Beyond Carbon: A Call for Research on the Impacts of Computing Systems on Human Health

Dorota Kopczyk
University of Minnesota
Minneapolis, Minnesota, USA
kopcz004@umn.edu

Abhishek Chandra
University of Minnesota
Minneapolis, Minnesota, USA
chandra@umn.edu

Abstract

Sustainability research for large-scale computing has converged on carbon as the primary accounting unit, and much work has gone into reducing it. However, carbon emissions are only one possible indicator of computing's impact. The human health impacts of computing systems have not received the same attention. U.S. data centers are projected to drive premature deaths and asthma cases from operational air pollution, and there is an asymmetry between carbon-optimal scheduling and the regions where human health is impacted. Embodied impacts extend upstream into mining and manufacturing and downstream into e-waste processing. Neither registers in a CO_2e number. This paper maps documented health impacts across the modern large-scale computing hardware lifecycle. We call for systems research into minimizing the human health impacts incurred by large-scale computing. We propose a number of possible research directions towards designing *health aware systems*: extending computational impact models and operational scheduling, lifetime extensions and repair, and a measurement infrastructure to track the impact through the hardware lifecycle.

CCS Concepts

• **Social and professional topics** → **Sustainability**; • **Hardware** → **Impact on the environment**; • **Computer systems organization** → *Cloud computing*.

Keywords

sustainable computing, human toxicity, hardware life cycle, supply chain, e-waste, rare earth elements, embodied impact

ACM Reference Format:

Dorota Kopczyk and Abhishek Chandra. 2026. Beyond Carbon: A Call for Research on the Impacts of Computing Systems on Human Health. In *Proceedings of HotCarbon'26, July 16-17th, 2026, Seattle, WA. (HotCarbon '26).* ACM, New York, NY, USA, 7 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Sustainability research for large-scale computing has largely focused on carbon as the primary accounting unit, which is typically correlated to the energy consumed. U.S. datacenters consumed 176 TWh in 2023, representing 4.4% of national electricity demand, up from 1.9% in 2018. Projections reach 325–580 TWh by 2028 [26]. Carbon-aware scheduling [1, 30], embodied carbon estimation [4], and lifecycle analysis [9] have all improved our understanding of the climate costs associated with large-scale computing systems. Industry has matched the framing. Google has targeted 24/7 carbon-free operation by 2030 [22], Microsoft has committed to being carbon-negative by 2030 [28], and Apple has pledged carbon neutrality across its entire footprint by 2030 [3]. However, carbon remains a single-dimensional aggregate metric. It compresses mineral extraction, semiconductor fabrication, assembly, operation, and end-of-life processing into a common global unit of CO_2e, while obscuring where these processes occur and the human conditions under which they take place.

HotCarbon '26, 978-1-4503-XXXX-X/2026/07

ACM ISBN 978-x-xxxx-xxxx-x/YYYY/MM
https://doi.org/XXXXXXX.XXXXXXX

Recent work suggests that carbon emissions and human health impacts often capture different system impacts. A physical teardown analysis of the Nvidia A100 found that manufacturing accounted for 94.5% of the device's lifetime human toxicity (cancer) impact, while contributing less than 3% of its total climate change impact [7]. At the operational level, carbon intensity and health-impact intensity across U.S. power grids correlate only weakly ($r = 0.292$) [10]. In practice, routing workloads to the lowest-carbon grid may increase public health costs in many regions. Prior analysis projects that U.S. datacenter expansion could contribute to approximately 1,300 premature deaths and 600,000 asthma cases annually by 2028 [10]. Water shows the same divergence. Drawing down a stressed aquifer concentrates the contaminants already present [15, 29]. Data centers add a large new withdrawal that local water systems never planned for. These impacts are difficult to capture through carbon accounting alone.

We identify human health impact as an under-addressed dimension of sustainable computing and argue that future systems research should incorporate health and social externalities alongside carbon and energy efficiency. This paper makes the following contributions:

- We extend the health impact framing across the full hardware lifecycle rather than operations alone.
- We identify the key shortcomings in the current sustainability models and mechanisms from a human health perspective.
- We propose a number of possible research directions towards designing *health aware systems*: extending computational impact models and operational scheduling, lifetime extensions and repair, and a measurement infrastructure to track the health impacts through the hardware lifecycle.

LAYER	**MATERIALS** Mining & extraction	**FABRICATION** Semiconductor fabs	**OPERATION** Data centers	**END-OF-LIFE** E-waste recycling
IMPACT	• Lung cancer [17] • Radiation exposure [34] • Heavy-metal poisoning [14]	• Cancer [13] • Birth defects [13] • Arsenic exposure [16]	• Asthma [10] • Drinking-water contamination [15, 27, 29]	• Dioxin exposure [23] • Childhood lead poisoning [21] • Take-home lead exposure [5]
LOCALE	*Montana (U.S.), Inner Mongolia (China), DRC*	*Taiwan, South Korea, mainland China*	*Across the U.S.*	*China, Ghana, Uruguay, Pakistan*

Figure 1: Health impacts across the computing hardware lifecycle. Signatures shift by stage (occupational vs. community, acute vs. chronic); affected communities are decoupled from the corporations and carbon frameworks governing them.

2 Background: Human Health Impacts Across The Hardware Lifecycle

This section synthesizes peer-reviewed findings from environmental epidemiology and occupational health, along with existing LCA work on hardware, to provide the background on human health impacts of computing across the full hardware lifecycle.

