basel action network

Champions of environmental health & justice

The Coming AI Waste Wave

A White Paper in 4 Parts

PREPARED BY

Jim Puckett, Basel Action Network · September 2026

BASEL ACTION NETWORK

INTRODUCTION — We know better, don't we?

Over the course of the last fifty years we have learned much about chemicals and waste management. The lessons included the principle of entire life-cycle management of products, from cradle to grave. We next learned that even that principle was inappropriately linear, and ideally we would manage materials/waste from cradle-to-cradle in a circular waste = food paradigm. We learned that the keystone of true circularity is product design and that first we should assess societal need for products, and then, once necessity is established, use only those materials and methods that will eliminate harm (e.g. toxicity). We learned that products should be designed for end-of-life and not for the dump, and that we would seek designs ensuring product longevity, durability, upgradability, and repairability for reuse. Secondarily, when the originally intended lifespan of a product was fully spent, our products would be designed to be safely and completely recycled so that resources could be easily recovered and remade into new products for a second, third, and ideally endless series of reincarnations of other well-designed products. This working our way down the ladder of preferential steps of managing product life-cycle is known as the waste management hierarchy. These are industry lessons learned.

They were learned the hard way. The world had borne witness in the 1970s and 1980s to a legacy of accumulated toxic waste dump sites that had to be cleaned up at enormous cost. We had also witnessed what became the phenomenon of global waste dumping or "waste colonialism", when actors sought to avoid the high costs of proper waste management by exporting the effluent of the affluent to populations and environments in developing countries. E-waste certainly became central to that horror show.

In the last three decades, the electronics industry also had plenty of hard bumps in the road based on their own design catastrophes. The industry lurched from the reckless use of PCBs in the 1970s, to the use of leaded solders, to the design of cathode ray tubes, emitting so much radiation that leaded glass had to be deployed to prevent household harm, while utilizing toxic cadmium phosphors in those same CRTs. And then later, we realized nobody had planned the mass disposal of the same devices as we had to find a means to dispose of them by the millions upon the arrival of flat-screen monitors and TVs. And then, of course, the industry turned around and repeated the same toxic pattern by placing mercury phosphors into the fluorescent backlighting of LCD monitors, which were then inserted into the screens in a manner making them almost impossible to remove without breaking the lamps and releasing the mercury.

In hindsight, a case can be made that the electronics industry's track record of learning from the lessons of design-for-no-harm has not been stellar, despite a push in the last two decades that has resulted in positive gains to remove brominated flame retardants, leaded solders, and mercury batteries and switches. The recent initiatives to rid products of problematic plastics like PVC and single-use packaging is a positive step and should be applauded, even if these gains were largely driven by legislation or pressure from civil society and not by corporate planning or desire to succeed in the ESG realm. We can remember the scorecard advocacy by Greenpeace and the Computer TakeBack Campaign ranking Apple last among 14 major electronics OEMs in sustainability principles and green design.[i] Apple famously responded in 2007 with the corporate manifesto "A Greener Apple", which highlighted their aim to catch up to the general movement towards Green design with progressive toxic material phase-outs.

The electronics industry, although a reluctant participant perhaps, was nevertheless marching in competition to "win" on design for environment principles.

Today, as the AI horserace bell has rung, and the major hyperscalers are already many strides from the starting gate of one of the most competitive and dramatic industrial transformations in history, the question must be asked: have we learned anything at all about waste and toxicity prevention, or are these late-learned sustainability principles now considered an expendable handicap, and, in practice, the first casualty, in the gallop towards AI dominance? That is the question this white paper series puts to the evidence.

Part I of the AI E-Waste Wave Series will cover the amount of e-waste anticipated from the coming AI buildout.

White Paper Part I: How Big Is The AI Waste Wave?

Executive Findings:

1. Industry insiders project an expenditure of 7 trillion dollars from 2025 to 2030 just to begin the build-out of the necessary infrastructure to power AI. $7 trillion is roughly the cost of ending world hunger for the next 75 years being spent in five. Or, the sum of money could instead fund a $10,000 higher-education trust for every child born on Earth (700 million children) during those 5 years.
2. That buildout to 2030 will equate, in energy terms, to roughly 219 gigawatts of total data center capacity — 80% more than what existed as recently as 2025. That is more than enough energy to power every home in the United States combined.
3. The AI buildout, now well underway, represents a paradigm shift in computing's material footprint. The exhaustion of Moore's Law, which delivered exponential performance gains by the steady improvement of individual chips (vertical scaling), has forced the industry to scale horizontally, deploying massive quantities of associated electronic infrastructure. This architectural shift across the planet will generate an unprecedented tsunami of rapidly obsolescent electronic equipment.
4. Until now, environmental scrutiny of AI data centers has focused almost entirely on carbon footprint and freshwater consumption. The e-waste impacts, its likely toxicity, and its environmental justice implications have to date largely been ignored. This head-in-the-sand outlook exists even as we know that currently the world safely manages only one-fifth of the e-waste it produces, with much of it exported to developing countries to be mismanaged in substandard conditions in poorer communities.
5. The data center boom is well underway. 2030 is but 4 years away. Roughly 100 GW of new data center capacity is anticipated between 2026 and 2030, effectively doubling today's installed base. The U.S. alone has 4,871 facilities in the pipeline (821 under construction and 4,050 announced). Do the communities where these are being built realize that each of these centers will become point sources of hazardous waste, requiring disposal of massive amounts of this toxic waste every 2–3 years?
6. Previous quantitative AI e-waste estimates have underestimated the coming volumes as they focused overwhelmingly on servers and accelerators, which represent just 13% of a data center's electro-mechanical infrastructure. This study identifies five equipment categories (including networking, power distribution, storage/backup, and cooling), all of which total approximately 70,000 tonnes per GW of capacity. The previously uncounted 87%, the vast majority of which is also defined as e-waste, has never appeared in any AI e-waste projection of which we are aware.

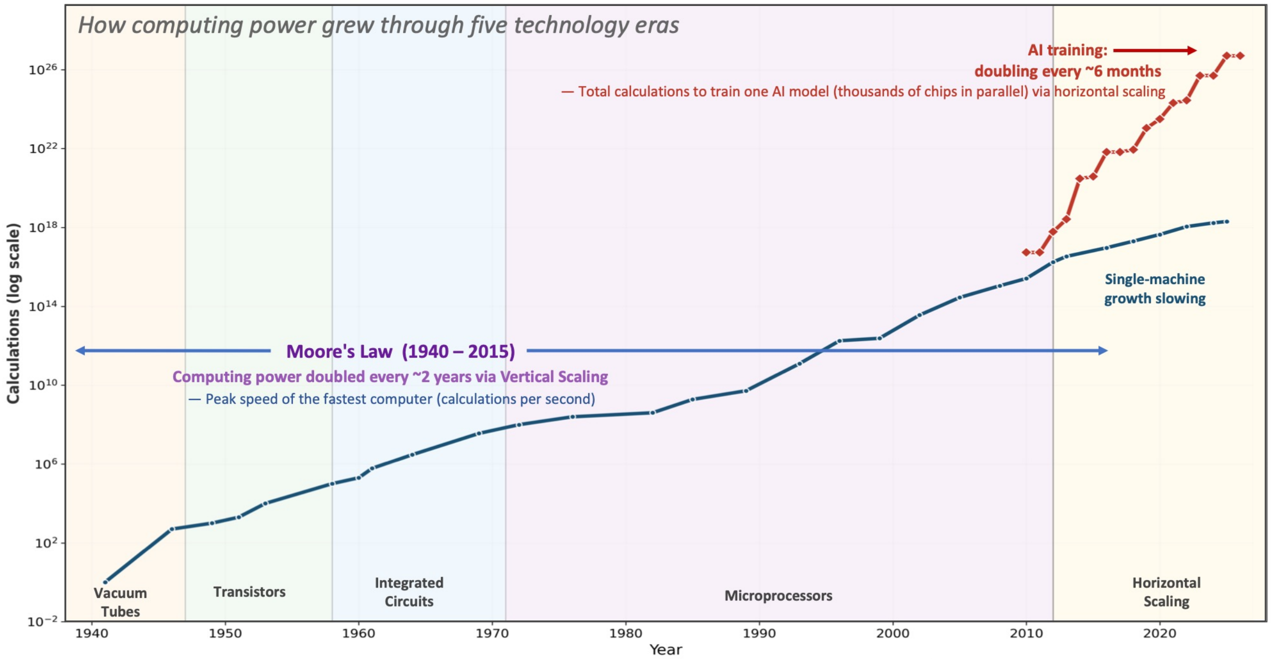
7. The data center industry's "cattle not pets" operational doctrine and rapid GPU generational turnover are already compressing AI equipment lifespans to 2.5–5 years, shorter than the normal lifespan of servers and related equipment would require, meaning the same infrastructure becomes waste far sooner than conventional replacement cycles would predict.
8. The paper introduces another overlooked AI e-waste category we call the AI Waste Contagion. This is electronic equipment retired outside data centers, induced when global AI adoption forces rapid replacement of PCs, phones, edge devices, and telecommunications infrastructure at home and at work. In both our conservative and aggressive models of this phenomenon, this downstream contagion (16.2–30.7 Mt/yr) is expected to exceed the waste generated inside data centers themselves (15.5 Mt/yr).
9. By 2030, this study's estimate of AI-driven electronic equipment being retired (8.6–13.1 Mt/yr) is roughly 40–60 times higher than the most widely cited academic projection (de Vries and Gao, 2026), primarily because prior work counted only servers and GPUs and did not consider the Contagion effect.
10. Counting all sources (conventional and AI-driven), this study projects a total global e-waste generation of 196–211 Mt/yr by 2050 – more than tripling the roughly 67 million tonnes the world is producing today. Of the projected 2050 totals, 31–46 Mt/yr is attributable to AI waste alone. The cumulative AI-driven electronic equipment being retired between 2025 and 2050 (395–617 Mt) would fill 15–23 million 40-foot shipping containers. Placed end to end, they would circle the Earth about six times.

SECTION 1 — From Moore’s Law To Hyperscaling

Historically, computing increased its power by making individual components faster and more capable. The effort to improve involved a continual push for faster CPUs, more RAM per chip, and higher clock speeds. The guaranteed performance predictor, which became observable even before 1940 and extended to about 2015 of IT growth, was known as Moore's Law. This "law" stated that component improvement in terms of the number of components that could be packed into a single circuit, and thus raw computing power, using such inventions as vacuum tubes, transistors, integrated circuits, and microprocessors would roughly double every two years. This type of power efficiency growth came to be known as vertical scaling and it stood the test of time for almost 80 years.

But then we reached the barrier of physics – where the use of photolithography to print transistor circuits into silicon chips reached the limit of size, and electron tunneling and the wavelengths of light itself became barriers – the Moore's Law wall. It was no longer possible to print more transistors in a confined space. Further, the heat generated with so much electronic power operating at higher and higher densities also became a growing physical barrier. The rate of improvement slowed dramatically, and the cost per transistor stopped falling. And yet industrial demand required ever more computing power. How could industry continue to scale as necessary to get faster and more powerful, in particular to satisfy the needs of Artificial Intelligence (AI)?

Figure 1. 80 Years of Moore's Law and Then – The Pivot to Horizontal Scaling



© Basel Action Network

Sources: Computing milestones from Top500.org and Wikipedia. AI training compute from Epoch AI Notable AI Models dataset (CC BY).l

The industry response to not being able to improve each hardware unit by virtue of faster chips has run along two pathways. First, a shift in chip type became necessary: general-purpose Central Processing Units (CPUs) are now augmented by specialized silicon-based products, including Graphics Processing Units (GPUs), Tensor Processing Units (TPUs), and other custom AI accelerators, purpose-built for the very narrow use of maximizing AI-specialized workloads. Second, and most consequential, is a paradigm shift in scaling strategy: rather than each generation of a single chip getting faster (vertical scaling), AI infrastructure scales by deploying incomprehensible quantities of these specialized chips working in parallel (horizontal scaling).

The intelligence emerges from the quantity and interconnection of hardware, not the sophistication of any single unit. The industry is able to muscle-up simply by replicating the most powerful chips horizontally "arm-in-arm", "hyperscaling" across the planet. The only limit is the size of our world.

And so it is that "data centers" or "server farms" are planned to be built rapidly, and seemingly everywhere.

Both of these shifts signal a devastating consequence of increased generation of electronic waste.

1. **Specialty, Proprietary Obsolescence:** New, proprietary, specialized chips and related software will not easily find secondary markets for reuse and when a new generation arrives, it doesn't retire one better machine, but hundreds of thousands of parallel machines all at once, in a wave of hyper-obsolescence.

2. **Volume:** The sheer number of units planned to be deployed horizontally is staggering compared to anything previous.

A recent article from McKinsey and Company[ii], has placed the total Capital Expenditure needed to build-out the planned data centers across the globe just from 2025 to 2030 at almost 7 trillion dollars. 7 trillion dollars.

To grasp what $7 trillion means: it is roughly the cost of ending world hunger for the next 75 years — being spent in just five.[iii] Or consider that the same $7 trillion, distributed evenly across those same five years, could instead fund a $10,000 higher-education trust for every child born on Earth (700 million children).[iv]

That buildout to 2030 will equate, in energy terms, to roughly 219 gigawatts of total data center capacity — 80% more than what existed as recently as 2025. That is more than enough energy to power every home in the United States combined.[v] And as we will learn, gigawatts correlate closely with hardware consumption as more capacity means more servers, and more servers means more decommissioning — with more decommissioning meaning more waste looking for proper ethical management.

The $7 trillion 5-year data center buildout figure from McKinsey, which includes not just electronics hardware (that alone counts as $4.3 trillion), is the monetary expression of the paradigm shift that we are facing with AI on many fronts. And we are talking here about scaling a build-out across many years, not just the next 3.5. We are talking about many societal impacts – on the social compact of labor and wealth, on global governance, on warfare. And from an environmental perspective, on energy consumption, on land use, on water consumption, on climate change. But seriously overlooked to date is the impact from the mass of AI waste to be generated. This white paper intends to focus on this impact, which until now has been a little-explored subject.

The horse race we are each now experiencing to dominate AI may be creating a culture of throwing caution and wisdom to the wind as some of the most powerful companies on earth kick furiously towards a hyper-profitable finish line. To get there most expediently the cloud industry hasconsistently argued in their own jargon for a hardware creed of spin them up, and when they break, spin them down and replace them. It sounds eerily like the "take, make, and dispose" mentality rightly disparaged as obsolete linear thinking, now discredited in favor of the "circular economy".

But it was formalized as cloud doctrine by Bill Baker, a Distinguished Engineer and senior technical leader at Microsoft, who coined the phrase "cattle v. pets" to describe the new era of cloud computing. His point was later explained and popularized by another cloud computing guru named Randy Bias, CEO of Cloudscaling (later acquired by EMC/Dell) in this way:

> ***"In the old way of doing things, we treat our servers like pets, for example Bob the mail server. If Bob goes down, its all hands on deck. The CEO can't get his email and its the end of the world. In the new way, servers are numbered like cattle in herd...when one server goes down, its taken out back, shot, and replaced on the line."*** [vi]

The doctrine of "cattle v. pets" became arguably the single most influential metaphor in cloud computing.

While the metaphor originally describes the interchangeability of logical instances rather than the physical disposal of hardware, the operational culture it created — in which individual machines receive no special maintenance, repair, or life-extension — has material consequences for waste generation. It signifies a departure from the circular-economy principle that equipment should be maintained, repaired, and reused.

