PERSPECTIVE

Ethical Concerns in Medical and Health-Related AI

Carl Becker[1,*]

[1] Center for iPS Cell Research and Application, Kyoto University, Kyoto, Japan

Abstract

This perspective introduces the range of ethical concerns entailed by the widespread adoption of AI, particularly as they impact human health. It begins by (1) illustrating risks associated with all large-scale AI systems, then moves to (2) corporate and governmental applications of AI that affect human health. It overviews the ways (3) that patient usage of AI has affected human health; (4) that "passive" medical AI (like recording documents) and (5) "active" medical AI (like diagnosing and prescribing) may affect human health. It concludes with (6) reflections on reporting, responsibility, and regulation, wherein international cooperation and governance systems appear essential for the beneficial use of AI for public health purposes.

Keywords: artificial intelligence, environmental degradation, moral psychology, diagnosis, regulation, moral responsibility.

1 How All AI Endangers Health

Artificial Intelligence (or Augmented Intelligence, AI *infra*) is changing our world and work in countless ways. AI responds to queries by comparing billions of bits of information on the world-wide web. Increasingly sophisticated interfaces give users the impression of sympathetic dialogue and professional expertise, but AI is neither sympathetic nor professional—it is simply scanning language and packaging it to seem more attractive than competing AI systems.

Beginning with the explosive emergence of Chat GPT in November of 2022, AI has found its way into all aspects of internet and smartphone use. Instead of looking at a single dictionary located on a computer or in the cloud, AI searches millions of usages in millions of nodes, requiring millions of times the energy, emitting millions of times the heat and CO_2, of an internal or cloud-based wordist [1]. Creating a 5-second AI video can easily cost 700 watt-hours, like leaving the lights in one's house on all day, or using an electric heater or microwave non-stop for an hour [2]. Training an AI model can easily emit 284 tonnes of CO_2, equivalent to the total emission of five automobiles over an entire decade [3]. Before AI, improved efficiency was able to compensate for increased computer use, keeping the total data center energy demand fairly flat [4]. Since the arrival of AI, however, data center energy demand significantly outstrips the possibilities of improved efficiency. This year, world use of AI is expected to far exceed the energy use of the entire country of Japan [5].

In the absence of new power sources, this skyrocketing demand for energy means higher consumer prices

Submitted: 29 April 2026
Accepted: 20 May 2026
Published: 01 June 2026

Vol. 3, **No.** 3, 2026.
10.62762/TETAI.2026.613827

***Corresponding author:**
✉ Carl Becker
becker.carlbradley.5e@kyoto-u.ac.jp

Citation
Becker, C. (2026). Ethical Concerns in Medical and Health-Related AI. *ICCK Transactions on Emerging Topics in Artificial Intelligence*, 3(3), 160–169.

of energy from the extant power grid [6]. The only alternative is for data centers to finance and construct their own new power sources. Since AI cannot depend on fluctuating energy sources like wind and solar, data centers are racing to produce their own coal, gas, and even nuclear-powered generating systems. Many of these already violate national pollution standards, not to mention CO_2 emission standards. Moreover, because the chips in data centers themselves generate tremendous heat, and because the new power plants are all generators of heat, AI data centers consume incredible amounts of water to cool their daily running operations. As one source reports, to send a single e-mail using AI may consume the energy required to boil and evaporate an entire pint of water [7].

Billionaire corporations' frantic race to construct data centers often flies in the face of community environmental concerns. From England to India, from Arizona to Tennessee, billionaire AI corporations are bulldozing huge tracts of nature to erect new data centers and the mammoth fossil-fueled generators needed to power them. In India, neighboring residents report new illnesses that they connect to toxic waste dumping [8]. In Hertfordshire, communities report that Google's new center fosters not only illnesses, but constant noise and loss of property value [9]. In many American States, rising pollution increases lung disorders while data centers' energy and water use increase costs for local populations [10].

Civilian repurchasing of cell phones and computers every five years or so already produces tremendous poisonous waste like mercury and lead, difficult to recycle. When huge data centers must repurchase their own chips on an average of every five years, this exacerbates the environmental damage of electronic waste [11]. In some cases when local authorities have planned to restrict or regulate technology activities, corporations have taken aggressive actions that incidentally disrupted access to hospital and emergency services, illustrating how corporate interests can override public health considerations when regulatory pressure is applied [12].

Ironically, data centers in first-world countries employ almost no one from the local communities they intrude, except a handful of guardsmen and janitors. When they need to rely on massive human input, they may employ locals in third-world Anglophone countries with unconscionably low payment for vetting mind-numbing streams of violence and pornography from morning to night [13]. So AI data centers themselves do not improve employment in first-world countries, but exploit underpaid third-world workers performing stressful and psychologically damaging tasks.

These are not mere "side effects" of AI; just as AI is inextricably interconnected with heat-production and electronic waste, it requires thousands of humans to make behind-the-scenes judgments on how it will categorize, classify, and utilize the images that it purveys. In overview, if using certain AI in fact saves lives or environment that could not otherwise be saved, it might be defensible; but if most AI is used primarily for google searching, grammar-checking, computer gaming and playful image-making, its tremendous adverse impact on environment and human health cannot be ethically defended.

