

Large Language Model

Souveraineté et supply chain

François PAUPIER • Septembre 2025

Agenda

- * Partie 1 - Supply chain hardware: composants, acteurs et enjeux géopolitiques
- * Partie 2 - Supply chain model - qui développe les architectures SOTA ?
- * Partie 3 - **(TP)** - Supply chain data - quelle représentation du monde exposée au modèle

Qui suis-je ?

- ♦ ECN Promo 2018
- ♦ Option DataSim
- ♦ Club Raid

♦ Data engineer / scientist

♦ Grands groupes 

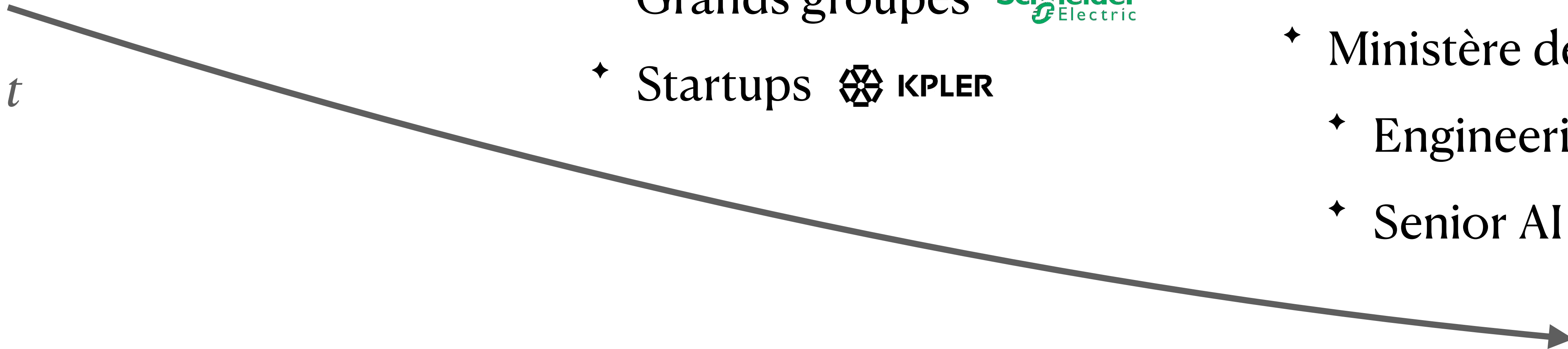
♦ Startups  KPLER

♦ Ministère des Armées 

♦ Engineering manager

♦ Senior AI advisor

t



Hypothèses

Sont considérés comme connus

- ✓ Bases de machine learning et deep learning
- ✓ Niveau python intermédiaire
- ✓ Commandes UNIX de base
- ✓ Accès à VM via SSH (TP)

Hors périmètre

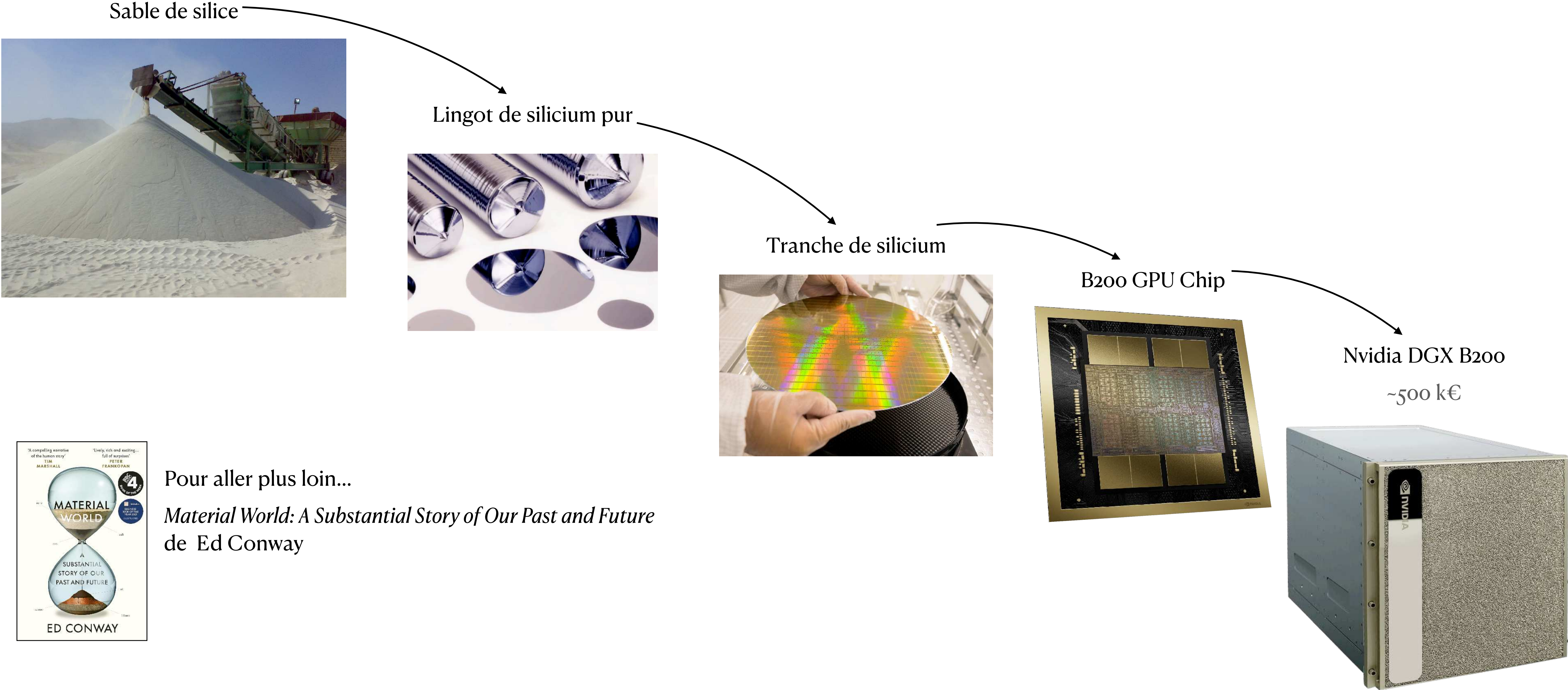
Ne sont pas couverts dans ce cours

- Introduction au LLM
- Fonctionnement interne d'un LLM (voir {Vaswani et al., 2017})
- Computer vision / Audio / Signal processing
- Agent IA
- Droit d'auteur, propriété intellectuelle

