AI and Sustainability at Akila:

A Carbon Assessment

WHITE PAPER

Preferred Partner

Contents

Introduction

AI and Sustainability at Akila – a carbon assessment

AI consumes electricity, and that electricity carries emissions — a cost now under real scrutiny from clients, regulators and the public. But a footprint is only one side of the ledger. The question worth asking is not simply what AI costs, but what it returns.

This paper assesses Akila's AI-related carbon impact in 2025 and its trajectory through 2026, as we deepen the deployment of agents across the built world.

The finding is consistent, and robust to the assumptions behind it: Akila's AI already pays back its own footprint by a wide margin. In 2025 that payback came overwhelmingly from energy-optimization agents; in 2026 we are extending the same approach to compliance, risk and territorial benchmarking and more, which we expect to broaden it further.

Every figure here is built from the ground up wherever possible — metered usage and measured savings, with published benchmarks only where direct measurement isn't yet available.

- **Carbon footprint** — the emissions from the electricity that trains and runs our models, measured bottom-up from actual usage and each location's real grid.
- **Carbon payback** — the emissions avoided elsewhere because the AI did its job: in 2025, chiefly energy not consumed in the buildings it optimises. In carbon-accounting terms, these are avoided emissions.

Context of study

Drawn from 2025 data across seven clients in different sectors — 32 sites in France, Singapore and China.

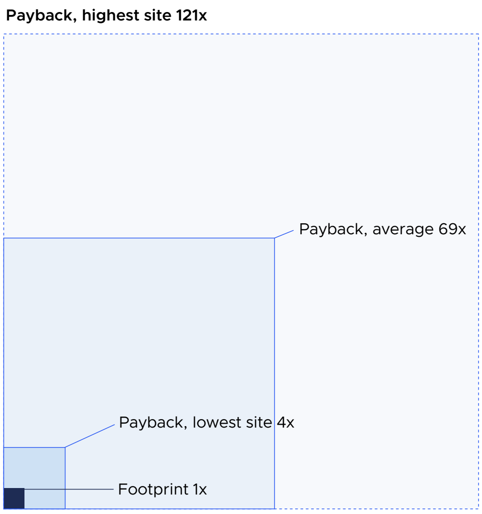
- at least ~**4×** the footprint on the cleanest grid
- up to ~**121×** on the most carbon-intensive
- **69×** on average across sites

The payback from energy agents is clear:

In every deployment, the emissions avoided comfortably exceeded the AI's own footprint.

How did we arrive at these figures?
Find our methodology on page 08-09

Sectors included in our data
Chemicals | Retail | Logistics | MedTech | Commercial Real Estate



AI consumption breakdown

The AI deployed in client buildings is deliberately lightweight and already high-impact — across live sites it returned far more carbon than it consumed, measured on each site's real grid.

Half of our current AI consumption however is something else: R&D. Spending here is key because we're actively building the next generation of agents for the built world — a deliberate investment, de-risked by the return our deployed AI has already proven.

We Measure It	It Pays Back	More Is Coming
Rigorous, bottom-up carbon accounting	Proven energy return from HVAC AI	A new generation of agents in build

Footprint shares from Akila's bottom-up AI carbon model (2026).
Shown as proportions: full methodology and data available on request.

Share of our carbon footprint	Scope of activity		Rationale & benefits
51% / 13.3 t	**Facilities Optimization**	Hardware deployed at client site to optimize their facility equipment control. Footprint includes manufacturing and electricity usage.	**PROVEN TODAY** HVAC optimization 4-121x more carbon saved than the AI spends. Across seven live deployments, our heating & cooling AI returned far more energy than it consumed measured on each site's real grid, not projected.
45% / 11.7 t	**R&D AI tools**	Use of agentic tools provided by world's leading AI companies to build our own platform.	**BUILDING TOMORROW** a new generation of agents Today's heavy AI use is investment - our R&D is building agents that extend value across the whole building: • 3D digital-twin assistant • Site-risk & maintenance • Portfolio energy analysis • Audit & due-diligence prep • Refrigeration optimization • Territory benchmarking *...and beyond - every client can bring their own question.*
3% / 0.85 t	**Platform AI Insights**	Already deployed AI features on the platform, such as energy AI insights at site and portfolio level.	In line with previous category, those existing features already accelerate identification of energy waste and cost reduction opportunities.
<1% / 0.03 t	**Office / Other Internal Use**	Internal use of AI tools for non-dev daily office tasks (less token-intensive than development).	Those tools strongly increase productivity for our teams in their daily tasks.

Footprint shares from Akila's bottom-up AI carbon model (2026).
Shown as proportions: full methodology and data available on request.