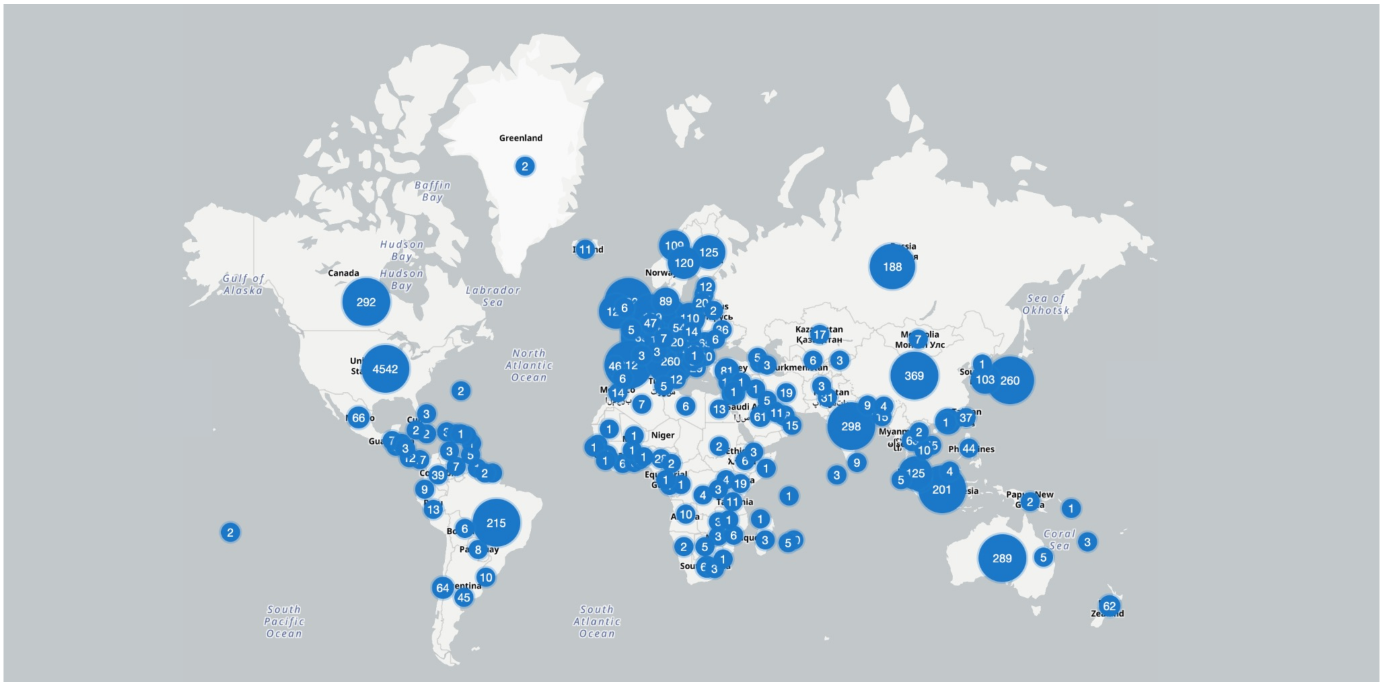
And as wrong-headed as this might be, it is the way we are all heading pell-mell. And, we all find ourselves unwitting participants in the AI horserace, as IT users everywhere find themselves using AI whether they would like to or not, and yet almost always, without being aware of the greater consequences of doing so. The entire social exercise/experiment has, despite our collective hard-earned wisdom, reasserted the primacy of hardware as consumable rather than needing to be durable and then reusable. And, the amount of AI e-waste we can now logically predict is truly massive in comparison to what we already have realized to be an e-waste crisis.

So how much are we talking about? To answer that question, we first need to understand just what is inside of a data center – its anatomy and how many we intend to build. What is the machinery needed to make this new revolution tick? What is in these things? And then, what does hyperscaling across the planet really look like?

The 6 leading hyperscalers representing 6 of some of the most powerful companies on earth.

SECTION 2 — Data Centers: Anatomy And Footprint

The Global Footprint



Active Data Centers Globally as of August 2026. Not all of these are used for AI but many are. Source: Programs

As of August 2026, there are 12,259 active data centers across 179 countries. There are roughly 11,800 data centers operating worldwide as of early 2024, including at least 1,189 large hyperscale AI facilities (Q1 2025). The U.S. leads with 5,388 (~45% of global total), followed by Germany (529), the UK (523), and China. Global installed capacity stood at 122.2 GW in 2024, with the U.S. and China together accounting for 70%; the EU holds 11.9 GW.[vii]

There are over 8.9 GW worth, currently expected to be under construction across 105 projects by the end of 2026, with 47 already under construction.[viii] In total, over 23 GW was under construction globally as of September 2025, three-quarters in the U.S.[ix] Five facilities at 1 GW or more are expected to come online in 2026.[x]

Roughly 100 GW of new data center capacity is anticipated between 2026 and 2030 — effectively doubling today's active centers.[xi]

The U.S. has 4,871 facilities in the pipeline (821 under construction, and 4,050 announced).[xii] But Europe and Asia are accelerating too. The UK and France have announced major AI infrastructure projects, and Asia-Pacific's data center development pipeline is accelerating sharply, with one of the strongest half-year increases on record.[xiii]

The Site Footprint

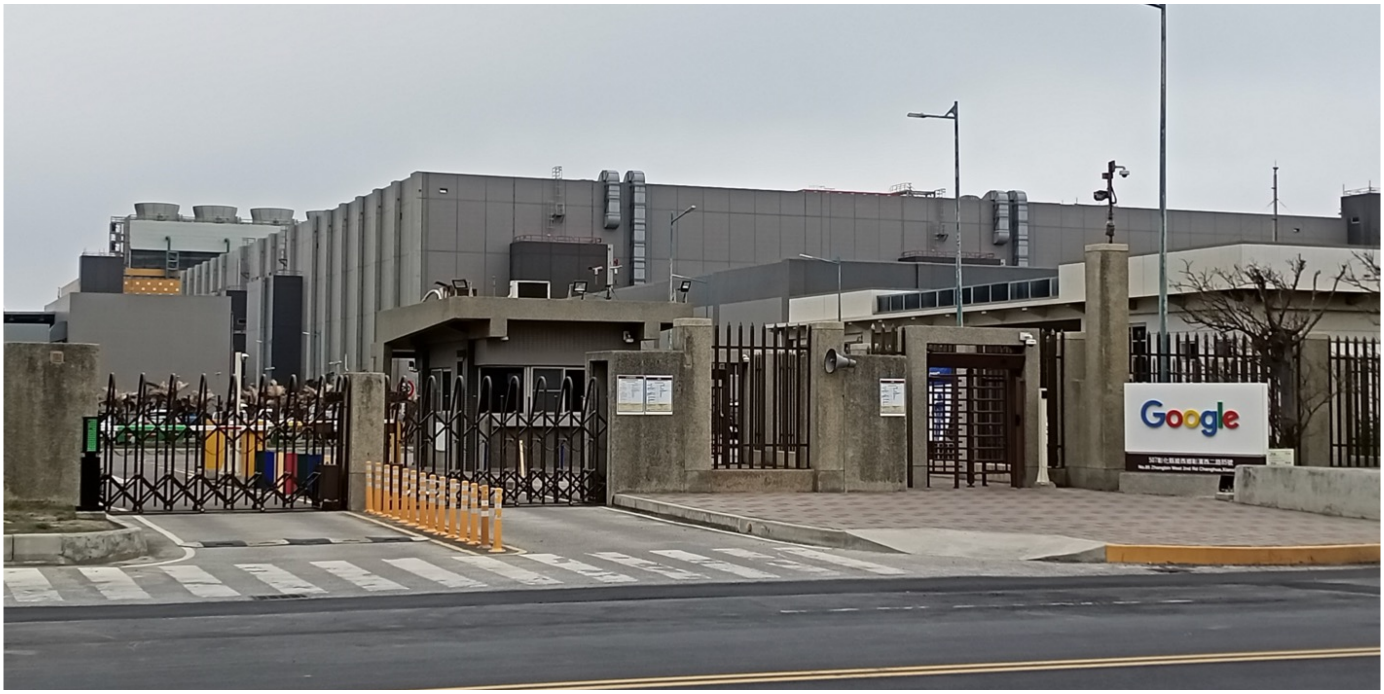
A typical data center building today covers roughly 100,000 square feet, but that average is rising fast as new AI-focused facilities come online at far larger scale. The average data center parcel grew 144% between 2022 and 2024, reaching 224 acres.[xiv] Modern hyperscale campuses now routinely exceed 200–500 acres, with some crossing 1,000 acres to future-proof for decades of expansion.

Meta's Richland Parish campus in northeast Louisiana spans 2,250 acres with over 4 million square feet of space across nine planned buildings, at a cost of $10 billion. It is projected to be the world's largest individual campus by the end of 2027.[xv] Vantage Data Centers' "Frontier" campus in Shackelford County, Texas, covers 1,200 acres with 3.7 million square feet across 10 planned data centers, at a cost of $25 billion, with the first building scheduled for delivery in the second half of 2026.[xvi] Microsoft's Mount Pleasant campus in Wisconsin, built on land originally set aside for a Foxconn manufacturing hub,[xvii] started at 315 acres; the company now owns over 2,000 acres in Mount Pleasant[xviii] and in January 2026 received approval for 15 additional data center buildings across two new campuses covering nearly 9 million square feet, with a taxable value exceeding $13 billion.[xix]

Google Data Center in Council Bluffs, Iowa, USA. 2017. [Chad Davis.]

These latest examples are not anomalies but represent the new baseline for AI-era data center construction. Inside each of these campuses, once completed, will be embedded thousands of tonnes of copper, steel, concrete, cooling and power distribution infrastructure, networking, energy storage, and computing hardware — much of which will require decommissioning within years, not decades, given the anticipated pace of AI hardware obsolescence.

Core Components of a Data Center That Will Become Waste



Exterior view of the main gate, Google's Taiwan data center in Xianxi, Changhua, 7 March 2021, cropped top and bottom.
Source: Kai3952.

AI data centers differ meaningfully from a more traditional data center, which runs general-purpose workloads on CPU-based servers with air cooling. In contrast, AI data centers typically run GPU clusters drawing 40 to 140 kilowatts per rack, with each rack able to accommodate 8-36 GPU servers, versus 3 to 5 kilowatts for standard server racks.[xx] An AI rack typically includes GPU servers, high-wattage intelligent power distribution units (PDUs), high-speed networking switches, and integrated liquid or hybrid immersion cooling systems to manage the heat output of the AI chips. Existing traditional facilities generally can't simply be repurposed for AI because AI workloads need roughly 10 times more computing power than traditional applications,[xxi] as well as liquid cooling systems.

While most people expect AI data centers to be full of servers on racks and the connections between them, the hardware inside the building is made up of far more than this - and most of this will face rapid obsolescence due to technical innovation, and most will be considered as e-waste under the Basel Convention and its recent amendments[xxii] at end-of-life if it cannot be reused (more on this in Part 2 of this series). This hardware we categorize in 5 baskets:

1. **Accelerators, servers, racks**
2. **Networking equipment**
3. **Power supply and distribution**
4. **Data storage and backup**
5. **Cooling systems**

Accelerators, Servers, Racks

AI workloads demand massive parallel computation — the same mathematical operation performed billions of times simultaneously. The processors in the form of chips purpose-built to do this are called accelerators.[xxiii] As of early 2026, NVIDIA's H100 and B200 GPUs chips are the most common accelerators currently in deployment, though Google uses its own chips optimized for AI known as TPU v5p hardware, while Meta uses its own MTIA chips, and Amazon uses Trainium.[xxiv] What they have in common is that they are purpose-built, application-specific, and largely non-interchangeable – thus those designed for frontier applications have a narrower capacity for secondary-use options and a shorter window of useful life than general-purpose servers.[xxv]

Modern Data Center Server Room. Glass-fronted Server Cabinets Running Artificial Intelligence Workloads in a Data Center Corridor, 2026. Source: Gorodenkoff

Training a single large model program requires thousands of accelerators running in parallel for weeks. Physically, inside each server are these accelerator chips, on printed circuit boards, as well as motherboards with CPUs, DIMM memory modules (RAM), SSD/NVMe storage drives, network interface cards, power supply units, and to reduce heat – heatsinks/cold plates.

A single NVIDIA DGX H100 server (8 GPUs) weighs approximately 130 kg;[xxvi] a complete GB200 NVL72 rack weighs roughly 1,360 kg (3,000 lbs).[xxvii] AI racks are significantly heavier than standard racks due to dense GPU configurations and liquid cooling components, often requiring floor reinforcement.[xxviii]

Networking equipment

An AI data center's networking infrastructure includes top-of-rack switches, spine switches, InfiniBand or Ethernet fabric switches, and optical transceivers, along with the cabling that connects them all. Networking accounts for only about 5% of a data center's electricity consumption.[xxix] But, high-speed networking gear is also on a compressed replacement cycle: the industry-wide upgrade from 40/100G to 400/800G equipment is forcing replacement of switches, optics, and cabling well before the hardware has physically failed. By mass, networking equipment is the lightest of the four categories. Individual switches weigh 10–30 kg, and cabling adds roughly 225–360 kg of copper per rack[xxx] — but it will likely need replacement or "refreshing" more frequently than power or cooling infrastructure.[xxxi]

Power supply and distribution

The power chain runs from utility-scale transformers and medium-voltage switchgear through UPS (uninterruptible power supply) systems, power distribution units (PDUs), and busbars down to individual rack-level power supplies. This is by far the heaviest hardware category to be found in a data center.

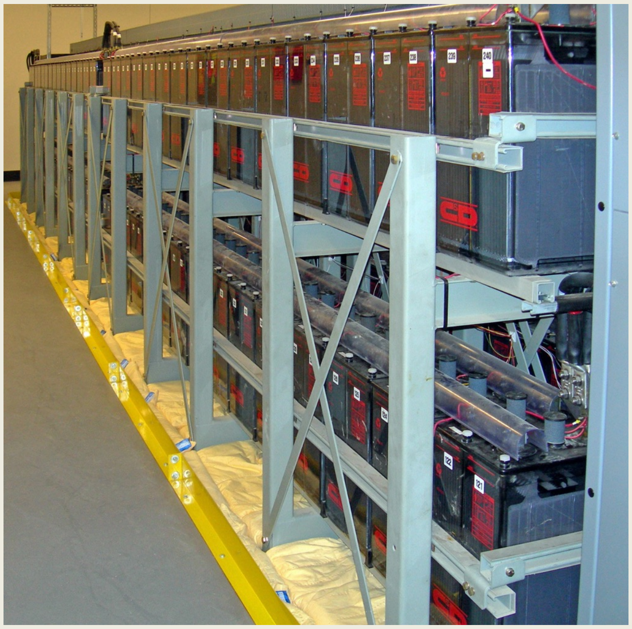
A single 100 MW data center requires approximately 2,700 tonnes of copper for cabling and busbars alone.[xxxii] UPS systems weigh 800–3,000 kg per unit,[xxxiii] and a megawatt-scale facility deploys many, not counting the battery banks behind them.

AWS Expanding Data Center, Boardman, Oregon, 2023.   Source: [Tedder]

When a facility upgrades from 5–15 kW rack density to the 50–140 kW required for AI workloads,[xxxiv] much of this power infrastructure must be ripped out: it cannot simply be uprated in place.

Backup Power Systems

Current data centers run battery-backed-up UPS systems, a significant potential e-waste stream separate from server hardware. The industry standard has been VRLA (valve-regulated lead-acid) batteries, but AI facilities are shifting to lithium-ion (LiFePO4), which lasts 10+ years versus 3–6 for lead-acid, and weigh 60–80% less.[xxxv] Further, they handle fluctuating AI demands more effectively than lead batteries. Hyperscaler facilities require 10–50 MWh of battery storage each.[xxxvi] At 30–50 Wh/kg,[xxxvii] a 10 MWh VRLA installation weighs 250–330 tonnes, predominantly made up of lead and sulfuric acid. An equivalent lithium-ion installation weighs roughly one-third as much but introduces thermal runaway risk and cells containing lithium, cobalt, nickel, and manganese compounds[xxxviii] requiring specialized handling. Both chemistries are hazardous e-waste under the Basel Convention unless a case can be made for their direct reuse.[xxxix] Given the short replacement cycles, a large facility may generate hundreds of tonnes of battery waste every few years. Backup diesel generators (typically 40 to 56 per 100 MW facility, each weighing approximately 16–18 tonnes)[xl] and bring additional regulated waste streams: diesel fuel, lubricating oils, and diesel exhaust fluid.



A bank of batteries in a large datacenter, used to provide emergency power until diesel generators can be started. 2007. Source: Jelson25

Roof of a data center featuring cooling towers and backup generators, November 2025 by Rsparks3

A technician working on an immersion cooling system in a data center for cryptocurrency mining, cloud services and AI computing in a large, temperature controlled warehouse in a remote location in Stutsman County, North Dakota. The server is submerged in a tank of a non-conductive fluid that draws away heat more efficiently than air, which lowers the energy and water consumption of the facility. 2024. Source: halbergman

Cooling Systems

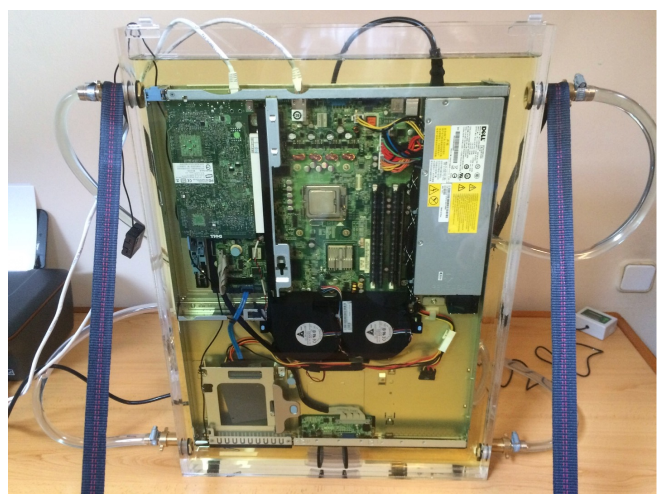
Cooling is the second-heaviest infrastructure category after power, and the one evolving most rapidly. A traditional air-cooled data center uses CRAC (Computer Room Air Conditioning) or CRAH (Computer Room Air Handler) units connected to chillers and cooling towers. A 100-ton air-cooled chiller weighs roughly 4,900 kg.[xli] These systems use refrigerants, commonly R-410A (GWP 2,088) and R-134a (GWP 1,430), both HFCs (hydrofluorocarbons) and potent greenhouse gases regulated under the Kigali Amendment to the Montreal Protocol,[xlii] the EU F-Gas Regulation,[xliii] and in the United States under the AIM Act.[xliv] These refrigerants are targeted for phasedown, with the EU mandating complete HFC elimination by 2050.

