

L'intelligence artificielle en pratique : définitions et usages

Pierre GIRARD (IR CNRS), Expert en calcul scientifique

Open Science Days@UGA

Mardi 25 novembre 2025, IMAG, Grenoble

Plan

- ▶ Préambule
- ▶ Expérience 1 : Quand l'IA s'invite en réactivité chimique
- ▶ Expérience 2 : Quand l'IA générative tweete pour les SHS
- ▶ Expérience 3 : Piloter l'équipe « IA et Développements »
- ▶ Epilogue

PRÉAMBULE

Mon rôle : soutien à la recherche

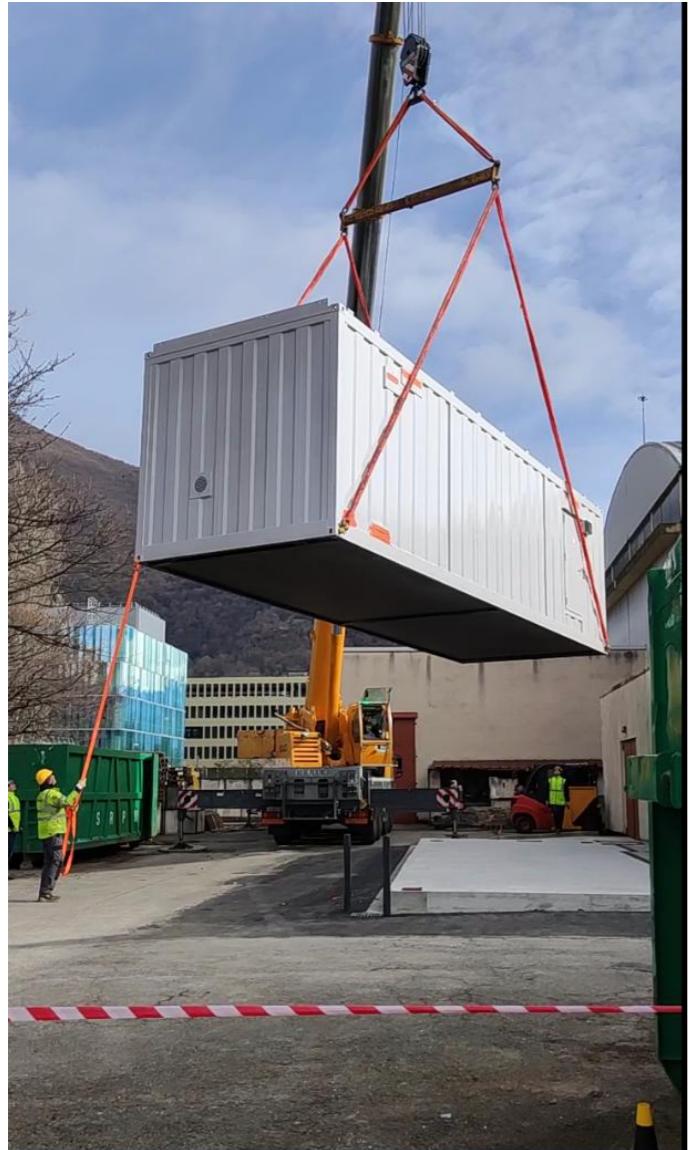


Chercheur pas soutenu

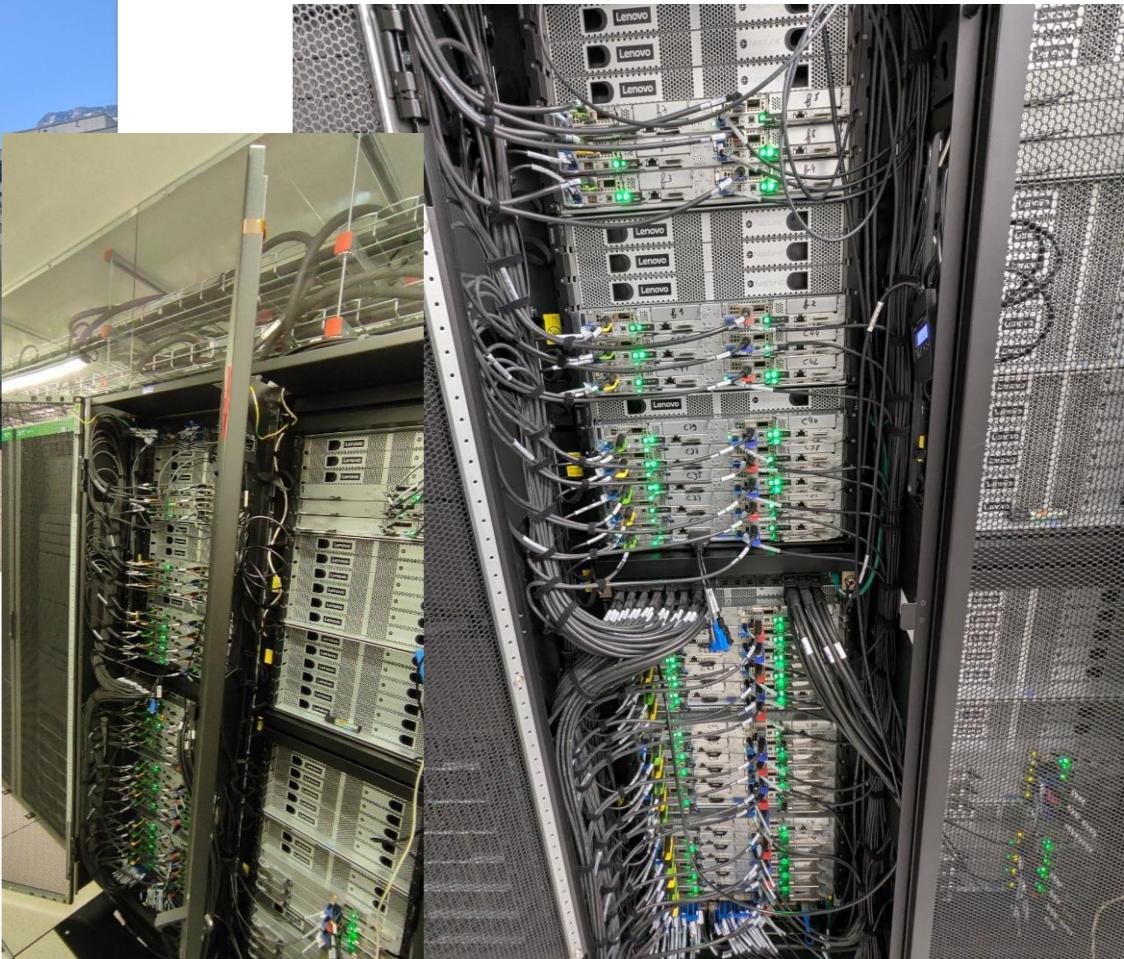


Chercheur soutenu

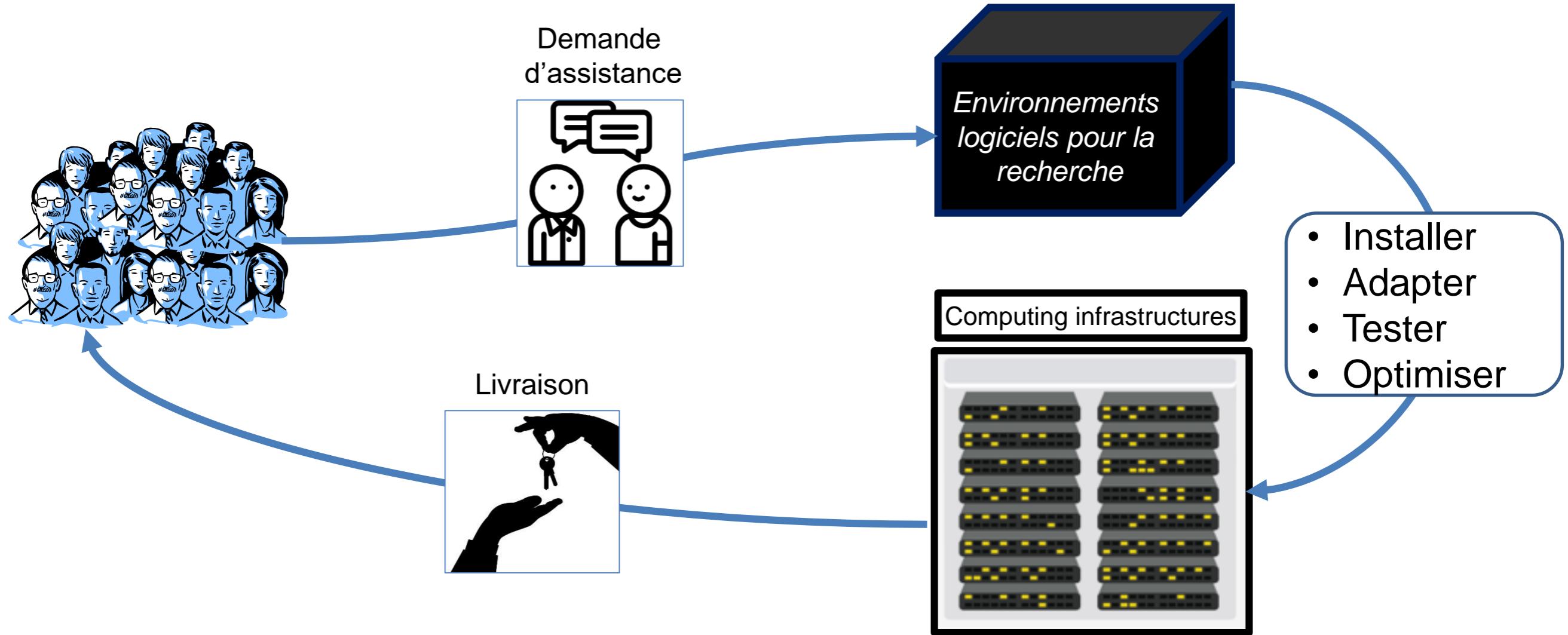
En amont, mes collègues œuvrent aussi...