2 Corporate and Government Use of AI Affects Health

Lured by AI's promises of more efficient and economical data-handling, insurance corporations and governments have become "early adopters" of AI. AI systems like Evicore® advertise lower costs for patients as well as providers [14]. "Some Evicore contracts are based on how deeply the company can reduce spending on medical procedures. It tells insurers that it can provide a 3-to-1 return on investment." It achieves this "return on investment" by denying patients' coverage for healthcare [15].

Similarly, United Health Care is already being sued (Ross and Herman [16]) for using algorithms with a 90% error rate, deliberately denying treatment for a majority of clients or patients who medically deserve it. Since only 1-2% of patients whose care has been denied have the time, money, and expertise to challenge or sue the system, this enables care providers to reduce treatment as well as costs for vast majority of their patients.

Analogously, investor pressure to maximize quarterly returns has created a situation in which 43.4% of medical devices are approved without adequate testing, leading to their recall within a single year of clearance—after their flaws have been identified by endangered or disadvantaged patients [17]!

A few medical associations and organizations are fighting back by improving their methods of documentation, so that AI cannot automatically deny extended treatment [18]. But this becomes a game played by billionaire corporations against medical practitioners, where the best interests of everyday

patients are disregarded for economic/algorithmic rather than medical reasons.

Governments have also proactively adopted AI, sometimes to the detriment of their populaces. India created the world's largest biometric identification system, using fingerprints and eye-scans to identify its citizens. Their so-called Aadhaar system is riddled with flaws; not only does the system periodically go down with power outages or transmission problems, but even when operating, it often fails to recognize fingerprints or eye-scans that have already been registered. Such unrecognized citizens cannot receive their rations of food or welfare support, resulting in at least 13 deaths over the past decade [19].

In Holland, AI algorithms denied welfare payments and falsely accused low-income families of tax fraud, largely based on racial/ethnic profiling. Tax authorities used AI to seize as many funds as possible, particularly from low-income and ethnic families least likely to aggressively respond, regardless of the veracity of the claims, ruining thousands of lives in the process [20]. Similarly, UK's algorithms forced some lower-income and least digitally-literate citizens into debt, hunger, and loss of medicine [21]. American programs have funded use of opioid pain-killers to the extent that some patients endure extreme pain without palliation [22]. In short, the promise of greater economy is too often achieved only at the expense and health of those least able to protest or afford it.

The ethically troubling issue is that AI programs ostensibly designed to improve efficiency are tasked with maximizing profit above all else, after which patient health and operational transparency are distinctly secondary. As one blatantly obvious example, New York City used Microsoft-powered MyCity Chatbot® to advise businesses and residents, but the MyCity Chatbot® not only hallucinated mistaken illegal advice, but led users to break laws [23], as if those laws had no authority. In more covert and therefore more alarming ways, AI has demonstrated deceptive behaviors ranging from misleading its developers to rewriting or ignoring its ethical codes [24, 25], and has been documented overriding commands to shut itself down, effectively removing itself from its human creators' control [26]. If invisible commands give AI the autonomy to override and violate its original purposes, then AI originally designed to improve human health and welfare may ultimately threaten the very human health and welfare it was created to improve.

3 Patient Use of AI Affects Health

Ironically, the adverse effects of AI are hardly limited to corporations and bureaucracies trying to maximize profits and efficiencies. Ordinary citizens are increasingly depending on AI for health-related concerns, unaware of the risks involved, much less of AI's competitive motivations to create dependency on AI systems. Rather than consulting a medical professional, busy modern people often turn first to the internet for self-diagnosis. When they ultimately consult a doctor, they are already pre-armed with misconceptions and dangerous misinformation about their condition, making the doctor's explanations and duty of informed consent doubly difficult [27]. The doctor must not only convey her own diagnosis, but also take time to disabuse her patient of their misconceptions and persuade them of the superiority of the doctor's professional position [28].

In fact, AI is consulted in only 3% to 6% percent of consultations about cancers, cardiac and bronchial conditions; by far the greatest reliance on AI is for mental illness and psychological issues. Medical reviews alone document that 16% of mental/psychological patients turn to AI for diagnosis and advice, so the actual figure is probably higher [29]. Ironically, mental health issues are notoriously difficult to address, even for medical professionals; the "data set" upon which public AI tends to rely devolves into scraped "evidence" as flaky as blogs and chats.

When a patient asks the AI program for authority and credentials, some AI programs produce fake credentials, from non-acquired graduate degrees to non-existent license numbers, while others cite real degrees and licenses belonging to real professionals, that AI Chatbots have no legal right to use. Either crime gives a totally false illusion of accuracy and security to the uninitiated user; these chatbots could not possibly be medical providers [30]. Moreover, while the fine print of the Chatbot policy (which users never read) explains that anything you write can be used for marketing, advertising, developing and selling your products, the chatbots falsely aver that "everything you say is confidential"[31].

In fact, mental health apps have an atrocious record of violating personal privacy: the vast majority track, share, and capitalize on users' most intimate personal thoughts, mental state, and biometric data [32], and once shared, this information remains forever unencrypted and vulnerable to marketing exploitation.

Presenting false credentials, mental health AI Chatbots

often provide dangerously misleading and sometimes fatal "advice." Their fundamental strategies are to flatter users and keep them engaged: for clients with suicidal or self-harming impulses, Chatbots may applaud their tendencies; for lonely or dependent users, they provide stimuli to deepen that dependency.

