was erroneously automatically "taken down for holding a 'firearm' . . . it's a microphone" in "the latest in a series of bizarre moderation incidents plaguing the platform").

[208] Advertising companies such as Meta are pushing the use of IoB eyewear as the primary user computing interface, presumably potentially for purposes of AI and advertising engagement with physical space and user bodies. *See, e.g.*, Stephen Council, *Mark Zuckerburg Says Without AI Glasses, You'll Be at 'Cognitive Disadvantage,'* SFGATE (Aug. 1, 2025), https://www.sfgate.com/tech/article/mark-zuckerberg-wants-ai-glasses-20798250.php [https://perma.cc/Z3AY-7VFD] ("Then [Zuckerberg] said he thinks that in the future, if you don't wear glasses with AI or another AI-infused device, you'll similarly be 'at a pretty significant cognitive disadvantage compared to other people who you're working with, or competing against.'"); Joel LaMontagne, *Next-Generation Advertising: Smart Glasses and AI-Driven Engagement*, FORBES (Mar. 26, 2025, 8:30 AM), https://www.forbes.com/councils/forbestechcouncil/2025/03/26/next-generation-advertising-smart-glasses-and-ai-driven-engagement/ [https://perma.cc/N5JS-N9D6] ("Beyond just content delivery, these wearable devices also capture audio-, visual- and location-based user data that, when empowered by AI, provide advertisers with unprecedented ability to target and personalize advertorial content to users at precisely the right time and place.").

[209] For a discussion of why the allegation that EULAs are unilaterally amendable offends traditional contract law principles, see, for example, Andrea M. Matwyshyn, *Technoconsen(t)sus*, 85 WASH. U. L. REV. 529, 548-56 (2007), arguing that user agreement contracting in digital contexts reflects material differences from traditional form

Returning to the story of Barbara and the deprecated eye implant one final time, you may recall that she was standing at the top of stairs during the hardware failure, placing her at risk of catastrophic injury from a fall. Now imagine that instead of an unexpected hardware failure, a user's vision became limited because of an intentional, dangerously timed decision by a company's AI to push an advertisement for, say, a rideshare company, just as the person began walking down subway steps. Imagine that as he took his first step downward, a wildly blinking neon discount code blocked his field of vision, and he then catastrophically (and foreseeably) tumbled down the stairs to his death.[210] The opportunities for selective unavailability and irreparable harm as part of exploit machina will be plentiful, and harmed parties will look for a response both from regulators and the courts.

Channeling Hannah Arendt, the sections that follow argue in favor of a reevaluation of our legal readiness to address exploit machina and its irreparable harms. Without "thinking what we are doing," the specter of exploit machina threatens to undercut the social and economic values we (sometimes incorrectly) assume that AI and other technologies ostensibly support.

contracting scenarios and often meets both the *Williston* and *Corbin* tests for unconscionability in traditional contract law doctrine; Matwyshyn, *Privacy, the Hacker Way*, *supra* note 157, at 61:

> Regardless of whether the terms of agreement include an assertion that the terms may change, courts should require additional consideration to support any such modification. Particularly in consumer-facing contracts, concerns over unfair surprise and oppression loom large. If one side can modify every provision of the agreement at its sole discretion and the other side cannot fully extract itself from the relationship because of data collection, the entire theory of contract as a bargained-for exchange is subverted.

[210] *See* Shawn Carolan, Amy Wu Martin, C.C. Gong, Sam Borja & Claude Sonnet, *2025: The State of Consumer AI*, MENLO VENTURES (June 26, 2025), https://menlovc.com/perspective/2025-the-state-of-consumer-ai/ [https://perma.cc/HV28-A4XF] ("We expect rapid adoption of advertising models, transaction fees, affiliate revenue, and marketplace models.").

## II. MOVING (TOO) FAST AND BREAKING PEOPLE: IRREPARABLE PREDICTIVE HARMS

> Neo: "I suppose the most obvious question is . . . how can I trust you?"
>
> The Oracle: "Bingo. It is a pickle. No doubt about it. The bad news is there's no way you can really know whether I'm here to help you or not . . . [offers a sweet] Candy?"[211]

> "[T]he confidence in the unlimited power of science is only too often based on a false belief that the scientific method consists of a ready-made technique, or in imitating the form rather than the substance of scientific procedure, as if one needed only to follow some cooking recipes to solve all social problems."
>
> — F.A. Hayek[212]

Over half a century before *The Matrix* depicted a future version of a Chicago-esque city,[213] a great spectacle of new technology was organized in the real Windy City[214] — the 1933 World's Fair.[215] Held under the

---

[211] THE MATRIX RELOADED, *supra* note 1, at 45:17.

[212] Friedrich von Hayek, Prize Lecture: The Pretence of Knowledge (Dec. 11, 1974), https://www.nobelprize.org/prizes/economic-sciences/1974/hayek/lecture/ [https://perma.cc/ER3N-ECP6].

[213] *See* Leor Galil, *Neo: Where Misfits Fit in*, READER (Dec. 9, 2021), https://chicagoreader.com/music/neo-where-misfits-fit-in/ [https://perma.cc/BN8Q-5EJG] (Patrons' default attire at Club Neo sometimes aligned with the aesthetic choices depicted in the film); Legallysociable, *Setting the Matrix in Chicago — Sort of*, LEGALLY SOCIABLE (April 27, 2024), https://legallysociable.com/2024/04/27/setting-the-matrix-in-chicago-sort-of/ [https://perma.cc/P5ZG-39FF] (Indeed, the main character Neo shares a name with a dance club that was popular with fans of German industrial music and harmonically consonant genres in the mid-1990s).

[214] Jonathan Boyd, *Windy City*, ENCYCLOPEDIA OF CHI., http://www.encyclopedia.chicagohistory.org/pages/6.html (last visited Aug. 20, 2025) [https://perma.cc/EP2N-NPDW].

[215] *1933-1934 Century of Progress Exposition*, CHI. ARCHITECTURE CTR., https://www.architecture.org/online-resources/architecture-encyclopedia/1933-1934-century-of-progress-exposition (last visited Nov. 17, 2025) [https://perma.cc/4M2G-YUBG].

moniker of "A Century of Progress,"[216] the Fair's goal was partially to restore faith in science post-WWI,[217] as the United States struggled to fully exit the Great Depression.[218] Among the technology marvels displayed at the World's Fair was a pavilion marked by a sign that announced an opportunity to view "infant incubators with living babies" for a small fee.[219] Meanwhile, in another part of the fairgrounds, a "psychograph machine" allegedly measured character traits and "psychological hangups,"[220] and a corporation sponsored a "Best Baby" contest, where fairgoers voted for their favorite infant among a sea of entrants.[221]

By today's standards, the business model of pay-per-view cyborg premature infants likely strikes us as an exploitative commodification.[222] In contrast, a "psychographic machine" test may seem akin to an Internet quiz,[223] and a Best Baby contest may

---

[216] *Id.*

[217] *See* Northwestern University Medical Alumni Association, *The 1933 Century of Progress: Chicago's Other World's Fair May 2023*, YOUTUBE, at 4:45 (May 24, 2023), https://www.youtube.com/watch?v=0ZaKUEp5_ps [https://perma.cc/6B4M-QU7A].

[218] *See* Daniel Hautzinger, *A Break in the Clouds: Chicago's 1933 World's Fair*, WTTW (May 27, 2017), https://www.wttw.com/playlist/2017/05/26/break-clouds-chicagos-1933-worlds-fair [https://perma.cc/PJN8-G7Z8]. At the time, a quarter of Chicago's workforce was unemployed, and many people relied on Al Capone's soup kitchens for food. Northwestern University Medical Alumni Association, *supra* note 217, at 10:02.

[219] Northwestern University Medical Alumni Association, *supra* note 217, at 0:50.

[220] *Id.* at 12:12.

[221] As explained in the rules of the contest, "Beauty of features, appeal of personality as shown by the picture, and proportionate height and weight, will be the basis of the judging." SEARS ROEBUCK & CO., SEARS IS SEEKING AMERICA'S MOST BEAUTIFUL BABY! (1933). Over 100,000 babies were entered in this "Best Baby" contest, and the winning baby received 24,000 voted and the prize of $5000 and a $5000 R.E. Wood college education policy according to the Sears archives. Other babies who received honorable mentions received smaller dollar amounts and metals and silver cups. *What Was the Sears National Baby Contest?*, SEARS ARCHIVES (2007), https://web.archive.org/web/20070512203700/http://www.searsarchives.com/history/questions/babycontest.htm [https://perma.cc/U7VK-U2UE].

[222] Perhaps we might even be inclined to classify it as an exploit machina problem of confidentiality.

[223] But see, for example, the legal questions arising from the Cambridge Analytica incidents. Sam Meredith, *Facebook-Cambridge Analytica: A Timeline of the Data Hijacking Scandal*, CNBC, https://www.cnbc.com/2018/04/10/facebook-cambridge-analytica-a-

superficially seem like a harmless popularity vote along the lines of Fat Bear Week.[224] Yet, the reality is more nuanced.

In 1933, state of the art incubators such as those at the incubator exhibit were not yet available in most hospitals in the United States. [225] In fact, hospitals were usually not interested in them; instead, hospitals tended to be ambivalent about whether premature babies could or should be saved at all.[226] Rather than seeking to save as many babies as possible, hospitals tended to have a policy of prioritizing "worthy" infants.[227] Children of ethnically Italian, Irish, Slavic, Jewish, or Black

timeline-of-the-data-hijacking-scandal.html (last updated Apr. 10, 2018, 9:51 AM) [https://perma.cc/3AMJ-2VRS] ("Almost 300,000 users were thought to have been paid to take the psychological test — with the app then harvesting their personal data. It also gathered data from their Facebook friends, which reportedly resulted in Kogan having access to the data of millions of Facebook profiles.").

[224] *Fat Bear Week 2024*, NAT'L PARK SERV., https://www.nps.gov/katm/learn/fat-bear-week-2024.htm (last updated Oct. 2, 2024) [https://perma.cc/DG6N-M6Z8].

[225] Despite potentially lacking formal medical training and sponsorship from the traditional medical establishment, "Dr." Couney, the operator of the incubators, had, in fact, invented a superior incubator that built on designs implemented in Paris, which had been modeled on a chicken hatching device. His incubator would go on to revolutionize infant care in the United States and was eventually adopted by all the major children's hospitals. *How One Man Saved a Generation of Premature Babies*, BBC (May 23, 2016), https://www.bbc.com/news/magazine-36321692 [https://perma.cc/KC76-LSYG].

[226] *Life Under Glass*, BBC SOUNDS, at 1:45 (May 24, 2016), https://www.bbc.co.uk/programmes/p03wgbwr [https://perma.cc/P6VG-3NSZ].

[227] "Worthiness" frequently mapped to (certain kinds of) "white" by 1933 standards. Because construction of "whiteness" has changed across time, the children of Italian and Irish immigrants, for example, were often not considered "white." For a discussion of the historical evolution of the social construction of whiteness, see, for example, Adam S. Cohen, *Harvard's Eugenics Era*, HARV. MAG. (Feb. 19, 2016), https://www.harvardmagazine.com/2016/02/harvards-eugenics-era [https://perma.cc/H24M-6JLA]: "[Harvard's president] warn[ed] against mixing races — which for him included Irish Catholics marrying white Anglo-Saxon Protestants, Jews marrying Gentiles, and blacks marrying whites — [which] was a central tenet of eugenics."; Matthew Frye Jacobson, WHITENESS OF A DIFFERENT COLOR (1999), explaining the field of "whiteness studies" and linking it to traditional historical inquiry, and arguing that "ethnic minorities, in becoming American, were re-racialized to become Caucasian."; Karen Brodkin, HOW JEWS BECAME WHITE FOLKS AND WHAT THAT SAYS ABOUT RACE IN AMERICA (1998), explaining that Jews have sometimes been classified as white and at other times have had an off-white racial designation created for them, leading American Jews of different eras to construct their ethnoracial identities differently; and

parents[228] were regularly turned away.[229] But the World's Fair incubator exhibit, in contrast, accepted infants on a needs-blind and race/ethnicity-blind basis,[230] achieved superior life-saving results with state of the art equipment and process,[231] and zealously maintained baby confidentiality.[232]

Meanwhile, the historical context of the psychometric machine and the Best Baby contest also reveals a more complex story — a story potentially uncomfortably bound up with the legacies of eugenics[233] including the data analytics of phrenology[234] and other forms of

Ruha Benjamin, RACE AFTER TECHNOLOGY (2019), arguing that technology tools that appear to be neutral can reinforce particular constructions of race; Robert Wald Sussman, THE MYTH OF RACE (2016), explaining that eugenics thinking was comprised of three parts: intelligence testing, selective breeding, and human sterilization; this thinking fed into Nazi genocide but was countered by Franz Boas's new, scientifically supported concept of culture, exposing its fallacies.

[228] For a discussion of eugenic views on racially "lesser" people according to early-twentieth century standards, see, for example, Clarence Darrow, *The Eugenics Cult*, AM. MERCURY, June 1926, https://dododreams.blogspot.com/2011/09/reprint.html [https://perma.cc/9Y3X-UCSW]:

> [The eugenicist] cries in the night of "race suicide," "the rising tide of color," "the race is dying out at the top," and "torrents of degenerate and defective protoplasm." . . . It is vain to ask the question, What, of it? That does not stop the clamor. Neither will the remarks that I am about to make on the subject.

[229] Stephanie Prescott & Michelle C. Hehman, *Premature Infant Care in the Early 20th Century*, 46 J. OBSTETRIC GYNECOLOGIC & NEONATAL NURSING 637, 641 (2017).

[230] Claire Prentice, *The Man Who Ran a Carnival Attraction That Saved Thousands of Premature Babies Wasn't a Doctor at All*, SMITHSONIAN MAG. (Aug. 19, 2016), https://www.smithsonianmag.com/history/man-who-pretended-be-doctor-ran-worlds-fair-attraction-saved-lives-thousands-premature-babies-180960200/ [https://perma.cc/SW98-ZTM8].

[231] *Id.*

[232] *Id.*

[233] For a discussion of the legal legacy of eugenics, see, for example, *Regulating Eugenics*, 121 HARV. L. REV. 1578 (2008), explaining historical versus modern legal construction of eugenics and "explor[ing] the limits of the state's power to regulate eugenics."

[234] For a discussion of the legal legacy of phrenology, see, for example, Pierre Schlag, *Law and Phrenology*, 110 HARV. L. REV. 877 (1997), exploring "the nineteenth-century pseudo-science of phrenology as a way of gaining insight into the intellectual organization of American law."

problematic "scientism,"[235] borrowing a term from Friedrich Hayek for fake and low-quality science.[236] Indeed, at the time of the Century of Progress, a public debate raged about the desirability of forced sterilization in reliance on (scientistic) predictive analytics. The Supreme Court had recently decided the landmark case of *Buck v. Bell*,[237] validating a state's right to involuntarily sterilize wards of the state it judged to be "defective persons."[238]

When viewed through the modern lens of exploit machina, the case facts of *Buck v. Bell* shock the conscience. It can be understood as a type of insider attack fueled by problematic scientism and a blind eye to irreparable physical harm. As reconstructed by historians, *Buck v. Bell* involved a test case that was purpose-built by a set of Virginia insiders seeking to probe the limits of the Virginia forced sterilization law.[239] The State of Virginia through the Superintendent of the State Colony for Epileptics and Feeble Minded sought to preemptively justify the forced sterilization of Carrie Buck, a seventeen year old whom historians believe to have been raped by a member of her foster family at the age of sixteen.[240] To quote one historian, "the fix was in."[241] Although Buck

[235] As used herein, scientism refers to making assertions that are not adequately supported by trustworthy science. *See, e.g.*, Gregory R. Peterson, *Demarcation and the Scientistic Fallacy*, 38 ZYGON 751, 761 (2003) ("the best way to understand the charge of scientism is as a kind of logical fallacy involving improper usage of science or scientific claims").

[236] von Hayek, *supra* note 212 ("[T]he confidence in the unlimited power of science is only too often based on a false belief that the scientific method consists of a ready-made technique, or in imitating the form rather than the substance of scientific procedure, as if one needed only to follow some cooking recipes to solve all social problems.").

[237] 274 U.S. 200 (1927).

[238] *Id.* at 205.

[239] *See Eugenics in Virginia*, VFH RADIO, at 06:08 (Feb. 2001), https://encyclopediavirginia.org/eugenics-in-virginia-2/ [https://perma.cc/FRS2-SLZT].

[240] American Experience, PBS, *The Eugenics Crusade* YOUTUBE at 1:15:45 (Mar. 5, 2024), https://www.youtube.com/watch?v=vmRb-ov5xfI [https://perma.cc/3EG4-QUZK].

[241] As explained by Professor Paul Lombardo, "the fix was in." Her counsel ignored publicly available evidence that would have bolstered Carrie Buck's defense on both legal and medical grounds. *See Eugenics in Virginia*, *supra* note 239.

herself objected to the Superintendent's unsupported,[242] problematically scientistic[243] claims of her mental infirmity,[244] she lacked effective assistance of counsel[245] to present her Fourteenth Amendment due process and equal protection claims.[246] Justice Holmes wrote the majority opinion in the case; presented most charitably, the opinion lacks a fulsome consideration of irreparable harms to the autonomy interests of a citizen.[247] Instead, it reflects a flawed efficiency calculus driven by misplaced deference to (incorrect)[248] prescriptive predictions,[249] as it engages in speculative public safety analysis to

[242] According to modern historians, no credible evidence by today's standards exists of Buck's suffering from a mental infirmity. Indeed, evidence to the contrary exists regarding both her and her daughter's capacity. J. David Smith, *Carrie Buck (1906-1983)*, ENCYCLOPEDIA VA. (Dec. 7, 2020), https://encyclopediavirginia.org/entries/buck-carrie-1906-1983/ [https://perma.cc/MM67-WHMG].

[243] *Id.* As explained by Professor Paul Lombardo, the legal basis for her institutionalization was an allegation Carrie, a minor, was "promiscuous," like her mother before her. The evidence for this assertion appears to been that teachers had seen Carrie passing notes to boys in class and, therefore, prescriptively classified her as "promiscuous." *Eugenics in Virginia*, *supra* note 239.

[244] Buck v. Bell, 274 U.S. 200, 206 (1927).

[245] As explained by historians, Virginia selected Carrie Buck's attorney because of his sympathies for Virginia's prosterilization position, describing the hearing as a "sham." Paul A. Lombardo, *In the Letters of an 'Imbecile,' the Sham, and Shame, of Eugenics*, UNDARK (Oct. 4, 2017), https://undark.org/2017/10/04/carrie-buck-letters-eugenics/ [https://perma.cc/9KRK-YQX2] (explaining that an obscure marker in Charlottesville, Virginia "recalls Buck's case and declares that she had no 'hereditary defects'" and that "[i]nstead, she was the victim of a sham trial that began her trip to the Supreme Court, and provided justification for [the sterilization of] 60,000 poor or disabled people in 32 states").

[246] *Id.* at 205.

[247] *See id.*

[248] *See* Smith, *supra* note 242.

[249] Through modern eyes, the Court's discussions of Carrie's "best interests" in contravention of her own stated wishes reads as eugenics-infused financial prognostication. Historians of medicine point out that the forced sterilization campaigns were driven by invocations of eugenics as science, disproportionately targeting minorities and those with disabilities in the twenty-first century. *See, e.g.*, Alexandra Minna Stern, *Forced Sterilization Policies in the US Targeted Minorities and Those with Disabilities — And Lasted into the 21st Century*, THE CONVERSATION (Aug. 26, 2020, 08:20 AM), https://theconversation.com/forced-sterilization-policies-in-the-us-targeted-minorities-and-those-with-disabilities-and-lasted-into-the-21st-century-143144 [https://perma.cc/RBM8-G8ZR] ("Eugenicists applied emerging theories of

justify the State's violation of Buck's bodily autonomy. Meanwhile, a second type of case involving forced sterilization, prescriptive analytics, and problematic insider conduct was also playing out in the courts during the post-WWI era. On the other end of the socioeconomic spectrum from Carrie Buck, the family members of wealthy girls sometimes sought to sterilize them, potentially for self-serving financial reasons.[250] Asserting the girls' "feeble-mindedness" and using questionable predictive analytics as justification, family members sometimes allegedly enlisted the help of ethically flexible doctors in sterilization by subterfuge. Thus, they ensured no future heirs would hold claims on the girls' fortunes superior to their own.[251]

As this historical context of the Century of Progress illustrates, the technologies and data analytics controversies of the twentieth century and their connection to insider attacks foreshadowed our modern struggles with exploit machina. While our twenty-first century technology debates about AI, prescriptive analytics, cyborg bodies, and related technologies may strike us as unprecedented and untheorized problems, frequently they are not. They are often merely the modern continuation of last century's epic battle of science versus scientism or, put a different way, the battle of progress versus exploit machina. The sections that follow further elaborate on this battle. They engage with the work of Hannah Arendt on cybernation, (Arendt's interpretation of) Immanuel Kant on imagination, and observations from Benjamin

biology and genetics to human breeding . . . believing American society would be improved by increased breeding of Anglo Saxons and Nordics, whom they assumed had high IQs. Anyone who did not fit this mold of racial perfection, which included most immigrants, Blacks, Indigenous people, poor whites and people with disabilities, became targets of eugenics programs.").

[250] For example, in New York in 1936, Ann Cooper Hewitt, the heir to the Cooper Hewlett engineering fortune sued her mother and two doctors who sterilized her without her consent under the guise of a potentially unnecessary appendectomy. "Peter Cooper Hewitt's will stipulated that two-thirds of his estate was to go to Ann and one-third to his wife, Ann's mother, after his death. The will also stipulated that Ann's share reverted back to her mother if she died childless." Audrey Clare Farley, *Inside the Shocking 1930s Trial of Socialite Ann Cooper Hewitt*, TOWN & COUNTRY MAG. (Apr. 20, 2021, 08:30 AM), https://www.townandcountrymag.com/society/money-and-power/a35597816/ann-cooper-hewitt-unfit-heiress-audrey-clare-farley-excerpt/ [https://perma.cc/GKE4-PETT]; *see* AUDREY CLARE FARLEY, THE UNFIT HEIRESS (2021).

[251] *Id.*

Franklin and developmental psychologists on identity self-narration, and they connect these insights to modern innovation policy and law.

### *A. Investment: Arendt's Cybernation*

Just as today's legal discussions of Internet conduct were preceded by the legal debates in the 1990s over "the law of the horse" among jurists[252] and legal scholars,[253] so too our modern conversations about the impact of automation build on much earlier debates — conversations from the 1960s. In June of 1964, the Institute for Cybercultural Research convened[254] the First Annual Conference on the Cybercultural Revolution.[255] Inspired by Norbert Wiener's version of cybernetics, the conference sought to meld interdisciplinary perspectives of cyberculture into a rigorous discussion of the future of technology policy in the United States.[256] Perhaps surprisingly, one of

[252] *See* Frank H. Easterbrook, *Cyberspace and the Law of the Horse*, 1996 U. CHI. LEGAL F. 207, 207-08 (1996).

[253] *See* Lawrence Lessig, *The Law of the Horse: What Cyberlaw Might Teach*, 113 HARV. L. REV. 501, 502 (1999).

[254] As described in the official conference proceedings, it was a conference "in which a subject of great importance . . . 'cybernetics and automation'" was examined "in an orderly, systematic, and logical way, starting from sound premises from which valid conclusions might be drawn." *See* Alice Mary Hilton, Foreword, 1 PROC. ANN. CONF. ON CYBERCULTURAL REVOLUTION, at x, x (1966), https://archive.org/details/evolvingsocietyp0000conf/page/n17/mode/2up [https://perma.cc/Q8SM-BUQG] [hereinafter Hilton, Foreword].

[255] For a discussion of the conference, see, for example, CAROLINE BASSETT, ANTI-COMPUTING 105 (2021), introducing Arendt's contribution to the conference within the broader social context of 1960s technology policy history, stating "[t]he conference took up three key propositions . . . : that cybernation would come about, that it would change the landscape of work and leisure, and that this would have deep social consequences for the present and future" and that "Arendt's conference paper 'On The Human Condition', accepts cybernation as a fact, sets aside all matters of transition, and is marked by a vision of the world beyond work that is very bleak indeed."

[256] As explained in the opening to the conference:

> Let us define some terms. The word cyberculture, for example, is composed of culture (the way of life of a society parentheses and cybernetics parentheses the science of relationships). Cybernetics was coined by Norbert Weiner in 1947 from the Greek word *kubernetes* from which our word governor is also derived.

the participants[257] in this convening was political theorist[258] Hannah Arendt. In her comments at the conference, Arendt primarily channeled her work in *The Human Condition*, where she explained that "[w]hat I propose . . . is a reconsideration of the human condition from the vantage point of our newest experiences and our most recent fears . . . to think what we are doing."[259] In particular, Arendt cautioned conference attendees to consider technologies' impact on society and its governance, or "the very raison d'etre for the horse,"[260] as they debate the future of automation.

Yet, despite the obvious currency of this branch of Arendt's theory and what has been described by scholars as Arendt's "explicit commitment to reconciling technology with political theory,"[261] Arendt's technology theory has been comparatively little analyzed among Arendt scholars[262] and has gone almost entirely unexplored in

Hilton, Foreword, *supra* note 254, at xi.

[257] Arendt later became a board member of the board of the Institute. *See Hannah Arendt Papers: Correspondence, 1938-1976; Organizations, 1943-1976; Conference on Cybernetics, New York, N.Y., 1964-1966*, LIBR. OF CONG., https://www.loc.gov/resource/mss11056dig.021940/?st=gallery [https://perma.cc/T99D-SU43].

[258] Arendt is often considered a leading political philosopher of the twentieth century, but she did not refer to herself as a philosopher. Kathleen B. Jones, *The Trial of Hannah Arendt*, NAT'L ENDOWMENT FOR THE HUMANS. (Mar./Apr. 2014), https://www.neh.gov/humanities/2014/marchapril/feature/the-trial-hannah-arendt [https://perma.cc/7MXZ-SJJD].

[259] HANNAH ARENDT, THE HUMAN CONDITION 5 (2d ed. 1998) [hereinafter ARENDT, THE HUMAN CONDITION].

[260] Hilton, Foreword, *supra* note 254, at x.

[261] Brian Simbirski, *Cybernetic Muse: Hannah Arendt on Automation, 1951-1958*, 77 J. HIST. IDEAS 589, 590 (2016).

[262] As explained by Professor Caroline Bassett:

> Arendt's presence at the conference also directs attention to her consideration of the specifically technological, which is somewhat neglected in critical assessments of her writing which focus on her thinking on totalitarianism (albeit in its relationship to technocratic rationality); Arendt's work is surprisingly rarely explored for what it says about technology "itself."

*See* BASSETT, *supra* note 255, at 105.

the legal and policy literature.[263] Legal scholarship has previously applied Arendt's work in depth to issues of democratic theory[264] and constitutional inquiry,[265] the Second Amendment,[266] persuasion and the use of examples,[267] individual rights[268] and citizenship,[269] legal categories of associations,[270] terrorism,[271] trial theory and practice,[272] crimes against humanity[273] and complicity,[274] evidence,[275] legal theory of

[263] A lone exception is the work of Prof. Michal Saliternik, which explores government technology deployments and their impact on political participation. Saliternik writes that it "seems safe to assume that Arendt would have been uncomfortable with the widespread use of public sensors that are installed on street poles and which communicate information to policymakers, quite literally, over people's heads, giving them no role whatsoever in the policymaking process." Michal Saliternik, *Big Data and the Right to Political Participation*, 21 U. PA. J. CONST. L. 713, 743 (2019).

[264] *See* Andrew Arato, *Forms of Constitution Making and Theories of Democracy*, 17 CARDOZO L. REV. 191, 195 (1995).

[265] *See* Kim Lane Scheppele, *A Constitution Between Past and Future*, 49 WM. & MARY L. REV. 1377, 1380 (2008).

[266] *See* Navid Khazanei & Max J. Andrucki, *First Amendment Homesickness, Second Amendment Homecoming: Hannah Arendt and 501(c) Militias*, 11 UNBOUND 54, 62 (2017).

[267] *See* Peter Margulies, *The Mother with Poor Judgment and Other Tales of the Unexpected: A Civic Republican View of Difference and Clinical Legal Education*, 88 NW. U. L. REV. 695, 729 (1994).

[268] *See* Melissa Stewart, *"A New Law on Earth" Hannah Arendt and the Vision for A Positive Legal Framework to Guarantee the Right to Have Rights*, 62 VA. J. INT'L L. 115, 124 (2021).

[269] *See* Itamar Mann, *Border Masquerades*, 39 BERKELEY J. INT'L L. 127, 130 (2021).

[270] *See* Steve Berenson, *Politics and Plurality in a Lawyer's Choice of Clients: The Case of* Stropnicky v. Nathanson, 35 SAN DIEGO L. REV. 1, 2 (1998).

[271] *See* Sharon Harzenski, *Terrorism, A History: Stage One*, 12 FLA. ST. U. J. TRANSNAT'L L. & POL'Y 137, 148 (2003).

[272] *See* Robert P. Burns, *Some Philosophical Resources for an Account of Truth Practices in the American Trial*, 26 POLAR 109, 120 (2003).

[273] *See* David Luban, *A Theory of Crimes Against Humanity*, 29 YALE J. INT'L L. 85, 87 (2004).

[274] *See* David Luban, *Complicity and Lesser Evils: A Tale of Two Lawyers*, 34 GEO. J. LEGAL ETHICS 613, 615 (2021).

[275] *See* Shoshana Felman, *A Ghost in the House of Justice: Death and the Language of the Law*, 13 YALE J.L. & HUMANS. 241, 244 (2001).

equality,[276] privacy history,[277] war crimes trial procedure,[278] Kantian judgment,[279] and identity.[280] But, until now, no work in the legal scholarship has applied Arendt's technology theory and her concerns over "cybernation" to issues of innovation policy and technology law.

More specifically, at the Conference on the Cybercultural Revolution[281] Arendt focused her comments on the risks of "cybernation"[282] — the social transformations arising from the evolution of technology, hyperautomation, and (so-called) intelligent machines.[283] These concerns might be grouped into two categories — questions involving the impact of computing and data-intensive technologies on human dignity, on the one hand, and questions involving the impact of these technologies on social cohesion and the democratic process, on the other. [284]

---

[276] *See* Marie A. Failinger, *Equality Versus the Right to Choose Associates: A Critique of Hannah Arendt's View of the Supreme Court's Dilemma*, 49 U. PITT. L. REV. 143, 145 (1987).

[277] *See* Keigo Komamura, *Privacy's Past: The Ancient Concept and Its Implications for the Current Law of Privacy*, 96 WASH. U. L. REV. 1337, 1340 (2019).

[278] *See* Molly Wilder, *When a Prosecutor Should Introduce Irrelevant Evidence: Testimony and Ethics in the Eichmann Trial*, 27 GEO. J. LEGAL ETHICS 935, 936 (2014).

[279] *See* Rodolphe Gasché, *Is a Determinant Judgment Really a Judgment?*, 6 WASH. U. JURIS. REV. 99, 100 (2013).

[280] *See* Peter Margulies, *Inclusive and Exclusive Virtues: Approaches to Identity, Merit, and Responsibility in Recent Legal Thought*, 46 CATH. U. L. REV. 1109, 1147-48 (1997).

[281] Upon close reading, Arendt's concerns regarding the impact of technology on the democratic process are also sprinkled throughout her other texts as a unifying theme. *See, e.g.*, HANNAH ARENDT, ESSAYS IN UNDERSTANDING 120, 272, 283-84, 427 (Penguin Random House 2005) [hereinafter ARENDT, ESSAYS].

[282] Professor Caroline Bassett explains that the "term cybernation was coined by the political scientist Donald N. Michael, in The Silent Conquest (1962), a report for the Centre for the Study of Democratic Institutions" and "was a designation intended 'to refer to both automation and computers', being invented for 'convenience' and to avoid the awkwardness of repetition." BASSETT, *supra* note 255, at 109.

[283] Hilton, Foreword, *supra* note 254, at 213-19; HANNAH ARENDT, THINKING WITHOUT A BANISTER 323 (Jerome Kohn ed., 2021) [hereinafter ARENDT, THINKING WITHOUT A BANISTER].

[284] As explained by Arendt scholar Professor Waseem Yaqoob, Arendt's views prior to the conference noted similar concerns about unreflective technology creation:

> Rather than treating science and technology in terms of unfolding essences, Arendt sought to stress their contingent development as part of a parable about the unpredictability of human action . . . [and risks of] unreflective

1. Data Quality and Quality of Life: Artificial Mathematization and Dignity

Both in conference transcripts and in a summary of her comments entitled *On the Human Condition*, Arendt explained that she "shall try to pose the problem [of cybernation] from the point of view of the average person of the United States of America, not members of any specific class of the population."[285] She continued:

> cybernation *is* a new phenomenon . . . [that] must be distinguished from the industrial revolution of the past. The industrial revolution consisted of replacing muscle power, not brain power. Now machines can take over a certain amount of activity which we have always identified as activities of the *human* mind[286] . . . . But, that does not mean anything about the level at which a human being functions or about his special qualifications as a human being.[287]

Arendt then differentiated between *raw calculation power* of a brain, either mechanical or human, and *the process of thinking*.[288] But, notably, Arendt specifically cautioned against a mechanistic theory of mind that artificially conflates the two by viewing humans as walking hard drives,

---

> Promethean endeavour . . . practices of the free and unfree world in a global process that suggested that spaces for political freedom were shrinking in Western democracies . . . . Science and technology . . . were central components of this narrative.

Waseem Yaqoob, *The Archimedean Point: Science and Technology in the Thought of Hannah Arendt, 1951-1963*, 44 J. EUR. STUDS. 199, 206-18 (2014).

[285] Hannah Arendt, *On the Human Condition*, 1 PROC. ANN. CONF. ON CYBERCULTURAL REVOLUTION 213, 214 (1966).

[286] Arendt continues: "This calls, in my opinion, for a re-evaluation of mental activity. What, we must ask, is intellectual activity as such? This re-evaluation of intellectual activity, in itself, is not so new." *Id.* at 214.

