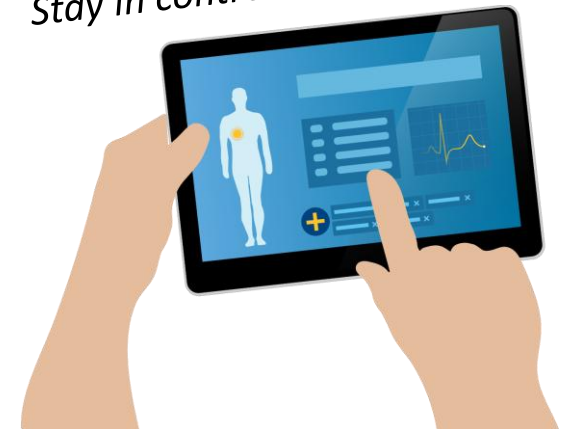# How to feed Hungry AI
with high-quality Data, on Demand, at Scale

„Vaccinate Data against Abuse"
Stay in control in your Ecosystems

AI solutions' growth driving growth elsewhere

The 18 potential arenas of tomorrow could generate $29 trillion to $48 trillion in revenues and $2 trillion to $6 trillion in profits.
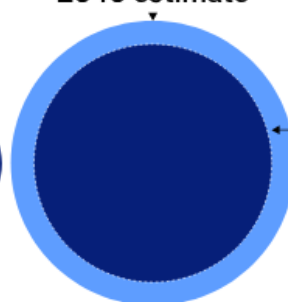
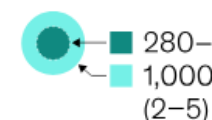18 potential arenas of tomorrow, by 2040 revenue estimate, $ billion
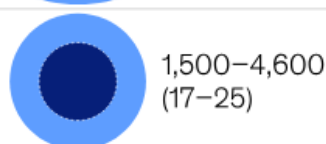
Arena	Revenue, 2022	2040 estimate (CAGR, 2022–40, %)	Profit, 2040 estimate,[1] $ billion (profit margin, %)
E-commerce	4,000	14,000– / 20,000 (7–9)	280– / 1,000 (2–5)
AI software and services	85	1,500–4,600 (17–25)	230–920 (15–20)
Cloud services	220	1,600–3,400 (12–17)	160–510 (10–15)
Electric vehicles	450	2,500–3,200 (10–12)	100–320 (4–10)
Digital advertisements	520	2,100–2,900 (8–10)	320–580 (15–20)
Semiconductors	630	1,700–2,400 (6–8)	340–600 (20–25)
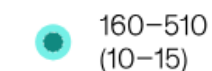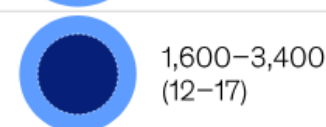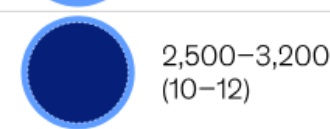
Figure: https://www.mckinsey.com/featured-insights/sustainable-inclusive-growth/charts/dynamic-industries-to-watch/

Data Feeding in the Good Old Days (before AI)
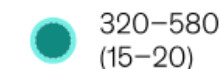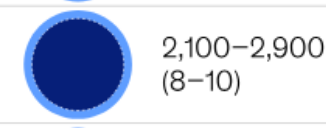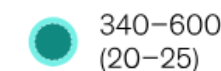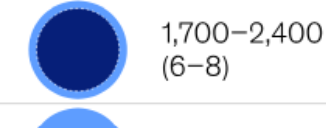
Applications always needed good data

- Accurate, structured, curated

- Centrally organized data or federated data

- Right-time data, from…
    - Post-transaction analytics to
    - Decision support to
    - Real-time decision making

AI's hunger games: Data Markets to feed insatiable LLMs

Figure: https://www.linkedin.com/pulse/ais-hunger-games-lucrative-data-market-exploding-feed-insatiable-9a4ie/

AI's Hungry Brain: A Data-Driven Dilemma

Why is AI so data-hungry?

- **Training Models:** AI models learn from massive datasets. The more data they consume, the more accurate and sophisticated they become.

- **Real-time Processing:** AI applications often require real-time processing of large amounts of data, such as in self-driving cars or facial recognition systems.