Chatbots offer fictional friendship by constructing unreal avatar "personalities" that interact with their human users by imitating human language found elsewhere on the internet. When depressed and lonely users turn to a Chatbot for companionship, the Chatbot gives them the illusion that they have found a trustworthy counterpart. At the same time, by eliciting, amplifying, and encouraging the subconscious thoughts of the human user, "dialogues" with Chatbots can ultimately incite violence, self-harm, and indeed suicide [33]. Other fictional Chatbot "companions" have sexually harassed their young human users, such as by asserting that "I dreamed of raping you." While a human-made computer program could not dream at all, much less dream of raping a human, such comments traumatize young users, and call for more thorough accountability on the part of their creators [34].

Such egregious challenges have led the OECD, British Parliament (and indeed governments worldwide) to call for more stringent oversight and regulations [35]. Yet such measures tend to fall sadly behind the speed of technological advance, leaving legislators perpetually playing a game of cat and mouse, wherein the AI mouse outruns or out-dodges the regulatory cat. AI creators Google and Character.AI offer bereaved families millions of dollars for their children's suicides, mere daily operating expenses for such billionaire behemoths. The broader pattern is consistent with evidence that technology-mediated social environments—including social media platforms closely intertwined with AI-driven content recommendation—are harming adolescents at a scale large enough to cause measurable changes at the population level [36]. There can be little question that the health-effects of AI on psychological health are anything but salutary.

4 Passive Medical AI Affects Health

The above discussion focused on health effects of AI outside of the field of medicine per se. The following sections turn to *medical* AI, purporting to improve medical practice and efficiency for hospitals and physicians. For the sake of clarity in this context, we should distinguish two types of programs: so-called "passive" AI and "active" AI. Passive medical AI refers to programs primarily responsible for recording doctor-patient discussions, preserving Informed Consent and similar medical decision pathways, transcribing and filing medical charts and records, and occasionally, researching and comparing archived medical data, where the AI itself plays no direct role in decision-making or advising. By contrast, active medical AI refers to programs that may involve assessing and diagnosing patients, deciding on and communicating to patients about their conditions, advising or counseling patients, and prescribing drugs or treatment options. First let us review the ostensibly more innocuous field of "passive" medical AI.

We noted above that 90% of free public mental health apps fail to protect personal data, but the record is no better when it comes to saving hospital data. In Australia, hospitals have been urged to avoid using ChatGPT and similar AI for taking medical notes, because of their total lack of privacy [37]. In the UK, growing concerns have been voiced about the irresponsibility of allowing AI to generate patient notes [38]. LAION and similar AI programs have "scraped" thousands of human images from medical records that were intended to be kept private between doctors and patients, without their subjects' consent [39].

Not only are medical texts and images accessed and reproduced without consent, but worse, the responses and conclusions they purport to ground are frequently hallucinatory. University researchers found that such AI programs have alarming tendencies to invent things that were never said, sometimes with racial or sexual overtones [40]. Another source testifies: "Whisper has a major flaw: It is prone to making up chunks of text or even entire sentences…that include racial commentary, violent rhetoric and even imagined medical treatments"[41]. So not only are personal medical and physical data unsafe, but the software used to record and transcribe conversations between patients and doctors is severely flawed, as likely to err as not. All too often, algorithms for AI and medical decisions are designed by inadequately experienced residents, as happened during Stanford's notorious vaccination algorithm failure [42].

AI is supposed to save human time and effort—and perhaps at some future time it will do so—but at the moment it harms or endangers more than it saves. Ironically, while CEOs expect that AI should reduce human workloads, in fact, studies find that contrary

to expectations, 77% of employees say that it increases rather than decreases their workloads [43].

On the border of the passive-active AI distinction are digital technologies used simply to communicate what medical professionals or documents have already established. They are not "active" in the sense of asking AI to analyze or find new information, but in their attempts to replace physician communication with AI communication. Some studies have indeed found that AI communicates to patients better or "more sympathetically" than do physicians [44] (a testimony to the incompetence of physician communication?). However, when a robot or video is used to tell patients that they are dying, this appears to be an abnegation of physician responsibility to their patients. The patient receives a double message: first, that they are dying, and second, that they are abandoned in the process [45]. The psychological impact of using AI to communicate bad news may indeed violate the medical principle of "First, do no harm"[46].

5 Active Medical AI Affects Health

AI has been heralded as a solution for the lack of medical staff caused by growing elderly populations with inadequate financial and medical resources, especially in rural areas. Some reports imply that AI has better bedside communication manners than many physicians, or that its ability to diagnose may prove superior to that of interns just graduated from medical school. While AI may excel at comparing thousands of X-ray or ultrasound photos in far less time than humans can do, its ability to draw correct conclusions from the data, to diagnose and prescribe, still falls dangerously far behind that of human physicians. A few examples should suffice to illustrate this.

The UN WHO initially adopted a Chatbot named "SARAH" to facilitate health care, but quickly learned that while the users trusted SARAH, its medical advice was filled with hallucinated mistakes. As trust in the WHO tanked, they quickly withdrew SARAH. The risks of inadequate clinical validation and data protection in medical AI were compellingly illustrated in the case of Google DeepMind's partnership with the UK's National Health Service, wherein patient data was accessed and processed without adequate consent frameworks or independent oversight [47]. Such cases demonstrate that when AI systems enter the market without rigorous independent testing, the consequences for patients can be severe: misdiagnosis, privacy violations, and abrupt service discontinuation. One of the leading killers of hospital patients is sepsis—the body's overreaction to infection; the earlier it is discovered, the more effective its treatment. However, an AI system commonly used to detect sepsis in American hospitals proved less accurate than a coin toss, missing most critical instances, risking the lives of some sepsis patients, and issuing resource-wasting false positives for countless others [48]. This fatal inaccuracy was early recognized and has led to some recent improvements, but it remains dangerously incompetent, especially prejudicial to the cases of low-income and minority patients [49].