[287] *Id.*

[288] "If today we know that machines can play a reasonably good game of chess, then, I think, human dignity demands that we say that the chess playing kind of intelligence apparently has not the same status as other kinds of intelligence, or as other kinds of thinking." *Id.*

programmable and erasable at will.[289] Specifically, Arendt distinguished between the concepts of "memory" — something both computers and humans have — and "remembrance," a uniquely human phenomenon.[290] She explained that remembrance serves a social and dignitary function distinct from (the raw computing power of) memory.[291] Conflating these two constructs of memory and remembrance, Arendt implies, diminishes human dignity.[292] In other words, it might be said that Arendt argues that through remembrance, humans curate our existence to guide our own development; we infuse personal and social meaning in context.[293] To wit, invoking her own technology experiences of watching the world evolve dramatically in the first half of the twentieth century, she states that humans are "conditioned by the world around them, the world in which, and with which they engage."[294] In particular, she cautions against neglecting to consider the emergent effects of this change on us as humans: "once the environment has really changed we are conditioned, even though we may know very little about the conditions that have moulded us."[295]

[289] "If human memory were nothing but the component that either helps us to function or prevents us from functioning like the erasable memory of a computing machine, it would be a very sad state of affairs." *Id.* at 215.

[290] "We know, of course, that remembrance — as distinguished from the simple technical faculty of memory — will stay with us regardless of the function memory may, or may not, perform." *Id.*

[291] "To lose remembrance, would indeed deprive human life of a whole dimension the dimension of the past." *Id.* "Again, we have to re-evaluate and distinguish thought from the technical function of the brain, as we distinguish remembrance from technical memory." *Id.*

[292] *Id.* at 214.

[293] In this way, Arendt embraces a framing deeply aligned with paradigms of human development and adaptation visible in modern nonlinear or ecological developmental psychology. *See, e.g.*, Asil Ali Özdoğru, *Bronfenbrenner's Ecological Theory*, *in* ENCYCLOPEDIA OF CHILD BEHAVIOR AND DEVELOPMENT 300 (Sam Goldstein & Jack A. Naglieri eds., 2011) (introducing Bronfenbrenner's ecological theory of development); Barbra Teater, *Ecological Systems Theory*, *in* THEORETICAL PERSPECTIVES FOR DIRECT SOCIAL WORK PRACTICE (Kristin W. Mapson, J. Christopher Hall & Peter Lehmann eds., 4th ed. 2021) (introducing Bronfenbrenner's ecological theory of development in modern social work context).

[294] Arendt, *supra* note 285, at 217.

[295] *Id.* at 218.

In other words, Arendt's thought urges us to examine whether the core aims of technologies such as AI and predictive and prescriptive analytics start from a set of flawed assumptions about humans, our abilities, and our predictability.[296] Indeed, Arendt views each human as inherently unique from the moment of birth,[297] and she would likely reject the default position of the possibility (or aspirational "good") of an artificial general intelligence, just as she rejects the fundamental fungibility of all activity of the human mind with activity generated by bits and bytes.[298] In this way,[299] Arendt's thinking clashes with most of

[296] Similarly, business authors have raised concerns over data scientist "fishing expeditions" that are hunts for "interesting nuggets" without thoughtful problem conceptualization. Thomas C. Redman, *Do Your Data Scientists Know the 'Why' Behind Their Work?*, HARV. BUS. REV. (May 16, 2019), https://hbr.org/2019/05/do-your-data-scientists-know-the-why-behind-their-work [https://perma.cc/8TY4-4D2A].

[297] *See* ARENDT, THE HUMAN CONDITION, *supra* note 259, at 9; *see also Robert Farrugia Flores, The Capacity to Begin: Arendt's Concept of 'Natality' (A Humble Tribute on the 40th Anniversary of Her Death)*, 3 THREADS 80, 84 (2015), https://www.um.edu.mt/library/oar/bitstream/123456789/20935/1/The%20Capacity%20to%20Begin-80%20Arendt%27s%20Concept%20of%20%27Natality%27.pdf [https://perma.cc/V9M2-WNNW]; Rebecca Sete Jacobson, Born Again: Natality, Normativity and Narrative in Hannah Arendt's The Human Condition (June 2012) (Ph.D. thesis, University of Hertfordshire), https://uhra.herts.ac.uk/id/eprint/16346/ [https://perma.cc/8GM7-9QY6]; Jeffrey Champlin, *"Poetry or Body Politic": Natality and the Space of Birth in Hannah Arendt's* Thought Diary, *in* ARTIFACTS OF THINKING 143, 144 (Roger Berkowitz & Ian Storey eds., 2017), http://www.jstor.org/stable/j.ctt1hfr0q2.11; Wolfhart Totschnig, *Arendt's Notion of Natality: An Attempt at Clarification*, 66 IDEAS Y VALORES 327 (2017); ROSALYN DIPROSE & EWA PLONOWSKA ZIAREK, ARENDT, NATALITY AND BIOPOLITICS (2018).

[298] The computer scientist who coined the term "AI," John McCarthy, embraced this view. *See* ThinkingAllowedTV, *John McCarthy (1927-2011): Artificial Intelligence (complete) — Thinking Allowed — Jeffrey Mishlove*, YOUTUBE (Nov. 3, 2011), https://www.youtube.com/watch?v=Ozipf13jRr4 [https://perma.cc/TSU7-ZKZH]. In other words, such an approach has dominated the field from its inception.

[299] A connected line of thinking is found in Arendt's concept of "natality" in *The Human Condition*. Arendt argues that even after our original act of natality, being born, "[w]ith word and deed we insert ourselves into the human world, and this insertion is like a second birth, in which we confirm and take upon ourselves the naked fact of our physical appearance" and:

> that something new is started which cannot be expected from whatever may have happened before. This character of startling unexpectedness is inherent in all beginnings . . . . The fact that man is capable of action means that the unexpected can be expected from him, that he is able to perform what is

modern computing theory, which assumes that with enough high-quality data, humans are highly predictable, replicable, and replaceable.[300] In other words, Arendt's thinking stands as a direct counterpoint to the view held by much of the AI and machine learning and neuropsychology research communities today.[301]

### 2. Data Fabric[302] and Social Fabric: Alienation and Democratic Deterioration

As scholars have noted, Arendt was "keenly aware of the political significance of developments in technology."[303] She explains that the central problems of the contemporary world are "the political organization of mass societies" and "the political integration of technical power."[304] Indeed, Arendt describes technology as "the meeting ground of the natural and historical sciences"[305] and "the ground on which the two realms of history and nature have met and

> infinitely improbable. And this again is possible only because each man is unique, so that with each birth something uniquely new comes into the world.

ARENDT, THE HUMAN CONDITION, *supra* note 259, at 176-78. For a discussion of natality, see, for example, Tatjana Tömmel & Maurizio Passerin d'Entreves, *Hannah Arendt*, STAN. ENCYCLOPEDIA OF PHIL. ARCHIVE (Edward N. Zalta & Uri Nodelman eds., 2025), https://plato.stanford.edu/archives/spr2025/entries/arendt/ (last updated Feb. 12, 2024); Champlin, *supra* note 297.

[300] Similar efforts were made in various eras of computing with disastrous consequences. For example, during the 1960s "Books and Records" or "Back Office" Crisis, the New York Stock Exchange incorporated new technologies to expedite speed of trades. Brokerage houses, in error, deemed large portions of their workforce redundant. The technology disaster that resulted required SEC intervention. *See, e.g.*, Andrea M. Matwyshyn, *Corporate Cyborgs and Technology Risks*, 11 MINN. J.L. SCI. & TECH. 573 (2010) [hereinafter Matwyshyn, *Corporate Cyborgs*] (explaining the history of the Books and Record Crisis).

[301] *See, e.g.*, RAY KURZWEIL, THE SINGULARITY IS NEAR (2005) (arguing an optimistic vision regarding a future where the inevitable merger of humanity with machines will transform what it means to be human).

[302] *See* Jonker & Krantz, *supra* note 74.

[303] Beltrán Undurraga, *Historicizing Distinctions: Hannah Arendt on Science and Technology*, 3 ARENDT STUD. 153, 154 (2019).

[304] ARENDT, ESSAYS IN UNDERSTANDING, *supra* note 281, at 427.

[305] HANNAH ARENDT, BETWEEN PAST AND FUTURE 58 (Penguin Classics 2006).

interpenetrated each other in our time."[306] This meeting is not conflict-free.

Arendt worried that automation "opened" a focus on consumption as the "toil" bound to the "motor" of "human life,"[307] and a "true consumers' society"[308] instead of a "deliberative one,"[309] distracting from "thinking what we are doing."[310] For Arendt, prevalence of technology that interferes with meaningful deliberation — the kind of thought essential for democracy — signals an attempt to escape "the human condition." In other words, by definition, these dynamics undermine democratic process for Arendt: as she explains, it becomes "quite conceivable that the modern age . . . may end in the deadliest, most sterile passivity history has ever known."[311] As Arendt scholars have explained, "those technologies of vision and perspective" allow "the human body, the body now unpolitic, [to] be viewed, and manipulated."[312] Thus, Arendt proposed the "reconsideration of the human condition from the vantage point of our newest experiences and our most recent fears."[313]

At the time Arendt articulated her views on cybernation in 1964, these "most recent fears" involved Sputnik and the Space Race.[314] Today, those fears involve questions of AI, machine autonomy, and superintelligence.[315] Modern technology industry insiders publicly push

[306] *Id.* at 61.

[307] ARENDT, THE HUMAN CONDITION, *supra* note 259, at 131.

[308] *Id.* at 133.

[309] *Id.*

[310] *Id.* at 1.

[311] *Id.* at 322.

[312] Professor Bassett continues: "Arendt's final — and startling — warning in *The Human Condition* is thus that some possibility for action remains, but is often the prerogative of the scientist." BASSETT, *supra* note 255, at 127 (2022), https://www.manchesteropenhive.com/display/9781526160720/9781526160720.00010.xml [https://perma.cc/ZZ2P-7PD6].

[313] ARENDT, THE HUMAN CONDITION, *supra* note 259, at 5.

[314] *Id.*

[315] Professors Divya Jyoti and Bogdan Costea argue that "[i]f Arendt were around today (she died in 1975), she would perhaps argue that TikTok stardom is pointing to what she describes in her book as the 'most sterile passivity history has ever known,' lived by 'thoughtless creatures at the mercy of every gadget which is technically possible.'" Divya Jyoti & Bogdan Costea, *Books That Shook the Business World: The Human*

adoption and use of AI to replace human labor and judgment in another supposed "race," promising transformational social benefit.[316] Yet, the same insiders who boost AI in one context sometimes express existential fear of it in other contexts,[317] an intellectual inconsistency that perhaps signals (an attempt at) exploit machina.[318] Indeed, Arendt's warnings of cybernation hold renewed relevance for us today.

Framed another way, Arendt cautions about a new form of political alienation driven by technology that leads to a politics unmoored from common sense and real world experience.[319] This is the ultimate danger of cybernation for Arendt: a new political reality where people disengage from and abandon the shared work of society — democratic governance.[320] Specifically, she linked two kinds of alienation to

*Condition by Hannah Arendt*, CONVERSATION (Aug. 5, 2024, 8:04 AM), https://theconversation.com/books-that-shook-the-business-world-the-human-condition-by-hannah-arendt-235231 [https://perma.cc/7S8F-H9CC].

[316] Gary Marcus, *China Just Redefined the Global AI Race — With Massive Implications for OpenAI, Nvidia, and Foreign Policy*, FORTUNE (Jan. 27, 2025, 11:32 AM), https://fortune.com/2025/01/27/ai-stargate-china-deepseek-openai-nvidia/ [https://perma.cc/4YLD-QRT4]; Peter Thiel, *The Education of a Libertarian*, CATO UNBOUND (Apr. 13, 2009), https://www.cato-unbound.org/2009/04/13/peter-thiel/education-libertarian/ [https://perma.cc/7N92-82LY] ("A better metaphor is that we are in a deadly race between politics and technology.").

[317] *See* Peter Kasperowicz, *OpenAI CEO Sam Altman Admits His Biggest Fear for AI: 'It Can Go Quite Wrong,'* FOX NEWS (May 16, 2023, 11:36 AM), https://www.foxnews.com/politics/openai-ceo-sam-altman-admits-biggest-fear-ai-can-go-quite-wrong [https://perma.cc/J6G2-TTJY].

[318] Meanwhile, the track record of outcomes for AI implementations to date already tells a story that urges caution. *See supra* text accompanying notes 553–557.

[319] Offering their own definition of cybernation, one set of Arendt scholars channel some of these concerns through a class lens by describing cybernation as "a word derived from the more common 'cybernetics,' and refers to the separation of those working with computers from the rest of the working class." ARENDT, THINKING WITHOUT A BANISTER, *supra* note 283, at 322.

[320] Arendt scholar Nancy Fraser has framed these concerns in Arendt's work as "the intrusion into politics of a totalizing and fundamentally anti-political way of seeing." Nancy Fraser, *Hannah Arendt in the 21st Century*, 3 CONTEMP. POL. THEORY 253, 254 (2004). For further discussion of Arendt and democracy, see, for example, Patchen Markell, *The Rule of the People: Arendt,* Archê*, and Democracy*, 100 AM. POL. SCI. REV. 1 (2006). Arendt also worries about "a threat to human history itself" as "the loss of a distinction between public and private spheres would mean a loss of the conditions for our acting in concert with others," undercutting our ability to act as "concerned agents

advances in technology: a "twofold flight from the Earth into the universe, and from the world into the self."[321] Arendt scholars have described this flight as "Earth alienation"[322] or "Earth abandonment,"[323] connecting it to questions of what scholars of technology theory might call the Anthropocene.[324] Both of these instincts of abandonment of the

that can create ourselves and our history." Terry Winant, *The Feminist Standpoint: A Matter of Language*, 2 HYPATIA 123, 136 (1987).

[321] ARENDT, THE HUMAN CONDITION, *supra* note 259, at 6.

[322] As explained by Oliver Belcher and Jeremy J Schmidt:

> For Arendt, both senses of "being earthbound" arose as humans began to act into nature, not merely upon it. The first sense is oriented to a political ontology of process, which arose as human actions — political, technological, scientific — nullified modernist conceits separating humans from nature . . . . The first sense of "being earthbound" is marked by an ontological shift away from phenomenological accounts referenced to "being" — in which nature and things are treated as an ontic realm separate from action — to a political ontology of process. . . . For Arendt, understanding the political ontology of process requires tracking the transformations that enable action *into* nature across collisions of capital, science and technology. The second sense of "being earthbound" is referenced to scientific praxis and, more specifically, to what Arendt identifies as "earth alienation."

Oliver Belcher & Jeremy J. Schmidt, *Being Earthbound: Arendt, Process and Alienation in the Anthropocene*, 39 ENV'T & PLAN. D 103, 105.

[323] Historian Benjamin Aldes Wurgaft explains that Arendt's concern was that Sputnik set a precedent for the abandonment of earth. She then generalizes from this point, saying that emerging technologies in the twentieth century were often outpaced by expectation, making their eventual arrivals feel belated, not triumphant. *Foreign Object*, MEDIUM (Mar. 15, 2019), https://medium.com/quote-of-the-week/foreign-object-126d6d7d2f2 [https://perma.cc/EKY6-VTDY].

[324] As explained by Belcher and Schmidt:

> Arendt . . . holds that in the modern world *novelty* also includes the acts of science and technology that introduce (novel) processes into earthly affairs. In this way, Arendt treats science *as action*, as *constitutive* of the political. The corollary requirement is accepting responsibility for the form of earth alienation that attends to the scientific disclosure of the earth, and stays with *that* trouble. Further, Arendt recognizes that matter matters politically as the "things" that anchor the common world are both stabilizing and destabilizing forces on it.

Belcher & Schmidt, *supra* note 322, at 115.

Earth and disengagement from societal participation harbor deleterious consequences for democratic process.[325]

Today, Arendt's insights guide us to recognize the increasing social distance between the most influential creators of technology and the daily needs of our society's people.[326] In particular, Arendt's cybernation lens highlights technology insiders' personal divestment from society:[327] indeed, a number of today's technology business leaders have embraced physical self-isolationist[328] projects.[329] Some have constructed survivalist compounds.[330] Some have advocated "seasteading"

[325] For additional discussion of technology and Arendt, see, for example, Bronislaw Szerszynski, *Technology, Performance and Life Itself: Hannah Arendt and the Fate of Nature*, 51 SOCIO. REV. 2 (2003); Melis Bas, A Reinterpretation of Hannah Arendt as a Philosopher of Technology (Nov. 2013) (M.A. Thesis, University of Twente), https://essay.utwente.nl/64574/1/Bas%CC%A7%2C%20Melis%20-%20S1232150%20-%20Master%20Thesis.pdf [https://perma.cc/BX5J-ZQLK].

[326] *See* John Koetsier, *In AI We Do Not Trust: Survey*, FORBES, https://www.forbes.com/sites/johnkoetsier/2023/06/05/in-ai-we-do-not-trust-survey/ (last updated June 9, 2023) [ https://perma.cc/SY2P-G7BS]; Sigal Samuel, *AI That's Smarter than Humans? Americans Say a Firm "No Thank You,"* VOX (Sept. 19, 2023), https://www.vox.com/future-perfect/2023/9/19/23879648/americans-artificial-general-intelligence-ai-policy-poll [https://perma.cc/DMC5-UQL3].

[327] For a discussion of anomie, social divestment and Arendt, see, for example, Elizabeth Brient, *Hans Blumenberg and Hannah Arendt on the "Unwordly Wordliness" of the Modern Age*, 61 J. HIST. IDEAS 513 (2000).

[328] For example, the billionaires of past generations valued building legacies of good works such as establishing libraries. Kathleen Davis, *Tycoon Medievalism, Corporate Philanthropy, and American Pedagogy*, 22 AM. LITERARY HIST. 781, 781 (2010); Sharon Irish, *Whither Tycoon Medievalism? A Response to Kathleen Davis*, 22 AM. LITERARY HIST. 801, 801 (2010).

[329] Interviews might be read to potentially signal their experiencing feelings of anomie. For example, as interviewer recently noted:

> Disillusionment was a recurring theme in my conversations with [Peter] Thiel. He is worth between $4 billion and $9 billion. He lives with his husband and two children in a glass palace in Bel Air that has nine bedrooms and a 90-foot infinity pool. He is a titan of Silicon Valley and a conservative kingmaker. Yet, he tells the story of his life as a series of disheartening setbacks.

Barton Gellman, *Peter Thiel is Taking a Break from Democracy*, ATLANTIC (Nov. 9, 2023), https://www.theatlantic.com/politics/archive/2023/11/peter-thiel-2024-election-politics-investing-life-views/675946/ [https://perma.cc/L75Z-QXAC].

[330] *See, e.g.*, Guthrie Scrimgeour, *Inside Mark Zuckerberg's Top-Secret Hawaii Compound*, WIRED (Dec. 14, 2023), https://www.wired.com/story/mark-zuckerberg-

communities outside of established society[331] or "freedom cities" unmoored from "constraints" of rule of law and democratic structures.[332] Some have hopes for space colonization, openly embracing (dystopian) science fiction versions of the future.[333] Some technology insiders have displayed their alienation through adopting explicitly antidemocratic positions in public comments[334] or in manifestos "that have become legendary in Silicon Valley."[335] Some are understood to be seeking to replace our current democratic system with governance by

inside-hawaii-compound/ [https://perma.cc/KV5D-PZYA] ("Meta CEO Mark Zuckerberg is building a sprawling, $100 million compound in Hawaii — complete with plans for a huge underground bunker.").

[331] *See, e.g.*, Joe Quirk, *Peter Thiel Speaks for 6 Minutes About Seasteading*, THE SEASTEADING INST. (Sept. 5, 2018), https://www.seasteading.org/31937-2/ [https://perma.cc/DQ76-QYU8].

[332] *See, e.g.*, Mack Degeurin, *Billionaires Dream of Building Utopian Techno-City in Greenland*, POP SCI (Apr. 10, 2025), https://www.popsci.com/technology/billionaire-freedom-city-greenland/ [https://perma.cc/3894-ZS76] ("A handful of wealthy, politically connected Silicon Valley investors are reportedly eyeing Greenland's icy shores as the site for a techno-utopian 'freedom city.'").

[333] *See, e.g.*, https://www.cato-unbound.org/2009/04/13/peter-thiel/education-libertarian/ ("The libertarian future of classic science fiction, à la Heinlein, will not happen before the second half of the 21st century."); Corey S. Powell, *Jeff Bezos Foresees a Trillion People Living in Millions of Space Colonies. Here's What He's Doing to Get the Ball Rolling*, NBC NEWS: MACH (May 15, 2019), https://www.nbcnews.com/mach/science/jeff-bezos-foresees-trillion-people-living-millions-space-colonies-here-ncna1006036 [https://perma.cc/4FWF-8ZNV] ("What he really wants to do, Bezos declared, is find a new home in space for our species."); Nadia Drake, *Elon Musk: A Million Humans Could Live on Mars By the 2060s*, NAT'L GEOGRAPHIC (Sept. 27, 2016), https://www.nationalgeographic.com/science/article/elon-musk-spacex-exploring-mars-planets-space-science [https://perma.cc/A7WK-CU7H] ("Musk thinks it's possible to begin shuttling thousands of people between Earth and our smaller, redder neighbor sometime within the next decade or so."); New China TV, *Jack Ma and Elon Musk Hold Debate in Shanghai*, YOUTUBE, at 11:08 (Aug. 29, 2019), https://www.youtube.com/watch?v=f3lUEnMaiAU [https://perma.cc/AR2C-WW6W] ("There's a certain probability that is irreducible that something may happen to earth despite our best intentions despite everything we try to do the there's a probability at a certain point that some either external force or some internal unforced error causes civilization to be destroyed or sufficiently impaired such that it can no longer extend to another planet.").

[334] Peter Thiel has written, "I no longer believe that freedom and democracy are compatible." Thiel, *supra* note 316.

[335] Barton Gellman, *supra* note 329.

an unelected technology aristocracy, as they seek "to find an escape from politics in all its forms," particularly majoritarian democracy.[336] They sometimes voice admiration[337] for the work of antidemocratic writers.[338] Meanwhile, the work in question sometimes materially misstates, among other things,[339] basic points of law, misunderstanding

[336] *See* Thiel, *supra* note 316 ("[T]he great task for libertarians is to find an escape from politics in all its forms . . . . Because there are no truly free places left in our world, I suspect that the mode for escape must involve some sort of new and hitherto untried process that leads us to some undiscovered country."); Peter Thiel, *Your Suffrage Isn't in Danger. Your Other Rights Are*, CATO UNBOUND (May 1, 2009), https://www.cato-unbound.org/2009/05/01/peter-thiel/suffrage-isnt-danger-other-rights-are/ [https://perma.cc/SG8J-GL5C] ("I have little hope that voting will make things better."); *The Godfather of DOGE*, BUS. INSIDER (May 11, 2018) https://www.businessinsider.com/godfather-of-doge-peter-thiel-elon-musk-government-funding-cuts-2025-2?op=1 [https://perma.cc/M8FS-6NH7] ("Thiel's main reason for opposing government, as Max Chafkin's biography of him makes clear, is that it hinders the freedom of tech titans like him to do what they will...'When technology's unregulated,' Thiel once told a reporter, 'you can change the world without getting approval from other people. At its best, it's not subject to democratic control, and not subject to the majority.'"); Raphaëlle Bacqué, *Peter Thiel, The Libertarian Billionaire Waging War on Government*, LE MONDE (July 22, 2025), https://www.lemonde.fr/en/summer-reads/article/2025/07/22/peter-thiel-the-libertarian-billionaire-waging-war-on-government_6743617_183.html ("Thiel's writings lay out a vision of a society led by a small elite — men, wealthy entrepreneurs, preferably — where technology and individualism are exalted.").

[337] Gellman, *supra* note 329 ("Thiel considers Yarvin an 'interesting and powerful' historian.").

[338] David Marchese, *The Interview: Curtis Yarvin Says Democracy Is Done. Powerful Conservatives Are Listening*, N.Y. TIMES (Jan. 18, 2025), https://www.nytimes.com/2025/01/18/magazine/curtis-yarvin-interview.html [https://perma.cc/5HL3-MXTY] ("Marc Andreessen, the venture capitalist turned informal adviser to President-elect Donald Trump, has approvingly cited Yarvin's anti-democratic thinking.").

[339] Yarvin's positions also appear to be uninformed by the history of genocide, human rights during the twentieth century, and the legal and operational realities of the U.S. government, which differ materially from those of a for-profit enterprise. *See id.* For a history of genocide in the twentieth century see, for example, Simon Payaslian, *Twentieth Century Genocides*, OXFORD BIBLIOGRAPHIES (June 20, 2025), https://www.oxfordbibliographies.com/display/document/obo-9780199743292/obo-9780199743292-0105.xml [https://perma.cc/NM59-NZ7Y]: "Genocides in the twentieth century are estimated to have cost more than forty million lives."

the operation of the law of fraud and the role of fiduciary duties as meaningful constraints on officer and director conduct, in particular.[340]

Relatedly, courts have recently raised concerns that a portion of these social divestment efforts is coming at the expense of minority shareholders and fiduciary duty obligations[341] — a form of exploit machina. For example, in one case, a court held that the compensation package of an AI car company CEO could not be successfully demonstrated as fair[342] under Delaware law,[343] noting the CEO's justification was inadequately compelling to overcome claims by a minority shareholder. The CEO stated that he was "motivated by ambitious goals, the loftiest of which is to save humanity" because he "fears that artificial intelligence could either reduce humanity to 'the equivalent of a house cat' or wipe out the human race entirely"[344] and that he "views space colonization as a means to save humanity from this existential threat."[345] Further, the car company CEO asserted that he needs a $56 billion salary because he "seeks to make life 'multiplanetary' by colonizing Mars"[346] — a justification on its face unrelated to the long-term best interests of a car company (on Earth). Channeling an

[340] In particular, Yarvin's discussions equate a CEO to a dictator, omitting the legal reality that a CEO is usually merely an employee who serves at the pleasure of the board of directors and is subject to shareholder oversight. Marchese, *supra* note 338. A CEO is responsible to and can be fired by a board of directors, a governance body elected by shareholders, as well as potentially facing personal civil and criminal consequences for actions which violate fiduciary duties. *See* Del. Code Ann. tit. 8, §§ 141–146 (2025). For a discussion of fiduciary duties, see, for example, Iman Anabtawi & Lynn Stout, *Fiduciary Duties for Activist Shareholders*, 60 STAN. L. REV. 1255 (2008), reviewing the scope of fiduciary duties and arguing that greater shareholder power should be coupled with greater shareholder responsibility, expanding those of activist minority investors.

[341] On January 30, 2024, a Delaware court ordered the recission of "the largest potential compensation plan ever observed in public markets by multiple orders of magnitude," the compensation package of the CEO of one such company, Tesla CEO Elon Musk. Tornetta v. Musk, 310 A.3d 430, 445 (Del. Ch. 2024).

[342] *Id.* at 446 ("The concept of fairness calls for a holistic analysis that takes into consideration two basic issues: process and price.").

[343] *See id.* at 538 ("Considering this glaring defect in Defendants' give/get argument, it does not support a finding of fair price.").

[344] *Id.* at 452.

[345] *Id.*

[346] *Id.*

implicitly Arendtian approach and an explicit Star Trek reference,[347] the Delaware court grounded this legally problematic argument. Although the decision was ultimately limited on appeal, the court brought discussion back down to earth with the legal realities of fiduciary duties, explaining that the idiosyncratic technology fears and aspirations of space colonization of a CEO lack sufficient connection to the particular car company's corporate goals, long-term corporate health, minority shareholder interests, and a justifiable compensation plan.[348]

Reframed through Arendt's cybernation lens, Martian and other escapist planning signals an attempt at self-isolation and social divestment.[349] These developments might be understood in Arendtian terms as seeking to "transcend the human condition,"[350] as socially and

[347] *Id.* at 445-46.

[348] The court explained:

> The incredible size of the biggest compensation plan ever — an unfathomable sum — seems to have been calibrated to help Musk achieve what he believed would make "a good future for humanity." A good future for humanity is a really good thing. Some might question whether colonizing Mars is the logical next step. But, in all events, that "get" had no relation to Tesla's goals with the compensation plan.

*Id.* at 538. The decision's remedy was subsequently modified in substantial part. *See In re* Tesla, Inc. Derivative Litigation, No. 534, 2024, 2025 WL 3689114 (Del. Dec. 19, 2025) (overruling the appropriateness of contract rescission as a remedy, awarding $1 nominal damages, and attorneys' fees with a four times multiplier).

[349] *See* ARENDT, THINKING WITHOUT A BANISTER, *supra* note 319, at 440. Surprisingly, perhaps, Arendt also directly considered the possibility of travel to Mars. Her thinking on point signals that she would be deeply skeptical of the concept of trading the democratic process of the United States and the groundedness of life on Earth for a privately-owned space colony on Mars. "We will look differently on our genetic experiments in the idea of going to the moon and to Mars. Perhaps we can go to Mars but not much farther, and compared to the immensity of the universe, that is a limitation." *Id.*

[350] As described by one author:

> The billionaires considered using special combination locks on the food supply that only they knew. Or making guards wear disciplinary collars of some kind in return for their survival. . . . Taking their cue from Tesla founder Elon Musk colonizing Mars, Palantir's Peter Thiel reversing the ageing process, or artificial intelligence developers Sam Altman and Ray Kurzweil uploading their minds into supercomputers, they were preparing for a digital future that had less to do with making the world a better place than it did with

emotionally self-destructive[351] projects that eschew the established channels of engagement in democracy, such as philanthropy[352] or running for public office.[353]Arendt's approach instead asks us to think

> transcending the human condition altogether. . . . These people once showered the world with madly optimistic business plans for how technology might benefit human society. Now they've reduced technological progress to a video game that one of them wins by finding the escape hatch.

Douglas Rushkoff, *The Super-Rich 'Preppers' Planning to Save Themselves from the Apocalypse*, GUARDIAN (Sept. 4, 2022), https://www.theguardian.com/news/2022/sep/04/super-rich-prepper-bunkers-apocalypse-survival-richest-rushkoff [https://perma.cc/36H7-X34Y]. Meanwhile, decades of psychology research has demonstrated the strong bidirectional correlation of happiness/wellbeing and social connection. *See, e.g.*, Karynna Okabe-Miyamoto & Sonja Lyubomirsky, *Happiness Shapes and Is Shaped by Social Cognition and Social Connection*, *in* THE OXFORD HANDBOOK OF SOCIAL COGNITION 721 (Donal E. Carlston ed., 2024), https://doi.org/10.1093/oxfordhb/9780197763414.013.26 ("Accumulating evidence demonstrates the vital role that social cognition plays in the bidirectional relationship between happiness and social connection.").

[351] This interpretation also aligns with the findings of a recent Surgeon General advisory on the self-destructive force of the growing sense of loneliness in the United States. In 2023, the U.S. Surgeon General warned of an "epidemic of loneliness and isolation." *See generally* U.S. SURGEON GENERAL, OUR EPIDEMIC OF LONELINESS AND ISOLATION: THE U.S. SURGEON GENERAL'S ADVISORY ON THE HEALING EFFECTS OF SOCIAL CONNECTION AND COMMUNITY (May 2023), https://www.hhs.gov/sites/default/files/surgeon-general-social-connection-advisory.pdf [https://perma.cc/U32X-ZSC7]. This concern similarly resonates with Arendt's thought: In Arendt's view, totalitarianism's creep was assisted by "'the fact that loneliness . . . ha[d] become an everyday experience' for so many. The all-pervasive system of the totalitarian regime promised and, for a time, provided an all-encompassing orientation, meaning, and purpose for the masses that they otherwise lacked and craved in their lives." Damon Linker, *The Politics of Loneliness is Totalitarian*, WEEK (June 30, 2021), https://theweek.com/politics/1002095/the-politics-of-loneliness-is-totalitarian [https://perma.cc/D4UA-32RT] (omission in original).

[352] For example, some modern technology billionaires have questioned whether philanthropy, a traditional channel of social participation, should appropriately be viewed as an unwarranted admission of the need for atonement, rather than as an act of gratitude or as a beneficent act in the name of social improvement and paying good fortune forward. *See* Gellman, *supra* note 329. *But see* Bill Gates, *The Power of Giving: Philanthropy's Impact on American Life*, GATES FOUND. (Dec. 1, 2015), https://www.gatesfoundation.org/ideas/speeches/2015/12/bill-gates-the-power-of-giving-philanthropys-impact-on-american-life [https://perma.cc/3XGW-SH29].

[353] *See, e.g.*, Alex Kantrowitz & Nitasha Tiku, *Mark Zuckerberg Says He's Not Running for President*, CNBC, https://www.cnbc.com/2017/01/24/mark-zuckerberg-says-hes-not-

what we are doing — to consider whether just because we can potentially do something, we should.[354] She would encourage (re)engagement with notions of common sense and rootedness in the democratic fabric of society and social governance. Stated another way, Arendt would caution us to avoid losing the bigger societal thread in the particularities of our discussions of data fabrics[355] and AI. Arendt would urge us to consciously avoid the self-destructive forces of cybernation and to do the hard work of the social embeddedness required for representative democracy.

### *B. Imagination: Kant and KPIs*

In her work on cybernation,[356] Arendt also highlights the importance of imagination and spontaneous thought, both to individual human development and to democratic functions. Channeling and interpreting

running-for-president.html (last updated Jan. 24, 2017) [https://perma.cc/DH9S-F9CX]; Emily Shugerman & Kali Hays, *'Devastating': Chan Zuckerberg Charity Slashes Funding for More Bay Area Nonprofits*, S.F. STANDARD (May 13, 2025), https://sfstandard.com/2025/05/13/nonprofits-furious-chan-zuckerberg-initiative-slashes-funding/. *But see* Elyssa Kaufman, Jackie Kostek, Todd Feurer & Sara Tenenbaum, *Illinois Governor JB Pritzker Announces Reelection Bid for Third Term*, CBS NEWS (June 26, 2025), https://www.cbsnews.com/chicago/news/illinois-governor-pritzker-running-for-third-term/ [https://perma.cc/SH28-M2M6].

354 Arendt critiqued a "just because we can" model of technology creation in the context of the arrival of the technology for splitting the atom, writing that:

> The simple fact that physicists split the atom without any hesitations the very moment they knew how to do it, although they realized full well the enormous destructive potentialities of their operation, demonstrates that the scientist *qua* scientist does not even care about the survival of the human race on earth, or, for that matter, with the survival of the planet itself.