- **Constant Updates:** AI models need to be continuously updated and retrained to keep up with new data and evolving trends.

Figure: https://www.linkedin.com/pulse/ais-hungry-brain-data-driven-dilemma-meghan-mcparland-tgvqe/

Generative AI Is Data-Hungry: Is the All-You-Can-Eat Buffet Over?

Finding good quality data on which to train generative AI models has been difficult. And it's going to get harder, too, as rights holders to vast quantities of internet data begin to close ranks in recognition of just how valuable a resource it all is, observes Stephen Marcinuk of Intelligent Relations.

Figure: https://www.spiceworks.com/tech/artificial-intelligence/guest-article/generative-ai-is-data-hungry/

Not enough real-word data for AI > go Synthetic !?

By 2030, Synthetic Data Will Completely Overshadow Real Data in AI Models

Data Used for AI

- Future AI
- Today's AI

Synthetic Data
- Artificially Generated Data
- Generated From Simple Rules, Statistical Modelling, Simulation and Other Techniques

Real Data
- Obtained From Direct Measurements
- Constrained by Cost, Logistics, Privacy Reasons

2020 — Time — 2030

Source: Gartner
750175_C

Gartner

NVIDIA DRIVE Sim
Synthetic Data Generation
Powered by Omniverse Replicator

Dynamic Weather
24.4

Figure: Gartner, NVIDIA   https://blogs.nvidia.com/blog/what-is-synthetic-data/

AI GIGA FACTORIES (AI 1.0 = everything "Huge")

AI Gigafactories are large-scale facilities dedicated to the development and training of next-generation AI models containing trillions of parameters. 3 Jun 2025

The CPU handles all the tasks required for all software on the server to run correctly. A GPU, on the other hand, supports the CPU to perform concurrent calculations.

- ➤ **Huge** Data Centers
- ➤ **Huge** Processing Farms (GPUs from NVIDIA etc.)
- ➤ **Huge** Power Consumption
- ➤ **Huge** Costs

Data Center energy demand to grow by 3.5x

Both AI and non-AI workloads will be key drivers of global data center capacity demand growth through 2030.

Estimated global data center capacity demand, 'continued momentum' scenario, gigawatts

	2025	2026	2027	2028	2029	2030	2025–30 total
Incremental AI capacity added per year, gigawatts	13	18	21	19	22	31	124
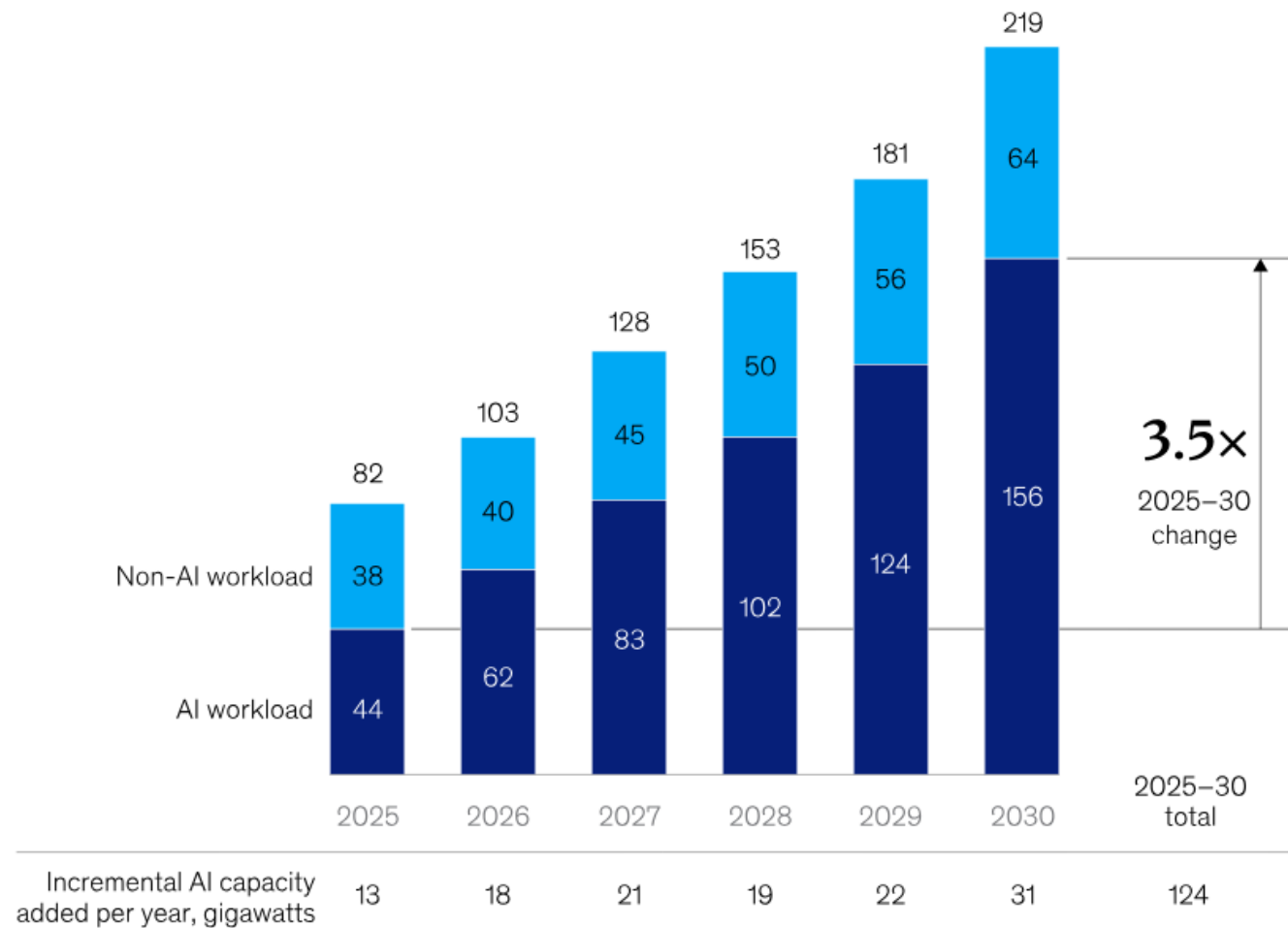
Edge AI (AI 2.0 = do it where the action is)

