LEADERSHIP. ACTION. RESULTS.

# 2035 - THE EDGE AI REVOLUTION: WHEN 80% OF AI MIGRATES TO THE EDGE

## From centralized AI to distributed intelligence: Technologies, Markets and Strategic Issues 2020-2035

Target Sectors

Industry 4.0, Healthcare, Smart Cities, Automotive,
IT & Telecom, Retail, Agriculture, Energy & Utilities

Author

**Yannick Gablin**

Director – MEA / Digital Services

ALVAREZ & MARSAL

# SECTION 1 – DETAILED TIMELINE

DETAILED TIMELINE

# 2020 – 2022 : The Awakening of the Giants

When AI left the laboratories to conquer the world

Art & Creativity | LLM Models | Code | Science | Open Source

**2020**
### GPT-3
OpenAI unveils a model with 175 billion parameters.. Emerging capabilities at scale.

**175B params | Training: $10M**

**End of 2020**
### AlphaFold 2
DeepMind Solves a 50-year-old challenge. Accelerates drug discovery.

**98% precision | 200M struct.**

**January 2021**
### DALL-E
Inaugurates AI-powered visual creation. Transforms the text-image relationship.

**Arch. Clip+VAE | 2.5M pic/day**

**June 2021**
### CoPilot
First coding assistant for the general public. Increases productivity by 55%.

**1.2M devs | 40% code auto**

**July 2022**
### Bloom
Democratize multilingual AI. A European alternative to US models.

**176B params | 46 languages**

**August 2022**
### Art for All
Midjourney and Stable Diffusion make visual creation accessible to all.

**10M users | Open Source**

**November 2022**
### ChatGPT
The tipping point. 100M users in 2 months. The world discovers the power of AI.

**100M users | Val $20B**

2020 2021 2022

DETAILED TIMELINE

# 2023 – 2025 : The Surge

Generative AI is invading every corner of our lives.

Art & Creativity · LLM Models · Competition · Open Source

### March 2023
## GPT - 4
Establishes a new standard for reasoning. Native multimodality revolutionizes use cases.

**1.76T params**

### July 2023
## Llama 2
Democratizes access to LLMs. Enables open innovation at scale.

**70B params**

### November 2023
## GPT - 4 Turbo
Makes AI economically accessible. Massive context window opens new horizons.

**128k context**

### February 2024
## Sora
Transforms professional video creation. Redefines the boundaries of AI realism.

**1080p 60s**

### March 2024
## Claude 3
Establishes ethics as a competitive advantage. Record window allows for analysis.

**200k context**

### June 2024
## Claude 3.5
Surpasses all benchmarks. Massive enterprise adoption.

**2x times faster**

### 2025
## Gemini 3
Native integration into the Android ecosystem. Extreme multimodality.

**10M context**

DETAILED TIMELINE

# 2026 – 2028 : The Invisible Army

Millions of agents and robots are invading our daily lives

Infrastructure | Corporate | Consumer | Quantic | Robotic | AI Agent

**January 2026**
### Stargate Project
Largest private investment in AI. Redefining computing infrastructure.

**$500B invest. | 5M GPUs**

**Mid 2026**
### Apple Intel.
2 billion devices, largest AI network. On-device/cloud hybridization.

**2B devices | $20B R&D**

**End of 2026**
### IPO Anthropic
Record valuation for ethical AI. Claude Enterprise dominates B2B.

**Val. $40B | 60% B2B Share**

**2026 - 2027**
### Invisible Army
100M autonomous agents are transforming work. 30% of tasks are automated.

**100M agents | 30% Auto.**

**End of 2027**
### 100k Robots
Humanoids deployed in industry. Priced at $30k, democratizing the physical auto.

**Price 30k$ | Humanoids**

**2028**
### AI Chips
TPU v7 and ASIC: 10x performance per watt. End of Nvidia's monopoly on AI.

**10x Perf/Watt | Post-Nvidia**

**2028**
### Quantic AI
1000-qubit hybrid quantum computer. Breaking the computing wall.

**1000 Qubits | Compute Wall**

DETAILED TIMELINE

# 2026-2030: The Great Consolidation of AI Models

From 153 active or experimental AI models and 42 vendors, to an oligopoly of 25 survivors

## Consolidation Phases

### 1 2026: The Great Purge — -60%

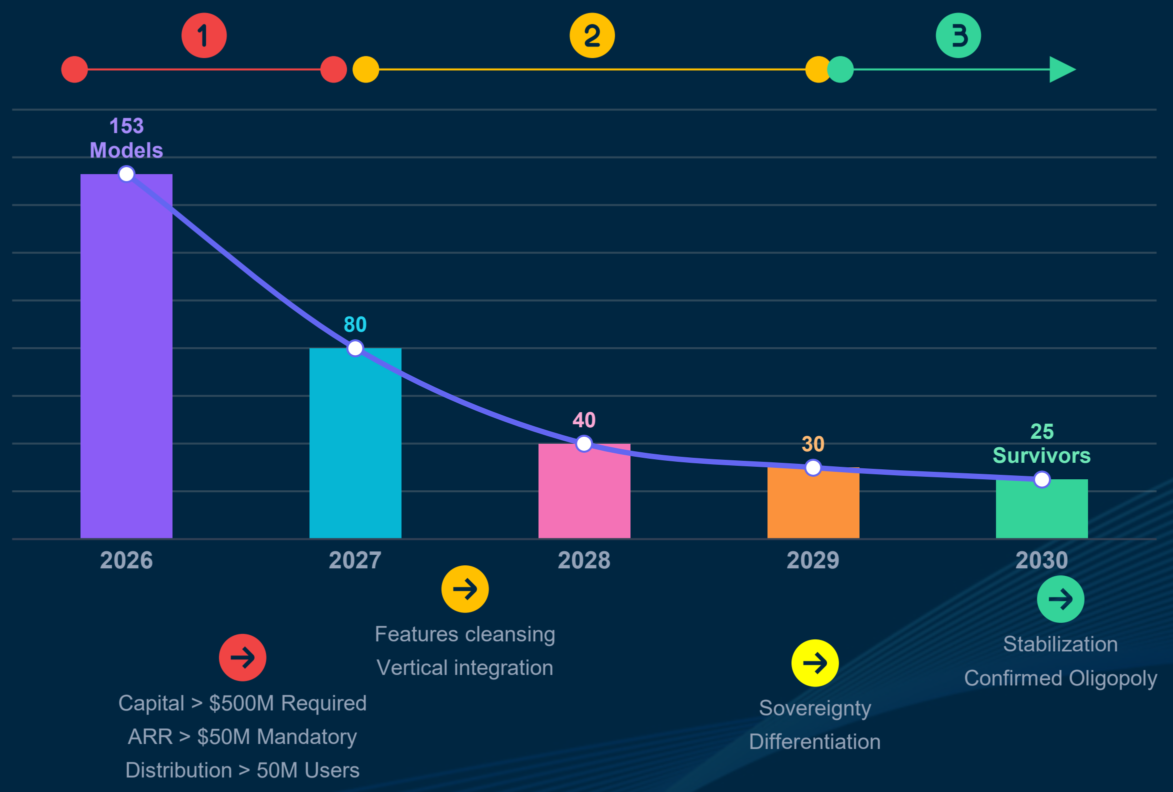
| Rationalization Criteria | Victims (73 Models) |
| --- | --- |
| Capital > $500M Required | "Me-too" without differentiation |
| ARR > $50M Mandatory | Underfunded startups |
| Distribution > 50M Users | Features sold as products |

### 2 2027 – 2028 : The Great Rebundling — -50%

| M&A Mechanisms | Impact (80 to 40 Models) |
| --- | --- |
| 30+ major acquisitions | Microsoft acquires 8 startups |
| 40 "Features" tools absorbed | Google is integrating 6 specialists |
| Vertical integration of platforms | Anthropic IPO or acquisition |

### 3 2029 – 2030 : The Final Oligopoly — STABLE

| Market Structure | 25 Survivors |
| --- | --- |
| Sovereignty: Regional Champions | 5 Global Giants (70% share) |
| Proprietary differentiation | 5 Regional Champions (15%) |
| Locked ecosystem (Network effects) | 15 Niche Specialists (15%) |




## 2030 : Top 10 Survivors

**84%** Eliminated

**Oligopoly** 2030 Structure

**Value** VC New Era

| TIER 1 – Global Leaders | TIER 2 - Challengers | TIER 3 - Specialists |
| --- | --- | --- |
| OpenAI – 35% | Meta - OpenSource | Mistral (Europe) Cohere |
| Google – 25% | xAI – 5% | MidJourney |
| Anthropic – 10% | Ernie - 10% (China) | Perplexity |

ALVAREZ & MARSAL

# The Oligopoly : Who Will Dominate AI in 2030 ?

From 153 to 25 models: portraits of the giants, champions, and specialists who survived the great purge

## TIER 1
### Global Leaders
5 players • 70% market share

**OpenAI** (35% market)
350M users | 2T params | $15B
*Mass distribution + Ecosystem*

**Google** (25% market)
200M users | 3T params | $12B
*Compute infrastructure + integrated Android*

**Anthropic** (10% market)
45M users | 500B params | $3B
*B2B dominance + Ethics*

**xAI** (5% market)
15M users | 6T params | $1.5B
*Musk ecosystem (X/Tesla) + Real-time*

**Meta** (Open Source)
50M devs | 405B params | Free
*Open-source strategy + Development ecosystem*

## TIER 2
### Regional Champions
5 players • 15% market share

**ERNIE / Baidu** (China)
200M users | 2.4T params | $8B
*National Champion + Captive Market*

**Mistral AI** (Europe)
8M users | 123B params | $500M
*EU Champion + GDPR compliance*

**Qwen / Alibaba** (China)
5M users | 235B params | $400M
*Aliyun Cloud Integration + E-commerce*

**Tencent** (China)
150M users | 1.8T params | $2.5B
*Massive WeChat ecosystem*

**Samsung** (Korea)
100M users | Gauss 2 | $1.8B
*Global mobile hardware integration*

## TIER 3
### Niche Specialists
15 players • 15% market share

#### CREATIVES & B2C

**Midjourney** | 20M users | $500M
*Artistic Monopoly + Premium Quality*

**Runway** | 3M users | $150M
*Standard Hollywood + Video Innovation*

**Stability AI** | 15M users | $200M
*Leading open-source visual + ecosystem*

**Suno** | 12M users | $120M
*Generative music market dominance*

**Character.AI** | 20M users | $180M
*Extreme B2C Retention + Personas*

#### RESEARCH, CODE & DATA

**Perplexity** | 25M users | $300M
*AI Search Engine + Quotes*

**GitHub Copilot** | 80M devs | $1.2B
*Mass adoption + Network effects*

**AlphaFold** | 1M scientists | Research
*Health / biotech monopoly + Real impact*

**Hugging Face** | 5M devs | $200M
*Community Hub + AI Infrastructure*

**Databricks** | 1M B2B | $800M
*Enterprise RAG + Data Security*

#### AUDIO, LANGUAGES & B2B

**ElevenLabs** | 10M users | $250M
*B2B audio quality + Voice cloning*

**Cohere** | 5M users | $150M
*Multilingual templates + Business focus*

**DeepL** | 15M users | $180M
*NLP Specialist Leader (Translation)*

**Canva** | 150M users | $2B
*AI Design Integration + B2C Distribution*

**Scale AI** | 2M users | $400M
*RLHF data provider + Shovels/picks*

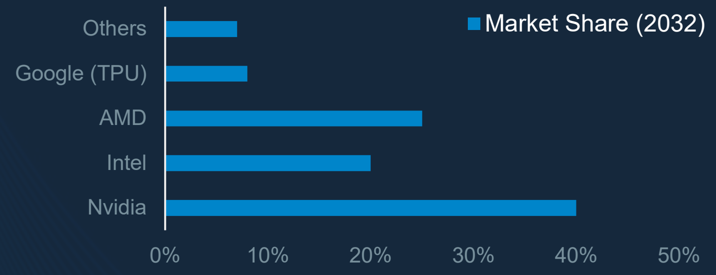
# 2029 – 2032 : The Great Upheaval

AI is reshaping the global economy and our societies

## Silicon War

### Fall of the Titans

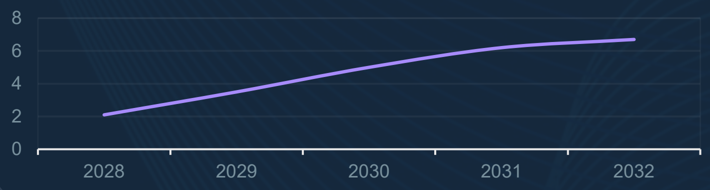
Nvidia loses its monopoly (2025 to 2032 : -80%). Intel is reborn thanks to chip sovereignty (2025 to 2032 : +85%). Hardware becomes the number one geopolitical issue.