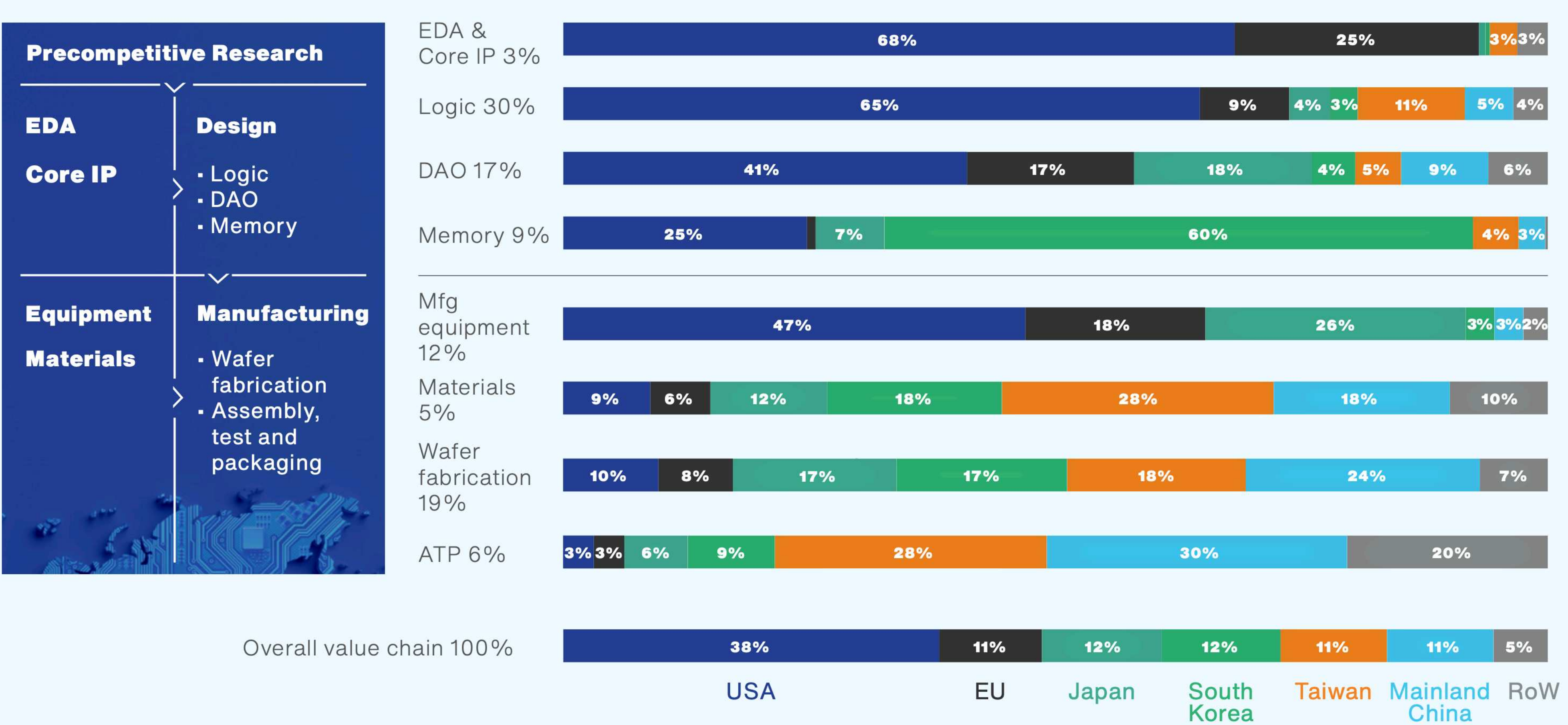
Partie 1 - Supply chain hardware

Composants, acteurs et enjeux géopolitiques

Du sable vers un GPU, quelques transformations clés

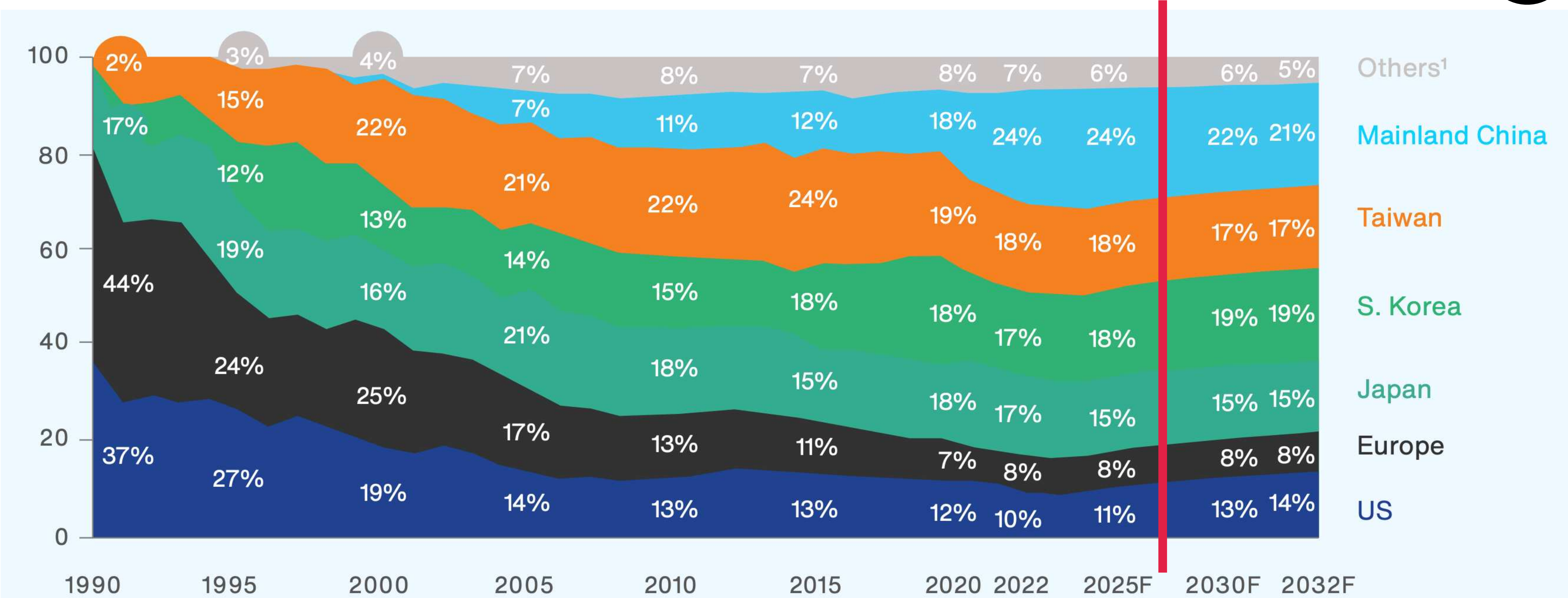


Plus-value par activité et par région dans l'industrie des semi conducteurs



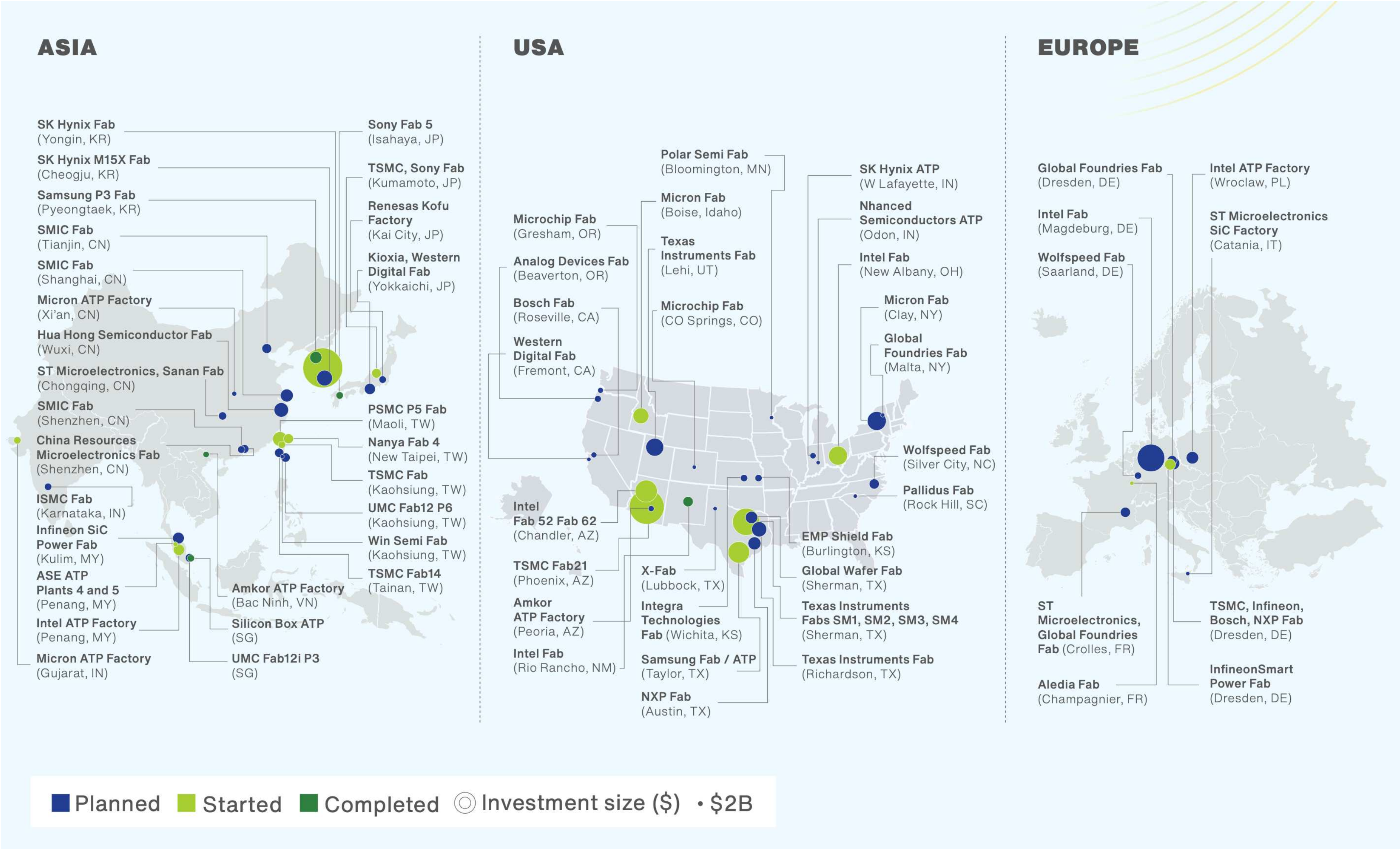
Notes on regional breakdown: EDA, design, manufacturing equipment, and raw materials based on company revenues and company headquarters location. Wafer fabrication and Assembly & testing based on installed capacity and geographic location of the facilities.
1. Includes Israel, Singapore, and the rest of the world
Source: IPnest; Wolfe Research; Gartner; SEMI; BCG analysis

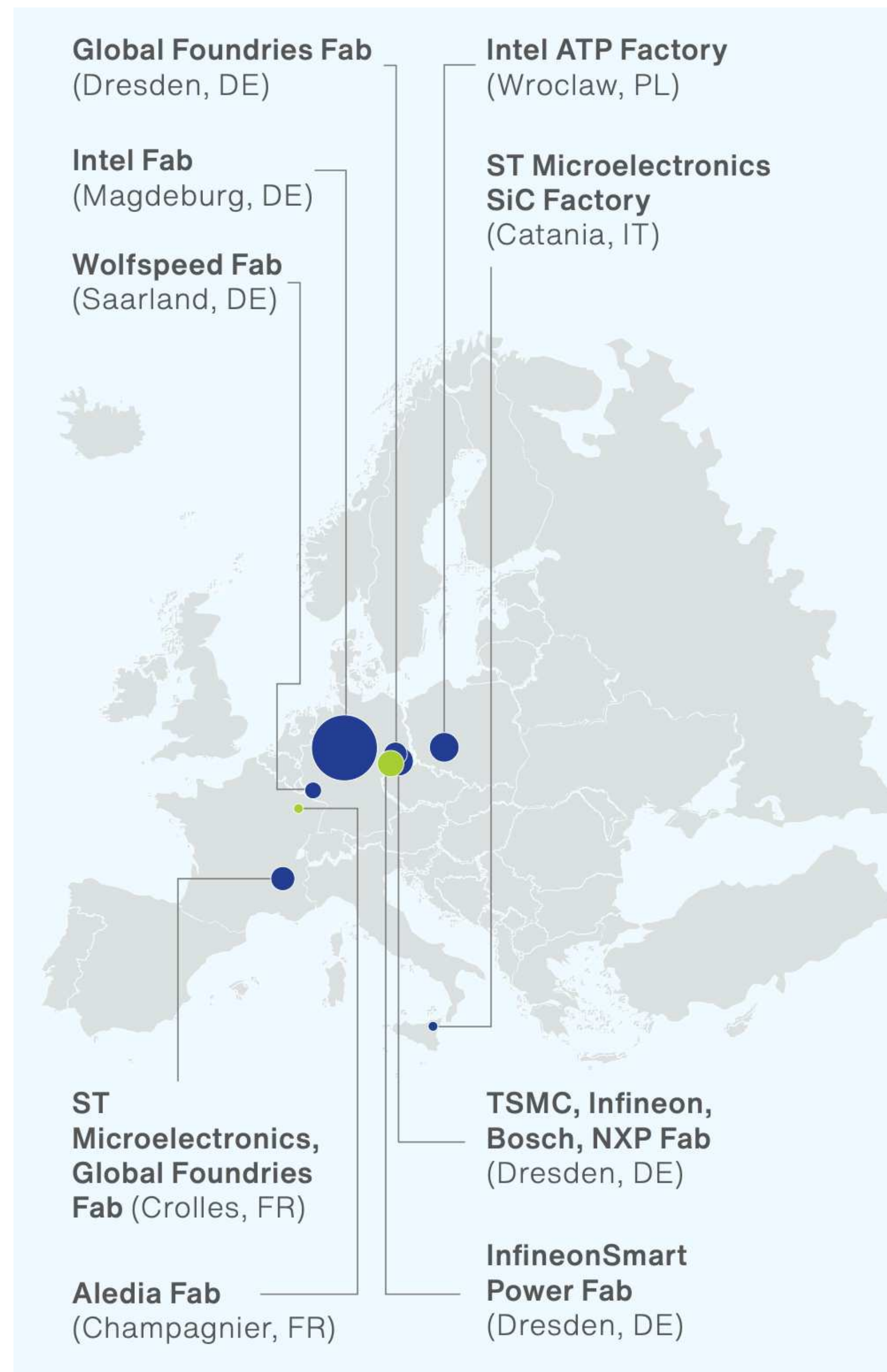
Evolution des capacités des fab à l'échelle mondiale



1. Others includes Malaysia, Singapore, India, and the rest of the world
Note 1: Rounding errors. All values shown in 300 mm (12") equivalents; excludes capacity below 5K wspm or produced on wafer sizes less than 8"
Note 2: May not total 100% due to rounding.
Source: SEMI; BCG analysis

Principaux clusters d'activités semi-conducteurs à l'échelle mondiale





Principaux clusters d'activité zone Euro

- Principalement des fab pour design des chips
- Peu d'acteur de droits européens

Fusion Franco/Italienne



ETI Française innovante



ASML

Leader européen sur l'industrie du semi conducteur

@fpaupier



Des positionnements variés sur la chaine de valeur...

Pt 1

...Menant à des écarts forts de valorisation



- Market Cap 25B
- Euronext Paris

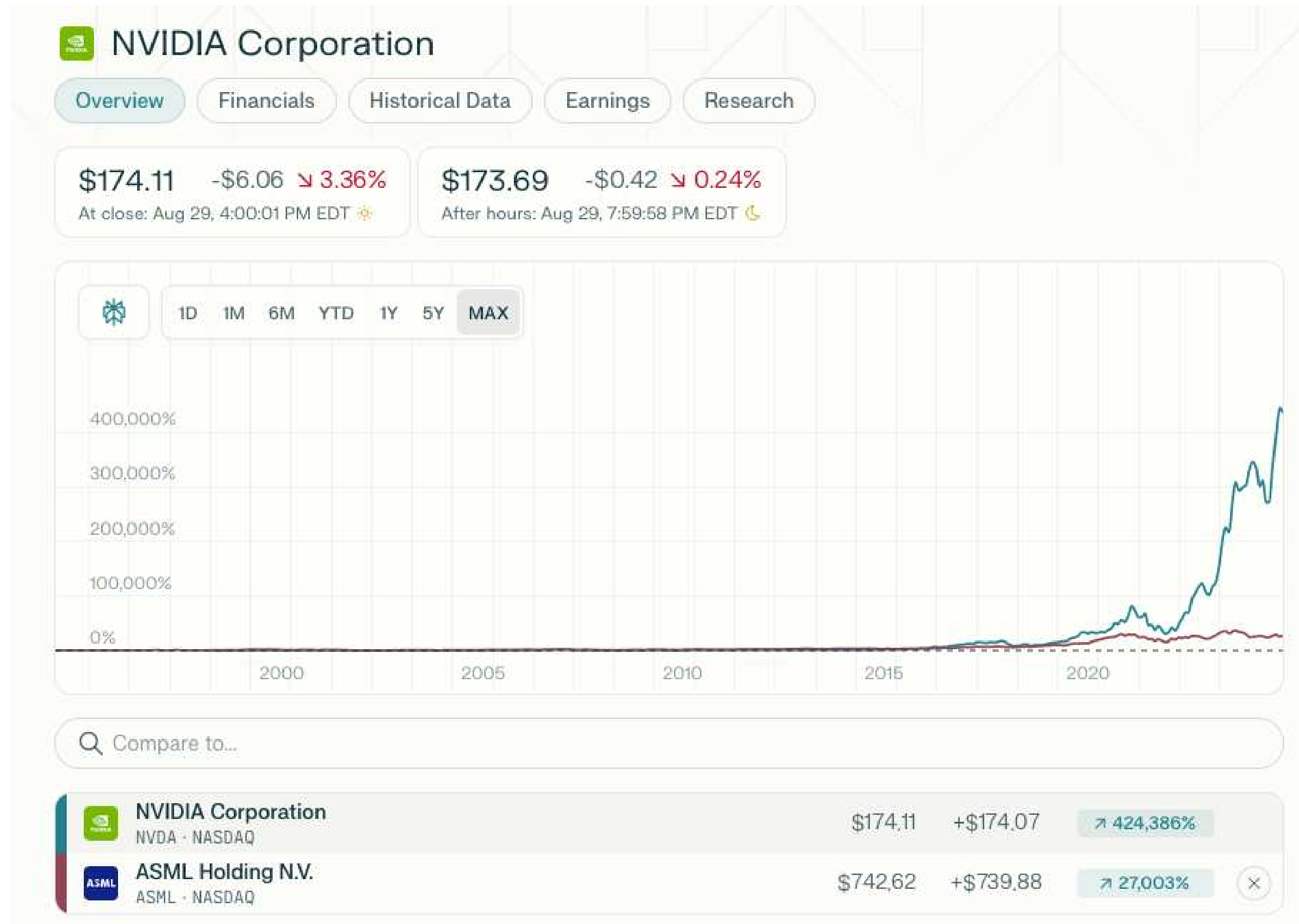


- Market Cap 300B
- Euronext Amsterdam

ASML

Les géants européens loin derrière les américains

Pt 1

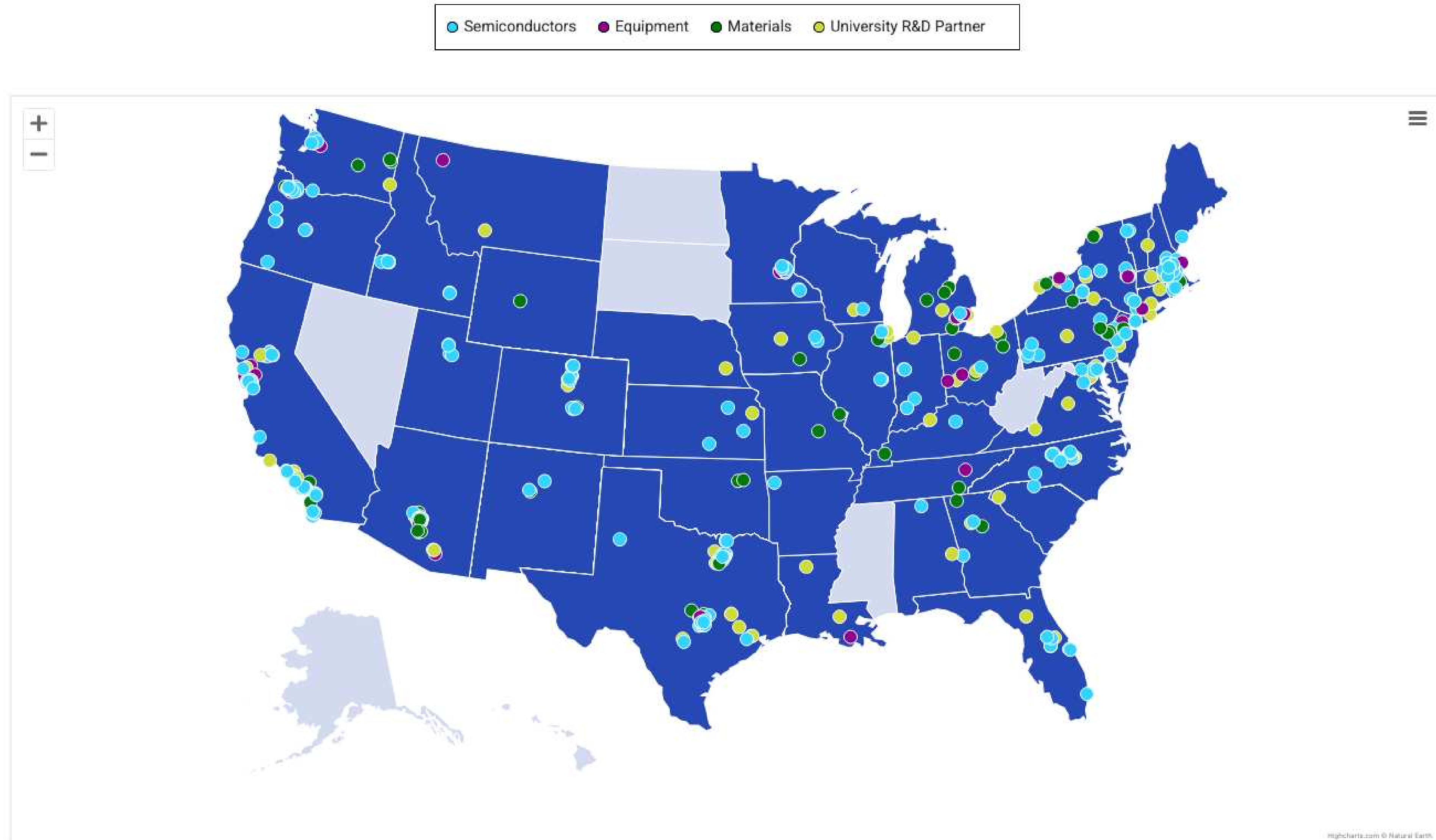


nVIDIA®

- Market Cap 4 250B
- NASDAQ

L'écosystème des semi conducteurs aux US

Pt 1



ZOOM sur

La lithographie

La Caricature morale, politique et littéraire / A. Audibert, rédacteur en chef ; [directeur : Ch. Philipon].
(1840, October 11). Retrieved from <https://gallica.bnf.fr/ark:/12148/bpt6k10517164/f5.item>



(L'ATELIER DU LITHOGRAPHE.)