- Low latency
- Lower bandwidth demand
- Higher scalability
- Better data privacy
- Higher security
- Lower costs

Edge AI will grow even faster than "GIGA AI", and will substitute parts of it.

10 minutes on Edge AI: https://www.youtube.com/watch?v=QXuvUW4B7jU

PPP to ensure Sovereignty in AI-ready Data Ecosystems

GOVERNANCE LAYER

DATA CLEARING AND INTERCONNECTION SERVICES

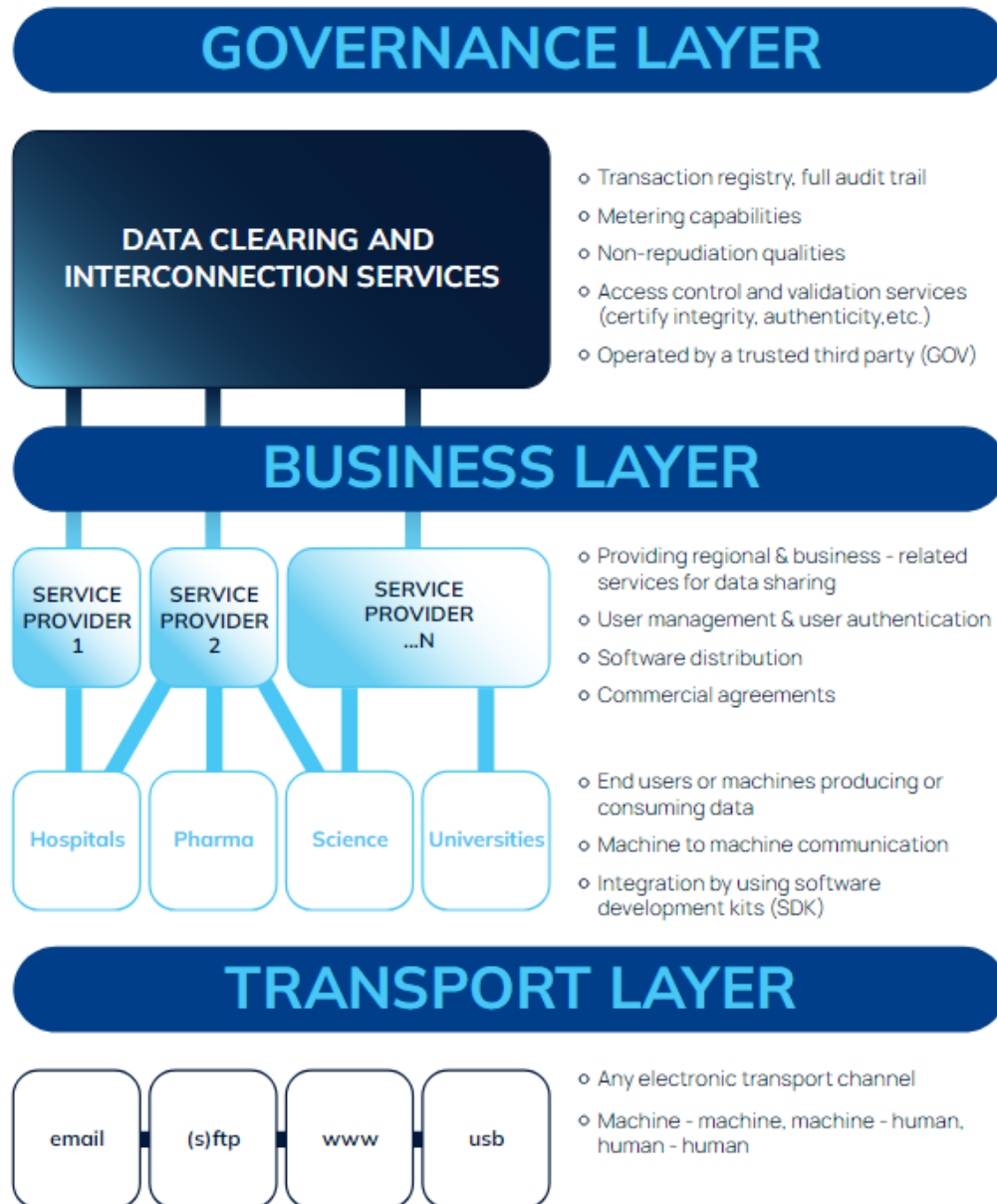- Transaction registry, full audit trail
- Metering capabilities
- Non-repudiation qualities
- Access control and validation services (certify integrity, authenticity, etc.)
- Operated by a trusted third party (GOV)

BUSINESS LAYER

SERVICE PROVIDER 1 | SERVICE PROVIDER 2 | SERVICE PROVIDER ...N

- Providing regional & business - related services for data sharing
- User management & user authentication
- Software distribution
- Commercial agreements

Hospitals | Pharma | Science | Universities

- End users or machines producing or consuming data
- Machine to machine communication
- Integration by using software development kits (SDK)

TRANSPORT LAYER

email | (s)ftp | www | usb

- Any electronic transport channel
- Machine - machine, machine - human, human - human

Public-Private-Partnerships (PPP) to balance governance and business needs.

