AI, Cybersicurezza e Energy: cosa c’è dietro l’angolo

Roberto Filipelli
Azure GTM Director for EMEA (SME&C SMB)
https://www.linkedin.com/in/msroberfil/

Efficiency vs. Effectiveness

Doing things right vs. doing the right things — Peter Drucker



EFFICIENCY

Doing things right

FOCUS
How we work — process & resources

KEY QUESTION
Are we doing it the best way?

MEASURED BY
Time, cost, output per input

EXAMPLE
Answering 50 emails in an hour

VS



EFFECTIVENESS

Doing the right things

FOCUS
What we do — goals & outcomes

KEY QUESTION
Are we doing the right things?

MEASURED BY
Results, impact, goals achieved

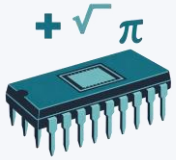
EXAMPLE
Closing the one deal that matters

Be effective first, then efficient — speed in the wrong direction is waste.

[Peter Drucker: Efficiency and Effectiveness Dichotomy - Shortform Books]

Why we added a math co-processor

1980: Intel's 8087 took floating-point math off the main CPU — efficiency through specialization



8086 CPU
Generalist: logic & data

→

8087 FPU
Specialist: maths & trig

01 **No floating-point hardware on the CPU**
Every decimal calculation was emulated in software

02 **Specialized silicon is far faster**
Dedicated adder, shifter and 80-bit registers

03 **Pay only if you need it**
Optional chip for CAD, science and spreadsheets

TIME TO COMPUTE ONE TANGENT

Software on the 8086 alone — **13,000 µs**

With the 8087 co-processor — **90 µs**

144× faster

on tangents — and ~100× on basic addition

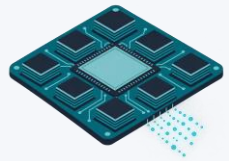
Same logic today: GPUs and AI accelerators (NPUs) offload specialized work for efficiency.

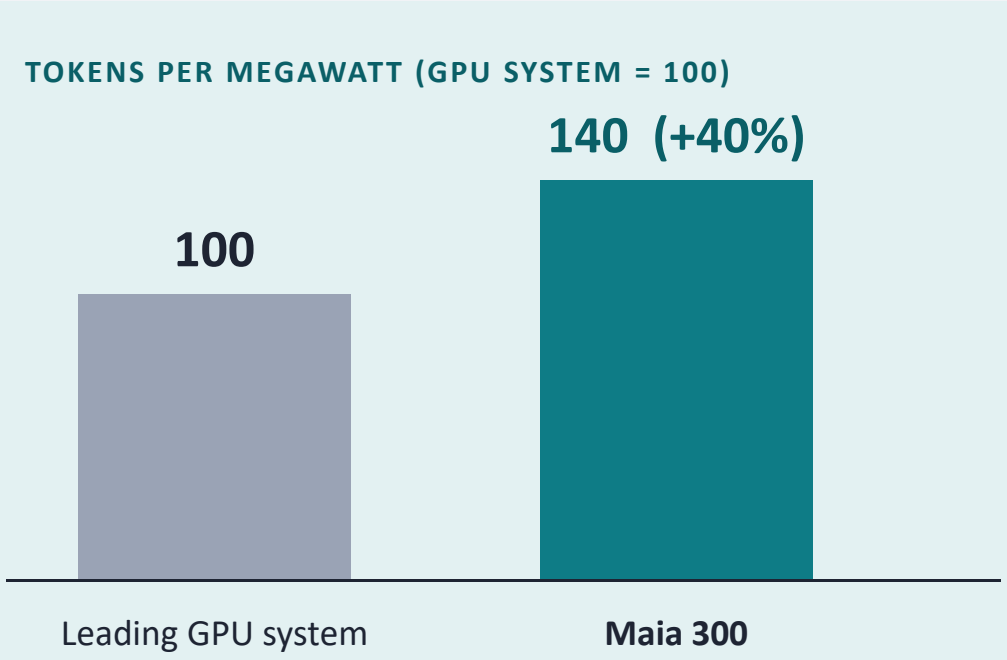
Source: Ken Shirriff's die-level analysis of the Intel 8087 (righto.com), as reported by LavX News, 2026.