What this means for AI data centers is the risk of venting and releases when cooling systems are retired or refrigerants are replaced.

This is especially relevant as AI facilities are now shifting to direct-to-chip liquid cooling, using propylene glycol/water coolant circulated through CDUs (Coolant Distribution Units).[xlv] When a traditional facility is retrofitted for AI,[xlvi] much of the existing cooling infrastructure becomes waste: chillers, CRAC/CRAH units, ductwork, raised-floor plenums, the HFC refrigerants they contain before a new liquid cooling system is installed.

The replacement coolant itself requires managed disposal. Propylene glycol is generally considered low-toxicity, but the corrosion inhibitors and biocides blended into data center coolant formulations may push classification toward regulated industrial wastewater, requiring testing before discharge.[xlvii]

Facilities that adopted two-phase immersion cooling face a far worse situation. The fluorocarbon dielectric fluids used (such as 3M's Novec line, discontinued in 2025) are PFAS "forever chemicals." In April 2024, the EPA designated PFOA and PFOS as hazardous substances under CERCLA (Superfund), and PFOS is listed under Annex B of the Stockholm Convention on Persistent Organic Pollutants.[xlviii]



Dell server immersed in coolant product 3M Novec 7100 — which is a PFAS based fluorocarbon fluid. 2015. Source: Submer

The Basel Convention provides additional guidance on the management and destruction of PFAS-containing wastes.[xlix] In the EU, a sweeping restriction on PFAS under the REACH regulation — covering manufacture, use, and placing on the market — is expected to proceed through the REACH regulatory process, with a broad restriction proposal under development.[l] Operators routing fluorinated immersion fluid through conventional waste channels face potential enforcement exposure under both TSCA and CERCLA in the United States.[li]

AI Waste Categories by Percentage

Based on facility-level data for a reference 100 MW AI data center, we estimate a total equipment mass of approximately 7,000 tonnes, distributed across five categories as follows, in order of relative mass, largest to smallest:

1. **Power Supply and Distribution (~34%, ~2,380 tonnes).** This is the heaviest category. It includes transformers, switchgear, UPS units, PDUs, cabling, and busbars.[lii] Together with Backup Power Systems (Category 4), these two categories account for approximately 49% of total equipment mass, based on our facility-level analysis.
2. **Cooling Systems (~35%, ~2,450 tonnes).** The second-heaviest category. In air-cooled legacy facilities, this includes chillers, CRAC/CRAH units, cooling towers, ductwork, and raised-floor plenums. In liquid-cooled AI facilities, it includes CDUs, piping, manifolds, and coolant. As detailed above, this category is evolving most rapidly, and facility retrofits for AI turn much of the existing cooling plant into waste.
3. **Accelerators, Servers, and Racks (~13%, ~910 tonnes).** Computing hardware is dense and valuable per kilogram but lighter in aggregate than most assume. At roughly 1,360 kg (3,000 lbs) per fully populated AI rack,[liii] across approximately 650 racks in a 100 MW facility,[liv] total computing mass is approximately 900 tonnes – less than half the mass of either power or cooling infrastructure.
4. **Backup Power Systems (~15%, ~1,050 tonnes).** Battery banks (VRLA and increasingly lithium-ion) account for 250–330 tonnes at 10 MWh capacity; diesel generators (40–56 units at 16–18 tonnes each) account for 640–1,008 tonnes.[lv]
5. **Networking Equipment (~3%, ~210 tonnes).** The lightest category in absolute mass but the one with the most frequent replacement cycle, driven by rapid bandwidth transitions (400G → 800G → 1.6T).[lvi]

While public discussion of AI waste focuses almost entirely on computing hardware — servers and GPUs — it is the first and second categories above that dominate the mass hierarchy. Power distribution, cooling systems, backup power systems, and networking equipment all are almost certain to qualify as controlled electronic waste under the Basel Convention,[lvii] now newly amended to include recent e-waste amendments that went into force in 2025. The new amendments are comprehensive and now define controlled e-wastes as any whole piece of equipment, or a component thereof containing electronic circuits, as well as wastes arising out of processing such equipment or components, with the only exception being fully tested, fully functional equipment or wastes for which, after cleaning, sorting and analysis, it can be demonstrated they meet the criteria of a Basel Annex IX (non-hazardous) listing such as B1010 (non-dispersible, uncontaminated scrap metal).

Regarding the fully functional and Annex IX exemption, we acknowledge that some percentage of the weight of the totality of data center infrastructure, after being sorted, cleaned and prepared (e.g. uncontaminated coolant hoses, all metal racks without power strips) may eventually become exempt from the legal definition of e-waste. But that will only occur after it arrives at the waste management facility. That cannot be expected to happen at the Data Center. Rather, the refresh will all go on to the truck as a mixed load. Under Basel, a mixed load containing hazardous e-waste is managed as controlled until demonstrated it is otherwise.

The burden of proof is on the holder to demonstrate that specific materials qualify for the Annex IX exemption — and that determination can only happen after processing, including testing, and separation at a licensed and ideally certified facility. In Part 2 of this paper we will discuss that fraction, as well as possible wastes that can be deemed non-waste via reuse, including via repair, refurbishment and repurposing.

The implication, however, is abundantly clear that "AI e-waste" is much larger than people might first assume. Even the informed public imagines discarded GPUs. The reality is that for every rack of GPUs, there are many hundreds of kilograms of copper cabling and copper busbar (2,700+ tonnes in a 100MW facility),[lviii] tonnes of cooling infrastructure, power distribution equipment, networking gear, and batteries — all of which turn over on compressed refresh cycles driven by the same AI upgrade treadmill. And virtually all of it must, at least at the outset, be considered electronic waste.

SECTION 3 — Calculating The Scale Of The Problem

The UN Global E-Waste Monitor Baseline

For many years now, we have all read that e-waste is the fastest-growing waste stream on earth. The UN Global E-Waste Monitor's latest estimates place the amount generated in 2022 at 62 million metric tonnes of e-waste and, if extrapolated[lix], projects approximately 165 million metric tonnes per year by 2050 under their conventional e-waste growth scenario. Of the current amount produced, according to the Global E-Waste Monitor, only 22.3 percent is said to be properly recycled.

These traditional e-waste volumes, current and projected, can already be said to be a crisis in full swing even before a single AI server even enters the calculation. As noted in the Monitor:

> ***e-waste management remains a cause for concern and requires immediate attention and action, as the amount of e-waste has grown 5 times faster than compliant collection and recycling since 2010....most e-waste is managed outside formal collection and recycling schemes. As a result of non-compliant e-waste management, 58 thousand kg of mercury and 45 million kg of plastics containing brominated flame retardants are released into the environment every year. This has a direct and severe impact on the environment and people's health.***[lx]

Nevertheless, as bad a situation as described, the Global E-Waste Monitor 2024 provides no assistance in quantifying the additional layer of crisis due to the AI waste wave. While the Monitor does not explicitly state that AI or data center hardware is excluded from its estimates, a direct reading of the report's own methodology confirms that AI-driven infrastructure is not separately analyzed or projected.

Although its 54-category UNU-KEY classification system — the taxonomy underlying every figure in the report — does include "professional IT equipment" (UNU-KEY 0307, covering servers, routers, and data storage) it does not disaggregate data center infrastructure at the granularity needed to isolate AI-driven waste. 'Artificial Intelligence' is mentioned exactly once in the report, in a single sentence in the Foreword, with no connection to the report's own data or methodology. This silence appears to underscore that the Monitor's authors have not considered hyperscaled AI infrastructure as a distinct category requiring separate analysis.[lxi]

Indeed, Cornelis (Kees) Baldé, a senior scientific specialist at UNITAR and a co-author of the Global E-Waste Monitor itself, described early attempts (Wang et al) to quantify AI's specific contribution to e-waste as 'novel'[lxii] — an implicit acknowledgment that this quantification sits outside, not within, the Monitor's own existing baseline figures.

This distinction matters for the calculation that follows. Band 1 extrapolates the Monitor's conventional e-waste trajectory, which is built on placed-on-market data for commercial equipment. Band 2 is a separate bottom-up estimate of hyperscale AI infrastructure — custom-built, procured outside commercial channels, and dominated (87 percent by mass) by power distribution, cooling, networking, and battery systems that the Monitor does not identify or project as integrated data center stock. The two bands measure different things from different data sources.

Thus, while the Monitor is an excellent baseline on the conventional global e-waste crisis we have become accustomed to assessing and calls it out correctly as a very serious matter, we must turn to other work to begin to assess the size of the coming AI waste wave that will come as an addition to the well-documented conventional crisis.

AI Waste Studies / Wang et al.

The first such study we can find in the literature was published in Nature Computational Science in 2024 by Wang et al.[lxiii] The study's methodology has been the top-down extrapolation of rising rates of demand for servers over 10 years (2020-2030). It has been peer-reviewed and is transparent about its four indicated scenarios. Its weaknesses are that it may have assumed a 3-year server lifespan when later literature assumed longer lifespans. It also treated AI workloads as running on dedicated hardware. Both of these assumptions have been challenged by later literature's (Hintemann et al) hyperscaler stock data[lxiv] and (de Vries-Gao's chip-lifespan analysis).[lxv] These later assessments conclude that Wang overestimated AI e-waste by 90%.

Further, by the authors' own account, Wang et al.'s study captures only large language model-driven hardware demand. It does not account for the rapidly growing compute requirements of video generation, robotics, autonomous vehicle systems, recommender engines, scientific simulation, or the many other AI applications now driving data center buildout, workloads that collectively may already rival LLMs in hardware demand. This limitation applies equally to de Vries-Gao, whose supply-chain analysis is similarly scoped to LLM-class accelerators.

AI Waste Studies / De Vries-Gao

A 2026 recalibration of Wang's figures, published by data scientist Alex de Vries-Gao, took a different approach: instead of projecting server counts from AI demand, it checked how many AI servers the chip supply chain as it is today could actually produce, using known manufacturing capacity limits. Combined with real-world hyperscaler lifespan data (4–6 years, not 3), this produced a much lower estimate: 131 to 224.8 kilotonnes of AI server e-waste annually by 2030, roughly one-tenth of Wang's figure. This appears to be a methodological improvement, checking projections against physical production limits rather than demand alone. But it shouldn't be read as the final word: de Vries-Gao's prior work has consistently landed on more conservative estimates than others in this field. For example, averaging server lifespan can mask hardware that's pulled from frontier AI use early, but is kept running in a lesser role rather than it would normally be, but still not being completely scrapped; and the manufacturing capacity used as a ceiling reflects one snapshot in time in a market expanding specifically because that capacity keeps proving insufficient. This is a serious challenge to Wang's numbers, but not a settled replacement for them.

AI Waste Studies / Gröger et al.

The most recent attempt to quantify AI waste comes from a 2025 report commissioned by Greenpeace Germany and authored by Gröger et al. of the Öko-Institut.[lxvi] It's worth being precise about what this is: mainly a literature review of about 100 sources, not an original model like Wang's or de Vries-Gao's, and the authors themselves call their figures "rough estimates." Its widely quoted "up to 5 million tonnes" isn't an independent finding at all: it's Wang's own number, repeated. The report does contain one original calculation worth citing: using an estimated global server stock and a four-year average lifespan, the authors project roughly 4.2 million tonnes of e-waste from servers and storage alone between 2023 and 2030. Two caveats matter here.

The authors themselves admit their method — scaling hardware mass to power draw — is "crude," since different components scale with power differently. And their figure leaves out networking gear, cabling, batteries, and cooling infrastructure entirely, meaning it likely understates rather than overstates the real number. The report was reviewed before publication by other researchers in this field, including de Vries-Gao, which lends it some credibility despite its admittedly rough methodology.

BAN's Approach: A Bottom-Up Infrastructure Model

Rather than project from AI computing demand alone, a method that, as de Vries-Gao demonstrated, can produce estimates untethered from physical manufacturing realities, this paper starts from from what the industry says it intends to build, and calculates the waste that buildout implies. The method tracks how much equipment is in use at any given time, and how quickly it is being retired and replaced with respect to five categories of equipment. Every gigawatt of installed data center capacity requires hardware of known mass for these categories: servers, power distribution, cooling infrastructure, networking, and backup power systems. Section 2 quantifies this as averaging about 70 kilotonnes per gigawatt. Again, recall that due to the recent definitions of the Basel Convention, any whole unit or component that has electronics as part of its makeup is considered electronic waste (e.g integrated cooling devices) as are process residues from managing these wastes.

As capacity grows, so too does the installed equipment accordingly. Each of the five equipment categories has its own replacement cycle, ranging from 2.5 years for AI accelerators to 8 years for power distribution. Dividing the installed mass in each category by its lifespan yields the annual flow of waste.

The model's inputs are therefore three things: how fast capacity is growing, how much equipment of each category each gigawatt requires, and how long that equipment lasts before disposal.

Global installed data center capacity stood at approximately 122 gigawatts in 2024. Growth projections from McKinsey (171–219 GW by 2030), JLL (200 GW by 2030), and ABI Research (277 GW by 2035) yield an average implied compound annual growth rate of approximately 8.8 percent. This paper applies that rate through 2050, noting that sustaining 8.8% compound growth for 25 years might be viewed as a provocative assumption; actual growth could prove higher or lower.[lxvii]

It is true that the 122 GW in 2024 includes lighter legacy non-AI facilities, but virtually all new capacity is being built to AI specifications, with existing facilities increasingly being retrofitted. By 2050, the installed base will, in our view, overwhelmingly reflect the heavier AI profile. This makes the mass intensity assumption below a reasonable forward-looking estimate.

Gigawatts to Mass Intensity: 70 kilotonnes/GW

Using the detailed component data from Section 2 above, we derived a mass intensity of approximately 70 kilotonnes per gigawatt of installed capacity, built bottom-up from five equipment categories. This figure is independently corroborated by the World Economic Forum, which estimates that each megawatt of data center capacity embeds approximately 60–75 tonnes of minerals, mainly in power and cooling systems rather than servers.[lxviii] Microsoft's reported copper consumption of 26 tonnes per megawatt at its 80 MW Chicago facility also closely matches Section 2's figure of 27 tonnes per megawatt, though the WEF notes that facility-only copper is approximately 12 tonnes per megawatt while 26 tonnes per megawatt includes on-site and near-site power connections. A comprehensive bottom-up derivation yields a range of 62–77 kt/GW, with 70 kt/GW as the midpoint.

No efficiency discount is applied due to the fact that Lean Research (2026) reports that infrastructure cost per watt has risen 40 percent in the past two years, and AI racks are heavier than traditional ones, not lighter.[lxix]

Equipment Lifespans

Each of the five equipment categories carries its own replacement cycle. Where published data exists, we use it directly. Where it does not, we state our estimate and explain our reasoning. The five categories and their assumed lifespans:

Accelerators, Servers, and Racks: 2.5 years

AI accelerators are on 2–3 year replacement cycles, compressed from the traditional 5–7 years for general-purpose servers. Nvidia's biennial architecture cadence (Hopper → Blackwell → Rubin) drives this pace, with each generation offering performance gains large enough to make the prior generation economically uncompetitive for frontier workloads. Sources: Gartner, Dell'Oro, Resource Recycling, de Vries-Gao (ScienceDirect, 2026). This category represents 13 percent of facility mass.[lxx]

Networking Equipment: 3.5 years

Hyperscale networking is being retired at 3–4 years, driven by the 400G → 800G → 1.6T → 3.2T transition: four generations of switch and optical transceiver technology in approximately four years. Source: ROC Telecom (2026). This category represents 3 percent of facility mass, the lightest, but cycles through disposal more frequently than heavier infrastructure.[lxxi]

Power Supply and Distribution: 8 years

This is a BAN estimate. Traditional power infrastructure lasts 15–20 years, but two forces are compressing this: (a) AI density upgrades forcing wholesale replacement of power delivery systems designed for 5–15 kW racks to handle 50–140 kW (Dataconomy, Aug. 2026; JLL), and (b) incoming superconductor technology — VEIR's commercial rollout begins in 2027, with HTS cables 15 times lighter than copper, and design-in expected by 2028–2030 (Microsoft partnership announced). No single source gives a specific shortened lifespan; 8 years is the midpoint of a 7–10 year range. This category represents 34 percent of facility mass — the heaviest, driven by copper cabling and busbars.[lxxii]

Backup Power Systems: 5 years

This reflects industry-standard battery replacement cycles. VRLA (lead-acid) batteries last 3–6 years; lithium-ion lasts 10+ years. Five years is a reasonable blended average during the current technology transition. Additionally, sodium-ion and solid-state batteries arriving in 2027–2030 may trigger a second rip-out cycle within the decade. This category represents 15 percent of facility mass.[lxxiii]

Cooling Systems: 5 years

This figure is a BAN estimate, a blended average of two observed dynamics: (a) air-cooled systems being scrapped early as facilities convert to liquid cooling (70 percent of data centers remain air-cooled as of 2026), but are converting rapidly — (TechPowerUp, Grand View Research), and (b) liquid cooling systems themselves lasting 5–7 years. The conversion effect is significant. Cooling infrastructure is adding substantial mass, not saving it. The shift from air to liquid cooling appears to increase material intensity even as it reduces energy overhead costs. A single coolant distribution unit can weigh three tonnes when flooded, and liquid cooling requires physically modifying servers and racks to accommodate cold plates and manifolds. When a facility retrofits from air to liquid cooling, much of the existing cooling infrastructure becomes waste regardless of remaining useful life. This category represents 35 percent of facility mass.[lxxiv]

Why Previous Studies Found Less Potential E-Waste

Wang et al., de Vries-Gao, and Gröger et al. counted only AI servers and, in Gröger's case, storage. None included the power distribution infrastructure (34 percent of facility mass), cooling systems (35 percent), networking equipment, or battery backup. Gröger et al. acknowledged these as additional, unquantified sources, an important recognition that the scope of prior work was incomplete. Together, these non-server categories constitute 87 percent of the total equipment mass in a data center. They are most likely to be considered potential e-waste due to integrated electronics within each category of equipment. It is important to note that the Basel Convention now defines controlled e-waste as being whole equipment containing electronic circuits, components from such equipment, as well as residues from processing equipment or components.[lxxv] Counting servers alone is like estimating the scrap value of an automobile factory by weighing only the engines.