The hardware in large-scale computing (in datacenters, running AI) passes through four distinct stages (Figure 1), from the elements extracted to build it, to the factories that manufacture its components, to the data centers that operate it, to the recyclers. Each stage produces a different human-health signature. Understanding where those impacts originate is a prerequisite for building systems that can account for them.

2.1 Materials: Mining and Extraction

The hardware underpinning modern AI draws on aluminum (refined from bauxite ore), copper, rare earth elements (REEs), gallium, germanium, and tantalum [7, 8]. Each is extracted under conditions that impose documented health costs on the workers and communities nearest to the mine.

Copper smelting releases arsenic trioxide, which is an IARC Group 1 carcinogen. A longitudinal study of 8,014 Montana copper smelter workers spanning from 1938 to 1989 found respiratory cancer mortality rising with cumulative arsenic exposure, reaching 2.4× baseline (SMR 2.40, 95% CI 1.9–3.0) in the highest-exposure category [17]. At Bayan Obo in Inner Mongolia, the world's largest REE deposit and source of roughly 70% of global REE production [8], mine workers show elevated rates of lung disease and cancer from REE particle inhalation. Additionally, residents of mining areas inhale 101–430 μSv/year from thorium-bearing $PM_{2.5}$ [34]. The pattern is widespread. Artisanal and small-scale mining of the metals essential to AI hardware, including cobalt, copper, tin, tantalum, and lithium, operates with PPE deficits and elevated exposures to mercury, lead, arsenic, and cadmium [14].

2.2 Fabrication: Semiconductor Fabs

Semiconductor fabrication is a chemically intensive manufacturing process. A 2014 epidemiological review of US, UK, Taiwanese, and Korean fabrication cohorts [13] documents a broad pattern of worker and community consequences. Wafer fabrication workers are exposed to organic solvents, hydrofluoric acid, and glycol ethers. The group finds excess risk for non-Hodgkin's lymphoma, leukemia, brain tumors, and breast cancer. Workers in photolithography and diffusion show spontaneous abortion rates up to 3.21× above non-fabrication peers, and children born to fathers in semiconductor production are three times more likely to die from congenital malformation. The effects leach out of the factories too. Chlorinated solvents from semiconductor plants have been found in drinking water in San Jose, California and in Taiwan.

Gallium arsenide (GaAs), surging in demand due to ML and 5G, is an IARC Group 1 carcinogen that dissolves in lung tissue, releasing arsenic. Current exposure limits cover arsenic alone, missing the compound's synergistic toxicity. A 2026 risk assessment of Taiwanese semiconductor workers found that even administrative staff, the lowest-exposure group in the building, had a mean hazard quotient of 61, more than sixty times the acceptable threshold [16].

These hazards compound as the industry scales. Denser process nodes improve performance-per-watt but require more chemical process steps, more mineral-intensive manufacturing, and higher fabrication toxicity per device. Specialized components such as Power-on-Package (PoP) units have no dedicated life cycle assessment data [7]. These impacts occur in semiconductor manufacturing facilities, like those in Taiwan.

Fabrication accounts for 94.5% of a GPU's lifetime cancer human toxicity [7], geographically concentrated in the communities surrounding semiconductor facilities in Taiwan, South Korea, and mainland China. This A100 figure is the only toxicity breakdown published for any device. Comparable numbers for the rest of the server stack do not yet exist, and that absence is itself a measurement gap we identify.

2.3 Operation: Data Centers

Once running, data centers impose their own burden on downwind communities. Criteria air pollutants ($PM_{2.5}$ and NO_x from grid electricity and on-site diesel backup generators) reach surrounding neighborhoods regardless of how efficiently the hardware runs. U.S. data centers cause roughly 1,300 premature deaths and 600,000 asthma cases per year. Training a single Llama-3.1-scale model produces criteria air pollutant emissions equivalent to more than 10,000 cross-country car round trips [10]. The intuitive remedy is to shift load to greener grids. That does not translate cleanly to health outcomes. Carbon intensity and health-impact intensity across U.S. grids are weakly correlated ($r = 0.292$). In 96% of regions the spatial-temporal variability of public health costs exceeds that of carbon emissions, so carbon-optimal scheduling routinely lands on health-suboptimal regions [10].

Data centers also withdraw water, both directly for cooling and indirectly through the electricity they consume. Unlike air pollution, water reaches health only indirectly. A data center emits no contaminant of its own. Any large, unplanned draw on a stressed aquifer concentrates the contamination already in it, and a data center is one more such draw. U.S. data centers used 66 billion liters directly in 2023. Grid-side consumption adds nearly 800 billion liters more. Hyperscale facilities are projected to reach 60–124 billion liters of direct withdrawals annually by 2028 [26]. One fifth of that direct footprint comes from moderately to highly water-stressed watersheds [26, 27], where additional withdrawals pull contaminated shallow groundwater (nitrates from agriculture) into drinking-water depths. Across roughly 6,000 wells in California's Central Valley, such withdrawals degraded water quality 3–5× faster during drought [15]. Drinking-water nitrate at elevated levels is associated with infant methemoglobinemia, colorectal cancer, thyroid disease, and neural tube defects. Recent epidemiology finds increased risk even at concentrations below the EPA maximum contaminant level of 10 mg/L NO_3-N [35]. The mechanism is not specific to nitrate. In the same Central Valley, over-pumping releases naturally occurring arsenic from aquifer clays, as much as tripling the risk of hazardous levels [29]. None of this registers in carbon metrics or aggregate water-volume disclosures.