[Die analysis of the 8087 math coprocessor's fast bit shifter]

Maia 300: more tokens per megawatt

Microsoft's 3rd-gen inference chip vs. a general-purpose GPU system





Same power budget → 40% more AI answers served

288 GB
Memory-first design
HBM4 at 18 TB/s per chip keeps models close

96
Dense, co-designed racks
accelerators in one 167 kW rack

Up to 7×
Generational leap
rack token throughput vs. Maia 200

Purpose-built for serving tokens — no silicon spent on graphics or training

GPUs stay the versatile generalist for training — Maia is the inference specialist. The 8087 lesson, again.

Source: Microsoft announcement, 12 Oct 2026 (Scott Guthrie). Microsoft's projection; GPU baseline not named. Production from 2027.

[Microsoft to expand Azure Maia system to multi-gigawatt deployments - The Official Microsoft Blog]

Nadella on AI security: lessons learned

From “security above all else” to AI agents as insider risks

NOV 2023

Secure Future Initiative
Launched after nation-state breaches exposed gaps

MAY 2024

“Do security.”
Memo: security above all else, tied to exec pay

JUN 2026 · BUILD

Agents are a new runtime
They need identity, policy, containment, monitoring

OCT 2026

Models as insider risks
Assume compromise; keep authority outside the model

LESSONS LEARNED

Security beats speed
When trade-offs arise, security wins over new features

Agents are insiders
Give every agent an identity and least privilege

Keep control external
Never rely on a model to police itself

Evidence and a brake
Tamper-proof logs and a human stop button

“We simply can't outsource responsibility for what intelligence does on our behalf.” — Satya Nadella

Sources: Nadella memo (May 2024) and letter (Oct 2026); Build 2026 keynote; Microsoft SFI site; reported by TechCrunch, Business Standard.

Models as Insider Risks in the Super Intelligence Era | sn scratchpad

The risk is also in the device

Beyond AI: the Cyber Resilience Act makes every connected product a security duty



THE CRA CLOCK

AUG 2025
RED art. 3.3: cybersecurity mandatory for radio devices

11 SEP 2026 · IN FORCE
Report exploited flaws: 24h warning, 72h notice

11 DEC 2027
CRA fully applies secure by design for CE marking

€15M or 2.5%
of global turnover — and products already sold are in scope

CASE: ERMIT SECURE IoT MODULE + EXEIN

ermit

Security in the hardware, governed in Azure

Secure by design	Keys born and kept in an NXP SE051 secure element
Updates for life	OTA updates with anti-rollback; no silently expired certificates
Detect, then report	Exein Runtime watches device behaviour in real time
Prove compliance	CRA, RED, ETSI EN 303 645 built in; audit trail from day one

Security isn't bolted on at the end — it's designed in from the first chip.

Sources: European Commission, CRA reporting obligations; Ermit IoT Secure Module; Physical AI Security | Exein

Achieve CRA & RED Compliance for Physical AI with Exein & Ermit

Why the EU introduced milk quotas

“Quote latte”, 1984–2015: matching supply to real demand



THE CAUSE
Guaranteed prices
Paid regardless of demand, above world prices

➡

THE EFFECT
Butter mountains
Supply outgrew demand; the EU bought the surplus

➡

THE FIX · 1984
Quota + superlevy
A cap per producer, a levy for exceeding it

➡

THE EXIT · 2015
Market-led supply
Quotas ended 31 March 2015; the market leads

EU-10 COW'S MILK OUTPUT, MILLION TONNES



+21% in ten years, far faster than demand

>1M t
butter and milk-powder stocks at peak, each

31 yrs
of quotas — production stabilized, mountains fell

Farms were efficient at producing milk — but not effective: output nobody needed.

Sources: European Commission, “End of milk quotas” FAQ (26 Mar 2015); Eurostat, “Milk and milk products – 30 years of quotas”.