Figure 3 of this report makes this visible. Wang's, de Vries-Gao's server-only projections appear as thin lines at the base of our chart, ending at 2030. Our model's full infrastructure estimate towers above these estimates and continues the trajectory through 2050. The gap is not so much a disagreement about methodology: it is a difference in scope. We are simply counting more of what will be there.

Just as we were completing Part 1 of this White Paper, Amoah et al. (2026) published a mineral demand study in Resources Policy that independently corroborates our finding that power infrastructure, not semiconductors, dominates the material mass of AI data centers. From the study: "The AI buildout's demand for critical materials is driven primarily by bulk infrastructure requirements." This new study focuses on mineral supply chains; ours translates that same infrastructure into projected waste flows by component category and replacement cycle through 2050.[lxxvi]

Finally, neither study accounts for the very significant, overlooked AI contagion effect — the downstream obsolescence of consumer, enterprise, and telecommunications hardware driven by AI innovation and requirements. While this segment of waste is not data center waste per se, it is nevertheless created by AI and is very much an impact of developments derived directly from the hyperscaling taking place in data centers.

The AI Contagion Waste

Data center infrastructure e-waste is only part of the story, and, as it turns out, the smaller part. The four projections reviewed above, including our own model, all share a common limitation: they count only what is generated from inside themselves. None of the assessments above asks what happens outside as a result of these same data centers. And yet of course, data centers do not exist in isolation. They are a new engine for the entire computational universe. Everything the world will purchase and eventually discard in order to use what those data centers produce — the consumer devices, the telecommunications infrastructure, the edge computing hardware — constitutes a second, larger wave of AI-driven waste that no published study has attempted to quantify as yet.

We call this the AI waste contagion: the downstream effect of data center buildout on the obsolescence rate of the much broader electronics ecosystem. But in doing so, we have realized that very little data exists to quantify this segment. It has not, to our knowledge, been attempted. This is nevertheless an essential segment of the AI waste spectrum that is coming and simply cannot be ignored. We challenge others to improve upon our estimations.

The AI waste contagion operates primarily through three channels.

> **First,** AI creates hard hardware thresholds that render existing consumer and enterprise devices functionally obsolete.

Microsoft's Copilot+ requires neural processing units capable of 40 trillion operations per second and 16 GB of RAM (specifications that exclude many PCs in current use).[lxxvii] Gartner projects that 100 percent of enterprise PC purchases will be AI PCs by the end of 2026;[lxxviii] IDC projects that 93 percent of PCs shipped will be AI-capable by 2028.[lxxix] This is not a forecast of what consumers will want to buy — it is a forecast of what manufacturers will offer. The market is shifting wholesale to AI hardware, and the non-AI options are disappearing from the shelf. Apple Intelligence requires an iPhone 15 Pro or later.[lxxx] Google's Gemini requires recent Tensor processors.[lxxxi] These are not soft recommendations — they are thresholds for making IT equipment obsolete.

Second, entirely new categories of AI hardware are entering the world that the UN Global E-Waste Monitor does not track at all. Edge AI devices (smart cameras, IoT sensors with on-device inference, autonomous vehicle computing modules) represent an installed base that MarketsandMarkets estimates at $26 billion in 2025 and $59 billion by 2030, growing at 17.6 percent annually.[lxxxii] These devices have short lifecycles, typically four years,[lxxxiii] and no established end-of-life management pathway.

Third, the telecommunications infrastructure connecting users to AI services must be massively upgraded. The optical networking equipment market is growing at 10.5 percent annually, roughly double its historical rate.[lxxxiv] The Fiber Broadband Association projects a 2.3-fold increase in fiber miles by 2029.[lxxxv] Each generation of network equipment (from 400G to 800G to 1.2 terabit speeds) retires the previous one.

Quantifying these three channels requires different approaches for each.

For the first channel above, one must first calculate the fraction of each of the 6 traditional e-waste categories from large appliances to smartphones, identified by the Global e-Waste Monitor, that are likely to be pushed into early retirement due to AI. A small refrigerator is not likely to be, but a laptop almost certainly will.

We estimate this by assessing each category's susceptibility to AI hardware requirements. For small IT and telecom — phones, laptops, tablets — the susceptibility is high. For screens and monitors, the susceptibility is also significant as smart displays and AI-capable TVs replace conventional units. For large appliances, it is minimal. We have assigned factors to each of these 6 potential e-waste categories. Once those are established, we need to assign a factor for the rate of replacement due to AI-induced obsolescence. The critical insight is that this is a supply-side contagion: manufacturers have already decided that AI hardware is the future. Gartner forecasts 100 percent of enterprise PC purchases will be AI PCs by the end of 2026. Consumers and businesses will replace their devices not because they choose to, but because non-AI alternatives are ceasing to exist at reasonable prices. We use this market-wide hardware transition as an anchor to estimate the acceleration across other susceptible categories.

In the second channel, we are recognizing the rapid growth in AI "edge" hardware reported by MarketsandMarkets, that did not even exist until very recently and, by definition, is 100% AI hardware. For a replacement rate, we have assigned a four-year lifespan, reflecting the rapid pace of innovation in this quickly evolving field.

Finally, in channel three above, we need to measure the leap in telecom hardware growth induced by AI queries across the planet. Based on annual global telecom equipment revenues of approximately $100 billion,[lxxxvi] an estimated average equipment value of approximately $450 per kilogram,[lxxxvii] and a 7-year equipment lifespan, we can arrive at an installed base of carrier and enterprise telecom infrastructure of approximately 1.5 million tonnes. We then can measure the difference between the AI-driven growth rate (9% CAGR)[lxxxviii] and the historical baseline growth rate (4% CAGR).[lxxxix] That gap is attributed to AI. We assign an equipment lifespan of 7 years[xc] to derive our final AI waste generation calculation from this category.

Conservative and Aggressive Contagion

There is a distinction, however, between what a consumer is forced to make obsolete (conservative contagion) and the obsolescence they elect by virtue of the expected user "candy" — the AI desires that peers and advertising will entice us to choose to upgrade to (aggressive contagion). Our conservative estimate counts only devices rendered functionally unable to run AI features when it is demanded of a basic user experience (e.g. internet). We are talking about the laptop that cannot run Copilot+, the phone excluded from Apple Intelligence. While cloud-based AI services remain accessible from older hardware, the full performance and feature set — on-device inference, real-time processing, offline capability increasingly requires local AI hardware, creating unavoidable market pressure toward replacement. Under these conservative assumptions, the downstream AI contagion adds approximately 16.2 million tonnes of retired electronic equipment annually by 2050. Our aggressive estimate adds the cultural/psychological replacement effect. Devices that still work but are replaced because AI features become irresistible the way the iPhone made the flip phone obsolete not because it couldn't make calls, but because it no longer felt adequate. The aggressive scenario similarly increases growth assumptions for edge devices and telecommunications infrastructure, bringing the additional downstream waste to approximately 30.7 million tonnes by 2050.

A Note on the Concern over Contagion "Double-Counting"

A likely objection to our contagion category is that counting AI-driven consumer device replacement as "AI waste" simply relabels ordinary e-waste. This misunderstands what we are actually seeking to measure in Bands 3 and 4. Band 1 already assumes every phone, laptop, and server on earth will eventually become e-waste at historical replacement rates. This Global Monitor 2024 projection discussed earlier bakes in normal obsolescence, but it did not anticipate the AI effect. Bands 3 and 4 represent non-data center derived (Band 2) but nevertheless AI-accelerated obsolescence above the baseline trend provided by the Global Monitor that was set before AI contagion existed.[xci] Bands 3 and 4 do not count those devices again — they count only the delta created when AI forces devices to be replaced sooner than they otherwise would have been. The test is not "does this device use AI?" It is "was this device replaced earlier than baseline because of AI?" The historical parallel is the analog-to-digital TV transition, the 3G→4G→5G migrations, and Windows 11 dropping support for older processors, each creating a documentable acceleration in retired electronic equipment above the trend line. AI is doing the same thing across phones, PCs, networking, and telecom infrastructure simultaneously.

The BAN Model: The Four-Banded Framework

Applying these parameters — 122 GW baseline, 8.8 percent CAGR, 70 kt/GW mass intensity, and the category-specific lifespans above — the model projects four bands of AI-driven retired electronic equipment through 2050.

Band 1 — UN Baseline: ~165Mt/yr by 2050 (BAN's extrapolation beyond 2030). The conventional e-waste growth trajectory from the UN Global E-Waste Monitor, before any AI-generated hardware is added.

Band 2 — Data Center E-Waste Infrastructure: ~15.5 Mt/yr by 2050. The annual retired electronic equipment from all five equipment categories across the global installed data center base, at sourced or derived replacement rates.

Band 3 — AI Contagion E-Waste, Conservative: 16.2 Mt/yr by 2050. The downstream effect on consumer and enterprise devices forced into early replacement by hard AI hardware thresholds (NPU requirements, RAM minimums, connectivity upgrades).

Anchored to Gartner's projection that 100 percent of enterprise PC purchases will be AI PCs by the end of 2026 and IDC's forecast that 93 percent of all PCs shipped will be AI-capable by 2028 — a supply-side transition that leaves consumers no non-AI alternative at reasonable prices. Counts only "forced" technical obsolescence, not marketing-driven replacement.

Band 4 — AI Contagion, Aggressive Additional: 14.5 Mt/yr by 2050. The additional waste if AI features become culturally/psychologically desirable — the smartphone effect — driving mass replacement beyond what technical requirements alone would demand. No published data exists for the rate of psychologically driven replacement; this band provides an upper bound to bracket the plausible range. This is a BAN estimate.

All e-waste totals: (see Figure 4)

Conservative total (Bands 1+2+3): 196 Mt/yr by 2050.

Aggressive total (Bands 1+2+3+4): 211 Mt/yr.

AI-driven waste alone: (See Figure 2)

31 Mt/yr conservative (Bands 2+3)

46 Mt/yr aggressive (Bands 2+3+4)

Figure 3 shows the BAN projections alongside those of Wang et al. and de Vries-Gao until the year 2030.

Why These Projections May Understate the Problem

The four-band framework largely projects waste cycling through its expected lifespan and being disposed of on schedule. The amounts we project are already staggering. However, several independent factors lacking data will likely push actual volumes above these projections. We explore some of these factors below.

Mandated Redundancy: Data sovereignty laws in the EU, China, India, Indonesia, and a growing number of jurisdictions require that data generated within their borders be stored domestically. This means the same AI capability must be physically duplicated across multiple countries rather than served from a single global pool. Each duplicate facility carries its own full complement of power distribution, cooling, networking, and computing hardware — all of which will eventually require disposal. The cost of this fragmentation is measurable: one systems-research study found that workloads constrained to a single continent cost 34–46% more than the same workload run with full global placement freedom.[xcii] Military and intelligence agencies impose similar requirements. The result is that global installed capacity (and therefore future waste) scales faster than global demand alone would require.

Geopolitical and security risk: Data centers are increasingly recognized as critical infrastructure. In a conflict or act of terrorism targeting these facilities, the destruction would convert installed equipment to waste instantaneously and in bulk, outside any replacement cycle. This is not modeled and cannot easily be, but the concentration of billions of dollars of hardware in known, fixed locations creates a vulnerability that has no precedent in the history of electronic equipment and would likewise create greater incentive for redundancy.

Telecom infrastructure: While we mentioned telecom obsolescence in the contagion discussion above, it has likely been grossly underestimated. The telecommunications infrastructure connecting users to AI services is massive and poorly quantified to date. While AI is a significant accelerant, 5G/6G rollout and general data demand growth also drive telecom infrastructure replacement. AI inference at the edge requires low-latency 5G and eventually 6G connectivity, driving the same kind of hard network cutoffs that retired 3G in 2022 — but across optical transport, cell tower radio equipment, backhaul infrastructure, and last-mile access hardware simultaneously. The Fiber Broadband Association projects a 2.3-fold increase in fiber miles by 2029, bringing with it cable, conduit, splice enclosures, and distribution frames — physical infrastructure with real mass.

No published study has attempted to quantify AI-driven telecom waste in tonnes. This is a very large data gap. It will not be a small number.

No room for efficiency gains where it matters most: The segment driving the largest share of new GW capacity, frontier AI training, runs continuously for weeks or months at a time.[xciii] Unlike general-purpose computing we are all used to, where workloads can be consolidated and servers virtualized, training clusters cannot be made to do more with less hardware. Efficiency improvements in chip design have historically been absorbed by larger and more complex models, not by reducing the hardware footprint.

Rapid Conversion Waste: Three Impending Waves

There is a further reason to treat these projections as a floor rather than any kind of ceiling. The stock-and-flow model assumes orderly replacement at sourced refresh rates: equipment reaching the end of its useful life and being decommissioned on schedule. Technology transitions do not work this way, and competitive races work this way least of all. What we can expect is that fully functional equipment will be scrapped not because it failed but because the technology needed to change and change rapidly to ensure the hyper-scalers maintain their competitive advantage.

The preceding section documented that when a traditional data center is retrofitted for AI, much of the existing power and cooling infrastructure must be removed — not at the end of its 15- or 20-year useful life, but now, to make room for equipment that can handle 10 to 30 times the power density. This conversion waste is not a scheduled refresh. It is a one-time dump of fully functional infrastructure that happens to be the wrong kind.

History offers a direct analogy. The CRT-to-LCD transition, cited in the Introduction, confronted the world with billions of cathode ray tubes containing leaded glass that nobody had planned to dispose of safely. The industry did not build disposal capacity before building flat screens. It chased the new technology and left the waste for someone else to figure out. The AI conversion is the same pattern at a vastly larger scale.

The competitive pressure of the current AI buildout makes orderly, planned decommissioning even less likely. Facilities are being gutted and rebuilt to stay competitive, with no prior thought to making the old infrastructure upgradable or reusable. This is not speculation; it is what happens, repeatedly, when an industry is in a technology bubble. The usual modeling captures steady-state waste generation rates. It does not capture the conversion surges that the historical record tells us will come. Three such "rip-and-replace waves" are approaching simultaneously:

Superconductor transition: VEIR's commercial high-temperature superconducting cables, 15 times lighter than copper, begin deployment in 2027. Microsoft has partnered with VEIR for data center applications. As superconducting power distribution enters facilities, existing copper cabling and busbars, which are the single heaviest component in a data center, will be removed regardless of condition.

Liquid cooling conversion: Seventy percent of data centers remain air-cooled as of 2026, but AI workloads above 40–50 kW per rack require liquid cooling. Current-generation AI accelerators such as Nvidia's GB200 ship as liquid-cooled-only, with cold plates and manifolds integrated at the factory level. A facility converting from air-cooled to liquid-cooled AI hardware cannot retain its existing servers — the cooling and compute are a single assembly. The conversion therefore scraps not only the facility's air-cooling infrastructure — CRAC units, raised-floor plenums, air handlers — but also the computing hardware itself, regardless of remaining useful life:

Battery chemistry transitions: The shift from lead-acid to lithium-ion is already underway. Sodium-ion and solid-state batteries are expected to reach commercial deployment in 2027–2030, potentially triggering a second significant replacement cycle within a single decade. The stock-and-flow model captures steady-state replacement.

It does not capture these conversion surges. History — from the CRT-to-LCD transition to the coal-to-gas transition in power generation — tells us they will come and when they do it will not be a pretty phase-out. It will be rip and replace.

Why These Projections May Overstate the Problem

Equipment Lifespan Estimates Prove Wrong: Our model assumes a 2.5 year lifespan for accelerators, but some hyperscalers like Google and Microsoft have published research that they could extend server lifespans to 5-6 years. However, published research on extending server lifespans has focused on general-purpose computing fleets, not AI accelerator clusters.[xciv] Whether similar extensions are achievable for AI hardware remains untested, but given the rapid generational performance gains that the horserace demands, we doubt lifespan extensions are realistic. The problem being that the very latest innovations in GPUs make them so much faster than their predecessors that it makes little economic sense to retain them even while perfectly functional. This leads to shorter commercial, if not physical, lifespans.