Carbon payback is proven in sector after sector

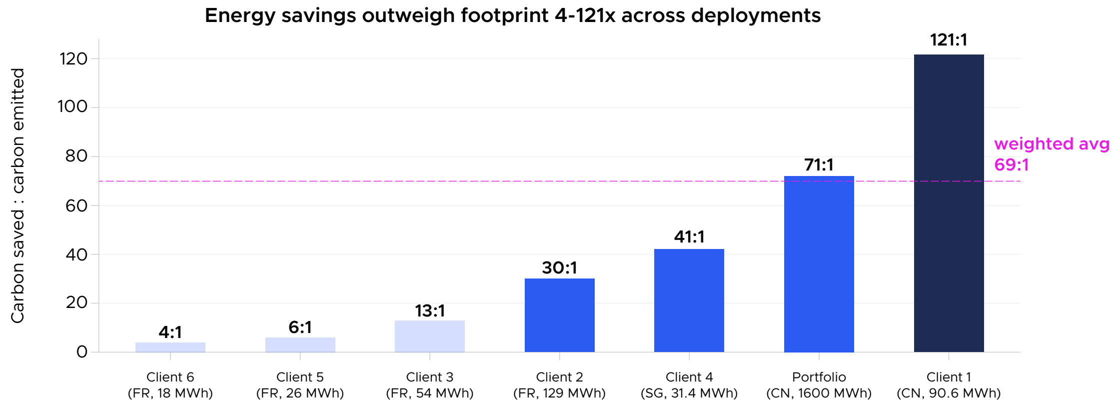
The proven return our first AI already delivered, and a concrete roadmap of what the next generation does. We are not promising value in the abstract; we can name where it comes from.

2025: AI that paid for itself, many times over

Our first production AI did one job: optimise HVAC — the heating, ventilation and cooling that dominates a building's energy use. The result was not subtle. For every unit of carbon spent running the AI, the optimization saved far more.

Across seven sample clients, savings outweighed the AI's own footprint by 4 to 121 times.

The return is largest on carbon-heavy grids and still strongly positive on clean ones such as France. This is measured, not projected — from live sites, using each location's real grid emissions, and accounting for embodied emissions of hardware.



Our results show that carbon payback is not restricted to any one sector or geography.

Client 1	China	Chemical manufacturing	Specialty R&D center combining office and laboratory environments. High cooling and climate-control needs due to sensitive research conditions.
Client 2	France	Retail site	Large-format retail environment with significant HVAC demand across customer-facing and operational spaces.
Client 3	France	Shopping center	Operated by a leading European commercial real estate group. Multiple functional zones with varied cooling, ventilation, lighting, and hot water loads.
Client 4	Singapore	Warehouse	Multi-level site, significant cooling demand across operating & storage areas. Chiller plant and air-handling systems, alongside solar PV and smart lighting.
Client 5	France	HealthTech	Life sciences manufacturing campus focused on biologicals and diagnostics. The site is part of a global operational footprint spanning multiple regions.
Client 6	France	Rail station	Major railway station combining passenger areas, technical rooms, and infrastructure systems with complex operational and comfort requirements.
Client 7	China	Big-box retail portfolio	A national portfolio of big-box retail stores and shopping malls. Together, the sites represent one of the client's largest country-level retail footprints.