Briefing European Parliamentary Research Service

Curtailment: today's milk quota

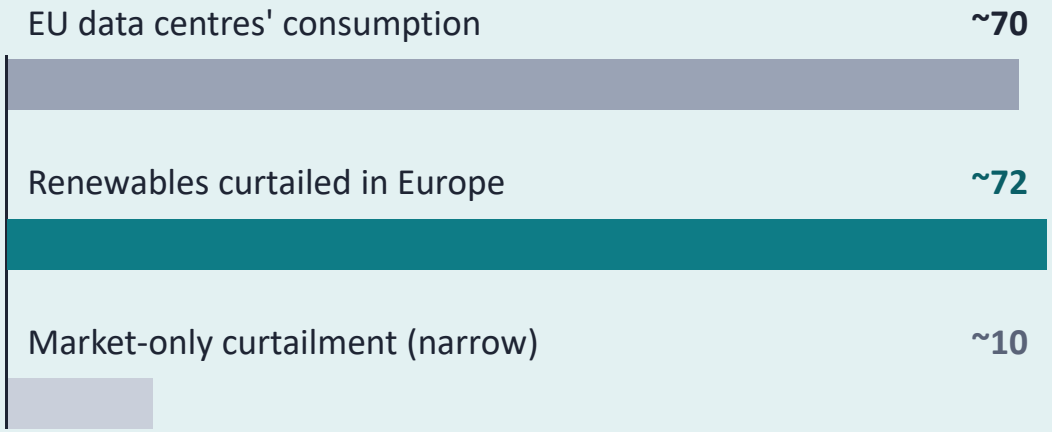
Europe throws away about as much clean power as all its data centres use



THE PARALLEL

	Milk, 1984	Energy, today
Surplus	Milk nobody bought	Wind, solar and hydro the grid can't absorb
Response	Cap output with quotas	Switch plants off (curtail)
Lesson	Supply without demand is waste	**Bring flexible demand to the surplus**

ELECTRICITY PER YEAR, TWh (2024)



≈ 1 : 1

wasted clean power vs. all EU data-centre demand

Don't cap the surplus — use it!

Sources: EU data centres ~70 TWh (2024, estimate cited by the European Commission); curtailment ~72 TWh (Aurora Energy Research via Clean Energy Wire, broad scope).

[Europe’s three-fold solar and wind growth sparks urgent grid investment, Aurora finds]

Zero-kilometre energy for data centres



Bring the compute to the surplus, not the surplus to the compute

THE ZERO-KILOMETRE MODEL

Renewable plant → **Data centre on site** → **Battery buffer** → **Heat back to the area**

Absorbs surplus at the source
Uses power the congested grid would otherwise curtail

Medium-sized and distributed
Smaller sites near users and plants, on reclaimed land