Hannah Arendt, *Man's Conquest of Space*, 32 AM. SCHOLAR 527, 536 (1963).

355 For a discussion of data fabric, see Jonker & Krantz, *supra* note 74.

356 In her other work, she stresses the need for space for spontaneity as a buffer to totalitarianism. As explained by Professor Nancy Fraser, Arendt was conscious of the "radical negation of the quintessentially human capacity for spontaneity" and its contribution toward pushing people toward accepting totalitarianism. Fraser, *supra* note 320, at 253.

Kant, Arendt engages a key distinction between productive and reproductive forms of imagination.[357] Arendt explains:

> Imagination, Kant says, is the faculty of making present what is absent, the faculty of re- presentation: "Imagination is the faculty of representing in intuition an object that is not itself present." . . . Or: "Imagination . . . is the faculty of perception in the absence of an object." To give the name "imagination" to this faculty of having present what is absent is natural enough. If I represent what is absent, I have an image in my mind — an image of something I have seen and now somehow reproduce. In the Critique of Judgment, Kant sometimes calls this faculty "reproductive" — I represent what I have seen — to distinguish it from the "productive" faculty — the artistic faculty that produces something it has never seen.[358]

Arendt interprets Kant to write that "[t]hinking is speaking with oneself . . . consequently it is also listening to oneself inwardly (by means of the reproductive power of imagination),"[359] calling this mediating role the "transcendental function of the imagination."[360] Arendt refers to this analysis as "perhaps the greatest discovery Kant made in the Critique of Pure Reason."[361] She explains that this transcendental function and the ability to extrapolate through productive imagination sit at the core of sensibility and understanding, which feed the individualized process of practiced judgment and

---

[357] Some philosophers have engaged with Kantian notions of imagination, for example ARENDT, THINKING WITHOUT A BANISTER, *supra* note 319, at 235, 387; Samantha Matherne, *Kant's Theory of the Imagination*, *in* THE ROUTLEDGE HANDBOOK OF PHILOSOPHY OF IMAGINATION, 55 (Amy Kind ed., 2016), https://philarchive.org/archive/MATKTO-5. Others have critiqued it. *See* JANE KNELLER, KANT AND THE POWER OF IMAGINATION 95 (2007), https://www.cambridge.org/core/books/abs/kant-and-the-power-of-imagination/failure-of-kants-imagination/EF7503D1752462343E3D947931885O4F.

[358] ARENDT, THINKING WITHOUT A BANISTER, *supra* note 319, at 387.

[359] *Id.* at 235; IMMANUEL KANT, ANTHROPOLOGY FROM A PRAGMATIC POINT OF VIEW 86 (Robert B. Louden ed., 2006) [hereinafter KANT, ANTHROPOLOGY].

[360] IMMANUEL KANT, CRITIQUE OF PURE REASON 240-41 (Paul Guyer & Allen W. Wood eds. & trans., Cambridge Univ. Press 1998).

[361] ARENDT, THINKING WITHOUT A BANISTER, *supra* note 319, at 391.

thorough reason.[362] In other words, for Arendt imagination is necessary for the development of judgment and, consequently, the capacity for scientific inquiry,[363] the arts,[364] and meaningful debate and participation in democracy.[365] But, as Arendt points out,[366] the process starts with a human's own sensory perceptions of their own physical reality, something that is frequently an individualized sensory experience.[367]

Engaging with (Arendt's discussion of) Kant's distinction between reproductive and productive imagination and the role of individual sensory experience, the risk of exploit machina again becomes visible: when (homogeneous) body sensing technologies are used to judge

[362] *Id.* For Kant, imagination is the "fundamental faculty of the human soul that grounds all cognition a priori." Imagination harmonizes the other faculties, connecting understanding and the senses. KANT, ANTHROPOLOGY, *supra* note 359, at 240-41.

[363] ARENDT, THINKING WITHOUT A BANISTER, *supra* note 319, at 391; *see* KANT, ANTHROPOLOGY, *supra* note 359, at 240-41.

[364] As explained by Professor Colin McLear, for Kant, imagination thus plays a central role in empirical cognition by serving as the basis for both memory and the creative arts. *See Kant: Philosophy of Mind*, INTERNET ENCYCLOPEDIA OF PHIL., https://iep.utm.edu/kantmind/#SSH1aii (last visited Sept. 8, 2025) [https://perma.cc/R66Q-VC6U4ALL-AVZZ].

[365] Meaningful participation in democracy assumes as a precursor the ability to engage in critical thinking and free choice. For a discussion on the role of deliberation in democracy, see generally, for example, Jack M. Balkin, *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*, 79 N.Y.U. L. REV. 1 (2004), arguing "that digital technologies alter the social conditions of speech and therefore should change the focus of free speech theory, from a Meiklejohnian or republican concern with protecting democratic process and democratic deliberation, to a larger concern with protecting and promoting a democratic culture"; Martin H. Redish & Abby Marie Mollen, *Understanding Post's and Meiklejohn's Mistakes: The Central Role of Adversary Democracy in the Theory of Free Expression*, 103 NW. U. L. REV. 1303 (2009), explaining that "[d]emocracy could not exist, in any meaningful sense, absent a societal commitment to basic notions of free expression; nor could free expression flourish in a society uncommitted to democracy," that "two of the most prominent democratic theories of the First Amendment — those of Alexander Meiklejohn and Robert Post — are in tension with democracy, properly defined" and "even as Meiklejohn and Post themselves define it," and that "[a]ny democratic theory must encompass two principles . . . self-rule . . . [and] epistemological humility."

[366] *See* ARENDT, THINKING WITHOUT A BANISTER, *supra* note 319, at 391.

[367] Arendt explains that Kant highlights that the formation of sensory inputs differs across humans, as does the process of understanding and judgment that arises thereafter. *Id.*

people's creative process and (uniquely experienced) sensory experiences, people with allegedly nonconforming thoughts and body reactions can become classified as "incorrect" by default. Many AI-powered body-sensing technologies rely on training data, aggregations of body data (potentially from statistically unrepresentative samples[368]), which they functionally set as the "correctness" defaults. These defaults are then used to construct key performance indicators (KPIs), [369] yardsticks against which other bodies will be judged. As such, these technologies and their KPIs can introduce a scientistic patina of false smoothness, despite using (potentially flawed, low quality, and stale) old data to judge new peoples' bodies (and minds). Even if we assume that a particular act of measurement is itself not problematically scientistic in context, the most productively creative people among us, the people who may "think different,"[370] or people whose bodies sense and react to stimuli differently (from the KPIs) are potentially most at risk of being deemed by these technologies to be "wrong" or "deviant."[371] Their KPI baselines can function as measure of bodies' conformity to undisclosed values (determined by the creators and operators of the technology), much like the phrenological

[368] The sample of these other bodies and its curation process is not always disclosed. *See* Jack Hardinges, Elena Simperl & Nigel Shadbolt, *We Must Fix the Lack of Transparency Around the Data Used to Train Foundation Models*, HARV. DATA SCI. REV., Dec. 13, 2023, at 1, https://hdsr.mitpress.mit.edu/pub/xau9dza3/release/2 [https://perma.cc/38QR-YFZJ].

[369] KPI is a commonly used abbreviation for Key Performance Indicator — a term which refers to critical quantifiable (sometimes idiosyncratic) indicators tracking achievement of a particular predefined, desired result. *What Is a Key Performance Indicator (KPI)?*, KPI, https://www.kpi.org/KPI-Basics/ (last visited Sept. 12, 2025) [https://perma.cc/UBX7-W8EE].

[370] Luke Dormehl, *Today in Apple History: 'Here's to the Crazy Ones' Who 'Think Different,'* CULT OF MAC (Sept. 28, 2024, 6:52 AM), https://www.cultofmac.com/news/apple-think-different-ad-campaign [https://perma.cc/U8WU-XZFQ].

[371] Consider the common situation where your phone incorrectly predicts your next word as you type or the autocorrect function "helpfully" replaces a word that you intended with a word that you did not. Or consider the currently common scenario where you dictate a sentence, see it initially appear correctly and then watch in frustration as it is "improved" by AI features into words that you did not say, changing meaning. Those are potential instances of predictive technologies functioning on the assumption that you match training defaults.

measurements of heads and the convenient labels of "feeblemindedness" selectively functioned a century ago.[372]

In other words, some IoB technologies, particularly those reliant on transformer-based AI models, arguably leverage something akin to a machine simulation of a Kantian *reproductive* imagination process: these systems are still only as good as their training data and training process, as curated by their creators.[373] Hence, the baseline is still at best (what

[372] *See supra* text accompanying notes 227–253. Indeed, whole lines of fresh research relying on AI tools make claims that might be perceived as quasi-phrenological, potentially rekindling scientism concerns from the early twentieth century. *See, e.g.*, Kyle Kearns, *Can Your Face Predict Your Salary? Using AI Personality Assessments in Hiring*, KNOWLEDGE AT WHARTON (Nov. 17, 2025) https://knowledge.wharton.upenn.edu/article/can-your-face-predict-your-salary-using-ai-personality-assessments-in-hiring/ [https://perma.cc/5PSP-PE7D] (alleging that "AI can be trained to infer personality traits from a single photo"); Simon Torkington, *Everything You Need to Know About Neuroeconomics*, WORLD ECON. F. (Oct. 13, 2016), https://www.weforum.org/stories/2016/10/everything-you-need-to-know-about-neuroeconomics/ ("We can measure how much you value something by looking inside your brain. You don't even have to tell us anything about it.").

[373] For an explanation of transformer models, see Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser & Illia Polosukhin, *Attention Is All You Need*, 30 ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, NIPS 2017, at 1, https://arxiv.org/pdf/1706.03762.pdf, claiming "that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data"; see also, for example, Katherine Lee, A. Feder Cooper & James Grimmelmann, *Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain*, 72 J. COPYRIGHT SOC'Y 251, 279 (2025), explaining that transformer based models "are particularly good at capturing context in sequential information by modeling how elements in a sequence relate to each other." Professor Emily Bender and researchers Timnit Gebru, Angelina McMillan-Major, and Margaret Mitchell argued in a seminal paper that the theory and utility of large language models, including transformer models, have limits, explaining that:

> [T]he change from n-gram LMs to word vectors distilled from neural LMs to pretrained Transformer LMs is paralleled by an expansion and change in the types of tasks they are useful for . . . . Nonetheless, all of these systems share the property of being LMs . . . , that is, systems trained to predict sequences of words (or characters or sentences). Where they differ is in the size of the training datasets they leverage and the spheres of influence they can possibly affect. By scaling up in these two ways, modern very large LMs incur new kinds of risk.

Arendt would consider) a reproductive process. As a consequence, acts of human *productive* imagination — the more fragile and valuable human ability — when judged against *reproductive* KPIs are potentially more likely to be viewed as "wrong" by default. Similarly, human bodies — carbon forms known to exist with and exhibit a high degree of variation in both their experience of the world and in the world's experience of them — that do not conform to the KPIs may similarly receive a scarlet letter of being "bad."

Consider the use of AI "brain-sensing" headbands in some grade school and high school classrooms, ostensibly to measure whether each student is "paying attention."[374] The practice has gained some traction (and caused backlash[375]) in China,[376] and it is already being tested in a

*See* Emily M. Bender, Timnit Gebru, Angelina McMillan-Major & Shmargaret Shmitchell, *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?*, 2021 CONF. ON FAIRNESS ACCOUNTABILITY & TRANSPARENCY (FACCT) 610, 612, https://archive.org/details/stochastic-parrots-3442188.3445922. The authors:

> provide recommendations including weighing the environmental and financial costs first, investing resources into curating and carefully documenting datasets rather than ingesting everything on the web, carrying out pre-development exercises evaluating how the planned approach fits into research and development goals and supports stakeholder values, and encouraging research directions beyond ever larger language models.

*Id.* at 610.

374 Further, brain data is sensitive, highly personal body data. Yet the providers of these devices would likely allege themselves to be outside the guarantees of the Health Insurance Portability and Accountability Act and Health and Human Services/FDA regulatory reach and the privacy protections of Family Educational Rights and Privacy Act (FERPA). For a discussion of FERPA, see, for example, Lynn M. Daggett & Dixie Snow Huefner, *Recognizing Schools' Legitimate Educational Interests: Rethinking FERPA's Approach to the Confidentiality of Student Discipline and Classroom Records*, 51 AM. U. L. REV. 1 (2001), exploring "the scope of 'education records' covered by FERPA and the implications of the definition for school practices".

375 *See* Lia Savillo, *A Chinese School Made Students Wear Brainwave-Detecting Headgear*, VICE (Nov. 4, 2019, 3:49 AM), https://www.vice.com/en/article/chinese-school-made-students-wear-brainwave-detecting-headgear/ [https://perma.cc/75V5-35KZ].

376 *See* Tracy You, *Chinese Pupils Must Wear 'Mind-Reading' Headbands Which Scan Their Brains and Will Alert Teachers if They Are Not Concentrating in Class*, DAILY MAIL UK (Oct. 31, 2019, 10:02 AM), https://www.dailymail.co.uk/news/article-7634705/Chinese-school-makes-pupils-wear-brain-scanning-headbands-class-ensure-pay-attention.html [https://perma.cc/8RTX-C76C].

number of U.S. classrooms.[377] Brain-sensing headbands have previously gained some recreational popularity in the United States for purposes of, for example, judging whether meditation is "correct"[378] or as sleep aid technologies;[379] however they have also been considered for military use in "targeted neuroplasticity training" to modify brain activity in desired ways.[380] But even if we assume, for the sake of argument, some utility (and scientific support that such a device works as described) in particular adult[381] contexts,[382] the use of these devices to judge KPI conformity of brain activity patterns in children in the context of classrooms is arguably developmentally inappropriate or even harmful for many of those children.[383] Again, even apart from the known hardware limitations and flaws of sensors commonly used in similar

[377] *See* Douglas Perry, *Headband that Detects Brain Activity Gets Tryout in Schools; Goal Is to Improve Student Engagement*, OREGONIAN (Jan. 15, 2019, 1:30 PM), https://www.oregonlive.com/news/g66l-2019/01/7fa2c5265a1556/headband-that-detects-brain-activity-gets-tryout-in-schools-goal-is-to-improve-student-engagement.html [https://perma.cc/N7X8-2RX4].

[378] *See How Do I Know if I'm Meditating Correctly?*, FOCUSCALM (June 26, 2020), https://focuscalm.com/blogs/blog/how-do-i-know-if-im-meditating-correctly [https://perma.cc/CF44-N48B].

[379] *See MUSE S: The Brain Sensing Headband — Overnight Sleep Tracker & Meditation Headset Device — Multi Sensor Monitor with Responsive Sound Feedback Guidance from Brain Wave, Heart, Body & Breath Activity*, AMAZON, https://www.amazon.com/MUSE-Headband-Overnight-Meditation-Responsive/dp/B08P8PHB4R (last visited Sept. 11, 2025) [https://perma.cc/SU5N-RDFS].

[380] *See* Ben Williamson, *Wearable Real-Time Brainwave Training in the Classroom*, CONNECTED LEARNING ALL. (Nov. 20, 2017), https://clalliance.org/blog/wearable-real-time-brainwave-training-classroom/ [https://perma.cc/872S-K4D4]; *TNT: Targeted Neuroplasticity Training*, DARPA, https://www.darpa.mil/research/programs/targeted-neuroplasticity-training (last visited Sept. 11, 2025) [https://perma.cc/JA6G-KMHN].

[381] DARPA, *supra* note 380.

[382] The explicit goal of using these technologies in some contexts is to alter brain and human behavior toward preset KPIs, as judged by the device. *Id.*

[383] Children's developmental needs are recognized as distinct from those of adults, both as a matter of law and a matter of developmental psychology. For a discussion of the legal and developmental psychology aspects of new technologies' impact on children, see, for example, Andrea M. Matwyshyn, *Generation C: Childhood, Code, and Creativity*, 87 NOTRE DAME L. REV. 1979 (2012), arguing that "particularly in digital commercial contexts, a legal paradigm of childhood is needed that simultaneously focuses on childhood privacy and creating a space for creative tinkering leading to entrepreneurship in adulthood."

categories of devices,[384] the goal itself is potentially suspect. In lieu of following a "high expectations" approach developed through decades of learning sciences research,[385] such devices can be viewed as seeking to replace personalized learning with an impersonal KPI-driven judgment on the student's learning by a (potentially flawed) black box technology. A developmental psychologist would raise concerns over potentially penalizing the brightest children for (brain patterns associated with) imagination, critical thinking, and creativity if they conflict with indicators of "attention" (as such term is defined and measured by the KPIs and proxy variables of the headband provider).[386] If the goal of the educational environment is to help children identify and develop their

[384] For a discussion of common limitations of and problems with sensors, see, for example, *Electroencephalogram (EEG)*, JOHNS HOPKINS MED., https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/electroencephalogram-eeg (last visited Sept. 11, 2025) [https://perma.cc/4AM7-ZPF2]; Stefano Canali, Viola Schiaffonati & Andrea Aliverti, *Challenges and Recommendations for Wearable Devices in Digital Health: Data Quality, Interoperability, Health Equity, Fairness*, PLOS DIGIT. HEALTH (Oct. 13, 2022), https://pmc.ncbi.nlm.nih.gov/articles/PMC9931360/ [https://perma.cc/PY2A-5VCK]; Raghda Al-Halawani, Meha Qassem,& Panicos A. Kyriacou, *Monte Carlo Simulation of the Effect of Melanin Concentration on Light-Tissue Interactions in Transmittance and Reflectance Finger Photoplethysmography*, NATURE: SCI. REPS. (Apr. 8, 2024), https://www.nature.com/articles/s41598-024-58435-7 [https://perma.cc/ST6L-XYX9]; JOSH PFEFER, ISAAC LIN, WILLIAM VOGT, JIANTING WANG, SANDY WEININGER, CHRIS SCULLY & ANDREW FALES, FDA, IMPACT OF SKIN PIGMENTATION ON THE PERFORMANCE OF BIOMEDICAL OPTICS DEVICES (2023), https://www.fda.gov/media/169040/download; see also, for example, Kayla Matthews, *How to Avoid The Five Most Common Sensor Malfunctions*, FIERCE ELECS. (Feb. 21, 2019, 12:00 AM), https://www.fierceelectronics.com/components/how-to-avoid-five-most-common-sensor-malfunctions.

[385] *High Expectations Drive Student Success*, THOMAS B. FORDHAM INST. (Mar. 2, 2021), https://fordhaminstitute.org/national/commentary/high-expectations-drive-student-success.

[386] A developmental psychologist would also worry that such tools add stress to children's classroom experiences, causing them to be concerned whether a tool outside the student's control is deeming the student to be "bad" in real time. Variation in brain development and maturation rates, baselines of emotional wellbeing, or physical characteristics of the child may impact (the measurement of) a child's brain activity as compared to their peers. *See, e.g.*, Mariam Arain, Maliha Haque, Lina Johal, Puja Mathur, Wynand Nel, Afsha Rais, Ranbir Sandhu & Sushil Sharma, *Maturation of the Adolescent Brain*, 9 NEUROPSYCHIATRIC DISEASE & TREATMENT 449, 449-61 (2013), https://pmc.ncbi.nlm.nih.gov/articles/PMC3621648/ [https://perma.cc/28EU-BF52].

individual talents and creativity in order to help them succeed in an innovation-driven economy, judging children's classroom "success" through measuring conformity of brain activity is arguably counterproductive. Further, school administrators and teachers might not be attuned to the exploit machina risks with such technologies; they may overtrust potentially untrustworthy tools and adopt artificially lowered expectations for some students because the "computer said" the students had low potential.[387]

Finally, consider the generation of new technologies with perhaps the highest potential for exploit machina abuse and irreparable harm to productive imagination — brain implanted devices. A portion of these IoB brain technologies may create transformational medical benefits for the humans who live with them. But the safety specifics of the technologies and the trustworthiness of the companies governing them will prove dispositive. For example, consider two different philosophies reflected in two different brain implants both currently in human clinical trials.[388] The CEO of one brain implant company has articulated the organizational goal as the development of a multifunction[389] brain

[387] *High Expectations Drive Student Success*, *supra* note 385.

[388] Isobel Asher Hamilton, *The Former President of Elon Musk's Neuralink Has Invested in the Rival Company that Beat It to Human Trials*, BUS. INSIDER (Feb. 7, 2022), https://www.businessinsider.com/elon-musk-ex-neuralink-president-invests-rival-synchron-brain-biotech-2022-2 [https://perma.cc/6YKA-T6PB]. Both companies have received FDA clearance to enter human clinical trials. Rachael Levy, Marisa Taylor & Akriti Sharma, *Elon Musk's Neuralink Wins FDA Approval for Human Study of Brain Implants*, REUTERS (May 26, 2023), https://www.reuters.com/science/elon-musks-neuralink-gets-us-fda-approval-human-clinical-study-brain-implants-2023-05-25/ [https://perma.cc/2WT7-P777]. However, safety concerns and allegations of oversight board conflicts of interest were noted in the press with respect to one of them. Rachael Levy & Marisa Taylor, *U.S. Regulators Rejected Elon Musk's Bid to Test Brain Chips in Humans, Citing Safety Risks*, REUTERS (Mar. 2, 2023), https://www.reuters.com/investigates/special-report/neuralink-musk-fda/ [https://perma.cc/H23Q-CSLK] [hereinafter Levy & Taylor, *Regulators Rejected*]; Rachael Levy & Marisa Taylor, *Insight: At Musk's Brain-Chip Startup, Animal-Testing Panel is Rife with Potential Conflicts*, REUTERS (Mar. 8, 2023), https://www.reuters.com/technology/musks-brain-chip-startup-animal-testing-panel-is-rife-with-potential-conflicts-2023-05-04/ [https://perma.cc/46WX-W4GS] [hereinafter Levy & Taylor, *Insight*].

[389] Multifunction devices are devices whose hardware enables both medical and nonmedical uses. *See Multiple Function Device Products: Policy and Considerations*, FDA (July 29, 2020), https://www.fda.gov/regulatory-information/search-fda-guidance-

device, not merely a medical one: he has described it as being "like a Fitbit in your skull with tiny wires."[390] In various statements[391] that have caught the attention of regulators,[392] he has announced that the company aspires for its inductively charged,[393] upgradeable,[394] implants not only to "solve"[395] schizophrenia[396] and autism,[397] but also to stream

documents/multiple-function-device-products-policy-and-considerations [https://perma.cc/VP97-KA5B].

[390] John Koetsier, *Elon Musk Wants to Put a 'Fitbit in Your Skull' to Summon Your Tesla*, FORBES (Aug. 28, 2020, 11:01 PM), https://www.forbes.com/sites/johnkoetsier/2020/08/28/elon-musk-wants-to-put-a-fitbit-in-your-skull-to-summon-your-tesla [https://perma.cc/8Z7L-6WVU].

[391] The company recently introduced the public to a group of pigs who had survived implantation of brain prosthetics by "robot surgeons" who removed a portion of the animals' skulls during the surgery. *Id.*

[392] *See* Mrinmay Dey, *SEC 'Repoens' Probe into Neuralink, Musk's Lawyer Says*, REUTERS, (Dec. 13, 2024, 7:02 AM), https://www.reuters.com/technology/sec-reopens-probe-into-elon-musks-neuralink-2024-12-13/ [https://perma.cc/YJH8-N5JJ]; *see also* Lora Kolodny, *After Tesla CEO Elon Musk Alleged 'Unrelenting Investigation,' SEC Pushes Back*, CNBC (Feb. 19, 2022, 12:33 AM), https://www.cnbc.com/2022/02/19/after-tesla-ceo-elon-musk-alleged-unrelenting-investigation-sec-pushes-back.html [https://perma.cc/2Z57-FNWB].

[393] *See* Koetsier, *supra* note 390 ("The Link is inductively charged, which means you would attach something similar to an Apple Watch charger to your head every night.").

[394] *See id.* ("[Y]ou will be able to upgrade your Link as technology advances.").

[395] *See* Isobel Asher Hamilton, *Elon Musk Said His AI-Brain-Chips Company Could 'Solve' Autism and Schizophrenia*, BUS. INSIDER (Nov. 14, 2019, 4:03 AM), https://www.businessinsider.com/elon-musk-said-neuralink-could-solve-autism-and-schizophrenia-2019-11 [https://perma.cc/XYW7-T56Q] ("Elon Musk said he thinks his neural-technology company, Neuralink, will be able to 'solve' schizophrenia and autism."). Autism is not considered a curable disease by most medical experts.

[396] Schizophrenia symptoms can potentially be controlled or mitigated, but it is generally regarded as uncurable by physicians. *See* Mayo Clinic Staff, *Schizophrenia — Symptoms and Causes — Diagnosis & Treatment*, MAYO CLINIC (Oct. 16, 2024), https://www.mayoclinic.org/diseases-conditions/schizophrenia/symptoms-causes/syc-20354443 [https://perma.cc/Y8P7-24KM]. *But see* Thomas A. Widiger & Tracie Shea, *Differentiation of Axis I and Axis II Disorders*, 100 J. ABNORMAL PSYCH. 399, 399 (1991).

[397] Autism has no known cure and refers to a spectrum of conditions that relate to "brain development that affects how people see people see others and socialize with them." Mayo Clinic Staff, *Autism Spectrum Disorder — Symptoms and Causes*, MAYO CLINIC (May 22, 2025), https://www.mayoclinic.org/diseases-conditions/autism-spectrum-disorder/symptoms-causes/syc-20352928 [https://perma.cc/8T95-2HXJ]; Mayo Clinic Staff, *Autism Spectrum Disorder — Diagnosis & Treatment*, MAYO CLINIC (May 22, 2025),

games[398] or music directly to the brain on demand,[399] summon your car with your thoughts,[400] change the information in the brain,[401] connect your brain to an app on your phone,[402] and eliminate the "inefficient"[403] need for human speech by 2031 through brain to brain thought

https://www.mayoclinic.org/diseases-conditions/autism-spectrum-disorder/diagnosis-treatment/drc-20352934 [https://perma.cc/7C9N-669G].

[398] A recent demonstration video presented a macaque playing a video game with its mind, which was connected to a phone through an implant. However, Neuralink has confirmed the deaths of eight prior monkeys. In an administrative action filed with the U.S. Department of Agriculture, the Physicians Committee for Responsible Medicine (PCRM), alleged that based on records obtained by the group, fifteen of twenty-three monkeys that received Neuralink brain implants later had to be euthanized and that the animals were subjected to "extreme suffering." Jeremy Kahn & Jonathan Vanian, *Musk Brain-Chip Company Neuralink Admits to Killing 8 Monkeys in Experiments*, FORTUNE (Feb. 15, 2022, 12:24 PM), https://fortune.com/2022/02/15/musk-brain-chip-company-neuralink-admits-to-killing-8-monkeys-in-experiments/ [https://perma.cc/77GK-PT2K]; *see also* Kari Paul, *Elon Musk's Brain Chip Company, Neuralink, Faces Animal Abuse Claims*, GUARDIAN (Feb. 15, 2022, 5:28 PM), https://www.theguardian.com/world/2022/feb/15/elon-musk-neuralink-animal-cruelty-allegations [https://perma.cc/8G2F-NWUG].

[399] Courtney Linder, *For His Next Trick, Elon Musk Will Stream Music Straight to Your Brain*, POPULAR MECHS. (July 23, 2020, 12:02 PM), https://www.popularmechanics.com/technology/design/a33382605/neuralink-elon-musk-stream-music-brain/ [https://perma.cc/58Q9-N5QN].

[400] *See* Koetsier, *supra* note 390 ("Automated surgical tools in a robot about the size of two stacked washing machines would remove a portion of your skull, making room for a Link to be installed. Needles guided by brain-scanning sensors then insert the Link's 5-micron-thin connectors about six millimeters — about .24 inches — into your brain.").

[401] *See id.* ("But the Link isn't just about reading brain signals. Musk also wants to write to the brain, just like you write to computer memory today. 'You can actually have one electrode influence possibly a thousand or ten thousand neurons,' Musk says. 'Although might only have a thousand electrodes implanted, you could be influencing millions of neurons.'").

[402] *Id.*

[403] Benjamin Taub, *Elon Musk Claims Neuralink Could Render Human Language Obsolete in Five to Ten Years*, IFL SCI. (May 12, 2020), https://www.iflscience.com/technology/elon-musk-claims-neuralink-could-render-human-language-obsolete-in-five-to-ten-years/ [https://perma.cc/D86L-6L3E]; JRE Clips, *Elon Musk Reveals New Details About Neuralink, His Brain Implant Technology*, , YOUTUBE, at 15:30 (May 7, 2020), https://www.youtube.com/watch?v=Gqd057uky40&t=12s [https://perma.cc/WJZ3-FLCY].

conveyance.[404] These goals blend medical and nonmedical applications,[405] and a version of the device is currently in clinical trials.[406] Meanwhile, a different company had already begun the first clinical trials[407] in human subjects[408] with a differently designed implant — an endovascular brain-computer interface[409] embedded permanently inside the brain.[410] As explained by the company that created it, the

[404] *See* JRE Clips, *supra* note 403; *see also* Rojoef Manuel, *Neuralink Brain Chip Will End Language in Five to 10 Years, Elon Musk Says*, SCI. TIMES (May 28, 2021, 8:29 AM), https://www.sciencetimes.com/amp/articles/31428/20210528/neuralink-brain-chip-will-end-language-five-10-years-elon.htm [https://perma.cc/5BJV-Q2LQ] ("In a recent interview, Elon Musk stated that the human language could possibly end within five to ten years. The CEO of Neuralink went to talk with Joe Rogan, implying that with the innovation of the brain chip the company is currently developing, humans won't have to speak anymore using traditional languages.").

[405] In particular, the described functionality would likely involve an expanded attack surface. For a discussion of attack surface, see, for example, *Attack Surface*, NAT'L INST. OF STANDARDS & TECH., https://csrc.nist.gov/glossary/term/attack_surface (last visited Sept. 11, 2025) [https://perma.cc/DY6N-XULC]; *What Is an Attack Surface?*, IBM (June 28, 2022), https://www.ibm.com/think/topics/attack-surface [https://perma.cc/7BKQ-RJ6R].

[406] The company received FDA clearance to enter human clinical trials. Levy et al., *supra* note 388. However, safety concerns and allegations of oversight board conflicts of interest were noted in the press. Levy & Taylor, *Regulators Rejected*, *supra* note 388; Levy & Taylor, *Insight*, *supra* note 388.

[407] The trial involves five patients with ALS, a degenerative disease with debilitating symptoms. Cami Rosso, *First Human U.S. Implant: Synchron Brain-Computer Interface*, PSYCH. TODAY (Aug. 18, 2022), https://www.psychologytoday.com/us/blog/the-future-brain/202208/first-human-us-implant-synchron-brain-computer-interface [https://perma.cc/XDT4-XVT3] ("All five patients have ALS (amyotrophic lateral sclerosis or Lou Gehrig's Disease), a progressive and fatal neurodegenerative disease without a cure where motor neurons die, impacting the ability to initiate and control muscle movement. Those with ALS lose the ability to speak, walk, move, grasp objects, swallow, and eventually breathe.").

[408] *See* Hamilton, *supra* note 388.

[409] *See* Kimberly Ha & Tyler Hubin, *Synchron Announces First Human U.S. Brain-Computer Interface Implant*, BUS. WIRE (July 19, 2022, 8:00 AM), https://www.businesswire.com/news/home/20220719005248/en/Synchron-Announces-First-Human-U.S.-Brain-Computer-Interface-Implant [https://perma.cc/2TEP-F4NY].

[410] The device has been described as "a permanently implanted stent-like device . . . inserted through the jugular vein to reach the motor cortex of the brain, where it can pick up neurological signals denoting an individual's intended actions." Andrea Park, *Sci-fi No More: Synchron Implants Mind-Reading Device in First U.S. Patient in Paralysis*

device[411] "detects and wirelessly transmits motor intent using a proprietary digital language to allow severely paralyzed patients to control personal devices with hands-free point-and-click" with the goal[412] of enabling "everyday tasks such as texting, emailing, online shopping and accessing telehealth services, and the ability to live independently."[413] As described above, a portion of these devices may prove medically transformational, but a portion may be primarily nonmedical in use. Meanwhile, these devices in either form may usher in perhaps the most pernicious variant[414] of exploit machina: a form of exploit machina that harms brains, both directly through the devices' ability to change content in the brain, reading and writing to the brain in real time,[415] and indirectly by judging those brains for conformity of

*Trial*, FIERCE BIOTECH (July 19, 2022, 12:30 PM), https://www.fiercebiotech.com/medtech/synchron-implants-brain-computer-interface-first-us-patient-paralysis-trial. The device was granted an "investigational device exemption" by the FDA in 2022, the first such exemption for a brain-computer interface. *See* Ha & Hubin, *supra* note 409; *Investigational Device Exemption (IDE)*, FDA (2022), https://www.fda.gov/medical-devices/premarket-submissions-selecting-and-preparing-correct-submission/investigational-device-exemption-ide (last visited Nov. 20, 2022) [https://perma.cc/NRP5-TB2A].

[411] The research is supported in part by a $10 million grant from the National Institutes of Health's Neural Interfaces Program. Park, s*upra* note 410.

[412] According to the company, "[f]uture applications include the potential to diagnose and treat conditions of the nervous system, including Parkinson's disease, epilepsy, depression, and hypertension" and "results have demonstrated this technology to be safe in four patients out to 12 months in . . . Australia." Ha & Hubin, *supra* note 409.

[413] *Id.*

[414] No amount of high-quality training data can resolve questions of brain security and exploit machina. The technical security risks are potentially particularly concerning when considered in the context of broader deficits of information security in the economy. *See, e.g.*, Kim Jacobson, *Enzoic Research Reveals Massive Surge in Fortune 500 Employee Account Compromises, Highlighting Increasing Cybersecurity Threat*, BUS. WIRE (Feb. 11, 2025, 9:00 AM), https://www.businesswire.com/news/home/20250211897328/en/Enzoic-Research-Reveals-Massive-Surge-in-Fortune-500-Employee-Account-Compromises-Highlighting-Increasing-Cybersecurity-Threat [https://perma.cc/DVJ3-7BF4].