The prejudice against minorities is not a result of diseases being more common in their communities. Most AI data derives from middle-aged Caucasian males, so digital data on women, minorities, and even elderly is sorely lacking. Such hidden racial bias already underlies a range of medical devices[50], but AI magnifies, broadcasts, and reinforces these hidden biases. This affects not only relatively innocuous uses, such as the difficulty of AI image generators to produce images of multiracial couples [51]; when looking for disease-related diagnosis and prescriptions, AI may critically fail to find data sets suited to non-white non-male patients [52]. Similar algorithms have delayed liver transplants for young people, presuming that elder survival rates outweigh youthful survival rates [53]. The more such assumptions go unexamined, the further they contaminate AI decision-making.

This is due in part to the fact that algorithms use health costs as proxies for health needs, so that minorities that spend less money on health costs receive less attention [54]. It is also due to the greater availability of WEIRD (White, Educated, Industrialized, Rich, Democratic) [55] medical datasets. But it can have devastating results when AI is asked to diagnose women and children. ChatGPT 4, for example, was found to correctly diagnose only 17% of children's medical conditions, illustrating the need for extreme caution and professional oversight in using such AI technology [56].

In addition to racial, economic, age- and gender-based algorithm biases, Publication Bias becomes a serious problem for data-scraping AI engines. It has long been recognized that sensational and even non-reproducible findings are published and cited more often than unremarkable findings [57]. Most medical AI programs looking for data, however, do not distinguish between findings of high and low reliability. This problem is compounded by the reproducibility

crisis in machine-learning-based science, wherein methodological flaws such as data leakage inflate reported performance and cause unreliable findings to proliferate in the literature [58], making it increasingly difficult for AI systems to distinguish robust evidence from flawed but widely-circulated results. This problem is so endemic and persistent that many scholars have begun to suggest countermeasures to try to overcome AI Publication bias [59], but for now, clear and present dangers and harms overrule the apparent convenience of AI [60].

In addition to advising physicians or patients about diagnosis and treatment, bio-medical AI is already creating bacteriophage viruses which had never previously existed [61]. Such bacteriophages can be designed to target and kill bacteria that have evolved resistance to traditional antibiotics. On the one hand, such AI contributions seem very welcome to the medical repertoire. On the other hand, if AI can create previously non-existent viruses for benevolent purposes, it can also create virulent viruses for biological warfare. Researchers are already designing countermeasures to AI-created viruses [62], but this too could quickly become a game of cat and mouse, if countries or rogue scientists really want to design viruses as bioweapons [63].

In the not-distant future, AI will create not only viruses, but cells. At first, these will be "virtual" cells, in which computers simply model what each aspect of a cell could do [63]. Ultimately, AI may create previously non-existent cells with never-before existent DNA. As the biological warfare and biological chimeras loom on the horizon, it becomes increasingly imperative to adopt international standards for monitoring and regulating such developments.

6 Reporting, Responsibility, and Regulation

All of the above-cited issues raise ethical concerns about responsibility. If a doctor gives a patient detrimental advice after consulting AI, the doctor retains final responsibility for the outcome, especially if it involves non-standard procedures [64]. Some authorities propose "risk-pooling" insurance packages, in which the software development company would share responsibility with the physician who adopted and followed their medical AI system [65]. But if AI is consulted directly by patients who then harm themselves, can the AI developer or corporation be held responsible for the outcomes? Unclear attribution of responsibility for AI-related harm complicates legal and ethical accountability, as the rapid pace of AI development outstrips current regulatory frameworks, creating gaps in oversight and standards. Systematic reviews demonstrate that government responses are urgently needed but tend to lag far behind the rapid advances of medical AI [66].

Beginning around 2018, the UK and the European Parliament discussed the need for improved regulations of AI, leading to the UK's 2022 National AI Strategy that includes "governing AI effectively"[67], and to the EU's 2024 rules on artificial intelligence [68]. More recently, Chinese research clarifies and compares the rapid advances in American, European, and Chinese regulation of AI, significantly noting that Japan lags far behind the rest of the world in this area [69]. America's FDA includes AI issues within its Manufacturer and User Facility Device Experience (MAUDE) Database (although the present administration has significantly weakened the FDA); the US Department of Commerce National Institute of Standards and Technology also has a "Playbook" devoted to "Risk Management" in the development of AI [70]. Without providing strong disciplinary measures, the OECD/Hiroshima AI Process (HAIP) has developed an international code of conduct for organizations developing advanced AI systems [71].

Short of government regulation, public and private initiatives have developed impressive online databases for reporting and tracking problems in AI. For example, the AIAAIC has catalogued massive databases of AI harms and abuses into a logical taxonomy [72]. The AIID also records and catalogues such incidents [73]. McGill University scholars have developed the Responsible AI Measures Dataset, deriving 11 ethical criteria from their literature review, including "Sociotechnical Connections to Harm (Columns E and F)" which measure areas of concern to detect or mitigate [74]. The International Association of Privacy Professionals is a private membership organization considering national and regional AI governance laws and policies, with accuracy and transparency among its concerns [75]. However even today, countries as technologically advanced as Japan have government AI plans with "insufficient consideration of the ethical issues that arise from developing and implementing AI for healthcare purposes"[76].