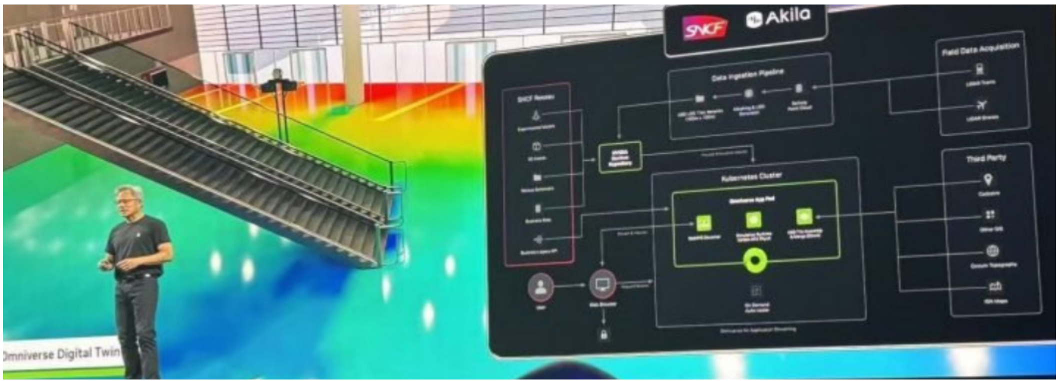
2026

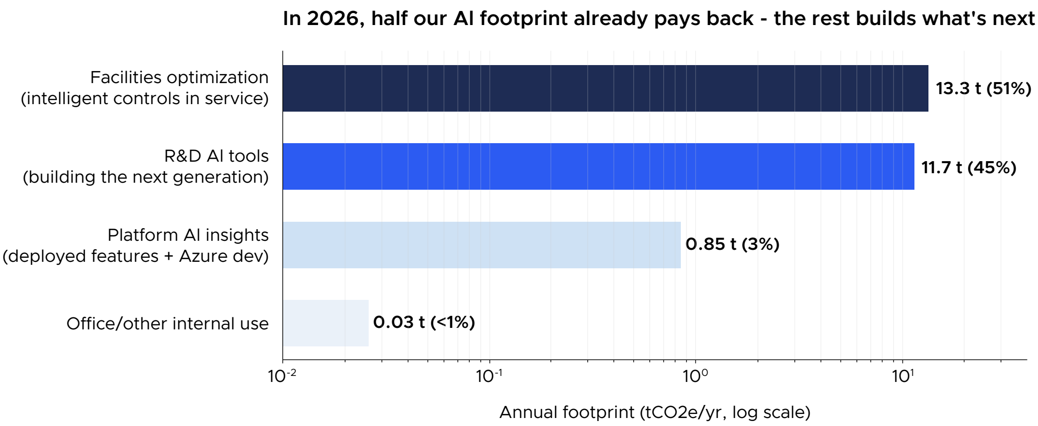
A system that already pays back, plus the build for what’s next

In 2026 our AI consumption rose, and it splits into two kinds of spend. The first is the operating cost of AI that already works in the field — chiefly the intelligent facilities control running on client sites — which sits firmly on the profitable side of the ledger.

The second is our R&D team building the next generation of the platform: tools that work continuously against the entire Akila building model, which is computationally heavy. Deployed-feature usage beyond the proven controls is still small — most features are not yet live at scale.

So today our AI footprint is split almost evenly between two halves: systems already deployed and already paying back, and the R&D build of what comes next. Only the second half is a forward-looking bet — and it is now a bounded one, on the same order as our proven, profitable base rather than dwarfing it. We are honest about what this does and does not show: the next generation’s returns will be measured once these tools reach client sites at scale, the way the first ones were.





Optimism here is not hope — it rests on two things: a proven return we are still earning today from AI already in the field, and a concrete roadmap of what the next generation does. We are not promising value in the abstract; we can name where it comes from.

Where the next return comes from

Selected AI agent projects

Where 2025 AI did one thing (save energy), the 2026 generation spreads across building management — cutting not only energy, but repair costs, consultancy and legal fees, analysis time, and energy in systems beyond HVAC. A selection of what we are building:

Agent	What it does	Value created
3D digital-twin assistant	Natural-language questions answered directly against the building's live data and 3D model (equipment, occupancy, air quality, energy, work orders) — already in production testing.	**Faster operational decisions; less time hunting through dashboards**
IFM site-risk agent	Surfaces recurring equipment failures, inspection gaps and mis-prioritised work orders; warns before issues escalate.	**Lower repair cost; longer asset life; better preventive maintenance**
Territory benchmarking	Combines portfolio data with public building databases (energy performance, solar potential, accessibility, safety compliance).	**Sharper positioning and investment decisions per building**
Portfolio energy analysis	Compares sites across a client's portfolio, flagging abnormal or inconsistent patterns for HQ directly.	**Energy savings identified at portfolio scale**
Audit & due-diligence prep	Centralises unstructured records (contracts, safety, leases) against the digital twin for M&A, transactions, ISO certification.	**Reduced consulting and legal fees; faster transactions**
Refrigeration optimization	Applies machine learning to refrigeration control, extending the optimization approach beyond HVAC.	**Significant energy savings on a new system type**

...and beyond. This list is not exhaustive. Once a building's live and historical data sits in one place alongside external public data — weather, climate, territory and government databases — each client can bring their own question. That is the point of an AI-native platform.

The throughline: we measure, then we trust the measurement

The same discipline runs through both years. We measured our first AI's return honestly, including its footprint, using real grid data rather than convenient assumptions — and we hold the next generation to the same test. What began as intelligent control of energy is becoming an AI-native platform for the cognitive building: one that reduces not just the energy a building consumes, but the time and cost of running it. The proven return gives us the standing to build toward that horizon — and the method is what makes the claim worth trusting.

Methodology Appendix

How we calculate Akila's AI carbon footprint · companion to 'AI, Energy and Climate at Akila' · June 2026

Approach. A bottom-up model. Software AI = tokens × energy-per-token × data-centre overhead × grid factor, plus a uniform training-amortisation uplift. On-site hardware = estimated electricity × grid factor, plus amortised manufacturing carbon. Emissions follow the grid where each workload physically runs, not where the user sits. Public figures are proportions and ratios; the full workbook is available on request.