Making Servers Modular so they Can Be Upgraded: Some might argue that hyperscalers can swap out the old GPU chips and preserve the vast majority of the server and infrastructure, such as the chassis, the power supply, cooling, and networking in place. This is, of course what is the desirable outcome and a truly circular way of thinking. However, current AI hardware design is moving in the opposite direction. NVIDIA's GB200 NVL72 rack, for example, ships as a single integrated unit where the liquid cooling system, the compute boards, and the networking fabric are physically built together at the factory.[xcv] You can't pull out the GPUs and drop in next-generation chips because the cooling manifolds, cold plates, and power delivery are all designed around the specific thermal and electrical profile of that particular chip.

When the next generation arrives with different power draw, different heat output, and different physical dimensions, the whole assembly has to go. Sadly, the trend in AI hardware is toward more integration, not less — which means more waste per upgrade cycle, not less.

Software Efficiency Reducing Hardware Demand: AI software is getting more efficient — would this not mean fewer servers and therefore less waste? In what has often been called Jevons' Paradox, adding more lanes to the freeway does not free up traffic if the extra lanes are immediately filled. The industry example is when in January 2025, DeepSeek, a maverick Chinese AI company, released DeepSeek-R1, an AI model that matched OpenAI's best reasoning model on math and coding benchmarks — trained for about $5.6 million, a fraction of what comparable models cost.[xcvi] The market reaction was dramatic: NVIDIA lost roughly $589 billion in market value in a single day — the largest single-day loss in stock market history[xcvii]— on fears that algorithmic efficiency could reduce demand for chips. But that fear proved short-lived. Demand for NVIDIA GPUs kept climbing. The industry's response wasn't "we need fewer chips"; it was "now we can do more with each chip," which drove greater deployment than ever before. It may be that the only thing that will reduce hardware consumption will be a decline in demand for AI. That may be a long time coming.

Demand plateau before 2050: What if the AI buildout slows dramatically and our 8.8% CAGR is not sustained for 25 years straight, that perhaps demand will saturate at some point (see Jevons Paradox argument above) or the supply chain is unduly disrupted. What if the growing concern over the threat of AI generally, or Data Centers specifically, grows to the point that democracies across the globe demand and succeed in slowing down the AI horserace? But again, we are projecting just 25 years, and during that time if the curve tapers, the cumulative waste from the buildout is already enormous, with the equipment already installed during the boom, so even if the bottom drops out and the horserace is over, that still will not diminish the waste already produced during the boom. A slowdown in new deployment won't reduce the waste already generated; it just will change when the peak waste year arrives.

It is our strong belief that the waste pipeline will still be quite full in 2050 (not so far off) and will still present a critical environmental concern.

Reuse/refurbishment via Secondary Markets Will Ensure Extended Use: Many in the ITAD industry are viewing the AI horserace as a goldmine for business, the most lucrative of which will be refurbishing and reselling the surplus hardware following the buildout boom and likely rapid refresh cycles. Surely there will be a massive demand for fast but not fastest equipment in developing countries, in lower-tier data centers, and inference workloads. The question is a valid and open one. Will reuse save us from an e-waste tsunami? Surely the hyperscalers are planning to ensure such a future. Part 2 of this White Paper aims to probe that potential, that hope, that mitigating factor.

On balance, the moderating factors so far explored are plausible but unlikely to offset the amplifying ones at the scale we are projecting here. Uncertainty cuts both ways. In light of that, we have created a Sensitivity Analysis in Appendix B for the reader's better understanding.

SECTION 4 – Summary: Dimensions of a Tsunami

We have examined four different ways of approaching the question of just how much new retired electronic equipment can we expect from the AI e-Waste Wave — Wang et al.'s demand-driven model, de Vries-Gao's supply-constrained recalibration, Gröger et al.'s literature synthesis, and this paper's own infrastructure-based approach with its four-banded framework. Each provides a very different trajectory. All of them are alarming.

But our numbers, while dramatically exceeding the other modelled trajectories, are thus largely because we have counted what others failed to count, and not because of some new hyperbolic approach being taken. It is to be remembered that the societal concern we seek to measure is the extent to which the current e-waste crisis may be worsened by the AI revolution. In order to get our arms around this coming AI wave, we must measure that concern and scan the entire ecosystem of AI hardware, not just the servers. The international definition of e-waste defined by the Basel Convention is far broader and rightly so. It would certainly include most wastes from other data center infrastructure, including categories such as data storage, power supply and distribution, cooling systems, and networking systems, most of which are e-waste and subject to varying rates of obsolescence.

Certainly, there is a lot of speculation as to how steep the growth curve of the AI build-out will be. There are so many variables possible in that question that one can lose precious sleep and time gazing into that crystal ball. We chose to leave that question behind for the most part, and take the industry at its word: the numbers they are citing, with the full expectation that their ambitions have been tempered by their knowledge of anticipated restraints.

Our approach then is based on the premise that IF the industry grows as fast as they intend, what will the consequences be?

Next, we believe we are on very safe ground in correlating Gigawatts consumed with hardware utilized. As demonstrated above, GW capacity tracks hardware mass regardless of chip efficiency gains. Decoupling energy consumption from hardware weight is not likely to be possible. Based on Section 2's equipment mass data for a reference 100 MW AI data center facility, every gigawatt of installed capacity represents roughly 70,000 tonnes of physical equipment across all five categories.

The final variable, then, is rate/s of obsolescence. For this, we have taken each category of potential e-waste within a data center and assigned conservative lifespan estimates (from 2.5 years for servers and accelerators to 8 years for power distribution equipment), our waste obsolescence rates should err on the low side. Given the category-specific lifespan assumptions detailed above, this translates to approximately 14,000 tonnes of electronic equipment retired annually per gigawatt of installed capacity.

Then, as icing on this unfortunate AI waste cake, we have added what most analysts are completely ignoring: the waste from induced obsolescence taking place when conventional electronic hardware is replaced due to software updates demanding new AI processing capabilities or direct AI hardware adaptation requirements. This AI Contagion obsolescence as we call it, will not take place in data centers but with our own IT equipment in our homes, our businesses, and our gaming rooms. But this discarded obsolete e-waste cannot be overlooked if we wish to correctly calculate the scale of the AI e-Waste Wave to be unleashed upon a world which is already suffering from woefully inadequate and often unethical e-waste management infrastructure.

Taken together, and including the conventional predicted e-waste growth, the four bands of our global frame — 1) the UN Global Monitor baseline (extrapolated to 2050), 2) BAN's data center infrastructure projection based on GW installed, and the 3) conservative, and 4) aggressive contagion scenarios are laid out in the following table and the subsequent graphs.

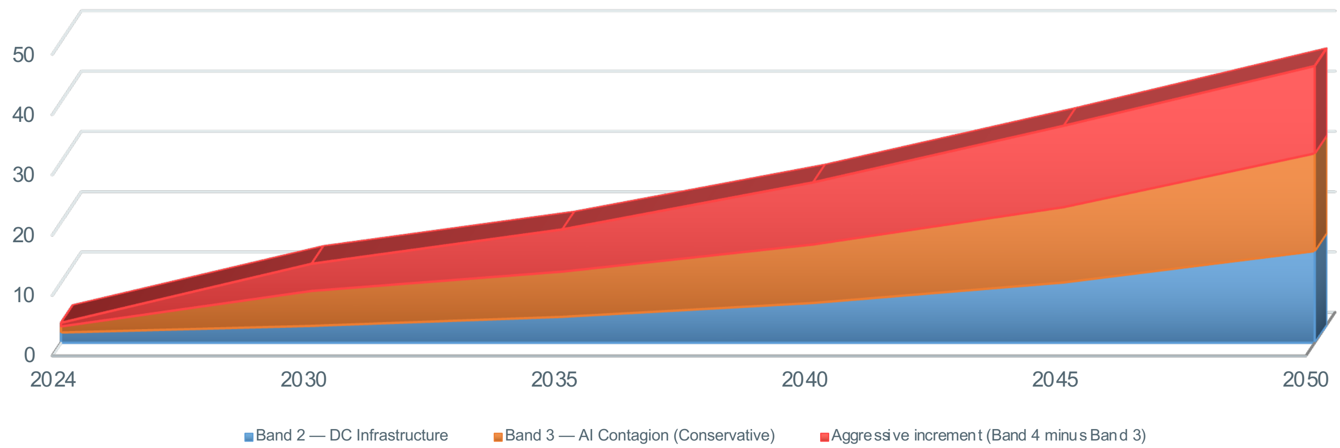
Table 1: E-Waste (Conventional) and Added AI Waste Projections per Annum

Year	Giga watts	Band 1 (Mt) Global e-Waste Monitor	Band 2 (Mt) BAN's AI Estimate (Data Centers Only)	Band 3 (Mt) Contagion AI Waste (Conservative)	Band 4 (Mt) Contagion AI Waste (Aggressive)	BAN's Estimate AI Waste with Contagion (Conservative)	BAN's Estimate AI Waste with Contagion (Aggressive)	Total e-Waste (Mt) (Conservative)	Total e-Waste (Mt) (Aggressive)
2024	122	66.5	1.7	1.0	1.6	2.7	3.3	69.2	69.8
2030	202	82.0	2.8	5.8	10.3	8.6	13.1	90.6	95.1
2035	308	97.7	4.3	7.5	14.5	11.8	18.8	109.5	116.5
2040	470	116.3	6.6	9.7	20.0	16.3	26.6	132.6	142.9
2045	717	138.5	10.0	12.5	26.0	22.5	36.0	161.0	174.5
2050	1,093	165.0	15.2	16.2	30.7	31.4	46.0	196.4	210.9

Band 1 = Global e-Waste Monitor projections (2050 extrapolated) not including AI waste)
Band 2 = BAN's AI Estimate including all electronic waste from AI Data Center Infrastructure based on GW growth rate and categorized life span rates.
Band 3 = BAN's Conservative estimate of e-waste generated by AI induced obsolescence outside of Data Centers
Band 4 = BAN’s Agressivee estimate of e-waste generated by AI induced obsolescence outside of Data Centers
Columns 7 = Band 2+3, Column 8 = Band 2+4, Column 9 = Band 1+2+3, and Column 10 = Band 1+2+4.

Figure 2. AI-driven e-waste

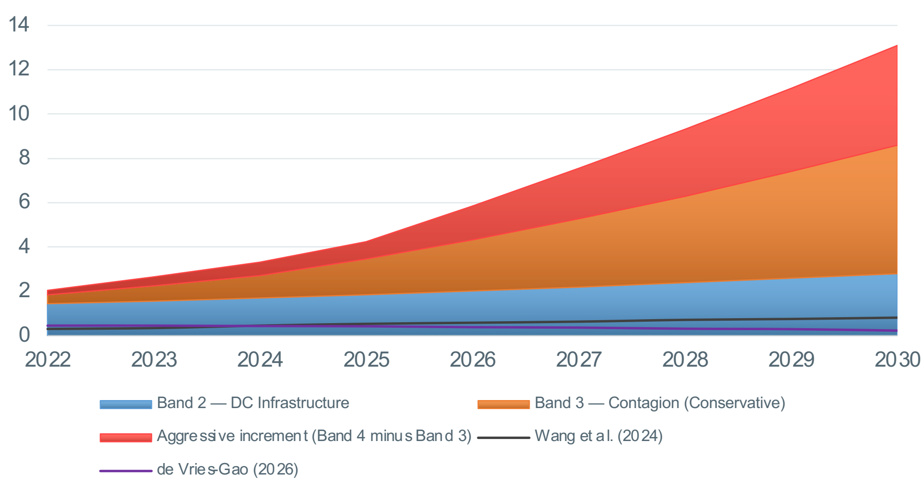
BAN's Full Infrastructure Model including Contagion outside of Data Centers



Note: See inset below to visualize Wang and de Vries-Gao quantification looking only at AI servers only. BAN includes all five equipment categories plus downstream contagion.

Figure 3. AI e-waste and earlier studies

Inset (2022-2030) showing amounts projected by earlier Studies (Wang et al. and de Vries Gao)



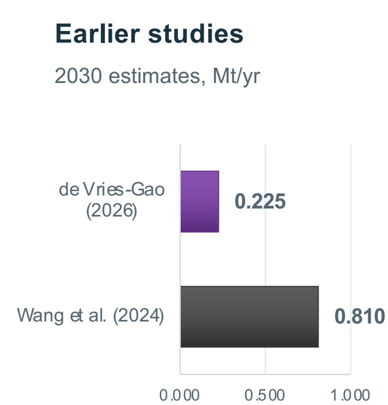
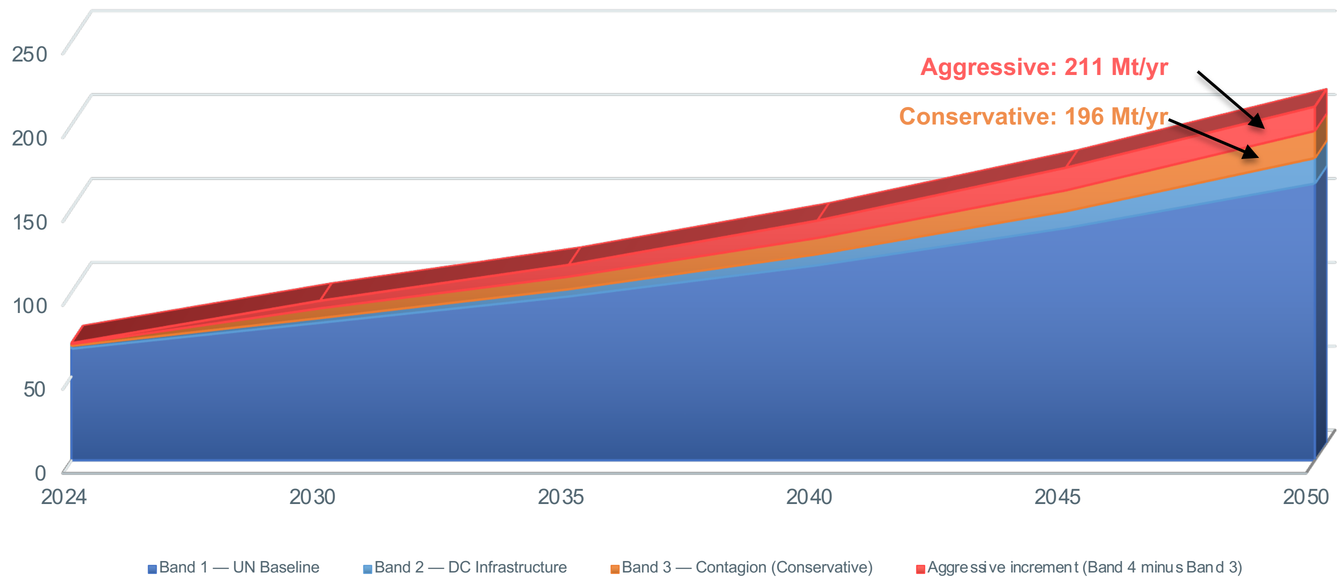


Figure 4. Global e-waste projections

Four-band framework • Annual e-waste, million tonnes



Adding It All Up

This paper estimates that unless significant efforts can be made to minimize the projected waste, for example by employing circularity principles such as reusing equipment, by 2050 we will be generating roughly **between 196 - 211 million tonnes of e-waste each year**. That is a 19-28 percent increase over what has been extrapolated from the Global E-Waste Monitor, as they did not count the AI E-Waste wave at all. That represents **more than a tripling** of today's annual global e-waste total of the roughly 67 million tonnes the world is producing today, based on the Monitor's own growth trajectory.

Cumulatively, the amounts generated of just AI-driven e-waste across the 25 years (2025 to 2050) will be roughly **395 to 617 million tonnes**, depending on how aggressively the contagion effect will play out.

That cumulative amount of AI e-waste we can likely expect, is equivalent to:

- roughly six to ten years of the world's entire current total e-waste output
- **15-23 million 40-foot intermodal shipping containers** (~27 tonnes per container). 20 million of these full-size containers placed end to end (~12.2 meters) **would stretch about 244,000 km — roughly** 6 times around the earth's circumference.

We further noted above that due to a variety of significant compounding reasons, these estimates are likely to be a floor, with the actual levels being higher.

This growth will not be evenly paced. **The steepest acceleration in data center waste is expected in the 2027–2035 window,**[xcviii] the period when hardware installed during the current buildout surge reaches end of life in large volumes. The downstream contagion effect follows on a slight delay from that window, as AI hardware requirements propagate from data centers through consumer, enterprise, and telecommunications markets.

The Rest of the Story

While this paper has painted a stark picture of an e-waste tidal wave of frightening proportions, sheer tonnage alone, the primary focus of Part I of the white paper, does not by itself capture the full gravity of the AI Waste threat. Nor have we examined what we can collectively do about it.

Part 2: In our next publication of this paper, we aim to quantify the potential for employing circularity strategies that might be employed to reduce the volumes we have discussed in Part 1. What is the potential for reuse, refurbishment, or repurposing that might be deployed and what are the challenges to doing so. Will reusing the equipment save us?

Part 3: While we have so far focused on the volumes of a veritable mountain of potential new e-waste. We have not discussed how dangerous this waste torrent is likely to be. Early indications tell us that this new AI infrastructure increasingly contains highly persistent organic pollutants known as PFAS – the so-called 'forever chemicals,' which on the global stage we have been seeking to phase out of existence. Part 3 of this study will examine the nature of this e-waste. Just how toxic will this new type of e-waste be?