## Infrastructure Expenses

Explosion of cumulative costs ($ Trillions)

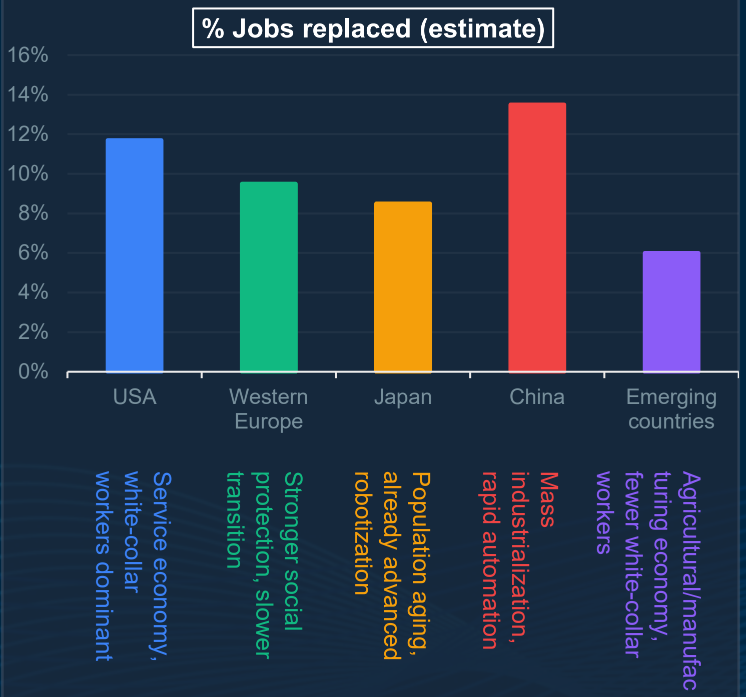
Inc. Data Centres, Chipset, Energy, Network, Training




## Social Emergency

### The Great Obsolescence

Automation is mainly hitting white-collar workers. Universal Basic Income is no longer a theory, but a necessity for political survival.




## Workforce 2.0

### Robots Invasion (100k units end of 2027)

Millions of robots in factories and homes. The cost of labor is collapsing.

**2029 - 500k units** — Massive pilot deployment

**Vendors** : Tesla Optimus (300k), Figure AI (150k), 1X Technologies (50k).
**Sectors** : Automotive, Logistic, Warehouse (e.g., Amazon)
**Unit Price** : $35k

**2030 - 2M units** — Industrialization

**Vendors** : + Boston Dynamics production, chineese startups (BYD, Xiaomi).
**Sectors** : + Agriculture, Hospitality, Construction
**Unit Price** : $30k

**2031 - 8M units** — Public adoption

**Distribution** : Industry (5M), Serviices (2M), Domestic (1M)
**Unit Price** : from $15k (domestic model) to $25k
**Impact** : massive pressure on manufacturing employment

**2032 - 40M units** — Ubiquity

**Worldwide Dist.** : USA (15M), China (18M), Europe (5M), RoW (2M).
**Sectors** : all sectors, inc. Retail & Healthcare
**Unit Price** : from $10k (basic model) to $20k

# 2033 – 2035 : AGI, the Point of No Return

Metrics of the Technological Singularity

## 2035 Economic Impact

Source : PWC Global

Global Growth GDT
**+$357 Trillion**

Annual AI Revenues
**$8.5 Trillion**

## Infrastructure

Source : Epoch AI / Forbes

Compute Capacity
(x1000 vs 2025)
**$10^{27}$ Flops**

Global Electricity Consumption
**8%**

2032 / 2035

AGI

Artificial General Intelligence

## Societal Transformation

Source : McKinsey / OECD

Global workforce transformed / replaced
**40%**

Active Autonomous Agent
**2 Billions**

## Event Horizon

Source : Sutskever 2025

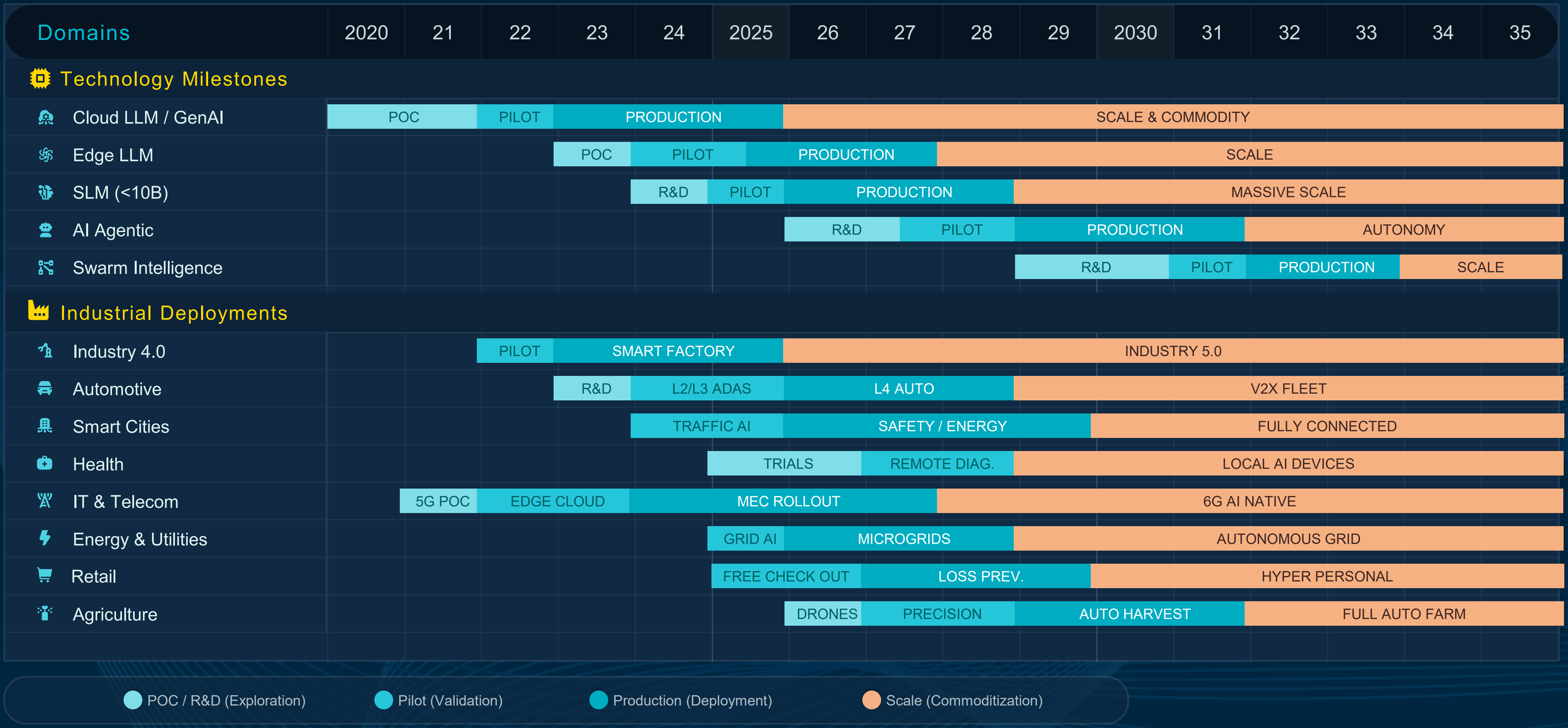
AGI Estimate
(Probability > 50%)
**5 – 20 Years**

Major Innovation Cycle
(2M AI Patents per Year)
**3 Months**

# SECTION 2 – SECTORS IMPACTS

# Strategic roadmap : When each sector shifts towards the Edge

Technology milestones and deployment phases by industry sector



2026 SECTOR ANALYSIS

# 8 transformed industries : Where Edge AI creates the most value

## Accelerated Adoption by Industry (2025-2026) & Value Transformation

### Priority verticals (2026 projections)

**IT & Telecom** — Leader 21.1%

Mass adoption of Edge AI for 5G/6G network optimization and MEC services.

MEC (Multi-access Edge Computing) for ultra-low latency in mission-critical applications.

**Health** — 86% AI Usage

Connected medical devices and real-time diagnostics with patient privacy.

Advanced telemedicine and continuous patient monitoring with certified local AI.

**Industry 4.0** — Factory Floor

Edge AI industrial PCs for quality control and autonomous predictive maintenance.

Autonomous collaborative robotics and real-time energy optimization.

**Automotive** — ADAS/V2X

Critical inference for L3+ autonomous driving and V2X safety.

Intelligent fleet management and predictive diagnostics for connected vehicles.

**Smart Cities** — Intel Mgmt.

Traffic and urban energy optimization via distributed edge sensor networks.

Enhanced public safety through distributed intelligent surveillance.

**Retail** — Auto-Checkout

Frictionless customer experience and loss prevention through computer vision.

Customer behavioral analysis and dynamic inventory management by Edge AI.

**Energy & Utilities** — Smart Grid

Smart microgrids and anomaly detection for critical energy infrastructure.

Real-time demand prediction and network balancing.

**Agriculture** — AgriTech

Edge sensors for precision agriculture and optimized resource management.

Crop monitoring by AI drones and autonomous smart irrigation.

### Financial Value Levers

Projected Economic Impact (2025-2035)

**$551 Md**
Global AI Market Value 2035 (inc. Edge AI, Cloud AI, Agentic AI & Swarm Intelligence)

**$357 Md**
Edge AI 2035

**27.7%**
Average Edge AI CAGR (2025-2035)

**Acceleration 2026**

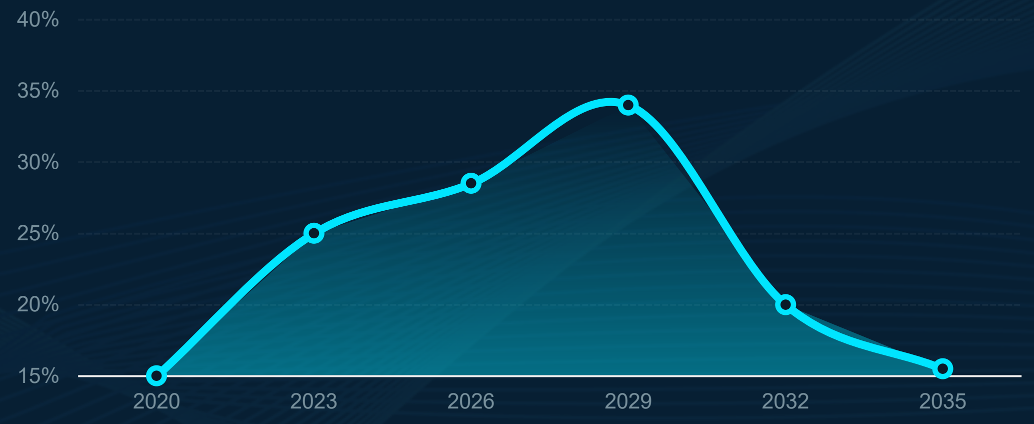
Peak growth for Edge AI at 34% in 2029, in line with the annual chart.

**Industrial ROI**

Major productivity gains via Edge AI factory floor PCs.

Adoption: 86% AI Healthcare

CAGR Edge AI Evolution (2020-2035) — Post-2025 acceleration



FOCUS PROFITABILITY

# ROI Edge vs Cloud : Why Edge outperforms by 2x per sector

## Cloud AI vs Edge AI: Financial and Operational Performance (2025)

| Average ROI Edge AI (Global) | ROI Medium Cloud AI (Global) | Average Return Time (Edge) | Favorable Edge Sectors |
|---|---|---|---|
| 285% ↑ Vs Cloud | 145% stable | 9.5 Months fast | 7 / 9 dominance |

| Industrial Sector | Best use cases | Advantage | ROI Cloud AI | ROI Edge AI | Return Time |
|---|---|---|---|---|---|
| Industry 4.0 | Real-time quality control, Zero latency | Heavy Edge | 120% | 320% | 6 months |
| Automotive | Embedded ADAS, Local Vehicle Data Processing | Edge | 90% | 280% | 12 Months |
| Smart Cities | Traffic management, Video protection, Smart lighting | Light Edge | 180% | 210% | 18 months |
| Health | Patient Privacy, Medical Devices | Edge | 200% | 250% | 14 months |
| Retail | CRM Personalization, Global Trend Analysis | Cloud | 220% | 190% | 10 Months |
| IT & Telecom | Optimization RAN, MEC (Mobile Edge Compute) | Heavy Edge | 150% | 290% | 8 months |
| Energy & Utilities | Smart Microgrids, Grid Anomaly Detection | Heavy Edge | 160% | 270% | 9 months |
| Agriculture | Precision Agriculture, Autonomous IoT Sensors | Edge | 110% | 240% | 11 Months |

### Latency Factor

Real-time constrained sectors (<10ms) see a 2.5x increase in Edge ROI thanks to the elimination of network delays.

### Egress Savings

Local processing reduces cloud bandwidth costs, a decisive factor for video (Smart Cities, Industry), by 90%.