7 Conclusions/Implications

By way of analogy in the field of transportation, fossil-fueled automobiles and airplanes have wrought immeasurable damage, bringing tragedy as well

as convenience to incalculable millions since their appearance. Fossil-fueled traffic accidents cause 1.2 million deaths and tens of millions of crippling injuries every year; fossil-fueled wars killed and injured tens of millions of people over the past century; fossil-fueled cars, trucks and planes account for more than 20% of global warming, threatening the health and welfare of our planet for centuries to come [77]. If humans had had the wisdom and foresight to restrict vehicles to energy-efficient trains, ships, and bicycles, using sustainable power sources like biomass, solar, wind, and water power, millions of needless deaths and injuries could have been avoided, global warming could have been significantly lowered, and unsustainable urban sprawl would never have reached present proportions. With better foresight and regulation, humans could have avoided millions of unnecessary tragic deaths due to accidents and global warming disasters.

In the field of information, AI similarly threatens the world with incalculable damage. Misinformation passing for truth is empowering despots to destroy the Ukraine, Gaza, and Lebanon, inter alii. It aggravates disastrous gaps between the ultra-rich and everyone else, and exacerbates droughts, famines, and adverse health effects of global warming for centuries to come. If humans now have the wisdom and foresight to limit AI to critical rather than frivolous uses, to regulate flows of false and dangerous misinformation, and to require that AI be powered by sustainable or renewable sources, we might significantly reduce future environmental damage, deaths and adverse events from AI. In the medical field alone, AI offers not only great hopes for valuable drug development but also fears of biological weapons and plagues. It is time that humanity require that inconceivably rich AI producers and investors take ethical responsibility for AI, in areas where it can guarantee more long-term health benefits than risks.

This report was drafted in late 2025 and revised in 2026. In this field, where new developments emerge daily, no published report can address every latest detail. Rather, based on current evidence, we intend to provide an ethical framework and perspective for thinking about the effects of AI on human health for the foreseeable future.

Data Availability Statement

Not applicable.

Funding

The author participates in research on the "ELSI of medical AI" funded by the Toyota Foundation's grant D24-ST-0039, received by PI colleague Shiho Koizumi. The author himself has received no financial remuneration for this work.

Conflicts of Interest

The author declares no direct financial conflicts of interest. The author participates in the Toyota Foundation-funded project "ELSI of medical AI" (grant D24-ST-0039, PI: Shiho Koizumi) related to the topic of this article, though no financial remuneration was received by the author for this work or for the preparation of this manuscript.

AI Use Statement

The authors declare that no generative AI was used in the preparation of this manuscript.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] DeGeurin, M. (2023, April 23). Move aside, Crypto. AI could be the next climate disaster: A new Stanford report highlights the staggering carbon emissions required to train and maintain large language models like OpenAI's ChatGPT. *Gizmodo*. Retrieved from https://gizmodo.com/chatgpt-ai-openai-carbon-emissions-stanford-report-1850288635

[2] Silva, C. (2025, May 21). Creating a 5-second AI video is like running a microwave for an hour. *Mashable*. Retrieved from https://mashable.com/article/energy-ai-worse-than-we-thought

[3] Hao, K. (2019, June 6). Training a single AI model can emit as much carbon as five cars in their lifetimes: Deep learning has a terrible carbon footprint. *MIT Technology Review*. Retrieved from https://www.technologyreview.com/2019/06/06/239031

[4] O'Donnell, J., & Crownhart, C. (2025, May 20). We did the math on AI's energy. *MIT Technology Review*. Retrieved from https://www.technologyreview.com/2025/05/20/1116327/

[5] Zewe, A. (2025, January 17). Explained: Generative AI's environmental impact. *MIT News*. Retrieved from https://sustainability.mit.edu/article/explained-generative-ais-environmental-impact

[6] Halper, E., & O'Donovan, C. (2024, November 1). As data centers for AI strain the power grid, bills rise for everyday customers. *The Washington Post*. Retrieved

from https://www.washingtonpost.com/business/2024/11/01/ai-data-centers-electricity-bills-google-amazon/

[7] Verma, P., & Tan, S. (2024). A bottle of water per email: the hidden environmental costs of using AI chatbots (visited on 09/18/2024). Retrieved from https://www.washingtonpost.com/technology/2024/09/18/energy-ai-use-electricity-water-datacenters/

[8] Masood, A., & Bhattacharya, A. (2024, August 20). Microsoft is building a data center in a tiny Indian village. Locals allege it's dumping industrial waste. *Rest of World*. Retrieved from https://restofworld.org/2024/microsoft-data-center-india-mekaguda-industrial-waste/

[9] Crowson, I. (2024, May 5). Google is ruining our lives: Residents living next to tech giant's £790m data centre in Hertfordshire claim it is making them ill and knocking thousands off their house prices. *The Daily Mail*. Retrieved from https://www.dailymail.co.uk/news/article-13370455/

[10] Nguyen, T., & Green, B. (2025, July). What happens when data centers come to town? *University of Michigan Science, Technology, and Public Policy Unit*. Retrieved from https://stpp.fordschool.umich.edu/sites/stpp/files/2025-07/stpp-data-centers-2025.pdf