Comme c'est léger!

Chez Beaupré & Co. Editeurs des dessins de la CARICATURE, du FIGARO et du CHARIVARI, Rue du Croissant, N° 16.

Semiconductor manufacturing equipment vendors, by HQ region revenue

Pt 1

Global geographical distribution of semiconductor manufacturing equipment¹

Wafer fabrication process

Lithography

US

EU

CN²

KR

JP

Others

Market size (\$B)

1%

89%

10%

\$17.5

Photoresist processing

3%

5%

92%

\$3.7

Ion implant & doping eqpt

83%

17%

\$2.5

Thermal processes

57%

4%

2%

1%

37%

\$2.9

Deposition

67%

11%

2%

5%

15%

\$22.8

Material removal & cleaning

56%

0%

5%

4%

35%

0%

\$30.5

Manufacturing automation

5%

11%

64%

21%

\$4.2

Process control

76%

7%

11%

6%

\$13.5

Other

1%

8%

5%

43%

42%

\$3.4

1-5%

6-10%

11-25%

26-50%

51-75%

75-100%

Top 3

1. Geographies based on company HQ's; distribution based on company revenues 2. Mainland China
Source: Gartner; BCG analysis

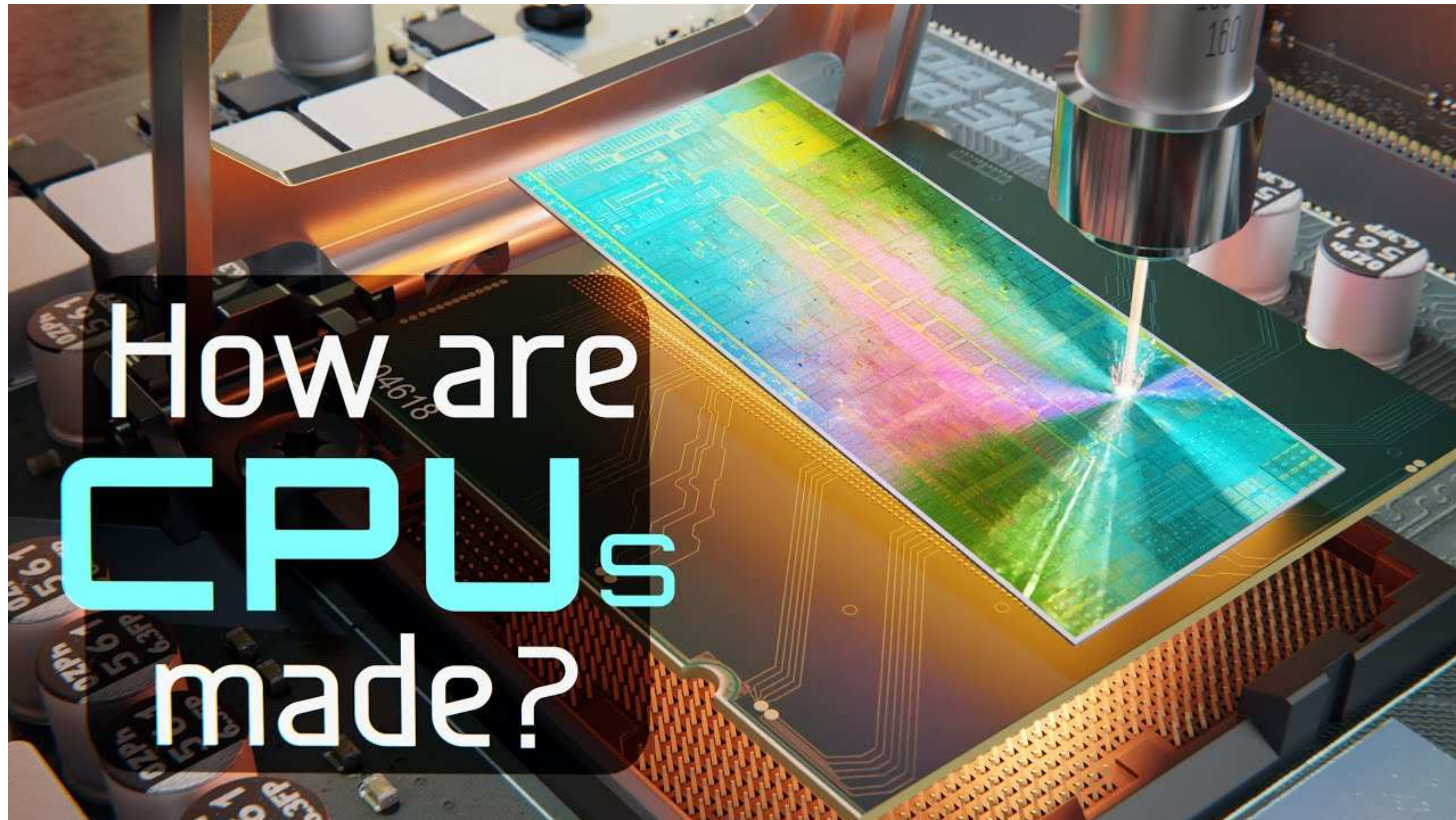
De quoi s'agit-il ?

Pt 1



Comment fonctionne la lithographie UV

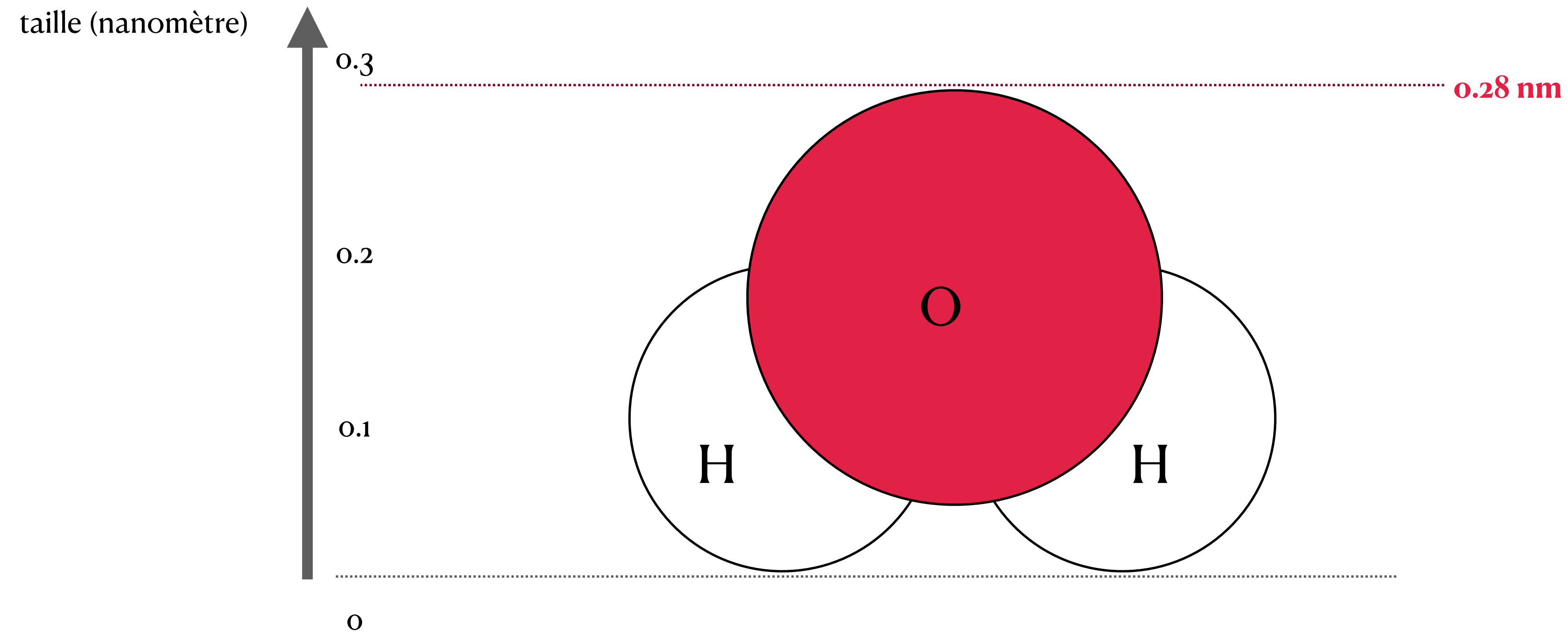
Pt 1



Source: Branch Education, [url](#)

Une molécule d'eau H_2O ~ 0.3 nm

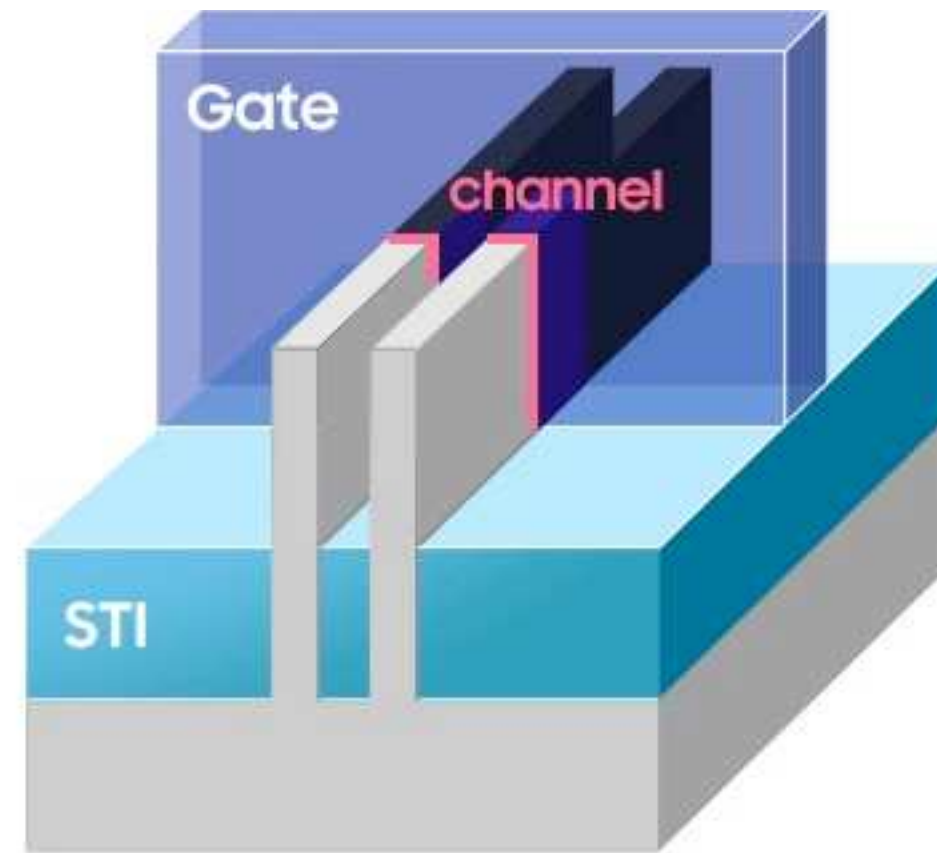
Pt 1



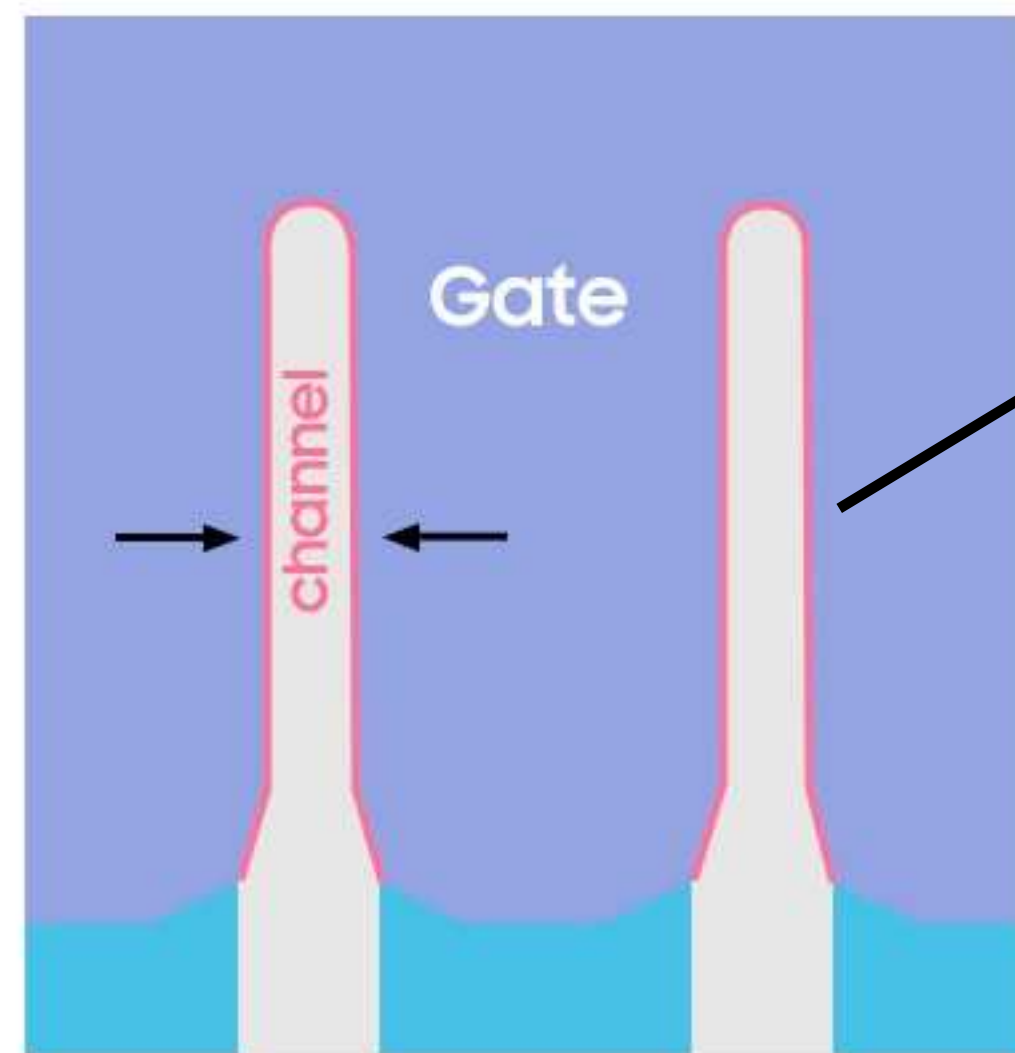
3nm d'ordre de grandeur de détails

Pt 1

taille (nanomètre)