1. Activity data

AI system	Activity metric & source	Data Quality Tier
Akila × GPT — features	Input/output tokens, Azure OpenAI API logs; Jan–Jun 2026 actuals annualised ×2.20	**T1**
R&D feature development	Input/output tokens, Azure OpenAI API logs; same annualisation	**T1**
R&D developer AI tools	Input, output AND cache tokens, R&D dashboard: 44 days of June 2026 actuals (token quantity provided by suppliers, majority are cache tokens), annualised ×8.30	**T1**
Claude — internal office	Seat proxy: 25 users × ~30k tokens/day	**T3**
Microsoft Copilot	Seat proxy: 80 users × 8 interactions/day, M365 Admin Centre	**T3**
HVAC control servers (7 sites)	Electricity: nominal power × 25% usage × running days; savings client-validated; + amortised embodied carbon	**T1**

Azure feature rows annualised from 166 days of 2026 actuals; the developer-tool row uses 44 days of directly-metered input, output and cache tokens (June 2026). Seat-proxy rows are interim.

2. Emission factors & energy parameters

Factor	Value	Source
Grid factors (FR/US/CN/SG/JP)	0.041 / 0.384 / 0.526 / 0.497 / 0.477 $kgCO_2e/kWh$	Ember Global Electricity Review 2025
Input token energy	0.22 Wh / 1k	Epoch AI (2025), from long-input figure
Output token energy	0.60 Wh / 1k	Epoch AI (2025): 0.3 Wh / 500-output query; cross-checked vs OpenAI & Jegham
Cache token energy	0.022 Wh / 1k (10% of input)	Provider cache: input pricing (DeepSeek/Anthropic); ~5.4% of decode energy (Plasmon 2024)
Training uplift	×1.125 (uniform)	3-mo training vs 24-mo inference at comparable power (Epoch AI 2025)
Data-centre PUE	1.15 (unified)	Hyperscaler-class published range
Server embodied carbon	2,100 $kgCO_2e$ → 157.5 kg/yr (4-yr, 30% non-use)	Dell PowerEdge R250 LCA
AI-fleet embodied carbon	Neglected (<5%)	Towards Sustainable LLM Serving, arXiv 2501.01990

Energy-per-token is applied uniformly across providers (no robust Western-vs-Chinese efficiency comparison exists at 2026).
Input, output and cache are priced separately as genuinely different operations — computing context, generating text, and re-reading cached context.

2.1 Cache tokens & training amortisation

Cache tokens are memory reads of pre-computed context (the KV cache), not generation — so they consume a small fraction of a token's energy. We anchor this to the provider cache: input pricing ratio (DeepSeek and Anthropic bill cache reads at 10% of input → 0.022 Wh/1k), independently corroborated by a hardware analysis (~5.4% of decode energy, Plasmon 2024).

Training is a one-off upstream cost every token benefits from equally, amortised as a uniform +12.5% uplift on all inference energy. The figure compares a GPT-4-class run (~3 months at 20–25 MW, Epoch AI 2025) against a ~24-month inference service life at comparable power (current ~2.5 bn queries/day implies inference power near training power). We investigated this explicitly rather than excluding it as most public estimates do; the result is consistent with the literature's finding that training is smaller than inference over a model's life.

Embodied hardware carbon for the AI fleet is estimated below 5% of operational emissions and omitted; on-premise HVAC control hardware does include it, as its lower utilisation makes manufacturing dominate, especially in low-carbon grid countries (§1).

3. Data-quality tiers & key assumptions

Confidence tags: **T1** measured/official · **T2** published literature · **T3** internal proxy.

- **Grid routing.** Workloads mapped to where the server runs, not the user. Azure workloads split evenly across FR/US/JP/SG; R&D personal-account tools route Chinese models → China grid, Western models → US grid.
- **Developer AI volume (T1).** Input, output and cache now taken directly from 44 days of dashboard actuals, replacing the earlier proxy; the same split applies to Chinese (90%) and Western (10%) shares.
- **Office proxies (T3).** Claude and Copilot office use are seat estimates, negligible in total (<1%).
- **Hardware (T1–T2).** One optimization server per site (portfolio scaled by site count); embodied carbon amortised straight-line over four years.

About Akila Digital Twin & AI platform

Akila is a digital twin and AI platform for the built world – buildings, cities, infrastructure and territories. Currently, Akila is deployed over 15 million square metres of property.

- Integrates HVAC, energy, occupancy, BMS and more data into a single operational layer
- Enables continuous energy and emissions optimization without retrofit projects
- Supports AI agents for regulatory compliance and performance-based incentive programs
- Deployed across critical infrastructure, public sector building portfolios, multinational operating portfolios and for institutional real estate owners.

Akila

akila3d.com
marcom@akila3d.com

New York
25 39th St
New York City 10018 USA

Paris
1-7 Cours Valmy
92800 Puteaux France

Singapore
1 Keong Saik Road,
089109 Singapore

Shanghai
218 South Xiangyang
Road Shanghai, China