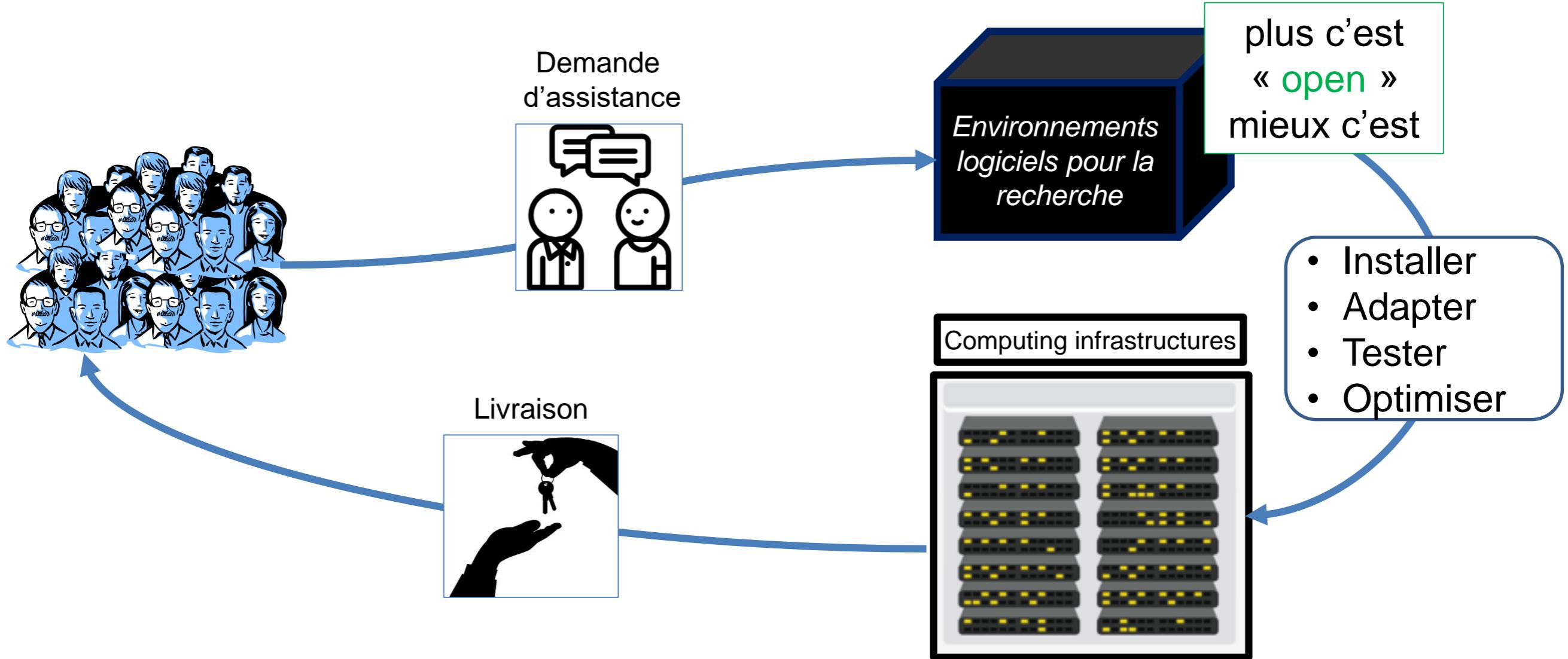
Je vous présente
KRAKEN...
Un superordinateur
portable !!