FinFET



3 nm

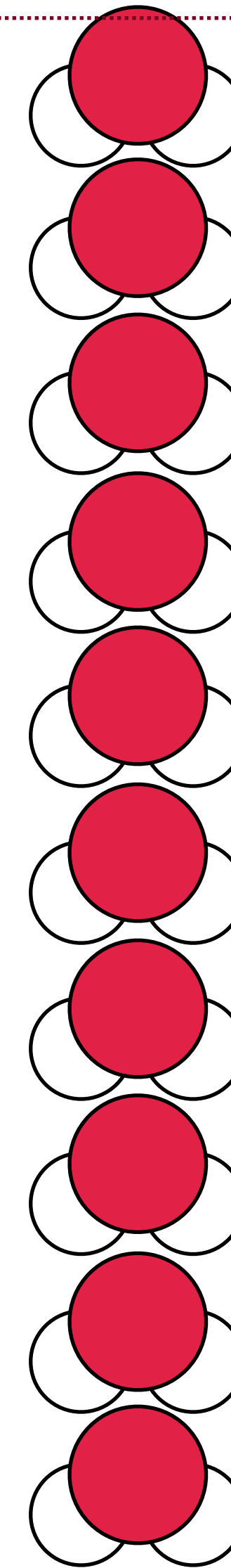
2 nm

1 nm

channel

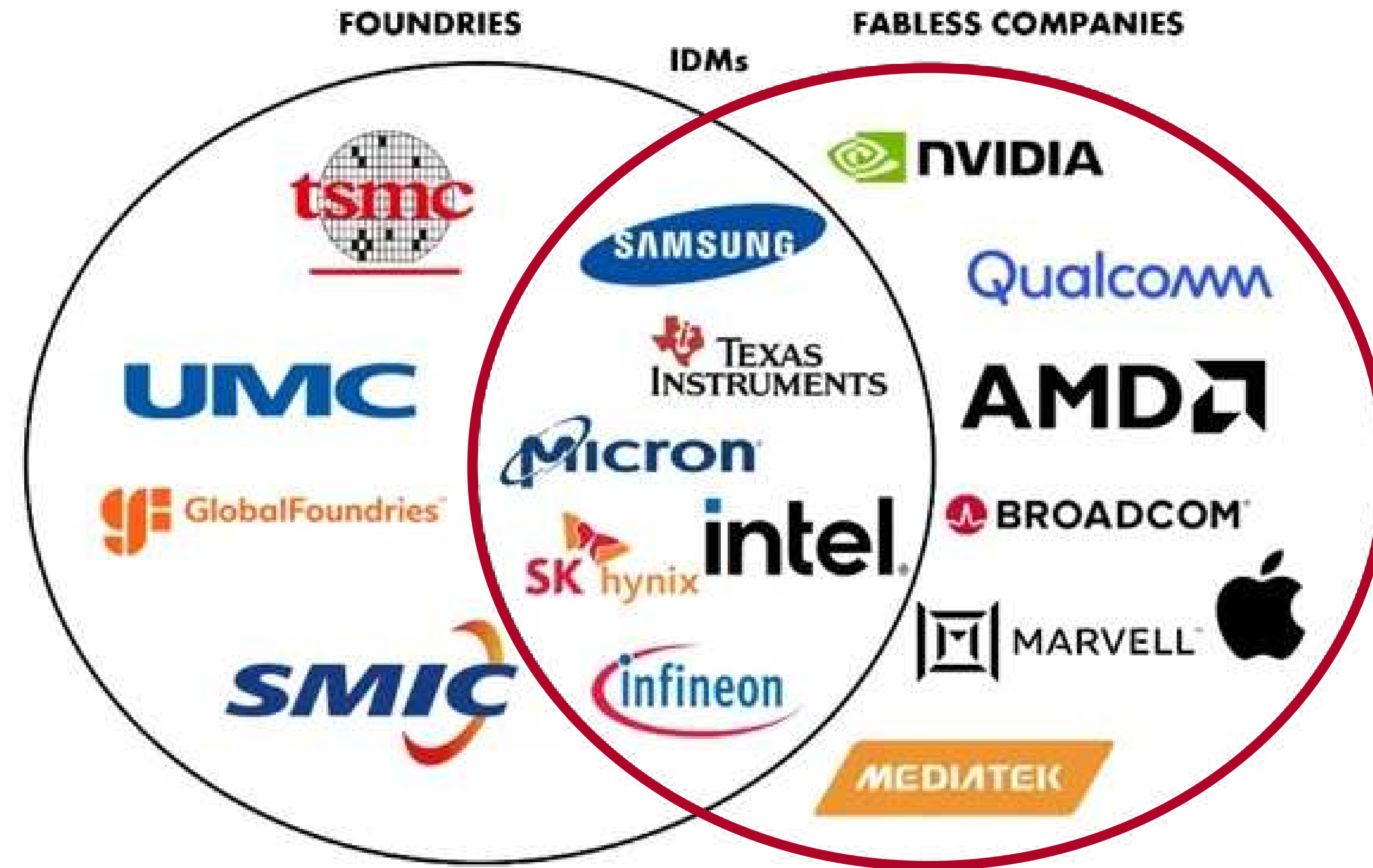
0

20



Zoom sur

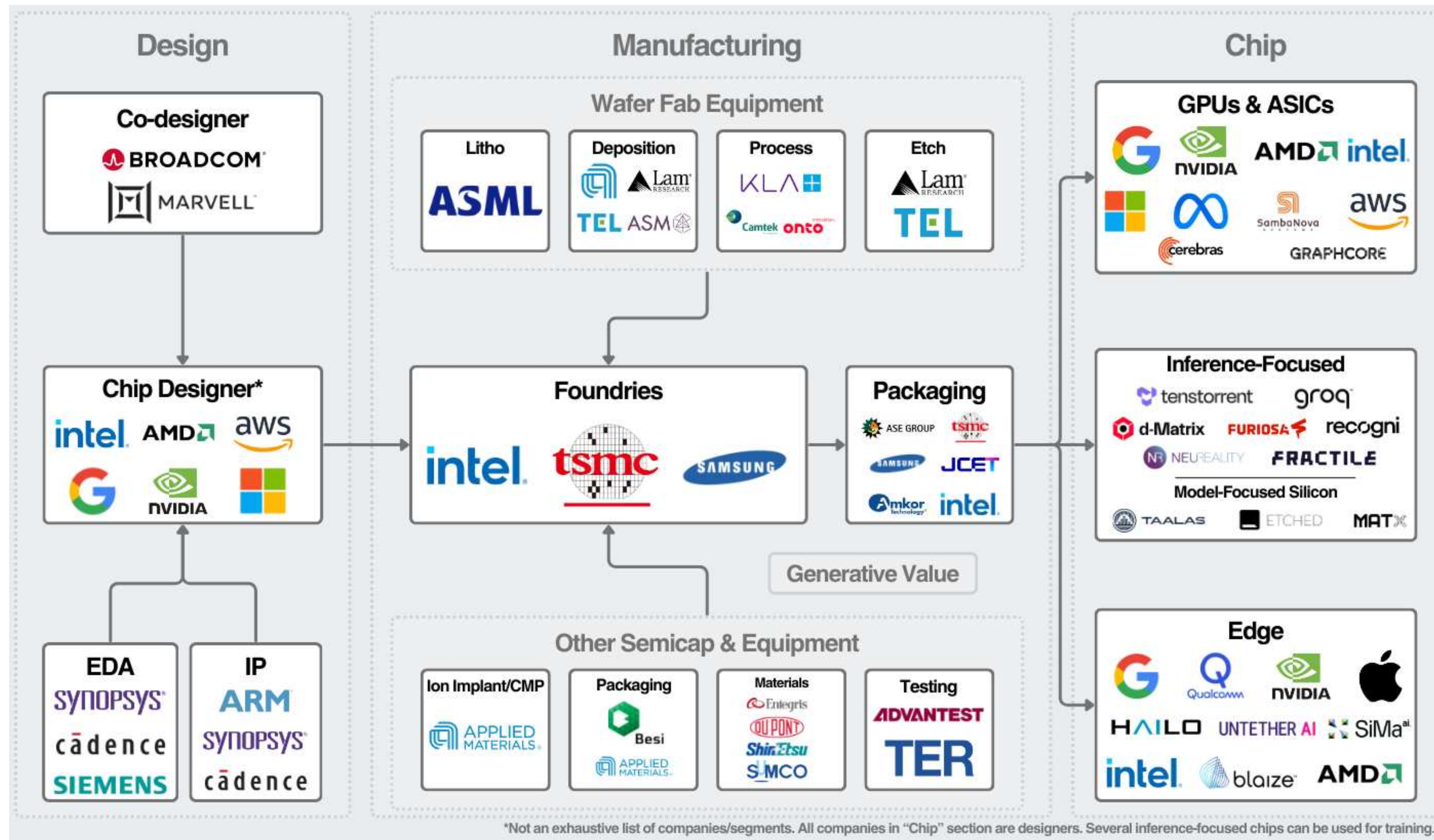
Le modèle “*Fabless*”