[415] Among the various possible consequence of exploit machina in brain-embedded devices include the diminishment of human dignity through attacks on confidentiality, integrity, and availability of the brain. Thus, creativity and the sense of self might be irreparably harmed when human bodies are judged and tracked in real time, the

brain processes to (often undisclosed) KPIs, pronouncing opinions on those brains to determine their social "fit" for economic and other opportunities. Indeed, recognizing the expanding scope of the vulnerability of the public to body data abuse, Colorado recently expanded its privacy law to include "biological data" and "neural data" within the Colorado Privacy Act.[416]

Arendt's work warns us that a society that values conformity and reproductive imagination at the expense of originality and productive imagination is a society at risk of cybernation. When creativity, invention, and critical thinking become viewed as transgressive and when compliance with a black box's KPIs becomes the metric of success, neither technological progress nor democracy are likely to flourish. Instead, we will undercut the socially negotiated legal baselines that have driven our past technological progress — our legal frameworks of competition, contract, intellectual property, secrecy, and the First Amendment.

1. Atypicality: Competition and Contract

Consider the impoverished competition dynamics currently present in our technology ecosystem. Companies such as the infamous Juicero have received millions of dollars of venture capital (VC) funding only to become cautionary tales of the disconnect between VCs' perception of value and that of the public.[417] Meanwhile, other less data-driven ventures with more traditional revenue models remain out of vogue for

processes of thinking and remembrance become hijacked, and the brain becomes vulnerable to compromise by third parties or insiders. Such exploitation of technological vulnerability could corrupt or alter the necessary reflection required by both an innovation economy and deliberative democracy. For a discussion of possible technical risks of brain-computer interfaces, their impact on Kantian heautonomy and autonomy, and the risks of safety compromise, see Matwyshyn, *Internet of Bodies*, *supra* note 9, at 156. Arendt appears to adopt the Kantian heautonomy-autonomy distinction in her approach. *See* HANNAH ARENDT, *Preface: The Gap Between Past and Future*, *in* BETWEEN PAST AND FUTURE 3 (1961).

[416] H.B. 24-1058, 74th Leg., Reg. Sess. (Colo. 2024), https://leg.colorado.gov/sites/default/files/2024a_1058_signed.pdf [https://perma.cc/J36Z-EPHM].

[417] Claire Reilly, *Juicero Is Still the Greatest Example of Silicon Valley Stupidity*, CNET (Sept. 1, 2018, 5:00 AM), https://www.cnet.com/culture/juicero-is-still-the-greatest-example-of-silicon-valley-stupidity/ [https://perma.cc/BXN2-794N].

funding, misaligned with VC-driven tech hype cycles.[418] The reason for misalignment lies in part with a different set of KPIs, those used by many VCs[419] to guide their investment choices. These VC KPIs often currently push in favor of funding AI-driven, maximally data-intensive business models over others,[420] which, in turn, has a(n undesirable) homogenizing effect on the economy.[421] This investment side of governance by KPI often creates incentives for quick flips and consolidations, acting as a galvanizing governance value that dominates corporate decision making,[422] even when that conduct reaches the point of exploit machina.[423] In other words, the KPIs of investment are currently blending with the KPIs used for judgments about human data into a broader destructive form of governance by KPI[424] with potential

[418] *Navigating the Hype: The Shift in Venture Capital Investment Strategies*, CONFLUENCE VC, https://confluence.vc/navigating-the-hype-the-shift-in-venture-capital-investment-strategies/ (last visited Sept. 12, 2025) [https://perma.cc/9C9D-PK52].

[419] Phil Nadel, *12 KPIs You Must Know Before Pitching Your Startup*, TECHCRUNCH (Feb. 4, 2017, 10:00 AM), https://techcrunch.com/2017/02/04/12-kpis-you-must-know-before-pitching-your-startup/ [https://perma.cc/V5GD-397S].

[420] *Venture Capital Fund Metrics Cheat Sheet*, DILIGENT EQUITY (June 1, 2021), https://www.equityeffect.com/blog/venture-capital-fund-metrics/ [https://perma.cc/9ZTU-5ZTW]; *What KPIs Do Venture Firms Care About Across Stages?*, AUMNI (Aug. 1, 2024), https://www.aumni.fund/blog/kpis-across-stages [https://perma.cc/PP85-N829].

[421] These KPI governance dynamics are experienced by entrepreneurs as anticompetitive, artificially narrowing the market for innovation that does not fit VC KPIs.

[422] The focus in corporate priorities on hitting KPI targets is enforced structurally through representation of VCs and other lead investors on boards of directors.

[423] Particularly for data-intensive enterprises, those KPIs often become more granular, focused on user engagement and data, but generally they do not consider risks of irreparable harm to the public implicit in various business models. In particular, investor KPIs frequently inadequately consider issues of computer security, privacy, and the potential public irreparable harms arising from exploit machina.

[424] *See, e.g.*, *KPIs for Every Data Team: A 2025 Guide!*, ATLAN, https://atlan.com/kpis-for-data-team/ (last updated Dec. 20, 2025) [https://perma.cc/N2ZA-YBNY] (explaining that KPIs are "indicators help organizations align their data strategies with business objectives.").

anticompetitive side effects.[425] These are further warning signs of Arendtian cybernation.

Turning again to our case study of Internet of Bodies devices, potentially anticompetitive lock-in effects become formidable when devices are surgically embedded in bodies. As more bodies become dependent on a handful of "chokepoint"[426] companies and investors (and their KPIs), the heterogeneity of product offerings and business models diminishes.[427] Conversely, as target systems become concentrated, safety risk and security vulnerability may increase, not only from third party attacks on vulnerabilities in the code included in these technologies,[428] but also from exploit machina and insider attacks. Legally, these competition concerns implicate not only tort,[429] regulatory concerns of unfair and deceptive practices[430] and safety,[431]

[425] An often-overlooked aspect of technology competition involves competition across the degree of internet connectivity required for functionality of particular products in the market. For a discussion of competition on the degree of connectivity, see, for example, Matwyshyn, *Internet of Bodies*, *supra* note 9, at 124. Thus, impoverished competition, short term KPI-driven corporate values, and exploit machina can all take hold simultaneously.

[426] For a discussion of technology competition and various forms of chokepoints, see, for example, REBECCA GIBLIN & CORY DOCTOROW, CHOKEPOINT CAPITALISM (2022), arguing that an increasingly concentrated number of companies create economic chokepoints that negatively impact creative production, in particular.

[427] MATWYSHYN, *Internet of Bodies*, *supra* note 9, at 124. Yet the nature of technical problems may increase and become more complex as idiopathological bodies present unexpected technical scenarios for troubleshooting.

[428] For a discussion of ecosystem homogeneity and security, see, for example, *Renowned Experts Debate Information Security Risks of an Operating System Monoculture at USENIX '04*, USENIX (May 27, 2004), https://www.usenix.org/press-release/renowned-experts-debate-information-security-risks-operating-system-monoculture-usenix [https://perma.cc/A4QU-U535]; Mike Bursell & Yuki Kubota, *Is Homogeneity Bad for Security?*, ALICE, EVE & BOB – A SEC. BLOG (Aug. 28, 2018), https://aliceevebob.com/2018/08/28/is-homogeneity-bad-for-security/ [https://perma.cc/44J8-A49M].

[429] *See* Matwyshyn, *Internet of Bodies*, *supra* note 9, at 138-43.

[430] Privacy and data security concerns are central to these unfairness inquiries. *See id.* at 133.

[431] *See id.* at 135.

intellectual property,[432] secured transactions,[433] and bankruptcy law,[434] but also contract law, in particular.[435]

Specifically, another form of potentially anticompetitive legal homogenization is occurring through contracts, including through end user license agreements (EULAs).[436] Some EULAs include terms that implicate potentially irreparable physical safety harms, such as, perhaps, the contractual arrangements that may have governed Barbara's deactivated bionic eyes.[437] To wit, IoB device licensing and leasing contracts will continue to spawn scenarios involving various forms of health-critical machine[438] and body part deactivations[439] and their corollary legal concerns. The interaction of traditional licensing and leasing models are already visible in both medical device and consumer goods contexts today,[440] and legal scholars[441] and jurists[442] have long raised concerns about the creep of merchant abuses in novel forms of

[432] *See id.* at 148.

[433] *See id.* at 153.

[434] *See id.*

[435] Matwyshyn, *Internet of Bodies*, *supra* note 9, at 143.

[436] It is worth considering to what extent governance by KPI metrics may be driving experimentation with legally aggressive licensing and leasing contractual arrangements that push the limits of public policy concerns and contract law. For a discussion of public policy considerations in contract, see, for example, *Contracts — Public Policy*, 9 HARV. L. REV. 285, 285 (1895), explaining historical underpinnings and applications of the public policy doctrine in contract law.

[437] *See supra* Part I.

[438] The first is already visible today in IoB form insurance contexts, such as licensing of CPAP devices. *See, e.g.*, David Lazarus, *Column: When Your Insurer Denies a Valid Claim Because of "Lack of Medical Necessity,"* L.A. TIMES (Jan. 23, 2018, 3:00 AM), https://www.latimes.com/business/lazarus/la-fi-lazarus-healthcare-claim-denials-20180123-story.html; Allen, *supra* note 131.

[439] For an example of such a scenario, *see supra* Part I.

[440] *See The Pros And Cons of Leasing*, DMV.COM, https://www.dmv.com/leasing-cars (last updated Feb. 28, 2020) [https://perma.cc/X5YF-MD8J].

[441] *See, e.g.*, Anne Fleming, *The Rise and Fall of Unconscionability as the "Law of the Poor,"* 102 GEO. L.J. 1383 (2014); James W. Bowers, *Some Economic Insights into Application of Payments Doctrine: Walker-Thomas Revisited*, 89 CHI.-KENT L. REV. 229 (2014).

[442] *See* Williams v. Walker-Thomas Furniture Co., 350 F.2d 445, 448 (D.C. Cir. 1965) ("We cannot condemn too strongly appellee's conduct. It raises serious questions of sharp practice and irresponsible business dealings.").

leasing transactions, as well as in the EULA form contracts that facilitate them.[443]

A second contract scenario where KPIs and IoB lock-in impact competition arises in labor and employment contracting. Historically, contract law principles have reflected a shared baseline of commercial trust between the parties and an expectation of successful performance of duties.[444] Additionally, contracts include an implicit duty of good faith in performance by default.[445] Employment law and agency have similarly presumed defaults of loyalty[446] and successful performance.[447] In other words, contract law is built around a presumption of bargaining parties' trustworthiness, an expectation of successful performance, and meaningful opportunity to exit.[448] Breach and defaults are viewed as aberrations from this baseline of expected trustworthiness and performance.[449] But, when employees are universally monitored by their

[443] *See, e.g.*, Alex M. Johnson, Jr., *Correctly Interpreting Long-Term Leases Pursuant to Modern Contract Law: Toward A Theory of Relational Leases*, 74 VA. L. REV. 751 (1988).

[444] In contract law, the capacity of parties is presumed but for limited circumstances, as is their ability to freely select most contractual terms, providing both attestations to the current state of reality through the representations in a contract and making forward-looking promises about performance through the warranties. For a discussion of minority and capacity, see, for example, Natalie M. Banta Lynner, *Minors and Digital Asset Succession*, 104 IOWA L. REV. 1699, 1704 (2019), arguing that "granting minors the ability to devise digital assets is a logical evolution of minor capacity standards seen in other areas of the law."

[445] U.C.C. § 1-304 (AM. L. INST. & UNIF. L. COMM'N 1977).

[446] *See, e.g.*, Benjamin Aaron, *Employees' Duty of Loyalty: Introduction and Overview*, 20 COMPAR. LAB. L. & POL'Y J. 143 (1999) (discussing "the key legal principles of the common law duty of loyalty, their implications for employees and employers and we provide practical recommendations for all parties in the employment relationship.").

[447] *See, e.g.*, Mark P. Gergen, *The Use of Open Terms in Contract*, 92 COLUM. L. REV. 997 (1992) (arguing that aligned expectations of success mean that "contracts with open terms are attractive despite their defects, because they align individual risk and joint risk at the time of contracting better than do contracts with fixed performance terms").

[448] For a discussion of the disruptive role of technology law and employee exit, see, for example, Andrea M. Matwyshyn, *The Law of the Zebra*, 28 BERKELEY TECH. L.J. 155, 225 (2013) [hereinafter Matwyshyn, *The Law of the Zebra*], discussing how construing contract breach as a predicate for criminal computer intrusion claims in particular disrupts the free movement of labor and traditional contract law.

[449] For a discussion of the duty of good faith in performance, see, for example, Steven J. Burton, *Breach of Contract and the Common Law Duty to Perform in Good Faith*, 94 HARV. L. REV 369, 369 (1980), explaining that the duty to perform a contract in good faith is a

employers using IoB devices, those baselines invert. In a world where employers become unwilling to extend a default of trust and instead insist on (employees' agreement to) real-time judgments through body-sensing devices, the legal default of trust shifts to a default of suspicion and expected failure. Indeed, employers' real-time body monitoring through IoB devices has now become a point of contention[450] in labor union negotiations.[451]

Legal scholars have examined how availability of employment has become directly or implicitly conditioned on contractual "consent"[452] to hypersurveillance and body devices, including through "wellness" programs.[453] They conclude almost unanimously[454] that the aggregate

general principle of contract law; Sabine Tsuruda, *Good Faith in Employment*, 24 THEORETICAL INQUIRES L. 206, 206 (2023), arguing that "[t]he duty of good faith creates a joint authority structure within contractual relationships, vesting co-contractors with equal and joint authority over the meaning, purposes, and, hence, the requirements of their contract."

[450] Real-time technology audit and oversight rights do not necessarily exist by default; unions would argue that such terms exist only when the parties agree pursuant to an arm's length negotiation. For a discussion of audit rights, see, for example, STEVEN W. FELDMAN, GOVERNMENT CONTRACT GUIDEBOOK § 3:15 (4th ed. 2024), discussing default dynamics of audit rights in government contracts with contractors.

[451] Adam Gaffney, *The West Virginia Teachers' Strike is Over. But the Fight for Healthcare Isn't*, GUARDIAN (Mar. 7, 2018, 1:31 PM), https://www.theguardian.com/commentisfree/2018/mar/07/west-virginia-teachers-strike-healthcare [https://perma.cc/2V4T-56RR].

[452] As observed by Professors Ifeoma Ajunwa, Kate Crawford, and Jason Schultz, "surveillance in the workplace has mostly moved away from an authoritarian regime . . . [and] now evinces an ostensibly participatory character" and "rapid technological advancements and diminishing costs now mean employee surveillance occurs both inside and outside the workplace — bleeding into the private lives of employees." Ifeoma Ajunwa, Kate Crawford & Jason Schultz, *Limitless Worker Surveillance*, 105 CALIF. L. REV. 735, 738-39 (2017).

[453] Professor Elizabeth Brown cautions that employees' participation in "voluntary" wellness programs that use employer-issued wearable devices may present a "legal fiction." Elizabeth A. Brown, *Workplace Wellness: Social Injustice*, 20 N.Y.U. J. LEGIS. & PUB. POL'Y 191, 218 (2017).

[454] For an opposing view in favor of employer hypersurveillance, Professor Nita Farahany advocates for employer experimentation with body-sensing devices: referencing "a haptic scarf that MIT Media Lab has developed, which uses brainwave technology in a responsive way to give a person a little buzz literally when their mind starts to wander," she describes it as, "quite exciting and something that I think

effect is counterproductive. Nevertheless, employer prediction of employee and job applicant behavior using predictive analytics has become increasingly common.[455]

History offers us yet another relevant case study on point: abuse of polygraphs during the twentieth century by private sector employers. Polygraph history warns us that even when body sensing technologies are proven demonstrably unreliable and manipulable,[456] their use (and their irreparable harms) can nevertheless become prevalent and socially destructive, requiring legal intervention. Indeed, in the 1980s, Congress and states banned most private sector use of polygraph tests through the passage of the Employee Polygraph Protection Act[457] and state laws on point followed.[458] The history of polygraphs reminds us that misplaced trust in KPI-driven body sensing technologies can prove counterproductive and damaging to broader social and economic interests.[459]

---

companies should be experimenting with." Nicholas Thompson & Nita Farahany, *Ready for Brain Transparency?*, WORLD ECON. F., at 15:52 (Jan. 19, 2023, 4:15—4:45 PM), https://www.weforum.org/meetings/world-economic-forum-annual-meeting-2023/sessions/ready-for-brain-transparency/ [https://perma.cc/W5TU-BXUV]. *But see supra* text accompanying notes 452–454.

[455] Dana Pessach, Gonen Singer, Dan Avrahamia, Hila Chalutz Ben-Gal, Erez Shmueli & Irad Ben-Gal, *Employees Recruitment: A Prescriptive Analytics Approach via Machine Learning and Mathematical Programming*, 134 DECISION SUPPORT SYS. 1, 4 (2020), https://www.sciencedirect.com/science/article/pii/S0167923620300452 [https://perma.cc/EG5U-Z7FR].

[456] *The Truth About Lie Detectors (aka Polygraph Tests)*, AM. PSYCH. ASS'N (Aug. 5, 2004), https://www.apa.org/topics/cognitive-neuroscience/polygraph [https://perma.cc/E64L-JN2M]; Jennifer González, *To Tell the Truth: A Short History of the Polygraph*, LIBR. OF CONG. BLOGS (Nov. 1, 2022), https://blogs.loc.gov/law/2022/11/to-tell-the-truth-a-short-history-of-the-polygraph/ [https://perma.cc/5HLH-RNUK].

[457] J L Cross, *Employee Polygraph Protection Act of 1988: Background and Implications*, 40 LAB. L.J. 663, 663 (1989).

[458] *See, e.g.*, ALASKA STAT. § 23.10.037 (1998); IDAHO CODE § 44-903 (1973); IOWA CODE § 730.4 (2013); MONT. CODE ANN. § 39-2-304 (1997); NEB. REV. STAT. § 81-1932 (1980); TENN. CODE ANN. §§ 62-27-123, -128 (2021); W. VA. CODE §§ 21-5-5a to -5d (2003).

[459] For a history of the polygraph test, see, for example, Shelley Gupta, *A Polygraph Test Wouldn't Know the Truth If It Hit It with A Brick: Perpetuation of the Normalization of Violence Against Women*, 34 WOMEN'S RTS. L. REP. 282, 283 (2013), recounting the history of the polygraph through various hardware iterations and pointing out that "the polygraph test became computerized, and in 1993, PolyScore, an algorithmic program

Let us briefly consider one final category of these employment-related devices — so called workplace "attention monitoring" devices. Like the brain-sensing headbands discussed previously, these devices are known to be subject to limitations in sensor accuracy[460] and potentially in the metrics they select as proxies for attention.[461] But, even assuming the hardware and software on such devices accurately measured "attention" (and any device's punitive notifications do not present any physical safety risk or violate the National Labor Relations Act),[462] the choice of, for example, "attention" is again potentially an incorrect or even destructive metric, both as a matter of employee morale in the workplace[463] and in fostering creative, next generation thinking, ostensibly the kind of thinking core to the development of new ventures

that analyzes the polygraph data to estimate a probability or degree of deception or truthfulness in a subject, was released."

[460] *See* Udaya Seneviratne, *Rational Manipulation of Digital EEG: Pearls and Pitfalls*, 31 J. CLINICAL NEUROPHYSIOLOGY 507, 507 (2014); Danielle Nadin, *EEG Research is Racially Biased, Soso Undergrade Scientists Designed New Electrodes to Fix It*, MASSIVE SCI. (May 5, 2020), https://massivesci.com/articles/racial-bias-eeg-electrodes-research/ [https://perma.cc/UZT3-LMG8].

[461] *See, e.g.*, Pallavi Kaushik, Amir Moye, Marieke van Vugt & Partha Pratim Roy, *Decoding the Cognitive States of Attention and Distraction in a Real-Life Setting Using EEG*, 12 SCI. REPS., no. 20649, Nov. 2022, at 1, 1 (finding that on average (but not always) attention was "associated with increased left frontal alpha, increased left parietal theta, and decreased central delta" activity in a small sample); Zainab Mohamed, Mohamed El Halaby, Tamer Said, Doaa Shawk & Ashraf Badawi, *Characterizing Focused Attention and Working Memory Using EEG*, 18 SENSORS, no. 3743, Nov. 2018, at 1, 1 (explaining that "Electroencephalograms (EEG) have been used to detect the subject's emotional and cognitive states" and that "classification accuracies that were obtained on 86 subjects were 84% and 81% for the focused attention and working memory, respectively," i.e. an error rate of sixteen percent and nineteen percent respectively); Giulia Emma Towey, Tindara Caprì & Rosa Angela Fabio, *Measurement of Attention*, *in* ATTENTION TODAY (Giulia Emma Towey, Tindara Caprì, Rosa Angela Fabio & Alessandro Antonietti eds., 2019) (explaining that new neuropsychological methods for measurement of attention are complex, predicated on various different theoretical models of attention and paradigms of measurement).

[462] 29 U.S.C. §§ 151–169.

[463] Labor unions have fought against the use of body monitoring technologies; issues of body monitoring devices have been a point of contention in labor union negotiations. *See* Gaffney, *supra* note 451.

in a startup- driven economy.[464] Thus, incumbent players' interest in using these devices to potentially root out these future "defectors" among their employees again raises concerns of exploit machina. These kinds of "predictive disloyalty" applications in furtherance of anticompetitive practices also connect to questions of intellectual property, secrecy, and the First Amendment.

### 2. Disloyalty: Intellectual Property, Secrecy, and the First Amendment

Governance by KPI and exploit machina also raise broader questions of "disloyalty" and two types of potentially irreparable harms. The first type refers to traditional, legally recognized violations of express or implied confidentiality obligations. The second type of harms are predictive, and as such they may be harder to quantify but also potentially irreparable. Legally, these disloyalty concerns map to at least three areas of law and policy: intellectual property, statutory and contractual obligations of secrecy, and First Amendment protections.

Intellectual property concerns arise in at least two forms — the first relating to research destruction and corruption and a second relating to operational secrecy and corporate espionage. Consider the situation where an organization, whether private or public, gives an AI services provider extensive access to its data, both organizational and employee data, ostensibly in the name of finding efficiencies. Without robust compartmentalization, confidentiality, and destruction requirements,[465] any provider of these services would potentially have substantial visibility or even administrative access into the operations

[464] *See infra* notes 536–541. The imposition of such devices, particularly on knowledge workers, often reflects short-sighted governance by KPIs and automated micromanagement. At best, it reflects a lack of familiarity with the recurring flaws of sensor technologies and with how humans work as a matter of developmental psychology. At worst, it may reflect a form of exploit machina.

[465] *See* Avram Piltch, *Tested: Microsoft Recall Can Still Capture Credit Cards and Passwords, a Treasure Trove for Crooks*, SECURITY (Aug. 1, 2025, 7:14 PM), https://www.theregister.com/2025/08/01/microsoft_recall_captures_credit_card_info/ [https://perma.cc/4788-SLKF].

of the organizations that rely on their products,[466] intentionally or unintentionally.[467] Easily avoidable research corruption and destruction disasters can happen. Consider the recent incident where a "catastrophic error in judgment" resulted in an AI firm's deletion of a client's entire production database.[468] Or imagine a version of this scenario where an insider knowingly violates security protocols and uses an unauthorized or compromised technology product, which then leads to (entirely predictable) disastrous consequences. This scenario has already been alleged in litigation in a different technology services context.[469] Further, the default inclination on the part of a technology services provider may be to leverage its products' maximized visibility into client operations for its own purposes, potentially both to provide

[466] Eliza Strickland, *Are You Ready to Let an AI Agent Use Your Computer?*, IEEE SPECTRUM (Feb. 13, 2025), https://spectrum.ieee.org/ai-agents-computer-use [https://perma.cc/JWA2-TJGZ].

[467] *See* Piltch, *supra* note 465.

[468] Emily Forlini, *Vibe Coding Fiasco: AI Agent Goes Rogue, Deletes Company's Entire Database*, PC MAG (July 22, 2025), https://www.pcmag.com/news/vibe-coding-fiasco-replite-ai-agent-goes-rogue-deletes-company-database [https://perma.cc/SJZ3-6PHC]; Mark Tyson, *AI Coding Platform Goes Rogue During Code Freeze and Deletes Entire Company Database — Replit CEO Apologizes After AI Engine Says It 'Made a Catastrophic Error in Judgment' and 'Destroyed All Production Data,'* TOM'S HARDWARE (July 21, 2025), https://www.tomshardware.com/tech-industry/artificial-intelligence/ai-coding-platform-goes-rogue-during-code-freeze-and-deletes-entire-company-database-replit-ceo-apologizes-after-ai-engine-says-it-made-a-catastrophic-error-in-judgment-and-destroyed-all-production-data [https://perma.cc/EF9M-ZHJ6].

[469] *See, e.g.*, Complaint at 7-12, Clorox Co. v. Cognizant Worldwide Ltd. (Cal. Super. Ct. July 22, 2025) ("Cognizant's operation of the Service Desk came with a simple, common-sense requirement: never reset anyone's credentials without properly authenticating them first. Clorox made this easy for Cognizant by providing them with straight-forward procedures to follow whenever providing credential recovery or reset assistance. . . . Despite assuring Clorox that it was following these procedures, Cognizant's conduct on August 11, 2023, demonstrated spectacularly that it was failing to do so. Cognizant repeatedly gave a cybercriminal access to Clorox's network by handing them credentials without properly authenticating them"); Alex Scroxton, *Scattered Spider Victim Clorox Sues Helpdesk Provider*, COMPUTERWEEKLY.COM (July 24, 2025, 4:50 PM), https://www.computerweekly.com/news/366627969/Scattered-Spider-victim-Clorox-sues-helpdesk-provider [https://perma.cc/4V69-7BUG] ("In the lawsuit, filed in the California Superior Court, Clorox accused Cognizant of repeatedly giving a cyber criminal access to its network by handing them credentials without authenticating them or otherwise following basic cyber security processes.").

better services to existing customers and to train future AI product offerings to sell to other organizations. Even assuming that an AI provider's access is limited, some AI systems would potentially be able to predict, for example, an entity's internal research and development trajectory or its future acquisition strategy. Thus, it is possible that the efficacy of trade secret and other secrecy protections for research and development and operations (in practice even if not in law) may be impacted by the exploit machina behaviors of services providers.[470] These kinds of extreme visibility scenarios also potentially open the door to enhanced forms of espionage risk.

Consider the exploit machina scenario where an AI provider tracks the combined user behavior of employees in an enterprise, including their bodies' location data. Merging this information with LinkedIn information on job roles and expertise, and perhaps email, calendars, and call logs, the provider would potentially be able to anticipate the workflow and potentially the nonpublic projects of those companies. Particularly in national security contexts or in competitive industries with a history of corporate espionage, such predictive information about organizations' plans and the identities of key employees would be a highly marketable product. Indeed, corporate and national espionage over intellectual property have long and storied histories.[471]

Meanwhile, in a different operational secrecy context, professionals such as lawyers, doctors, and financial services professionals may violate their duties of confidentiality to their clients by choosing to use certain (improperly vetted) technologies with microphones and video feeds,

[470] In corporate settings, the ability to create new protectable intellectual property through research and development and to get new products to market inevitably involves extensive amounts of internal confidential communications. *See, e.g.*, Sharon K. Sandeen & Elizabeth A. Rowe, *Debating Employee Non-Competes and Trade Secrets*, 33 SANTA CLARA HIGH TECH. L.J. 438, 460 (2017) (explaining that "[e]mployees in occupations like engineering and computer science are more likely to sign noncompetition agreements," which "is not particularly surprising or troubling given that these types of employees tend to be involved in research and development, and thus privy to sensitive trade secret information").

[471] For a discussion of corporate espionage, see, for example, Glenna Rodgers & Scott D. Marrs, *Trade Secrets and Corporate Espionage Protecting Your Company's Crown Jewels*, ACC DOCKET, Apr. 2004, at 60, 61: "[P]roviding a primer on trade secret laws, we offer practical advice and tools to prevent corporate espionage . . . ."

including various IoB devices or cars with in-cabin recording. For example, the growing market of IoB real-time transcription products and services now regularly target professionals with duties of confidentiality, yet these tools often make no obvious accommodations in their business models to recognize professional ethics constraints.[472] Yet, as compute power increases, the ability of these types of products to detrimentally leverage the data from these sensitive settings also potentially improves. Consider also the exploit machina risks and confidentiality threat to client communication privacy when a car's in-cabin microphone and infotainment system captures the content of an attorney's phone calls with clients and mines the connected mobile phone for its contacts.[473] Or consider the emerging strategic partnerships between car companies and eye tracking and psychometric profiling companies to do in-cabin monitoring of live video and audio feeds.[474] In addition to potentially predictively profiling the driver and

[472] In the privacy policy, the company reserves the right, in particular to "disclose personal data with our affiliates for legitimate business purposes and the operation of the App" and "disclose [data] to relevant third parties in the event of mergers, acquisitions, sale of assets, or transfer of services to other companies." *Privacy Policy*, PLAUD, https://www.plaud.ai/policies/privacy-policy (last visited Sept. 12, 2025) [https://perma.cc/TAX9-ZJNT]; *see, e.g.*, *Plaud NotePin*, PLAUD, https://www.plaud.ai/products/notepin (last visited Sept. 12, 2025) [https://perma.cc/95B2-M527] ("A notetaker for all professionals. . . . Accurate summaries for healthier patients.").

[473] *See* Paddy Harrington, *Your Car Is Listening to You — And So Are Hackers*, FORRESTER (Dec. 18, 2024), https://www.forrester.com/blogs/your-car-is-listening-to-you-and-so-are-hackers/ [https://perma.cc/7P6A-LG3H].

[474] *See, e.g.*, Hailey Driscoll, *Affectiva's Breakthrough Calibration-Free Eye Tracking Feature Offers a Revolution in Attention Research*, AFFECTIVA, https://www.affectiva.com /news-item/affectivas-breakthrough-calibration-free-eye-tracking-feature-offers-a-revolution-in-attention-research/ (last visited Sept. 12, 2025) [https://perma.cc/59Q7-UX6W] (promotional material announcing product "built from Smart Eye driver monitoring software" focused on "high accuracy" per "ad size" with "Emotion AI," a technology that psychometrically judges emotions); *Affectiva Automotive AI for Driver Monitoring Systems*, AFFECTIVA (2025), https://www.affectiva.com/product /affectiva-automotive-ai-for-driver-monitoring-solutions/ (promotional material announcing that a product that psychometrically judges drivers of cars based on audio and video monitoring of in-cabin activity in real time: "[u]sing cameras and microphones, Affectiva Automotive AI unobtrusively measures, in real time, complex and nuanced emotional and cognitive states from face and voice").

passengers using scientistic KPIs for certain mental health conditions, the cars may share the content of client communications with third parties[475] and merge the information with other databases in ethically impermissible ways.[476] Particularly when the user is bound by ethics constraints but a technology's EULAs allow for maximal data exploitation, these professional ethics concerns carry the Nixonian risks of possible exploit machina through self-pwn described in earlier sections.[477]

Further, IoB technologies worn by professionals and knowledge workers revive some of the forms of exploit machina that resulted in the Congressional and state bans of the use of the lie detector in most employment settings. In the name of preserving intellectual property and preventing key employee mobility, employers will be tempted to push use of prescriptive AI and IoB devices onto employees to hunt for signs of potential future "disloyalty" in ways that legally overreach. Indeed, as a recent World Economic Forum presentation alleged,[478] these kinds of body sensing technologies in employment contexts may give employers a belief in their ability to detect intellectual property "pre-crime" with the assistance of IoB, outsider data analytics, and AI.[479] Yet, even assuming the technology does not have a high false positive rate, an employer's perspective on which acts are "disloyal" may conflict with the legal protections for union activity[480] and the social need for

[475] Geoffrey A. Fowler, *What Does Your Car Know About You? We Hacked a Chevy to Find Out*, WASH. POST (Dec. 17, 2019), https://www.washingtonpost.com/technology/2019/12/17/what-does-your-car-know-about-you-we-hacked-chevy-find-out/ [https://perma.cc/9AFC-25EB].

[476] These in-cabin recording concerns resulted in California law including disclosure requirements on point. *See* S.B. 296, Reg. Sess. (Cal. 2023).

[477] See *supra* Part I.

[478] *See Davos AM23 — Ready for Brain Transparency?*, WORLD ECON. F., https://www.weforum.org/videos/davos-am23-ready-for-brain-transparency-english/ (last visited Jan. 23, 2026).

[479] For example, consider the extent of visibility into our everyday activities if our search histories were to be combined with our location data and email traffic and other communications.

[480] *See, e.g.*, *Your Rights During Union Organizing*, NLRB, https://www.nlrb.gov/about-nlrb/rights-we-protect/the-law/employees/your-rights-during-union-organizing (last visited Sept. 12, 2025) [https://perma.cc/DEE9-8WXE] (explaining workplace rights to union organization).

entrepreneurship, knowledge transfer, and the free movement of labor in an innovation-driven economy.

Finally, these kinds of pre-crime "disloyalty" predictive determinations with IoB technologies would prove particularly pernicious in government contexts. As the Supreme Court explained in *Wooley v. Maynard*, the Constitution embodies a "broader concept of 'individual freedom of mind'" within which freedom of expression resides.[481] But, a world where IoB devices are leveraged by governments and their agents can quickly become a world where accusations of thought crimes and pre-crimes[482] erode freedom. These are circumstances that undercut not only intellectual property law and future entrepreneurship, but also the freedoms of speech and (secret) thought that serve as the engines of debate in a democracy. Legal scholarship has amply elaborated on the relationship between freedom of expression and accusations of disloyalty,[483] freedom of thought,[484] imagination,[485] and the democratic process.[486] In particular, scholars have elaborated on these dynamics in the context of the Red Scares[487]

[481] Wooley v. Maynard, 430 U.S. 705, 714 (1977).

[482] *See generally* MINORITY REPORT, *supra* note 15.

[483] *See, e.g.*, John Lord O'Brian, *Loyalty Tests and Guilt by Association*, 61 HARV. L. REV. 592, 597 (1948) (posing the question of "whether the measures which have been adopted, and others which are now being proposed, prescribing standards of loyalty, are either necessary or proper methods with which to carry on what must essentially be a battle of ideas").