2.4 End-of-Life: E-Waste Recycling

Most electronic waste is not data center hardware. Retired servers and accelerators are a small but fast-growing share of a much larger global stream. What follows describes those pathways. An estimated 80% of electronic waste collected in developed countries is illegally exported to low and middle-income countries in Asia, Africa, and Latin America, despite the Basel Convention's prohibition on hazardous waste exports – a treaty the U.S. has not ratified [20, 23].

Should a device find itself in the informal recycling stream, the recyclers who process that waste use what is available. They burn wire insulation in the open, leach precious metals with acid, and dismantle boards by hand on bare ground. Each method releases specific toxicants (e.g. lead, cadmium, mercury, dioxins, brominated flame retardants) into the air, water, soil and food supply of whoever lives nearby [20]. In Guiyu, China, one of the world's largest informal e-waste operations, airborne dioxins have been measured at 15–56× the WHO's recommended maximum intake [23]. The pattern is global. In Montevideo, Uruguay, children near informal cable-burning sites had mean blood lead of 9.19 μg/dL, nearly twice the CDC's clinical reference value. Soil lead levels reached 17.7× the EPA residential limit [21]. Across literature, e-waste exposure is associated with 2.4× higher odds of ADHD diagnosis, suppressed vaccine response, and measurable DNA oxidative damage [20].

Formalization of recycling is not a silver bullet fix. A systematic review of 37 studies of permitted e-recycling facilities found worker air samples for lead and cadmium routinely exceeding occupational exposure limits, elevated brominated flame-retardant levels in workers' serum, and lead from dust carried home on work clothes [5]. The impact is smaller in magnitude than informal hubs but real.

Machine learning is accelerating this impact by volume. Under varying generative AI growth scenarios, cumulative AI-related e-waste could reach 1.2–5.0 million tonnes by 2030 [33], while rare earth elements are recovered at rates below 1% of global supply [2].

3 Where Current Systems Models Fall Short

The preceding section maps documented human health impacts across the hardware lifecycle. The sustainability frameworks governing AI infrastructure were not designed to track who gets sick. Three structural measurement failures follow. Section 4 proposes remedies for each.

3.1 Scope 3 Without Social Granularity

Gupta et al. [9] brought embodied carbon to systems researchers' attention by quantifying its shift to dominance. The share of lifecycle emissions attributable to hardware manufacturing rose from 49% for the iPhone 3GS (2009) to 86% for the iPhone 11 (2019), and for Facebook's renewables-powered data centers reached 82% of remaining emissions. Bashir et al. [4] subsequently cautioned that optimizing embodied carbon in isolation can mislead, overstating real reductions and conflicting with operational-carbon incentives. We argue the same systems-tractable framing extends to embodied *human health impact*. It accumulates per device, scales with refresh cadence, and demands its own accounting. Under the GHG Protocol, the standard for emissions accounting, Scope 3 covers indirect value-chain emissions outside a company's direct operations, including upstream manufacturing and downstream disposal. Embodied carbon aggregates these into a single CO_2e figure, but in doing so discards three dimensions relevant to health impact.

3.1.1 Carbon-to-health mapping. Identical units of CO_2e can carry very different human-toxicity burdens depending on which lifecycle phase produced them. Falk et al. found that A100 manufacturing accounts for 94.5% of the device's lifetime cancer human toxicity but under 3% of its climate impact. Operations is the inverse, with 97% of climate impact and only 5.5% of cancer toxicity [7]. The same gram of CO_2e represents a categorically different health impact signature depending on where in the lifecycle it was emitted.

3.1.2 Missing health accounting metrics. Occupational hazards in mining and fabrication are invisible to a metric that counts only the molecules emitted to the atmosphere. A smelter worker inhales arsenic. A fab worker handles glycol ethers. An artisanal miner works without protective gear. None of these absorbed exposures show up in a CO_2e accounting.

3.1.3 Missing local impact. Health impact is locally borne but globally aggregated. A tonne of CO_2e from a Taiwanese fab and a tonne from a Virginia data center are treated identically in the ledger, despite producing categorically different localized impacts on the communities exposed to them.

Two structural barriers hinder addressing this. Vendors do not disclose per-SKU elemental composition for AI server hardware. No regulatory standard requires this. Without that disclosure, no workload can be connected to its extraction region, its processing facility, or the community bearing the embodied health risks.

Defining an adequate schema is itself an open systems problem, including which materials, at what granularity, and across which supply-chain tiers. Separately, the GHG Protocol has no health impact analog. Scope 1 (on-site generators), Scope 2 (grid electricity), and Scope 3 (manufacturing and end-of-life) each correspond to a documented health burden mapped in Section 2, yet none is required reporting. Carbon optimization is a single-objective problem with a well-defined gradient, reducing gCO_2e. Human health impact is multi-objective, and the objectives are not co-located.

3.2 Carbon-Optimal ≠ Health-Optimal

3.2.1 Manufacturing tradeoffs. Process-node shrinks reduce per-FLOP operational carbon, but they require more chemical process steps and more mineral-intensive manufacturing per device [7]. What is saved in operational gCO_2e is paid in toxic by-products at fabrication.