Part 4: To our knowledge, no hyperscaler — including AWS, Azure, Oracle, Google, Alibaba, or Meta — nor any government or international body has published a plan for managing AI-driven e-waste at the scale projected here. While these companies have sometimes made commitments on carbon neutrality and water stewardship, none have addressed the impacts of the electronic waste their projected AI infrastructure will cause very soon and for generations to come. The infrastructure to safely process even current e-waste volumes does not exist — and so far nothing is being built to manage a more than tripling of today's global e-waste annual generation. What can we do to mitigate the dangers ahead?

Postscript: Using AI to Restrain AI

I used AI to write this paper. Anthropic's Claude was a most capable assistant and ChatGPT was a good peer reviewer. While AI did not write this paper it did prove invaluable for research, hunting and checking sources, critiques, playing devil's advocate, and finding errors. AI has already proven to be a remarkable tool for good. It currently warns 700 million people of floods a week before they hit, saving countless lives. It catches breast cancers that radiologists miss. quite a remarkable tool. However, unfortunately its enormous utility comes to humanity with many strings and chains attached. It comes with enormous risk about which too few are asking the necessary questions, let alone planning to mitigate risks upon their discovery

Certainly, we are seeing little evidence of critical risk assessment or mitigating strategies from the hyperscalers and their investors, lost in the lust of this generation's gold rush. We are not seeing it, and yet the pace we are hurtling along is already so fast that, yes, I will climb onto this runaway freight train, at least long enough to see if we can apply brakes as rational minds might find prudent. Because a tiny minority of humanity has the power to do so, all of us whether we asked for it or not, are collectively embarked on this amazing yet dangerous trajectory. We are on a frantic journey that risks planet and society as we have known them.

I am living in Greece now; I am drawn geographically to relearning the myths of this land. Zeus gave Pandora a box and warned her never to open it. But of course, curiosity got the better of her and she did so, releasing evils into the world. But Zeus left one other thing in the box which would not fly away – hope. It is our job now, as real-world Pandoras, in the name of this hope we have been handed, to first rationally identify the evils we have let fly, and then to devise a way to delay, diminish, or otherwise deal with them to preserve all that we hold dear. In Parts 1 - 3 of this paper, we seek to identify what we see as serious concerns about just this one facet of AI – a toxic e-waste tsunami. In Part 4, we'll try to give that hope a strategy.

– Jim Puckett, *Founder Basel Action Network*

APPENDICES

APPENDIX A — BAN Model Parameters and Inputs

The following table lists every input to the four-band model so that readers may reproduce, challenge, or refine the projections. Each parameter is marked as either independently sourced or a BAN estimate — original analysis where no published figure exists. Full citations appear in the main text endnotes at the reference indicated.

Data Center Infrastructure (Band 2)

- Baseline capacity: 122 GW (2024). Source: IEA. [See note lxvii]
- Growth rate: 8.8% CAGR, average of McKinsey (10.2%), JLL (8.6%), ABI Research (7.7%). [See note lxvii]
- Mass intensity: 70 kt/GW (range 62–77), derived from Section 2 component data and cross-validated against Microsoft's Chicago facility. [See note lxix]
- Category breakdown and lifespans:
    - Accelerators/Servers/Racks: 13% of mass, 2.5-year lifespan. Sourced. [See note lxx]
    - Networking Equipment: 3%, 3.5 years. Sourced. [See note lxxi]
    - Power Supply/Distribution: 34%, 8 years. BAN estimate. [See note lxxii]
    - Backup Power Systems: 15%, 5 years. Sourced. [See note lxxiii]
    - Cooling Systems: 35%, 5 years. BAN estimate. [See note lxxiv]

AI Contagion — Channel 1: Consumer/Enterprise Obsolescence (Band 3a)

- Six UN e-waste categories with assigned AI-susceptibility fractions (5%–85%) and acceleration factors (10%–45%). BAN estimates, anchored to the documented market-wide shift to AI hardware: Gartner projects 100% of enterprise PC purchases will be AI PCs by end of 2026; IDC projects 93% of all PCs shipped will be AI-capable by 2028. [See note lxxviii]

AI Contagion — Channel 2: Edge AI Hardware (Band 3b)

- Market size: $26B in 2025, 17.6% CAGR. Source: MarketsandMarkets. [See note lxxxii]
- Lifecycle: 4 years. BAN estimate. [See note lxxxiii]
- 100% AI hardware by definition.

AI Contagion — Channel 3: Telecom Infrastructure Premium (Band 3c)

- Installed base: ~1.5 Mt, derived from ~$100B annual revenue [see note lxxxiv] x 7-year lifespan ÷ $450/kg [see note lxxxvii]. Revenue is sourced (Dell'Oro); $/kg is a BAN estimate.
- AI-driven growth: 9% CAGR (BAN estimate) [see note lxxxviii] vs. 4% historical CAGR (Dell'Oro) [see note lxxxiv].
- Equipment lifespan: 7 years. Source: NTIA. [See note xc]

Aggressive Scenario (Band 4)

- Approximately doubles consumer acceleration factors from Band 3a. Increases edge device growth to 20% CAGR. BAN estimates. No published data exists for the rate of psychologically driven device replacement due to AI. These factors are designed as an upper-bound scenario to bracket the range of plausible outcomes, not as a forecast.

UN Baseline (Band 1)

- 62 Mt in 2022, 82 Mt projected for 2030. Source: UN Global E-Waste Monitor 2024. [See notes lix, lx]
- Extrapolated to 2050 at implied 3.56% CAGR. BAN extrapolation; the Monitor projects only to 2030.

Each parameter marked "BAN estimate" reflects original analysis where no published figure was found. All sourced parameters include citations found in the main text endnotes.

APPENDIX B — Sensitivity Analysis

Band 2 (Data Center Retired Electronic Equipment) Parameters

Looking at a Sensitivity Analysis of our assumptions for Band 2, we can see that our estimated Compound Annual Growth Rate (CAGR) is the dominant driver, but it's also our best-sourced parameter with three credible sources calculating gigawatts consumed per year as a barometer for hardware growth. Equipment Lifespan and our Data Center mass intensity assumptions for the differing types of equipment can shift the result modestly.

Under the most conservative combined scenario with every assumption pushed towards minimal waste, Band 2 (without the Contagion effect) still produces 4 Mt/yr of unplanned server and infrastructure e-waste.

Parameter varied	Test value	Band 2 (Mt/yr)	Change
Accelerator lifespan (baseline: 2.5 yr)			
	2.0 yr	16.5	+6%
	2.5 yr ◀	**15.5**	–-
	3.0 yr	14.9	−4%
	4.0 yr	14.0	−10%
Networking equipment lifespan (baseline: 3.5 yr)			
	2.5 yr	15.8	+2%
	3.5 yr ◀	**15.5**	–-
	5.0 yr	15.3	−1%
Power/distribution lifespan (baseline: 8 yr)			
	6 yr	16.6	+7%
	8 yr ◀	**15.5**	–-
	10 yr	14.9	−4%
	12 yr	14.5	−7%
Backup Power Systems lifespan (baseline: 5 yr)			
	3.0 yr	17.1	+10%
	5.0 yr ◀	**15.5**	–-
	7.0 yr	14.9	−4%
Cooling system lifespan (baseline: 5 yr)			
	4.0 yr	16.9	+9%
	5.0 yr ◀	**15.5**	–-

Parameter varied	Test value	Band 2 (Mt/yr)	Change
	7.0 yr	14.0	−10%
	8.0 yr	13.5	−13%
Capacity growth CAGR (baseline: 8.8%)			
	4.0%	4.8	−69%
	6.0%	7.9	−49%
	8.8% ◀	**15.5**	–-
	10.0%	20.7	+33%
	12.0%	33.0	+112%
Mass intensity (baseline: 70 kt/GW)			
	50 kt/GW	11.1	−29%
	60 kt/GW	13.3	−14%
	70 kt/GW ◀	**15.5**	–-
	80 kt/GW	17.8	+14%
	90 kt/GW	20.0	+29%
COMBINED SCENARIOS			
All parameters conservative	**3.8 Mt/yr**		
4 yr accelerator · 5 yr networking · 10 yr power · 7 yr backup · 8 yr cooling · 6% CAGR · 50 kt/GW			
Baseline	**15.5 Mt/yr**		
2.5 yr accelerator · 3.5 yr networking · 8 yr power · 5 yr backup · 5 yr cooling · 8.8% CAGR · 70 kt/GW			
All parameters aggressive	**35.5 Mt/yr**		
2 yr accelerator · 2.5 yr networking · 6 yr power · 3 yr backup · 4 yr cooling · 10% CAGR · 90 kt/GW			

APPENDIX C — Glossary of Terms

This glossary defines the acronyms and specialized terms used in Part I of The AI e-Waste Wave. It is divided into two sections: Acronyms and Abbreviations, followed by Subject-Matter Terms. Entries are alphabetized within each section.

Acronyms and Abbreviations

AI — Artificial Intelligence

AIM Act — American Innovation and Manufacturing Act — U.S. legislation directing the phase-down of hydrofluorocarbons (HFCs)

BAN — Basel Action Network — an international environmental-justice organization focused on the global trade in toxic waste

CAGR — Compound Annual Growth Rate — the annualized average rate of growth over a specified period

CDU — Coolant Distribution Unit — equipment that circulates liquid coolant to and from server racks in a liquid-cooled data center

CERCLA — Comprehensive Environmental Response, Compensation, and Liability Act — the U.S. "Superfund" law governing cleanup of hazardous-substance releases

CPU — Central Processing Unit — the general-purpose processor in a computer

CRAC — Computer Room Air Conditioning — a precision air-cooling unit designed for data center environments

CRAH — Computer Room Air Handler — similar to a CRAC but uses chilled water rather than a self-contained refrigerant loop

CRT — Cathode Ray Tube — the heavy, leaded-glass display technology used in older televisions and monitors

DIMM — Dual In-line Memory Module — the physical form factor for RAM sticks installed in servers

ESG — Environmental, Social, and Governance — a framework for evaluating corporate sustainability and ethical impact

EU — European Union

GW — Gigawatt — one billion watts; the standard unit used in this paper to measure data center power capacity

GWP — Global Warming Potential — a measure of how much heat a greenhouse gas traps relative to CO_2 over a given period

HBM — High Bandwidth Memory — a type of stacked DRAM used in AI accelerators, permanently bonded to the processor package, making chip-level reuse or repair effectively impossible

HFC — Hydrofluorocarbon — a class of synthetic greenhouse gases used as refrigerants in cooling systems

HTS — High-Temperature Superconducting — a class of materials that conduct electricity with zero resistance at relatively elevated (though still cryogenic) temperatures; used in next-generation power cables

IDC — International Data Corporation — a market-intelligence firm whose PC and server shipment data is widely cited

IEA — International Energy Agency

IoT — Internet of Things — the network of physical devices embedded with sensors, software, and connectivity

ITAD — IT Asset Disposition — the industry and processes for decommissioning, refurbishing, reselling, or recycling end-of-life IT equipment

kt — Kilotonne — one thousand metric tonnes

LCD — Liquid Crystal Display — flat-panel display technology that replaced CRTs

$LiFePO_4$ — Lithium Iron Phosphate — a lithium-ion battery chemistry increasingly used in data center backup systems

Mt — Million Metric Tonnes — unit used in this paper for large-scale e-waste volumes

MTIA — Meta Training and Inference Accelerator — Meta's custom AI chip

MW — Megawatt — one million watts
NPU — Neural Processing Unit — a specialized processor for on-device AI tasks; a key hardware threshold for AI-capable consumer devices
NVMe — Non-Volatile Memory Express — a high-speed storage interface used in modern solid-state drives
OEM — Original Equipment Manufacturer — a company that designs and builds finished hardware products
PCB — Printed Circuit Board — the fiberglass-and-copper board on which electronic components are mounted
PDU — Power Distribution Unit — equipment that distributes electric power to server racks within a data center
PFAS — Per- and Polyfluoroalkyl Substances — a large class of synthetic "forever chemicals" that are extremely persistent in the environment; found in some data center cooling fluids, cable insulation, and circuit board laminates
PFOA — Perfluorooctanoic Acid — a specific PFAS compound designated as a hazardous substance under U.S. CERCLA (Superfund)
PFOS — Perfluorooctane Sulfonate — a specific PFAS compound listed under Annex B of the Stockholm Convention on Persistent Organic Pollutants
PTFE — Polytetrafluoroethylene — a fluoropolymer (a type of PFAS) used in cable insulation and other data center components
PVC — Polyvinyl Chloride — a plastic polymer targeted for phase-out in electronics due to toxic combustion by-products
RAM — Random Access Memory — volatile working memory in a computer
REACH — Registration, Evaluation, Authorisation and Restriction of Chemicals — the EU's comprehensive chemicals regulation, under which a broad PFAS restriction is expected by 2027
SSD — Solid State Drive — a data storage device using flash memory rather than spinning disks
TPU — Tensor Processing Unit — Google's custom AI accelerator chip
TSCA — Toxic Substances Control Act — U.S. federal law governing the manufacture, import, and use of chemical substances
UNITAR — United Nations Institute for Training and Research — the UN body that co-authors the Global E-Waste Monitor
UNU-KEY — United Nations University Key — the 54-category product classification system used by the Global E-Waste Monitor to categorize electronic equipment
UPS — Uninterruptible Power Supply — battery-backed systems that provide emergency power to data center equipment during outages
VRLA — Valve-Regulated Lead-Acid — a sealed lead-acid battery type widely used in data center UPS systems, with a typical lifespan of 3–6 years

Subject-Matter Terms

Accelerator — A processor purpose-built for a narrow class of computation — in this paper, for AI workloads. Examples include NVIDIA GPUs, Google TPUs, and Meta's MTIA chips. Distinguished from general-purpose CPUs by their specialization and, consequently, their limited secondary-use options.
AI Waste Contagion — A term introduced in this paper for the downstream e-waste effect of data center buildout: the accelerated obsolescence of consumer devices, enterprise equipment, and telecommunications infrastructure caused by AI hardware requirements propagating outward from data centers into the broader electronics ecosystem.
Basel Convention — The Basel Convention on the Control of Transboundary Movements of Hazardous Wastes and Their Disposal — the principal international treaty governing the export and import of hazardous waste, including e-waste. Its recent amendments define controlled e-waste as any whole unit or component containing electronic circuits.

Cattle v. Pets — A cloud-computing metaphor coined by Bill Baker of Microsoft and popularized by Randy Bias. "Pets" are servers given individual attention and repair; "cattle" are interchangeable units replaced rather than maintained. The operational culture this doctrine created has material consequences for waste generation.

Circular Economy — An economic model that aims to eliminate waste by keeping products and materials in use as long as possible — through design for longevity, repair, reuse, and recycling — as opposed to the linear "take, make, dispose" model.

Cradle-to-Cradle — A design philosophy in which products are conceived so that, at end-of-life, their materials become inputs for new products of equal or higher value — an advance over the "cradle-to-grave" approach that merely manages disposal.

Data Sovereignty — Laws requiring that data generated within a country's borders be stored on servers physically located in that country, forcing duplication of data center infrastructure across jurisdictions.

Design for Environment — A product-design approach that considers environmental impacts across the entire life cycle — material selection, energy use, repairability, recyclability, and safe end-of-life management.

Direct-to-Chip Liquid Cooling — A cooling method in which liquid coolant is circulated through cold plates attached directly to processors, replacing or supplementing air cooling. Required for AI racks drawing 40 kW or more per rack.

Edge AI — AI processing performed on local devices (cameras, sensors, vehicles) rather than in a centralized data center. Edge AI hardware is a rapidly growing and entirely new e-waste category not tracked by the UN Global E-Waste Monitor.

EU F-Gas Regulation — European Union regulation governing the use and phase-down of fluorinated greenhouse gases, including the HFC refrigerants used in data center cooling systems.

Four-Banded Framework — The BAN model's structure for projecting total global e-waste through 2050: Band 1 (UN baseline conventional e-waste), Band 2 (data center infrastructure), Band 3 (conservative AI contagion), and Band 4 (aggressive AI contagion including cultural/psychological replacement effects).

Frontier Workloads — The most demanding AI tasks — typically training very large models — that require the latest-generation hardware. When a new chip generation arrives, older hardware is no longer economically competitive for frontier workloads, even if it still functions.

Global E-Waste Monitor — A flagship report published by UNITAR (formerly UNU/UNITAR) that provides the most widely cited estimates of global e-waste volumes. The 2024 edition estimated 62 Mt generated in 2022 and projected ~82 Mt by 2030.

Horizontal Scaling — Adding more hardware units working in parallel to increase computing capacity — the dominant scaling strategy for AI, as opposed to vertical scaling (making individual units faster). Also called "scaling out."

Hyperscaler — A company that operates data centers at massive scale — typically the major cloud and AI providers such as Google, Microsoft, Amazon, and Meta. The term reflects the practice of hyperscaling: deploying vast quantities of hardware horizontally.

Hyper-Obsolescence — A term used in this paper to describe the accelerated, simultaneous retirement of large quantities of specialized AI hardware. Because AI infrastructure scales horizontally, a new chip generation does not retire one better machine — it retires hundreds of thousands of parallel machines at once.