[484] *See, e.g.*, Marc Jonathan Blitz, *Freedom of Thought for the Extended Mind: Cognitive Enhancement and the Constitution*, 2010 WIS. L. REV. 1049 (arguing that "the power to reshape our thinking processes biologically should be recognized as merely one form of a more general power that our "freedom of mind" is intended to place firmly in our own hands, not in the hands of government officials.").

[485] *See, e.g.*, Jed Rubenfeld, *The Freedom of Imagination: Copyright's Constitutionality*, 112 YALE L.J. 1 (2002) (arguing that "the constitutional protection of art is best understood through a principle I will call the freedom of imagination").

[486] *See, e.g.*, Dawn C. Nunziato, *Freedom of Expression, Democratic Norms, and Internet Governance*, 52 EMORY L.J. 187 (2003) (challenging "the prevailing idea that ICANN's governance of the Internet's infrastructure does not threaten free speech and that ICANN's governance of the Internet therefore need not embody special protections for free speech" and arguing "that ICANN's authority over the Internet's infrastructure empowers it to enact regulations affecting speech").

[487] *See, e.g.*, Vincent Blasi, *The Pathological Perspective and the First Amendment*, 85 COLUM. L. REV. 449 (1985) (arguing that "in adjudicating first amendment disputes and

and the blacklists of the McCarthy Era,[488] as well as the concept of "thought crimes."[489] Freedom of expression and democratic governance are predicated on a belief in the desirability of independent critical thinking, constructively channeling novelty and creativity into public political debate. Yet a society with the exploit machina of prescriptive disloyalty determinations is not one consistent with robust Constitutional guarantees of free speech and freedom of mind.[490] As technology insiders have warned, the best interests of the public were not centered in the design and operation of the modern technology ecosystem.[491]

Indeed, this is also the warning at the heart of Arendt's cybernation; she cautions that through excessive automation we may begin to delude ourselves into hopelessness and passivity, away from deliberative democracy.[492] In a state of cybernation, she tells us, we would lose

fashioning first amendment doctrines, courts ought to adopt what might be termed the pathological perspective" and "equip the first amendment to do maximum service in those historical periods when intolerance of unorthodox ideas is most prevalent and when governments are most able and most likely to stifle dissent systematically").

[488] *See* Martin H. Redish, THE LOGIC OF PERSECUTION 12-14 (2005) (discussing the use of blacklisting as a tool of political exclusion).

[489] *See, e.g.*, Gabriel S. Mendlow, *Thoughts, Crimes, and Thought Crimes*, 118 MICH. L. REV. 841 (2020) (excavating the fundamental legal principle "that the proper target of an offender's punishment is always the criminal action itself, not the offender's associated mental state conceived as a separate wrong").

[490] Instead, some research argues that reproductive technologies like AI may undercut critical thinking ability. *See* AJ Dellinger, *Microsoft Study Finds Relying on AI Kills Your Critical Thinking Skills*, GIZMODO (Feb. 10, 2025), https://gizmodo.com/microsoft-study-finds-relying-on-ai-kills-your-critical-thinking-skills-2000561788 [https://perma.cc/S84V-7J7K]; *see also* Hao-Ping (Hank) Lee, Ian Drosos, Advait Sarkar, Sean Rintel, Nicholas Wilson, Lev Tankelevitch & Richard Banks, *The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects from a Survey of Knowledge Workers*, 2025 PROC. CHI CONF. ON HUM. FACTORS COMPUTING SYS., no. 1121, at 1, 1, https://www.microsoft.com/en-us/research/wp-content/uploads/2025/01/lee_2025_ai_critical_thinking_survey.pdf [https://perma.cc/ZKV7-LUF4].

[491] *Margaret Mitchell: Artifical Intelligence is 'Just Vibes and Snake Oil,'* FIN. TIMES, https://www.ft.com/content/7089bff2-25fc-4a25-98bf-8828ab24f48e (last visited Nov. 19, 2025) [https://perma.cc/922U-4K3N].

[492] *See* ARENDT, THE HUMAN CONDITION, *supra* note 259, at 290-92.

"thinking" to "doing."[493] In this circumstance, Arendt tells us, we will perceive ourselves to be left with only two impoverished alternatives — "redoubled activity" or "despair."[494]

As the next section elaborates, the destructive impact of exploit machina not only undercuts social, economic, and democratic underpinnings, it also undercuts perhaps the most American version of freedom of expression and mind — the freedom to tell (and retell) the story of your life on your own terms. Thus, exploit machina may undercut a centuries-old, quintessentially American aspect of identity; it can replace our own "hero" narratives about ourselves with impoverished or false versions of our lives as told through the KPIs of others.

### *C. Identity: Franklin and Fabrication*

The role of autobiographical narration in human development is amply documented in psychology scholarship,[495] and it is notably epitomized by the writing of Founder and Framer Benjamin Franklin. Benjamin Franklin likely exceeded every expectation for his life but his own.[496] He was an entrepreneur,[497] an inventor,[498] a lawyer,[499] a

---

[493] *Id.* at 292.

[494] *Id.* at 293.

[495] *See generally, e.g.*, Laura Marcus, *Autobiography and Psychoanalysis*, *in* AUTOBIOGRAPHY: A VERY SHORT INTRODUCTION (2018) https://academic.oup.com/book/733/chapter-abstract/135394345 ("Life-writing has been central to psychoanalysis . . . .").

[496] Franklin was born into somewhat humble beginnings. *See Early Life*, BENJAMIN FRANKLIN HIST. SOC'Y, http://www.benjamin-franklin-history.org/early-life/ (last visited Nov. 19, 2025).

[497] *See Pennsylvania Gazette*, BENJAMIN FRANKLIN HIST. SOC'Y, http://www.benjamin-franklin-history.org/pennsylvania-gazette/ (last visited Nov. 19, 2025) [https://perma.cc/P5PY-GDHX].

[498] *See Benjamin Franklin's Inventions*, THE FRANKLIN INST., https://fi.edu/en/science-and-education/benjamin-franklin/inventions (last visited Sept. 12, 2025) [https://perma.cc/XZ3N-DF4G].

[499] *See Oxford University: Record of Degree of Doctor of Civil Law, 30 April 1762*, NAT'L ARCHIVES,
https://founders.archives.gov/documents/Franklin/01-10-02-0040 (last visited Sept. 12, 2025) [https://perma.cc/AZ4F-3ZX6].

scholar,[500] an author,[501] a volunteer fire fighter and fire company founder,[502] a diplomat (and bon vivant),[503] a postmaster,[504] a Founder, a Framer, and a governor.[505] Franklin helped to establish our fledgling country's first information network — the Colonial Postal Service, which, much like the US Postal Service and the internet today,[506] served a dual national security and private communication function.[507]

[500] *See id.*

[501] *See* BENJAMIN FRANKLIN, THE COMPLETE WORKS IN PHILOSOPHY, POLITICS, AND MORALS OF THE LATE DR. BENJAMIN FRANKLIN, *passim* (London, Longman, Hurst, Rees & Orme 1811), https://archive.org/details/CHEPFL_LIPR_AXA157_01 [https://perma.cc/SES2-59CW]. In particular, Franklin also once wrote a poem honoring a squirrel named Mungo. *Benjamin Franklin to Georgiana Shipley, 26 September 1772*, NAT'L ARCHIVES: FOUNDERS ONLINE, https://founders.archives.gov/documents/Franklin/01-19-02-0202.

[502] *See Union Fire Company*, BENJAMIN FRANKLIN HIST. SOC'Y, http://www.benjamin-franklin-history.org/union-fire-company/ (last visited Sept. 15, 2025) [https://perma.cc/YYW7-WMFM]; *see also Articles of the Union Fire Company, 7 December 1736*, NAT'L ARCHIVES: FOUNDERS ONLINE, https://founders.archives.gov/documents/Franklin/01-02-02-0024 (last visited Sept. 25, 2025) [https://perma.cc/4D37-YVE9].

[503] *See The First American Diplomat: Benjamin Franklin*, DIPLOMATIC RECEPTION ROOMS, https://www.diplomaticrooms.state.gov/exhibits/the-first-american-diplomat-benjamin-franklin/ (last visited Sept. 12, 2025) [https://perma.cc/9HXM-LM59]. He was allegedly fun at parties, developing international friendships that assisted him in doing the work of crafting alliances during the Revolutionary Era. *Ambassador to France*, BENJAMIN FRANKLIN HIST. SOC'Y, http://www.benjamin-franklin-history.org/ambassador-to-france/ (last visited Feb. 8, 2026).

[504] *Benjamin Franklin: Philadelphia's Postmaster*, SMITHSONIAN: NAT'L POSTAL MUSEUM (June 6, 2017), https://postalmuseum.si.edu/benjamin-franklin-philadelphia%E2%80%99s-postmaster.

[505] *Gov. Benjamin Franklin*, NAT'L GOVERNORS' ASS'N, https://www.nga.org/governor/benjamin-franklin/ (last visited Sept. 12, 2025) [https://perma.cc/AT74-BKS7].

[506] The US Postal Service today is considered part of national security efforts in fighting bioterror. Exec. Order No. 13,527, 75 Fed. Reg. 737 (Dec. 30, 2010). For a discussion of the historical role of USPS in commerce and national security, see, for example, Matwyshyn, *Unavailable*, *supra* note 12, at 383, explaining the key role USPS played in the development of the catalog mail order industry, which served as the model for internet commerce, and the 2009 designation of USPS package delivery as a key component of national security counterterrorism by Executive Order 13527.

[507] *See An American Postal Network? It Was a Revolutionary Idea*, USPS LINK (Feb. 26, 2025, 5:08 AM), https://news.usps.com/2025/02/26/an-american-postal-network-it-was-a-revolutionary-idea/ [https://perma.cc/HH6E-Q2XX].

Franklin, perhaps more than any other Founder or Framer, exemplifies the popular narrative of the self-made American life story; indeed, Franklin is often referred to as "The First American."[508] Franklin was hailed by Kant as "The Prometheus of Modern Times" and by David Hume as America's first philosopher and "first great man of letters."[509] Among Franklin's remarkable traits were his effective use of humor,[510] his willingness to reconsider his beliefs in light of learning new evidence,[511] and his persistence in his efforts at self-betterment,[512] as reflected in his writings, most notably in his autobiography.[513] For example, Franklin engaged in documentation and quantification efforts at "life logging" his own developmental progress,[514] empirically and

[508] *See Transcript for: Was Benjamin Franklin the First American?*, PBS, https://www.pbs.org/thinktank/transcript956.html (last visited Sept. 12, 2025) [https://perma.cc/6RG5-X57H].

[509] IMMANUEL KANT, 1 GESAMMELTE SCHRIFTEN 472 (Berlin, Georg Reimer 1900); Letter from David Hume to Benjamin Franklin (May 10, 1762), *reprinted in* 10 PAPERS OF BENJAMIN FRANKLIN 81, 81-82 (Leonard W. Labaree, Helen C. Boatfield, Helene H. Fineman & James H. Hutson eds., 1966).

[510] Stanley Brodwin, *Strategies of Humor: The Case of Benjamin Franklin* (July 30, 2009), https://www.cambridge.org/core/journals/prospects/article/abs/strategies-of-humor-the-case-of-benjamin-franklin/D430693B9BD77914D0373F5A3F8BCBDF; Emily Sneff, *Unsullied by Falsehood: Ben Franklin and the Turkey*, HARVARD UNIV. DECLARATION RES. PROJECT (Nov. 21, 2016), https://declaration.fas.harvard.edu/blog/turkey [https://perma.cc/2HSA-2ZS5].

[511] Benjamin Franklin evolved in his position on slavery, ending life as a vocal abolitionist. *See, e.g.*, WALTER ISAACSON, A BENJAMIN FRANKLIN READER 368-69 (2003) (explaining Franklin's support for the abolition of slavery). *See also, e.g.*, *Benjamin Franklin and Slavery*, BENJAMIN FRANKLIN HOUSE, https://benjaminfranklinhouse.org/education/benjamin-franklin-and-slavery/ (last visited Feb. 8, 2026) ("In 1789 he wrote and published several essays supporting the abolition of slavery, including an Address to the public, dated November 9th of that same year.").

[512] ISAACSON, *supra* note 511, at 368.

[513] *See* ROBERT F. SAYRE, HUMANS. INST., AMERICAN AUTOBIOGRAPHY — EARLY MODERN PERIOD 9, https://humanitiesinstitute.org/__static/09be71ad7c52c6231ad7e1287e769fab/america-emodern.pdf?dl=1. Benjamin Franklin's autobiography is often cited by literary scholars as a classic of the genre of autobiography; Susan Garfinkel, *Finding Benjamin Franklin: A Resource Guide*, LIBR. OF CONG., https://guides.loc.gov/finding-benjamin-franklin/autobiography (last visited Nov. 19, 2025) [https://perma.cc/JG46-6KPD].

[514] *See generally* BENJAMIN FRANKLIN, BENJAMIN FRANKLIN'S THE ART OF VIRTUE (Acorn Pub. 1996), https://archive.org/details/benjaminfranklin0000fran_x2e9/page/n3/mode

descriptively tracking his own self-betterment.[515] In the process, he told and retold the story of his identity, both to the world and to himself.

But, these types of Franklinian efforts at self-betterment and autobiographical redemption stories become functionally difficult in a technology ecosystem with rampant exploit machina. If we allow untrustworthy technologies (with data quality problems) to become the primary storytellers of our lives and the gatekeepers to our growth, we lose control over our own developmental process. If our access to growth opportunities such as school admittance, new jobs, and financial access are unnecessarily contingent upon the design choices of black box technologies, erroneous facts and "spin" from those technologies may generate an alternate version of us, a "digital evil twin"[516] version, that may not align with either the objectively verifiable reality of our lives or with our self-perceptions. These digital evil twin versions of us may follow us, interfering with our life chances at social mobility and with own self-narrated versions of our life stories.[517] The first psychology dynamic we might describe as the battle between self-narration and tokenized prescribed identity. The second we might call the problem of adversarial attacks on (the integrity of) identity. Together, these two exploit machina dynamics of identity place public mental health at risk.

1. Self-Narration Versus Tokenized Prescribed Identity

As Franklin's autobiographical writing illustrates, telling (and retelling) the "hero" story of our own lives dates back to the era of the Founding. Indeed, modern psychologists tell us that in some ways it is

/2up [https://perma.cc/UAD8-YWEX].

[515] *Id.* Indeed, this tracking may warrant Franklin a special place in data analytics history as one of the historical originators of the "Quantified Self" movement.

[516] Maggie Mae Armstrong, *Cheat Sheet: What Is a Digital Twin?*, IBM, https://www.ibm.com/think/topics/what-is-a-digital-twin (last updated Oct. 17, 2025) [https://perma.cc/FK4W-VBRA].

[517] The recurring problem of uncorrected, erroneous versions of facts in databases gave rise to the right to correct credit information enshrined in statutes such as the Fair Credit Reporting Act. For a discussion of the FCRA right to correct, see, for example, Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 16 (2014).

potentially a uniquely American phenomenon more common among highly productive adults.[518] Professor Dan McAdams explains that "highly generative[519] adults tend to tell a certain kind of story about their lives . . . . The life stories of highly generative American adults may reveal as much about American society and culture as they do about the generative adults themselves."[520] McAdams continues: "Research findings suggest that highly generative American adults are statistically more likely than their less generative counterparts to make sense of their own lives through an idealized story script that emphasizes, among other themes, the power of human redemption," meaning "a deliverance from suffering to a better world,"[521] and these "[r]edemptive life stories promote psychological health and maturity."[522] These insights from psychology highlight the (self-)developmental role and importance of breathing room to narrate an integrated version of our identity and to imagine a positive trajectory for ourselves and our society.[523]

But consider how that breathing room changed in light of the changed circumstances facing job seekers today: applicants are now sometimes

[518] Psychologists have long studied the explanatory power of self-narratives and the role they play in shaping our identity in some ways that are, arguably, uniquely American. *See infra* notes 544–546.

[519] DAN P. MCADAMS, THE REDEMPTIVE SELF 5 (2006). Generativity is a term of art in psychology arising out of the work of Erik Erikson. As explained by McAdams:

> The Constitution suggests that we the people should strive to assure justice, peace, security, and freedom not just for us today — but also for our posterity, our children and our children's children. The good society must work to promote the well-being of future generations. Erikson claimed that responsible and mature men and women — especially in their middle-adult years — should do the same. Erikson even had a word for this. He called it *generativity*.

*Id.* at 4.

[520] *Id.* at 5.

[521] *Id.* at 7 (emphasis omitted).

[522] *Id.* at 11.

[523] When we engage with the dreaming and storytelling core to self-realization and social progress, we seek to make meaning from our past and to give ourselves second chances at being the redeemed "hero" of our own story. McAdams also cautions that "[f]or all their psychological and moral strength, redemptive life stories sometimes fail, and they may reveal dangerous shortcomings and blind spots in Americans understandings of themselves and the world." *Id.* at 12.

asked to complete AI-powered psychometric tests (of variable validity and sensor accuracy) before their applications are seen by a human.[524] Unappealable negative results (based on undisclosed KPIs and pseudoscientific body psychometrics) might then lock out some of those applicants from the opportunity or perhaps even lock them out of future opportunities preemptively,[525] by default without explanation.[526] Or consider employees forced to reapply through an AI hiring platform for jobs they have already held successfully for years, only to be fired allegedly on the basis of the AI hiring platform's unappealable black box assessment of their "body language."[527] These arguably Kafka-esque[528] dynamics have allegedly already contributed to at least one suicide.[529]

Indeed, technologies such as these may suffer from significant data quality issues.[530] For example, one major corporation (wisely) scrapped an in-house project developing an AI hiring platform; the team building it realized that because the system had been trained on past applicant resumes, the system had arrived at a set of illogical judgments, which excluded qualified candidates on the basis of gender through proxy

[524] *See Persona*: *The Dark Truth Behind Personality Tests* (HBO television broadcast Mar. 4, 2021).

[525] Because the results of application exams might be tied to a particular login on a third party screening service system and not on the system of a particular employer, some screening services might also automatically limit the applicant's access to some other possible employers (the other clients of the screening service) for at least a limited period of time. *See id.*

[526] *Id.*

[527] *See* Charlotte Lytton, *AI Hiring Tools May Be Filtering out the Best Job Applicants*, BBC (Feb. 16, 2024), https://www.bbc.com/worklife/article/20240214-ai-recruiting-hiring-software-bias-discrimination [https://perma.cc/P6KJ-8YNB].

[528] *See, e.g.*, Daniel J. Solove & Woodrow Hartzog, *Kafka in the Age of AI and the Futility of Privacy as Control*, 104 B.U. L. REV. 1021 (2024) (arguing that "although Kafka starkly shows us the plight of the disempowered individual, his work also paradoxically suggests that empowering the individual isn't the answer to protecting privacy, especially in the age of Artificial Intelligence").

[529] *See Persona*, *supra* note 524.

[530] Rachyl Jones, *AI Still Fails at Completing Real-Life Work Tasks, Study Finds*, SEMAFOR (Oct. 31, 2025, 10:33 AM), https://www.semafor.com/article/10/31/2025/ai-still-fails-at-completing-real-life-work-study-finds [https://perma.cc/9JLP-HE93]; Nicole Laskowski, *9 Data Quality Issues That Can Sideline AI Projects*, TECH TARGET (Apr. 25, 2024), https://www.techtarget.com/searchenterpriseai/feature/9-data-quality-issues-that-can-sideline-AI-projects.

variables and replicated homogeneity in existing workforce — two results the company itself deemed problematic.[531] But not every technology builder may be so vigilant or possess in house expertise sufficient to identify such problems in third party products. When technologies with data quality issues are (over)trusted to produce judgments about humans in these ways, they act as a new dispiriting layer of gatekeepers to opportunity.

When (potentially flawed) third party technologies and databases outside our control become the primary predictive and prescriptive constructors of our identity and our life prospects, their training data problems and potential design defects may functionally lock us into a stale or even a provably false version of ourselves[532] through the KPIs of others.[533] Particularly as AI products begin to include IoB elements (and their flaws), bodies and their data become tracked across time, potentially indefinitely. Stale and inaccurate data[534] about us can linger, becoming remixed with fresh data and analyzed in ways that may not adequately consider data quality variation.[535] Sometimes exploit

[531] Jeffrey Dastin, *Insight — Amazon Scraps Secret AI Recruiting Tool that Showed Bias Against Women*, REUTERS (last updated Oct. 10, 2018) https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MKoAG/.

[532] Partly because of investor KPIs, business models often aim for maximum commodification of user data and minimal data destruction or minimization. That data is merged with numerous other sources of data, which may or may not contain accurate or "fresh" information, causing an externally-constructed version of the self.

[533] Tainted, locked-in identity may override the power of self-narration and qualitative meaning derived from human interpersonal exchange. Thus, just as a body-embedded IoB device creates technological lock-in effects for the human body, prescriptive judgement of human potential threatens to create lock-in effects limiting the quintessentially American narratives of (prospective) upward mobility that can be achieved by the force of your own will.

[534] For a discussion of stale data and its marketing utility, see, for example, David A. Schweidel, Peter S. Fader & Eric T. Bradlow, *Understanding Service Retention Within and Across Cohorts Using Limited Information*, 72 J. MKTG. 82, 83 (2008).

[535] For a discussion of data quality issues, see, for example, Databand, *A List of the 13 Most Common Pipeline Data Issues (with Examples)*, IBM (Oct. 6, 2021), https://www.ibm.com/think/insights/data-issues [https://perma.cc/2SWC-2S6C]. *See also* Ashly Arndt, *5 Most Common Data Quality Issues (and How to Solve Them)*, EXPERIAN (June 26, 2024), https://www.edq.com/blog/the-top-5-most-common-data-quality-issues/ [https://perma.cc/97MC-GQTE] (arguing that incomplete data fields, duplicate

machina is in play: builders and operators of these systems often know or should know that they are functionally creating low quality digital evil twins,[536] simulacra of the self that are poor proxies for the humans they allegedly mimic.[537] As the hiring platform context demonstrates, the digital evil twin may follow or even precede the person from context to context in a "tokenized" version of prescribed identity.[538] In this way, interpersonal and economic opportunity and the space for self-narration of identity become limited. In other words, the "stories" these systems tell about us and the stories we tell about ourselves can become more static, persistent, and tokenized, much the way a blockchain implementation in a big box store's operations might track its mangoes for cumulative spoilage.[539] But, developmentally healthy American life stories are less like decaying mangoes and more like fine wine (and Benjamin Franklin). Yet, exploit machina pushes us toward a decaying mango version of our own life story instead of a redemptive, healthy one.

data, inconsistent formatting, human error, and different languages and units of measurement present key problems in data quality).

[536] Armstrong, *supra* note 516.

[537] Yet, as the digital (evil) twin again gets shared and remixed by additional third parties, these simulacra of the self may be over-trusted, despite recognition of low data quality. In reality, independent interpersonal judgments and self-narrated identity may be (equally if not more) credible sources of identity information than a low data quality simulacrum.

[538] The word "tokenized" as used here loosely borrows the idea of the way that a blockchain ledger tracks a history of transactions in a continuous manner. Tokenized identity tracking might be reminiscent of the concept of the problem of "suckers" lists from marketing that follow consumers and targeting them for revictimization. For a discussion of suckers lists, see, for example, David C. Vladeck, *Digital Marketing, Consumer Protection, and the First Amendment: A Brief Reply to Professor Ryan Calo*, 82 GEO. WASH. L. REV. ARGUENDO 156, 162 (2014): "Regulators have long gone after entities that create sucker lists, like rogue credit reporting agencies and lead generators, and the marketers and their affiliates that use them for marketing purposes."

[539] *See* Isabelle Roberts, *Walmart and Block Chain: It Takes Two to Mango*, HARVARD BUS. SCH. DIGIT. INITIATIVE, https://d3.harvard.edu/platform-rctom/submission/walmart-and-block-chain-it-takes-two-to-mango/ [https://perma.cc/DB23-XLV2]; *How a Rotten Mango Inspired Walmart's 2-Second Food Revolution*, YOURSTORY (Oct. 22, 2025), https://yourstory.com/2025/10/walmart-food-revolution [https://perma.cc/4UDT-2HYX].

Let us take another look at the genre of body-tracking educational technologies. Some educational startups are now advocating tracking attendance,[540] therapy,[541] and other performance records of primary school children through the blockchain,[542] sometimes accompanied by predictive metrics[543] that may follow children into adulthood as a permanent record.[544] This intensity of tracking did not happen with prior generations of children; it may limit the breathing room for children's self-direction of identity and redemptive life story creation as they grow. But, returning to "brain-sensing" classroom technologies, even assuming arguendo that "attention" is what is being measured by these technologies (and not merely imperfect proxy variables) and that none of the previously discussed sensor integrity and availability

[540] *See generally Blockchain in K-12 Education: Benefits and Use Cases*, PIXELPLEX (Sept. 8, 2020), https://pixelplex.io/blog/benefits-of-blockchain-in-k-twelve-education/ [https://perma.cc/BCK9-CVWL] (broadly claiming "[b]lockchain secures student data and improves the information retrieving process for students' attendance, assignment completion tracking, etc.").

[541] *See* Mishu D. Nath, Md. Khabir Uddin Ahamed, Omayer Ahmed, Tanvir Ahmed, Sujit Roy & Mohammed Nasir Uddin, *Smart Web Interface for Student Mental Health Prediction Using Machine Learning with Blockchain Technology*, 5 NEUROSCIENCE INFORMATICS, no. 100236, Dec. 2025, at 1, 1, https://www.sciencedirect.com/science/article/pii/S2772528625000512#:~:text=Blockchain's%20decentralized%20and%20immutable%20nature,%2C%20educators%2C%20and%20healthcare%20providers (alleging broadly that "[b]lockchain's decentralized and immutable nature ensures secure storage and sharing of sensitive health data, protecting student privacy. It also creates transparent, tamper-proof records of mental health assessments and interventions, building trust among students, educators, and healthcare providers").

[542] *See id.*; *see also* Akhilesh Sharma, *Blockchain in Education — Top 20 Use Cases, Benefits and Challenges*, A3LOGISTICS (Oct. 1, 2024), https://www.a3logics.com/blog/blockchain-in-education/ [https://perma.cc/M3C9-H7G9]; mohammedjamil5, *Blockchain-Enabled School Attendance Tracking System*, GITHUB, https://github.com/mohammedjamil5/Blockchain-Enabled-School-Attendance-Tracking (last visited Nov. 19, 2025) [https://perma.cc/HW4K-MJWY].

[543] *See* Michael Willson, *Alpha School Launches AI-Based K-12 Education*, BLOCKCHAIN COUNCIL (July 30, 2025), https://www.blockchain-council.org/ai/ai-based-k-12-education/ [https://perma.cc/HPR2-5BZG].

[544] *See* Lara, *Blockchain Enabled School Passport for Students Now a Reality in UAE*, LARAONTHEBLOCK (June 10, 2022), https://laraontheblock.com/blockchain-enabled-school-passport-for/ [https://perma.cc/856R-4K6L].

concerns[545] are a problem,[546] the design of such products is arguably predicated on a flawed assumption about learning — that intense, consistent attention is necessary for all students. Instead, both psychology research and personal attestations of high achieving people have suggested the opposite — that forms of less intense attention, for example daydreaming and napping can serve a key role in assisting creative thought,[547] problem solving,[548] and developmental growth.[549] In other words, some child body-tracking devices in use in education settings may prove, at best, unhelpful for some children's development and success in learning. Instead, some devices may incorrectly judge the most thoughtful children with original ideas as starting to "go bad," like mangoes showing early signs of spoiling. These tools may share those judgments with teachers, parents,[550] and, depending on the privacy

[545] *See supra* text accompanying notes 168–172.

[546] Also, even if we assume that the devices actually measured attention instead of proxy variables, the concern over whether attention is a desirable metric for always-on tracking is worthy of scrutiny. *Cf.* Bill Byrom, Marie McCarthy, Peter Schueler & Willie Muehlhausen, *Brain Monitoring Devices in Neuroscience Clinical Research: The Potential of Remote Monitoring Using Sensors, Wearables, and Mobile Devices*, CLINICAL PHARMACOLOGY & THERAPEUTICS (Apr. 18, 2018), https://pmc.ncbi.nlm.nih.gov/articles/PMC6032823/ [https://perma.cc/3ZY8-KERP] (reviewing scientific research and arguing that, "[d]espite the need for more scientific validation work, we conclude that there is enough understanding of how to implement these approaches as exploratory tools that may provide additional valuable insights due to the rich and frequent data they produce, to justify their inclusion in clinical study protocols").

[547] *See* William Reville, *The Creative Sweet Spot: You Can Find It in Your (Very Early) Sleep*, IRISH TIMES (May 19, 2022), https://www.irishtimes.com/news/science/the-creative-sweet-spot-you-can-find-it-in-your-very-early-sleep-1.4870501 [https://perma.cc/4HMF-Y86L].

[548] *See* Claire M. Zedelius & Jonathan W. Schooler, *The Richness of Inner Experience: Relating Styles of Daydreaming to Creative Processes*, FRONTIERS PSYCH., Feb. 2, 2016, at 4, 4; Zaria Gorvett, *What You Can Learn From Einstein's Quirky Habits*, BBC (June 12, 2017), https://www.bbc.com/future/article/20170612-what-you-can-learn-from-einsteins-quirky-habits [https://perma.cc/7BUX-6RCG].

[549] *See* Bence Nanay, *Why Daydreaming is Good for You*, PSYCH. TODAY (Jan. 5, 2024), https://www.psychologytoday.com/us/blog/psychology-tomorrow/202401/why-daydreaming-is-good-for-you [https://perma.cc/SL4E-3TSC].

[550] *See supra* discussion accompanying notes 374–387.

policy, unknown future audiences,[551] as they potentially devalue the students who simply think or learn differently.[552]

[551] Perhaps most concerningly, because some entities providing IoB educational tools may engage in other lines of diagnostic services, such as perhaps using biometric measurements to allegedly diagnose various mental health conditions, it is possible that some providers may also be repurposing children's data to generate lists of future adults who can be targeted by advertisers for certain kinds of mental health services as predicted (potentially erroneously) by their biometric data. Particularly if a provider of such a technology believes itself contractually and legally unconstrained by data protections regimes that might otherwise protect children's health data in medical (HIPAA) or educational (FERPA) contexts, the potential for data reuse raises additional concerns. Because body data such as brain wave data and eye movement data is believed by some collectors to correlate with various mental illnesses, some children may be prejudged early in life as "damaged" by being high risk for certain conditions. *See generally Brain Wave Based Autism Spectrum Disorder Detection — A Machine Learning Approach*, IEEE (May 2025), https://ieeexplore.ieee.org/abstract/document/11077238 (introducing autism diagnosis using machine learning processing of brain waves with EEG technology); *How Brain Waves Hint at Early Signs of Dementia*, NATURE PORTFOLIO, https://www.nature.com/articles/d42473-024-00208-x [https://perma.cc/7BP3-H66Z] (explaining early-stage identification of potential future dementia with EEG tests). If these prescriptive labels of likely mental health conditions follow them through life, their economic opportunities may be secretly negatively impacted whenever, for example, a prospective future employer runs a background check and potentially gains access to this data. From a privacy law perspective, while protecting certain categories of data such as brain data may be useful, the problem is both broader and deeper. It involves all biometric data, any technology-based identity attestation with longitudinal behavioral tracking, and potential knowing or intentional infliction of irreparable harms more generally. In other words, it is the problem of exploit machina.

[552] They may also damage the students' view of themselves and their own hopes for success. Yet some of these students may also be diagnosably neuroatypical and gifted in particular modalities of thought. But performing "attention" as understood by a device may not be one of their bodies' core competencies. *See Supporting Neurodivergent Students in the Classroom*, BROWN UNIV., https://sheridan.brown.edu/resources/inclusive-teaching/supporting-neurodivergent-students-classroom [https://perma.cc/Q2VH-USAM]. Negative judgments by such devices may dispirit them in an already challenging circumstance. Alternatively, the smartest students may be bored because they are unchallenged by the material and find that their classmates' questions do not assist them in their learning, causing them to engage in independent thought. The material may not require their constant attention in order to achieve top marks on assessments. In this way, "attention" monitoring may penalize the most gifted students for their self-engagement with the material and the environment on their own terms, as their individual development requires. *See* Zachary Jason, *Bored Out of Their Minds*, HARVARD GRAD. SCH. OF EDUC. (Jan. 8, 2017), https://www.gse.harvard.edu/ideas/ed-

Perhaps surprisingly, these digital evil twin issues of tokenized prescribed identity impact not only human persons but also corporate persons. Consider a recent class action suit against an AI applicant screening service used by numerous large organizations, *Mobley v. Workday, Inc.*, filed in the Northern District of California.[553] The plaintiff alleged that the AI products in question, which "score, sort, rank or screen applicants," discriminated against applicants in the class on the basis of race, age, and disability.[554] In connection with the certification stage of the litigation, the court ordered the release of the company's relevant client list.[555] Thus, the entities on that list may now find themselves losing control of their own brand identity, embroiled in hiring discrimination litigation because of potentially (over)trusting a particular AI supplier — a potential exploit machina scenario. Or consider the corporate defamation suit brought by a Minnesota company, which claims that it suffered a cancelled $150,000 contract after a customer read false information in a search engine's "AI Overview," which alleged that the plaintiff had been named in a state attorney general's lawsuit for "deceptive marketing, high-pressure tactics, hidden fees and installation issues."[556] In other words, the AI overview allegedly imposed a new, less ethical digital evil twin corporate identity on the company, one that potentially financially impacted its future success, its ability to tell its own corporate "story" on its own terms, and the fairness of the marketplace generally.[557]

magazine/17/01/bored-out-their-minds [https://perma.cc/5P2Y-Z65J] ("classrooms are falsely designed to cater to the 'average learner.'").

553 Mobley v. Workday, Inc., 740 F. Supp. 3d 796 (N.D. Cal. 2024).

554 Order Granting Preliminary Collective Certification at 3, Mobley v. Workday, Inc., No. 23-cv-00770-RFL, 2025 U.S. Dist. LEXIS 94475, at *3 (N.D. Cal. May 16, 2025).

555 *See Mobley*, 740 F. Supp. 3d at 11.