3.2.2 Scaling amplification. Per-device efficiency does not bound aggregate health impact. Each new GPU generation increases total demand for high-risk minerals, including gallium, germanium, and rare earth elements. The resulting hardware turnover is a primary driver of AI-related e-waste [7, 33]. Volume amplifies impact even when per-device intensity falls.

3.2.3 Geographic divergence. Geographic scheduling has the same problem at runtime. Water intensity and carbon intensity across U.S. counties follow different patterns [12]. Carbon-minimizing workload placement can route computations into water-stressed watersheds. Each additional withdrawal pulls nitrate-laden shallow groundwater into drinking-water depths and accelerates aquifer degradation 3–5× during drought [15].

In each case the missing objective is its own physical quantity, and a scheduler optimizing for carbon alone cannot satisfy it.

3.3 Lack of Attribution Mechanisms

For a systems researcher, attribution is the chain of identifiers that traces a compute action back to its physical determinants. A request points to a server SKU, a SKU to a die, a die to a fab, a fab to its inputs. Carbon attribution closes that loop reasonably well, drawing on grid mix, PUE, server power, and embodied carbon estimates. The same cannot be said for health impact attribution, where the chain breaks in three distinct places.

3.3.1 Disclosure gaps. Server-hardware vendors do not disclose per-SKU elemental composition. Apple's Full Material Disclosure program shows that the data can be gathered at scale, mapping 1,900 materials in iPhone Air from over 500 suppliers [3]. No analog exists in the large-scale computing supply chain. Falk et al. needed a physical teardown of an A100 to find that its heatsink, 98% copper, drives 91% of the GPU's cancer toxicity [7]. No observable connects a workload to the community bearing its embodied health risk. That a single component can dominate, though, hints the gap may be narrowable without full disclosure (Section 4.4).

3.3.2 Spatial averaging. Even where health impact is quantified, the data is averaged globally. Comparative Toxic Units for humans (CTUh), the standard human-toxicity factor in life cycle assessment (LCA), is reported as a single number per kilogram of emitted substance, with elementary flows "treated as global rather than spatially differentiated" [7]. A CTUh value for arsenic emitted near Bayan Obo is identical to a CTUh value for arsenic emitted in Virginia. The exposed populations and exposure pathways differ, and so do the downstream health outcomes. Collapsing these into a single global average discards the information a scheduler or procurement system would need to act.

Table 1: Embodied health impact of one Nvidia A100, converting the toxicity of manufacturing it [7] to DALYs via Huijbregts severity factors [11].

Toxicity	CTUh	DALY/case	DALY
Cancer	3.8×10^{-5}	11.5	4.4×10^{-4}
Non-cancer	6.8×10^{-6}	2.7	0.2×10^{-4}
Total			$\mathbf{4.6 \times 10^{-4}}$

3.3.3 End-of-life invisibility. Even with material composition and spatial resolution, the chain from workload to health outcome breaks at end-of-life. An estimated 80% of e-waste flows are informal and off-record [20]. Server counts are proprietary [33]. Safety thresholds for REE exposure remain largely unestablished. A comprehensive 2024 review of REE human health toxicity opens by acknowledging that the toxicity profile of REE particulate matter "remains largely unknown" [34]. Health impact that cannot be measured against an established baseline cannot be attributed even when every upstream identifier is intact.

4 Toward Health Aware System Design

The gaps identified above point to a research agenda. The following subsections address each in turn, extending embodied and operational impact models, redesigning scheduling signals, and building the measurement infrastructure both require.

4.1 Extending Embodied and Operational Impact Models

One direction is a workload-normalized human-health metric analogous to Schneider et al.'s Compute Carbon Intensity (gCO_2e/ExaFLOP) [25]. Falk et al. already derive per-GPU-hour human toxicity figures from lifetime LCA data [7], which provides a starting numerator. Extending it to CPUs, memory, storage, etc. remains open work.

Such a metric needs a health unit, the way carbon accounting settled on gCO_2e. The *disability-adjusted life year*, DALY, counts health loss as years of healthy life lost, combining years lost to early death with years lived in illness, and it underlies the WHO's Global Burden of Disease accounting. One DALY is one healthy year gone. This can be the natural common denominator for embodied toxicity and operational health impacts.

Turning that toxicity into DALYs is a settled step in life-cycle assessment, not one we invent. Huijbregts et al. fixed the severity factors two decades ago, 11.5 DALYs per cancer case and 2.7 per non-cancer case, now standard in mainstream LCA methods [11]. The embodied health impact of a device is then each toxicity count times its severity [1]

$$D_{\text{embodied}} = \text{CTUh}_{\text{cancer}} \times 11.5 + \text{CTUh}_{\text{non-cancer}} \times 2.7$$

A back-of-envelope illustration. Table 1 works this for one A100. Manufacturing a single card costs about 4.6×10^{-4} DALYs, roughly four hours of healthy life and almost all of it cancer. At the scale of a 100,000-GPU cluster, that climbs to about 46 DALYs, some 46 healthy years of life lost to build those accelerators. This is a floor. It captures one GPU and it stops before the e-waste stage. It carries no geography, a single global average blind to whether it lands on the mining region or the fab town.