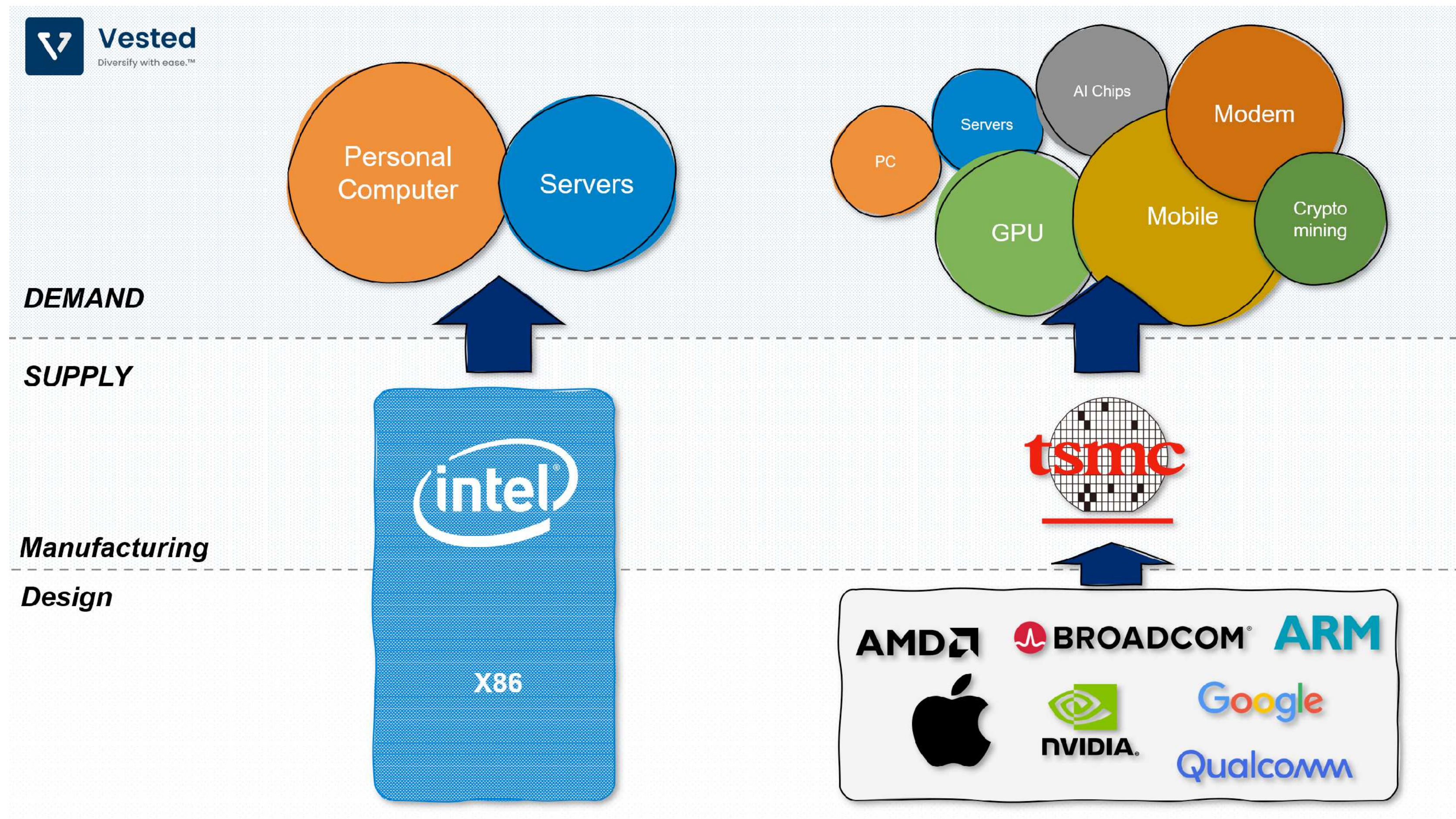
IDM: Integrated Device Manufacturer

Chaine de valeur des semi conducteurs

Pt 1

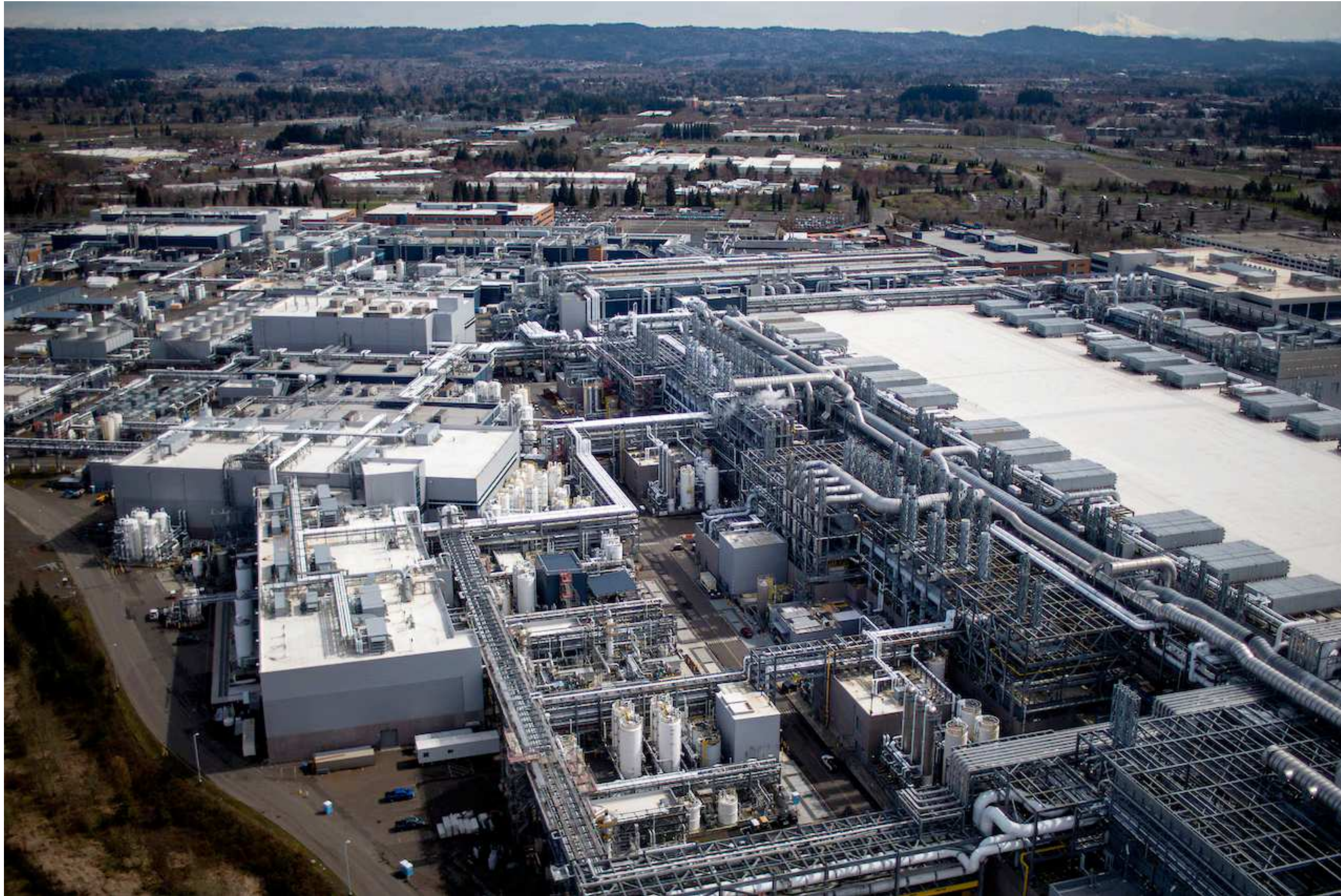


Le Fabless : Un modèle économique qui limite le besoin de capital



Intel: un des rares modèles intégrés

Pt 1

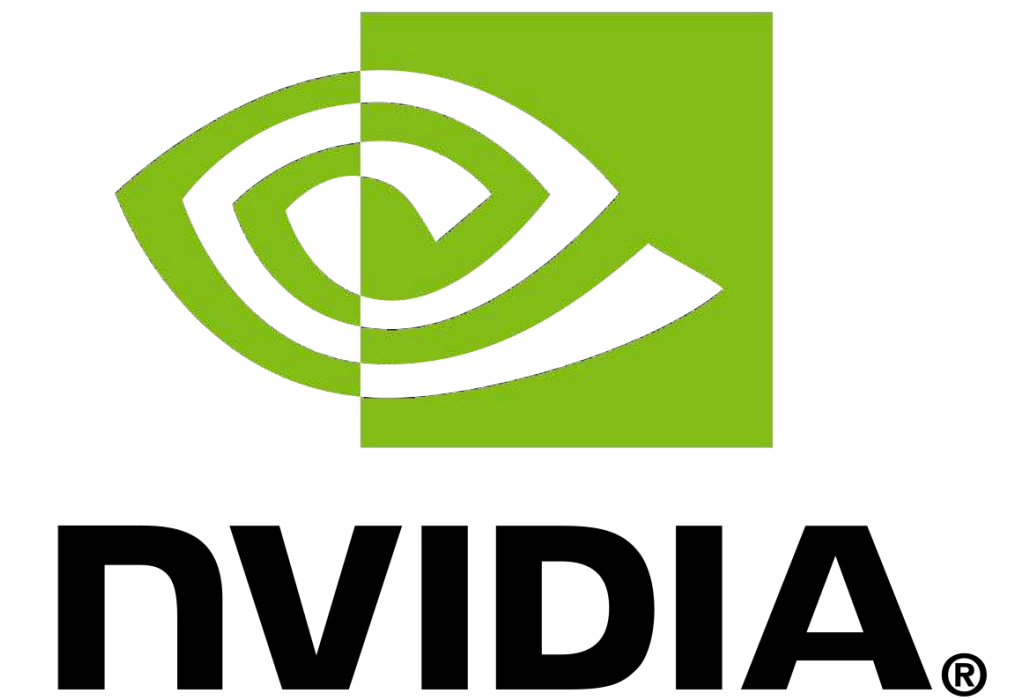


Uber

Has no cars



Has no house



Has no fab

Uber

Has no cars



Has no house



Has no fab

Uber

Has no cars



Has no house



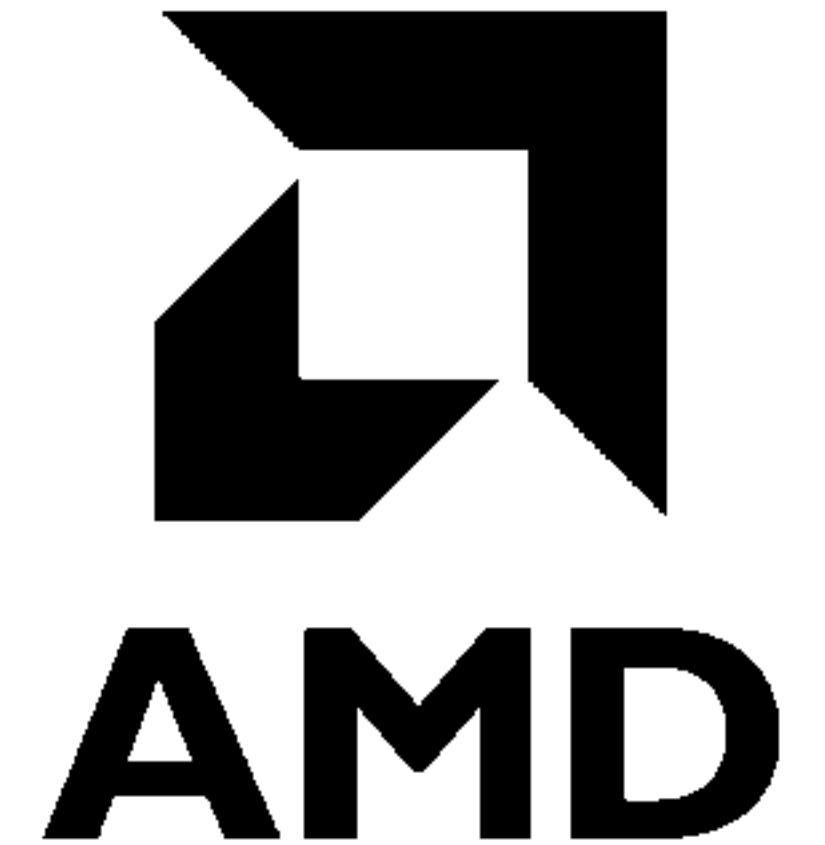
Has no fab

Uber

Has no cars



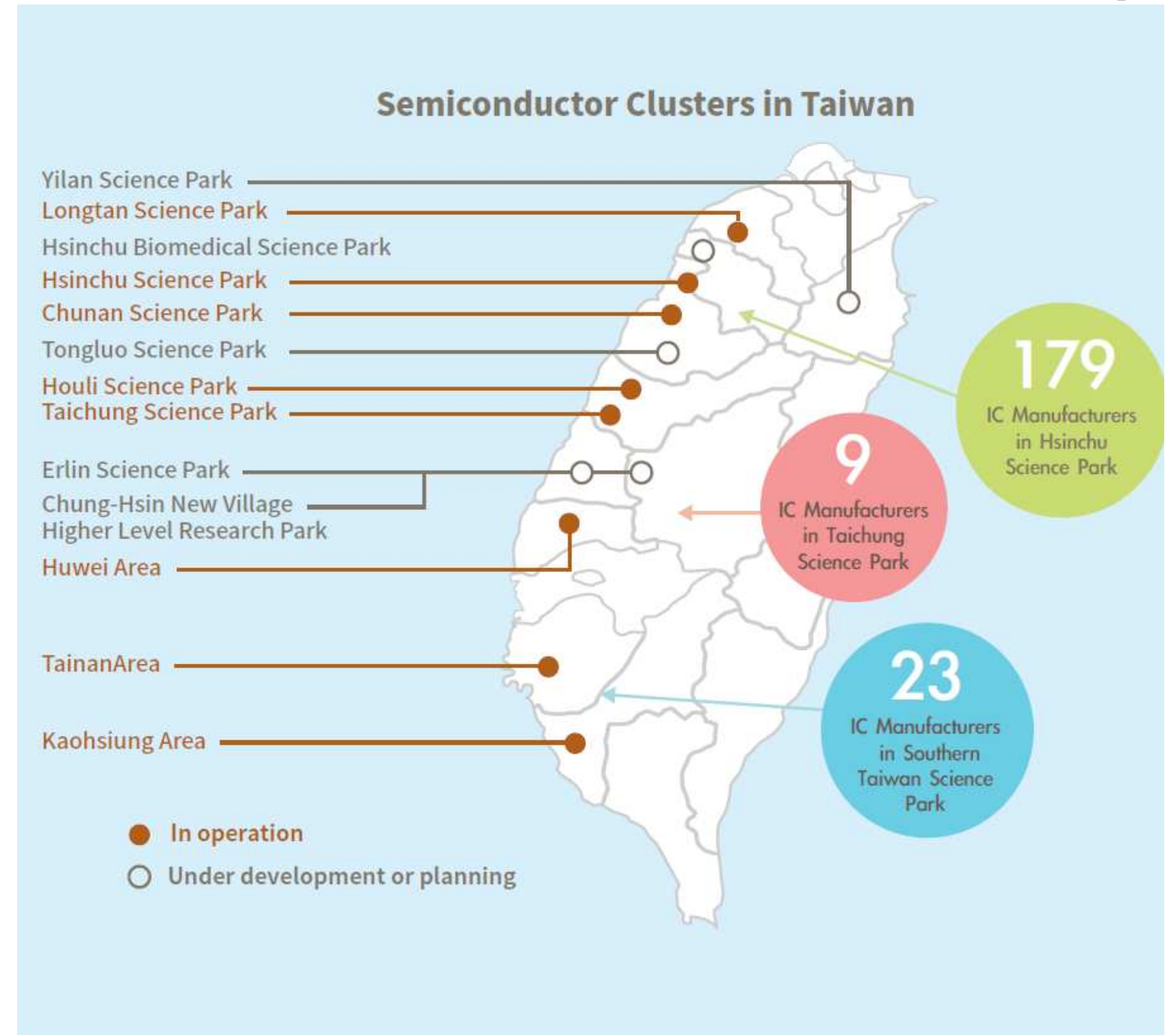
Has no house



Has no fab

ZOOM sur

Le cas Taïwan



Une zone indopacifique sous tension

Pt 1



4.3.6. Renforcer les échanges sur les technologies numériques et l'IA

La France défend une intelligence artificielle (IA) mise au service de l'intérêt général, du progrès économique, social et environnemental. À ce titre, elle a coprésidé avec l'Inde le Sommet pour l'Action sur l'IA organisé à Paris en février 2025 et soutient l'organisation d'un Sommet en Inde en 2026. Les deux pays ont lancé une Feuille de route conjointe pour le développement d'une IA sûre, ouverte, sécurisée et digne de confiance.

La France mène des discussions avancées pour intégrer ses partenaires de la région à la fondation Current AI, dédiée à l'IA d'intérêt général, avec l'objectif de remédier aux inégalités d'accès. Les investissements réunis lors du Sommet de l'IA (109 milliards d'euros) permettront notamment d'implanter des data centers souverains dans l'Indopacifique.

La France s'affirme comme un acteur de référence en matière de technologies numériques et d'innovation. Elle promeut la convergence des écosystèmes quantiques de recherche et veille à ce que les intérêts européens et indopacifiques soient représentés dans les forums comme le Groupe de développement quantique (Allemagne, Australie, Corée, Danemark, Etats-Unis, Finlande, France,

7. Antananarivo, Phnom Penh, Vientiane, Ho-Chi-Minh Ville, Nha Trang, Shanghai, Hong-Kong, Séoul, Nouméa. Le lancement de l'Institut Pasteur du Japon est en cours.