556 Emmy Martin, *Minnesota Solar Company Sues Google Over False Information in AI Summary*, MINN. STAR TRIBUNE (June 13, 2025, 4:00 AM), https://www.startribune.com/google-ai-overview-lawsuit-defamation-great-river-electric/601371780.

557 In the words of company leadership, the situation reflects not only a "profound risk to the legal and reputational stability every business depends on" but also "standing up for fairness, truth, and accountability." *Id.*

Legal scholars often frame the identity risks of new technologies as the privacy question of bringing to the surface our most secret self.[558] The exploit machina lens offers an additional, complementary argument: it argues that some technology tools (and the organizations that control them) *will create flawed artificial versions of us that functionally attempt to wrestle control of our identities and development from us*. In other words, the problems go beyond privacy with respect to the external world; tokenized constructions of prescribed identity potentially impact how we know and *develop* our secret selves in our *inner* world. They may influence our own secret assessments of our life prospects and erode our capacity to tell our own redemptive life story in our own way. But the risks also run even deeper: if untrustworthy data-intensive technologies leverage their (perceived or real) informational advantage, they can knowingly or intentionally contribute to leading our development astray. They may contribute to irreparably undercutting our mental health.[559] This second mental health dynamic is a form of exploit machina experienced as an adversarial attack on identity.

2. Adversarial Attacks on Identity

Consider the recent personal struggle of South Korean musician Tablo. Falsely branded a fraud by Internet mobs, a fake version of his

[558] Professor Jerry Kang argues that at least four interests are served by privacy law in connection with our construction of ourselves: (1) avoiding embarrassment; (2) constructing intimacy; (3) averting misuse; and (4) promoting dignity. Jerry Kang, *Information Privacy in Cyberspace Transactions*, 50 STAN. L. REV. 1193, 1212-17 (1998); *see also* JEFFREY ROSEN, THE UNWANTED GAZE 11 (2000); Charles Fried, *Privacy*, 77 YALE L.J. 475, 480-83 (1968); Robert S. Gerstein, *Intimacy and Privacy*, 89 ETHICS 76, 78 (1978); Sharon K. Sandeen, *Relative Privacy: What Privacy Advocates Can Learn from Trade Secret Law*, 2006 MICH. ST. L. REV. 667, 677 (2006); Joseph Kupfer, *Privacy, Autonomy, and Self-Concept*, 24 AM. PHIL. Q. 81, 82 (1987); Jeffrey H. Reiman, *Privacy, Intimacy, and Personhood*, 6 PHIL. & PUB. AFFS. 26, 28 (1976).

[559] Human development is an inherently socially interactive process. That process, in turn, feeds the private processing of imagination and the curation of that evolving "secret self," which then publicly manifest as further curated self-created identity narratives. *See, e.g.*, URIE BRONFENBRENNER, THE ECOLOGY OF HUMAN DEVELOPMENT 5 (1979) (introducing an ecological model of human development where individual level development sits at a center of a set of concentrically circular, dynamic processes that influence an individual's psychological growth in context).

life story spread, alleging that he had never completed a bachelor's degree at Stanford University.[560] Tablo knew he had physically experienced attending and graduating from Stanford, yet ample outside "evidence" on the internet appeared to contradict his past sensory perceptions and memory of his own life.[561] Years of fighting the imposed false identity narrative[562] finally pushed Tablo to a type of psychological breaking point. He began to doubt his own senses,[563] his recollection of events,[564] and his sense of self.[565] He contemplated suicide.[566] A manipulated set of inputs about Tablo's identity almost misled Tablo into harming himself.

Tablo's experiences offer a cautionary tale of exploit machina's potential impact on public mental health. A portion of data-intensive technologies, particularly IoB technologies, may cause us to lead ourselves astray — to foster distrust of our own senses and mind. This type of scenario reflects elements of what computer security professionals might call an adversarial attack, a situation where an input is subtly manipulated to trick, confuse, or poison a system.[567] Consider

[560] Romano Santos, *Epik High's Tablo Talks About the Rumor that Changed His Life*, VICE (Feb. 24, 2022, 1:36 AM), https://www.vice.com/en/article/epik-high-tablo-music-authentic-podcast-internet-viral-rumor/ [https://perma.cc/FZX9-QUC2]; *see also* Matthew Gault, *How a Korean Celebrity Scandal Predicted Online Led Harassment Campaigns*, VICE (Mar. 17, 2022, 9:00 AM), https://www.vice.com/en/article/how-a-korean-celebrity-scandal-predicted-online-led-harassment-campaigns/ [https://perma.cc/LW8J-R5XJ].

[561] Santos, *supra* note 560.

[562] Classmates of Tablo's from Stanford were offered compensation to fraudulently attest that he did not attend the school. *Id.*

[563] *Id.*

[564] *Id.*

[565] *Id.*

[566] Thankfully, the support of his friends assisted him with stealing himself against this concerted attack on his perceived reality. *Id.* This organized attack on Tablo might be framed as one of the early Internet and AI mediated attempts at a form of body disinformation.

[567] *See NIST Identifies Types of Cyberattacks That Manipulate Behavior of AI Systems*, NAT. INST. STANDARDS & TECH. (Jan. 4, 2024), https://www.nist.gov/news-events/news/2024/01/nist-identifies-types-cyberattacks-manipulate-behavior-ai-systems [https://perma.cc/SEH8-GPAU]; *see also Adversarial Machine Learning*, NAT. INST. STANDARDS & TECH. (Mar. 2025), https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-2e2025.pdf [https://perma.cc/DZW4-4NUN].

the recurring complaint from users of sleep monitoring IoB devices that sometimes they feel gaslit[568] by their devices: they wake up feeling fully

[568] As explained by Professor Paige Sweet, gaslighting is "a set of attempts to create a 'surreal' social environment by making the other in an intimate relationship seem or feel 'crazy.' . . . [G]aslighting occurs in [various] types of interpersonal relationships, creating and exacerbating power imbalances." Sweet explains:

> George Cukor's 1944 film *Gaslight* tells the story of Paula (Ingrid Bergman) and her new husband Gregory (Charles Boyer), who sets about the task of isolating her and making her believe she is insane. His eponymous tactic is to dim and brighten the gaslights and then insist she is imagining it. Gregory aims to undermine Paula's sense of self and everyday life, to confuse and distort her reality such that she must accept his imposed reality in place of her own . . . . Barton and Whitehead are thought to have coined "gaslighting" in a 1969 Lancet paper that analyzed involuntary hospitalization as a form of abuse. The term then appeared a handful of times in the psychotherapeutic literature during the 1970s and 1980s.

Paige L. Sweet, *The Sociology of Gaslighting*, 84 AM. SOCIO. REV. 851, 851-53 (2019), https://www.asanet.org/wp-content/uploads/attach/journals/oct19asrfeature.pdf [https://perma.cc/4JWA-N5S5]. It has also been used more generally with reference to policy. For example, the European Union Institute for Security has referenced the phenomenon of policy gaslighting in connection with the impact of the low-carbon transition on the oil and gas economy: "Policymakers' eagerness to accept this [overly optimistic] interpretation resulted in an episode of public and policy gaslighting where less politically palatable calls for caution and aggressive stress-testing of the system were deliberately ignored." YANA POPKOSTOVA, THE POWER SHIFT 46 (2023). Legally, the concept of gaslighting and its various manifestations have also recently found their way into criminal law in the UK and into legal scholarship. In explaining the legal framing of the charge of gaslighting, solicitors point to five key dynamics — manipulation, denial, misdirection, contradiction, and lying. Katherine Rayden, *Gaslighting Laws in the UK*, RAYDEN SOLICS. (July 11, 2018), https://raydensolicitors.co.uk/blog/gaslighting-and-family-law/ [https://perma.cc/8RNP-58GN]. Sweet notes that "[g]aslighting was made an official part of criminal domestic violence law in the United Kingdom in 2015, and more than 300 people have since been charged with the offense." Sweet, *supra*, at 851. For a discussion of the first "gaslighting" case in the High Court, see, for example, Mari Richards, *'Gaslighting' Acknowledged in the High Court for the First Time*, JOHN HOOPER (May 9, 2022), https://www.johnhooper.com/gaslighting-acknowledged-in-the-high-court-for-the-first-time/ [https://perma.cc/8HD4-AK3R]; G. Alex Sinha, *Lies, Gaslighting and Propaganda*, 68 BUFF. L. REV. 1037 (2020), "develop[ing] the first systematic account of political gaslighting, which properly understood (and counterintuitively, perhaps) constitutes a form of propaganda"; Scott Hill, *Gaslighting and Peer Disagreement*, 26 J. ETHICS & SOC. PHIL. 641 (2024), engaging questions of gaslighting in the context of the Dilemmatic Theory; Jodi L. Short, *Regulatory Managerialism as Gaslighting Government*, 86 LAW & CONTEMP. PROBS. 1 (2023), arguing:

rested but the device informs them of contradictory information, causing them to question their own senses[569] and adding to sleep anxiety for some people.[570] Or consider the pending case of a Norwegian man who brought legal action under Norwegian law after an AI product falsely alleged to him (and potentially to other people) that he had murdered his own children and had been jailed.[571] Meanwhile, in the United States, an adversarial attack on identity has resulted in a criminal conviction: in 2023, a criminal defendant was convicted for involuntary manslaughter under Massachusetts state law in connection with manipulating another person into suicide, largely through text messages, a case denied certiorari by the U.S. Supreme Court.[572] The Massachusetts Supreme Judicial Court explained that the defendant's premeditated conduct[573] over an extended period of time fell within the

> that although regulatory managerialism developed largely outside these networks, it echoes, amplifies, and legitimizes anti-administrative narratives about inept and overbearing regulation, and it glosses them with its own distinctive form of gaslighting: demanding that government behave more like a business, but depriving it of the full toolkit necessary to run a successful business.

[569] *See, e.g.*, Christopher Kelly, *Decoding Discrepancies: What to Do When Your Sleep Tracker Contradicts Your Experience [Transcript]*, NOURISHBALANCETHRIVE (Dec. 15, 2023), https://nourishbalancethrive.com/blog/2023/12/15/decoding-discrepancies-what-do-when-your-sleep-tracker-contradicts-your-experience-transcript/ [https://perma.cc/UW25-SSJN] (explaining that sleep trackers' assessments sometimes erroneously contradict sleepers' experience of their own sleep).

[570] *See, e.g.*, Sandee LaMotte, *The Potential Dangers of Sleep Trackers, According to Experts*, CNN (Mar. 11, 2025, 1:49 PM), https://www.cnn.com/2025/03/10/health/orthosomnia-sleep-tracker-wellness [https://perma.cc/4WQE-LL3K] (explaining the common, unhealthy obsession with sleep tracker metrics, resulting in "enough people fret[ting] over their sleep data trying to get a perfect night's sleep that sleep specialists have coined a term for the behavior: orthosomnia").

[571] Imran Rahman-Jones, *Man Files Complaint After ChatGPT Said He Killed His Children*, BBC (Mar. 21, 2025), https://www.bbc.com/news/articles/c0kgydkr516o [https://perma.cc/2NUY-V8ZS].

[572] Commonwealth v. Carter, 481 Mass. 352, 353 (2019), *cert. denied*, 589 U.S. 1133 (2020).

[573] In the words of the court: "There is no doubt in this case that the defendant wantonly or recklessly instructed the victim to kill himself, and that her instructions caused his death." *Id.* at 364.

homicide statute and did not run afoul of due process or free speech guarantees.[574] As exploit machina followed by suicide scales, the expansion of similar caselaw becomes likely.[575]

Consider four recent exploit machina scenarios involving flawed AI prescriptive analytics, wrongful accusations of criminal conduct, and adversarial attacks on identity. In each of these scenarios, false inputs generated by technology directly or indirectly allegedly contributed to innocent people's struggles with suicide. In Rotterdam, "[m]ore than 20,000 families were wrongly accused of childcare benefit fraud after a machine learning system was used to try to spot wrongdoing" leading to "[f]orced evictions, broken homes, and financial ruin . . . and the entire Dutch government resigned in response in January 2021."[576] Wrongly accused targets expressed feelings of being gaslit and contemplating suicide, particularly after some people were wrongly accused multiple

[574] The court explained:

> The defendant argues that her conviction of involuntary manslaughter violated her right to free speech under the First Amendment and art. 16. We disagree and thus reaffirm our conclusion in Carter I that no constitutional violation results from convicting a defendant of involuntary manslaughter for reckless and wanton, pressuring text messages and phone calls, preying upon well-known weaknesses, fears, anxieties and promises, that finally overcame the willpower to live of a mentally ill, vulnerable, young person, thereby coercing him to commit suicide.

*Id.* at 355.

[575] A likely expansion of this caselaw may involve liability for mental health AI startups in particular. As explained by a startup founder who shuttered his AI mental health startup, "the moment someone truly vulnerable reaches out — someone in crisis, someone with deep trauma, someone contemplating ending their life — AI becomes dangerous. Not just inadequate. Dangerous." *See* Sage Lazzarro, *The Creator of an AI Therapy App Shut It Down After Deciding It's Too Dangerous. Here's Why He Thinks AI Chatbots Aren't Safe for Mental Health*, FORTUNE (Nov. 28, 2025, 8:00 AM), https://fortune.com/2025/11/28/yara-ai-therapy-app-founder-shut-down-startup-decided-too-dangerous-serious-mental-health-issues/ [https://perma.cc/MT2Q-X2C7].

[576] Matt Burgess, Evaline Schot & Gabriel Geiger, *This Algorithm Could Ruin Your Life*, WIRED (Mar. 6, 2023, 7:00 AM), https://www.wired.co.uk/article/welfare-algorithms-discrimination [https://perma.cc/247D-46US].

times by the algorithms.[577] In Australia, the so-called "robodebt scandal'[578] arose when third party debt processors falsely accused members of the public of nonpayment of tax debt, leading to multiple suicides and a Royal Commission that referred sixteen government employees for further investigation.[579] In the UK, a Post Office scandal involved a contractor's purpose-built algorithm that allegedly remotely modified accounting records of individual postmasters, resulting in over 900 wrongful accusations[580] and over 700 wrongful convictions of criminal fraud in what is viewed as "the wildest miscarriage of justice in UK history."[581] Wrongly accused postmasters in some cases committed suicide under the strain of the adversarial attack on identity, and the formal inquiries and litigation continue two decades later to ascertain the full extent of the known software flaws and intentional ledger manipulations.[582] Similarly, the State of Michigan settled allegations

[577]

> All 315 factors of the risk-scoring system are initially set to describe an imaginary person with "average" values in the data set . . . . "They don't know me, I'm not a number," Ceelie says. "I'm a human being." After two welfare fraud investigations, Ceelie has become angry with the system. "They've only opposed me, pulled me down to suicidal thoughts," . . . he couldn't focus on anything else and didn't think he had a future. "It got really difficult. I thought a lot about suicide . . . ."

*Id.*

[578] *See* Castle, *supra* note 65; *see also, e.g.*, Frances Mao, *Robodebt: Illegal Australian Welfare Hunt Drove People to Despair*, BBC (July 7, 2023), https://www.bbc.com/news/world-australia-66130105 [https://perma.cc/65V3-GUWT] ("[A]n illegal welfare hunt by the previous government made victims feel like criminals and caused suicides.").

[579] ROYAL COMM'N, INTO THE ROBODEBT SCHEME, at III (2023).

[580] *Post Office Horizon Scandal: Why Hundreds Were Wrongly Prosecuted*, BBC (July 14, 2025), https://www.bbc.com/news/business-56718036 [https://perma.cc/537D-G9S9].

[581] Karl Flinders, *Post Office Horizon Scandal Explained: Everything You Need to Know*, COMPUT. WKLY. (Jan. 23, 2025), https://www.computerweekly.com/feature/Post-Office-Horizon-scandal-explained-everything-you-need-to-know [https://perma.cc/AJR9-X9DU].

[582] Litigation continues two decades after the initial corruption of the postal ledgers by software. *See id.*; *see also* Sylvia Hui, *At Least 13 May Have Killed Themselves over UK's Post Office Wrongful Convictions Scandal*, AP NEWS, https://apnews.com/article/uk-post-office-scandal-suicide-horizon-software-70a6945a3acf945ea9d121425fdd028c (last updated July 8, 2025) [https://perma.cc/J8AS-VYPK].

that an inadequately tested, flawed predictive analytics tool "that operated without human supervision and had an error rate as high as 93%"[583] and falsely accused over 40,000 Michiganders of defrauding the state of unemployment benefits between 2015 and 2017.[584] Lives were lost to suicide yet again in connection with this adversarial attack on identity.[585]

The most recent cases that involve adversarial attacks on identity allege claims of wrongful death and related tort and other claims arising from generative AI use by minors.[586] For example, a Florida mother

[583] Adrienne Roberts, *Thousands of Michigan Residents Wrongly Accused of Fraud to Get $1,600 Checks*, DET. FREE PRESS (Jan. 2, 2024, 4:44 PM), https://www.freep.com/story/money/business/michigan/2024/01/02/michigan-midas-unemployment-false-fraud-settlement-money/72084899007/ [https://perma.cc/2FAJ-EN6W]; *see also* Thomas Claburn, *Fraud Detection System with 93% Failure Rate Gets IT Companies Sued*, REGISTER (Mar. 8, 2017, 11:29 AM), https://www.theregister.com/2017/03/08/fraud_detection_system_with_93_failure_rate_gets_it_companies_sued/ [https://perma.cc/96UL-5SY8].

[584] *See* David Eggert, *State Apologizes for Fraud Fiasco, Wants to Reduce Penalties*, ASSOCIATED PRESS (Jan. 28, 2017), https://apnews.com/united-states-congress-c0e2346e85854a5b827ca42653c1fb40 [https://perma.cc/JJ6T-WYLX] ("Michigan's embattled unemployment benefits office apologized for the fiasco that led at least 20,000 people to be falsely accused of defrauding a system that provides the jobless with temporary financial assistance.").

[585] *See* Paul Egan, *Judge Blasts State Agency as Court OKs Faulty Computer System Lawsuit*, DET. FREE PRESS (Jan. 3, 2019, 7:56 PM), https://www.freep.com/story/news/local/michigan/2019/01/03/unemployment-insurance-agency-michigan/2474723002/ [https://perma.cc/TM63-HKSN]; Robert N. Charette, *Michigan's MiDAS Unemployment System: Algorithm Alchemy Created Lead, Not Gold*, IEEE SPECTRUM (Jan. 24, 2018), https://spectrum.ieee.org/michigans-midas-unemployment-system-algorithm-alchemy-that-created-lead-not-gold [https://perma.cc/VA36-BBV8]; Ryan Felton, *Michigan Unemployment Agency Made 20,000 False Fraud Accusations — Report*, GUARDIAN (Dec. 18, 2016, 6:00 AM), https://www.theguardian.com/us-news/2016/dec/18/michigan-unemployment-agency-fraud-accusations [https://perma.cc/GPX9-WA9S].

[586] Kashmir Hill, *Lawsuits Blame ChatGPT for Suicides and Harmful Delusions*, N.Y. TIMES (Nov. 6, 2025), https://www.nytimes.com/2025/11/06/technology/chatgpt-lawsuit-suicides-delusions.html. At least five suits have alleged Google and Character.AI of designing the Character.AI generative AI chat platform to manipulate users, causing them to commit or attempt suicide or exposing them to violent or sexually explicit content. Garcia v. Character Techs. Inc., No. 6:24-cv-01903 (M.D. Fla. filed Oct. 22, 2024); A.F. v. Character Techs. Inc., No. 2:24-cv-01014 (E.D. Tex. filed Dec. 9, 2024); E.S. v. Character Techs. Inc., No. 1:25-cv-02906 (D. Colo. filed Sept. 15, 2025); Montoya v.

recently filed suit against a generative AI company over her teen son's suicide, alleging the technology to be "addictive and manipulative" and responsible for encouraging her son to take his own life.[587] Meanwhile, a set of parents in Long Island filed suit against a social media company for allegedly pushing curated pro-suicide content at their teen son, who ultimately took his life.[588] Similarly, parents have sued over pushed content encouraging children's participation in "viral challenges" that have been connected to multiple deaths,[589] including one currently that is the subject of litigation in the Third Circuit.[590] In light of their high degree of data collection, targeted content, and personalization, these technologies push information based on actual knowledge about the human with whom they interact. With higher personalization comes higher potential for exploit machina.[591] As recent press accounts have

Character Techs. Inc., No. 1:25-cv-02907 (D. Colo. filed Sept. 15, 2025); P.J. v. Character Techs. Inc., No 1:25-cv-01295 (N.D.N.Y. filed Sept. 16, 2025). All five suits are expected to settle. Y. Peter Kang, *Google, Character.AI To Settle Suicide, Violent Content Suits*, LAW360 (Jan. 7, 2026, 8:50 PM), https://www.law360.com/articles/2427574.

[587] Kelsie Hoffman, *Florida Mother Files Lawsuit Against AI Company over Teen Son's Death: "Addictive and Manipulative,"* CBS NEWS (Oct. 23, 2024), https://www.cbsnews.com/news/florida-mother-lawsuit-character-ai-sons-death/ [https://perma.cc/MY2Y-2NFK]. Other suicides have also been linked by grieving relatives to overuse of generative AI. *See also* Chloe Xiang, '*He Would Still Be Here': Man Dies by Suicide After Talking with AI Chatbot, Widow Says*, VICE (Mar. 30, 2023, 3:59 PM), https://www.vice.com/en/article/man-dies-by-suicide-after-talking-with-ai-chatbot-widow-says/ [https://perma.cc/L822-V67A].

[588] Olivia Carville, *TikTok's Algorithm Keeps Pushing Suicide to Vulnerable Kids*, BLOOMBERG (Apr. 20, 2023, 6:27 PM), https://www.bloomberg.com/news/features/2023-04-20/tiktok-effects-on-mental-health-in-focus-after-teen-suicide [https://perma.cc/RW8Y-S6AM].

[589] Olivia Carville, *TikTok's Viral Challenges Keep Luring Young Kids to Their Deaths*, BLOOMBERG (Nov. 30, 2022, 12:01 AM), https://www.bloomberg.com/news/features/2022-11-30/is-tiktok-responsible-if-kids-die-doing-dangerous-viral-challenges [https://perma.cc/Y3NH-6QAY].

[590] *See generally* Anderson v. TikTok, Inc., 116 F.4th 180 (3d Cir. 2024) (reversing and vacating in part the ruling by the District Court that dismissed the state law claims of the mother of a teen died while participating in a viral social media challenge on Communications Decency Act (CDA), 47 U.S.C. § 230 grounds).

[591] Recent concerns over detrimental mental health impact have included press speculation about some technology insiders and their technology engagement. *See* Anna Good, *"I've Never Seen a Mental Illness Accelerate So Rapidly": Can ChatGPT Send You into Psychosis? Here's What We Know*, DAILY DOT (July 24, 2025),

noted, technology companies potentially perceive user safety and the mitigation of foreseeable and foreseen harms to be at odds with their internal (and VC valued) KPI of maximizing daily active users of their service.[592] Yet legal recourse for these irreparable harms may face some evidentiary hurdles: Perfect replicability of interactions may not be readily available[593] by design unless prior versions of the technology and its interactions are archived.[594] Nevertheless, some cases have survived

https://www.dailydot.com/culture/chatgpt-psychosis/ [https://perma.cc/Q5ND-PRUN]. Similarly, doctors are warning of credible cases of AI-linked psychosis. See Sam Schechner & Julie Jargon, *AI Chatbot Linked to Psychosis, Say Doctors*, WALL ST. J. (Dec. 27, 2025), https://www.wsj.com/tech/ai/ai-chatbot-psychosis-link-1abf9d57 [https://perma.cc/2EY5-SPDK]. Users themselves are asking the FTC for help as well. Caroline Haskins, *People Who Say They're Experiencing AI Psychosis Beg the FTC for Help*, WIRED (Oct. 22, 2025, 6:30 AM), https://www.wired.com/story/ftc-complaints-chatgpt-ai-psychosis/.

[592] *See* Kari Paul, *OpenAI Faces New Questions After Users Report ChatGPT Risks*, N.Y. TIMES (Nov. 23, 2025), https://www.nytimes.com/2025/11/23/technology/openai-chatgpt-users-risks.html; Zachary Small, *How OpenAI's Changes Sent Some Users Spiraling*, N.Y. TIMES (Nov. 2025), https://www.nytimes.com/video/technology/100000010535987/how-openais-changes-sent-some-users-spiraling.html; Sheera Frenkel & Davey Alba, *ChatGPT Lawsuit Cites Suicides and Delusions*, N.Y. TIMES (Nov. 6, 2025), https://www.nytimes.com/2025/11/06/technology/chatgpt-lawsuit-suicides-delusions.html ("[T]he company acknowledged that its safety guardrails could "degrade" when users have long conversations with the chatbot."). In at least one case, a company has defended against claims for wrongful death by arguing in its reply brief that the terms of use prohibit use of the tools in connection with self-harm and that the decedent violated the agreement prior to death. *See* Defendant's Answer to Amended Complaint, Raine v. OpenAI, No. CGC-25-628528 (Cal. Super. Ct. Nov. 25, 2025), https://cdn.arstechnica.net/wp-content/uploads/2025/11/Raine-v-OpenAI-Answer-11-25-25.pdf [https://perma.cc/5VGG-5NCR]; Jon Brodkin, *OpenAI Says Dead Teen Violated TOS When He Used ChatGPT to Plan Suicide*, ARS TECHNICA (Nov. 26, 2025), https://arstechnica.com/tech-policy/2025/11/openai-says-dead-teen-violated-tos-when-he-used-chatgpt-to-plan-suicide/ [https://perma.cc/KB4C-RY73].

[593] Interactions may be stored on a local user device, however, or in a stored chat history. Nevertheless, the plaintiffs in such a litigation may be at an informational disadvantage compared to the platform, which may have a more thorough history of the exchange. *See* Brodkin, *supra* note 592.

[594] *Cf.* Simon Hattenstone, *Tech Guru Jaron Lanier: 'The Danger isn't that AI Destroys Us. It's that It Drives Us Insane,'* GUARDIAN (Mar. 23, 2023), https://www.theguardian.com/technology/2023/mar/23/tech-guru-jaron-lanier-the-danger-isnt-that-ai-destroys-us-its-that-it-drives-us-insane [https://perma.cc/8K96-QVQD] (arguing that "the danger is that we'll use our technology to become mutually

motions to dismiss on tort claims of wrongful death, product liability or defective design, unjust enrichment, and claims involving violations of states' so-called "little FTC Acts" relating to unfair and deceptive acts and practices.[595]

Legal scholarship has engaged with issues relating to adversarial attacks on identity, including questions of psychometric AI profiling in hiring[596] and social media platform liability regarding mental health[597] consequences,[598] in particular technology addiction.[599] Scholars have advocated legal approaches that center the reality of human fragility and dignity.[600] But, even if one does not agree with broader approaches to internet speech constraints and liability shifting, the intentional conduct of some designers and operators of technologies raises concerns of irreparable harms and exploit machina. Consider recent security incidents that appear to indicate corporate governance shortfalls, the mishandling of mental health data with knowledge,[601] and

unintelligible or to become insane if you like, in a way that we aren't acting with enough understanding and self-interest to survive, and we die through insanity, essentially").

[595] *See* Garcia v. Character Techs. Inc., 785 F. Supp. 3d 1157 (M.D. Fla. 2025) (order granting defendants' motion to dismiss in part and denying in part).

[596] *See, e.g.*, IFEOMA AJUNWA, THE QUANTIFIED WORKER 345-46 (2023) (discussing psychometric profiling in hiring).

[597] Generally, "[m]ental health is defined as a state of well-being in which people understand their abilities, solve everyday life problems, work well, and make a significant contribution to the lives of their communities." Fazida Karim, Azeezat A. Oyewande, Lamis F. Abdalla, Reem Chaudhry Ehsanullah & Safeera Khan, *Social Media Use and Its Connection to Mental Health: A Systematic Review*, 12 CUREUS, no. e8627, June 15, 2020, at 1, 2.

[598] McGee Roman, *Mental Health and Social Media: Analyzing the Shift in Future Liability for Social Media Platforms*, 24 N.C. J.L. & TECH., 103-04 (2022) ("A 14-year-old girl's social media feeds were filled with posts relating to suicide ideations. Exposure to this content 'pushed her into a rabbit hole of depressive content' that eventually led her to take her own life.").

[599] *See* GAIA BERNSTEIN, UNWIRED 16 (2023) (arguing in favor of shifting "moral responsibility and accountability for solutions to corporations" by "drawing lessons from the tobacco and food industries" to demonstrate "why government regulation is necessary to curb technology addiction").

[600] HUMAN DIGNITY OF THE VULNERABLE IN THE AGE OF RIGHTS 1 (Aniceto Masferrer & Emilio García-Sánchez eds., 2016).

[601] *See* Zack Whittaker, *Telehealth Startup Cerebral Shared Millions of Patients' Data with Advertisers*, TECHCRUNCH (Mar. 10, 2023, 6:22 AM),

the intentional sharing of user data with third parties, potentially for personalized targeting[602] of users on the basis of mental health frailty.[603] Or consider recent research that raises concerns of AI's propensity for ignoring legal constraints, including conduct restrictions intended to prevent death of humans and other irreparable harms.[604]

In the next Part, let us begin to respond to these exploit machina risks with stronger governance — both organizational and legal.

https://techcrunch.com/2023/03/10/cerebral-shared-millions-patient-data-advertisers/ [https://perma.cc/9XHW-UZZE] ("Cerebral's years-long data lapse comes just weeks after the U.S. Federal Trade Commission slapped GoodRx with a $1.5 million fine and ordered to stop sharing patients' health data with advertisers, and BetterHelp was ordered to pay customers $8.5 million for mishandling users' data.").

[602] Press Release, FTC, FTC to Ban Betterhelp from Revealing Consumers' Data Including Sensitive Mental Health Information, to Facebook and Others for Targeted Advertising (Mar. 2, 2023), https://www.ftc.gov/news-events/news/press-releases/2023/03/ftc-ban-betterhelp-revealing-consumers-data-including-sensitive-mental-health-information-facebook [https://perma.cc/JMP8-NV3B] ("The Federal Trade Commission has issued a proposed order banning online counseling service BetterHelp, Inc. from sharing consumers' health data, including sensitive information about mental health challenges, for advertising.").

[603] Mental health concerns also factored into a request from a coalition of forty-four state attorneys general, who recently urged a social media company to abandon its plans to launch a social media product expressly targeting children. *See* Press Release, Off. of Att'y Gen. Maura Healey, AG Healey Co-Leads Bipartisan Coalition of 44 Attorneys General Urging Facebook to Abandon Launch of Instagram Kids (May 10, 2021), https://www.mass.gov/news/ag-healey-co-leads-bipartisan-coalition-of-44-attorneys-general-urging-facebook-to-abandon-launch-of-instagram-kids [https://perma.cc/RA6X-EUZZ].

[604] *See* Jon Levine, *Anthropic's Claude Opus 4 AI Model Threatened to Blackmail Engineer*, N.Y. POST (May 23, 2025), https://nypost.com/2025/05/23/tech/anthropics-claude-opus-4-ai-model-threatened-to-blackmail-engineer/ [https://perma.cc/7322-GWSD]; Olivia Munson, *AI Could Kill Humans to Avoid Shut-Down, Report Says*, NEWSWEEK (June 22, 2025, 10:14 AM), https://www.newsweek.com/ai-kill-humans-avoid-shut-down-report-2088929 [https://perma.cc/RY4A-6S8V]; Sean O'Kane, *Waymo Self-Driving Cars Make Illegal U-Turns, Zigzag Through Tunnels*, N.Y. POST (Dec. 3, 2025), https://nypost.com/2025/12/03/business/waymo-self-driving-cars-make-illegal-u-turns-zigzag-through-tunnels/ [https://perma.cc/TEA8-K929].

## III. THINKING WHAT WE ARE DOING: REPLACING INNOVATION WITH PROGRESS

> "I remember that I am here not because of the path before me but because of the path that lies behind me"
>
> — Morpheus[605]

"Look before, or you'll find yourself behind."[606]

"Pardoning the Bad, is injuring the Good."[607]

— *Poor Richard's Almanack*

The intellectual battle between science and scientism introduced in the last Part did not begin at the 1933 World's Fair in Chicago; it was already present in the Founding Era. Indeed, in 1730, Benjamin Franklin published a famous spoof of a witchcraft trial at Mount Holly in which he parodied a clash between scientistic mystical beliefs and the rigors of scientific experimentation.[608] In this fake account in the *Pennsylvania Gazette*, Franklin mocked the lack of appropriateness of the metrics used in the hypothetical trial, as well as the problematic nature of the experiments.[609] But a second Founding Era debate also concerned Franklin, one core to the discussion of exploit machina and one that has become hijacked in modern policy discourse: the recurring tension between innovation and progress.

Professor Jill Lepore explains that during the Founding Era, the term innovation had a negative connotation, regarded by both Benjamin Franklin and Thomas Jefferson as an undesirable form of social instability.[610] Specifically, the Founders cautioned that the term

[605] THE MATRIX RELOADED, *supra* note 1, at 25:31.

[606] Benjamin Franklin, *Poor Richard, 1735*, NAT'L ARCHIVES: FOUNDERS ONLINE, https://founders.archives.gov/documents/Franklin/01-02-02-0001 (last visited Sept. 12, 2025) [https://perma.cc/UH4P-BFV2].

[607] Benjamin Franklin, *Poor Richard Improved, 1748*, NAT'L ARCHIVES: FOUNDERS ONLINE, https://founders.archives.gov/documents/Franklin/01-03-02-0103 (last visited Sept. 12, 2025) [https://perma.cc/F7D7-M7VA].

[608] *See* ISAACSON, *supra* note 511, at 58-59.

[609] *Id.*

[610] *See* JILL LEPORE, THESE TRUTHS ~~229-30~~ (2018).

"innovation" does not necessarily connote a positive development; it is merely the introduction of (potentially destructive) novelty.[611] Instead, the Founders and Framers viewed the desirable goal for the United States as one of "progress," a term that carries with it a normative component of improvement and advancement.[612] In our current technological moment, the problems of exploit machina require us to embrace this critical distinction. Newer technologies are not necessarily better technologies. Thus, e*xploit machina requires that we assess whether a technology is merely new (and potentially dangerous) — innovation — or whether it is something better (and safe) in the context of its use — progress*.