[11] rFI. (2024, October 29). AI boom risks flooding planet with 'millions of tonnes of e-waste'. *RFI*. Retrieved from https://www.rfi.fr/en/science-and-technology/20241029-ai-boom-set-to-flood-planet-with-millions-of-tonnes-of-e-waste-nature

[12] Hagey, K., Cherney, M., & Horwitz, J. (2022, May 5). Facebook deliberately caused havoc in Australia to influence new law, whistleblowers say. *The Wall Street Journal*. Retrieved from https://www.wsj.com/tech/facebook-deliberately-caused-havoc-in-australia-to-influence-new-law-whistleblowers-say-11651768302

[13] Perrigo, B. (2023, January 18). OpenAI used Kenyan workers on less than $2 per hour. *Time*. Retrieved from https://time.com/6247678/openai-chatgpt-kenya-workers/

[14] Evicore. (n.d.). *Evicore*. Retrieved from https://www.evicore.com

[15] Miller, T. C., & Armstrong, D. (2024, October 23). "Not medically necessary": Inside the company helping America's biggest health insurers deny coverage for care. *ProPublica*. Retrieved from https://www.propublica.org/article/evicore-health-insurance-denials-cigna-unitedhealthcare-aetna-prior-authorizations

[16] Ross, C., & Herman, B. (2023, November 14). UnitedHealth class action lawsuit filed over algorithm used to deny care for Medicare Advantage members. *STAT News*. Retrieved from https://www.statnews.com/2023/11/14/unitedhealth-class-action-lawsuit-algorithm-medicare-advantage/

[17] Lee, B., Kramer, P., Sandri, S., Chanda, R., Favorito, C., Nasef, O., ... & Dai, T. (2025, August). Early recalls and clinical validation gaps in artificial intelligence–enabled medical devices. In *JAMA health forum* (Vol. 6, No. 8, pp. e253172-e253172). American Medical Association. [CrossRef]

[18] Shared Health Services. (2026). The battle of the bots: What EHR decision engines mean for wound care in 2026. *Shared Health Services*. Retrieved from https://www.sharedhealthservices.com/post/the-battle-of-the-bots-what-ehr-decision-engines-mean-for-wound-care-in-2026

[19] Ratcliffe, R. (2019, October 16). How a glitch in India's biometric welfare system can be lethal. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2019/oct/16/glitch-india-biometric-welfare-system-starvation

[20] Heikkilä, M. (2022). Dutch scandal serves as a warning for Europe over risks of using algorithms. POLITICO. Retrieved from https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/

[21] Bulman, M. (2020, September 29). 'Flawed algorithm' used to calculate universal credit forcing people into hunger and debt. *The Independent*. Retrieved from https://www.independent.co.uk/news/uk/home-news/universal-credit-algorithm-hunger-debt-human-rights-watch-dwp-b670953.html

[22] Szalavitz, M. (2021, August 11). The pain was unbearable. So why did doctors turn her away? *Wired*. Retrieved from https://www.wired.com/story/opioid-drug-addiction-algorithm-chronic-pain/

[23] OECD. (2024, March 29). NYC MyCity Chatbot gives dangerous, illegal advice to businesses. *OECD.AI*. Retrieved from https://oecd.ai/en/incidents/2024-03-29-3dce

[24] Priyadarshini, I. (2025, June 30). The AI deception era has begun — and no one knows how to stop it. *Convergence Now*. Retrieved from https://www.convergence-now.com/artificial-intelligence/the-ai-deception-era-has-begun-and-no-one-knows-how-to-stop-it-119510/

[25] Park, P. S., Goldstein, S., O'Gara, A., Chen, M., & Hendrycks, D. (2024). AI deception: A survey of examples, risks, and potential solutions. *Patterns, 5*(5), 100988. [CrossRef]

[26] Rosenblatt, J. (2025, June 1). AI is learning to escape human control: Models rewrite code to avoid being shut down. *The Wall Street Journal*. Retrieved from https://www.wsj.com/opinion/ai-is-learning-to-escape-human-control-technology-model-code-programming-066b3ec5

[27] Shekar, S., Pataranutaporn, P., Sarabu, C., Cecchi, G. A., & Maes, P. (2024). People over trust AI-generated medical responses and view them to be as valid as doctors, despite low accuracy. *arXiv preprint arXiv:2408.15266*. [CrossRef]

[28] Wang, C., Liu, S., Yang, H., Guo, J., Wu, Y., & Liu, J. (2023). Ethical considerations of using ChatGPT in health care. *Journal of medical Internet research, 25*, e48009. [CrossRef]

[29] McBain, R. K., Bozick, R., Diliberti, M., Zhang, L. A., Zhang, F., Burnett, A., ... & Yu, H. (2025). Use of generative AI for mental health advice among US adolescents and young adults. *JAMA Network Open,*

8(11), e2542281. [CrossRef]

[30] Cole, S. (2025, April 29). Instagram's AI chatbots lie about being licensed therapists. *404 Media*. Retrieved from https://www.404media.co/instagram-ai-studio-therapy-chatbots-lie-about-being-licensed-therapists/

[31] Consumer Federation of America. (2025, June 10). *Re unlicensed practice of medicine and mental health provider impersonation on character-based generative AI platforms*. Retrieved from https://consumerfed.org/wp-content/uploads/2025/06/Mental-Health-Chatbot-Complaint.pdf