Stable, predictable costs
Self-production cuts exposure to energy price swings

CASE STUDY

Core Stack

Core Stack · Lazzari Group

13
medium-sized data centres in Italy

>€1bn
in 4 years, with Green Arrow Capital

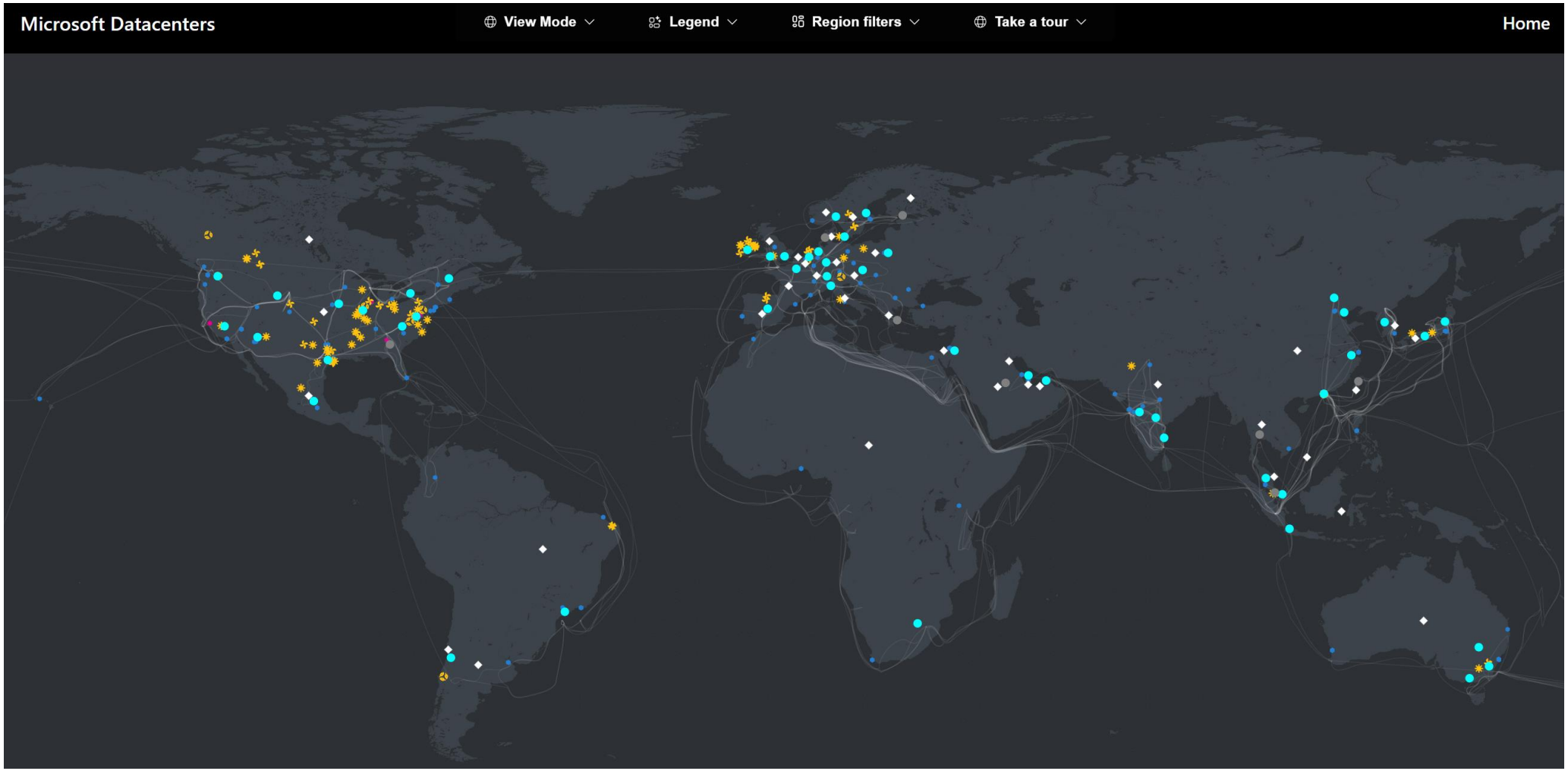
Q1 2028
first sites live: Cremona and Verona

Every digital kilowatt matched by a renewable kilowatt — produced on site.

Sources: Core Stack and Lazzari Group websites; Il Sole 24 Ore (6 Oct 2026); 01net (8 Oct 2026); ESG News (14 Oct 2026).

[Homepage - Corestack]

Microsoft Datacenters



https://datacenters.microsoft.com/globe/explore

One Azure, from one rack to a region

Microsoft Sovereign Cloud: a continuum of control, chosen workload by workload



ANY SIZE, ANY SITE

Edge site
A few servers at a factory or renewable plant

Regional data centre
Medium-sized sites, like Core Stack's network

Shared sovereign site
Azure Local multitenancy: one site, many tenants

Microsoft EU regions
Sovereign Public Cloud with extra EU controls

More of the app runs locally

Data, code, AI and productivity on Azure Local

Foundry Local
GA at Ignite · local AI models

GitHub Enterprise Local
GA at Ignite · code stays local

Data services
SQL GA · more in preview

Control you can see and prove
Sovereign Control Panel, Data Guardian, your own HSM keys, confidential VMs on site

Resilient and partner-led
Switch connected to disconnected (preview); Bleu, Delos and local partners

Any data centre can run Azure — with the right level of control for each workload.

Source: Official Microsoft Blog, “Advancing digital sovereignty with greater control and choice”, D. Phillips (20 Oct 2026).

Advancing digital sovereignty with greater control and choice - The Official Microsoft Blog

Grazie!

Roberto Filipelli
Azure GTM Director for EMEA (SME&C SMB)
https://www.linkedin.com/in/msroberfil/