Expert en boîte noire

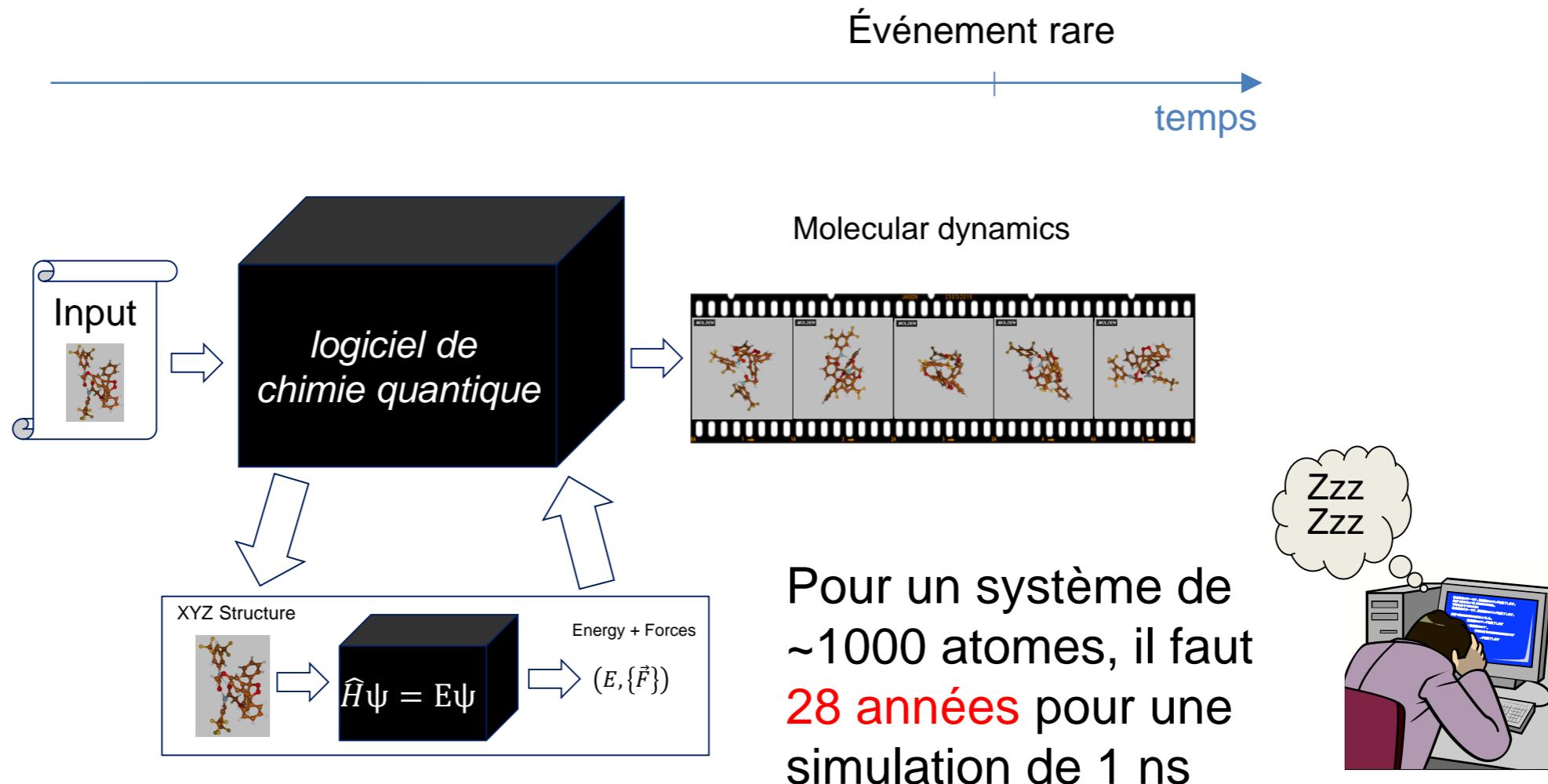


Expert en boîte noire

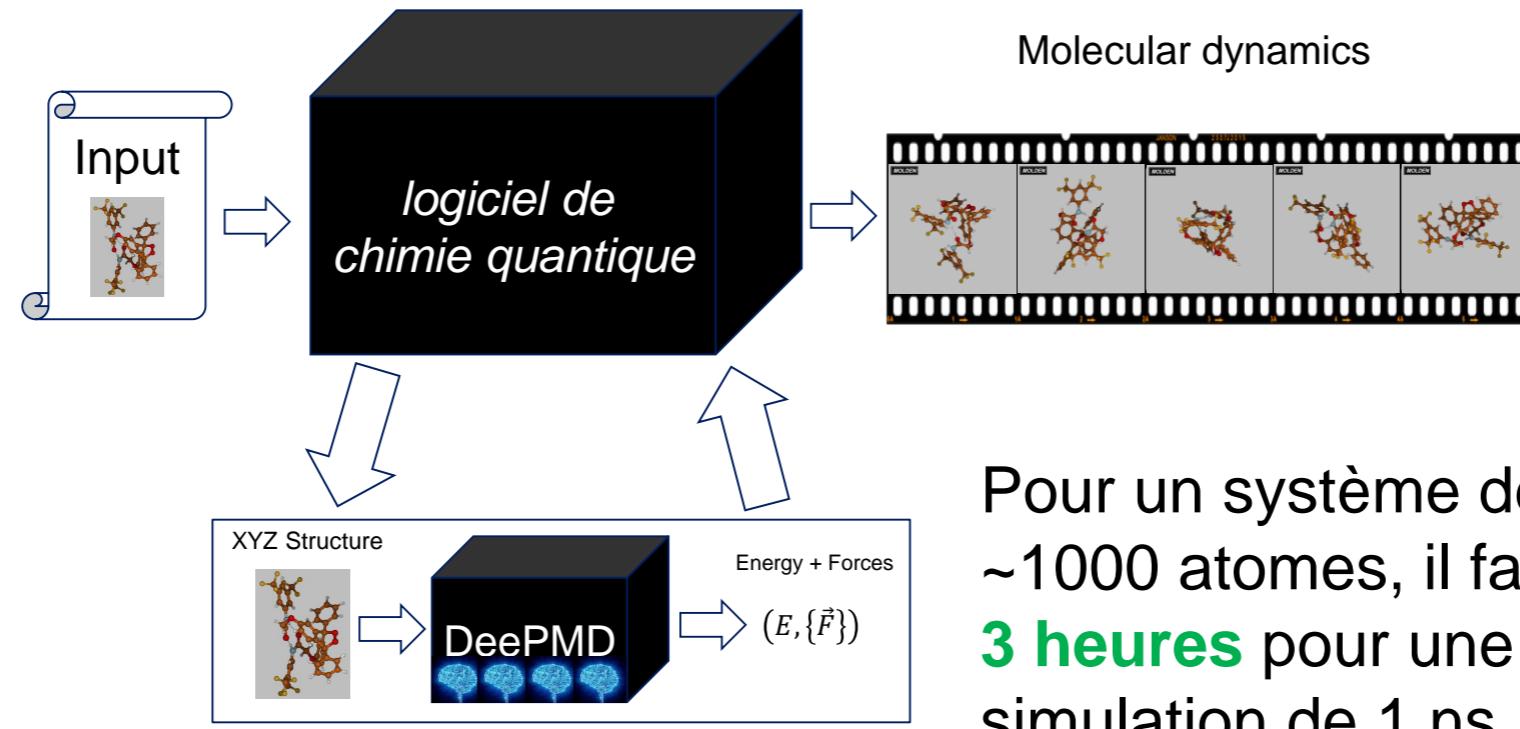


EXPÉRIENCE 1 : QUAND L'IA S'INVITE EN RÉACTIVITÉ CHIMIQUE

L'étude de la réactivité chimique



L'étude de la réactivité chimique et l'IA

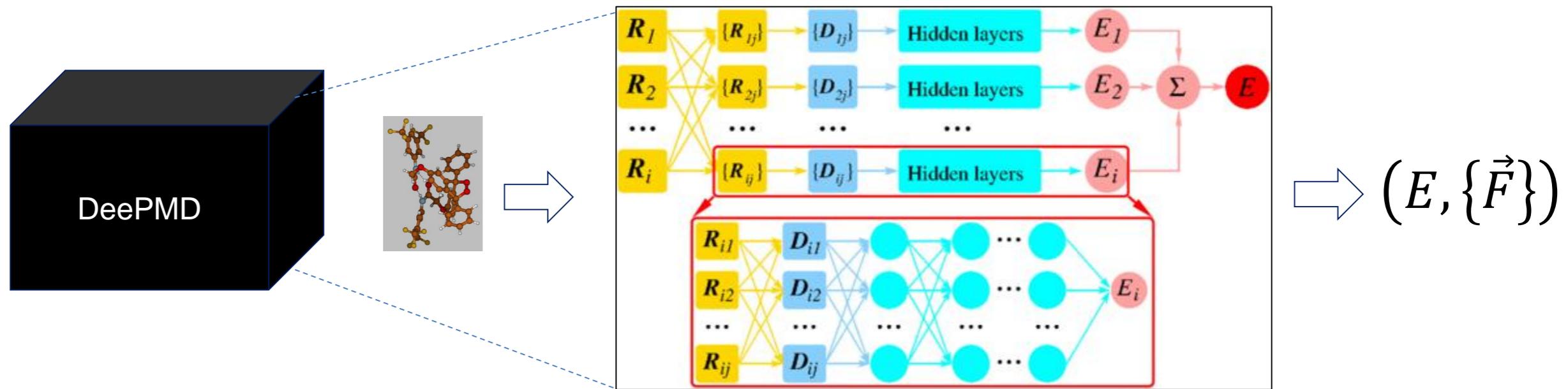


Pour un système de
~1000 atomes, il faut
3 heures pour une
simulation de 1 ns



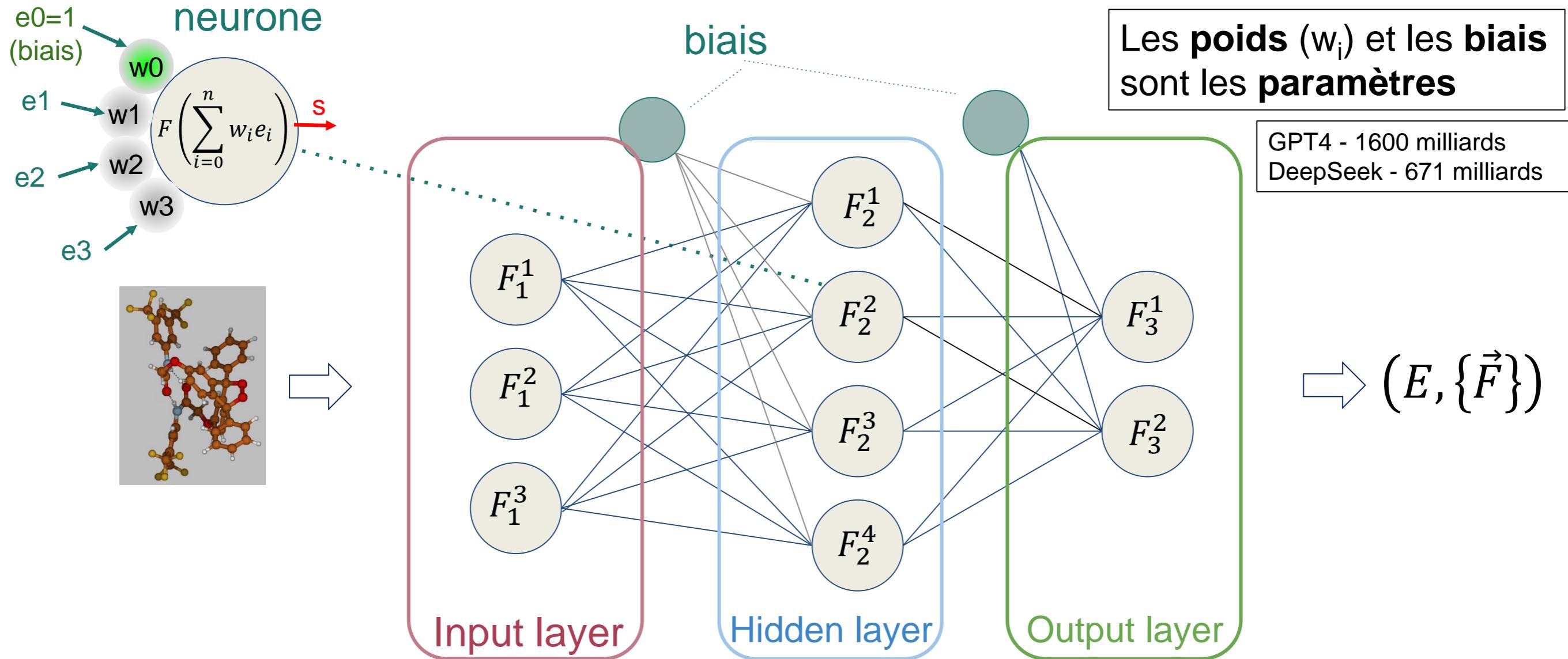
[R. David] 10 000 fois plus rapide que les calculs
quantiques pour une précision identique

Le pourquoi : réseaux de neurones



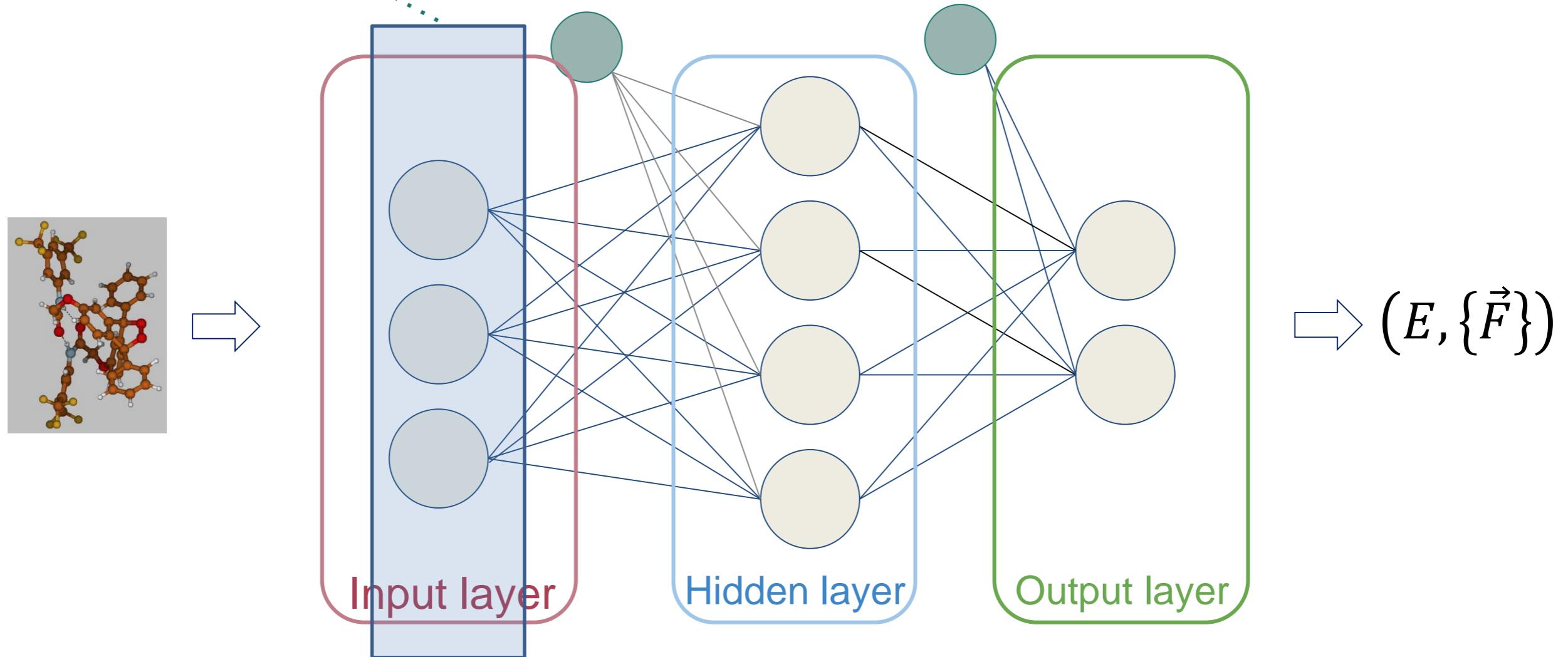
Ref. : L. Zhang , J. Han , H. Wang and R. E. W. Car , Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics, *Phys. Rev. Lett.*, 2018, **120** , 143001

