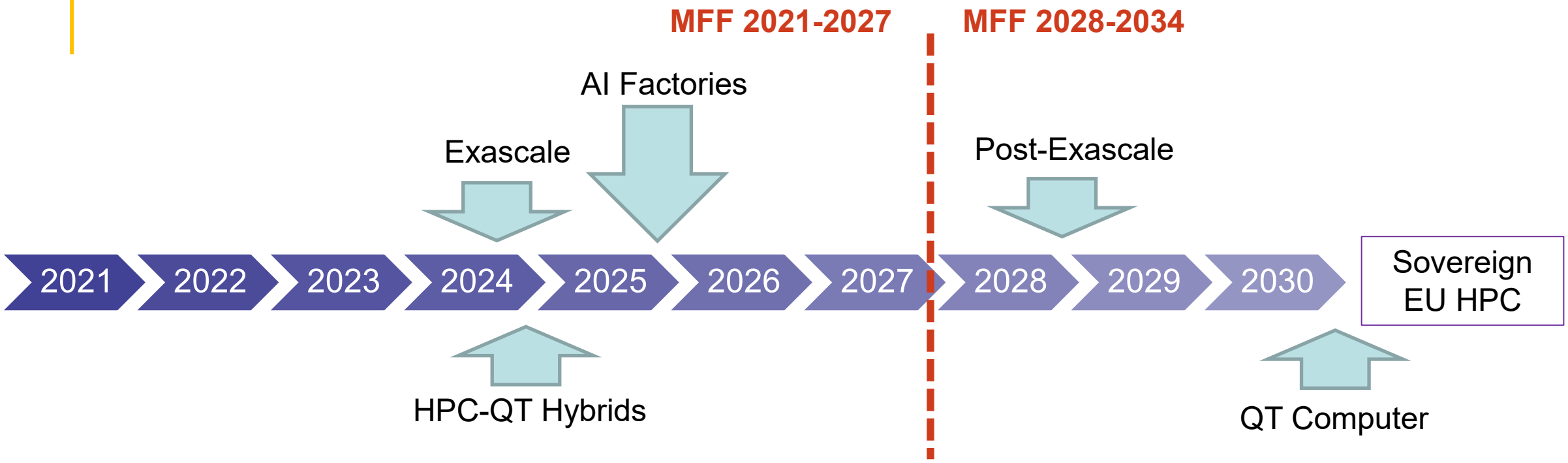
Leonardo Flores Añover - ***In my humble opinion…***

Where are we now?