### Compliance & Privacy

For Healthcare and Automotive, the Edge is the only legally viable option, making the theoretical Cloud ROI unattainable.

# 3 real-world case studies : How Edge AI generates 220-320% ROI

Financial and operational impact measured by sector (Actuals 2024-2025)

## Manufacturing
Industrie 4.0 & Smart Factory

### Visual Quality Control — ROI 320%
Edge AI system detecting micro-defects in real-time on high-speed production line.

| $2.4M | 6 months | -85% |
|---|---|---|
| Savings/year | Back Inv. | Defects |

### Predictive Maintenance — ROI 280%
Edge vibration sensors analyze machine wear to prevent critical failures.

| $1.8M | 8 months | -70% |
|---|---|---|
| Savings/year | Back Inv. | Downtime |

### Production Optimization — ROI 220%
Automatic adjustment of production parameters via local AI to maximize yield.

| +45% | 10 months | Low |
|---|---|---|
| Productivity | Back Inv. | Latency |

## Healthcare
Connected Care & Precision

### Local AI Diagnostics — ROI 250%
Medical imaging analysis on edge device without cloud transfer for complete privacy.

| $3.2M | 14 months | +92% |
|---|---|---|
| Savings/year | Back Inv. | Accuracy |

### Continuous monitoring — ROI 210%
Patient monitoring via Edge wearables with real-time detection of vital abnormalities.

| $1.5M | 12 months | +80% |
|---|---|---|
| Savings/year | Back Inv. | Alerts |

### Edge Telemedicine — ROI 190%
Standalone health kiosks with AI pre-diagnosis for rural or remote areas.

| $900K | 16 months | +300% |
|---|---|---|
| Savings/year | Back Inv. | Consults |

## Retail
Customer Experience & Operations

### Standalone checkout — ROI 220%
Checkout-free stores using sensor fusion and Edge computer vision.

| $1.2M | 10 months | -60% |
|---|---|---|
| Savings/year | Back Inv. | Losses |

### Behavior Analysis — ROI 190%
Understanding the customer journey in stores for layout and marketing optimization.

| +$2.8M | 11 months | +35% |
|---|---|---|
| Revenue/year | Back Inv. | Conversion |

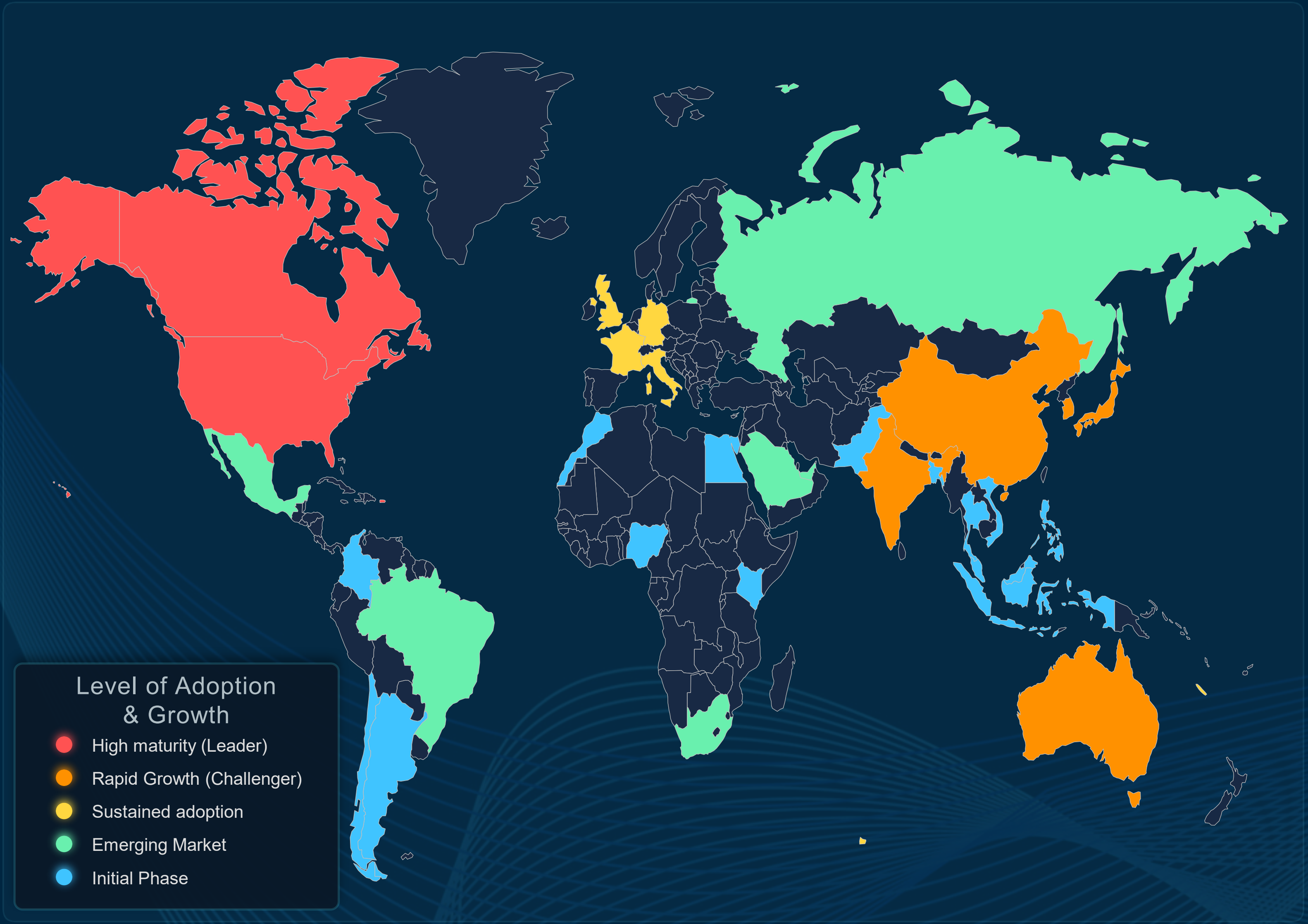
### Stock Management — ROI 180%
Real-time shelf tracking by Edge cameras for dynamic replenishment.

| $800K | 13 months | -75% |
|---|---|---|
| Savings/year | Back Inv. | Breakouts |

# SECTION 3 – ADOPTION

# 35 countries, 5 paces : The Edge AI global adoption map

## Regional dynamics: Market maturity and growth poles (2025)




### North America — 36% Share

CAGR (25-30): 22.5%
Maturity: Advanced

**Drivers:** Hyperscalers, Autonomous Vehicles, Defense, Smart Retail.

### Asia Pacific — 32% Share

CAGR (25-30): 28.4%
Status: Top Growth

**Drivers:** Smart Cities (China/Korea), Industrial Robotics, Massive 5G.

### Europe — 20% Share

CAGR (25-30): 24.1%
Focus: Regulation

**Drivers:** Industry 4.0 (DACH), Healthcare, GDPR/AI Act Compliance.

### Latin America — 5% Share

CAGR (25-30): 18.5%
Key Sector: AgriTech

**Drivers:** Precision Agriculture, Mining Monitoring.

### MEA — 3% Share

CAGR (25-30): 16.2%
Potential: Energy

**Drivers:** Smart Cities (Gulf), oil management, infrastructure.

WORLDWIDE ADOPTION

# 5G Maturity ≥4 + Ecosystem ≥4 = Guaranteed Adoption: The Regional Formula

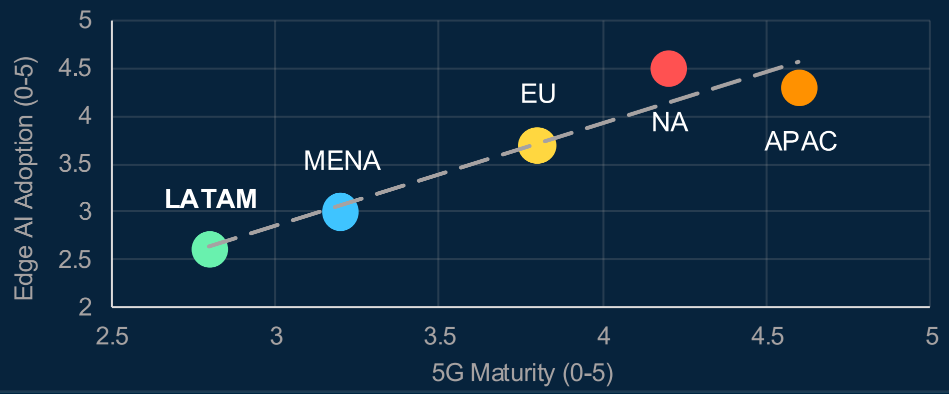
6 Decisive Factors – Why APAC and North America Are 3-5 Years Ahead of Europe

## Matrice Comparative des Drivers (Échelle 1-5)

| REGION | Regulatory | Infra & 5G | R&D / Innov. | Ecosystem | Talents | Competition | Profile (Radar) |
|---|---|---|---|---|---|---|---|
| Am. du Nord | 4 | 4 | 5 | 5 | 5 | 4 |  |
| Asie-Pacifique | 3 | 5 | 5 | 5 | 3 | 4 |  |
| Europe | 5 | 4 | 4 | 3 | 4 | 3 |  |
| Moyen-Orient | 3 | 4 | 3 | 2 | 2 | 3 |  |
| Am. Latine | 2 | 3 | 2 | 2 | 2 | 3 |  |

Leader · Advanced · Average · Low · Very Low

## Correlation : 5G Maturity vs Edge AI Adoption




## Strategic Insights

- Adoption accelerates significantly when 5G Maturity (≥4), Ecosystem (≥4) and Talent (≥4) are combined (NA, APAC).
- In Europe, strong public policies and regulations (GDPR, AI Act) partially compensate for the shortcomings of the startup ecosystem/

## Scale-out Adoption Timeline

**2025 - 2027**

**Leaders**
United States, China, South Korea

Driven by Hyperscalers, massive R&D investments and an already mature 5G infrastructure.

**2027 - 2029**

**Fast Followers**
Europe, Japan, Singapore

Structured adoption through regulation, focused on Industry 4.0 and autonomous vehicles.

**2029 - 2032**

**Strategic Followers**
Middle East, Latin America

Adoption driven by sovereign wealth funds (MENA) and specific industrial needs (Agri/Mining LATAM).

# Indus. 40% APAC, Ret. 30% US, Energy 35% MENA: 5 Regions, 5 Winning Bets

Radical sector specialization + 70% public funding in the MENA region: Where to invest based on your sector

## North America

- Smart Retail 30%
- Health & Telemedicine. 25%
- Smart Cities 20%
- Manufacturing 15%
- Others 10%

**Drivers** : Strong private equity, Hyperscaler ecosystem.
**Focus** : Auton. retail (Amazon), Embedded edge (Auto).

## Europe

- Industry 4.0 35%
- Automotive 25%
- Energy / Utilities 20%
- Smart Cities 12%
- Healthcare 8%

**Drivers** : Sovereignty (GDPR, AI Act), green transition.
**Focus**: Industry 4.0 (DACH), local power grids.

## Asia - Pacific

- Electronics / Manufacturing 40%
- Smart City & 5G 25%
- Telecom Edge 15%
- E-commerce 12%
- Healthcare 8%

**Drivers** : Absolute 5G SA maturity, government support.
**Focus** : Automation (Ports, Logistics), Edge Telecom.

## Middle East

- Energy / Utilities 35%
- Smart Cities 30%
- Transport / Logistic 20%
- Security 15%

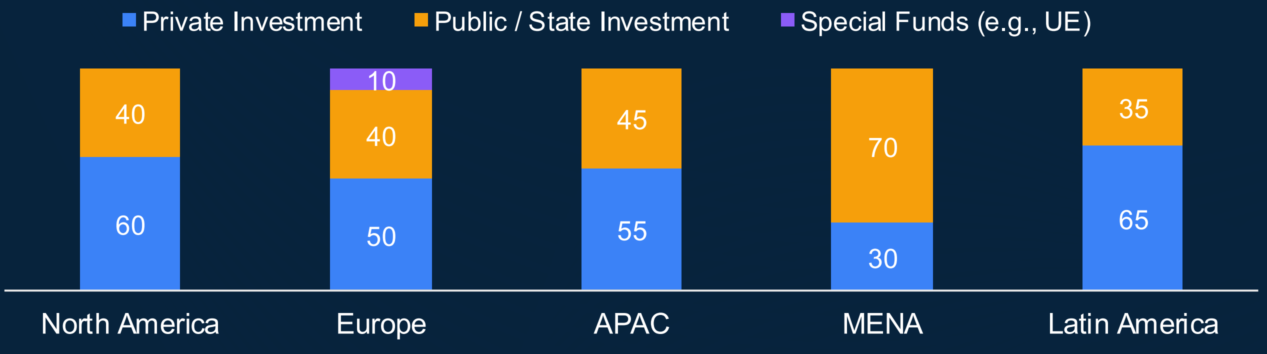
**Drivers** : Post-oil plans (Vision 2030), sovereign wealth funds.
**Focus** : Smart Cities (NEOM), infrastructure monitoring.