Immersion Cooling (Two-Phase) — A cooling method in which electronic components are submerged in a dielectric fluid that absorbs heat by evaporating. Some fluids used are PFAS "forever chemicals," creating hazardous waste disposal challenges.

Inference — The phase of AI operation in which a trained model processes new inputs and produces outputs (e.g., answering a question, generating an image). Distinguished from training, which builds the model. Inference runs continuously and at scale once a model is deployed.

Jevons' Paradox — The observation that improvements in the efficiency of resource use tend to increase, rather than decrease, total consumption — because lower cost per unit of output stimulates greater demand. In this paper, applied to AI chip efficiency: more efficient chips have historically led to larger models and more deployment, not less hardware.

Kigali Amendment — A 2016 amendment to the Montreal Protocol that phases down the production and consumption of HFCs — the refrigerants used in many data center cooling systems.

Mass Intensity — The total mass of physical equipment per unit of installed data center capacity, expressed in this paper as kilotonnes per gigawatt (kt/GW). BAN's estimate is ~70 kt/GW across all five equipment categories.

Moore's Law — The 1965 observation by Intel co-founder Gordon Moore that the number of transistors on an integrated circuit doubles approximately every two years. This pattern held for nearly 80 years but has since reached physical limits, driving the shift to horizontal scaling.

Placed-on-Market Data — Sales and shipment data for electronic products entering the consumer and commercial market — the basis for the UN Global E-Waste Monitor's projections. Does not capture custom-built data center infrastructure procured outside commercial channels.

Refresh Cycle — The interval at which data center operators replace a category of equipment with newer-generation hardware, whether or not the existing equipment has physically failed. AI hardware refresh cycles (2.5–5 years) are significantly compressed relative to traditional IT (5–7+ years).

Rip-and-Replace — The wholesale removal of functional infrastructure to make way for a fundamentally different technology — as distinct from a scheduled refresh. Examples in this paper include the conversion from air cooling to liquid cooling and the anticipated replacement of copper power distribution with superconducting cables.

Stockholm Convention — An international treaty (2001) targeting the elimination or restriction of persistent organic pollutants (POPs). PFOS is listed under Annex B. Relevant to AI e-waste because some data center cooling fluids contain PFAS compounds covered by the Convention.

Stock-and-Flow Model — A modeling approach that tracks the total installed base of equipment (stock) and the rate at which equipment enters and exits the base (flow). BAN's projection uses this method, with equipment entering as new capacity is built and exiting as each category reaches its replacement cycle.

Thermal Runaway — An uncontrolled increase in temperature in a battery cell, potentially leading to fire or explosion. A risk specific to lithium-ion battery chemistries used in data center backup systems, requiring specialized handling and disposal protocols.

Training (AI) — The computationally intensive process of building an AI model by running massive datasets through neural networks. Training a single large model requires thousands of accelerators running in parallel for weeks, and is the primary driver of frontier AI hardware demand.

Vertical Scaling — Increasing computing capacity by making individual hardware units faster and more powerful — the traditional approach governed by Moore's Law. Contrasted with horizontal scaling.

Waste Colonialism — Phrase (with "toxic colonialism" coined by BAN's Jim Puckett in 1988). It means the practice of exporting hazardous or problematic waste from wealthy nations to developing countries to avoid the costs of proper management — a central concern of the Basel Convention and of BAN's work.

Waste Management Hierarchy — The preferential ordering of waste management strategies: prevention first, then reuse, recycling, recovery, and disposal as a last resort. A foundational principle of environmental policy referenced throughout this paper.

Endnotes

i In December 2006 — Apple was first reported ranked last in Greenpeace's Guide to Greener Electronics. In April 2007 — Greenpeace reaffirmed Apple in last place among 14 major with Greenpeace explicitly citing Apple's failure to make progress on recycling and toxic-content policies, and its refusal to provide phase-out timelines for its most toxic substances. In May 2, 2007 — Steve Jobs responded with "A Greener Apple," publicly committing Apple to phase out PVC and brominated flame retardants (BFRs) by 2008 — a more aggressive timeline than competitors like Dell had set. In July 2007 — Greenpeace's next guide nearly doubled Apple's score, from 2.7 to 5.3, crediting the increased transparency and the phase-out commitments.

ii Mark Patel et al., "The Cost of Compute: A $7 Trillion Race to Scale Data Centers," McKinsey Quarterly (McKinsey & Company, April 28, 2025), https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-cost-of-compute-a-7-trillion-dollar-race-to-scale-data-centers

iii UN World Food Programme, 2026 Global Outlook (November 18, 2025), reporting that ending hunger by 2030 would cost $93 billion annually. At that rate, $7 trillion funds approximately 75 years of ending world hunger.

iv Calculation based on UN Department of Economic and Social Affairs, Population Division, World Population Prospects 2024, estimating approximately 140 million global births per year. Over a five-year period (2025–2030, matching McKinsey's own AI data center buildout window), this yields approximately 700 million children; $7 trillion divided evenly among them equals $10,000 per child.

v U.S. Energy Information Administration, "How much electricity does an American home use?," https://www.eia.gov/tools/faqs/faq.php?id=97&t=3 (10,500 kWh/yr average); U.S. Census Bureau, American Housing Survey (approximately 144 million housing units). At an average continuous load of approximately 1.2 kW per home, all U.S. homes require roughly 172 GW — less than the 219 GW of total data center capacity projected by McKinsey.

vi https://cloudscaling.com/blog/cloud-computing/the-history-of-pets-vs-cattle/

vii Programs: https://programs.com/resources/data-center-statistics/; Synergy Research Group, "The World's Total Data Center Capacity is Shifting Rapidly to Hyperscale Operators," Q1 2025, https://www.srgresearch.com/articles/the-worlds-total-data-center-capacity-is-shifting-rapidly-to-hyperscale-operators (1,189 hyperscale data centers at end of Q1 2025); Cloudscene, "Data Centers in United States," https://cloudscene.com/market/data-centers-in-united-states/all; IEA Energy and AI Observatory (2024), https://www.iea.org/data-and-statistics/data-tools/energy-and-ai-observatory.

viii Sightline Climate (now Currence), "Data Center Powering Models Report," August 2025 (data as of July 1, 2025). See also CTVC, "The Data Center Report We Promise You Haven't Read," August 8, 2025, https://www.ctvc.co/the-data-center-report-we-promise-you-havent-read/

ix BloombergNEF, September 2025. "AI Data Center Build Advances at Full Speed," https://about.bnef.com/insights/data-centers/

x Epoch AI, Frontier Data Centers Hub, November 2025, https://epoch.ai/data/data-centers. See also The Data Center Engineer, "Five AI Data Centers to Reach 1 GW Power Capacity in 2026, New Analysis Shows," November 25, 2025, https://thedatacenterengineer.com/news/five-ai-data-centers-to-reach-1-gw-power-capacity-in-2026-new-analysis-shows/

xi JLL, "2026 Global Data Center Outlook," January 2026, https://www.jll.com/en-us/newsroom/global-data-center-sector-to-nearly-double-to-200gw-amid-ai-infrastructure-boom

xii Aterio, "US Data Center Database: 7,800+ Facilities by State & Provider," updated daily, data as of August 28, 2026, https://www.aterio.io/insights/us-data-centers

xiii Cushman & Wakefield, "APAC Data Centre Update: H1 2026," August 2026, https://www.cushmanwakefield.com/en/singapore/insights/apac-data-centre-update

xiv Datacenters.com, "How Much Land Do Hyperscalers Need? Exploring Data Center Real Estate Trends in 2025," 2025. https://www.datacenters.com/news/space-for-servers-how-much-real-estate-do-hyperscalers-really-need

xv Meta Richland Parish (2,250 acres, 4 million sq ft, nine buildings, $10 billion, largest campus by 2027) — Meta/PR Newswire, "Meta Selects Northeast Louisiana as Site of $10 Billion Artificial Intelligence Optimized Data Center That Will Be Company's Largest in the World," November 19, 2024. https://www.prnewswire.com/news-releases/meta-selects-northeast-louisiana-as-site-of-10-billion-artificial-intelligence-optimized-data-center-that-will-be-companys-largest-in-the-world-302322851.html

xvi Vantage Frontier (1,200 acres, 3.7 million sq ft, 10 buildings, $25 billion, first delivery H2 2026) — Vantage Data Centers, "Vantage Data Centers Unveils Plans for 'Frontier,' a $25B Mega-campus in Texas to Meet Unprecedented AI Demand," 2025. https://vantage-dc.com/news/vantage-data-centers-unveils-plans-for-frontier-a-25b-mega-campus-in-texas-to-meet-unprecedented-ai-demand/

xvii Microsoft Mount Pleasant (315 acres originally, Foxconn site) — Microsoft Source, "Microsoft Completes Construction on First Datacenter Facility in Mount Pleasant, Wisconsin," June 23, 2026. https://news.microsoft.com/source/2026/06/23/microsoft-completes-construction-on-first-datacenter-facility-in-mount-pleasant-wisconsin/

xviii Microsoft (2,000+ acres owned) — AOL/Milwaukee Journal Sentinel, "Microsoft Expands Racine County Data Center Holdings. It Just Spent $38 Million on 135 Acres," 2024. https://www.aol.com/articles/microsoft-expands-racine-county-data-170646070.html

xix Microsoft (15 additional buildings, two new campuses, 9 million sq ft, $13 billion+) — Spectrum News 1, "Mount Pleasant Village Board Approves Microsoft Data Center Expansion," January 27, 2026. https://spectrumnews1.com/wi/milwaukee/news/2026/01/27/mount-pleasant–microsoft–data-center-expansion–approval

xx NVIDIA GPU-Ready Data Center technical overview: https://www.nvidia.com/content/g/pdfs/gpu-ready-print-data-center-tech-overview-update-final-web.pdf

xxi Data Center Hardware: The Core Components (2026), Jeff Howell, ENCOR Advisors, October 15, 2025. https://encoradvisors.com/data-center-hardware/

xxii See definitions of Basel controlled e-waste listings Y49, and A1181 in Annex II and VIII respectively in the Basel Convention. While e-waste is not defined in Basel text, it is defined in the Technical guidelines on transboundary movements of electrical and electronic waste and used electrical and electronic equipment, in particular regarding the distinction between waste and non-waste under the Basel Convention adopted on an interim basis, defines electronic equipment as: Electrical and electronic equipment that is dependent on electric currents or electromagnetic fields in order to work properly, including components that can be removed from equipment and can be tested for functionality and either be subsequently directly reused or reused after repair or refurbishment. And Wastes from Electric and Electronic Equipment is defined as: Electrical or electronic equipment that is waste, including all components, sub-assemblies and consumables that are part of the equipment at the time the equipment becomes waste.

xxiii IBM, "What Is an AI Accelerator?," https://www.ibm.com/think/topics/ai-accelerator. See also Amazon, "10 AI Chip Terms to Know," https://www.aboutamazon.com/news/aws/ai-chip-terms-explained.

xxiv CNBC, "Nvidia Sales Are 'Off the Charts,' but Google, Amazon and Others Now Make Their Own Custom AI Chips," Nov. 21, 2025, https://www.cnbc.com/2025/11/21/nvidia-gpus-google-tpus-aws-trainium-comparing-the-top-ai-chips.html.

xxv Epoch AI, "Leading AI Chip Designs Are Used for Around Four Years in Frontier Training," https://epoch.ai/data-insights/gpu-frontier-lifespan (median lifespan 3.9 years from release to final use in frontier training).

xxvi NVIDIA, "DGX SuperPOD Data Center Design Guide — White Space Infrastructure," https://docs.nvidia.com/dgx-superpod/design-guides/dgx-superpod-data-center-design-h100/latest/infrastructure.html (287.6 lbs / 130.45 kg).

xxvii Sunbird DCIM, "Is Your Data Center Ready for the NVIDIA GB200 NVL72?," https://www.sunbirddcim.com/blog/your-data-center-ready-nvidia-gb200-nvl72 (1.36 metric tons / 3,000 lbs).

Xxviii Data Center Knowledge, "Heavy Compute: AI Data Centers Have a Weight Problem," https://www.datacenterknowledge.com/build-design/heavy-compute-ai-data-centers-have-a-weight-problem.

xxix IEA, "Energy Demand from AI," 2025. https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai

xxx EABEL, "AI Data Center Server Rack: Design, Power, and Cooling Explained," July 2026. https://www.eabel.com/ai-data-center-server-rack/

xxxi InvRecovery, "Data Center Decommissioning in 2026: Navigating AI Infrastructure Upgrades," 2026. https://invrecovery.org/data-center-decommissioning-in-2026-navigating-ai-infrastructure-upgrades/

xxxii Discovery Alert, "Copper Demand from Data Centers: AI Infrastructure Growth," 2026. https://discoveryalert.com.au/electrical-infrastructure-requirements-data-center-2026/; "Scaling metals to secure the data centre materials backbone," December 2025, https://www.weforum.org/stories/2025/12/securing-data-centre-materials/ (60–75 tonnes of minerals per MW; facility-only copper ~12 tonnes/MW vs. ~26 tonnes/MW including near-site connections).

Xxxiii Eaton, "Eaton 9395 UPS Specifications," https://www.eatonguard.com/eaton-9395-spec.asp.

xxxiv ROC Telecom, "How GPU Decommissioning Differs from Standard Server Retirement," May 26, 2026. https://roctelecom.com/insights/gpu-decommissioning-vs-server-retirement/

xxxv Delta Power Solutions, "The Pros and Cons of Lithium-Ion Batteries for Data Centers," 2025. https://www.deltapowersolutions.com/en/mcis/technical-article-the-pros-and-cons-of-lithium-Ion-batteries-for-data-centers.php

xxxvi Introl, "UPS and Battery Systems: Power Protection for AI Data Centers," 2025. https://introl.com/blog/ups-battery-systems-lithium-ion-ai-data-center-2025

Xxxvii Schneider Electric, "Battery Technology for Data Centers and Network Rooms," https://download.schneider-electric.com/files?p_Doc_Ref=SPD_SADE-5TNRJR_EN (VRLA at 30–50 Wh/kg; 250–330 tonnes is derived from this energy density for a 10 MWh installation).

Xxxviii MCF Environmental Services, "What Hazardous Materials Are Found in a Data Center?," August 28, 2025. https://mcfenvironmental.com/what-hazardous-materials-are-found-data-center/

xxxix The relevant Basel Convention provisions are Annex VIII, List A, entries A1160 (lead-acid batteries) and A1170 (unsorted waste batteries). Any attempts to use Annex IX listings such as B1090 cannot accommodate the Annex IX chapeau as Lithium Ion batteries contain solvents, which are flammable and/or explosive so will not pass the threshold of having hazardous characteristics and hazardous constituents at the same time. https://www.basel.int/TheConvention/Overview/TextoftheConvention/tabid/1275/Default.aspx

xl Latitude Media, "The data center boom is a diesel generator boom," 2026, https://www.latitudemedia.com/news/the-data-center-boom-is-a-diesel-generator-boom/. Amazon's Manassas, VA facility deploys 93 Caterpillar 3516C-HD generators (2.5 MW standby each) for ~232 MW total backup capacity. See also California Energy Commission, "Sequoia Data Center," SPPE Docket 19-SPPE-03, https://www.energy.ca.gov/powerplant/reciprocating-engine/sequoia-data-center (54 diesel generators for 96.5 MW). Generator weight: Caterpillar, "3516C (60 Hz) Diesel Generator Set Specifications," https://www.cat.com/en_US/products/new/power-systems/electric-power/diesel-generator-sets/1000028912.html (dry weight 16,275–17,590 kg).

xli MacAllister Power Systems, "100 Ton Chiller Unit Spec Sheet," https://www.macallisterpowersystems.com/wp-content/uploads/sites/4/2022/11/100-Ton-Chiller-Unit-Spec-Sheet.pdf (operating weight ~10,723 lbs / 4,865 kg).