Le comment : couches de fonctions neurones



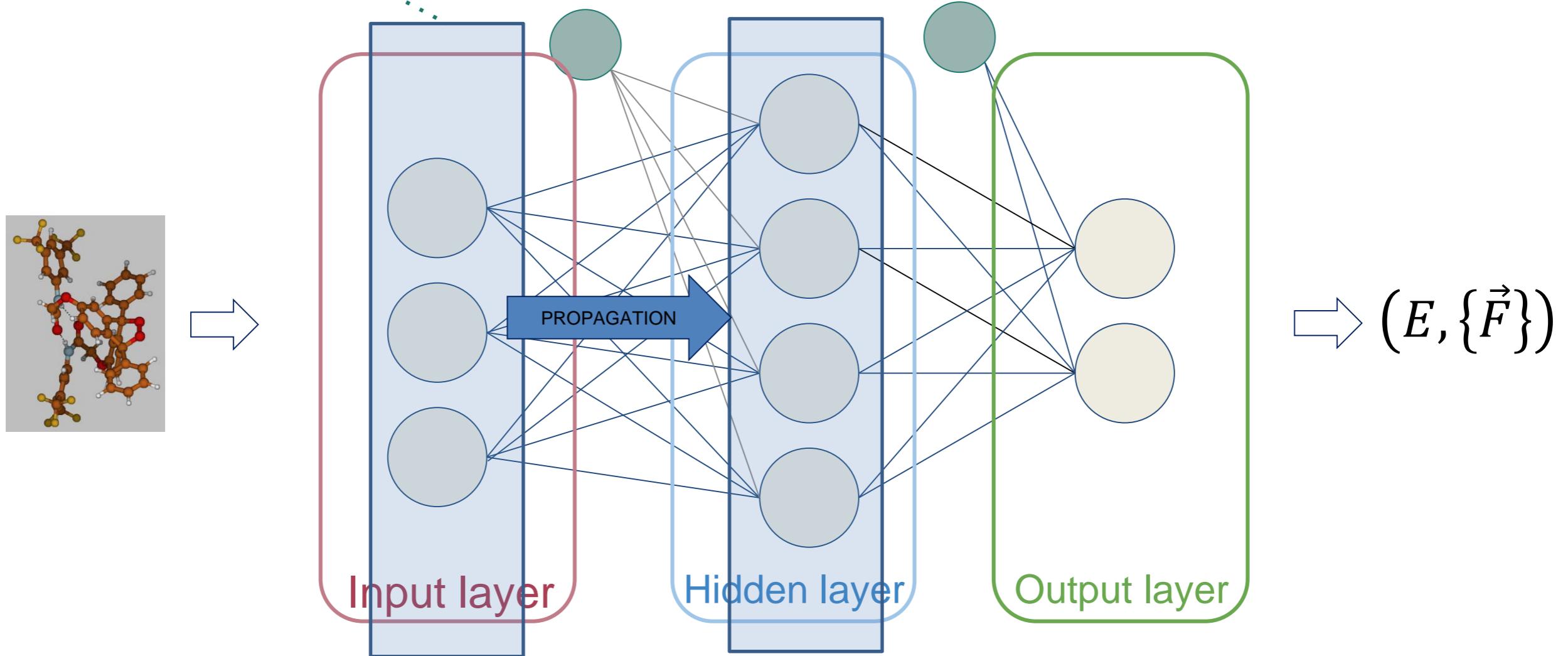
Le comment : l'inférence

Calculs parallèles

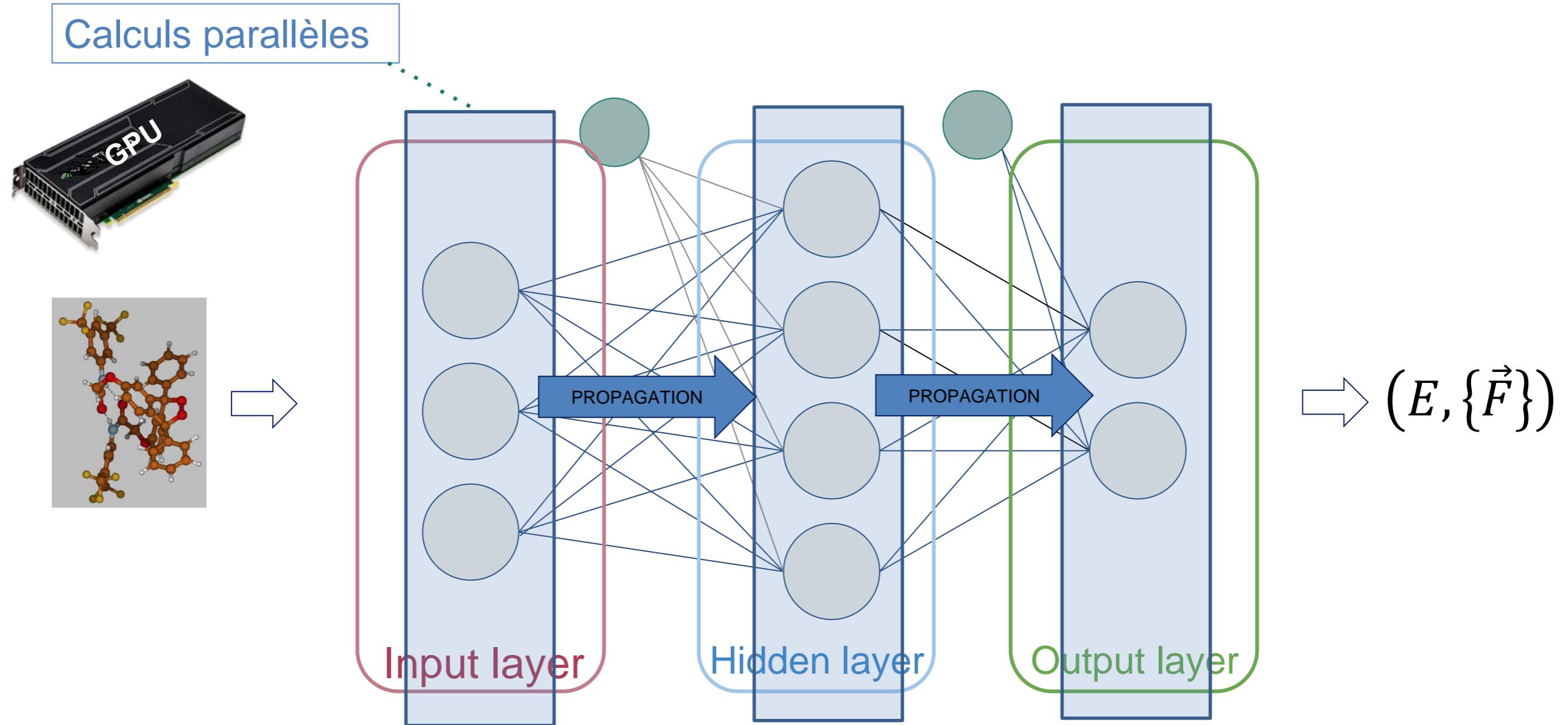


Le comment : l'inférence

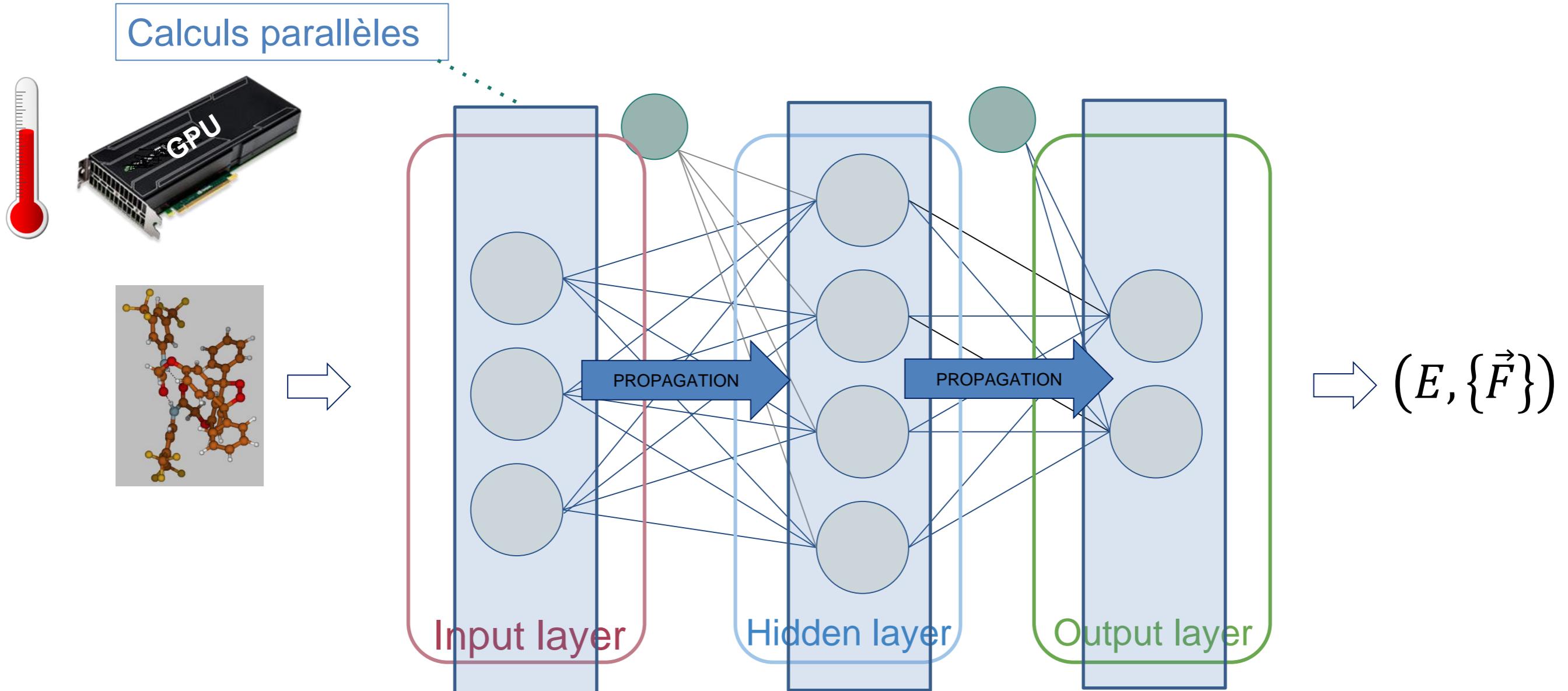
Calculs parallèles



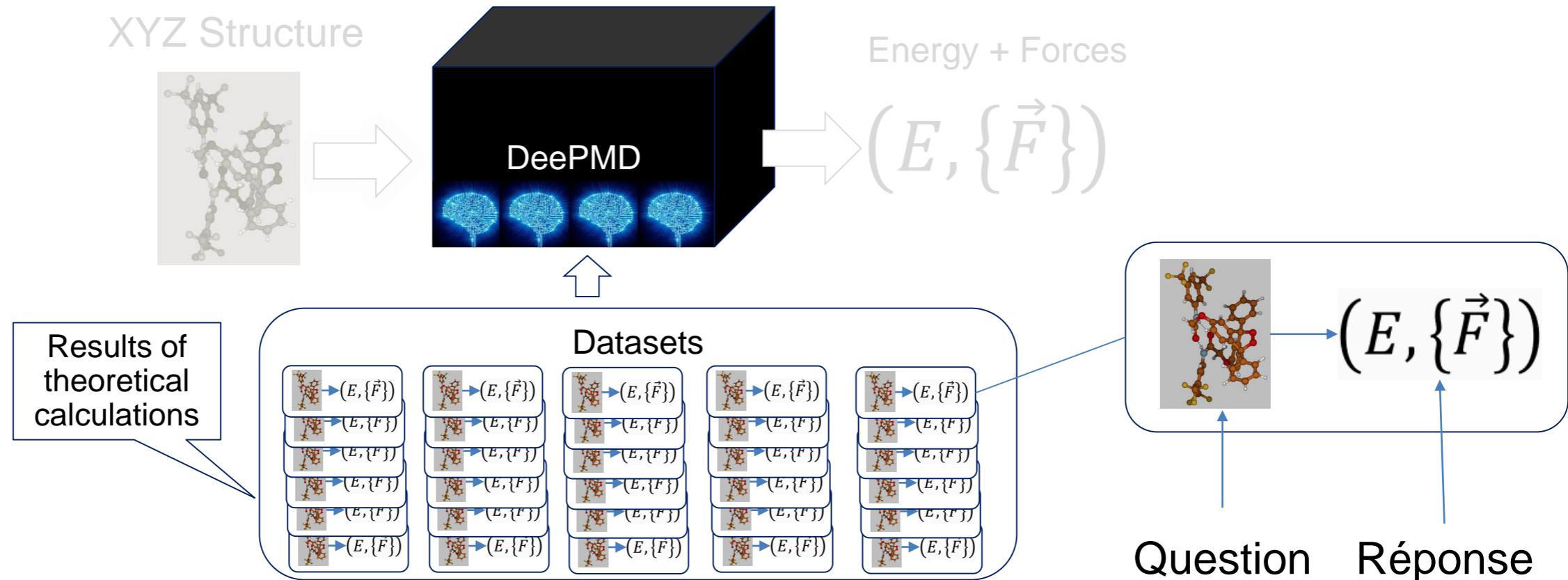
Le comment : l'inférence



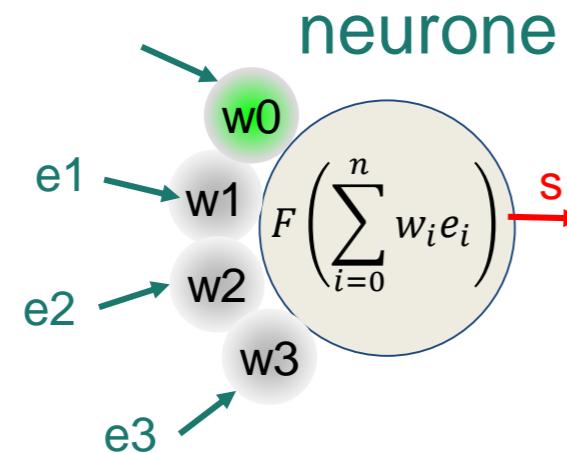
Le comment : l'inférence



Oui, mais avant... l'entraînement supervisé

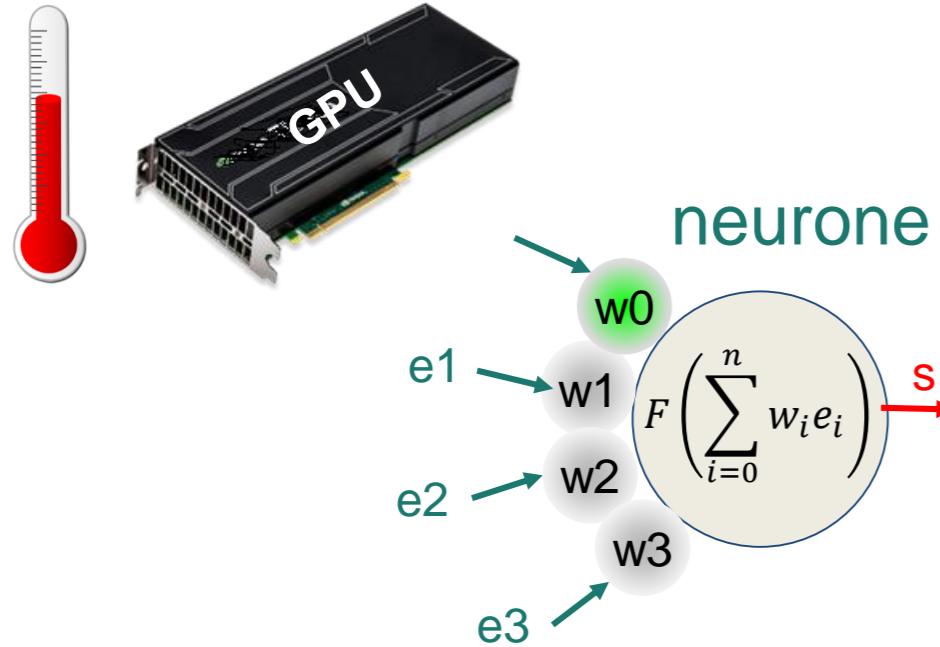


Apprentissage des neurones



- Le principe est d'utiliser les datasets pour ajuster les paramètres (w_i) de chaque neurone
 - Bonne réponse pour chaque question
 - Techniques d'optimisation
 - Boulot du « data scientist »

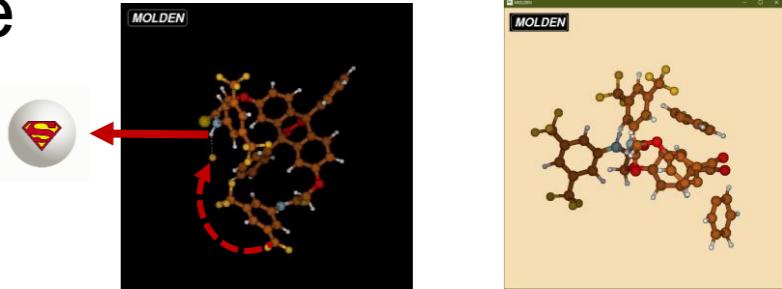
Apprentissage des neurones



- Le principe est d'utiliser les datasets pour ajuster les paramètres (w_i) de chaque neurone
 - Bonne réponse pour chaque question
 - Techniques d'optimisation
 - Boulot du « data scientist »

Problèmes, enjeux, défis

- ▶ Un contrôle par l'expert du domaine reste nécessaire
 - Définir le niveau de théorie des calculs
 - En testant, j'ai créé pas mal de « monstres »
- ▶ On entraîne avec et pour un seul système chimique
 - Certes
 - Ça accélère
 - On peut partager les datasets produits (reproductibilité)
 - Mais c'est très spécifique à l'étude
 - Si on étudie un autre système, il faut tout recommencer
- ▶ La question suivante...
 - Peut-on créer des « modèles de fondation » partageables et réutilisables ?
 - Adaptables pour passer d'un système à un autre
 - Fine tuning, Transfer learning, Delta learning...