The operational term D_{op} takes the same shape, a count of health outcomes times the healthy years each costs. Each count still needs a severity to become DALYs. A death needs the years of life lost ℓ behind it, which depends on the age at death. An illness needs a disability weight w and a duration d, which depend on what each case actually is, a brief attack or an ongoing condition, and the two differ in duration by orders of magnitude. Han et al. provide the counts n, roughly 1,300 premature deaths and 600,000 asthma cases a year across U.S. data centers [10]. In full,

$$D_{\text{op}} = n_{\text{death}} \times \ell + n_{\text{illness}} \times w \times d,$$

but those severities are not reported with the counts and remain open work.

Likewise, mineral-intensity per-workload metrics could serve as a companion signal to energy intensity, surfacing the supply-chain demand that energy reporting alone hides.

Operationally, water-quality health impact metrics are needed. Volume-based water accounting (liters withdrawn per GPU-hour) misses the health impact signal. The same withdrawal from a water-stressed aquifer with pre-existing nitrate contamination produces different outcomes than withdrawal from an unstressed source [15, 35]. An operational water health impact metric would incorporate watershed stress, nitrate loading, and proximity to drinking water sources as spatially-resolved inputs. This would let schedulers treat water quality health impact as a traceable parameter alongside carbon and energy.

4.2 Health Aware Operational Scheduling

The operational phase offers immediate scheduling leverage because workload placement is a runtime decision with no hardware changes required, just an updated objective function.

Geographical load balancing already routes workloads across data centers to minimize carbon or cost [1, 30]. Han et al. [10] extend it with public health signals such as local air quality, grid emission composition, and downwind population exposure. They demonstrate a 26% reduction in public health cost while preserving roughly 3% electricity savings and more than 1% carbon reduction relative to baseline. Pure carbon-aware load balancing applied to a major U.S. data center fleet would cut carbon emissions by 7.2% but *increase* public health cost by 2.8% [36]. Carbon and health are not collapsible into a single optimization signal.

[1] CTUh expresses human toxicity as estimated disease cases [24], which we convert to DALYs using the population-average severities of [11], 11.5 per cancer case and 2.7 per non-cancer case. These averages span disease types rather than the specific conditions each chemical causes, so the result is an order-of-magnitude illustration, not a precise total.

Health impact is also distributed unevenly, not just incurred in total. Recent health aware load balancing treats long-term fairness over cumulative, location-dependent health damage as a first-class objective, capping the worst-case burden any single region absorbs [**?**]. These are operational air-pollution results. Extending the same distributional framing to water and to embodied impact remains open.

Analogously, training jobs routed away from data centers drawing on water-stressed watersheds reduce the competition between compute cooling and local water. The case is strongest during drought, periods of high demand, or in areas with prior water contamination. The open research problem is integrating watershed and water quality signals into schedulers alongside carbon and energy.

The same health-routing idea generalizes to procurement. Hardware can be scored by supply-chain health impacts, and procurement systems can prefer the cleaner options.

4.3 Lifetime, Repair, and Partial Failure in Consideration of Health Impacts

Lifetime extension is a primary lever for reducing embodied health impacts. Human cancer toxicity exhibits the greatest sensitivity to lifespan of any LCA impact category, varying by up to 300% across a 1–4 year lifespan window [7]. This is the health impact analog of Gupta et al.'s framing of embodied carbon as a quantity that amortizes over device life [9], and the intervention is the same. Doubling effective hardware lifetime nearly halves embodied impacts per compute-year. If manufacturing drives 94.5% of an A100's cancer-toxicity impact [7], keeping that GPU in service one extra year delivers a larger toxicity reduction than any single-year operational-efficiency gain. Lifetime extension targets manufacturing where human health impacts are so concentrated; a carbon-only optimizer would underweight such an improvement.

The systems community has built the carbon version of this lever, and the same infrastructure can carry health impact signals. Junkyard Computing repurposes discarded smartphones as cloud microservice nodes and introduces Computational Carbon Intensity as a metric that measures lifetime carbon per unit of useful compute [31]. At hyperscale, GreenSKU designs cloud servers around energy-efficient cores, old DRAM repurposed via CXL, and reused SSDs, reducing emissions per core by 28% relative to currently-deployed Azure servers [32]. A health impact analog is direct. CTUh-per-compute-year tracks the same quantity these systems already amortize.

Extending SSD or HDD lifetime raises device failure rates, and the resulting erasure-coding overhead can erase the embodied-emissions savings [18]. Partial-failure tolerance is the systems-side answer. A GPU with one dead SM or a server with one failed DIMM, kept in service rather than retired, recovers compute-years that would otherwise be spent on new hardware. Both lifetime extension and partial-failure tolerance are existing systems research questions on the carbon side [4]. Manufacturing's share of cancer toxicity (94.5%) far exceeds its share of climate impact (5.5%) [7], so these interventions shift health impacts more than they shift carbon.

Table 2: GHG Protocol carbon scopes mapped to a proposed three-scope health impact framework, with documented impact sources at each level.

	Carbon (GHG Protocol)	Health (proposed)
Scope 1	On-site combustion (diesel, gas)	On-site particulate matter, water-contaminant concentration [15, 27, 29]
Scope 2	Purchased electricity	Premature deaths and asthma cases from grid-sourced fine particulate matter and nitrogen oxides [10]
Scope 3	Supply chain (manufacturing, EOL)	Worker and community exposure across mining, fabrication, and recycling [5, 7, 13]

4.4 Measurement Infrastructure

The interventions above assume a measurement layer that has not been built. Health aware scheduling requires signals. Health impact signals require attribution, which requires disclosure. Four primitives are needed.