xlii Kigali Amendment to the Montreal Protocol on Substances that Deplete the Ozone Layer, adopted October 15, 2016, entered into force January 1, 2019. See UNEP OzonAction, "The Kigali Amendment to the Montreal Protocol: HFC Phase-down," https://www.unep.org/ozonaction/resources/factsheet/kigali-amendment-montreal-protocol-hfc-phase-down

xliii Regulation (EU) 2024/573 of the European Parliament and of the Council of 7 February 2024 on fluorinated greenhouse gases, Official Journal of the European Union, 20 February 2024. https://eur-lex.europa.eu/eli/reg/2024/573/oj/eng

xliv American Innovation and Manufacturing Act of 2020 (AIM Act), 42 U.S.C. § 7675. Mandates 85% reduction in US HFC production and consumption by 2036. See EPA, "Enforcement of the American Innovation and Manufacturing Act of 2020," https://epa.gov/enforcement/enforcement-american-innovation-and-manufacturing-act-2020

xlv Introl, "Cooling Distribution Units (CDU) for Liquid Cooling in AI Data Centers," 2025, https://introl.com/blog/cooling-distribution-units-cdu-liquid-cooling-ai-data-center-2025. See also Vertiv, "CoolChip CDU 350 kW," https://www.vertiv.com/48ee82/globalassets/products/thermal-management/high-density-solutions/vertiv-coolchip-cdu/vertiv-coolchip-cdu-350kw-data-sheet-sl-71957.pdf

xlvi Introl, "Retrofitting Legacy Data Centers for AI," https://introl.com/blog/retrofitting-legacy-data-centers-ai-liquid-cooling-integration

xlvii Dober, "PG 25 Coolant for Data Centers," https://www.dober.com/performance-fluids/resources/pg-25-coolant-data-centers. See also Hydratech, "The Use of Glycol in Data Centre Cooling Systems," https://www.hydratech.co.uk/technical-blog/The-Use-of-Glycol-in-Data-Centre-Cooling-Systems/128

xlviii Schneider Electric Blog, "PFAS Phase-Out and Liquid Cooling: What US Data Center Operators Must Do," May 22, 2026, https://blog.se.com/datacenter/2026/05/22/pfas-phase-out-liquid-cooling-us-data-center-operators-must-do/. PFOS listed under Annex B of the Stockholm Convention by Decision SC-4/17 (2009). See Stockholm Convention, "PFAS Overview," https://chm.pops.int/Implementation/IndustrialPOPs/PFAS/Overview/tabid/5221/Default.aspx

xlix Basel Convention guidance on management and destruction of PFAS-containing wastes. See also EESI, "Data Centers Are Contributing to PFAS Forever Chemical Pollution," https://www.eesi.org/articles/view/data-centers-are-contributing-to-pfas-forever-chemical-pollution

l European Chemicals Agency (ECHA), updated PFAS restriction proposal under REACH, published 20 August 2025. https://www.echa.europa.eu/-/echa-publishes-updated-pfas-restriction-proposal

li Schneider Electric (2026), ibid. See also ServeTheHome, "2-Phase Immersion Cooling Halted Over Multi-Billion Dollar Health Hazard Lawsuits," https://www.servethehome.com/2-phase-immersion-cooling-halted-over-multi-billion-dollar-health-hazard-lawsuits/

lii See xxxii, xxxiii.

liii See xxvii

liv nVent, "100 MW Hyperscale AI Reference Architecture Data Center," 2025. https://www.nvent.com/en-us/data-solutions/100-mw-hyperscale-ai-data-center-blueprint. See also Introl, "High-Density Racks: 100kW+ Designs for AI Data Centers," 2025. https://introl.com/blog/high-density-racks-100kw-ai-data-center-ocp-2025

lv Falcon UPS, "Lithium-Ion UPS Battery vs. VRLA: The TCO Equation for Modern Data Centers." https://falconups.com/lithium-ion-ups-battery-vs-vrla-tco/. See also MANLY Battery, "When Should Data Centers Plan a Lithium Battery Replacement for Aging UPS Battery Banks," https://manlybattery.com/when-should-data-centers-plan-a-lithium-battery-replacement-for-aging-ups-battery-banks/

lvi Introl, "800G Networking for AI," 2025. https://introl.com/blog/800g-networking-ai-gpu-fabric-planning-2025. See also FS.com, "Upgrading Data Center Networks: When to Upgrade from 100G to 400G/800G Switches." https://www.fs.com/blog/data-center-switches-timing-the-move-to-400g800g-9152.html

lvii Basel Convention, "E-waste Amendments Overview." https://www.basel.int/Implementation/Ewaste/EwasteAmendments/Overview/tabid/9266/Default.aspx. As of 1 January 2025, all e-waste is subject to Basel Convention control: hazardous e-waste under A1181 (Annex VIII), all other e-waste under Y49 (Annex II). See also US EPA, "New International Requirements for Electrical and Electronic Waste." https://www.epa.gov/hwgenerators/new-international-requirements-electrical-and-electronic-waste

lviii See xxxii

lix 62 Mt (2022) → 82 Mt (2030) implies a 3.56% CAGR, giving ~165 Mt by 2050 (extrapolating provided by BAN)

lx While the Global e-Waste Monitor 2024 (ITU/UNITAR, Geneva/Bonn, 2024). https://www.itu.int/en/ITU-D/Environment/Documents/Publications/2025/d-gen-e_waste.01-2024-pdf-e.pdf, itself does not use the word 'crisis' to describe the e-waste problem, the author of this paper has not such qualms and notes with great concern, that the scope and level of the crisis has yet to be fully comprehended.

lxi Cornelis P. Baldé et al., The Global E-waste Monitor 2024 (ITU/UNITAR, Geneva/Bonn, 2024), Chapter 1 ("What is EEE and E-waste?") and Chapter 2 ("Methodology"), https://www.itu.int/en/ITU-D/Environment/Documents/Publications/2025/d-gen-e_waste.01-2024-pdf-e.pdf

lxii Kees Baldé, quoted in James O'Donnell, "AI Will Add to the E-Waste Problem. Here's What We Can Do About It," MIT Technology Review, October 28, 2024, https://www.technologyreview.com/2024/10/28/1106316/ai-e-waste/

lxiii Peng Wang, Ling-Yu Zhang, Amir Tzachor, and Wei-Qiang Chen, "E-waste Challenges of Generative Artificial Intelligence," Nature Computational Science 4, no. 11 (2024): 818–823, https://doi.org/10.1038/s43588-024-00712-6. The study explicitly notes it "considers only large language models, not other forms of generative AI."

lxiv Ralph Hintemann, Simon Hinterholzer, and Fedor Konrat, "Server Stock Data — A Basis for Determining the Energy and Resource Requirements of Data Centres" (conference paper, Borderstep Institute for Innovation and Sustainability, Berlin, June 2024). Available via ResearchGate and Borderstep publications at https://www.borderstep.de/facts-and-figures/zahlen-aus-der-welt-der-rechenzentren/

lxv Alex de Vries-Gao, "Recalibrating Global Artificial Intelligence E-Waste Estimates," Resources, Conservation and Recycling 215 (April 2026), https://www.sciencedirect.com/science/article/pii/S0921344926000960. The article was published online February 24, 2026.

lxvi Jens Gröger et al., Environmental Impacts of Artificial Intelligence: Evaluation of Current Trends and Compilation of an Overview Study (Öko-Institut, commissioned by Greenpeace Germany, May 2025), https://www.oeko.de/fileadmin/oekodoc/Report_KI_ENG.pdf. The report reviewed more than 95 sources. Its "up to 5 million tonnes" headline figure repeats the Wang et al. estimate; the original Gröger et al. calculation of approximately 4.2 million tonnes uses a separate methodology based on estimated global server stock and a four-year average lifespan, which the authors themselves describe as a "rough estimate."

lxvii Baseline of 122 GW from IEA Energy and AI Observatory (2024), https://www.iea.org/data-and-statistics/data-tools/energy-and-ai-observatory. Projections: McKinsey, 171–219 GW by 2030 ('AI Power,' 2025), https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/ai-power-expanding-data-center-capacity-to-meet-growing-demand; JLL, 200 GW by 2030 ('Global Data Center Outlook,' 2026), https://www.jll.com/en-us/newsroom/global-data-center-sector-to-nearly-double-to-200gw-amid-ai-infrastructure-boom; ABI Research, 277 GW by 2035 ('Hyperscaler Data Center Capacity,' May 2026), https://www.abiresearch.com/press/hyperscaler-data-center-capacity-to-surge-more-than-6x-by-2035-as-ai-and-cloud-expansion-reshape-global-infrastructure. The 8.8% CAGR is the average of the implied growth rates from 122 GW to each target.

lxviii World Economic Forum, "From Minerals to Megawatts: Building Resilience for EVs, Data Centres," December 2025, https://reports.weforum.org/docs/WEF_From_Minerals_to_Megawatts_2025.pdf. See also WEF, "Scaling metals to secure the data centre materials backbone," December 2025, https://www.weforum.org/stories/2025/12/securing-data-centre-materials/ (60–75 tonnes of minerals per MW; facility-only copper ~12 tonnes/MW vs. ~26 tonnes/MW including near-site connections).

lxix Mass intensity of 70 kt/GW derived from Section 2 component data (range 62–77 kt/GW). Cross-validated against Microsoft's Chicago facility copper data: World Economic Forum, 'Scaling Metals to Secure the Data Centre Materials Backbone,' 2025, https://www.weforum.org/stories/manufacturing-and-value-chains/securing-data-centre-materials/. On rising mass intensity: Lean Research, 2026, https://www.leanrs.com/insights/data-center-equipment-growth; DataX Connect, 2025, https://dataxconnect.com/insights-why-are-data-centres-getting-so-heavy/.

lxx Gartner, 'x86 Server Life Cycle Planning,' https://www.gartner.com/en/documents/924414; Dell'Oro Group, 'Hyperscale AI Investment Cycle Anchors Accelerator-Led Infrastructure Spending,' 2026, https://www.delloro.com/news/hyperscale-ai-investment-cycle-anchors-accelerator-led-infrastructure-spending/; Resource Recycling, 'AI Servers Reshape ITAD Sector,' March 2026, https://resource-recycling.com/e-scrap/2026/03/09/ai-servers-reshape-itad-sector-recyclers-brace-for-new-wave/; de Vries-Gao, 'Recalibrating Global Artificial Intelligence E-Waste Estimates,' ScienceDirect, 2026, https://www.sciencedirect.com/science/article/pii/S0921344926000960. Amazon shortened server depreciation from six to five years effective January 2025, citing AI's pace: SEC Form 10-Q, https://www.sec.gov/Archives/edgar/data/1018724/000101872425000086/amzn-20250630.htm.

lxxi ROC Telecom, 'How Long Does Networking Equipment Really Last?,' 2026, https://roctelecom.com/insights/how-long-does-networking-equipment-really-last/; ROC Telecom, 'Why AI Is Shortening Your Hardware Lifecycles,' 2026, https://roctelecom.com/insights/why-ai-is-shortening-your-hardware-lifecycles-and-what-to-do-about-it/."

lxxii Dataconomy, 'The Power Problem Behind the AI Boom,' August 2026, https://dataconomy.com/2026/08/04/the-power-problem-behind-the-ai-boom-why-data-centers-are-racing-to-upgrade-their-electrical-infrastructure/; JLL, '2026 Global Data Center Outlook,' https://www.jll.com/en-us/insights/market-outlook/data-center-outlook. On VEIR: Microsoft Azure Blog, 'Can High-Temperature Superconductors Transform the Power Infrastructure of Datacenters?,' https://azure.microsoft.com/en-us/blog/can-high-temperature-superconductors-transform-the-power-infrastructure-of-datacenters/. Eight years is a BAN estimate; no published source gives a specific shortened lifespan.

lxxiii Battery replacement cycles are industry standard (VRLA 3–6 years, lithium-ion 10+). On sodium-ion commercial deployment in 2027: Peak Energy/Sacramento plant, https://www.intelligentliving.co/first-sodium-ion-battery-plant-sacramento/; BYD solid-state and sodium-ion for 2027, https://carnewschina.com/2026/02/09/byd-breaks-ground-with-solid-state-battery-and-10000-cycle-sodium-tech-power-2027-evs/.

lxxiv TechPowerUp, 'Liquid Cooling to Scale in AI Data Centers, Penetration to Surpass 30% in 2025,' https://www.techpowerup.com/340164/liquid-cooling-to-scale-in-ai-data-centers-penetration-to-surpass-30-in-2025; Grand View Research, 'Data Center Liquid Cooling Market Report 2026–2033,' https://www.grandviewresearch.com/industry-analysis/data-center-liquid-cooling-market-report. Five years is a BAN estimate.

lxxv See lvii.

lxxvi Amoah et al., "Mineral demand from AI data centers," Resources Policy, 2026

Lxxvii Microsoft, "Copilot+ PCs," https://www.microsoft.com/en-us/windows/learning-center/copilot-plus-pcs-windows-pcs-differences

Lxxviii Gartner, "Gartner Forecasts Worldwide Artificial Intelligence Chips Revenue to Grow 33% in 2024," May 29, 2024, https://www.gartner.com/en/newsroom/press-releases/2024-05-29-gartner-forecasts-worldwide-artificial-intelligence-chips-revenue-to-grow-33-percent-in-2024. See also CIO Dive, "Enterprises Will Go All-In on AI PCs by 2026, Gartner Predicts," May 29, 2024, https://www.ciodive.com/news/AI-PC-enterprise-windows-dell-gartner/717437/.

lxxix IDC, "AI-Enabled Workforce," Analyst Brief, 2024, https://www.amd.com/content/dam/amd/en/documents/products/processors/business-systems/idc-ai-pc-analyst-brief.pdf

lxxx Apple, "How to Get Apple Intelligence," https://support.apple.com/en-us/121115

lxxxi Google, Gemini Intelligence hardware requirements, 2026, https://www.techradar.com/phones/android/gemini-intelligence-hardware-requirements-revealed

Lxxxii MarketsandMarkets, "Edge AI Hardware Market," 2025, https://www.marketsandmarkets.com/Market-Reports/edge-ai-hardware-market-158498281.html

Lxxxiii BAN estimate. No published lifecycle study for edge AI devices exists.

Lxxxiv Dell'Oro Group via IEEE ComSoc, March 2023, https://techblog.comsoc.org/2023/03/20/delloro-worldwide-telecom-equipment-market-growth-3-in-2022-mtn-2-growth-in-2022/; The Business Research Company, "Optical Networking Global Market Report," 2026, https://www.thebusinessresearchcompany.com/report/optical-networking-global-market-report

lxxxv Fiber Broadband Association, July 2025, https://fiberbroadband.org/2025/07/31/fiber-broadband-association-research-reveals-critical-need-for-fiber-in-ai-and-data-center-growth/

Lxxxvi Dell'Oro Group, "Worldwide Telecom Equipment Up 3 Percent in 2022," March 2023, via IEEE ComSoc Technology Blog, https://techblog.comsoc.org/2023/03/20/delloro-worldwide-telecom-equipment-market-growth-3-in-2022-mtn-2-growth-in-2022/. Total worldwide telecom equipment revenues approached $100 billion in 2021, with 3 percent growth in 2022.

Lxxxvii A blended estimate reflecting equipment ranging from base stations at approximately $250 per kilogram to optical transponders at several thousand dollars per kilogram.

Lxxxviii A BAN estimate, slightly above the sourced optical networking range, based on the judgment that the full telecom equipment market (not just optical) is being pushed harder by AI demand than the optical sub-segment alone; Mordor Intelligence, "Optical Networking and Communications Market," 2025, https://www.mordorintelligence.com/industry-reports/optical-networking-and-communications-market. Market forecast at 7.19 percent CAGR to 2030, with AI-driven data center interconnect as the fastest-growing sub-segment.

Lxxxix See lxxxvi, lxxxiv

xc National Telecommunications and Information Administration (NTIA), U.S. Department of Commerce, "Fact Sheet: Useful Life Schedule," November 26, 2024, https://www.ntia.gov/other-publication/2024/fact-sheet-useful-life-schedule-0. Assigns seven-year useful life to broadband switching, routing, and transport equipment, wireless base stations, and antennas.

xci The Monitor's 82 Mt projection for 2030 is based on historical EEE placed-on-market trends (see Annex 1: Methodology Details). AI is mentioned once in the Foreword but plays no role in the projection model. BAN's contagion bands therefore represent acceleration above a pre-AI baseline.

xcii Zhifei Li, Tian Xia, Ziming Mao, Zihan Zhou et al., "SkyNomad: On Using Multi-Region Spot Instances to Minimize AI Batch Job Cost" (UC Berkeley, arXiv:2601.06520, January 2026), https://arxiv.org/abs/2601.06520. Figure 12 in the paper evaluates AI workloads under geographic constraints: Europe-only placement cost 46% more than global placement, and Asia-only cost 34% more than global placement, demonstrating the efficiency loss from geographic fragmentation.

xciii The Next Platform / The Register, "Stop Measuring AI Training Costs in GPU Hours," April 23, 2026, https://www.nextplatform.com/cloud/2026/04/23/stop-measuring-ai-training-costs-in-gpu-hours/

xciv Resetting GPU depreciation: Why AI factories bend, but don't break, useful life assumptions — SiliconANGLE, November 22, 2025

xcv NVIDIA, "NVIDIA GB200 NVL72," 2024, https://www.nvidia.com/en-us/data-center/gb200-nvl72/

xcvi DeepSeek, "DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning," January 20, 2025. DeepSeek reported $5.6 million in GPU compute costs for training the V3 base model; R1 matched OpenAI's o1 on math and coding benchmarks. See Epoch AI, https://epoch.ai/gradient-updates/what-went-into-training-deepseek-r1

xcvii CNBC, "Nvidia Sheds Almost $600 Billion in Market Cap, Biggest One-Day Loss in U.S. History," January 27, 2025, https://www.cnbc.com/2025/01/27/nvidia-sheds-almost-600-billion-in-market-cap-biggest-drop-ever.html

Xcviii BAN inference based in large part on many sources above and the following serge point. Deloitte: "AI Data Center Power Demand Could Surge 30x by 2035, Amid Power and Grid Capacity Constraints" June 24, 2025. https://www.prnewswire.com/news-releases/deloitte-ai-data-center-power-demand-could-surge-30x-by-2035-amid-power-and-grid-capacity-constraints-302488877.html