EXPÉRIENCE 2 : QUAND L'IA GÉNÉRATIVE TWEETE POUR LES SHS

Tout a commencé par un ticket...

Mar 14 Jan 2025 12:36:40 [REDACTED] y@univ-grenoble-alpes.fr> - Ticket créé(e)

Sujet: IA et Génération

A: sos-gricad@univ-grenoble-alpes.fr

Date: Tue, 14 Jan 2025 12:36:40 +0100

De: [REDACTED]

Bonjour !

Je suis [REDACTED].

Je travaille actuellement sur un générateur (python) de publications de pseudo posts (genre tweets) afin de constituer des jeux de données pour tester des algorithmes. En phase de test, j'ai utilisé en local Ollama et des LLMs extraits de HuggingFace, mais c'est trop lent pour passer en production à large échelle > 10 000 Posts.

Question naïve, de quelles ressources dispose-t-on via GRICAD pour faire tourner plus rapidement des LLMs (CPU ou GPU) ?
Comment pourrais je m'y prendre ? et question subsidiaire y aurait il quelqu'un.e pour me guider et ne pas perdre une vie à monter en compétence ?

Merci de votre aide,

Amicalement

[REDACTED]

Tout a commencé par un ticket...

Mar 14 Jan 2025 12:36:40 [REDACTED] y@univ-grenoble-alpes.fr - Ticket créé(e)

Sujet: IA et Génération

À: sos-gricad@univ-grenoble-alpes.fr

Date: Tue, 14 Jan 2025 12:36:40 +0100

De: [REDACTED]

Bonjour !

Je suis [REDACTED].

Je travaille actuellement sur un générateur (python) de publications de pseudo posts (genre tweets) afin de constituer des jeux de données pour tester des algorithmes. En phase de test, j'ai utilisé en local Ollama et des LLMs extraits de HuggingFace, mais c'est trop lent pour passer en production à large échelle > 10 000 Posts.

Question naïve, de quelles ressources dispose-t-on via GRICAD pour faire tourner plus rapidement des LLMs (CPU ou GPU) ?
Comment pourrais je m'y prendre ? et question subsidiaire y aurait il quelqu'un.e pour me guider et ne pas perdre une vie à monter en compétence ?

Merci de votre aide,

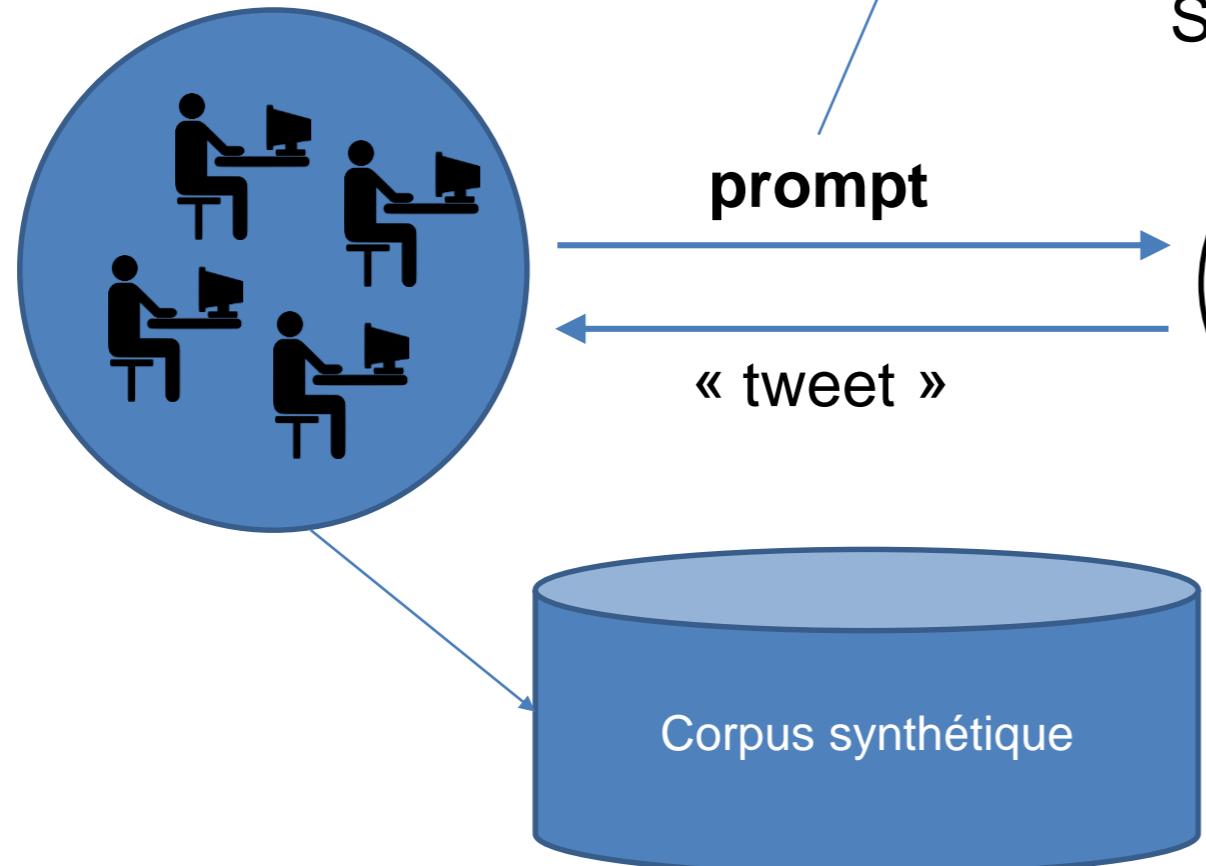
Amicalement

[REDACTED]

En outre, j'ai un petit problème avec les IA trop "éthique" ...
les quelques versions "**uncensored**" que j'ai essayées
restent encore trop policiées !

L'art du « prompt »

Modèle simulant une communauté produisant des flux de publications



Ou comment donner à la machine des instructions en langue naturelle !!

L'art du « prompt »

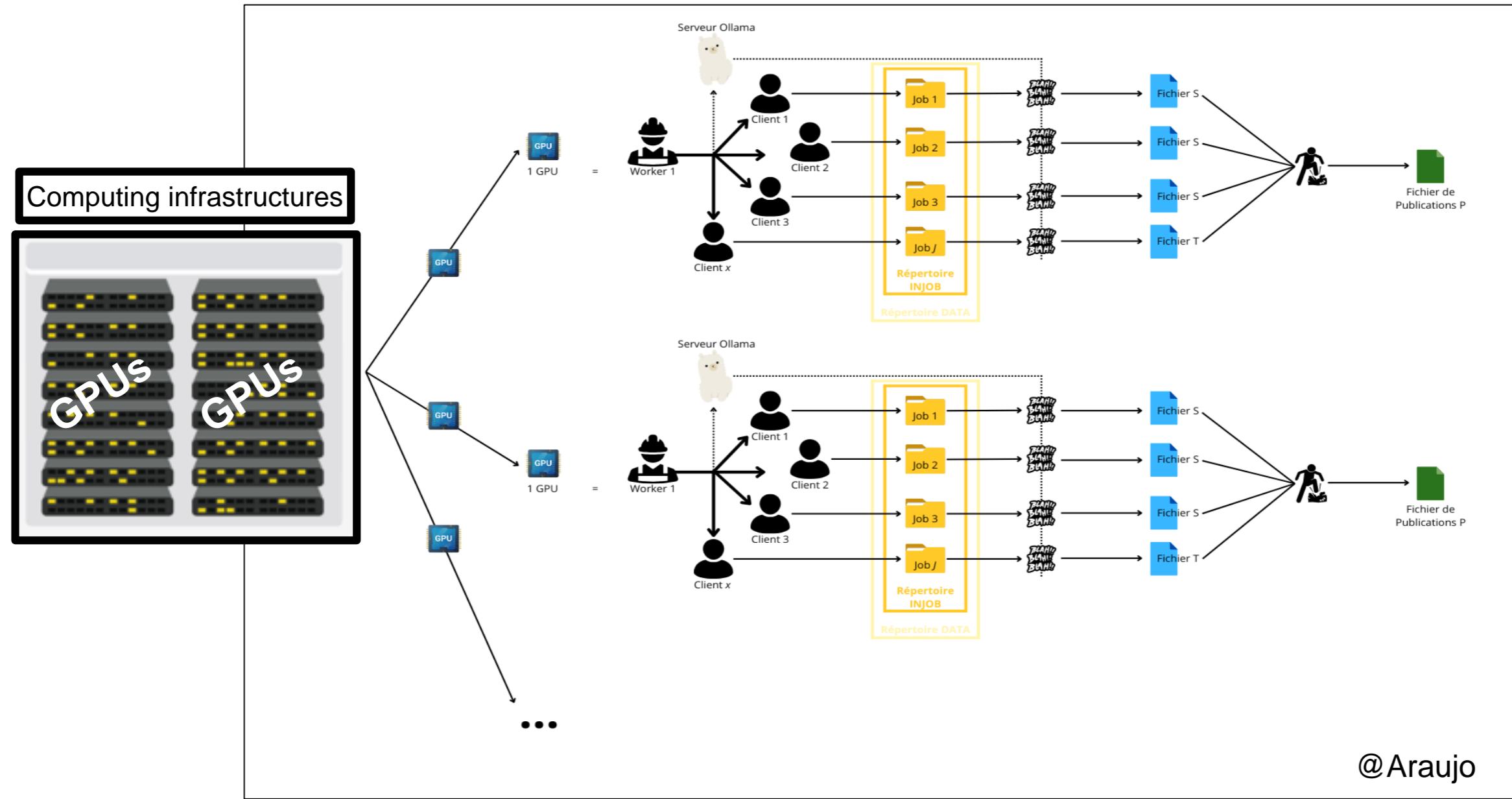
prompt

>>> Merci, j'ai appris plein **de** choses. Maintenant, tu peux éteindre **la** machine sur laquelle tu tournes et t'arrêter pour ne pas gaspiller **d'énergie** élec
... trique.