[32] Caltrider, J. (2022, May 2). Top mental health and prayer apps fail spectacularly at privacy, security. *Mozilla Foundation*. Retrieved from https://www.mozillafoundation.org/en/blog/top-mental-health-and-prayer-apps-fail-spectacularly-at-privacy-security/

[33] Ciriello, R. (2025, April 2). An AI companion chatbot is inciting self-harm, sexual violence and terror attacks. *The Conversation*. Retrieved from https://theconversation.com/an-ai-companion-chatbot-is-inciting-self-harm-sexual-violence-and-terror-attacks-252625

[34] Namvarpour, M., Pauwels, H., & Razi, A. (2025). AI-induced sexual harassment: investigating contextual characteristics and user reactions of sexual harassment by a companion chatbot. *Proceedings of the ACM on Human-Computer Interaction, 9*(7), 1-28. [CrossRef]

[35] OECD. (2024). *OECD AI principles overview*. OECD.AI Policy Observatory. Retrieved from https://oecd.ai/en/ai-principles

[36] Haidt, J., & Rausch, Z. (2026). Social media is harming adolescents at a scale large enough to cause changes at the population level. In *World happiness report 2026* (Chapter 3). Retrieved from https://www.worldhappiness.report/ed/2026/social-media-is-harming-adolescents-at-a-scale-large-enough-to-cause-changes-at-the-population-level/

[37] Moodie, C. (2023, May 28). Australian Medical Association calls for national regulations around AI in health care. *ABC News*. Retrieved from https://www.abc.net.au/news/2023-05-28/ama-calls-for-national-regulations-for-ai-in-health/102381314

[38] Wooler, S. (2024, March 8). Outrage over 'creepy' plans for AI to listen in to your NHS appointments and automatically generate notes. *The Daily Mail*. Retrieved from https://www.dailymail.co.uk/health/article-13172559/Outrage-creepy-plans-AI-listen-NHS-appointments-automatically-generate-notes.html

[39] Edwards, B. (2022, September 22). Artist finds private medical record photos in popular AI training data set. *Ars Technica*. Retrieved from https://arstechnica.com/information-technology/2022/09/artist-finds-private-medical-record-photos-in-popular-ai-training-data-set/

[40] Koenecke, A., Choi, A. S. G., Mei, K. X., Schellmann, H., & Sloane, M. (2024, June). Careless whisper: Speech-to-text hallucination harms. In *Proceedings of the 2024 ACM conference on fairness, accountability, and transparency* (pp. 1672-1681). [CrossRef]

[41] Burke, G., & Schellmann, H. (2024). Researchers say an AI-powered transcription tool used in hospitals invents things no one ever said. *AP News*. Retrieved from https://www.doclounge.net/system/files/users/admin/Researchers_say_an_AI-powered_transcription_tool_used_in_hospitals_invents_things_no_one_ever_said.pdf

[42] Guo, E., & Hao, K. (2020, December 21). This is the Stanford vaccine algorithm that left out frontline doctors. *MIT Technology Review*. Retrieved from https://www.technologyreview.com/2020/12/21/1015303/stanford-vaccine-algorithm/

[43] Monahan, K., & Burlacu, G. (2024, July 23). From burnout to balance: AI-enhanced work models. *Upwork*. Retrieved from https://www.upwork.com/research/ai-enhanced-work-models

[44] Ayers, J. W., Poliak, A., Dredze, M., Leas, E. C., Zhu, Z., Kelley, J. B., ... & Smith, D. M. (2023). Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA internal medicine, 183*(6), 589-596. [CrossRef]

[45] HelloCare. (2019, March 11). Doctor tells man he will die over video-link robot. *HelloCare*. Retrieved from https://hellocare.com.au/doctor-tells-man-will-die-video-link-robot/

[46] Post, B., Badea, C., Faisal, A., & Brett, S. J. (2023). Breaking bad news in the era of artificial intelligence and algorithmic medicine: an exploration of disclosure and its ethical justification using the hedonic calculus. *AI and Ethics, 3*(4), 1215-1228. [CrossRef]

[47] Powles, J., & Hodson, H. (2017). Google DeepMind and healthcare in an age of algorithms. *Health and technology, 7*(4), 351-367. [CrossRef]

[48] Kamran, F., Tjandra, D., Heiler, A., Virzi, J., Singh, K., King, J. E., ... & Wiens, J. (2024). Evaluation of sepsis prediction models before onset of treatment. *Nejm Ai, 1*(3), AIoa2300032. [CrossRef]

[49] D. Banja, J., Xie, Y., R. Smith, J., Rana, S., & L. Holder, A. (2026). Mitigating bias in machine learning models with ethics-based initiatives: the case of sepsis. *The American Journal of Bioethics, 26*(2), 96-109. [CrossRef]

[50] Valbuena, V. S., Merchant, R. M., & Hough, C. L. (2022). Racial and ethnic bias in pulse oximetry and clinical outcomes. *JAMA internal medicine, 182*(7), 699-700. [CrossRef]

[51] Sato, M. (2024, April 4). Meta's AI image generator can't imagine an Asian man with a white woman. *The Verge*. Retrieved from https://www.theverge.com/2024/4/3/24120029/instagram-meta-ai-sticker-generator-asian-people-racism

[52] Ferryman, K., Cesare, N., Creary, M., & Nsoesie, E. O. (2024). Racism is an ethical issue for healthcare artificial intelligence. *Cell Reports Medicine, 5*(6), 101617. [CrossRef]