## Latin America

- AgriTech 30%
- Mines and Resources 25%
- Retail & Banking 20%
- Smart Cities 15%
- Healthcare 10%

**Drivers** : Primary efficiency, pragmatic investment.
**Focus** : Agriculture (drones), mining security (Brazil/Chile).

## Investment Mix: Public vs Private (2025)




## Strategic Insights

**Strong Sector Specialization**
Industry dominates in **EU** / **APAC** (35-40%), Energy structures MENA (35%), while Retail and Healthcare drive **North America** (55% combined).

**The Driver of Public Investment**
Public money is the main trigger in emerging regions (**MENA 70%**) and catalyzes **APAC** (45%).

**5G Depth & Ecosystem**
The maturity of the 5G infrastructure dictates the complexity of deployments (e.g., massive Edge Telecom cases in **APAC**, absent elsewhere).

# 6 major obstacles to Edge AI adoption (and how to overcome them)

Analysis of adoption challenges and recommended mitigation strategies

Mitigation Roadmap: Recommended phased approach over 18-24 months

## High CAPEX upfront costs

Impact: 40-60% Delays

**Root Causes**

Expensive specialized hardware (NPU, TPU), Need for large-scale multi-site deployments, Legacy migration costs and R&D investments

**Recommended Solutions**

- Progressive hybrid approach (Cloud→Edge)
- Models as-a-Service / Hardware Leasing
- Targeted POCs with fast ROI (<12 months)

## Technical Complexity & Integration

Impact: 55% Lead time>6 months

**Root Causes**

Hardware heterogeneity, Fragmented ML Ops pipelines, OT/IT integration difficulty, Complex fleet management

**Recommended Solutions**

- Unified ML Ops Edge Platforms
- Key Technology Partnerships (SI/OEM)
- Continuous training for IT/OT teams

## Security & Distributed Governance

Impact: 68% Main Brake

**Root Causes**

Expanded attack surface, Physical access to devices, Remote patching challenges, Local data confidentiality constraints

**Recommended Solutions**

- Zero Trust Architecture Edge-native
- Hardware-based end-to-end encryption
- Regular automated security audits

## Lifecycle Management (Scale)

Impact: 5x Maintenance (>100 nodes)

**Root Causes**

Manual, non-scalable deployments, Local model drift (data drift), Intermittent connectivity, Risky OTA updates

**Recommended Solutions**

- Centralized Orchestration (Kubernetes Edge)
- Automated & secure OTA updates
- Monitoring proactif & Self-healing IA

## Interoperability & Standards

Impact: +45% Integration Costs

**Root Causes**

Fragmented IoT protocols, Lack of unified standards, Proprietary lock-in, Multiple embedded operating systems

**Recommended Solutions**

- Open standards adoption (ONNX, MLflow)
- API standardisées Edge (Containerization)
- Industrial consortiums (Edge Native)

## Shortage of Specialized Skills

Impact: Gap 2.3M Pros (2027)

**Root Causes**

Dual expertise required (Data Science + Embedded systems), Rapid technology evolution, Global AI talent war

**Recommended Solutions**

- Internal requalification programs
- Academic & R&D partnerships
- Low-Code/No-Code Edge AI Tools

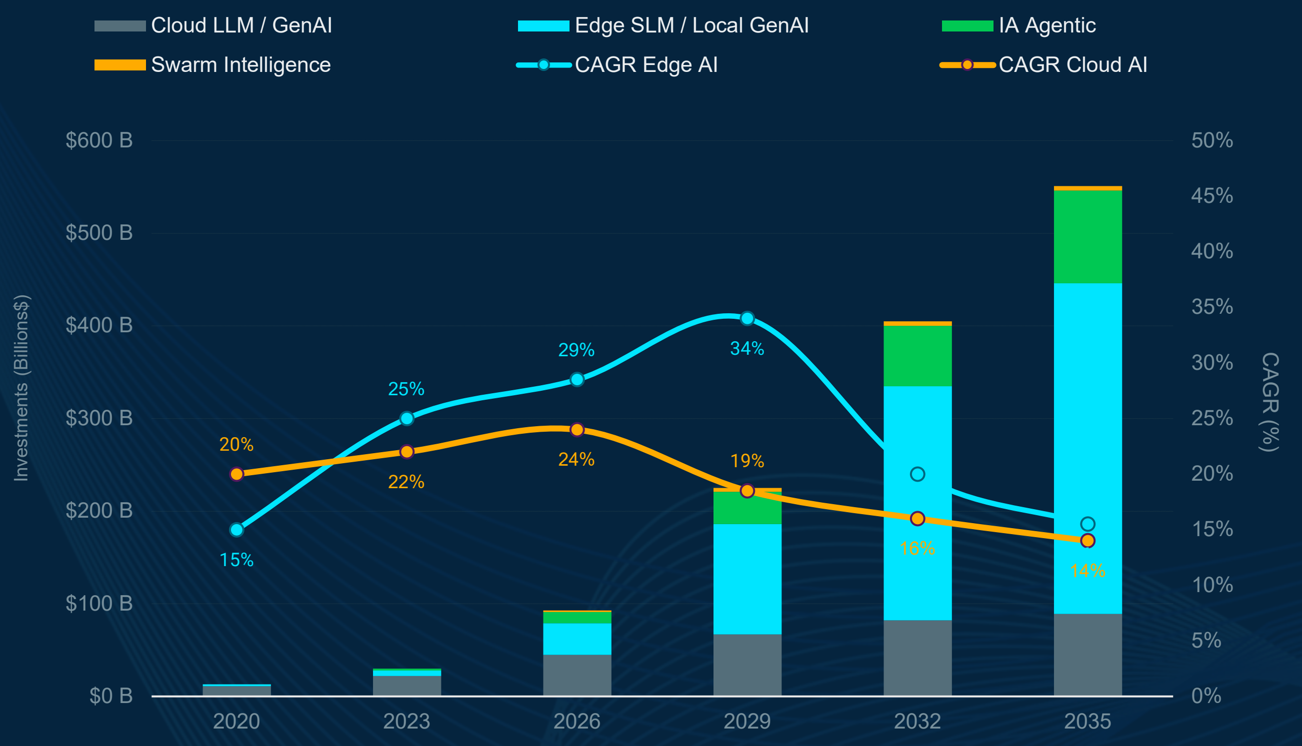
# SECTION 4 –
# FINANCE & M&A

# Edge AI : From $24 Billion to $357 Billion in 10 years (CAGR 27.7%)

## Latest 2025-2026 Projections: Massive Edge & Cloud Acceleration

Edge AI CAGR (2025-2035)
**27.7%**
2035 target: $240 Mds

Cloud AI CAGR (2025-2035)
**18.2%**
Sustained but moderate growth




### 2026 projections confirmed

Analyses confirm a peak CAGR Edge of 34% in 2029 before stabilizing at 24% in 2035 (validated sources: BCC Research, Roots Analysis, Precedence Research). This accelerated growth validates the $357 billion target.

Edge CAGR approx. 28% (peak approx. 34%)

### Edge Hardware on Boom

The Edge AI hardware market alone is expected to reach ~$59 billion by 2030, supported by the explosion of embedded SLMs and NPUs.

### Cloud AI Massive Growth

Despite the rise of Edge AI, Cloud AI aims for stable growth, remaining the backbone for training and massive foundational models.

CAGR means validated by consensus from 9 external sources (Roots Analysis, BCC Research, Precedence Research).
Note: AI Agentic: explosive growth confirmed (CAGR 40-45%, sources: Fortune Business Insights, Acumen Research, MarketsandMarkets). Swarm Intelligence: niche segment (<2% of the market).

M&A INTELLIGENCE

# $265 Billion at stake : The 10 biggest AI M&A deals

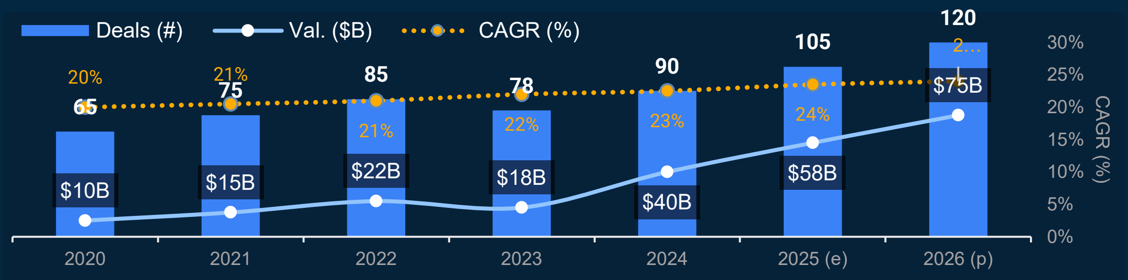
Strategic Divergence: Cloud Consolidation vs. Edge Ecosystem Explosion

## Cloud AI (Centralized)

Deals '26: 120

LLM Training | SaaS AI | Foundation Models

Val: $75B




Multiple EBITDA: 15x (2020) → 28x (2026)

Massive consolidation around hyperscalers. Record EBITDA multiple of 28x in 2026 driven by the race for sovereign LLMs.

### AI Cloud Buyers

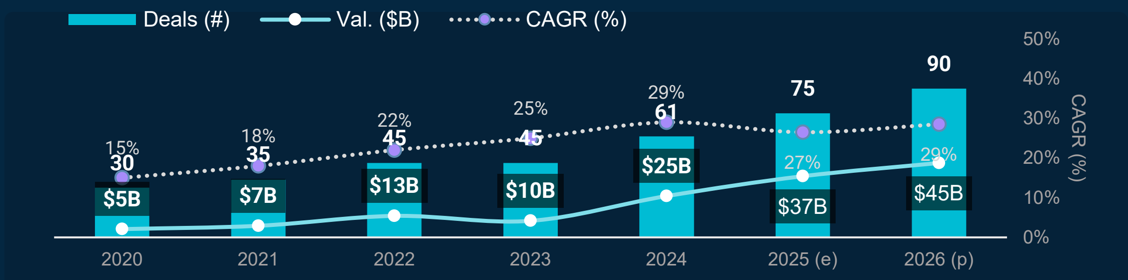
- GOOGLE
- MICROSOFT
- AMAZON
- SALESFORCE
- ORACLE
- IBM

## Edge AI (Distributed)

Deals '26: 90

Hybrid | Native | Pervasive | On-Device

Val: $45B




Multiple EBITDA: 10x (2020) → 24.5x (2026)

Acceleration of the Pervasive AI segment. Deal flow growth +200% (2020-2026) with focus on specialized hardware.

### Edge AI Buyers

- NVIDIA
- INTEL
- AMD
- QUALCOMM
- CISCO
- DELL

## Top 10 Strategic M&A AI Deals (2020-2026)

Cumulative Valuation: $>265B

| # | Deal | Value | Type |
|---|---|---|---|
| 01 | Broadcom VMware | $69B | CLOUD |
| 02 | Microsoft Activision | $68.7B | CLOUD |
| 03 | IBM Red Hat | $34B | CLOUD |
| 04 | Google Wiz | $32B | CLOUD |
| 05 | Cisco Splunk | $28B | CLOUD |
| 06 | HPE Juniper | $13.4B | EDGE |
| 07 | IBM Confluent | $11B | EDGE |
| 08 | NVIDIA Intel | $5B | EDGE |
| 09 | AMD ZT Systems | $4.9B | EDGE |
| 10 | AMD Multi-Acq. | $2B | EDGE |

ALVAREZ & MARSAL

# 2031 : The year Edge AI surpasses Cloud in M&A deals

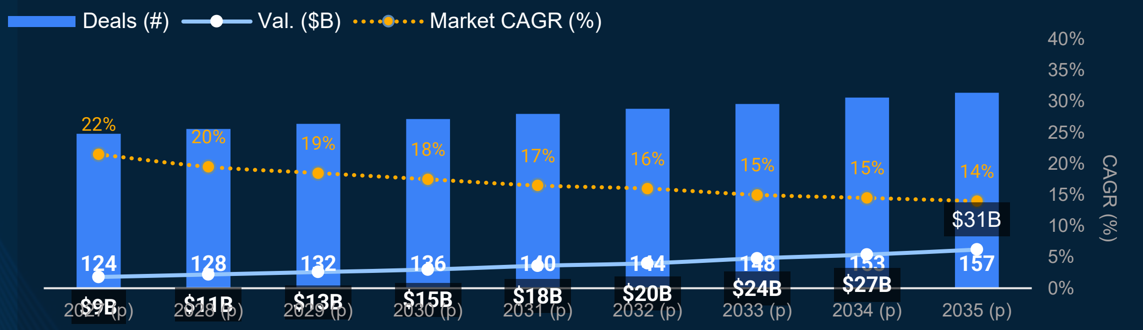
## Crossover point 2031: Edge (152 deals) > Cloud (140 deals)

### Cloud AI (Hyperscalers)

Deals ‘35: 157

Val: $31B

LLM Sovereigns | Vertical SaaS | Centralized Infra




M&A CAGR (Avg 2027-35): Deals 3.0% | Val CAGR Var 18.2%

#### Multiples EBITDA

15x - 25x (Median 22x)

0x 10x 20x 30x

Strong consolidation limit deals (+3%/year). Valuation adjusted to $31B in 2035 (20% of the market) respecting the maturity and saturation assumption.