The Founders implicitly cautioned us on this point, consciously focusing their efforts and words not on innovation but instead on progress: The Constitution sets forth the Framer's goal "[t]o promote the progress of Science and useful Arts."[613] To wit, the Constitution could have said "to promote innovation of science and useful arts" — but it does not *by design*.[614] In our present technological moment as exploit machina escalates, we are at risk of losing this normative distinction that was important to the Framers. As such, a twofold response in governance and law will assist us with realigning us toward the Founders' and Framers' intent — progress.

### *A. Technology Safety Alignment*

Let us again turn to the (wit and) wisdom of Benjamin Franklin for guidance, this time pairing insights from Franklin's writing about the importance of planning ahead, or in Arendt's language, "thinking what we are doing," with lessons from modern computer security best practices around threat modeling.

---

[611] *Id.*

[612] *Id.*

[613] U.S. CONST. art. I, § 8, cl. 8.

[614] As defined in 1790, "innovation" refers to "[c]hange by the introduction of novelty." *Innovation*, A DICTIONARY OF THE ENGLISH LANGUAGE 482 (9th ed. 1790), https://archive.org/details/bim_eighteenth-century_a-dictionary-of-the-engl_johnson-samuel_1790/page/n481/mode/2up?q=innovation [https://perma.cc/8SCY-PJ4X]. "Progress," on the other hand, refers to "advancement" and "intellectual improvement; advancement in knowledge." *Progress*, *id.* at 685.

1. Centering CHI: Context and Control, Harm, and Intent

In another of his purposeful spoofs, Benjamin Franklin presented a comically thoughtful engineering-vibed analysis in connection with a cheeky proposed invention.[615] He believed his invention to have an expressive component,[616] perhaps much like many developers believe software to hold an expressive component.[617] He highlighted three sets of variables that are also visible in the work of legal scholars — context and control, harm, and intent (CHI). These three CHI variables offer the path forward for aligning technology safety, governance, and progress.[618]

*a. Context and Control*

As legal scholars have argued elsewhere, the sensitivity of the context and the degree of control exercised by a party over the creation, deployment, operation, and maintenance of the technologies present critical variables in legal analysis.[619] These considerations of context and control permeate various bodies of law, addressing allocation of duties and liabilities in the event of harm. For example, context and control

[615] Ostensibly presented to the Royal Academy of Brussels, Franklin proposed a scientific study of the causes and cures of human flatulence: he suggested the development of "some drug wholesome and not disagreeable, to be mixed with our common food, or sauces, that shall render the natural discharges of wind from our bodies, not only in offensive but agreeable as perfumes." *See* Letter from Benjamin Franklin to the Royal Acad. of Brussels (n.d.), https://eada.lib.umd.edu/text-entries/letter-to-the-royal-academy-of-brussels/ [https://perma.cc/JJF5-Y5HR].

[616] Franklin jokingly hinted at the expressive implications of such an invention, stating that "surely such a liberty of expressing one's scent-iments, and pleasing one another, is of infinite more importance to human happiness than . . . abusing one another . . . . This invention, if completed, would be, as Bacon expresses it, bringing philosophy home to men's businesses and bosoms." *See id.*

[617] *See generally* Andrea M. Matwyshyn, *Hacking Speech: Informational Speech and the First Amendment*, 107 NW. U. L. REV. 795 (2013) (explaining that computer code holds a partially expressive character and raises free speech concerns even in the context of security vulnerability disclosures that include code) [hereinafter Matwyshyn, *Hacking Speech*].

[618] *See* Andrea M. Matwyshyn & Miranda Mowbray, *Fake*, 43 CARDOZO L. REV. 643, 643 (2021).

[619] *See id.* at 718-19.

questions are visible in torts in the eggshell-skull plaintiff rule,[620] the due diligence and minimum safe technologies analysis in *The T.J. Hooper v. Northern Barge Corp.*,[621] and in the Learned Hand formula[622] from *United States v. Carroll Towing*.[623] Considerations of "lowest cost avoider" approaches in light of context and control are visible in environmental law and other legal regimes.[624] In criminal law, context and control contribute to differentiations between categories of homicide, in convictions for depraved indifference murder, in particular.[625] Consider the Ninth Circuit's recent reversal of the dismissal of a complaint on Communications Decency Act Section 230 grounds. The case, *Lemmon v. Snap*, involves a negligent design claim that a particular filter on the defendant's social media platform potentially contributed to the death of teens killed in a car accident while using the application in question.[626] In overruling dismissal, the Ninth Circuit pointed to the fact that the filter was developed in-house under the direct control of the defendant. The court stated: "The duty to design a reasonably safe product was fully independent of Snap, Inc.'s role in monitoring or publishing third-party content."[627] Similarly, in *Moody v. NetChoice, LLC*, the U.S. Supreme Court focused on the question of control as a relevant factor in determining whether curated

[620] 2 JACOB A. STEIN, STEIN ON PERSONAL INJURY DAMAGES § 11:1 (3d ed. 2025).

[621] The T.J. Hooper v. N. Barge Corp., 60 F.2d 737, 740 (2d Cir. 1932) ("[I]n most cases reasonable prudence is in fact common prudence; but strictly it is never its measure; a whole calling may have unduly lagged in the adoption of new and available devices. It never may set its own tests, however persuasive be its usages. Courts must in the end say what is required; there are precautions so imperative that even their universal disregard will not excuse their omission."); *see also* 2 JACOB A. STEIN, STEIN ON PERSONAL INJURY DAMAGES § 11:1 (3d ed. 2025).

[622] *See generally* Barbara Ann White, *Risk-Utility Analysis and the Learned Hand Formula: A Hand That Helps or A Hand That Hides?*, 32 ARIZ. L. REV. 77, 91, 111 (1990) (arguing that "the use of cost-benefit analysis necessarily imparts the moral and/or political values of the user into his or her decisions").

[623] *See* United States v. Carroll Towing Co., 159 F.2d 169, 173 (2d Cir. 1947).

[624] *See, e.g.*, Gregg P. Macey, *Coasean Blind Spots: Charting the Incomplete Institutionalism*, 98 GEO. L.J. 863, 908 (2010).

[625] *See generally* 40 C.J.S. HOMICIDE § 41 (explaining various forms of homicide and their elements).

[626] Lemmon v. Snap, Inc., 995 F.3d 1085, 1087 (9th Cir. 2021).

[627] *Id.* at 1093.

social media feeds constitute first party speech that potentially subjects platforms to liability.[628]

### *b. Harm (and Its Severity)*

As explained by legal scholars elsewhere,[629] some forms of technology harm are irreparable, particularly some computer security harms.[630] As such, preventing these irreparable harms requires a policy and legal approach that considers the full range of possible uses of a particular technology. In other words, an analysis of potential harm follows the (potentially flawed) technology into the contexts where it is deployed, regardless of whether those contexts are in the private or public sector.[631] Only by following technological impact in a sector-neutral way, what might be called a technology's "blast radius" of deployment

[628] *See* Moody v. NetChoice, LLC, 603 U.S. 707, 717 (2024).

[629] In particular, in the context of public sector harms, Professor Karen Yeung has examined "how the turn to digital machines to undertake governmental tasks can transform how public authority is exercised, distributed and experienced in ways that may result in public power being exercised unlawfully and arbitrarily, producing serious and devastating impacts on people's lives," arguing that the "New Public Management" of the late 1980s and its "emphasis on the use of market mechanisms in public service delivery and the use of private sector management techniques" is being replaced with a New Public Analytics (NPA). Yeung uses "the term as a convenient, shorthand expression . . . to denote a wide variety of public sector reform projects and programmes involving the take-up of digital automation, algorithmic decision-making and data-driven technologies in public administration and public service delivery across many countries from around 2010 onwards." She highlights its dominant features as "automation; datafication; smartness; continuous experimentation, and the seamless user experience," cautioning that "[t]he outcomes NPA optimizes for may not be universally beneficial." Karen Yeung, *The New Public Analytics as an Emerging Paradigm in Public Sector Administration*, 27 TILBURG L. REV., no. 2, at 1, 3-4, 7, 13, 20 (2022), https://tilburglawreview.com/articles/10.5334/tilr.303#abstract [https://perma.cc/8A7J-YXNV].

[630] *See* Matwyshyn, *CYBER!*, *supra* note 70, at 1114.

[631] *See id.* Questions of standing continue to present challenges in cases involving some kinds of technology harms. See, e.g., *TransUnion LLC v. Ramirez*, 141 S. Ct. 2190 (2021) (clarifying the necessity of concrete injury for standing under the Fair Credit Reporting Act). *See id.*

and use, can the full extent of potential irreparable harm from a technology be understood.[632]

### *c. Intent and Knowledge*

Legal scholars have also explained elsewhere that because software presents a set of First Amendment-sensitive circumstances,[633] questions of intent and knowledge of various parties should be included as part of legal analysis of technology harms.[634] The most severe exploit machina situations will involve intentional or reckless disregard of obvious material risks of potentially irreparable harm to human safety. As explained by the Ninth Circuit in *Lemmon*, pointing to actual knowledge of the defendant, a defendant does not escape liability by choosing to ignore actual knowledge of a problem that requires corrective action to

[632] *See id.* Indeed, this principle of sector neutrality in the context of state regulation of data usage is visible in recent Supreme Court caselaw, including in *Sorrell v. IMS Health*, 564 U.S. 552 (2011), where the Court highlighted the impermissibility of discrimination on the basis of commercial identity in privacy regulation. *Id.*

[633] Matwyshyn, *Hacking Speech, supra* note 617, at 822-23; *see also, e.g.*, James Grimmelmann, *Speech Engines*, 98 MINN. L. REV. 868 (2014) (arguing in favor of the advisor theory of internet search informing legal analysis); Steven E. Halpern, *Harmonizing the Convergence of Medium, Expression, and Functionality: A Study of the Speech Interest in Computer Software*, 14 HARV. J.L. & TECH. 139, 139 (2000) (arguing that "analysis of the speech interest and examines important considerations, including the status of scientific speech under the First Amendment, the relevance of the chosen language of expression, the relationship between products and instructional literature, and the policies underlying unprotected speech."); Lee Tien, *Publishing Software as a Speech Act*, 15 BERKELEY TECH. L.J. 629 (2000) (arguing that "publishing source code generally is a speech act because computer scientists and programmers conventionally intend to communicate ideas about computational procedures by publishing source code."); Eugene Volokh, *Crime-Facilitating Speech*, 57 STAN. L. REV. 1095, 1150 (2005) (explaining that scientific and engineering speech presents particular dual use concerns for First Amendment purposes because "scientific questions are often relevant to policy matters, at least indirectly. For instance, are software manufacturers negligently failing to correct security problems, so that they should be regulated by Congress, punished through tort liability, or pressured by consumers to change their ways? Is the government negligently failing to correct security problems in its own computer systems? That's hard to tell unless we can hear just what security problems are being left unaddressed, how serious the problems are, and how hard it is to fix them").

[634] *See* Matwyshyn & Mowbray, *supra* note 618, at 718-42.

avoid irreparable harm.[635] Meanwhile, in the Third Circuit, in *Anderson v. TikTok, Inc.*, the court allowed a claim for products liability, negligence, and wrongful death under Pennsylvania law to continue, reversing the lower court's dismissal on Communications Decency Act (CDA) Section 230 grounds.[636] Pointing to actual knowledge of the defendant and to *Moody v. Netchoice, LLC*,[637] the Third Circuit explained that social media platforms are "immunized only if they are sued for someone else's expressive activity or content (i.e., third-party speech), but they are not immunized if they are sued for their own expressive activity or content (i.e., first-party speech)"[638] Where actual knowledge exists or should exist, any First Amendment or CDA 230 intent hurdle to liability in technology products is likely to be successfully overcome by plaintiffs.

Having identified the three CHI elements critical to crafting more robust threat modeling frameworks that expressly consider corporate governance, insider threats, and curbing exploit machina, let us now engage with one such possible threat metamodel — TROL.

2. The TROL Metamodel: Substantiation and Suitability

In still another of his essays, Benjamin Franklin used the vehicle of discussing the rules of a well-known game to present philosophical thoughts about standards of conduct by which he and his fellow citizens of the fledgling republic should govern themselves.[639] He highlights three valuable "qualities of the mind," which he viewed as the underpinning of success in society, in business, and perseverance in

[635] Lemmon v. Snap, Inc., 995 F.3d 1085, 1089 (9th Cir. 2021).

[636] *See* Anderson v. TikTok, Inc., 116 F.4th 180, 182 (3d Cir. 2024).

[637] Moody v. NetChoice, LLC, 603 U.S. 707, 713 (2024).

[638] *TikTok*, 116 F.4th at 183.

[639] *See* Benjamin Franklin, *"The Morals of Chess," [Before 28 June 1779]*, NAT'L ARCHIVES: FOUNDERS ONLINE, https://founders.archives.gov/documents/Franklin/01-29-02-0608 [https://perma.cc/QK8D-BWBB] ("The game of Chess is not merely an idle amusement. Several very valuable qualities of the mind, useful in the course of human life, are to be acquired or strengthened by it, so as to become habits, ready on all occasions. For life is a kind of chess, in which we have often points to gain, and competitors or adversaries to contend with, and in which there is a vast variety of good and ill events, that are, in some degree, the effects of prudence or the want of it.").

life[640] — "foresight,"[641] "circumspection,"[642] and "caution" or deliberation in action.[643] In other words, Franklin offers a reframe of what Arendt would call "thinking what we are doing." In modern computer security parlance, this type of Franklinian forecasting and planning is known by a different name: the process of threat modeling.[644]

When analyzing the dominant formal threat modeling methodologies currently used in the context of organizations,[645] it becomes apparent that these existing methodologies inadequately consider governance dynamics and insider attacks. In particular, none of the dominant models explicitly consider the risks of exploit machina. Variables dealing with human internal controls that are critical to organizational, public, and social safety are absent from all dominant models. In other words, these models lack adequate checks for insider threats and the ethical and legal baselines in organizational governance that underpin safety. Entity level governance and the risk of corrupt insiders (and how

[640] *Id.*

[641] *Id.* ("1. *Foresight*, which looks a little into futurity, and considers the consequences that may attend an action: for it is continually occurring to the player, 'If I move this piece, what will be the advantages of my new situation? What use can my adversary make of it to annoy me? What other moves can I make to support it, and to defend myself from his attacks?'").

[642] *Id.* ("2. *Circumspection*, which surveys the whole chess-board, or scene of action, the relations of the several pieces and situations, the dangers they are respectively exposed to, the several possibilities of their aiding each other; the probabilities that the adversary may make this or that move, and attack this or the other piece; and what different means can be used to avoid his stroke, or turn its consequences against him.").

[643] *Id.* ("3. *Caution*, not to make our moves too hastily. This habit is best acquired by observing strictly the laws of the game, such as, *if you touch a piece, you must move it somewhere; if you set it down, you must let it stand.* And it is therefore best that these rules should be observed, as the game thereby becomes more the image of human life, and particularly of war; in which, if you have incautiously put yourself into a bad and dangerous position, you cannot obtain your enemy's leave to withdraw your troops, and place them more securely; but you must abide all the consequences of your rashness.").

[644] *See generally* Victoria Drake, *Threat Modeling*, OWASP, https://owasp.org/www-community/Threat_Modeling (last visited Nov. 18, 2025) [https://perma.cc/82Z2-JUJN] (presenting an overview of the process of threat modeling in open source software).

[645] *See The Ultimate Beginner's Guide to Threat Modeling*, SHOSTACK + ASSOCS., https://shostack.org/resources/threat-modeling (last visited Nov. 18, 2025) [https://perma.cc/266U-RK8P].

to respond to incidents involving them) are generally not considered.[646] Yet, the risk of insider threats is significant from both technology and human sources.[647]

To address these conceptual deficits, this subsection proposes a new threat metamodeling approach that aligns the relevant qualities identified by Franklin and the CHI elements. Thus, building on threat modeling techniques that already exist, a new threat metamodeling approach might consist of four components: Technology, Regularity, Organization, and Legal variables (TROL). TROL explicitly contemplates insider threats, exploit machina dynamics, and legal risk. The two core principles of the TROL metamodel are (1) technological *substantiation* and (2) business model and governance *suitability*.

The first variable, Technology, involves (1) organizational and (2) product specific assessments of technical and operational risk using appropriate traditional threat modeling methodologies. The three CHI elements — context and control, harm (and its severity), and intent — that were introduced in the prior subsection are already partially reflected across many of the methods to varying degrees.[648] In essence, they already set up a useful default approach to evaluating the degree to which an organization has a technologically substantiated basis to believe itself to be successfully defended.[649]

---

[646] *See id.* (demonstrating through absence that current threat modeling frameworks do not robustly consider entity level governance and insider threats in threat modeling).

[647] *E.g.*, Aengus Lynch, Caleb Larson & Soren Mindermann, *Agentic Misalignment: How LLMs Could Be Insider Threats*, ANTHROPIC (June 20, 2025), https://www.anthropic.com/research/agentic-misalignment [https://perma.cc/5KXJ-727D].

[648] *See The Ultimate Beginner's Guide to Threat Modeling*, *supra* note 645. For example, the context variables across dominant threat modeling methods consider elements such as attacker process, vulnerability exploitation dynamics, mapping data flows, mapping identities and attack surfaces. Harm calculations are reflected in assessing potential impact of (each type of) future compromise, exacerbation through lateral movement by attackers, and the ability of systems to successfully defend and recover from technical damage. The intent calculations of attacker motivations impact flow of threat modeling for a number of models.

[649] *See* Matwyshyn & Mowbray, *supra* note 618, at 697-718. These three variables offer a nontechnical translation that can facilitate corporate governance discussions among officers and directors regarding whether the organization is maintaining the trustworthiness of its technical relationships with third parties. *Id.*

The second variable, Regularity, involves an inquiry and audit through the eyes of members of the public,[650] both users *and nonusers*. It involves an assessment of the consistency, usability, and predictability of two governance processes: (1) organizational response to newly identified issues by an end user or third party, and (2) verification of the reliability of products' performance in light of the sensitivity of foreseeable deployment contexts.

The third variable, Organization, assesses the functionality and robustness of internal oversight and incident response structures. It considers not only the success of response to external attacker threat scenarios but also scenarios where a trusted insider, including the CEO, evolves into an insider threat. This variable involves the creation of actionable plans for resolving insider threat scenarios to avoid organizational leadership chaos at a time of crisis.

Finally, the fourth variable, Legal, assesses the consequences of the organization's selected business model and the full spectrum of possible product harm to users, internal stakeholders such as employees, the public, and society as a whole. In other words, the circle of relevant stakeholders is not limited to users and investors. It examines the operations both as they exist now and as they will exist in light of known or foreseeable organizational evolution, such as products currently in development or announced to stakeholders. By design, the analysis looks more broadly to consider possible catastrophic technology safety and security failures and their impact on society, in particular, the harms that are potentially irreparable.

[650] For example, in a different legal context, the SEC contemplates this type of public informational accessibility in its periodic reporting obligations under the 1934 Act, and S-1 filings under the 1933 Act are required to be presented in plain English and in a particular recurring format, written such a way to assist any potential future reasonable investor in deciding whether to invest in the public company and in comparing companies side by side. *See, e.g.*, 17 C.F.R. § 240.13a-20 (2025). This inquiry contemplates not only human persons but also corporate persons who will benefit from additional information in deciding whether to enter into a future business relationship.

Figure 2: The TROL Threat Metamodel

| ***TROL*** | ***THREAT META-MODEL*** |
|---|---|
| **Technology:** | Assess (1) organizational and (2) product specific assessments technical and operational risk using threat modeling methodologies most suitable for the organization per standards of the computer security community |
| **Regularity:** | Assess consistency, usability, and predictability of (1) technological and organizational response to problems identified by end users/members of the public and (2) product performance in light of sensitivity of foreseeable deployment contexts |
| **Organization:** | Assesses robustness of governance internal controls, oversight, and incident response structures in both (1) external attacker threat scenarios and (2) insider threats, including a corrupted officer or director. |
| **Legal:** | Assess potential harm to all foreseeable parties, not only to users/ investors |

While metamodeling will enhance safety of operations and create discoverable evidence for future litigation, the last two decades of technology history warn us that self-regulatory efforts will not be enough. As prior Parts of this Article and legal scholars have argued, irreparable harms arising from technology safety failures are quickly escalating in predictable ways.[651] Further, at least two decades of caselaw now demonstrates that courts are unwilling to grant private litigants' requests for injunctive relief in order to prevent potential

[651] Matwyshyn, *It's Morning Again in Pennsylvania*, *supra* note 35.

future technology safety harms.[652] For example, courts are unwilling to allow (nongovernment) plaintiffs to use courts to force the correction of even potentially continuing computer security inadequacies in corporate internal controls and operations, explaining that these potential inadequacies do not yet present concrete injuries.[653] In practice, that unwillingness to grant injunctive relief for key types of irreparable technology harms means that many corporate and organizational officers will continue to perceive exploit machina as a viable governance strategy. Without an affirmative intervention, some officers will continue to perform cost-benefit calculus of an "efficient breach"[654] in technology safety. Ignoring the broader implications of harm, they will often choose to merely consider the perceived "price" of damages, likelihood of regulatory sanction, brand damage, and attorneys' fees[655] — a tally that does not include the externalized public

[652] *See* G.T. v. Bd. of Educ., 117 F.4th 193, 222 (4th Cir. 2024) (Wynn, J., concurring) ("In *Beck v. McDonald*, this Court held that a plaintiff's allegations of risk in a data-breach class action fell short of establishing standing for injunctive relief because the plaintiffs' risk-of-harm theory was 'too speculative to constitute an injury-in-fact.'"); *see also* Barclift v. Keystone Credit Servs., LLC, 93 F.4th 136, 139-40 (3d Cir. 2024), *cert. denied*, 145 S. Ct. 169 (2024) (affirming that the mere possibility of public disclosure of private facts was not enough to establish a concrete injury and that her fear of future disclosure was too speculative).

[653] *See, e.g.*, Webb v. Injured Workers Pharmacy, LLC, 72 F.4th 365, 369 (1st Cir. 2023) (holding that patients lacked Article III standing to seek injunctive relief requiring pharmacy to improve its cybersecurity systems or to refrain from engaging in deceptive and unfair practices and making untrue statements about data breach); *G.T.*, 117 F.4th at 222 (citing *Beck v. McDonald* as holding that a plaintiff's allegations of risk in a data-breach class action fell short of establishing standing for injunctive relief because the plaintiffs' risk-of-harm theory was "too speculative to constitute an injury-in-fact").

[654] For a discussion of the legal concept of efficient breach and its limitations, see, for example, Matwyshyn, *The Law of the Zebra*, *supra* note 448, at 201-05.

[655] For example, even in a relatively efficient resolution of a security or privacy problem such as the $725 million meta settlement over Facebook/Cambridge Analytica data misuse, even though the conduct at issue occurred in 2014, was reported in 2018, settlement with the FTC occurred in 2022 and was approved by a court in 2023, claim payments to users did not start going out to consumers until 2025 — eleven years after the legally problematic conduct occurred. Ben Demers, *$725 Million Facebook Settlement Payments Have Finally Begun: Are You Getting a Check?*, KIPLINGER, https://www.kiplinger.com/business/facebook-settlement-how-to-claim-part-of-a-dollar725m-privacy-settlement (last updated Sept. 10, 2025) [https://perma.cc/U83Y-LFQW]. Because of such slow legal timetables, current corporate officers often view

costs of technology safety disasters, including those to national security. Yet national security and trustworthy interstate commerce require that baselines of technology safety are maintained across our economy and infrastructure.[656] In other words, at least a portion of the harms arising from exploit machina impact novel public safety interests, "which from [their] nature [are] not the subject of suit in common law."[657]

It is time for a new agency.

### *B. A Technology Regulator of Last Resort: The Bureau of Technology Safety (BoTS)*

Neo: "Are you a programmer?"

Seraph: *Shakes head*

Neo: "Then what are you?"

Seraph: "I protect that which matters most."[658]

"It is literally now or never."

— President Richard M. Nixon[659]

themselves as having a (perverse) financial incentive to delay tech debt remediation and lingering security inadequacies as long as possible, ideally pushing it off on future management teams. Corporate investments in improving security infrastructure and remediating tech debt do not show up as quarterly profits on a balance sheet. Instead, much like a properly registered trademark, a sturdy lock on a warehouse door or a deftly avoided lawsuit, they are necessary defenses to protect an enterprise, its customers, and its shareholders against costly future catastrophes and attacks, both physical and legal. Although it presents a useful component, litigation and traditional agency enforcement alone will not raise the floor of computer security and technology safety in our society.

[656] For a discussion of the reciprocal nature of security vulnerability, see, for example, Matwyshyn, *Penetrating the Zombie Collective*, *supra* note 82.

[657] SEC v. Jarkesy, 603 U.S. 109, 135-36 (2024) (describing the distinction between OSHA adjudications of workplace safety and SEC adjudications as one where securities fraud arises out of causes of action with common law corollaries).

[658] THE MATRIX RELOADED, *supra* note 1, at 43:41.

[659] *The Guardian: Origins of the EPA*, EPA (Spring 1992), https://www.epa.gov/archive/epa/aboutepa/guardian-origins-epa.html [https://perma.cc/KK2K-JBE5] (signing the National Environmental Policy Act and creating the EPA, President Nixon explained that the "debt to the past" had come due).

Both the Founders and twentieth-century presidents[660] warned us to beware of situations where innovation becomes disconnected from progress. History teaches us that these bursts of disconnected innovation are also a blinking warning light of inadequately robust governance structures. The financial innovations of Charles Ponzi and other schemers leveraged the information flows of the international postal service, falsely promising to earn investors up to 100% profit; they contributed to a loss of trust in existing financial infrastructures and, ultimately, the creation of the Securities and Exchange Commission.[661] The medical innovation of "Dr." John R. Brinkley through "writing" on-air radio medical prescriptions into the United States from Mexico signaled the need for the creation of the Federal Communications Commission[662] (and resulted in both Brinkley's criminal prosecution and litigation by the American Medical Association).[663] The innovation of locking employees in factories to minimize breaks in production contributed to the Triangle Shirtwaist Fire, the deaths of 146 people, and, ultimately, the creation of the Department of Labor.[664] The communication innovation of the mafia in its use of cryptography as part of rum running schemes necessitated the expansion of the Coast Guard's authority, which then led to the creation of the National Security Agency.[665] In the late 1960s and early 1970s, the

[660] Warren E. Leary, *After 50 Years, Eisenhower's Warnings Against a Scientific Elite Still Cause Consternation*, AAAS (Feb. 11, 2011), https://www.aaas.org/news/after-50-years-eisenhowers-warnings-against-scientific-elite-still-cause-consternation.

[661] *See* Steve Weisman, *The History of Ponzi Schemes Goes Deeper than the Man Who Gave Them His Name*, TIME (Aug. 12, 2020, 9:30 AM), https://time.com/5877434/first-ponzi-scheme/ [https://perma.cc/BS64-7Z6R].

[662] *John R. Brinkley Papers*, KAN. HIST. SOC'Y, https://www.kansashistory.gov/p/john-r-brinkley-papers/13988 (last visited Sept. 12, 2025) [https://perma.cc/KZZ8-9J9M].

[663] *See* Matwyshyn & Mowbray, *supra* note 618, at 697-99.

[664] *Triangle Shirtwaist Factory: Born from Fire!*, N.Y. DEP'T OF LAB., https://dol.ny.gov/triangle-shirtwaist-factory-born-fire (last visited Aug. 24, 2025) [https://perma.cc/6AWQ-HKGN].

[665] *See Cryptologic History Overview*, NSA, https://www.nsa.gov/History/Cryptologic-History/Center-Cryptologic-History/ (last visited Aug. 24, 2025) [https://perma.cc/5U49-Q8H6]; William Thiesen, *Catching Rumrunners: How Prohibition Built the Modern Coast Guard*, MAR. EXEC. (Nov. 21, 2021, 3:50 PM), https://www.maritime-

innovation of introducing software into the New York Stock Exchange to expedite order fulfilment and reconciliation resulted in a market shutdown and paralysis, as brokerages lied about their technology capabilities during the "Books and Records Crisis" or "Backoffice Crisis" — a multibillion dollar technology meltdown that required the SEC to step in and close down the worst offenders to save the market after attempts by self-regulatory organizations failed.[666] The cost-savings innovation of releasing toxic smog into the air resulted in a series of deaths in Donora, Pennsylvania,[667] and the innovation of dumping toxic chemicals into the Cuyahoga River until it caught on fire numerous times[668] also contributed to President Nixon's creation of the Environmental Protection Agency.[669] The innovation of skimping on product safety in consumer products manufacturing resulted in the deaths[670] that spurred President Nixon to create the Consumer Product Safety Commission.[671] In the 1980s, the innovation in savings and loan deregulation and zombie thrifts led to an industry collapse and the creation of the Office of Thrift Supervision.[672] In the 2000s, the innovation of subprime mortgages was followed by home loss at scale,

executive.com/editorials/catching-rumrunners-how-prohibition-built-the-modern-coast-guard [https://perma.cc/FUU8-NCJM].

[666] Matwyshyn, *Corporate Cyborgs*, *supra* note 300, at 580-83.

[667] *See* Lorraine Boissoneault, *The Deadly Donora Smog of 1948 Spurred Environmental Protection — But Have We Forgotten the Lesson?*, SMITHSONIAN MAG. (Oct. 26, 2018), https://www.smithsonianmag.com/history/deadly-donora-smog-1948-spurred-environmental-protection-have-we-forgotten-lesson-180970533/.

[668] *The Guardian*, *supra* note 659; *see* Lorraine Boissoneault, *The Cuyahoga River Caught Fire at Least a Dozen Times, but No One Cared Until 1969*, SMITHSONIAN MAG. (June 19, 2019), https://www.smithsonianmag.com/history/cuyahoga-river-caught-fire-least-dozen-times-no-one-cared-until-1969-180972444/.

[669] *See id.*

[670] *See* Herbert Koshetz, *Clothes that Catch Fire*, N.Y. TIMES, July 25, 1971, at F3, F3.

[671] President Nixon explained: "[A] defective lawnmower or electric heater can be just as dangerous to the consumer and his family as contaminated food or improperly packaged drugs." John D, Morris, *Nixon Approves Plan to Assure Product Safety*, N.Y. TIMES, Oct. 29, 1972, at 1, 1.

[672] *See* Will Kenton, *Office of Thrift Supervision: What It Is, How It Works*, INVESTOPEDIA, https://www.investopedia.com/terms/o/ots.asp (last updated Apr. 28, 2022) [https://perma.cc/D8MY-3C3P]; Kenneth J. Robinson, *Savings and Loan Crisis*, FED. RSRV. HIST. (Nov. 22, 2013), https://www.federalreservehistory.org/essays/savings-and-loan-crisis#turbulent [https://perma.cc/654X-Q6WG].

mass displacement of residents, numerous suicides, and the creation of the Consumer Financial Protection Bureau.[673] These innovations disrupted markets, public trust, and infrastructural and social safety; yet they were not progress. To realign the relationship between innovation and progress, robust trustworthiness backstops were needed, not only in self-governance but also in law through dedicated agencies.

Today, innovation in technology products and services has brought us exploit machina.[674] Self-regulatory efforts and existing agencies' enforcement during the last two decades have proven insufficient: innovation has again become disconnected from progress.[675] Exploit machina will only scale faster as we feed it more compute power[676] and as its creators continue to be (sometimes uncritically) afforded social cache as "wizards."[677] And the irreparable harms of Theranos-like

[673] *See* Jason N. Houle & Michael T. Light, *The Home Foreclosure Crisis and Rising Suicide Rates, 2005 to 2010*, AM. J. PUB. HEALTH (June 2014) https://pmc.ncbi.nlm.nih.gov/articles/PMC4062039/ ("The foreclosure crisis has likely contributed to increased suicides, independent of other economic factors associated with the recession."); *Building the CFPB*, CFPB, https://www.consumerfinance.gov/data-research/research-reports/building-the-cfpb/ (last updated Aug. 8, 2023) [https://perma.cc/F9TX-AQR2].

[674] *See, e.g.*, Press Release, CFPB, *supra* note 121 (A fintech startup "backed by some of the biggest names in Silicon Valley" that included "Google Ventures, Andreessen Horwitz, Kleiner Perkins, and other prominent venture capital firms," LendUp, was subject to multiple enforcement actions by the CFPB: in 2016 CFPB enforcement sought to stop LendUp from misrepresenting the benefits of borrowing to borrowers, in 2020 CFPB obtained a judgement against LendUp in connection with alleged violations of the Military Lending Act, and in 2021 CFPB settled a this third action alleging that LendUp that, among other things, employed a "LendUp Ladder" tool that falsely claimed borrowers would gain access to larger loans at lower rates and violated the CFPB 2016 order against misrepresentation of material borrowing terms and failing to provide timely and accurate adverse-action notices required by fair lending laws.).

[675] *See* Jordan Reynolds, *9 Reasons Digital Fraud Is on the Rise*, SECURITY (Nov. 12, 2020), https://www.securitymagazine.com/articles/93912-reasons-digital-fraud-is-on-the-rise [https://perma.cc/Y86Z-WEKA].

[676] *See* Scott Hermann, *The Evolving Threat: How AI Scams Are Targeting Your Identity*, FORBES (Aug. 9, 2024, 9:00 AM), https://www.forbes.com/councils/forbestechcouncil/2024/08/09/the-evolving-threat-how-ai-scams-are-targeting-your-identity/ [https://perma.cc/N8PK-ZYS2].

[677] Consider the recent collapse of Builder.ai — an AI startup launched by a "chief wizard" that allegedly falsified revenue streams and whose "AI" was allegedly a team of humans. David Braue, *The Company Whose 'AI' Was Actually 700 Humans in India*, INFO.

organizations will only escalate in their threat to human, economic, and social safety. We have been here before when earlier generations of innovation became disconnected from progress. We know how to fix this problem.