EuroHPC Roadmap 2025 - 2027



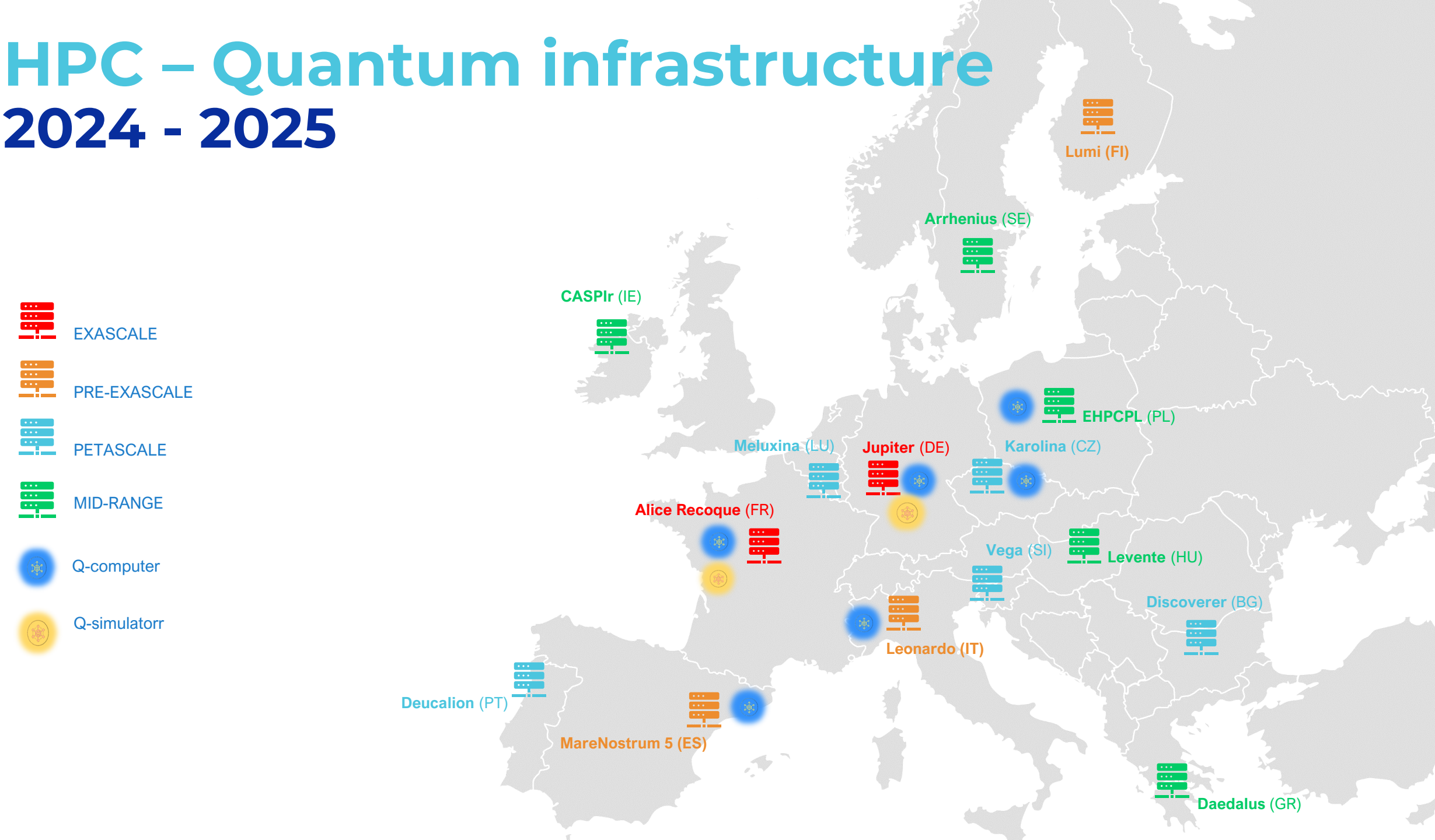
A- Deploying Infrastructure

B- Developing underlying EU technology

C- Making supercomputing useable

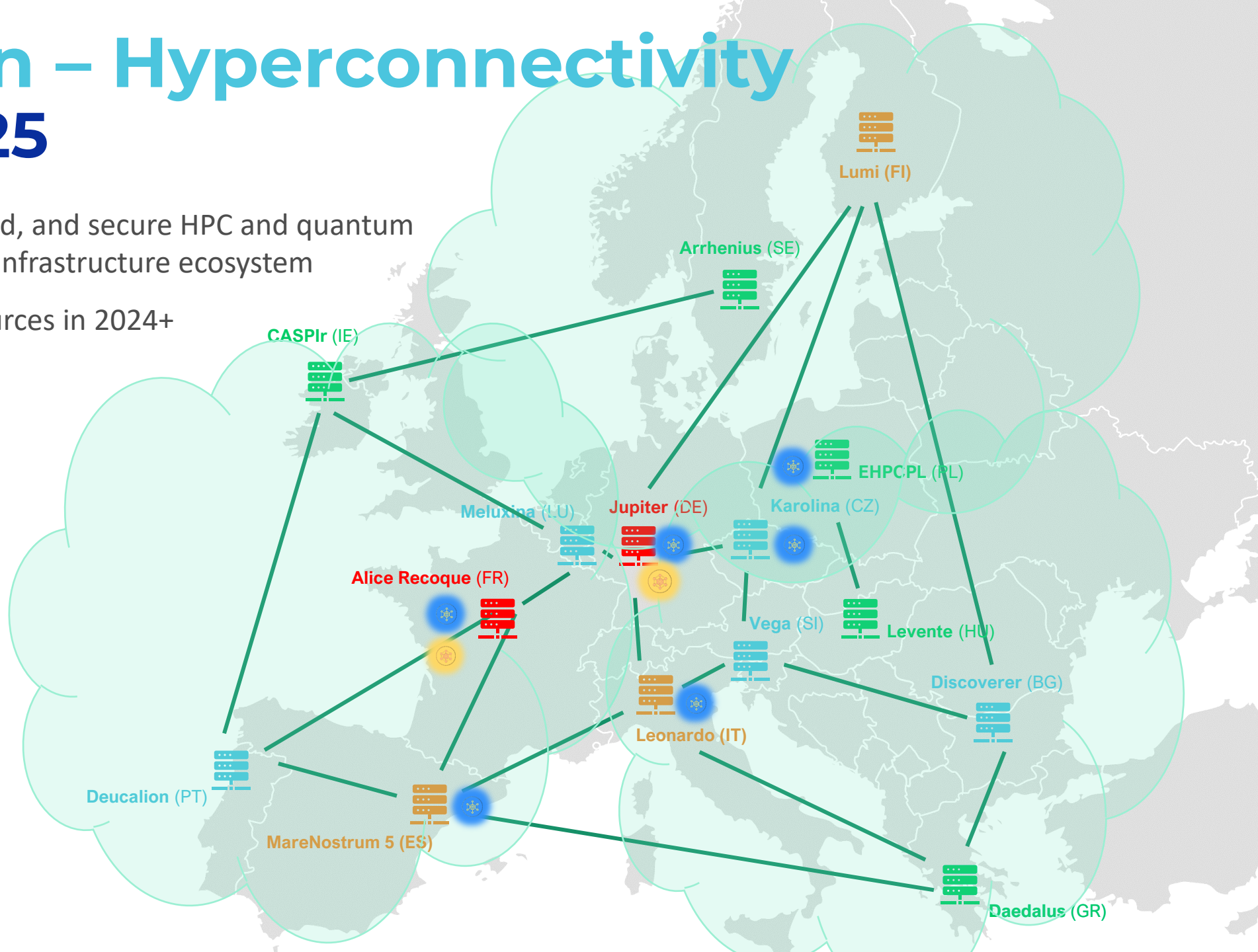
European Commission

HPC – Quantum infrastructure
2024 - 2025



Federation – Hyperconnectivity 2024+ / 2025

- A hyper-connected, federated, and secure HPC and quantum computing service and data infrastructure ecosystem
- Federation of EuroHPC resources in 2024+
- Hyperconnectivity in 2025