8. Pondichéry, New Dehli, Hong Kong, Bangkok, Tokyo.

Développer des partenariats de souveraineté • 29

Japon, Royaume-Uni). Elle invite ses partenaires à participer à la définition de futures normes en matière de technologies quantiques. Elle s'engage, dans le domaine des semi-conducteurs, en faveur de chaînes d'approvisionnement résilientes en s'appuyant sur la diversification des sources, la coopération et les normes communes.



Source: France Diplomatie, [url](#)

Une zone indopacifique sous tension

La France réaffirme sa politique constante d'une seule Chine et exprime son attachement à la préservation de la paix et de la stabilité dans le détroit de Taïwan. Elle s'oppose à toute modification unilatérale du statu quo, y compris par le recours ou la menace de recours à la force ou la coercition, et appelle à une résolution pacifique des différends. En mers de Chine méridionale et orientale, la France ne prend pas partie sur les litiges territoriaux, mais appelle au respect du droit international tel que défini par la CNUDM. Elle s'oppose à tout recours à la force, ainsi qu'à toute forme de coercition ou d'action déstabilisatrice, y compris la militarisation d'éléments contestés, et affirme la nécessité d'adopter un comportement sûr et professionnel dans tous les domaines du droit maritime.



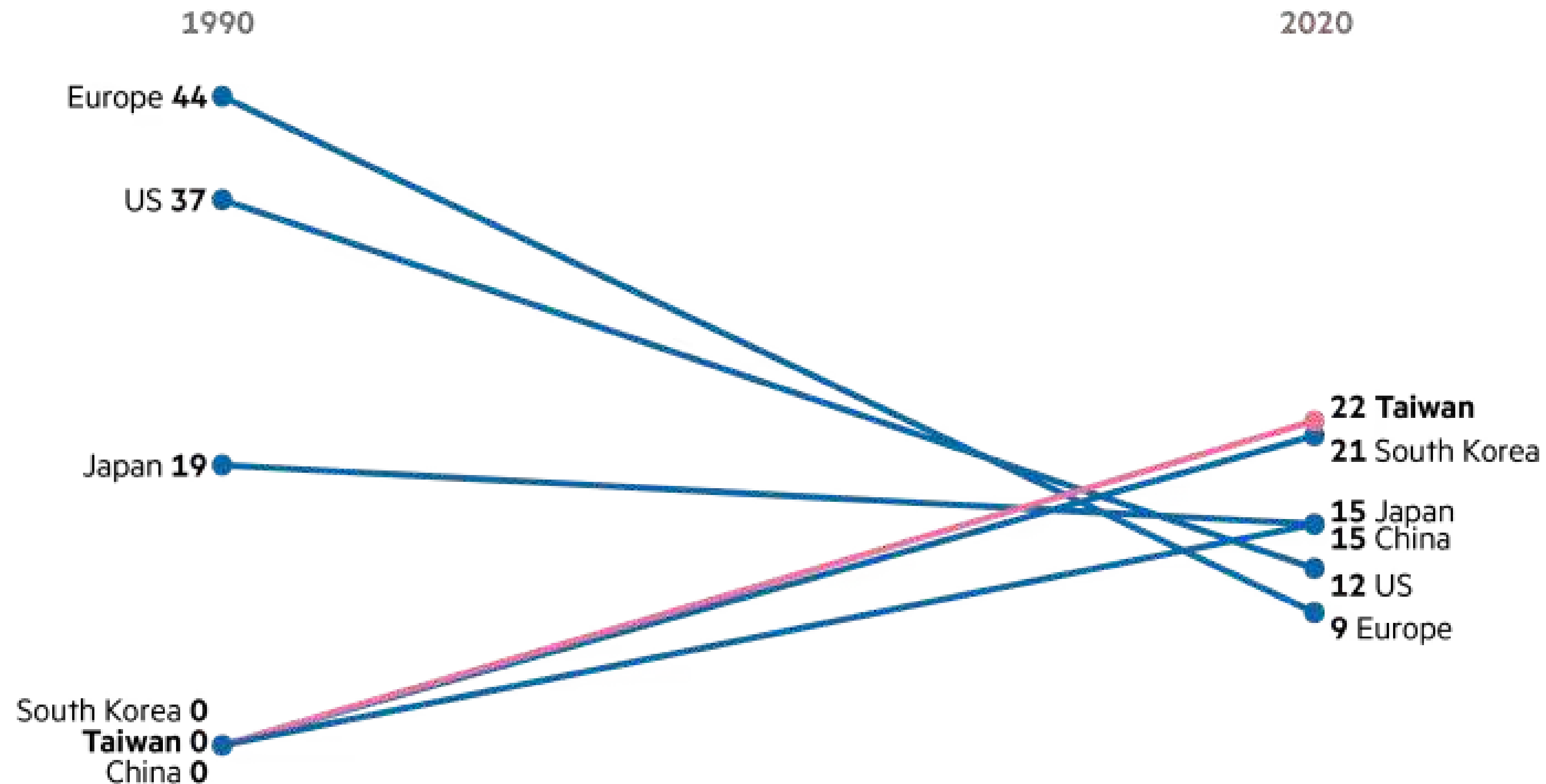
Source: France Diplomatie, [url](#)

Taiwan, ascension fulgurante comme usine du monde de chips

Pt 1

Taiwan develops into the leader of the semiconductor industry

Semiconductor manufacturing market share (%)

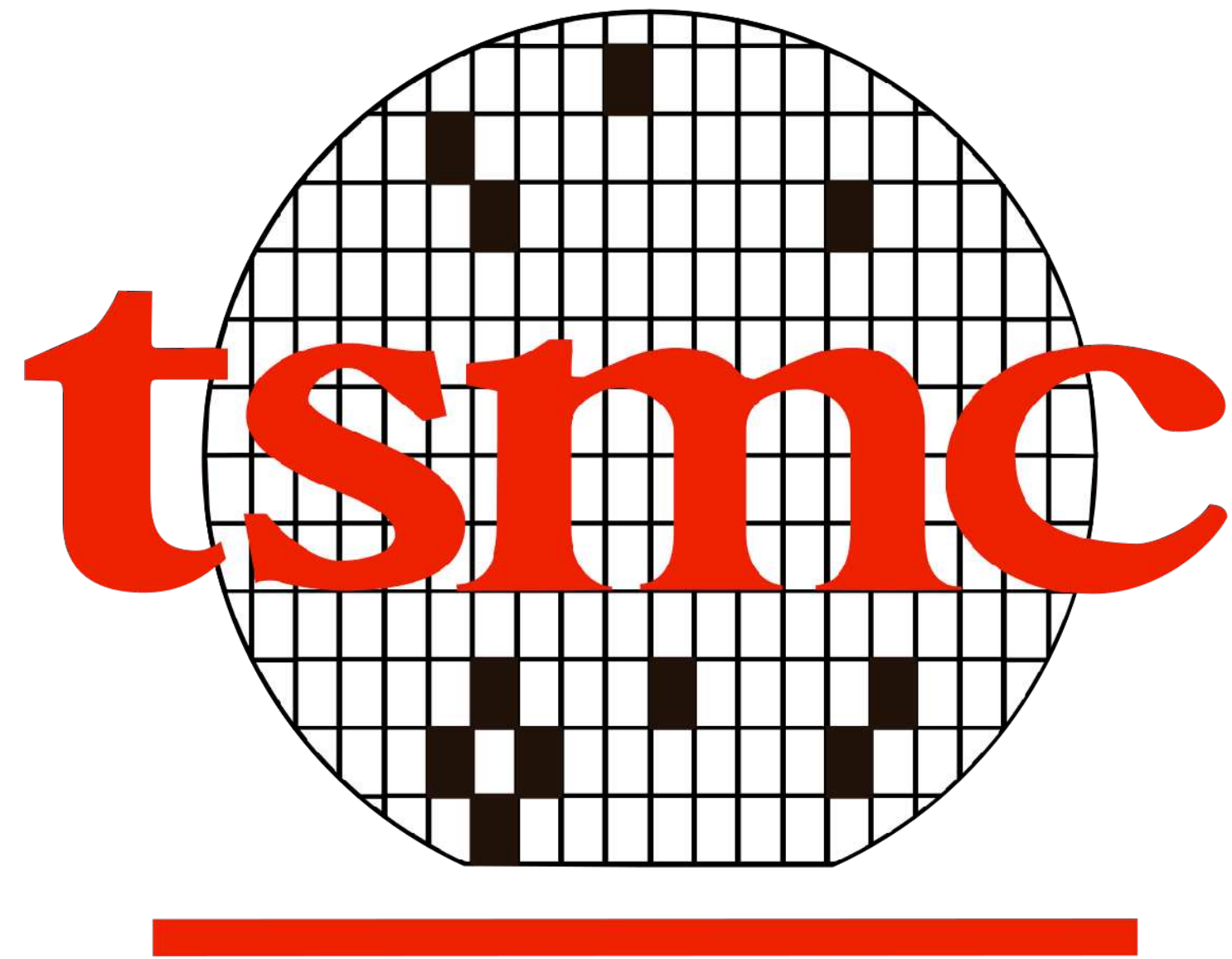


Sources: Credit Suisse; VLSI Research; SEMI Q2 2020 database; BCG analysis

© FT

Le *Silicon Shield* de Taiwan

Pt 1



Taiwan Semi Conductor Manufacturing Company

Morris Chang
Fondateur de TSMC



TSMC Global research and development Center in Hsinchu, Taiwan

CHINA'S WAR GAMES AROUND TAIWAN

STRAIT THUNDER-2025A

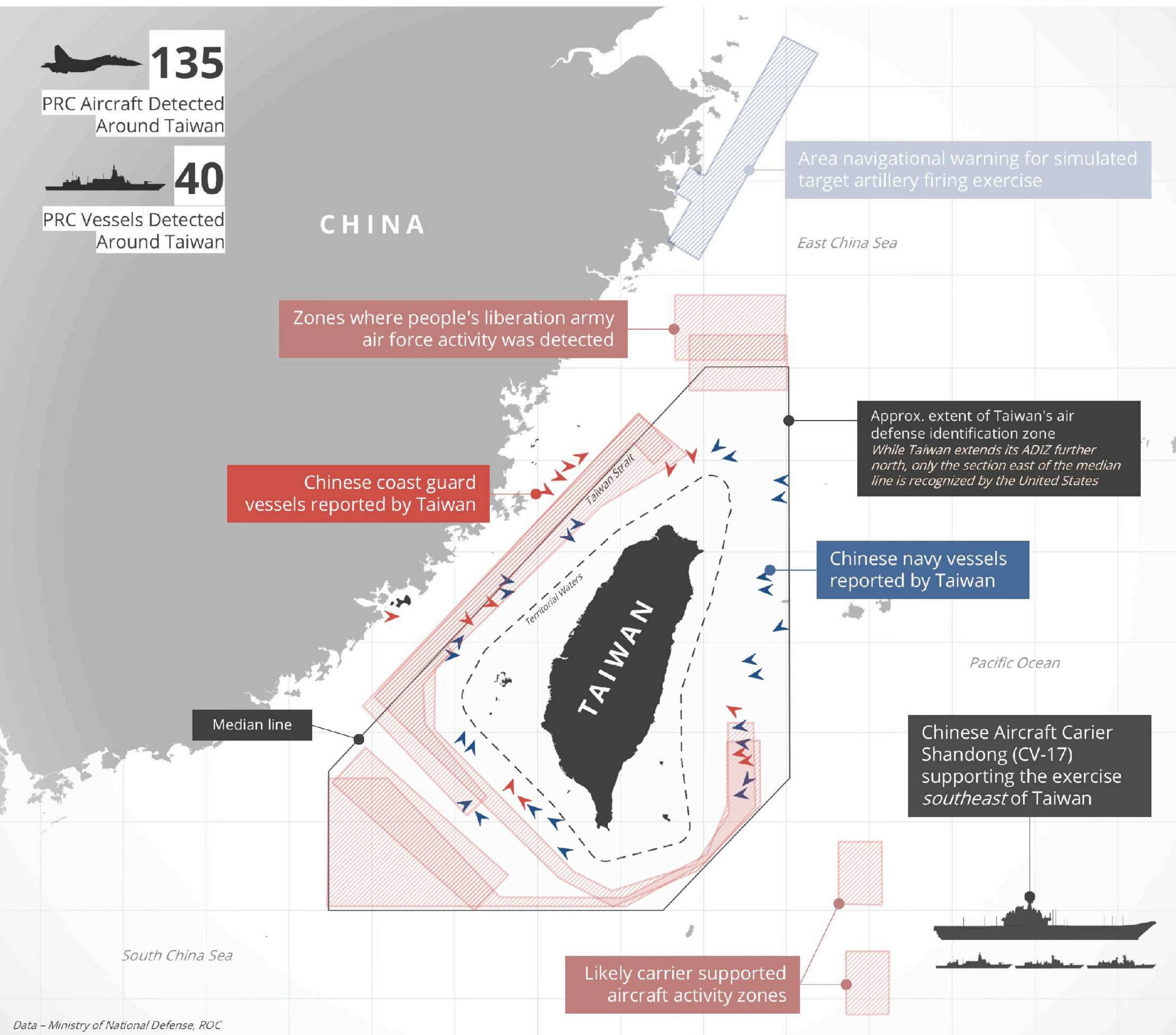
LATEST POWER PROJECTION EXERCISE BY CHINA, 48-HOUR EXERCISES ENCIRCLING TAIWAN, INCLUDING EXTENDED-RANGE ARTILLERY FIRINGS—AIR/NAVAL DEPLOYMENTS



Pt 1

Strait Thunder 2025-A

Un exercice de 48h le 2 avril 2025
mené par la People Liberation Army
(Chine) avec encerclement de Taiwan



China practises hitting Taiwan key ports, energy sites in 'live-fire' drills

ASIA / PACIFIC

China's military carried out "live-fire" drills targeting crucial ports and energy sites directed at Taiwan. Washington condemned the drills as "intimidation tactics" and Taiwan's President Lai Ching-te labeled Beijing a "foreign hostile force".