Material composition disclosure is foundational. Without it, no workload can be connected to the materials it was built from. A per-SKU elemental requirement, like a nutrition label on a cereal box, would expose that link and let supply-chain health impacts be attributed back to specific extraction regions. Apple's Full Material Disclosure program shows the data can be gathered at consumer-electronics scale, mapping 1,900 materials in iPhone Air [3]. No analog exists in the large-scale computing supply chain at all. Defining the schema is a systems research problem, including which materials, at what granularity, and across which supply-chain tiers. The legal mandate is a policy question. Researchers specify what goes in it. Where full disclosure is not provided, impact could instead be approximated from the few materials that dominate it rather than the complete bill, sidestepping the proprietary detail vendors guard.

Health impact accounting needs spatial resolution. CTUh [24] reports one value per emission. Han et al. [10] resolve operational pollutants spatially. The manufacturing phase has no equivalent notion [7].

Digital lifecycle passports close the loop from mineral source to decommission. End-to-end tracking lets hardware carry its own provenance through deployment and disposal [33]. The EU Battery Regulation 2023/1542 [6] mandates digital product passports for EV and industrial batteries with phased rollout from 2027. The mechanism exists at policy scale. Extending it to AI hardware is a regulatory and engineering question, not a conceptual one.

Three-scope health reporting parallels the GHG Protocol. Table 2 maps each carbon scope to its health analog, with the documented impact sources at each level. The breakdown can be added to previously recommended model cards [19], the per-model reporting artifact already used to disclose intended use, evaluation conditions, and fairness considerations. Han et al. [10] propose extending model cards to carbon and public health, and the three-scope health structure fits the same fields.

5 A Call to Action

Sustainable computing has made real progress tracking carbon. This paper maps documented human health impacts across the large scale computing hardware lifecycle and identifies the measurement infrastructure needed to change that. We call for research into health impact aware systems and for the systematic health impact accounting those systems must optimize against.

We identify four research directions. Define Comparative Toxic Units for humans (CTUh) as a standard health impact metric and extend it beyond GPUs to the full server stack. Section 4.1 outlines what such a metric could look like. We develop per-SKU material composition disclosure standards enabling supply-chain health attribution and show a future vision in Section 4.4, inspired by ingredient labels and Apple's work in consumer electronics. We build spatially-aware health impact signals into schedulers and procurement frameworks. Han et al.'s health-informed load balancing (Section 4.2) shows this for operational air pollution. Extending the same framing to water pollution and embodied impacts and procurement remains open. Finally, we propose to close the end-of-life attribution gap through digital lifecycle passports and three-scope health reporting in model cards and sustainability disclosures. The EU Battery Regulation offers an existence proof at the policy level. Table 2 sketches the GHG-Protocol-to-health-impact mapping.

Much of this work is interdisciplinary, requiring inputs from epidemiologists, engineers, and public health researchers. Systems researchers are not positioned to produce the health evidence, but we are positioned to define the measurement primitives, build the attribution infrastructure, and integrate health impact signals into the design constraints that govern hyperscale computing.

Disclaimer

This paper is not an argument against data centers or AI, nor against any company or facility we name. Our aim is to surface human health as an under-measured dimension of sustainable computing and to motivate the systems research that would track and reduce it. We take no position on any specific project, which would need its own detailed, context specific evaluation. The figures here are illustrative, meant to demonstrate method and motivate measurement.

References

[1] Bilge Acun, Benjamin Lee, Fiodar Kazhamiaka, Kiwan Maeng, Udit Gupta, Manoj Chakkaravarthy, David Brooks, and Carole-Jean Wu. 2023. Carbon Explorer: A Holistic Framework for Designing Carbon Aware Datacenters. In *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2* (Vancouver, BC, Canada) *(ASPLOS 2023)*. Association for Computing Machinery, New York, NY, USA, 118–132. doi:10.1145/3575693.3575754

[2] Saleem H. Ali. 2014. Social and Environmental Impact of the Rare Earth Industries. *Resources* 3, 1 (2014), 123–134. doi:10.3390/resources3010123

[3] Apple Inc. 2026. *Environmental Progress Report 2026.* Technical Report. Apple Inc. https://www.apple.com/environment/pdf/Apple_Environmental_Progress_Report_2026.pdf

[4] Noman Bashir, David Irwin, and Prashant Shenoy. 2023. On the Promise and Pitfalls of Optimizing Embodied Carbon. In *Proceedings of the 2nd Workshop on Sustainable Computer Systems (HotCarbon '23)*. ACM, 1–6. doi:10.1145/3604930.3605710

[5] Diana Maria Ceballos and Zhao Dong. 2016. The formal electronic recycling industry: Challenges and opportunities in occupational and environmental health research. *Environment International* 95 (2016), 157–166. doi:10.1016/j.envint.2016.07.010

[6] European Parliament and Council of the European Union. 2023. Regulation (EU) 2023/1542 of the European Parliament and of the Council of 12 July 2023 concerning batteries and waste batteries, amending Directive 2008/98/EC and Regulation (EU) 2019/1020 and repealing Directive 2006/66/EC. Official Journal of the European Union, L 191, 28.7.2023, p. 1–117. https://eur-lex.europa.eu/eli/reg/2023/1542/oj

[7] Sophia Falk, David Ekchajzer, Thibault Pirson, Etienne Lees-Perasso, Augustin Wattiez, Lisa Biber-Freudenberger, Sasha Luccioni, and Aimee van Wynsberghe. 2025. More than Carbon: Cradle-to-Grave Environmental Impacts of GenAI Training on the Nvidia A100 GPU. (2025). arXiv:2509.00093 [cs.CY] doi:10.48550/arXiv.2509.00093

[8] FP Analytics. 2025. *Artificial Intelligence and the Critical Minerals Crunch*. Technical Report. Foreign Policy Analytics / JCDREAM. https://fpanalytics.foreignpolicy.com/2025/07/18/artificial-intelligence-critical-minerals-supply-chains/ 47-page policy brief.