EXASCALE

PRE-EXASCALE

PETASCALE

MID-RANGE

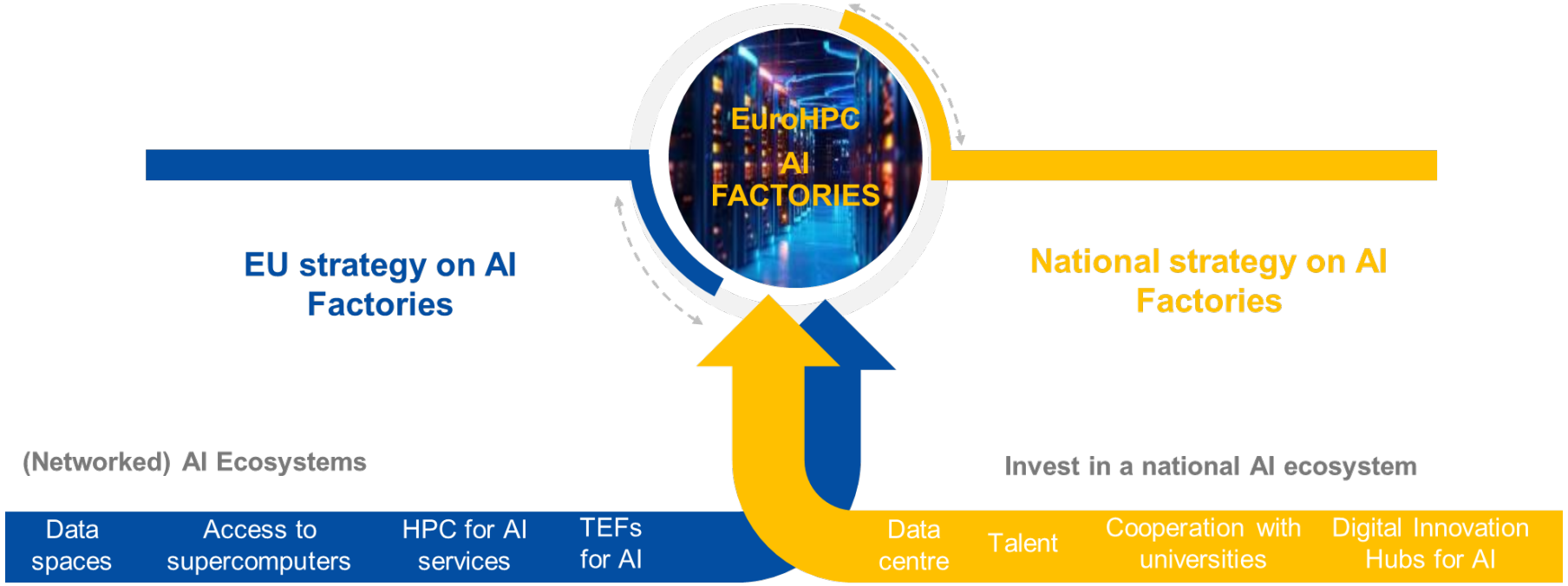
Q-computer

Q-simulatorr

EuroHPC AI Factories

GenAI and **AI foundation models** play a key role in the future of technology and society.

HPC + AI are contributing to innovative and competitive economy and improving citizens' life.



➲ AI Factories **Call for Expression of Interest** in July 2024.

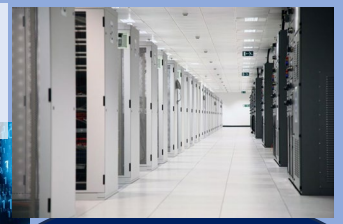
European Commission

AI Factories Ecosystem

TEFs for AI

Regulatory Sandboxes

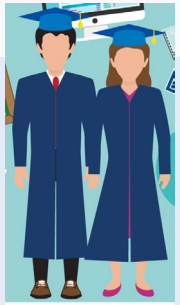
Data & Data Centre



Data Spaces

Digital Innovation Hubs for AI

Cooperation with Universities



Talent

HPC for AI Services

Start-ups

International initiatives

Drivers for the future EU HPC infrastructures?

The future of European competitiveness: Report by Mario Draghi

Pursuit of inclusive economic growth - Pillars of prosperity

- sustainable competitiveness
- economic security
- open strategic autonomy
- fair competition

Analysis: the EU is falling behind in providing state-of-the-art infrastructures for the digitalisation of the economy

Recommendations for closing the innovation gap:

A programme to tackle the innovation deficit, including

- **to lower the cost of AI deployment by increasing computational capacity and making available its network of high-performance computers**

In-depth analysis and recommendations 3.2 Computing and AI

Analysis: the EU has secured a strong international position in HPC – a unique advantage to exploit in areas such as AI, and to stimulate private investment

The EU should adopt a new 'EU Cloud and AI Development Act':

- enhancing European HPC, AI and quantum capabilities and infrastructure
- harmonising cloud architecture requirements and procurement processes
- coordinating priority initiatives to scale-up private involvement and financing

Recommendations HPC/AI/Quantum:

1. EuroHPC upgrade program
2. Launch an 'EU Vertical AI Priorities Plan'
3. Harmonise national 'AI Sandbox regimes'

EuroHPC upgrade : More money in the next MFF?