Source: France24 , [url](#)

@fpaupier



Quelles réponses étatiques

Pour incentiver son écosystème régional





A l'échelle européenne, le Chips Act

Pt 1



EN

HomeAbout usOur prioritiesNews and mediaTopicsResourcesEurope and youState of the Union

Home > Strategy and policy > Priorities 2019-2024 > A Europe fit for the digital age > European Chips Act

European Chips Act



The European Chips Act will bolster Europe's competitiveness and resilience in semiconductor technologies and applications, and help achieve both the digital and green transition. It will do this by strengthening Europe's technological leadership in the field. Following the approval by the Parliament and the Council, the regulation entered into force on 21 September 2023.

PAGE CONTENTS

The need for EU action

Strengthening Europe's technological leadership

Investments to support the Chips Act

Short video introducing the European Chips Act

The need for EU action

Chips are strategic assets for key industrial value chains. With the digital transformation, new markets for the chip industry are emerging such as highly automated cars, cloud, Internet of Things, connectivity, space, defence and supercomputers.

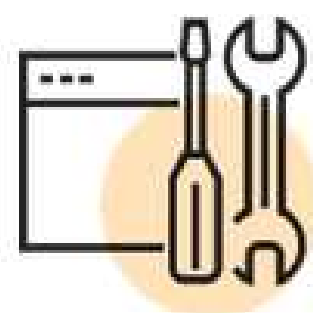
1 trillion microchips were manufactured around	10% EU's share of the global microchips
--	---



A l'échelle européenne, le Chips Act

Pt 1

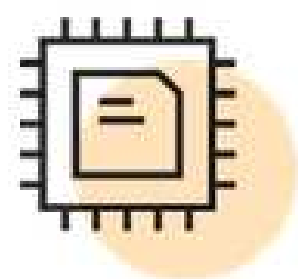
The aim is to:



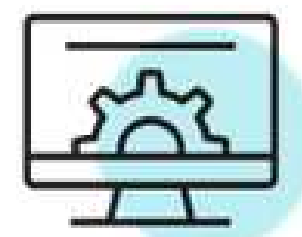
Strengthen Europe's research and technology leadership towards smaller and faster chips



Put in place a framework to increase production capacity to 20% of the global market by 2030



Build and reinforce capacity to innovate in the design, manufacturing and packaging of advanced chips



Develop an in-depth understanding of the global semiconductor supply chains



Address the skills shortage, attract new talent and support the emergence of a skilled workforce



THE WHITE HOUSE

Aux USA, le CHIPS and Science Act



[Administration](#)

[The Record](#)

[Briefing Room](#)

[Visit](#)

[Español](#)

[MENU](#)

Pt 1

AUGUST 09, 2022

FACT SHEET: CHIPS and Science Act Will Lower Costs, Create Jobs, Strengthen Supply Chains, and Counter China



[BRIEFING ROOM](#)

[STATEMENTS AND RELEASES](#)

In President Biden's first year in office, the Biden-Harris Administration has implemented an industrial strategy to revitalize domestic manufacturing, create good-paying American jobs, strengthen American supply chains, and accelerate the industries of the future. These policies have spurred an historic recovery in manufacturing, adding 642,000 manufacturing jobs since 2021. Companies are investing in America again, bringing good-paying manufacturing jobs back home. The construction of new manufacturing facilities has increased 116 percent over last year.

Comparatif des approches législatives EU / USA

Pt 1

sur le renforcement de leur industrie semi-conducteur

Aspect	US Chips and Science Act 	European Chips Act 
Financement	\$280 milliard d'aide fédérale, avec subventions et réductions d'impôts	Usage de fonds existants dans l'UE et des ressources des états membres; pas de nouveau financement dédié
Framework légal	Loi fédérale spécifiant les subventions et avantages fiscaux	Regulation UE avec aide des états membres; les subventions nécessitent accord de la Commission
Périmètre	Semi-conducteurs plus large périmètre science / tech R&D	Focus spécifique industrie des semi-conducteurs
Complexité d'implémentation	Agrément bilatéral état fédéral / acteur privé direct	Partenariat multi couche entre état membre - Commission européenne - acteur privé

Matérialisation de ces incentives aux USA



Ouverture d'une fab TSMC en Arizona avec début des opérations en 2025

A bénéficié de \$6.6 billion USD de financement via le CHIPS Act

Une accélération difficile en UE

L'USINE NOUVELLE

Secteurs ▾Obsessions ▾Rendez-vous ▾Régions ▾Vidéos & Podcasts ▾Data ▾

INTEL \ CARTE À PUCE \ ELECTRONIQUE - INFORMATIQUE

L'Allemagne espère rebondir après l'abandon par Intel de son projet de méga-usine à 30 milliards d'euros

Pour lire l'intégralité de cet article, [abonnez-vous à L'Usine nouvelle - édition Abonné](#)

La santé économique du géant des semi-conducteurs a eu raison de son projet d'implantation à Magdebourg, un investissement de 30 milliards d'euros, dont 10 milliards d'euros d'argent public. Ce revirement relance aussi le débat sur la stratégie allemande en matière de subventions et de réindustrialisation technologique.

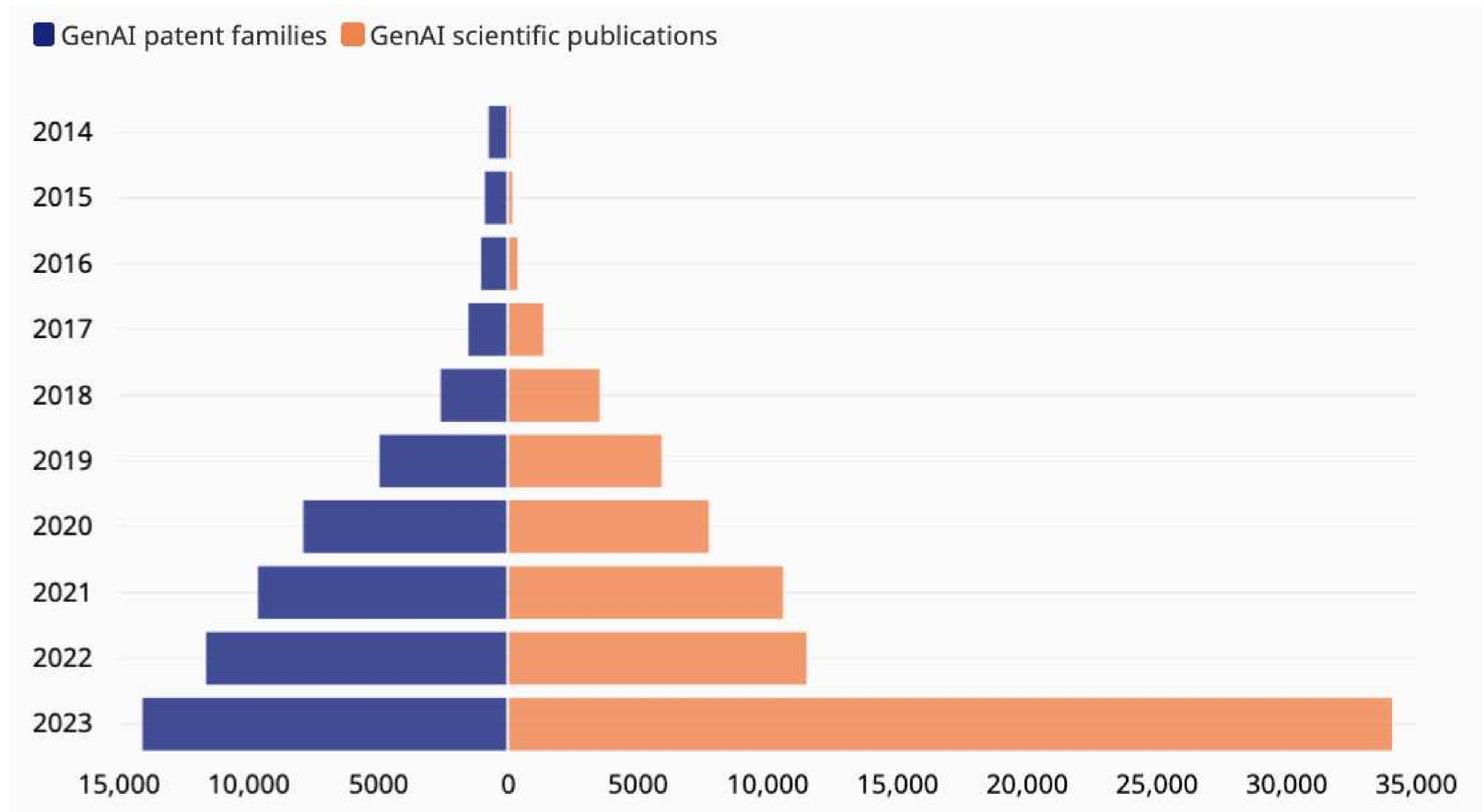
Abandon par intel de son projet de Megfab en Allemagne en fin juillet 2025

Partie 2 - Supply chain model

Qui développe les architectures SOTA ?

Une explosion du nombre de publications et de brevets

Pt 2



L'open source en IA gen - un gradient d'ouverture

Pt 2

	Base Model Data	End User Model Data	Base Model Weights	End User Model Weights	Training Code	Code Documentation	Hardware Architecture	Preprint	Paper	Modelcard	Datasheet	Package	API and Meta Prompts	Licenses
OLMo by AI2 Base models: OLMo-2-0325-32B	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	~	✓
YuLan by Gaoling School of Artificial Intelligence Base models: YuLan-Mini	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✗	✓
BLOOMZ by BigScience Workshop Base models: BLOOM	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	~
Poro by Silo AI and TurkuNLP and High Performance Language Technologies (HPLT) Base models: Poro-34B	~	~	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓
Open Assistant by LAION Base models: Pythia-12B	✓	✓	✓	✓	✓	✓	~	~	~	✓	~	✗	✓	✓

Model provider - tour d'horizon des prestataires

Industriel	Nationalité	Modèles/Produits
 OpenAI		GPT4, GPT4-o, GPT-5
 Meta		Llama, llama2, llama3, llama3.2, llama4
 Microsoft		Phi, Phi-2, Phi-3, Phi-3.5 - MSFT Copilot, Github Copilot
 Google		Gemini (ex Bard), Gemma
ANTHROPIC		Claude, Claude instant, 3.5 Sonnet, Claude 4, Opus
		Apple Intelligence
		Command-R, Embed, Rerank
 (twitter)		Grok, Grok-2, Grok3
 MISTRAL AI	 / 	Mistral, Mixtral, Pixtral, Ministral, Magistral
LightOn	 / 	Alfred, Paradigm
		NeMo Megatron
 deepseek		DeepSeek V3, R1
 Alibaba Cloud		Qwen 1,2,3,4
		OpenSeek, BGE

Les hubs - espaces privilégiés pour animer sa veille

Pt 2

The screenshot displays the Hugging Face homepage. At the top, the navigation bar includes the Hugging Face logo, a search bar, and links to Models, Datasets, Spaces, Community, Docs, Enterprise, and Pricing. The left sidebar contains a user profile for 'fpaupier' with links to Profile, Inbox (0), Settings, Billing, and Get PRO. Below this are 'Organizations' (huggingPartyParis) and 'Resources' (Getting Started, Documentation, Forum, Tasks, Learn). The main content area is titled 'Following 61' and shows a list of model updates. The first update is from 'danielhanchen' for 'unsloth/DeepSeek-V3.1-GGUF', updated about 1 hour ago, with 671B parameters, 33.4k downloads, and 69 likes. The second update is from 'NousResearch' for 'Hermes-4-70B', updated about 2 hours ago, with 71B parameters, 2.61k downloads, and 108 likes. The third update is from 'jinaai' for 'jina-embeddings-v4', updated about 2 hours ago, with 4B parameters, 49.4k downloads, and 334 likes. The fourth update is from 'NousResearch' for 'Hermes-4-405B', updated about 4 hours ago, with 406B parameters, 252 downloads, and 49 likes. The right sidebar shows a 'Trending' section for the last 7 days, listing models like 'microsoft/VibeVoice-1.5B', 'openbmb/MiniCPM-V-4_5', 'tencent/Hunyuan-MT-7B', 'meituan-longcat/LongCat-Flash-Chat', 'DeepSite v2', 'fka/awesome-chatgpt-prompts', 'Wan2.2 14B Fast', and 'FastVLM WebGPU'.

Hugging Face Search models, datasets, users...

Models Datasets Spaces Community Docs Enterprise Pricing

Following 61

All **Models** Datasets Spaces Papers Collections Community Posts Upvotes Likes Articles

danielhanchen updated a model about 1 hour ago

unsloth/DeepSeek-V3.1-GGUF
671B • Updated about 1 hour ago • 33.4k • 69

NousResearch model updated by teknium about 2 hours ago

NousResearch/Hermes-4-70B
Text Generation • 71B • Updated about 2 hours ago • 2.61k • 108

jinaai model updated by jupyterjazz about 2 hours ago

jinaai/jina-embeddings-v4
Visual Document Retrieval • 4B • Updated about 2 hours ago • 49.4k • 334

NousResearch models updated by teknium about 4 hours ago

NousResearch/Hermes-4-405B
Text Generation • 406B • Updated about 4 hours ago • 252 • 49

NousResearch/Hermes-4-405B-FP8
Text Generation • 406B • Updated about 4 hours ago • 1.45k • 16

NousResearch/Hermes-4-70B-FP8

Trending last 7 days

All Models Datasets Spaces

microsoft/VibeVoice-1.5B
Text-to-Speech • 3B • Updated about 1 hour ago • 107k • 1.23k

openbmb/MiniCPM-V-4_5
Image-Text-to-Text • 9B • Updated about 1 hour ago • 11.3k • 813

tencent/Hunyuan-MT-7B
Translation • 8B • Updated about 1 hour ago • 18 • 331

meituan-longcat/LongCat-Flash-Chat
Text Generation • 562B • Updated about 1 hour ago • 218 • 320