[9] Udit Gupta, Young Geun Kim, Sylvia Lee, Jordan Tse, Hsien-Hsin S. Lee, Gu-Yeon Wei, David Brooks, and Carole-Jean Wu. 2020. Chasing Carbon: The Elusive Environmental Footprint of Computing. arXiv:2011.02839 [cs.AR] https://arxiv.org/abs/2011.02839

[10] Yuelin Han, Zhifeng Wu, Pengfei Li, Adam Wierman, and Shaolei Ren. 2026. Health-Informed Computing: Estimating and Addressing the Public Health Impact of Data Centers. arXiv:2412.06288 [cs.CY] https://arxiv.org/abs/2412.06288

[11] Mark A.J. Huijbregts, Linda J.A. Rombouts, Ad M.J. Ragas, and Dik van de Meent. 2005. Human-toxicological effect and damage factors of carcinogenic and noncarcinogenic chemicals for life cycle impact assessment. *Integrated Environmental Assessment and Management* 1, 3 (2005), 181–244. doi:10.1897/2004-007R.1

[12] Yankai Jiang, Raghavendra Kanakagiri, Rohan Basu Roy, and Devesh Tiwari. 2025. ThirstyFLOPS: Water Footprint Modeling and Analysis Toward Sustainable HPC Systems. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC '25)*. ACM, 855–869. doi:10.1145/3712285.3759804

[13] Myung-Hee Kim, Hyun Kim, and Domyung Paek. 2014. The health impacts of semiconductor production: an epidemiologic review. *International Journal of Occupational and Environmental Health* 20, 2 (April 2014), 95–114. doi:10.1179/2049396713Y.0000000050 Epub 2013-12-19.

[14] Philip J. Landrigan, Stephan Bose-O'Reilly, Julia Elbel, Gunnar F. Nordberg, Roberto G. Lucchini, Claire Bartrem, Philippe Grandjean, Donna Mergler, David Moyo, Benoit Nemery, Joachim von Braun, Dennis Nowak, and Collegium Ramazzini. 2022. Reducing disease and death from Artisanal and Small-Scale Mining (ASM) - the urgent need for responsible mining in the context of growing global demand for minerals and metals for climate change mitigation. *Environmental Health* 21, 1 (Aug. 2022), 78. doi:10.1186/s12940-022-00877-5

[15] Zeno F. Levy, Bryant C. Jurgens, Karen R. Burow, Stefan A. Voss, Kirsten E. Faulkner, Jose A. Arroyo-Lopez, and Miranda S. Fram. 2021. Critical Aquifer Overdraft Accelerates Degradation of Groundwater Quality in California's Central Valley During Drought. *Geophysical Research Letters* 48, 17 (2021), e2021GL094398. arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021GL094398 doi:10.1029/2021GL094398 e2021GL094398 2021GL094398.

[16] Ying-Chen Lin, Su-Yin Chiang, Tsun-Jen Cheng, and Kuen-Yuh Wu. 2026. Assessing gallium arsenide health risk: critical insights for protecting workers and the environment. *Environment International* 211 (2026), 110253. doi:10.1016/j.envint.2026.110253

[17] Jay H. Lubin, Lee E. Moore, Joseph F. Jr. Fraumeni, and Kenneth P. Cantor. 2008. Respiratory cancer and inhaled inorganic arsenic in copper smelter workers: a linear relationship with cumulative exposure that increases with concentration. *Environmental Health Perspectives* 116, 12 (Dec. 2008), 1661–1665. doi:10.1289/ehp.11515 Epub 2008-07-23.

[18] Sara McAllister, Fiodar Kazhamiaka, Daniel S. Berger, Rodrigo Fonseca, Kali Frost, Aaron Ogus, Maneesh Sah, Ricardo Bianchini, George Amvrosiadis, Nathan Beckmann, and Gregory R. Ganger. 2024. A Call for Research on Storage Emissions. In *Proceedings of the 3rd Workshop on Sustainable Computer Systems (HotCarbon '24)*. ACM. https://hotcarbon.org/assets/2024/pdf/hotcarbon24-final126.pdf

[19] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. Model Cards for Model Reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19)*. ACM, 220–229. doi:10.1145/3287560.3287596

[20] S. M. Parvez, F. Jahan, M. N. Brune, J. F. Gorman, M. J. Rahman, D. Carpenter, Z. Islam, M. Rahman, N. Aich, L. D. Knibbs, and P. D. Sly. 2021. Health consequences of exposure to e-waste: an updated systematic review. *The Lancet Planetary Health* 5, 12 (Dec. 2021), e905–e920. doi:10.1016/S2542-5196(21)00263-1

[21] Ana Pascale, Alicia Sosa, Cecilia Bares, Alicia Battocletti, María J. Moll, Daniel Pose, Adriana Laborde, Hugo González, and Guillermo Feola. 2016. E-Waste Informal Recycling: An Emerging Source of Lead Exposure in South America. *Annals of Global Health* 82, 1 (Jan. 2016), 197–201. doi:10.1016/j.aogh.2016.01.016

[22] Sundar Pichai. 2020. *Our Third Decade of Climate Action: Realizing a Carbon-Free Future*. Google. https://blog.google/company-news/outreach-and-initiatives/sustainability/our-third-decade-climate-action-realizing-carbon-free-future/ Published 14 September 2020.