MFF 2021-2027

MFF 2028-2035

1 B

2021 2022 2023 2024 2025 2026 2027 2028 2029 2030

6 B

? B

#EuroHPC
Joint Undertaking

The European High Performance Computing Joint Undertaking (EuroHPC JU) will pool European resources to develop top-of-the range exascale supercomputers for processing big data, based on competitive European technology.

Member countries are Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Latvia, Lithuania, Luxembourg, Malta, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Türkiye and United Kingdom.

EuroHPC

“Mo money, mo problems.”

The Notorious B.I.G.

A rich and paradoxical environment for future HPC infrastructures

A rich and paradoxical environment for future HPC infrastructures

Draghi’s recommendations and possible implications for updated EuroHPC …

Regularly increase computational capacity dedicated to the training and algorithmic development of AI models in existing EU HPC centres, and for the development of tomorrow's exascale and post-exascale computing

Result of 2009 Workshop Compared to First Exascale System

			Predicted	Actual
Systems	2009	Pre-Exascale	Exascale	Frontier
System peak	2 Peta	100-200 Peta	1 Exa	> 1.5 Exa
Power	6 MW	~15 MW	~20 MW	29 MW (< 20MW per exaflop)
System memory	0.3 PB	5 PB	64 PB (+)	45 PB (9.6 PB + 36 PB NVM)
Node performance	125 GF	0.5 TF or 7 TF	1,2 or 15TF	NDA
Node memory BW	25 GB/s	1-2TB/s	2-4TB/s	3.2 TB/S
Node concurrency	12	O(100)	O(1k) or 10k	O(100K)
Total Node Interconnect BW	3.5 GB/s	100-200 GB/s 10:1 or 2:1 vs memory BW	200-400GB/s (1:4 or 1:8 from memory BW)	100 GB/s
System size (nodes)	18,700	50,000 or 500,000	O(100,000) or O(1M)	< 10,000
Total concurrency	225,000	O(100,000,000) *O(10)-O(50) to hide latency	O(billion) * O(10) to O(100) for latency hiding	O(billion)
Storage	15 PB	150 PB	500-1000 PB	716 PB
IO	0.2 TB	10 TB/s	60 TB/s	75 TB/s NVM and 9.4 TB/s HDD
MTTI	days	O(1day)	O(0.1 day)	O(0.3 day)

OAK RIDGE National Laboratory | OAK RIDGE LEADERSHIP COMPUTING FACILITY

ENERGY – DECARBONISATION - WATER

Supercomputer/System	Performance (Exaflops)	Current Power Consumption	Projected Power Consumption at 10 EF	Projected Power Consumption at 100 EF	Projected Power Consumption at 1000 EF
Frontier (Oak Ridge, USA)	1.206	22.8 MW	228 MW (≈0.23 nuclear plants)	2,280 MW (≈2.3 nuclear plants)	22,800 MW (≈23 nuclear plants)
Aurora (Argonne, USA)	1.012	38.7 MW	387 MW (≈0.39 nuclear plants)	3,870 MW (≈3.9 nuclear plants)	38,700 MW (≈39 nuclear plants)
Eagle (Microsoft Azure, USA)	0.561	15.5 MW	276 MW (≈0.28 nuclear plants)	2,760 MW (≈2.8 nuclear plants)	27,600 MW (≈28 nuclear plants)
Fugaku (RIKEN, Japan)	0.442	29.9 MW	676 MW (≈0.68 nuclear plants)	6,760 MW (≈6.8 nuclear plants)	67,600 MW (≈68 nuclear plants)
LUMI (CSC, Finland)	0.380	7.1 MW	187 MW (≈0.19 nuclear plants)	1,870 MW (≈1.9 nuclear plants)	18,700 MW (≈19 nuclear plants)