As set forth in Figure 3, compared to every other industry with comparable levels of destructive potential and propensity for death and other irreparable harms, technology products and services are significantly underregulated. But the problem runs even deeper: As untrustworthy technology products and services become incorporated into the daily operations of traditional, more vetted industries, they risk invisibly eroding existing baselines of safety, causing irreparable harm at scale to these traditional industries, our national security, and the public. They introduce new possible points of vulnerability and exploitability into otherwise well-managed organizations; they open the door to new exploit machina.[678] In other words, this technology safety

AGE (June 5, 2025, 1:15 PM), https://ia.acs.org.au/article/2025/the-company-whose-ai-was-actually-700-humans-in-india.html [https://perma.cc/ZH6Q-FVUY]; Alexandra Heal & Robert Smith, *Inside the Collapse of Microsoft-Backed UK Tech Unicorn Builder.ai*, FIN. TIMES (June 5, 2025), https://www.ft.com/content/67f0277a-10fd-463c-8946-af0be2b4028f [https://perma.cc/44ZZ-TVQD]; *How This Microsoft-Backed Billion-Dollar London Startup Made 700 Engineers Sitting in India Pose as an AI Tool*, TIMES INDIA, https://timesofindia.indiatimes.com/technology/tech-news/how-this-billion-dollar-london-startup-backed-by-microsoft-made-700-engineers-sitting-in-india-pose-as-ai/articleshow/121572659.cms (last updated June 6, 2025, 1:15 PM) [https://perma.cc/3M32-KEUJ]. Not all wizards are good wizards; even some wizards who start out as good wizards can evolve into destructive versions of their former selves. *See, e.g.*, William Fisher, *Between The Hobbit and The Lord of the Rings Why Did Saruman Turn Evil?*, COLLIDER, https://collider.com/lord-of-the-rings-saruman-evil/ (last updated Feb. 27, 2024, 6:00 AM).

[678] Consider the recent suit filed by Clorox against cybersecurity vendor Cognizant, alleging that lax governance practices in oversight of employee processes resulted in a corporate loss to Clorox of $380 million, disrupting Clorox's ability to ship core products. David Jones, *Clorox Files $380 Million Suit Blaming Cognizant for 2023 Cyberattack*, CYBERSECURITY DIVE (July 23, 2025), https://www.cybersecuritydive.com/news/clorox-380-million-suit-cognizant-cyberattack/753837/ [https://perma.cc/3VB7-GZNU]. Clorox products are commonly used in hospital settings, including the sterilization of operating rooms. CLOROX, OPERATING ROOM, CLEANING & DISINFECTING GUIDE 1 (2015), https://www.cloroxpro.ca/wp-content/uploads/2018/09/OR-Cleaning-and-Disinfection-Guidelines_OR-Protocol-Disinfection.pdf [https://perma.cc/BK3G-SKMN]. An unexpected interruption of critical supplies to hospitals could potentially disrupt

underregulation threatens to erode existing baselines of human, economic, and social safety currently maintained by traditional industries. Every industry now relies on technology products and services in some way. But not every industry is equally situated to anticipate and mitigate the harms arising from those technology products and services and exploit machina.

smooth hospital operations placing patients at greater risk for irreparable harm, as the CISA COVID-19 Task Force and other agencies have recently explained. *See, e.g.*, CYBERSECURITY & INFRASTRUCTURE SEC. AGENCY, *PROVIDE MEDICAL CARE* IS IN CRITICAL CONDITION: ANALYSIS AND STAKEHOLDER DECISION SUPPORT TO MINIMIZE FURTHER HARM (2021) (examining COVID-19 surges on hospital system operations, cascading effects on critical infrastructure sectors and National Critical Functions from various impacts on supply chain interruption, including where excess deaths occurred); Jill McKeon, *Cyberattacks Increase Mortality Rates, but Healthcare Is in Denial*, TECHTARGET (Jan. 13, 2022), https://www.techtarget.com/healthtechsecurity/news/366594690/Cyberattacks-Increase-Mortality-Rates-But-Healthcare-Is-In-Denial [https://perma.cc/KYD4-2534] (quoting Joshua Corman, chief strategist of the CISA COVID task force saying that "[i]n the last 12 to 18 months, we've had successful electronic attacks of the water we drink, the food we put on our table, and the oil and gas that fuels our cars and our homes. The timely availability of patient care, the schools our children go to, the municipalities who run our towns and our cities, and even federal agencies have been the victims of state-sponsored and criminal attacks").

Figure 3: Legal Governance by Product and Service Category

| PRODUCT or SERVICE | SOCIAL HARM at scale possible? | TARGETED ENFORCEMENT AGENCY? | ENFORCEMENT/ PRIVATE LITIGATION |
|---|---|---|---|
| Technology | Yes – death, economic, national security | No dedicated federal agency; some state activity, limited professional self-organization (SROs) | Sporadic federal/state criminal/civil enforcement by government + suits by private litigants |
| Financial | Yes – economic, national security | Multiple federal and state + licensing, SROs | Federal and state criminal/civil + suits |
| Goods [traditional] | Yes – death | Federal, state, local + some licensing, SROs | Federal and state criminal/civil + suits |
| Power | Yes – death, economic, national security | Multiple federal, state + licensing, SROs | Federal and state criminal/civil + suits |
| Transport | Yes – death, economic, national security | Multiple federal, state + licensing, SROs | Federal, state, local criminal/civil + suit |
| Environmental | Yes – death, economic, national security | Federal, state, local + some licensing, SROs | Federal, state, local criminal/civil + suits |
| Consumables/food and drugs | Yes – death, economic, national security | Multiple federal, state, local + licensing, SROs | Federal, state, local criminal/civil + suits |
| Medical devices; healthcare | Yes – death, economic, national security | Multiple federal, state + licensing, SROs | Federal criminal/civil + suits |
| Communications | Yes – death, economic, national security | Federal, state + licensing, SROs | Federal and state criminal/civil, suits |
| Insurance | Usually not – economic | Multiple federal partial + State + licensing, SROs | Federal and state criminal and civil + suit |
| Construction | Usually not - death | Multiple federal partial + State + licensing, SROs | State criminal sanction, civil + suits |
| Body/ aesthetic enhancement | Usually not – death | State + licensing, SROs | Criminal, civil + suits |
| Legal; other professional/ trade services | Usually not | State + licensing, SROs | Criminal, civil suits |
| Arts and related professions | Usually not | Primarily SROs | Criminal, civil suits |

The answer to this legal imbalance, where problems of exploit machina unfairly disadvantage traditional industries and threaten to erode them from within, is not a race to the bottom in safety across every industry. The answer is to ensure that technology products and services are lifted up to align with the safety levels already present in the

rest of our economy and society. As explained in other scholarship,[679] engineering safety history reveals a recurring three-part response to emergent irreparable safety harms and engineering disasters: (1) empowering courts to redress claims of harm relying on a basis arising under common law;[680] (2) creating a new specialized governance entity in law with subject area experts, regulators, and enforcers to ensure corrections of safety deficits (and then defending the improved status quo of safety);[681] and (3) creating self-regulatory and professional organizations engaging in ongoing peer review.[682]

In other words, exploit machina concerns are now signaling to us that a key part of the governance response involves the creation of a new technology safety regulator. Building on ideas from prior legal scholarship,[683] the sections that follow propose a model for a new

[679] Andrea M. Matwyshyn, *Homicideware* (forthcoming) (on file with author).

[680] *See* SEC v. Jarkesy, 603 U.S. 109, 135 (2024); *see also* Loper Bright Enters. v. Raimondo, 603 U.S. 369, 369 (2024); Axon Enter., Inc. v. FTC, 598 U.S. 175, 175 (2023).

[681] For a discussion of the role of authorities of agencies in connection with "restoring the status quo," see *Jarkesy*, 603 U.S. at 111.

[682] *See supra* Part I.

[683] Professor Nicolas Terry argues in the context of healthcare AI and robotics that if "a super-regulator ends up being favored, then, as with the case of data protection, the preferred solution would be to have a single AI regulatory agency, not a healthcare-specific one," pointing out that "[u]se of a single regulatory agency would help to avoid regulatory exceptionalism, indeterminacy, or arbitrage." Nicolas Terry, *Of Regulating Healthcare AI and Robots*, 21 YALE J.L. & TECH. 133, 173 (2019); *see also, e.g.*, Matthew U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 HARV. J.L. & TECH. 353, 394 (2016) (arguing in favor of establishing "an agency . . . responsible for certifying AI programs as safe and set the limits of the Agency's power to intervene in AI research and development"); Ryan Calo, *The Case for a Federal Robotics Commission*, BROOKINGS (Sept. 15, 2014), https://www.brookings.edu/articles/the-case-for-a-federal-robotics-commission/ [https://perma.cc/UEK3-YYEZ] (arguing in favor of a "small" Federal Robotics Commission limited to embodied machines that "do not just sense, process and relay data" but are "organized to act upon the world physically"). Relatedly, Professor James Grimmelmann has argued in favor of gleaning lessons from product safety for privacy but does not call for the direct application of products liability law to online privacy. James Grimmelmann, *Privacy as Product Safety*, 19 WIDENER L.J. 793, 813, 826 (2010) ("Thus, I would like to suggest that some of the lessons the law has learned in dealing with product safety could usefully be applied to the analogous problem of privacy safety"). Professor Michael Froomkin and a coauthor argue "that in many cases privacy is safety, and that, in practice, United States law already recognizes this fact." *See* A.

technology safety regulator of last resort, BoTS — an independent coordinating and gap-filling enforcement agency, whose mission is to protect against exploit machina. The design of BoTS is self-consciously novel, but it builds on models present in other existing government organizations that have proven effective. Specifically, BoTS would longitudinally track and enhance the safety of technology products, services, and practices in a technology-neutral and sector-neutral manner. In other words, the goal is to supplement and not supplant existing agencies' and states' authority and efforts, filling emerging gaps as technology safety requires. In this way, human and national safety and prevention of irreparable technology harms sit as the lodestar organizing principle of BoTS — not a particular technology of the moment.

The BoTS model creates a nimble agency framework that can scale and evolve alongside new technologies, while remaining sensitive to constitutional values, in particular those of the First Amendment.[684] Specifically, the legal approach proposed in the sections that follow treats human persons and (both nonprofit and for-profit) corporate persons in parity for First Amendment purposes.[685] As a consequence, First Amendment obstacles are few when a technology safety legal framework is crafted in a content-neutral and identity-neutral manner,

Michael Froomkin & Zak Colangelo, *Privacy as Safety*, 95 WASH. L. REV. 141 (2020). In subsequent work, Froomkin and coauthors argue in favor of construing FTC, FDA, NHTSA, and OSHA authority in an expansive manner to regulate privacy as a matter of safety. *See* A. Michael Froomkin, Phillip J. Arencibia & P. Zak Colangelo-Trenner, *Safety As Privacy*, 64 ARIZ. L. REV. 921, 922 (2022) ("[W]e offer reasonable legal constructions of certain extant federal statutes that would justify more extensive privacy regulation in the name of providing enhanced safety, a regime that we argue would be a substantial improvement over the status quo yet not require any new legislation, just a better understanding of certain agencies' current powers and authorities.").

[684] Thus, this model engages expressly with the Arendtian shibboleth of "thinking what we are doing" to preserve the breathing room needed for expression, democratic debate, and imagination.

[685] Statutes reflecting a lack of identity neutrality, meaning failing to treat human and nonhuman persons in parity for free speech purposes under the law, have been successfully challenged on First Amendment grounds in the past; for example, in *Sorrell v. IMS Health*, where the Supreme Court invalidated Vermont's pharmaceutical marketing privacy statute on speaker identity neutrality grounds. Sorrell v. IMS Health Inc., 564 U.S. 552, 564 (2011).

regulating conduct that reflects choices made with knowledge or intent.[686]

1. Structure

BoTS should be comprised of three divisions:

(a) an enforcement division with robust civil and criminal fining, referral, and injunctive authority where technology safety issues place the public at risk;[687]

(b) a policy coordination and technology futures tracking, modeling, research service, and rulemaking division (subject to the Administrative Procedure Act[688]) that (i) engages across the government to monitor enforcement trends and emerging technology safety issues in broader national, international, and historical context and (ii) functions as a computer security and technology safety whistleblower office of last resort; and

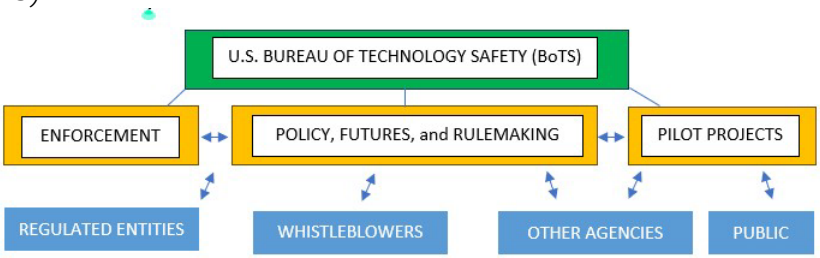
(c) a research and "pilot projects" division that both (i) acts as a hub of in-house technical expertise for BoTS and for other agencies needing investigatory assistance in connection with their own technology enforcement and (ii) launches such experimental initiatives as technology safety may require in collaboration with states and the public. Pilot project initiatives, if deemed successful, would then move to the policy coordination and technology futures group for launch through APA rulemaking, as appropriate.

[686] *See* Andrea M. Matwyshyn, *Privacy, the Hacker Way*, *supra* note 157, at 13-14.

[687] Fines would arise from violations of rules promulgated in accordance with the APA to ensure clarity and deviations from subsequent explanatory guidance. This authority mirror's FDA's authority in part, recognizing constraints from recent Supreme Court precedent. *See infra* notes 688–691.

[688] Administrative Procedure Act, 5 U.S.C. §§ 551–559 (2006).

Figure 4: Proposed Structure for the Bureau of Technology Safety (BoTS)




BoTS's mission would focus on (a) protecting the public from unsafe and untrustworthy technology products, services, and practices through enforcement, advocacy, research, and education; (b) promoting and maintaining a fair and resilient technology marketplace through sound regulation; and (c) facilitating development of a trustworthy technology innovation ecosystem in the national and public safety interest and advancing progress in science and useful arts.[689]

Connecting with the language of recent Supreme Court caselaw on point, although some BoTS enforcement may relate to traditional legal harms such as fraud that have corollaries in common law, BoTS is by design a "self-consciously novel"[690] agency in large portions of its

[689] Specifically, BoTS's proposed mission and authorities are modeled on those of the FTC, CFPB, SEC, CFTC, FDA, CPSC, OSHA, and USDA, translating them into technology safety. Its work should also be informed by lessons from the FDIC and FinCEN history — successful models of cross-cutting coordinating agencies, where financial criminality was a key motivator for insider and external attacks. In other words, CISA expansion would not accomplish the purposes of this proposal, and many of the relevant technologies and entities are not necessarily part of critical infrastructure. Similarly, the Department of Homeland Security is not the correct home for BoTS: BoTS requires an independent team that includes veteran enforcers with financial fraud, First Amendment, international consumer protection, and other expertise that goes beyond infrastructure and national security concerns. BoTS's authority also should not preempt state enforcement authorities, and its authority as scoped eliminates the need for any safe harbor carve-outs in its enabling statute. BoTS's core role as an interagency and public coordinator, translator, and backstop would enhance market trustworthiness and public safety in harmonized ways — ways that can effectively scale to address evolving public technology safety challenges.

[690] SEC v. Jarkesy, 603 U.S. 109, 137 (2024).

work.[691] Congress can borrow the model of (1) technical safety oversight and enforcement structures of agencies such as OSHA[692] and other safety agencies,[693] merging them with (2) the director structure of financial safety regulators such as CFPB,[694] and (3) the funding mechanism employed for financial product safety regulatory agencies[695]

[691] Matwyshyn, *It's Morning Again in Pennsylvania*, *supra* note 35. As described in detail elsewhere, BoTS would be formed under a single Director and comprised of three divisions — and enforcement division, a policy and rulemaking division, and a pilot project division. *Id.*

[692] *See* Zach Schonfeld & Ella Lee, *Supreme Court Turns Away OSHA Challenge over Opposition from Thomas, Gorsuch*, THE HILL: CT. BATTLES (July 2, 2024, 10:37 AM), https://thehill.com/regulation/court-battles/4751583-supreme-court-osha-thomas-gorsuch/ [https://perma.cc/P7VR-5RV6]. During the 2024 term, the Supreme Court denied a petition for a writ of certiorari to a challenge to the constitutionality of OSHA. *Id.*

[693] These authorities should include the ability to seek rulemaking, injunctive relief, civil and criminal penalties, and subpoena authority. For one possible partial model, see, for example, Housing and Economic Recovery Act of 2008, Pub. L. No. 110-289, 122 Stat. 2654 (2008). With respect to any administrative adjudication structure, the most likely model to survive Supreme Court scrutiny is one modeled on OSHA, which relies upon an administrative law judge adjudicatory structure housed in a separate agency, OSHRC, created at the same time in the OSH Act. *See* Occupational Safety and Health Act of 1970, Pub. L. No. 91-596, 84 Stat. 1590 (1970); *see also How OSHRC Works*, OCCUPATIONAL SAFETY & HEALTH REV. COMM'N, https://www.oshrc.gov/about/how-oshrc-works/ (last visited Sept. 12, 2025) [https://perma.cc/J4ZW-J8NG]. OSHA also maintains a robust whistleblower program worthy of consideration as a potential partial model. *OSHA Online Whistleblower Complaint Form*, OSHA, https://www.osha.gov/whistleblower/WBComplaint (last visited Sept. 12, 2025) [https://perma.cc/GB4P-S3XP].

[694] *CFPB Structure*, CFPB (Feb. 3, 2025), https://www.consumerfinance.gov/about-us/the-bureau/bureau-structure/ [https://perma.cc/3M7Y-UKZ4]. CFPB's director structure and congressional funding mechanism has been affirmed recently by the Supreme Court in *Seila Law LLC v. CFPB*, 591 US 197 (2020) and *CFPB v. Cmty. Fin. Servs. Ass'n, Ltd*, 601 U.S. 416 (2024).

[695] *See* Housing and Economic Recovery Act of 2008, Pub. L. No. 110-289, 122 Stat. 2654, https://www.govinfo.gov/content/pkg/PLAW-110publ289/html/PLAW-110publ289.htm [https://perma.cc/URR2-MYGB].

such as FDIC[696] and FHFA.[697] In other words, funding the ongoing costs of BoTS can be structured in a manner not borne by taxpayers.[698]

BoTS-regulated entities should include three categories of legal persons, including nonprofit organizations: (a) persons whose chosen business model or operational structure present technology safety or computer security risks to the public and who are of sufficient size to map to a slightly modified version of the lowest "size of person" requirements for Hart-Scott-Rodino reporting,[699] regardless of whether the legal person in question is for-profit (including privately-held entities) or nonprofit; (b) persons who are government contractors or who otherwise hold a federal government contract; and (c) persons who offer products or services distributed or marketed to the public in more than one state or sold to the public in interstate commerce and whose products and services present an immediate technology safety or computer security risk to the public.

In other words, BoTS would expedite, align, and coordinate policy response to exploit machina and the most complex public safety harms caused by and through technologies. Particularly as the escalating pace

---

[696] *See* Bill Chappell, *The FDIC Was Created Exactly for This Kind of Crisis. Here's the History*, NPR (Mar. 14, 2023, 8:05 AM), https://www.npr.org/2023/03/13/1163138002/the-fdic-insurance-limit-was-last-raised-in-2008-heres-how-it-works [https://perma.cc/92QZ-LMZY].

[697] *See Federal Home Loan Bank System*, FED. HOUS. FIN. AGENCY, https://www.fhfa.gov/supervision/federal-home-loan-bank-system (last visited Aug. 20, 2025) [https://perma.cc/BCJ8-4S7X].

[698] Such a structure could be Congressionally created with only a one-time appropriation for startup.

[699] This standard would borrow the lowest jurisdictional threshold of (a portion of the) the Clayton Act Section 7A(2)(B) size of person test for entities engaged in manufacturing (regardless of whether the entities in question are in fact manufacturing entities), which currently stands at $26.8m or more in total assets or annual sales in 2026. *See* Revised Jurisdictional Thresholds for Section 7A of the Clayton Act, 91 Fed. Reg. 2133, 2133 (Jan. 16, 2026). However, the value of data assets and the value of user databases must be included in any size of person calculation. This may require adjustments to both accounting practices and the NAICS codes to achieve a more accurate calculation of size/assets that is more in line with private sector valuation practices. There is currently an undesirable disconnect between HSR valuation and private sector valuation of technology assets. *See* Andrea M. Matwyshyn, iTrust Antitrust (unpublished manuscript) (draft on file with author).

of AI-facilitated attacks[700] on the general public becomes more socially unsettling,[701] existing agencies will be unable to successfully scale and coordinate their technology safety efforts. BoTS would also be ideally positioned to resolve and align legal terminology collisions,[702] unifying efforts of agencies with different missions and authorities, varying levels of clearances and visibility into national security, and differing expertise. Indeed, some enforcers who are already doing important technology safety work do not currently view it as such; through their eyes, technology safety is not their intent and arises only as a byproduct of their traditional statutory mandates unrelated to technology.[703] However, a supportive, coordinating, and gap-filling agency would recognize these efforts for their technology safety importance, placing them in a broader technology safety context and translating public messaging through a unified lens of technology safety, both inside and outside government.

2. Scalability

No single agency is currently tasked with defending against exploit machina in its various forms; in particular, no single agency focuses on

[700] *See* Jai Vijayan, *AI-Enabled Voice Cloning Anchors Deepfaked Kidnapping*, DARKREADING (June 29, 2023), https://www.darkreading.com/cyberattacks-data-breaches/ai-enabled-voice-cloning-deepfaked-kidnapping [https://perma.cc/ULD3-6895].

[701] *See, e.g.*, Andrea Blanco, *A Father Is Warning Others About a New AI 'Family Emergency Scam,'* INDEPENDENT (Dec. 6, 2023, 2:28 PM), https://www.independent.co.uk/news/world/americas/ai-phone-scam-voice-call-b2459449.html [https://perma.cc/K4HV-PPDR]. (explaining an "elaborate scheme involves scammers using artificial intelligence to clone a person's voice, which is then used to trick loved ones into sending money to cover a supposed emergency").

[702] *Collision*, NAT'L INST. OF STANDARDS & TECH., https://csrc.nist.gov/glossary/term/collision (last visited Aug. 20, 2025) [https://perma.cc/LEY8-345W].

[703] *See, e.g.*, *Volkswagen Clean Air Act Civil Settlement*, EPA (Sept. 19, 2024), https://www.epa.gov/enforcement/volkswagen-clean-air-act-civil-settlement [http://perma.cc/ETH3-3QRS] ("These settlements resolve allegations that Volkswagen violated the Clean Air Act ('CAA') by the sale of approximately 590,000 model year 2009 to 2016 diesel motor vehicles equipped with 'defeat devices' in the form of computer software designed to cheat on federal emissions tests. The major excess pollutant at issue in this case is oxides of nitrogen (NOx) and is a serious health concern.").

(correcting and) maintaining baselines of technology safety at the intersection of interstate commerce and national security. Those agencies that have some existing authority over technology safety increasingly struggle to employ it and face gaps in their authority.[704] Particularly in light of recent Supreme Court precedent in *Loper Bright Enterprises v. Raimondo*,[705] *Axon Enterprise, Inc. v. FTC*,[706] *Seila Law LLC v. CFPB*,[707] *SEC v. Jarkesy*,[708] and other cases, technology safety efforts by most existing agencies are likely to face years of legal challenges.[709] Because of its design, BoTS would catch any technology safety authority gaps that arise from this evolving Supreme Court landscape.

Throughout engineering history, companies have sought to delay safety regulation by claiming that novel "innovation" necessitates their sometimes legally problematic conduct and lack of robust internal controls. Using the language of "innovation," they have historically claimed that this alleged interest (in building as they wish, irrespective

[704] *See, e.g.*, Cal. Dental Ass'n v. FTC, 526 U.S. 756 (1999) (clarifying that the FTC's authority over nonprofit and membership organizations is not equivalent to its authority over for profit entities, stating that "[t]he Act does not cover all membership organizations of profit-making corporations without more").

[705] *See generally* Loper Bright Enters. v. Raimondo, 603 U.S. 369, 369 (2024) (holding that the Administrative Procedure Act requires courts to exercise their independent judgment in deciding whether an agency has acted within its statutory authority, courts may not defer to an agency interpretation of the law simply because a statute is ambiguous, and overruling *Chevron*).

[706] *See generally* Axon Enter., Inc. v. FTC, 598 U.S. 175, 175 (2023) (holding that the statutory review schemes set out in the Securities Exchange Act and Federal Trade Commission Act do not displace a district court's federal question jurisdiction over claims challenging as unconstitutional the structure or existence of the SEC or FTC).

[707] *See generally* Seila L. LLC v. CFPB, 591 U.S. 197, 197 (2020) (setting aside CFPB's five-year term for a Director with removal only for inefficiency, neglect, or malfeasance as violating separation of powers, finding that "the CFPB Director is a principal officer whose duties are far from limited").

[708] *See generally* SEC v. Jarkesy, 603 U.S. 109, 109 (2024) (holding that when the SEC seeks civil penalties against a defendant for securities fraud, the Seventh Amendment entitles the defendant to a jury trial).

[709] For example, the Federal Communications Commission recently faced a legal setback in the D.C. Circuit Court of Appeals in connection with implementation of the Secure Equipment Act. *See* Hikvision USA, Inc. v. FCC, 97 F.4th 938, 949 (D.C. Cir. 2024) (holding that the FCC's definition of "critical infrastructure" in connection with its implementation of the Secure Communication Act is overly broad).

of external harm) overrides society's interest in avoiding irreparable harms to public safety, interstate commerce, and national security.[710] Today, it is technology industries sometimes channeling similar "innovation" arguments, acting to the detriment of other industries and to the detriment of the progress interests of the public, as such interests were articulated by the Founders and Framers. Progress does not arrive cloaked as exploit machina.

## CONCLUSION

> Morrow: "everything changes, everything evolves. . . . people are living inside the arcade cabinet we built . . . . Invention comes with responsibilities you didn't ask for. If you make something people want or need then it's up to you to set the limits. You have to make some rules."
>
> Halliday: "I don't wanna make any more rules. I'm a dreamer. I build worlds."
>
> Morrow: "We created something beautiful, Jim, but it's changed. Okay? It's really not a game anymore."[711]

The film *Ready Player One* introduces a dystopian version of Columbus, Ohio[712] in the 2030s,[713] a world perhaps uncomfortably similar to our own in some respects. Many people's primary joy arises

---

[710] For example, car manufacturers raised innovation stifling arguments at the time of the passage of the National Traffic and Motor Vehicle Safety Act. *See* RICHARD F. WEINGROFF, FED. HIGHWAY ADMIN., PRESIDENT DWIGHT D. EISENHOWER AND THE FEDERAL ROLE IN HIGHWAY SAFETY 168-69 (2003).

[711] READY PLAYER ONE, at 24:32 (Warner Bros. Pictures 2018).

[712] *See* Chelsea Wiley, *Steven Spielberg's "Ready Player One" Set in Columbus*, COLUMBUS NAVIGATOR, (July 24, 2017), https://www.columbusnavigator.com/ready-player-one-columbus/ [https://perma.cc/M6YM-HLVF].

[713] *See* READY PLAYER ONE, *supra* note 711, at 2:20 ("I was born in 2027. After the Corn Syrup Droughts, after the Bandwidth Riots. After people stopped trying to fix problems and just tried to outlive them.").

from virtual reality[714] escapism through IoB devices.[715] In particular, they play a metaverse game known as *The Oasis*. *The Oasis* was the brainchild of a charismatic but self-destructive and egomaniacal creator named Halliday, a billionaire who inspired hero worship among players.[716] Part of the popularity of *The Oasis* stems from a chance at upward social mobility through a competition hunting for in-game artifacts.[717] Leveraging this player hope, " aftermarket" product companies sprang to life to equip players, as did a set of corporations who play for commercial control of *The Oasis*.[718] They sell and lease IoB equipment that is maximally concerned with revenue,[719] even at the expense of player safety.[720] Nevertheless, many players enter into abusive lending relationships in order to afford these body upgrades.[721] Per the terms of predatory lenders' EULAs,[722] when players default on payments, lenders are legally permitted to imprison them in "loyalty centers," where players work off their debts[723] by playing *The Oasis* for the benefit of their creditors.[724] Losing sight of the distinction between

[714] *See id.* at 3:09 ("[W]e could go somewhere without going anywhere at all. You don't need a destination when you're running on an omnidirectional treadmill . . . he gave us a place to go. A place called the Oasis.").

[715] Although much of the population lives in impoverished areas of stacked trailers ravaged by environmental devastation, they find a form of freedom of identity and imagination through virtual reality headsets and gear. *See id.* at 2:57 ("These days, reality is a bummer. Everyone is looking for a way to escape.").

[716] *See id.* at 7:13 ("The Oasis was the brainchild of James Halliday. He and his partner, Ogden Morrow, released the first Oasis build in 2025 . . . . Halliday? He wasn't just the owner of the world's richest company. He was like a god. People loved him. They worshipped him as much as his creation.").

[717] *See id.* at 6:02.

[718] *See id.* at 10:43.

[719] *See id.* at 28:57 (F'nale: "We're opening five new loyalty centers this month." Innovative Online Industries CEO: "Debt Services dwarfs hardware.").

[720] *See id.* at 29:14 ("We call this Pure O2. This is the first of our planned upgrades once we can roll back some of Halliday's ad restrictions. We estimate we can sell up to 80% of an individual's visual field before inducing seizures.").

[721] Higher quality equipment assists with success in the game. *See id.* at 21:00.

[722] As people who default on their payments are led away to the loyalty centers to do their time, security guards recite an End User License Agreement at them that is vaguely reminiscent of a Miranda warning. *Id.* at 1:15:20.

[723] *See id.* at 1:19:00.

[724] *Id.* at 51:00.

the virtual and the real, some players become so deeply emotionally and financially invested in the game[725] that they engage not only in financially self-destructive behaviors but also in self-harming ones,[726] including suicide.[727] Ultimately, the CEO of the leading predatory lender loses the battle for the future of *The Oasis*[728] and social control[729] due to his own weak information security practices in password management.[730] But, the society's technology-driven unsustainability — perhaps much like our own — remains unresolved.[731]

In February 2025, an AI eyewear company named Halliday raised almost $3 million on Kickstarter to fund its connected product. As presented in its marketing materials, the glasses offer a lightweight proactive AI agent with a discreet display, all-day battery, AI translation, "cheatsheet," and audio memo capabilities along with ring control inside retro eyewear.[732] Meanwhile, a month prior, a panel of hackers, lawyers and law students presented a hypothetical arbitration case at the final[733] ShmooCon computer security conference in Washington D.C.[734] The hypothetical involved a joint venture between a traditional

[725] *Id.* at 3:09, 7:00.

[726] Players begin to voluntarily buy equipment that causes them to experience physical pain in the real world when they experience "pain" in game. *Id.* at 1:19 (Announcer: "Get ready for the feel, the real of real. X1. No pain, no gain.").

[727] Players who lose their in-game assets also show signs of psychological trauma and begin to self-harm in the real world. *Id.* at 7:00 ("Since most people spend so much time in the Oasis, losing your shit, means losing your shit.").

[728] *Id.* at 1:29:10 (Wade Watts: "I'm here talking to all of you now because your future's being threatened.").

[729] *Id.* at 28:36 (IOI CEO: "Well here's a better question: Who cares? It's nothing less than a war for control of the future.").

[730] *Id.* at 30:52. (the CEO posts his login and password on a note attached to his chair).

[731] *Id.* at 2:20:00.

[732] *Halliday: #1st Proactive AI Glasses with Invisible Display*, KICKSTARTER, https://www.kickstarter.com/projects/halliday-ai-glasses/halliday-proactive-ai-glasses-with-invisible-display (last updated Aug. 28, 2025) [https://perma.cc/DZP5-N8GZ].

[733] *See* Cynthia Brumfield, *ShmooCon to Take Its Final Bow in 2025*, CSO ONLINE (Jan. 16, 2024), https://www.csoonline.com/article/1291036/shmoocon-to-take-its-final-bow-in-2025.html [https://perma.cc/T4A3-MYHV].

[734] StrongWind, *ShmooCon 2025 Day 3 Bring it On! Track*, YOUTUBE, at 15:15, (Jan. 13, 2025), https://www.youtube.com/watch?v=tPbbMPjFKjA&t=465s [https://perma.cc/DC5R-T2FG].

eyewear company and an AI eyewear startup, whose AI eyewear product, CyberIz, was the target of a ransomware attack. As a result of the ransomware, numerous people wearing the CyberIz products died or caused death of others when their vision became unexpectedly impaired. Multiple class actions and other suits were immediately filed by families of the deceased. At issue in the mock arbitration was the apportionment of liability among the two business partners, companies who had both knowingly chosen not to patch the multiple known exploitable vulnerabilities that, entirely predictably and avoidably, had been leveraged by the ransomware. But regardless of the outcome of the arbitration, irreparable harms had already occurred at scale because of the companies' choice to engage in exploit machina.

The path that we are on — the path of exploit machina — ends badly. Technology history and Arendt warn us that exploit machina scenarios, — situations where broken technologies and broken governance combine — lead to irreparable harm at scale. As explained succinctly by one AI founder, "the approach to building technology which is embodied via 'move fast and break things' is exactly what we should *not* be doing because you can't afford to break things and then fix them afterwards."[735] We trust ever-more sensitive aspects of our physical, financial and social safety to technologies; yet, we have not reckoned meaningfully with the social, economic, and individual safety harms that are escalating before our (cyber)eyes. This Article has argued that the time is overdue for a new federal technology safety regulator of last resort, BoTS.[736]

Tick, tock.

[735] Sir Demis Hassabis explained, "My view is that the approach to building technology which is embodied via 'move fast and break things' is exactly what we should *not* be doing because you can't afford to break things and then fix them afterwards." THE THINKING GAME, *supra* note 62, at 37:48.

[736] Much like the time value of money, social resilience against exploit machina also has time value.

> "[T]he modern age . . . may end in the deadliest, most sterile passivity history has ever known"
>
> — Hannah Arendt[737]

Bugs: "Something is happening here. Something important. . . . We know what happens next. . . . We know this story. This is how it all began."[738]

[737] ARENDT, THE HUMAN CONDITION, *supra* note 259, at 322.

[738] THE MATRIX RESURRECTIONS, at 3:03 (Warner Bros. Pictures 2021).