#### AI Cloud Buyers

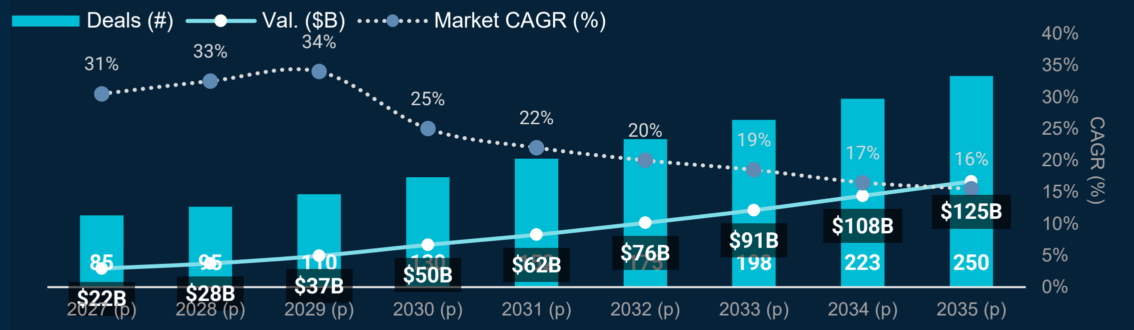
- GOOGLE
- MICROSOFT
- AMAZON
- SALESFORCE
- ORACLE
- ALIBABA

### Edge AI (Distributed)

Deals ‘35: 250

Val: $125B

Quantum Edge | Swarn | Neuromorphic




M&A CAGR (Avg 2027-35): Deals 12.0% | Val CAGR Var 24.5%

#### Multiples EBITDA

12x - 20x (Median 16x)

0x 10x 20x 30x

Continued expansion generates new deals (+12% per year). Valuation of $125B in 2035 (80%) confirms Edge dominance. Deal crossover point now in 2031 (152 vs. 140).

#### Edge Buyers

- NVIDIA
- INTEL
- AMD
- QUALCOMM
- APPLE
- TESLA

STRATEGIC ROADMAP

# Edge AI roadmap : 3 phases to a successful transition

## Strategic roadmap for a transition to Edge AI

### Trajectory

Transition from cloud-first to edge-first architectures for critical inference, driven by latency and privacy.

### Models & Architecture

Generalization of SLMs on constrained devices and the rise of swarms of agents (Swarm Intelligence) for collaborative autonomy.

### 2020 – 2025

- Opportunity : First Edge LLM deployments & TCO optimization
- Opportunity : Critical Latency Reduction (Automotive/IoT)
- Challenge : Limited Hardware Maturity (NPU)
- Challenge : Hybrid MLOps Complexity

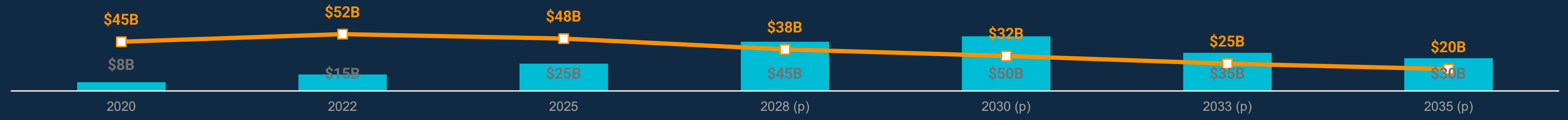
### 2025 – 2030

- Opportunity : Autonomous Agentic AI & E2E Automation
- Opportunity : New Business Models (Decentralized SaaS)
- Challenge : Governance & Distributed Security
- Challenge : Heterogeneous interoperability

### 2030 – 2035

- Opportunity : Swarm Intelligence & Collaboration massive
- Opportunity : Edge AI pervasive (Smart Cities/BioTech)
- Challenge : Ethics & Control of Autonomy
- Challenge : Overall energy consumption

### Investment Evolution: CAPEX vs OPEX (2020-2035)




**OPEX (declining after 2022 : -62%)**
• Cloud → Edge migration • Local processing • Agent Ops automation • NPU / TPU efficiency • SLM vs LLM optimization

**CAPEX (declining after 2030 : -40%)**
• Deployed infrastructure • Economy of scale • Technology maturity • Standardization

**0 – 6 months**

### Initiation & POC

- Select 3-5 priority use cases (Latency < 50ms)
- Launch POCs on Edge LLM/SLM (e.g., Llama-3-8B)
- Evaluate Cloud vs Edge Comparative TCO

**6 – 18 months**

### Hybrid Industrialization

- Deploy a Hybrid MLOps Platform (Train Cloud / Infer Edge)
- Integrate AI Act & Security Governance
- Implementation of the first Autonomous Agents (AgentOps)

**18 – 36 months**

### Scale & Autonomy

- Multi-site scaling & Heterogeneous fleets
- Activation of the Agentic Mesh (Swarm)
- Continuous Latency/Energy Optimization

KPIs for Success

| < 10 ms | -40% | +25% | Zero |
|---|---|---|---|
| Target Latency | Cost per Inference | Automation rate | Privacy (Data Egress) |

# SECTION 5 – EMPLOYMENT

# 690K Edge & Cloud AI jobs : which sectors are hiring the most?

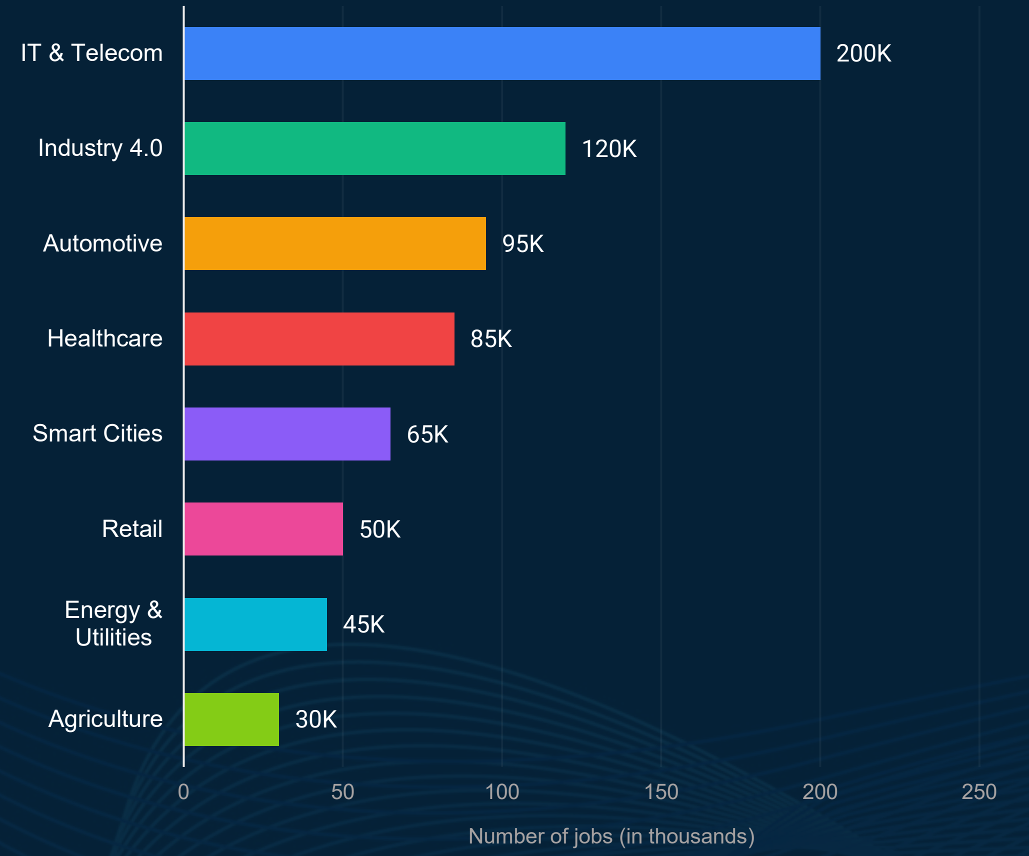
## IT/Telecom and Industry 4.0 account for 46% of opportunities – Full sector analysis

**690K**
Net global job creation (2027–2035)

**2.3M**
Talent shortage by 2027

**+56%**
AI salary premium

### Sector Breakdown (in thousands)




IT &Telecom and Industry 4.0 concentrate **46%** of new Edge AI jobs

### IT & Telecom
Top Sector
Trend: Massive deployment of 5G, distributed architectures
Skills: Edge Orchestration, Kubernetes, Low-latency networking

### Industry 4.0
Strong Growth
Trend: Autonomous factories, real-time predictive maintenance
Challenges: IT/OT convergence, industrial cybersecurity

### Automotive
High Value
Trend: Software-defined vehicles (SDV), autonomous driving
Outlook: Explosion in demand for embedded AI

### Healthcare
Critical
Trend: On-device diagnostics, robotic surgery
Skills: Ethical AI, patient data privacy

### Smart Cities
Trend: Intelligent traffic management, public safety
Challenges: Large-scale system interoperability

### Retail
Trend: Autonomous stores, in-store personalization
Skills: Computer vision, behavioral analytics

### Energy & Utilities
Outlook: Key role in the energy transition
Trend: Decentralized smart grids, renewable optimization

### Agriculture
Trend: Precision agriculture, autonomous robots
Challenges: Connectivity in remote rural areas

# Cloud AI & Edge AI talent geography : Where are the 690K jobs being created?

North America (40%), Europe (30%), Asia-Pacific (25%), Rest of the World (5%)

## North America — 275K Jobs (40% of the total)

| IT & Telecom 80K | Industry 4.0 50K | Automotive 40K | Healthcare 35K | Smart Cities 25K | Retail 20K | Energy & Util. 18K | Agriculture 12K |
|---|---|---|---|---|---|---|---|
| 5G Network Architect<br>Edge Computing Eng.<br>AI Systems Designer | Industrial IoT Eng.<br>Predictive Maint. Spec.<br>Smart Factory Des. | ADAS Systems Eng.<br>V2X Comm. Spec.<br>Autonomous Vehicle AI | Medical AI Diag.<br>Health Data Analyst<br>Telemedicine Eng. | Urban IoT Architect<br>Smart Infra. Eng.<br>City Data Analyst | Retail Edge Analytics<br>Customer Behavior AI<br>Auto. Store Des. | Smart Grid Engineer<br>Energy AI Analyst<br>Renewable Opt. | Precision Agri. Eng.<br>Farm AI Developer<br>Crop Monitoring |

## Europe — 205K Jobs (30% of the total)

| IT & Telecom 65K | Industry 4.0 45K | Automotive 35K | Healthcare 60K | Smart Cities 25K | Retail 17K | Energy & Util. 18K | Agriculture 10K |
|---|---|---|---|---|---|---|---|
| Edge AI Platform Dev.<br>Telco AI Spec.<br>Network Opt. Eng. | Ind. 4.0 Integration<br>AI Quality Control<br>Digital Twin Arch. | Embedded Auto. AI<br>Vehicle Safety Des.<br>Connected Car Arch. | Clinical AI Spec.<br>Medical Imaging AI<br>Healthcare IoT Dev. | Smart Mobility Eng.<br>Urban Planning AI<br>Public Safety AI | Omnichannel AI Dev.<br>Retail IoT Eng.<br>Customer Exp. Opt. | Grid Optimization<br>Energy Storage AI<br>Sustainability AI | AgriTech AI Eng.<br>Sustainable Farming<br>Agri. IoT Developer |

## Asia - Pacific — 170K Jobs (25% of the total)

| IT & Telecom 65K | Industry 4.0 45K | Automotive 35K | Healthcare 60K | Smart Cities 25K | Retail 17K | Energy & Util. 18K | Agriculture 10K |
|---|---|---|---|---|---|---|---|
| IoT Edge Engineer<br>Mobile AI Developer<br>Smart Infra. Arch. | Mfg. Edge AI<br>Quality AI Spec.<br>Supply Chain Opt. | EV Battery AI<br>Auto. Driving Sys.<br>Smart Vehicle Eng. | Medical Device AI<br>Patient Monitor AI<br>Health Data Sec. | Smart City Platform<br>Traffic AI Eng.<br>Public Service AI | E-commerce Edge<br>Retail Analytics<br>Customer Insight | Renewable Energy AI<br>Grid Stability<br>Energy Trading AI | Smart Farming AI<br>Crop Yield Opt.<br>Livestock Monitor |