SUPPORTING INNOVATION & COMPETITIVENESS

- **Reinforcing the pipeline from innovation into commercialisation**
- **Public infrastructures & private use for commercial access**

- Expansion of EuroHPC to additional cloud and storage capabilities for AI training, fine-tuning and inference
- Additional cloud and storage capabilities distributed throughout Europe
- Validate hosting in ‘regulatory compliant’ infrastructures (start-ups)
- EU-wide framework (inc. state aid rules) enabling public **‘computing capital’** for EU innovative SMEs in exchange of financial returns (e.g. equity options, royalties or dividends) for reinvestments in infrastructure

Access / Resource provisioning

Open up EuroHPC to a **‘federated AI model’** favouring cooperation of public-private infrastructure to provide AI training power, leveraging the joint capacity of public computing and private resources and increasing the EU’s competitive scale

How to move from AI Factories to “federated AI model” ?

Hybrid HPC/Quantum

Develop quantum labs or nodes attached to all EU HPC centres and launch public-private partnerships – involving large EU tech leaders as a priority – to co-invest in the whole frontier tech stack, including neuromorphic and quantum chips.

…. and more

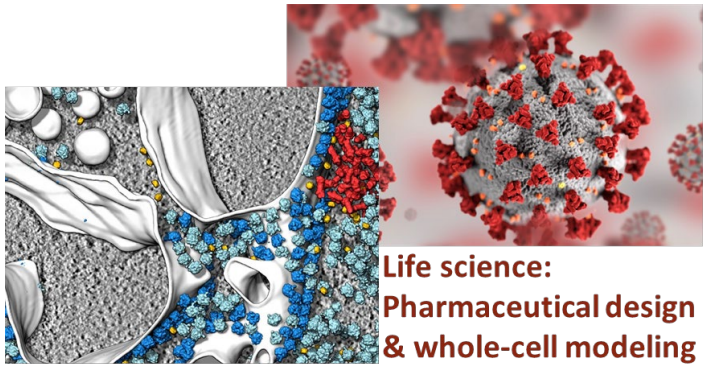
Data – applications - infrastructure

Data

Data infrastructures (AI models, data bridges, data lakes, European data spaces) !!

Applications

- Digital twins / Edge – cloud – digital continuum
- Public administration
- Manhattan-scale projects to solve societal problems

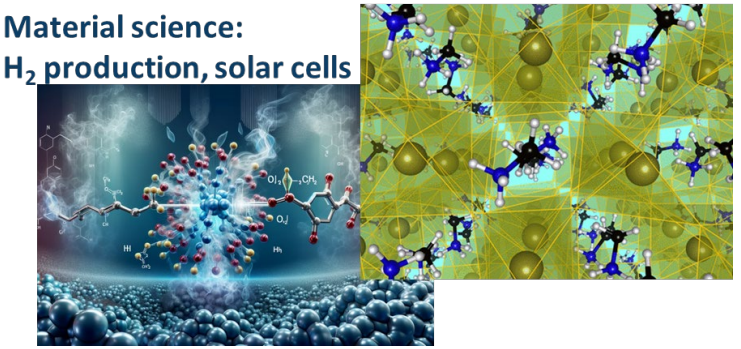


Life science:
Pharmaceutical design
& whole-cell modeling

Infrastructure

- AI vs “Traditional” HPC
- Dedicated infrastructures / flexible configuration of supercomputer resources?