[23] Brett H. Robinson. 2009. E-waste: An assessment of global production and environmental impacts. *Science of The Total Environment* 408, 2 (2009), 183–191. doi:10.1016/j.scitotenv.2009.09.044

[24] Ralph K. Rosenbaum, Till M. Bachmann, Lois Swirsky Gold, Mark A. J. Huijbregts, Olivier Jolliet, Ronnie Juraske, Annette Koehler, Henrik Fred Larsen, Matthew MacLeod, Manuele Margni, Thomas E. McKone, Jerome Payet, Marta Schuhmacher, Dik van de Meent, and Michael Z. Hauschild. 2008. USEtox—the UNEP-SETAC toxicity model: recommended characterisation factors for human toxicity and freshwater ecotoxicity in life cycle impact assessment. *International Journal of Life Cycle Assessment* 13, 7 (2008), 532–546. doi:10.1007/s11367-008-0038-4

[25] Ian Schneider, Hui Xu, Stephan Benecke, David Patterson, Keguo Huang, Parthasarathy Ranganathan, and Cooper Elsworth. 2025. Life-Cycle Emissions of AI Hardware: A Cradle-To-Grave Approach and Generational Trends. (02 2025). doi:10.48550/arXiv.2502.01671

[26] Arman Shehabi, Sarah J. Smith, Alex Hubbard, Alex Newkirk, Nuoa Lei, Md Abu Bakar Siddik, Billie Holecek, Jonathan Koomey, Eric Masanet, and Dale Sartor. 2024. *2024 United States Data Center Energy Usage Report*. Technical Report LBNL-2001637. Lawrence Berkeley National Laboratory. doi:10.71468/P1WC7Q Peer reviewed..

[27] Md Abu Bakar Siddik, Arman Shehabi, and Landon Marston. 2021. The Environmental Footprint of Data Centers in the United States. *Environmental Research Letters* 16, 6 (2021), 064017. doi:10.1088/1748-9326/abfba1

[28] Brad Smith. 2020. *Microsoft Will Be Carbon Negative by 2030*. Microsoft. https://blogs.microsoft.com/blog/2020/01/16/microsoft-will-be-carbon-negative-by-2030/ Published 16 January 2020.

[29] Ryan Smith, Rosemary Knight, and Scott Fendorf. 2018. Overpumping leads to California groundwater arsenic threat. *Nature Communications* 9, 1 (2018), 2089. doi:10.1038/s41467-018-04475-3

[30] Thanathorn Sukprasert, Abel Souza, Noman Bashir, David Irwin, and Prashant Shenoy. 2024. On the Limitations of Carbon-Aware Temporal and Spatial Workload Shifting in the Cloud. In *Proceedings of the Nineteenth European Conference on Computer Systems (EuroSys '24)*. ACM, 924–941. doi:10.1145/3627703.3650079

[31] Jennifer Switzer, Gabriel Marcano, Ryan Kastner, and Pat Pannuto. 2023. Junkyard Computing: Repurposing Discarded Smartphones to Minimize Carbon. In *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS '23)*. ACM, 400–412. doi:10.1145/3575693.3575710 Distinguished Paper Award.

[32] Jaylen Wang, Daniel S. Berger, Fiodar Kazhamiaka, Celine Irvene, Chaojie Zhang, Esha Choukse, Kali Frost, Rodrigo Fonseca, Brijesh Warrier, Chetan Bansal, Jonathan Stern, Ricardo Bianchini, and Akshitha Sriraman. 2024. Designing Cloud Servers for Lower Carbon. In *Proceedings of the 51st Annual International Symposium on Computer Architecture (ISCA '24)*. IEEE. doi:10.1109/ISCA59077.2024.00041

[33] Peng Wang, Ling-Yu Zhang, Asaf Tzachor, and Wei-Qiang Chen. 2024. E-waste challenges of generative artificial intelligence. *Nature Computational Science* 4, 11 (Nov. 2024), 818–823. doi:10.1038/s43588-024-00712-6

[34] Wenyu Wang, Yanfang Yang, Donglei Wang, and Lihua Huang. 2024. Toxic Effects of Rare Earth Elements on Human Health: A Review. *Toxics* 12, 5 (2024), 317. doi:10.3390/toxics12050317

[35] Mary H. Ward, Rena R. Jones, Jean D. Brender, Theo M. de Kok, Peter J. Weyer, Bridget T. Nolan, Cristina M. Villanueva, and Simone G. van Breda. 2018. Drinking Water Nitrate and Human Health: An Updated Review. *International Journal of Environmental Research and Public Health* 15, 7 (2018), 1557. doi:10.3390/ijerph15071557

[36] Adam Wierman and Shaolei Ren. 2025. *We Need to Talk About AI's Impact on Public Health*. IEEE Spectrum. https://spectrum.ieee.org/data-centers-pollution Accessed: 2026-05-04.