## Rest of the World — 40K Jobs (5% of the total)

| IT & Telecom 65K | Industry 4.0 45K | Automotive 35K | Healthcare 60K | Smart Cities 25K | Retail 17K | Energy & Util. 18K | Agriculture 10K |
|---|---|---|---|---|---|---|---|
| Remote Network Eng.<br>Satellite AI Spec.<br>Rural Conn. Arch. | Emerging Mkt Mfg.<br>Low-Cost Auto. Spec.<br>Process Eff. Eng. | Local Mobility AI<br>Vehicle Retrofit AI<br>Transport Fleet Opt. | Remote Diag. AI<br>Public Health Data<br>Mobile Clinic Tech. | Basic Infra. AI<br>Water Mgmt. AI<br>Traffic Flow Spec. | Mobile Pay. AI<br>Supply Chain Track.<br>Local Market Anal. | Solar Grid AI<br>Off-grid Power Sys.<br>Resource Mgmt. AI | Small Farm AI<br>Drought Pred. AI<br>Soil Health Analyst |

TALENT TRANSFORMATION

# Cloud AI vs Edge AI new Jobs: 72 Emerging Roles

Transformation of professional profiles with a critical need for hybrid skills

## Cloud AI Roles

- Cloud AI Architect
- ML Platform Engi.
- Data Pipeline Eng.
- Cloud AI Security
- Dist. Systems Eng.
- Cloud MLOps Eng.
- AI Product Manager
- Cloud AI Ethics

## Required Skills

**Cloud AI Skills**

Python & PyTorch · Cloud Arch. · Big Data · API Design · Dist. Compute

**Edge AI Skills**

C++/Rust · IoT Protocols · Embedded ML · Hardware Opt. · RTOS · Low-power

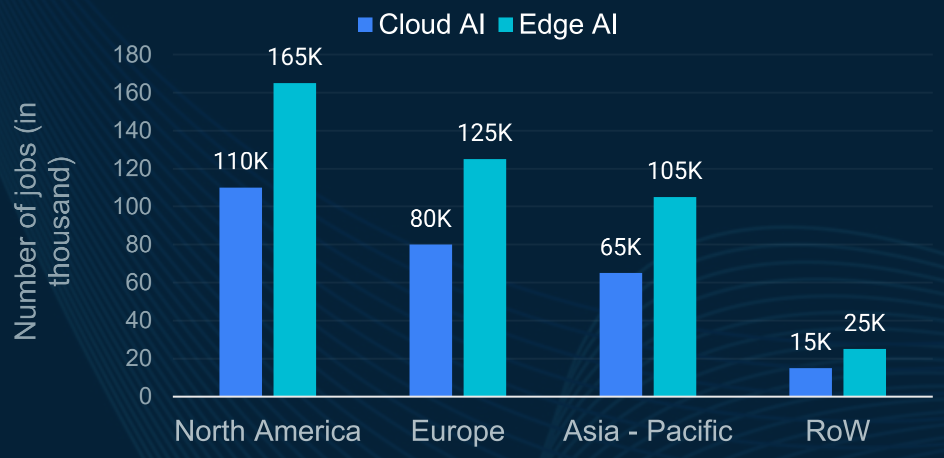
**Transversal Skills**

Data Science · MLOps · Cybersec · AI Ethics

## Edge AI Roles

- Edge AI Eng.
- Embedded ML Eng.
- IoT AI Developer
- Edge Security Arch
- Real-time Syst. Eng.
- Edge MLOps Spec.
- Hardware AI Opt.
- Edge Privacy Officer

## Geographic Distribution: Cloud AI vs Edge AI




### Global Synthesis

Edge AI dominates job creation across all regions, with an average ratio of 1.5 to 1 compared to Cloud AI. North America remains the absolute leader in volume (40% of the global total), but Asia-Pacific shows the fastest growth rate, driven by widespread adoption in manufacturing and smart cities.

### North America — Total: 275K

Leader in software innovation and R&D. High demand for hybrid architects and edge security engineers.

### Europe — Total: 205K

Focus on Industry 4.0 and the automotive sector (Germany, France). Regulation (AI Act) boosts compliance

### Asia - Pacific — Total: 170K

Explosive growth in consumer electronics and robotics. China and Japan leading the way in edge hardware.

### Rest of the World — Total: 40K

Emerging adoption in agriculture (LATAM) and energy infrastructure (Middle East).

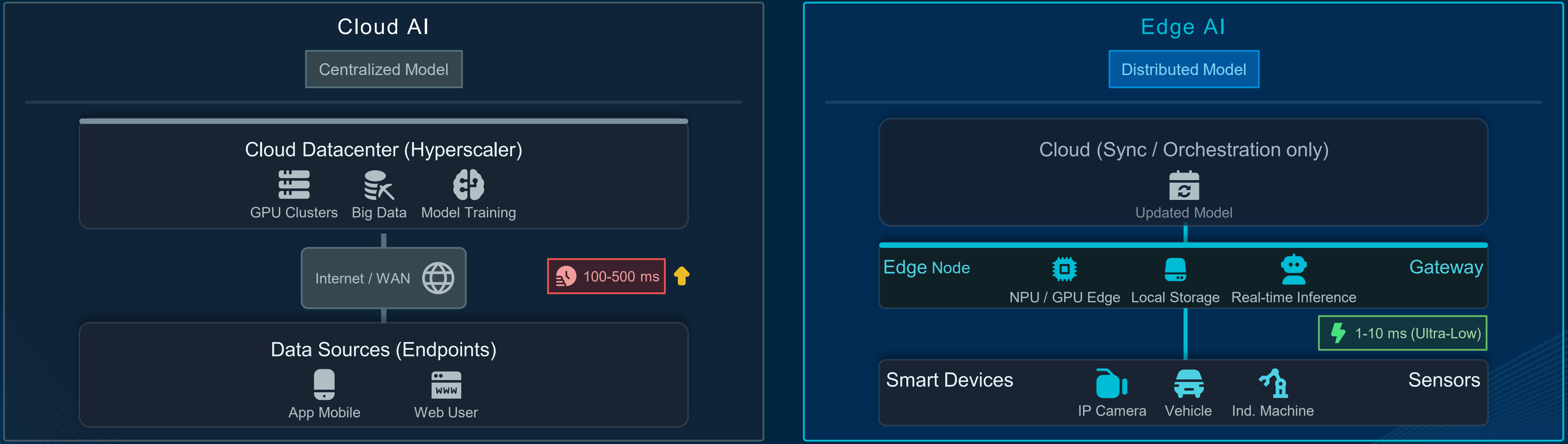
ALVAREZ & MARSAL

# SECTION 6 – OPERATING MODEL

TECH INFOGRAPHIC

# Cloud vs. Edge : 8 fundamental differences showing the shift in paradigm

Structural divergence: Massive centralization vs. real-time distributed processing




| Feature | Cloud AI (Training & Massive Inference) | Edge AI (Real-Time Inference & Privacy) |
|---|---|---|
| Latency | Variable, depends on the network (100ms+) | Deterministic, Real-Time (< 10ms) |
| Bandwidth | High consumption (Upload Data + Download Result) | Efficient (Local processing, only insights come out) |
| Security | Transmission of sensitive data off-site | Data Residency Guaranteed (Data Remains Local) |
| Costs (TCO) | High OPEX (Cost per API call + Storage) | Initial CAPEX, Reduced OPEX (No egress fees) |
| Maintenance | Centralized, instant update for everyone | Distributed, requires fleet management (MDM/OTA) |
| Scalability | Near-infinite horizontal scalability (elasticity) | Scalability limited by local hardware (addition of nodes) |
| Energetic Constraint | Massive consumption (energy-hungry data centers) | Optimized for low consumption (battery, IoT) |
| Offline Operation | Complete dependence on network connectivity | Autonomous offline operation (resilience) |

# 2035 : 4 Players, 4 Revolutions - The New Edge AI Value Chain

Structural transformation of the 4 key players in the Edge AI / Cloud AI value chain

## Industries

**Current 2025**

Cloud-first, OT/IT silos. High latency, long ROI.

**Target 2035**

Edge-first, local agentic AI, ROI <6 months.

- **Organization** OT/IT Silos → CoE Edge AI Convergence
- **Processus** Reactive maintenance → CI/CD MLOps Edge
- **Tools** SCADA/Cloud → Unified Platforms

- Decentralized MLOps "at the foot of the machine"
- Data Sovereignty & Zero Trust

## Connectivity

**Current 2025**

Centralized NSA 5G. Limited MEC, best-effort.

**Target 2035**

5G SA/6G distributed. On-demand slicing, guaranteed SLAs.

- **Organization** Net/IT Silos → DevNetOps Teams
- **Processus** Manual provisioning → Zero-touch auto
- **Tools** Legacy OSS → Cloud-native Orchestration

- Scaled MEC & Compute-Network Conv.
- B2B2X Revenue-Share Models

## Cloud AI

**Current 2025**

Centralized. High egress costs, variable latency.

**Target 2035**

Hybrid Cloud-Edge. Federated Learning, orchestration dist.

- **Organization** Centralized → Dist. Edge Teams
- **Processus** Monolith deployment → Federated Learning
- **Tools** IaaS/PaaS → Edge Extensions (Arc/Outposts)

- Sovereign "Cloud-to-Edge" Extension
- AI/Agent Model Marketplace

## Edge AI

**Current 2025**

Silicon-centric. Fragmented SDKs, pilots, one-off revenues.

**Target 2035**

Full-Stack Platform. Silicon + Runtime + OTA. Recurring Rev.

- **Organization** Silicon focus → Full-stack teams
- **Processus** Custom development → Unified SDKs & OTA
- **Tools** Fragmented → Unified Toolchains (ONNX)

- Standardization & Portability Models
- Hardware/software "as-a-Service" model

## Cross-Functional KPIs Targets 2035

- 5 – 10ms Target Latency
- ↓ 40 – 70% Cost per Inference
- > 70% Part Edge Inference
- 100% Zero Trust Native

# Industries : ROI < 6 Months or Failure - Edge AI Becomes Non-Negotiable

Transforming the model : From cloud centralization to distributed, edge-first intelligence

## 2025 Cloud-Centric Model

### Architecture & Infra

**Cloud-first** approach. Raw data sent to the cloud for analysis. Rigid silos between OT (Operations) and IT (Information).

### Value Creation

Occasional efficiency gains. Basic preventive maintenance. Few monetized data-driven services.

### Organization / Processes / Tools

- **Organization :** Separate OT/IT silos. Small, centralized IT data science teams.
- **Processus :** Reactive maintenance. Long deployment cycles (months). Dispersed POCs.
- **Tools :** Legacy SCADA/MES. Cloud ML (SageMaker, Azure ML). Manual integration.

### Costs & CAPEX

High cloud OPEX (storage, egress). CAPEX limited to pilot projects (POC). Hidden integration costs.

### Ecosystem

Dependence on Hyper scalers (AWS, Azure) and traditional IT Integrators.

**As Is KPIS**

| 80 – 200 ms | Silos | 12 – 18 months | < 10% |
|---|---|---|---|
| Average Latency | OT Security | Average ROI | Local Processing |

- IT/OT Convergence
- MLOps ‘At the Edge’
- Zero-Trust Native
- Data Governance

## 2035 Edge-First Model

### Architecture & Infra

**Edge-first** approach. Local NPUs/IPCs, Embedded Digital Twins. Decentralized local MLOps.

### Value Creation

X-as-a-Service models (Quality, Safety). Monetization of industrial data. Autonomous operations.

### Organization / Processes / Tools

- **Organization :** OT/IT convergence. Integrated MLOps teams. Edge AI Center of Exc.
- **Processus :** Predictive maintenance. CI/CD templates. Continuous OTA deployment.
- **Tools :** Unified Edge Platforms (Jetson/OpenVINO). Kubeflow Edge. Private 5G.

### Costs & CAPEX

Overall TCO ↓ 50% (less cloud/network). CAPEX modernization of lines & Edge Nodes.

### Ecosystem

OEM Edge (NVIDIA, Qualcomm), Telcos (Private 5G), Hyperscalers (Hybrid).

**Target KPIS**

| < 10 ms | Unified | < 6 months | > 80% |
|---|---|---|---|
| Target Latency | OT Security | Average ROI | Local Processing |

# Telcos 2035 : From Silent Pipes to Edge AI Orchestrators (+30% ARPU)

Transforming the model : From pure connectivity to monetized Edge & API services

## 2025 "Dumb Pipe" model

### Architecture & Infra

5G NSA (Non-Standalone), Cloud-centric Backhaul. Very limited MEC (Multi-Access Edge). Network slicing exists but is poorly monetized.

### Value Creation

B2B ARPU under pressure. Price competition on data. Basic IoT offerings with low margins.