DeepSite v2
Generate any application with DeepSeek
12.6k

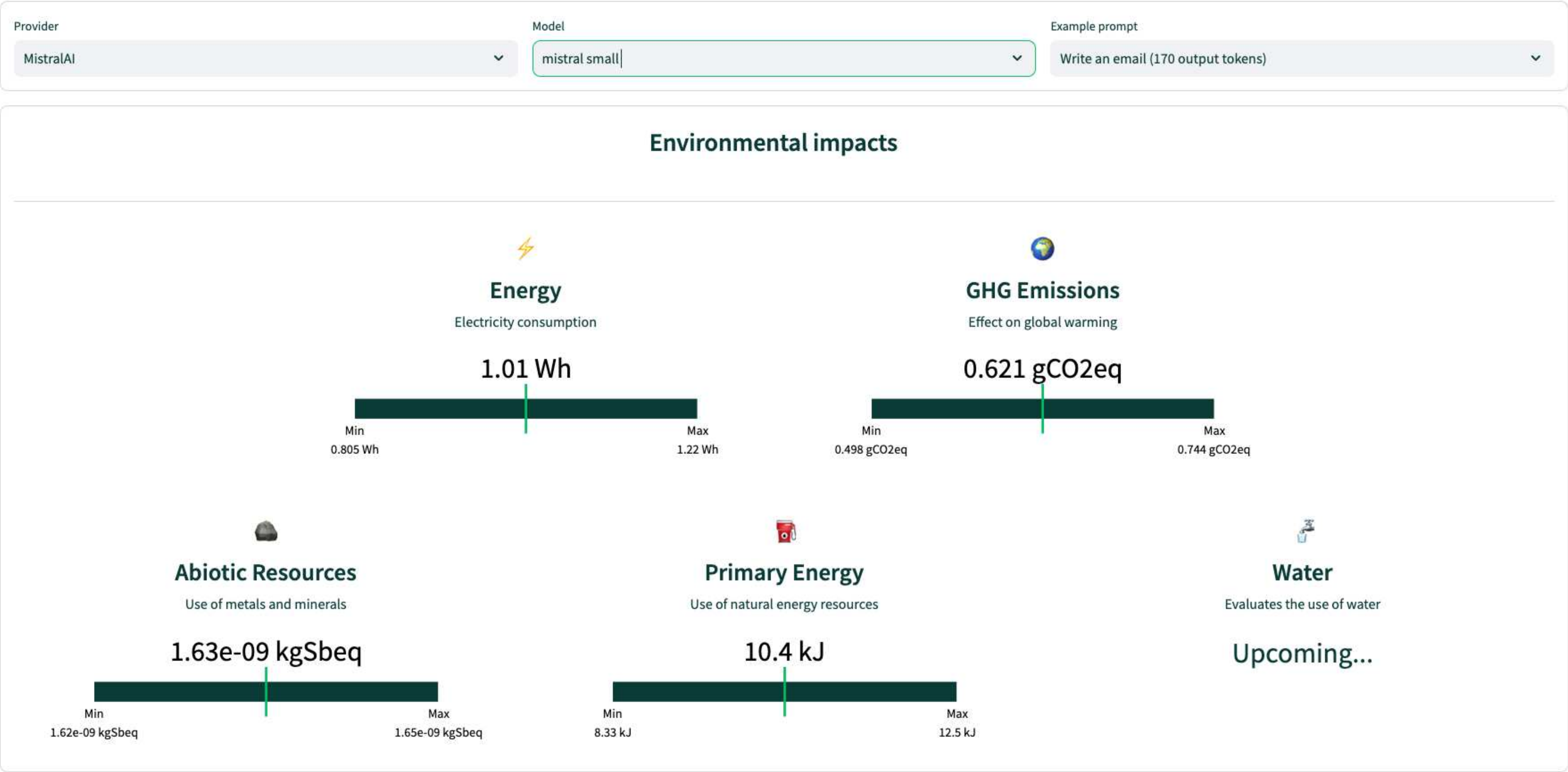
fka/awesome-chatgpt-prompts
Viewer • Updated Jan 6 • 203 • 41.3k • 8.94k

Wan2.2 14B Fast
generate a video from an image with a text prompt
631

FastVLM WebGPU
Real-time video captioning powered by FastVLM
257

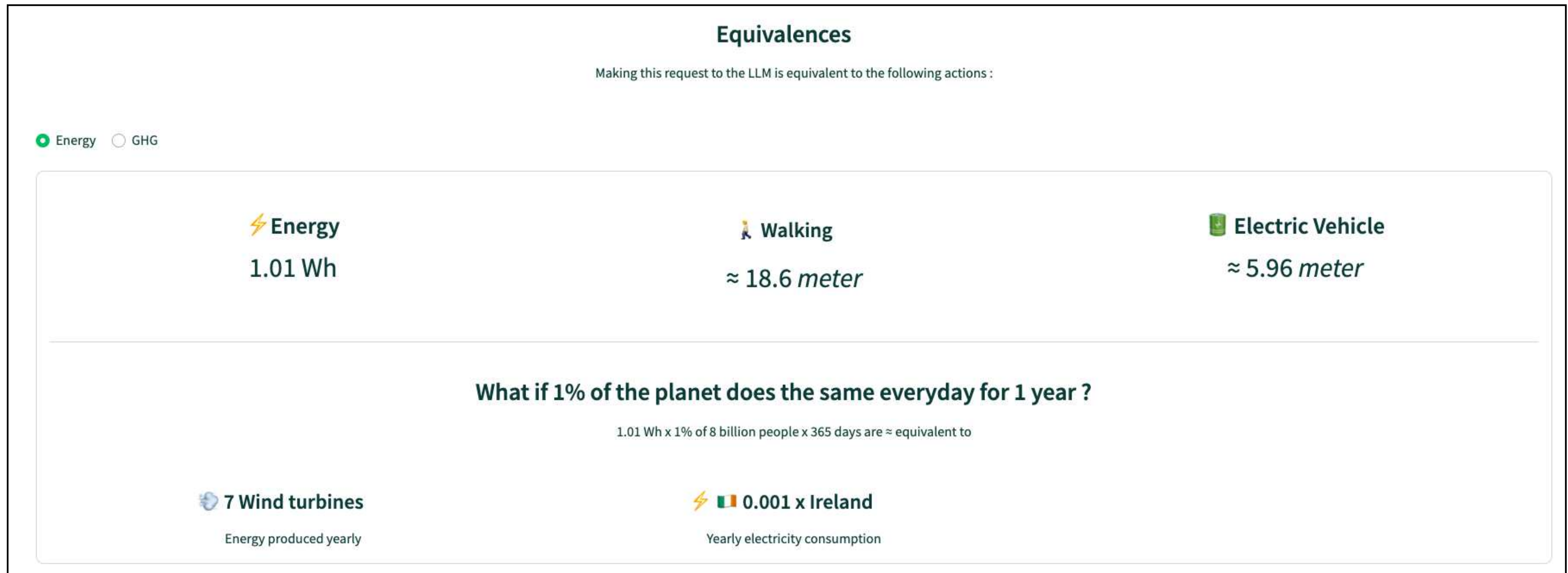
Des outils pour mesurer l'impact énergétique de l'inférence

Pt 2



Des outils pour mesurer l'impact énergétique de l'inférence

Pt 2



Pt 2

Source : Huggingface, [url](#)

Model card

Tentative de
standard de
l'industrie

Model Card

- **Model Details.** Basic information about the model.
 - Person or organization developing model
 - Model date
 - Model version
 - Model type
 - Information about training algorithms, parameters, fairness constraints or other applied approaches, and features
 - Paper or other resource for more information
 - Citation details
 - License
 - Where to send questions or comments about the model
- **Intended Use.** Use cases that were envisioned during development.
 - Primary intended uses
 - Primary intended users
 - Out-of-scope use cases
- **Factors.** Factors could include demographic or phenotypic groups, environmental conditions, technical attributes, or others listed in Section 4.3.
 - Relevant factors
 - Evaluation factors
- **Metrics.** Metrics should be chosen to reflect potential real-world impacts of the model.
 - Model performance measures
 - Decision thresholds
 - Variation approaches
- **Evaluation Data.** Details on the dataset(s) used for the quantitative analyses in the card.
 - Datasets
 - Motivation
 - Preprocessing
- **Training Data.** May not be possible to provide in practice. When possible, this section should mirror Evaluation Data. If such detail is not possible, minimal allowable information should be provided here, such as details of the distribution over various factors in the training datasets.
- **Quantitative Analyses**
 - Unitary results
 - Intersectional results
- **Ethical Considerations**
- **Caveats and Recommendations**

Figure 1: Summary of model card sections and suggested prompts for each.

Pt 2

Model Cards for Model Reporting

Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, Timnit Gebru
{mmitchellai,simonewu,andrewzaldivar,parkerbarnes,lucyvasserman,benhutch,espitzer,tgebru}@google.com
deborah.raji@mail.utoronto.ca

14 Jan 2019

ABSTRACT

Trained machine learning models are increasingly used to perform high-impact tasks in areas such as law enforcement, medicine, education, and employment. In order to clarify the intended use cases of machine learning models and minimize their usage in contexts for which they are not well suited, we recommend that released models be accompanied by documentation detailing their performance characteristics. In this paper, we propose a framework that we call model cards, to encourage such transparent model reporting. Model cards are short documents accompanying trained machine learning models that provide benchmarked evaluation in a variety of conditions, such as across different cultural, demographic, or phenotypic groups (e.g., race, geographic location, sex, Fitzpatrick skin

KEYWORDS

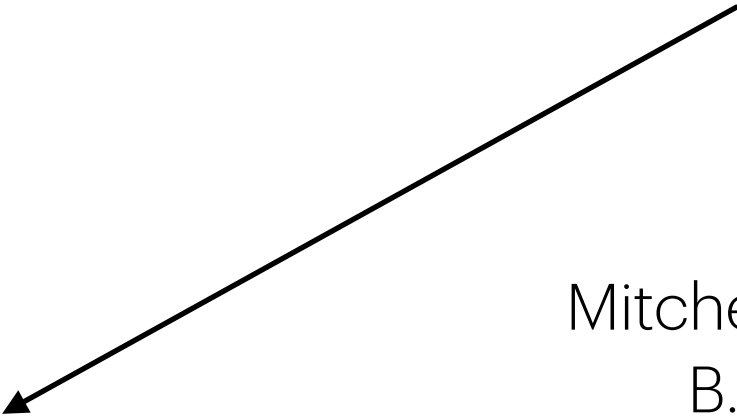
datasheets, model cards, documentation, disaggregated evaluation, fairness evaluation, ML model evaluation, ethical considerations

ACM Reference Format:

Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, Timnit Gebru. 2019. Model Cards for Model Reporting. In *FAT* '19: Conference on Fairness, Accountability, and Transparency*, January 29–31, 2019, Atlanta, GA, USA. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3287560.3287596>

1 INTRODUCTION

Currently, there are no standardized documentation procedures to



Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., ...Gebru, T. (2018). Model Cards for Model Reporting. arXiv, 1810.03993. Retrieved from <https://arxiv.org/abs/1810.03993v2>

ZOOM sur

Mistral



Frontier AI
in your hands

Focus sur les 3 fondateurs

Pt 2

- Arthur Mensch, ex- Google Deepming
- Guillaume et Timothée, ex Meta FAIR
- Lève **105 M€** seed le 13 juin 2023.
- Première modèle Mistral7B publié sep. 2023
- Lève **1.7 B€** (ASML majeur) en sep. 2025

Arthur Mensch



Guillaume Lample

Timothée Lacroix

Fonctionnement : Entre open weight et modèles Privés On-Prem /API

The screenshot displays the Mistral AI pricing page with a navigation bar at the top. The navigation bar includes links for 'Le Chat', 'API pricing', 'Mistral Code', and 'Enterprise deployments', along with currency selectors for 'USD (\$)' and 'EUR (€)'. The main content area features four pricing tiers:

- Free**: Your personal AI assistant for life and work. Get started with our highest-performing models. Price: 0€/mo. Button: Start now >
- Pro** (Popular): Unlock enhanced productivity with extended AI, and agentic capabilities. Price: 17.99€/mo*. Includes a checkbox for 'I am a student'. Button: Start now >
- Team**: Empower your team with a secure, collaborative, AI-powered workspace. Price: 29.99€/user/mo*. Includes a user selector (Add users: - 2 +) and a total price of 60€/mo. Button: Start now >
- Enterprise**: Private deployments powered by custom models, UI, and tools. Price: Custom. Button: Contact us >

Fonctionnement : Entre open weight et modèles Privés On-Prem /API

 Le Chat

 API pricing

 Mistral Code

 Enterprise deployments

USD (\$)

EUR (€)

Our most powerful models and APIs.



Mistral Medium 3

State-of-the-art performance. Simplified enterprise deployments. Cost-efficient.

Input (/M tokens)	0.4€
Output (/M tokens)	2€

mistral-medium-latest



Document AI & OCR

Introducing the world's best document understanding API.

OCR	1€ / 1000 pages
Annotations	3€ / 1000 pages

mistral-ocr-latest



Voxtral Small

State-of-the-art performance on speech and audio understanding.

Input (per min / per M tok)	0.004€ / 0.1€ (text)
Output (/M tokens)	0.3€

voxtral-small-latest

Partie 3 - Supply chain data

Quelle représentation du monde exposée au modèle

Lien du TP 

Dataset d'entraînement

Biais et enjeux de représentations

Nous explorerons dans le TP:

- Patrimoine de langue et biais culturel
- Génération de données synthétiques
- Finetuning d'un modèle sur un jeu de données d'intérêt

QUIZZ!

Bibliographie

- Chip War: The Fight for the World's Most Critical Technology, Chris Miller, October 2022, ISBN 978-1982172008
- ASML, How does EUV work? <https://www.asml.com/en/products/euv-lithography-systems>, retrieved on Sep. 1 2025
- The White House. (2023). FACT SHEET: CHIPS and Science Act Will Lower Costs, Create Jobs, Strengthen Supply Chains, and Counter China | The White House. White House. Retrieved from <https://web.archive.org/web/20250118011306/https://www.whitehouse.gov/briefing-room/statements-releases/2022/08/09/fact-sheet-chips-and-science-act-will-lower-costs-create-jobs-strengthen-supply-chains-and-counter-china>
- Ministère de l'Europe et des Affaires étrangères, (2025, September 01). L'espace indopacifique : une priorité pour la France. Retrieved from <https://www.diplomatie.gouv.fr/fr/dossiers-pays/strategies-regionales/indopacifique/l-espace-indopacifique-une-priorite-pour-la-france>
- Varadarajan, R., Koch-Weser, I., Richard, C., Fitzgerald, J., Singh, J., Thornton, M., & Casanova, R. (2024). Emerging Resilience in the Semiconductor Supply Chain. BCG Global. Retrieved from <https://www.bcg.com/publications/2024/emerging-resilience-in-semiconductor-supply-chain>