Material science:
H_2 production, solar cells



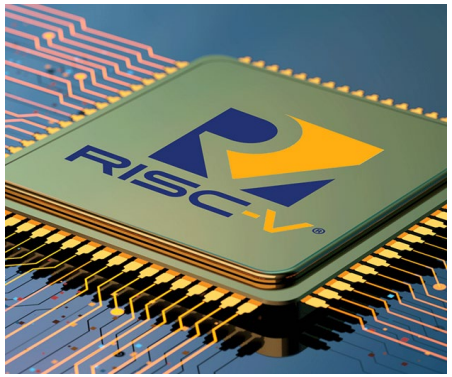
Defence / Security / Sovereignty

Defence & Cyber-security

- Support to dual use (EU Structural funds)
- Misuse of HPC infrastructures - disinformation, virus/weapons/terrorism… legal implications ?
- Security by Design and Zero Trust vs and open access to internet, relaxed authentication/authorization
- Location of critical infrastructure?

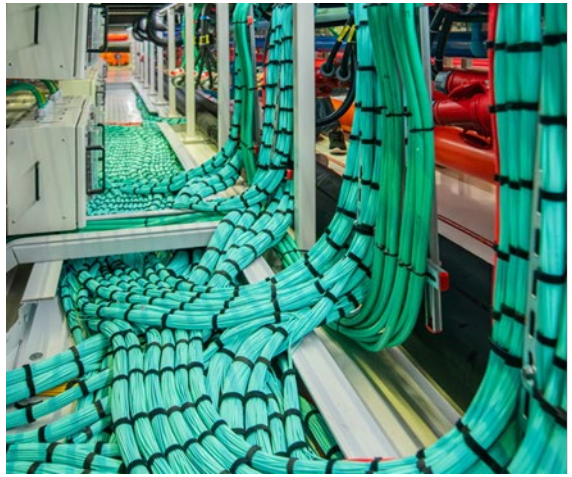
Sovereignty

- Risks of supply chain
- Geopolitics (control of fab facilities, export control etc.)
- Sovereign clouds (?)
- Technological development



Skills - ecosystems

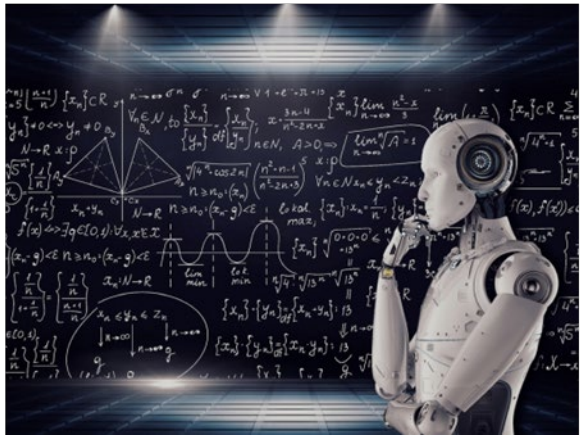
1. How to support the growth of an ecosystem ?
2. A great variety of new skills – infrastructure, computing, applications, civil engineering, legal…



International Cooperation

- Global challenges by collaboration vs EU/national interests (use, data, …) – going beyond roadmaps ?
- International cooperation with developing countries: Access… skills, applications, services, etc.

AI4Science – Science4AI



Technology - Architectures

- Heterogeneous and emerging architectures, Quantum, Neuromorphic …
- Incredibly flexible mechanisms for (re)configuration, virtualisation, federation of resources, resource allocation, monitoring…
- Optimisation tools (traces ☺)
- Interoperability and portability
- Software stack
- Workflows
- Software is the key? Algorithms ?
- Moore’s law

All is AI / AI is all (?)

Conclusions - personal

HPC Infrastructures

- HW/SW/Services/People/Policies, etc. cannot be addressed in isolation!

Users

- Representativity and structures
- Facing the new environments

Role of EU Supercomputing Centres (SC)

- Excellence poles
- Networking of AIF
- Actively contributing to the future of EU HPC infrastructures

What is the European model for future infrastructures?

Preparation in 2025?

THANK YOU!