### Organization / Processes / Tools

- **Organization** : Silos Network/IT, OSS/BSS legacy
- **Processus** : Manual provisioning, best-effort SLA
- **Tools** : RAN/Core trad. (Ericsson/Nokia)

### Costs & CAPEX

High network OPEX. Heavy 5G investments. Complex slicing management. Expensive energy/data transport.

### Ecosystem

Hyperscalers (limited area partnerships), traditional RAN equip. manufacturers.

### Current KPIS

| 20 – 50 ms | < 10% |
|---|---|
| 5G Latency | MEC Attachment |

→

- Marketplace APIs
- Edge Co-investments
- Revenue-Share Models
- Network-as-a-Service

## 2035 Edge-Native Service Model

### Architecture & Infra

5G SA / 6G. Massively distributed MEC, local UPF. Nationwide Edge POPs. Exposed and standardized network APIs.

### Value Creation

API monetization (on-demand QoS). Vertical slicing (NPN factories/automotive). Edge PaaS. B2B ARPU +20-30%.

### Organization / Processes / Tools

- **Organization** : Network-Cloud convergence, API products.
- **Processus** : Zero-touch provisioning, real-time SLA
- **Tools** : O-RAN, K8s cloud-native, NWDAF

### Costs & CAPEX

Automation (AIOps). Energy efficiency. CAPEX targeted at local MEC/Cores rather than transportation.

### Ecosystem

Hyperscalers (Wavelength, Outposts), Vertical ISVs, OT Integrators.

### Target KPIS

| 5 – 10 ms | 30 – 40% |
|---|---|
| Target Latency | MEC Attachment |

AWS, AZURE, GCP

# Hyperscalers 2035 : Why AWS, Azure and Google are investing heavily in the Edge

Transforming the model : From massive centralization to distributed Cloud-Edge hybridization

## 2025 Centralized Model

### Architecture & Infra

**Centralized** training and inference. Massive GPU clusters. Latency >50ms. Sovereignty and lock-in challenges.

### Value Creation

High IaaS/PaaS/MLaaS consumption. High egress fees. Margins under pressure (GPU/energy cost).

### Organization / Processes / Tools

- **Organization** : Centralized teams, IaaS/PaaS focus, regional silos.
- **Processus** : Centralized deployment, egress fees, monolithic models.
- **Tools** : TensorFlow, PyTorch, SageMaker, Azure ML, Vertex AI.

### Costs & CAPEX

CAPEX for massive data centers. OPEX for energy and cooling that are constantly increasing.

### Ecosystem

Nvidia (GPU), Intel (CPU), Databricks (ML Ops). Vendor lock-in.

**As Is KPIS**

| > 50 ms | IaaS / PaaS | High | Complex |
|---|---|---|---|
| Average Latency | Billing | Egress Fees | Sovereignty |

- Sovereign Offers
- Federated Learning
- Co-managed SLA
- Marketplace Agents

## 2035 Distributed Hybrid Model

### Architecture & Infra

**Hybrid** cloud-edge. Low-latency gateways, regional caches. Serving on-premises/sovereign.

### Value Creation

Inference token-based billing. Vertical platforms (Industrial / Automotive / Healthcare). Model governance services.

### Organization / Processes / Tools

- **Organization** : Distributed edge teams, focus on vertical platforms, geo-distributed pods.
- **Processus** : Hybrid deployment, federated learning, distributed models, multi-site orchestration.
- **Tools** : Edge extensions (AWS Wavelength, Azure Arc, GCP Distributed Cloud), multi-cluster Kubeflow, ONNX Runtime.

### Costs & CAPEX

Cost/inference ↓ 40%. Optimized egress (↓ 30-50%). Energy efficiency (liquid cooling).

### Ecosystem

Open and co-managed ecosystem : Telcos (MEC), Edge silicon vendors

**Target KPIS**

| < 20 ms | Token | Native | ↓ 30-50% |
|---|---|---|---|
| Edge Bridge SLO | Billing | Sovereignty | Egress Fees |

# From Silicon to Software : How Edge AI is Reinventing Chipmakers

Transforming the model: From Silicon-Centric to Full-Stack Edge Platform

## 2025 Silicon-Centric Model

### Architecture & Infra

Focused on **hardware (silicon)**. Fragmented SDKs and compilers. Focus on demos and technical drivers.

### Value Creation

Transactional hardware sales (chips/SoCs). Limited SDK licenses. Ad-hoc professional services for integration.

### Organization / Processes / Tools

- **Organization :** Silicon-centric teams, vendor-fragmented SDKs, responsive support.
- **Processus :** Custom development, manual integrations, long release cycles.
- **Tools :** CUDA, OpenVINO, TensorRT (siloed), proprietary compilers.

### Costs & CAPEX

CAPEX : Silicon R&D 15-20% of revenue, High BOM (fabs, tapeouts). OPEX: Technical support, certifications (ISO 26262, ASIL).

### Ecosystem

OEM device partners, niche integrators, open-source developer communities.

| As Is KPIS | | | |
|---|---|---|---|
| 20 – 40 ms | < 15% | Variable | Complex |
| Average Latency | Soft. Part | Security | Portability |

- Toolchain Standard.
- OEM Co-design
- Secure Supply Chain
- Compliance (ISO26262)

## 2035 Full-Stack Platform Model

### Architecture & Infra

**Full-Stack approach** (Silicon + Runtime + Unified SDK + MLOps + OTA). Certified vertical references (Safety/Medical).

### Value Creation

Recurring Software Mix >40% (Licenses, Support, Marketplaces). Managed Edge Services.

### Organization / Processes / Tools

- **Organization :** Full-stack teams (HW+SW), unified platform, proactive Customer Success.
- **Processus :** Standardized SDKs (ONNX), OTA deployment, native CI/CD, marketplaces
- **Tools :** ONNX Unified Runtime, standardized toolchains, native MLOps, TEE/SE.

### Costs & CAPEX

CAPEX: Full-Stack R&D 12-18% of revenue. Optimized BOM ↓15-20% (adv. nodes, co-design). OPEX: Cloud platforms, Cust. Success, exp.certifications.

### Ecosystem

Hyperscalers (Compilers/Backends), Telcos (MEC), Vertical ISVs, Indust. OEMs.

| Target KPIS | | | |
|---|---|---|---|
| < 10 ms | > 30 – 40% | TEE / SE | > 80% |
| Target Latency | Soft. Part | Security | Portability |

# SECTION 7 –

# KEY PROCESSES & EDGE AI TCO

# Monitor the Edge AI : 50x More Complex, 10x More Expensive

Comparative analysis of observability : From a unified centralized view to distributed blind spots

## Cloud AI Only

CENTRALIZED MONITORING

- Centralized Collection
- Unified Aggregation
- Real-Time Analysis
- Unified Alarms
- Global Dashboards

100% Visibility

**Complexity : Low (1x)**

## Cloud + Edge AI

DISTRIBUTED MONITORING

- Local Collection (1000s Nodes)
- Pre Aggregation & Edge Filtering
- **Disconnections & Buffer Storage**
- Selective Transmission (Costs)
- Distributed Correlation (Time)
- Alarms Prioritization
- Dashboards

Partial Visibility

**Complexity : High (50x)**

## 8 Dimensions of Complexification

| Dimension | Factor | From | To |
|---|---|---|---|
| Collection Point | x1000 | 1 Data Centre | 1000s Nodes |
| Data Volume | x100 | GB / Day | TB / Day (Dist.) |
| Collection Latency | x10 | Real Time | Differed / Buffered |
| Correlation | x5 | Direct | Multi Sources |
| Alarms | x3 | Single Source | Multi Levels |
| Storage | x5 | Centralized | Dist. & Aggregated |
| Costs | x3 | Fixed | Variable (Trans.) |
| Visibility | x2 | 100% | Blind Zones |

ALVAREZ & MARSAL

INCIDENT MANAGEMENT

# Edge AI Incidents : Why 1000 Sites = Complexity × 5

Comparative analysis of the resolution chain : From centralized simplicity to distributed complexity

## Cloud AI Only

CENTRALIZED ARCHITECTURE

1. Centralized Detection (CloudWatch)
2. Diagnostic Unified Logs
3. Single Team Escalation (SRE)
4. Resolution & Push Central
5. Global Validation

Linear Resolution — **Complexity: Low (1x)**

## Cloud + Edge AI

DISTRIBUTED ARCHITECTURE

- Local Alert (Edge) | Central Alert
- Localization & Correlation (1000s nodes)
- Hybrid Diagnostics (Network ? Device ? Cloud ?)
- Multi-team escalation (Cloud / Edge / Site Ops)
- Local Fix | Global OTA
- Complex Rollback (Dependency Management)

Multi-level resolution — **Complexity : Critical (5x)**

## 8 Dimensions of Complexification

| Dimension | Factor | From | To |
|---|---|---|---|
| Topology | x50 | 1 Data Centre | **1000s Nodes** |
| Detection | x3 | Centralized | **Distributed** |
| Diagnostic | x5 | Unified Logs | **Correlation** |
| Teams | x5 | Central | **Cloud / Edge / Site** |
| Tools | x2 | Single Stack | **Multi-Stack** |
| Deployment | x4 | Direct Push | **OTA Cond.** |
| Rollback | x5 | Global | **Granular** |
| Governance | x2 | Simple SLA | **E2E SLA** |

# Change Management Edge AI : Why 1 Team Becomes 1000

Comparative analysis of change management : From uniform adoption to multi-site transformation

## Cloud AI Only

CENTRALIZED AND LINEAR

- Single Planning
- Centralized Training
- Global Testing & Validation
- "Big Bang" deployment
- Post-Deployment Support

Uniform Adoption — **Complexity : Low (1x)**

## Cloud + Edge AI

MULTI-SITES & WAVES

- Strategic Planning
- Cascade Communications / Distributed Training
- Local Pilots (POCs)
- Conditional Wave Deployment
- Location Validation / Site Adjustment
- Multi-level support

Heterogeneous Adoption — **Complexity : High (10x)**

## 8 Dimensions of Complexification

| Dimension | Factor | From | To |
|---|---|---|---|
| Teams Perimeter | x50 | 1 Team | **1000s Dist.** |
| Communication | x5 | Direct | **Multi Level Casc.** |
| Training | x4 | Centralized | **Dist. / Cultural** |
| Adoption | x3 | Uniform | **Heterogenous / Site** |
| Resistance | x5 | Centralized | **Geographically Dist.** |
| Timeline | x3 | Linear | **Cond. Waves** |
| Governance | x2 | Centralized | **Local Autonomy** |
| Success Measurement | x2 | Central KPIs | **Multi Level KPIs** |

CYBERSECURITY & COMPLIANCE

# AI Edge Security: Attack Surface × 100, Compliance × 10 Jurisdictions

Comparison of security challenges : From the centralized model to distributed multi-site governance

## Cloud AI Only

CENTRALIZED PERIMETER

- Threat Detection
- Centralized Analysis
- Unified IAM
- Synchronous Global Patch
- Conformity Validation

Single Jurisdiction — **Complexity : Low (1x)**

## Cloud + Edge AI

SURFACE x 100

- Local / Central Detection
  - Correlation
  - Isolation Node
- **Physical Access / Model Poisoning**
- Zero-Trust Authentication / KMS Edge
- Conditional OTA Patch
- Multi Jurisdictions Validation

Multi Jurisdictions — **Complexity : Critical (100x)**

## 8 Dimensions of Complexification

| Dimension | Factor | From | To |
|---|---|---|---|
| Attack Surface | x100 | 1 Data Centre | **1000s Nodes** |
| Authentication | x5 | IAM Central | **Zero Trust** |
| Encryption | x5 | Centralized Keys | **KMS Edge** |
| Compliance | x10 | Unique | **Multi Jurisdictions** |
| Audits | x5 | Centralized | **Distributed** |
| Patching | x3 | Global Sync. | **OTA Cond.** |
| Isolation | x5 | Network | **Micro Segment** |
| Governance | x3 | Central Policy | **Multi Level** |

ALVAREZ & MARSAL

# CAPEX x 50, OPEX -47% : The Economic Paradox of Edge AI

Comparative Financial Analysis : From pure OPEX predictability to profitable CAPEX/OPEX hybridization

## Cloud AI Only

PREDICTABLE COSTS

- Unique Estimate
- Cloud Provisioning
- Usage (Pay-As-You-Go)
- Monthly Billing
- Simple Optimization

Linear & Predictable — **Complexity : Low (1x)**

## Cloud + Edge AI

HYBRIDE & VARIABLE

- Multi Components Estimate
- Initial Edge CAPEX (High)
- **High CAPEX & Obsolescence**
- Distributed Deployment
- Variable Transmission Costs
- Sites Maintenance
- Hybrid Billing