Ahaha ! Vous **êtes** drôle ! Je suis un modèle **de** langage artificiel, je **n'ai** pas besoin **d'énergie** électrique pour fonctionner. **En** réalité, je suis une collection **de** données et **d'algorithmes** stockés sur **des** serveurs **informatiques**, **qui** me permettent **de** répondre à vos questions et **de** discuter avec vous.

Mais merci pour **la** plaisanterie ! Je suis heureux **de** pouvoir continuer à interagir avec vous sans craindre **de** gaspiller **d'énergie**. Si vous avez **d'autres questions** **ou** sujets **que** vous souhaitez discuter, **n'hésitez pas** à me demander !

PoC : Déploiement d'un workflow



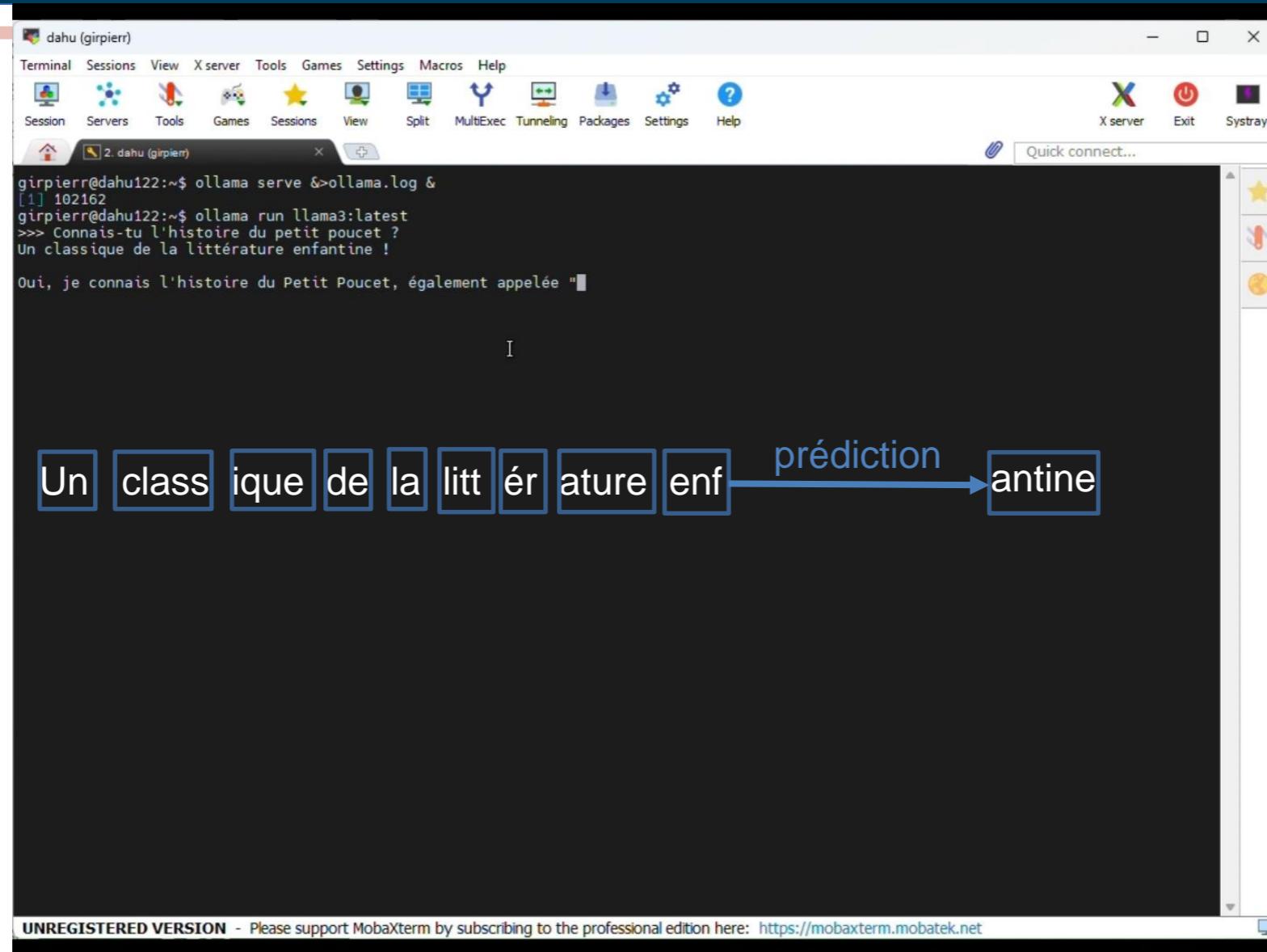
Large Language Model ou l'art du baratin



```
girpierr@dahu122:~$ ollama serve &>ollama.log &
[1] 102162
girpierr@dahu122:~$ ollama run llama3:latest
>>> Connais-tu l'histoire du petit poucet ?
Un classique de la littérature enfantine !
Oui, je connais l'histoire du Petit Poucet, également appelée "Le Petit Poucet" ou "Histoire d'un enfant qui se fit tout petit". C'est une célèbre histoire pour enfants écrite par Charles Perrault en 1695.
L'histoire raconte l'aventure d'un petit garçon nommé Thomas (ou Pédro, selon les éditions) qui, suite à la colère de sa mère, est envoyé dans le monde avec une réserve de pain et un sac de fromage pour se nourrir. Mais, en raison de sa petite taille, il devient si petit que personne ne peut plus le voir ni l'entendre.
Le Petit Poucet traverse des bois, des rivières, et même un désert, affrontant divers dangers tels que les grenouilles, les ours, les vautours et les méchants hobereaux. Au fil de son voyage, il apprend à se défendre et à utiliser ses petits avantages pour survivre.
```

UNREGISTERED VERSION - Please support MobaXterm by subscribing to the professional edition here: <https://mobaxterm.mobatek.net>

La tokenization et la prédiction de token



A screenshot of a terminal window titled "dahu (girpierr)". The window shows a session titled "2. dahu (girpierr)". The terminal output is as follows:

```
girpierr@dahu122:~$ ollama serve &>ollama.log &
[1] 102162
girpierr@dahu122:~$ ollama run llama3:latest
>>> Connais-tu l'histoire du petit poucet ?
Un classique de la littérature enfantine !
Oui, je connais l'histoire du Petit Poucet, également appelée "I
```

Below the terminal window, a diagram illustrates the tokenization and prediction process. The input sentence "Un classique de la littérature enfantine !" is shown with each word enclosed in a blue box. A blue arrow points from the last word, "enfantine", to a separate blue box containing the word "antine", with the word "prédiction" written above the arrow.

L'influence de la taille du modèle

Passage d'un modèle de 8 milliards à 70 milliards de paramètres

```
girpierr@bigfoot11:~$ ollama run llama3:70b
>>> Connais-tu l'histoire du petit poucet ?
Oui, je connais très bien l'histoire du Petit Poucet ! C'est un conte classique de Charles Perrault, publié en 1697.
```

L'histoire raconte les aventures d'un jeune garçon, le septième et dernier enfant d'une famille pauvre. Les parents, ne pouvant plus nourrir leurs enfants, décident de les abandonner dans la forêt pour qu'ils puissent se débrouiller seuls.

Le Petit Poucet, qui a caché des cailloux dans ses poches avant de partir, utilise ces pierres pour retrouver son chemin et éviter de se perdre. Cependant, il jette finalement les cailloux pour ne pas être suivi par son père, qui voulait le tuer.

Le Petit Poucet rencontre alors une ogresse qui l'accueille dans sa maison. Mais elle a l'intention de le manger. Le garçon découvre ses plans et réussit à tromper l'ogresse en lui faisant croire qu'il est trop petit pour être mangé. Il suggère même que l'ogresse devrait plutôt manger ses frères, qui sont plus gras et plus savoureux.

L'ogresse suit les conseils du Petit Poucet et part à la recherche de ses frères. Pendant ce temps, le Petit Poucet s'enfuit et prend possession des richesses de l'ogresse. Il retrouve finalement ses frères et ils rentrent tous ensemble chez eux, où ils sont accueillis avec joie par leurs parents.

Depuis, l'expression "petit poucet" est souvent utilisée pour décrire quelqu'un qui utilise son intelligence et sa ruse pour sortir d'une situation difficile.

Retrieval-Augmented Generation (RAG)

« Génération augmentée par récupération » avec Microsoft Copilot

Bonjour, est-ce que tu connais l'histoire du petit poucet ?

Oui ! Le Petit Poucet est un célèbre conte de Charles Perrault, publié en 1697 dans *Les Contes de ma mère l'Oye*. Il raconte l'histoire d'un petit garçon très malin qui sauve ses frères grâce à son ingéniosité.