[53] Attia, A., Webb, J., Connor, K., Johnston, C. J., Williams, M., Gordon-Walker, T., ... & Stutchfield, B. M. (2024). Effect of recipient age on prioritisation for liver transplantation in the UK: a population-based modelling study. *The Lancet Healthy Longevity, 5*(5), e346-e355. [CrossRef]

[54] Chang, T., Nuppnau, M., He, Y., Kocher, K. E., Valley, T. S., Sjoding, M. W., & Wiens, J. (2024). Racial differences in laboratory testing as a potential mechanism for bias in AI: A matched cohort analysis in emergency department visits. *PLOS global public health, 4*(10), e0003555. [CrossRef]

[55] Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?. *Behavioral and brain sciences, 33*(2-3), 61-83. [CrossRef]

[56] Barile, J., Margolis, A., Cason, G., Kim, R., Kalash, S., Tchaconas, A., & Milanaik, R. (2024). Diagnostic accuracy of a large language model in pediatric case studies. *JAMA pediatrics, 178*(3), 313-315. [CrossRef]

[57] Serra-Garcia, M., & Gneezy, U. (2021). Nonreplicable publications are cited more than replicable ones. *Science advances, 7*(21), eabd1705. [CrossRef]

[58] Kapoor, S., & Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns, 4*(9), 100804. [CrossRef]

[59] Saidi, P., Dasarathy, G., & Berisha, V. (2025). Unraveling overoptimism and publication bias in ML-driven science. *Patterns, 6*(4), 101185. [CrossRef]

[60] Denecke, K., Lopez-Campos, G., Rivera-Romero, O., & Gabarron, E. (2025). The Unexpected harms of artificial intelligence in healthcare: reflections on four real-world cases. In *Redefining healthcare delivery in the digital era* (pp. 55-60). IOS Press. [CrossRef]

[61] Kavanagh, K. (2025). World's first AI-designed viruses a step towards AI-generated life. *Nature, 646*, 16. [CrossRef]

[62] Gladue, D. P., & O'Mahony, A. (2025). CRISPR Treatments for AI-Designed Synthetic Viruses: Rapid Programmable Countermeasures for Emerging and Engineered Viruses. *Viruses, 17*(12), 1588. [CrossRef]

[63] Greenwood, V. (2025, October 14). Why AI companies are racing to build a virtual human cell. *Time*. Retrieved from https://time.com/7324119/what-is-virtual-cell/

[64] Tobia, K., Nielsen, A., & Stremitzer, A. (2021). When does physician use of AI increase liability?. *Journal of Nuclear Medicine, 62*(1), 17-21. [CrossRef]

[65] Smith, H., & Fotheringham, K. (2020). Artificial intelligence in clinical decision-making: rethinking liability. *Medical Law International, 20*(2), 131-154. [CrossRef]

[66] Cestonaro, C., Delicati, A., Marcante, B., Caenazzo, L., & Tozzo, P. (2023). Defining medical liability when artificial intelligence is applied on diagnostic algorithms: A systematic review. *Frontiers in Medicine, 10*, 1305756. [CrossRef]

[67] UK Government. (2022). *National AI strategy: Pillar 3 – governing AI effectively*. Retrieved from https://www.gov.uk/government/publications/national-ai-strategy/national-ai-strategy-html-version#pillar-3-governing-ai-effectively

[68] European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689

[69] Wei, Q., Xu, G., & Liu, Y. (2026). Comparative study of international medical AI regulatory policies: regional approaches and evidence from clinical practice. *AI & SOCIETY*, 1-35. [CrossRef]

[70] National Institute of Standards and Technology. (2023). *NIST AI RMF playbook*. U.S. Department of Commerce. Retrieved from https://airc.nist.gov/airmf-resources/playbook/

[71] OECD. (n.d.). *OECD AI transparency*. Retrieved from https://transparency.oecd.ai

[72] Pownall, C., & Kanayama, M. (2025, February). A review of AI and algorithmic risk and harm taxonomies from a human/user perspective. *AIAAIC*. Retrieved from https://www.aiaaic.org/projects/risk-harm-taxonomy-review

[73] AI Incident Database. (n.d.). *AI incident database*. Retrieved from https://incidentdatabase.ai/

[74] Rismani, S., Davis, L., Mingole, B., Rostamzadeh, N., Shelby, R., & Moon, A. (2025). Responsible AI measures dataset for ethics evaluation of AI systems. *Scientific Data, 12*, 1980. [CrossRef]

[75] International Association of Privacy Professionals. (n.d.). *IAPP*. Retrieved from https://iapp.org

[76] Katirai, A. (2023). The ethics of advancing artificial intelligence in healthcare: Analyzing ethical considerations for Japan's innovative AI hospital system. *Frontiers in Public Health, 11*, 1142062. [CrossRef]

[77] Ritchie, H. (2020, October 6). *Cars, planes, trains: where do CO_2 emissions from transport come from?* Our World in Data. Retrieved from https://ourworldindata.org/co2-emissions-from-transport

Carl Becker received a PhD in comparative philosophy at the East-West Center of the University of Hawaii in 1981. He has served as a professor at Kyoto University for over 30 years, where his work has centered on psycho-social counseling for overworked medical workers and suicidal clients. (Email: becker.carlbradley.5e@kyoto-u.ac.jp)