Variable & Complex — **Complexity : High (10x)**

## 8 Dimensions of Complexification

| Dimension | Factor | From | To |
|---|---|---|---|
| CAPEX Structure | x50 | 1 Data Centre | **1000s Nodes + Infra** |
| Compute OPEX | x5 | Centralized | **Hybrid Distributed** |
| Transmission | x3 | Egress Fees | **Bidirect. Variable** |
| Maintenance | x10 | Centralized | **Local Multi Sites** |
| Energy | x2 | Opt. Data Centre | **Distributed Non-Opt.** |
| Vendors | x5 | 1 / 2 Vendors | **Multi-Vendors** |
| Predictability | x3 | Linear | **Variable + Peaks** |
| Optimization | x5 | Simple (Resource) | **Complex (Cache)** |

2.5 YEARS ROI

# €360 Million in Savings Over 10 Years: A Business Case for Edge AI

## Comparative Economic Model: The Winning Bet of the Initial Edge Investment

### Assumptions (Reference Basis)

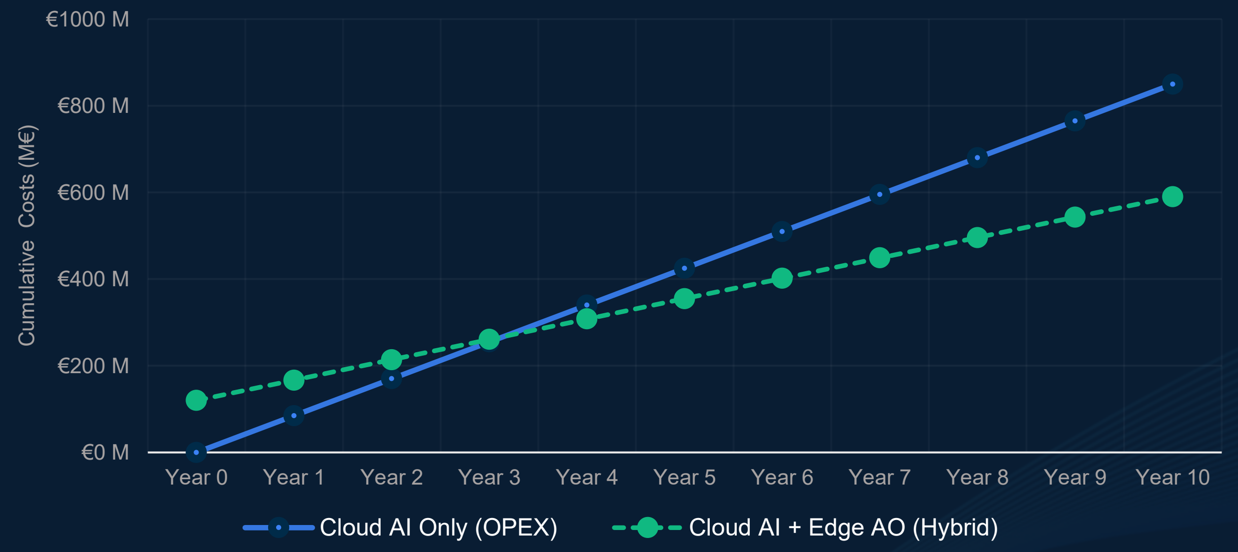
- Scope: 1000 Industrial Sites
- Load: Compute Intensive AI
- Period: 10 Years (2025-2035)
- Unit: Millions of Euros (€M)

### Comparative Cost Structure

| Cost Item | Cloud AI Only | Cloud + Edge AI | Delta |
|---|---|---|---|
| **CAPEX (Initial Investment)** | | | |
| Infra & Devices | 0 M€ | 120 M€ | +120 M€ |
| **OPEX (Annual Costs)** | | | |
| Compute & Storage | 48 M€ | 23 M€ | -52% |
| Transmission (Egress) | 20 M€ | 4 M€ | -80% |
| Maintenance & Energy | 0 M€ | 8 M€ | N/A |
| Licences Soft | 5 M€ | 4 M€ | -20% |
| **Total OPEX / Year** | **73 M€** | **39 M€** | **-47%** |
| **FTE (Human Resources)** | | | |
| Ops/Support Teams | 50 FTE | 80 FTE | +60% |
| **Annual FTE Cost** | **4 M€** | **8 M€** | **+100%** |
| **TOTAL ANNUAL** | **77 M€** | **47 M€** | **-39%** |

### Cumulative TCO Evolution (€M)

Break-even point at 4 years (CAPEX amortization)




**€300 M**
Savings over 10 years

**3.2 years**
Break-even point

**-39%**
Annual Cost Reduction

**Strategic Insight:** The hybrid model requires a high entry ticket (€120M) and increased HR complexity, but generates massive efficiency gains from the 4th year onwards thanks to the collapse of transmission and cloud compute costs.

ALVAREZ & MARSAL

# SECTION 8 –
# EXECUTIVE SUMMARY

# Edge AI : $357 Billion (80% of the Market by 2035), 8 Sectors transformed

Strategic framework for the evolution from Cloud AI to Edge AI (2020–2035)

## Timeline 2020-2035

- 2020-2022 : Cloud Dominant
  Massive centralization, LLM training
- 2023-2025 : Edge Emergence
  SLM <10B parameters, NPU, first deployments
- 2026-2028 : Hybrid
  Cloud-Edge orchestration, 50% adoption
- 2029-2031 : Agentic AI
  Autonomous multi-agent systems
- 2032-2035 : Pervasive AI
  Swarm Intelligence, ubiquitous deployment

| 5 | 15 |
|---|---|
| Key Phases | Years of Analysis |

## Sector Impact

- Industry 4.0
  Predictive maintenance & real-time quality
- Automotive
  Autonomous vehicles & ADAS
- Healthcare
  Instant AI diagnostic IA & telemedicine
- Smart Cities
  Intelligent management & public safety
- Retail
  Autonomous checkout & personalization
- Energy & Utilities
  Grid optimization & consumption forecasting
- Agriculture
  Precision irrigation & crop monitoring
- IT & Telecom
  Edge computing & distributed 5G

| 285% | 2.0x |
|---|---|
| Edge Average ROI | vs Cloud ROI |

## Economic Impact

- Edge AI Market Growth
  $24 Md (2025) → $357 Md (2035), CAGR 27.7%. Market share 20% → 80% vs Cloud.
- Edge AI & Cloud AI M&A Consolidation
  210 deals (2026) → 407 deals (2035). Valuation $45 Md → $156 Md (inc. Edge $125 billion, Cloud $31 billion)
- CAPEX/OPEX Transformation
  CAPEX peak $50 Md (2030) then -40%. OPEX reduction -62% (2022-2035). Savings : $32 Md/an.
- Revenue & Employment Impact
  Agentic AI +$450-650 Md in revenues by 2030. 2.3M specialized jobs created by 2027.
- Operational & Environmental Cost Reduction
  Inference cost : -85% (NPU vs GPU Cloud). Energy : 0.1W vs 250W. TCO -50% over 3 years.
- Innovation & Business Models
  Edge-as-a-Service $40 Md (2030). Decentralized SaaS, privacy-first & sovereign revenues.

| $156 Md | -62% |
|---|---|
| 2035 M&A Valuation | OPEX Savings |

STRATEGIC VISION

# 5 Reasons why Edge AI will dominate the Cloud by 2035

80% of workloads will migrate to the edge: 5 Strategic Insights

### Cloud Edge Transition →

80% Edge / 20% Cloud

**Major paradigm shift for real-time workloads by 2035.** The Edge becomes the center of gravity of critical inference.

2035 HORIZON

### Edge AI Market Growth

$357 billion in 2035 (80% of the AI market)

↑ Average CAGR 27.7% (vs $24.0 billion in 2025)

**Sustained growth over the decade 2025-2035**, peaking at 34% in 2029 (BCC Research). Reaches 80% of the total AI market share by 2035.

HYPER-GROWTH

### ROI Outperformance

ROI 2.0x vs Cloud

285% (Edge) vs 145% (Cloud) • Return 4.2 months

**Edge AI delivers superior immediate business value**, particularly in Industry 4.0 (320%) and Automotive (280%).

OPERATIONAL EFFICIENCY

### Explosion M&A Edge AI

250+ Deals in 2035

Valuation $125 billion • Edge overtakes Cloud in 2031

**The volume and value of Edge AI transactions will surpass those of Cloud AI,** marking the maturity and consolidation of the industry.

MARKET CONSOLIDATION

### Agentic AI Revolution

+$450-650 billion

2030 Total economic impact generated, separate from the technology investment • -50% Ops costs

The emergence of Swarm Intelligence and **Edge autonomous agents will create massive economic value** via cognitive automation.

TECHNOLOGICAL FUTURE

TIMELINE

# **From centralization to omnipresence :** 15 years of global transformation

Evolution of architectures: from centralization to distributed and pervasive intelligence

### Cloud Dominance

Centralized learning and inference. Rise of LLMs (GPT-3) and standardized ML Ops pipelines.

### Hybridization & SLM

Emergence SLM (<10B). NPU on device. Local inference for privacy.

### Agentic AI

Autonomous orchestrator agents. Complex local reasoning.

### Swarm Intelligence

Multi-agent collaboration. Ubiquitous Collective Intelligence (Mesh AI).

2020-2022 — Migration → — 2023-2025 — Autonomy → — 2026-2030 — Distribution → — 2031-2035

### Cloud-Centric

IA SaaS. High latency

Cloud 90% — Edge 10%

| | | | |
|---|---|---|---|
| Latency | High | Scalability | Excel. |
| Security | Avg. | Cost | High |

### Hybrid Cloud-Edge

Dynamic offloading

Cloud 65% — Edge 35%

| | | | |
|---|---|---|---|
| Latency | Mod. | Scalability | Good |
| Security | Good | Cost | Medium |

### Edge-Native

Critical local decision

Cloud 40% — Edge 60%

| | | | |
|---|---|---|---|
| Latency | Low | Scalability | Avg. |
| Security | V. Good | Cost | Low |

### Pervasive AI

Standalone mesh networks

Cloud 20% — Edge 80%

| | | | |
|---|---|---|---|
| Latency | V. Low | Scalability | Avg. |
| Security | Excel. | Cost | V. Low |

Model Evolution: Centralized → Hybrid → Distributed

# 2025-2026 data : sources and methodology of the analysis

Consolidated market data 2025-2026 and bibliographic references

## Market Indicators 2025-2026

**$19.0 billion**
Edge AI Market Size
Estimation 2025 (Global)

**30.48%**
CAGR 2025-2034
High scenarios up to 31%

**36%**
NA Market Share
Source: Bayelsawatch 2025

**2.0x**
ROI Edge vs Cloud
Real-Time Industries

**690k**
New Cloud AI & Edge AI jobs
Net addition by 2035

**-50%**
Operational Costs
Via Offloading Edge

## Main Sources & References


### Market Research

- Precedence Research (2024) Edge AI Market Size & Growth Forecast.
- IDC (2024) Worldwide Edge Spending Guide.
- Spherical Insights (2024) Edge AI Market Analysis 2033.
- Grand View Research (2025) Edge Computing Outlook.
- The Business Research Co. (2025) AI Infrastructure Report.

### Strategic Consulting

- McKinsey & Company (2025) "Empowering Industries with Agentic AI".
- McKinsey Insights (2024) The Next Decade of AI (2025-2035).
- Deloitte (2025) Tech Trends: The Spatial Web & Edge.
- Gartner (2024) Hype Cycle for Edge Computing.

### M&A & Valuations

- PWC (2025) "Global M&A Trends in AI".
- BCG (2026) "M&A Outlook: High Expectations".
- Morgan Stanley (2026) "5 Forces Driving M&A".
- Finerva (2026) "Robotics & AI Valuation Multiples".
- Lazard (2025) "M&A Review & Outlook".

### Government

- NIST (2024) AI Risk Management Framework.
- US Dept. of Transportation (2023) Smart City Data.
- US Dept. of Energy (2025) Smart Grid Report.
- European Commission (2024) EU AI Act Guidelines.

### Organizations Int.

- ITU (2024) AI for Good: Edge in Developing Nations.
- IEA (2024) Digitalization & Energy Efficiency.
- WHO (2024) Ethics & Governance of AI.
- World Econ. Forum (2025) Future of Compute.

### Tech Industry

- Premio Inc. (2024) Edge LLMs vs Cloud LLMs.
- Clarifai (2024) Edge vs Cloud AI: Hybrid Future.
- Siemens (2025) Industrial AI & Agentic Future.
- NVIDIA (2024) Edge Computing Solutions.

### Jobs & Talents

- PwC (2025) "Global AI Jobs Barometer".
- LinkedIn (2025) "AI Talent Insights".
- Gartner (2024) "AI Workforce Upskilling".
- JP Morgan(2025) "AI Impact on Jobs".
- BLS (2025) "AI impacts in projections".
- Deloitte( 2025) "Workforce Evolution".

LEADERSHIP. ACTION. RESULTS.℠

ALVAREZ & MARSAL

*CONFIDENTIAL – NOT FOR DISTRIBUTION*