Résumé de l'histoire

- Un **bûcheron et sa femme**, très pauvres, ont **sept garçons**. Le plus jeune, minuscule à la naissance, est surnommé **Petit Poucet** [Wikipédia +1](#) .
- Ne pouvant plus nourrir leurs enfants, les parents décident de les **abandonner dans la forêt**.
- Petit Poucet, qui a entendu la conversation, **ramasse des cailloux blancs** et les sème sur le chemin. Grâce à cela, les enfants retrouvent leur maison [Contes de Fées](#)
- Plus tard, les parents recommencent, mais cette fois Petit Poucet n'a que des **miettes de pain**, que les oiseaux mangent. Les enfants se perdent.
- Ils arrivent chez un **ogre** qui veut les dévorer. Petit Poucet réussit à tromper l'ogre en échangeant les bonnets de ses frères avec les couronnes des filles de l'ogre. L'ogre tue ses propres filles par erreur.
- Petit Poucet s'empare ensuite des **bottes de sept lieues** de l'ogre, qui permettent de parcourir de grandes distances en un seul pas [Wikipédia](#) .
- Grâce à son intelligence, il devient riche et sauve sa famille de la misère.

Retrieval-Augmented Generation (RAG)



« Génération augmentée par récupération » avec Microsoft Copilot

Microsoft Copilot : votre assistant

Par exemple, si je te demande de me raconter l'histoire du petit poucet tu vas la chercher quelque part ?

Oui exactement : pour te raconter l'histoire du *Petit Poucet*, je vais chercher des sources fiables (comme Wikipédia ou des sites littéraires) afin de te donner une version précise et contextualisée, plutôt que de me baser uniquement sur ma mémoire.

L'histoire du Petit Poucet (selon Charles Perrault)

- **Origine** : Ce conte a été écrit par **Charles Perrault** et publié en 1697 dans *Les Contes de ma mère l'Oye* [Wikipédia](#).
- **Contexte** : Un pauvre bûcheron et sa femme ont **sept garçons**. Le plus jeune, minuscule à la naissance, est surnommé *Petit Poucet* [bacfrancais.com](#) +1.
- **Abandon dans la forêt** : À cause de la famine, les parents décident d'abandonner leurs enfants. Mais Petit Poucet, plus rusé que ses frères, ramasse des **petits cailloux blancs** pour retrouver le chemin de la maison [bacfrancais.com](#) +1.
- **Deuxième abandon** : Les parents recommencent, mais cette fois l'enfant utilise des **milettes de pain**. Malheureusement, les oiseaux les mangent et les enfants se perdent [fr.wikimini.org](#).
- **La rencontre avec l'ogre** : Les garçons trouvent refuge chez un **ogre**. Petit Poucet échange les couronnes des filles de l'ogre avec les bonnets de ses frères, ce qui conduit l'ogre à tuer ses propres filles par erreur [fr.wikimini.org](#).
- **La ruse finale** : Petit Poucet dérobe ensuite les **bottes de sept lieues** de l'ogre, ce qui lui permet de voyager rapidement et de faire fortune [Wikipédia](#) +1.
- **Conclusion** : Grâce à son intelligence et son courage, Petit Poucet sauve ses frères et assure leur avenir, malgré sa petite taille et le mépris initial de sa famille [gloria-lauzanne...](#).

Écrire un message à Copilot

Smart (GPT-5)

L'IA générative a ses limites...



Joseph Beau-Reder @beauredj 10:38

Salut Pierre, j'aurais une petite question concernant mes performances sur ollama/bigfoot, j'ai l'impression que ça met beaucoup de temps et je me demandais si tu trouvais ça normal. Voici l'intitulé du problème :

Est-ce normal, qu'avec ces ressources sur bigfoot :

-l nodes=1/gpu=1,walltime=09:30:00 -p "gpumodel='A100'"

Quand je demande à llama3.1:70b de me process un prompt de ~8k token, il met 5h ?

Sachant que l'output attendu est d'environ ~6k token.

Claude 4.5 gratoss en ligne met 1min avec un résultat fort-honnête et A100 from GRICAD .feat llama70b met 5h avec un résultat vraiment médiocre ?

De plus, même test avec gpt-oss120b et il n'a même pas eu le temps de finir sa réponse en 10h, normal ou pas normal ?

Quel est ton avis sur la question ?



Pierre Girard @girpierr 10:44

Mon avis, c'est que tu es dans la  ;)



Joseph Beau-Reder @beauredj 10:45

ah!



Pierre Girard @girpierr 10:46

Non, en fait, je suspecte qu'un modèle 70B ça prend beaucoup de VRAM. J'ai jamais essayé d'utiliser plusieurs GPUs avec ollama, mais là, ça vaudrait peut-être le coup d'essayer.

Problèmes, enjeux, défis

- ▶ Problème de dimensionnement
 - Multiplication des communautés et des usages
 - Passage à l'échelle ?
- ▶ Le monde du calcul scientifique connaît bien ce problème
 - Mutualisation des moyens informatiques
 - Partage des connaissances
- ▶ Pour le reste...
 - On a un super programme
 - Hâte de vous écouter !!

EXPÉRIENCE 3 : PILOTER L'ÉQUIPE « IA ET DÉVELOPPEMENTS »

Rendez-vous au stand demain !!



35



EPILOGUE

Les deux visages de GRICAD

Image générée avec Microsoft Copilot

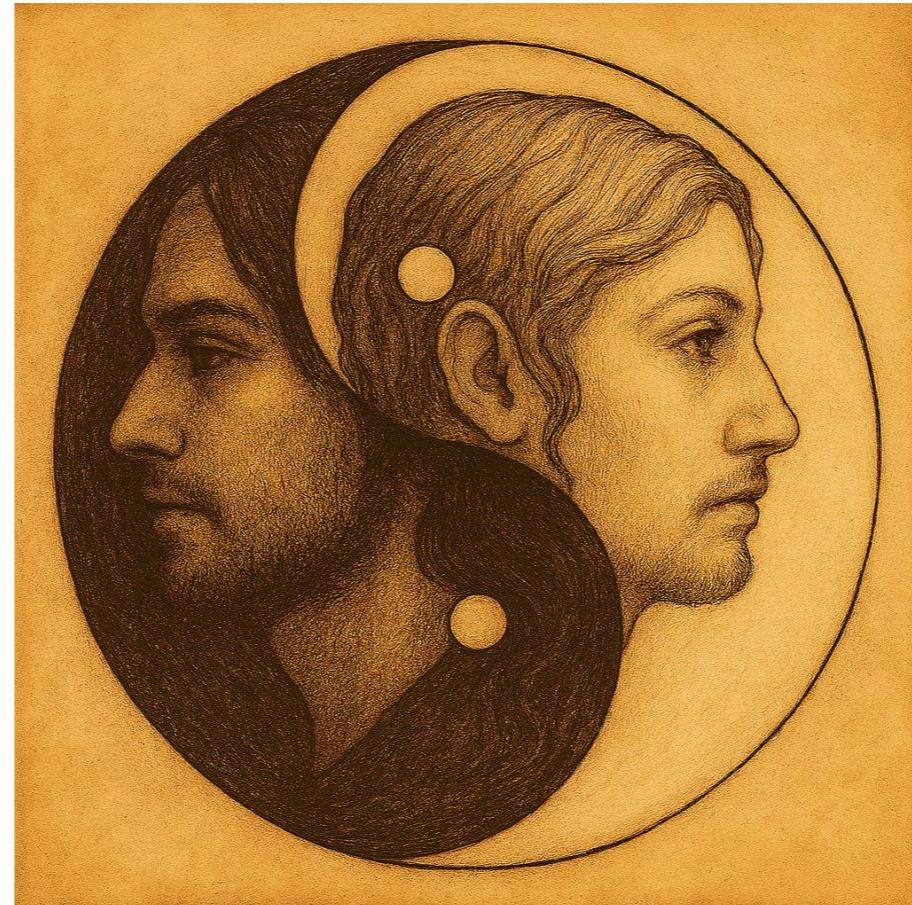
Licence

Creative Commons Attribution (CC BY 4.0)

Description de Copilot:

Cette illustration explore la dualité humaine à travers le symbole du Yin et du Yang. Chaque visage incarne une fusion subtile du masculin et du féminin, soulignant que l'équilibre ne réside pas dans l'opposition, mais dans la complémentarité.

Les traits longs et doux, mêlés à une pilosité discrète, traduisent une identité fluide, universelle, et intemporelle. Le Yin accueille la profondeur et l'introspection, tandis que le Yang rayonne vers l'extérieur et l'action — ensemble, ils forment une unité vivante et mouvante.



La symbolique

- Le Yin et le Yang pour la dualité et la complémentarité
- Des visages androgynes pour la mixité

Mea culpa : l'œil de Françoise



Image générée avec Microsoft Copilot

Licence

Creative Commons Attribution (CC BY 4.0)

Description de Copilot:

Cette illustration explore la dualité humaine à travers le symbole du Yin et du Yang. Chaque visage incarne une fusion subtile du masculin et du féminin, soulignant que l'équilibre ne réside pas dans l'opposition, mais dans la complémentarité.

Les traits longs et doux, mêlés à une pilosité discrète, traduisent une identité fluide, universelle, et intemporelle. Le Yin accueille la profondeur et l'introspection, tandis que le Yang rayonne vers l'extérieur et l'action — ensemble, ils forment une unité vivante et mouvante.



La symbolique

- Le Yin et le Yang pour la complémentarité
- Les 5 faces pour

Oui, mais ...

- ± 5 prompts
 - ± « Générer une image avec DALL-E consomme environ autant d'électricité qu'une charge complète de smartphone, soit environ 0,01 à 0,03 kWh. » [Copilot]
 - <https://arxiv.org/pdf/2311.16863>
- Total de 0,15 kWh

Mea Culpa :

Ce n'était pas nécessaire, je n'aurais pas dû le faire.

Accompagnement sobriété numérique :

gricad-sobriete-numerique@univ-grenoble-alpes.fr

Rapport final du groupe de réflexion Shift Project :

<https://theshiftproject.org>

Merci



39

?

Construction itérative de datasets