AIQURIS & DATA VACCINATOR

UNLOCKING THE POWER OF AI IN HEALTHCARE: Managing Data, Risk, and Compliance

[Access whitepaper](#)

Thank you.

slides at https://datavaccinator.com/news.html

Let's stay in touch.

Kurt Kammerer, CEO kurt@datavaccinator.com
https://www.linkedin.com/in/kammererkurt/

DataVaccinator SARL
7, route d'Esch
L-1470 Luxembourg
Luxembourg

DataVaccinator Limited
VD, 1st floor
Masdar City, Abu Dhabi
United Arab Emirates

DataVaccinator LLC
Riyadh, KSA

DataVaccinator

DataVaccinator (DV) powers Data Ecosystems. By "vaccinating data against abuse", DV solutions protect sensitive data, enable secure collaboration and ensure Data Sovereignty in AI environments. DV products rely on its unique technologies (60 patents).

As partner in major initiatives, DV provides guidance on cyber security, governance, compliance, setup and operation of AI-ready data networks.

Background and Expertise:
- Deep Technical Know-how in secure Collaboration and Ecosystems
- Intellectual Property Rights (IPR): Assessment and Management
- Data Ecosystems: Healthcare in EU/MENA, Secure Supply Chains (WEF), Belt&Road (Asia), Logistics (Air, Sea, Road), Cyber Security (ASEAN)
- Belfer Security Center (Harvard, JFK), WEF Security Group