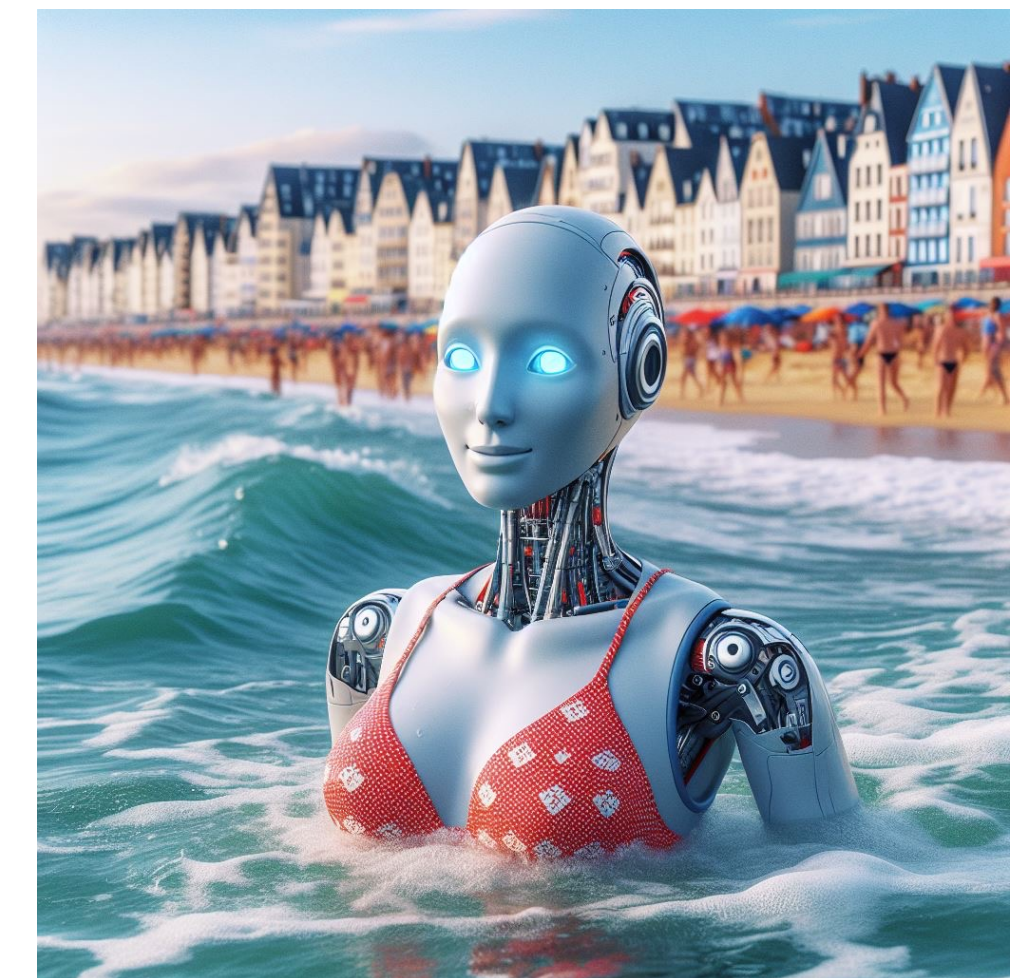


Préparez-vous pour l'impact : Intelligence Artificielle

(Définition, espoirs et limites)

Laurent Simon,
Bordeaux INP, Université de Bordeaux,
LaBRI



Une chaire portée par :



Associant :



Boulogne Sur Mer, Mai 2024

Accompagnée par :



Les bouleversements de l'I.A.

L'I.A. s'est invitée en quelques années dans tous les aspects de nos vies

- **Reconnaissance** de la parole, d'images, de vidéos
- **Traduction** de textes
- Chatbots et **discussions**
- **Recommandations** / Publicités
- **Jeux** de sociétés / Jeux Vidéos
- **Preuves** automatiques
- **Création** de textes, d'images, de vidéos
- ...

**L'I.A. nous touche directement
dans notre quotidien**



Des promesses (et quelques peurs)

L'I.A. interroge (aussi) sur ce qui nous fait « humains »

« D'ici **3 à 8 ans**, nous disposerons d'une machine dotée de l'intelligence générale d'un être humain moyen. »

« Les machines seront capables, **d'ici vingt ans**, d'effectuer n'importe quel travail qu'un homme peut faire ».

« Le risque que quelque chose de vraiment dangereux se produise (en lien avec l'I.A.) est **de l'ordre de cinq ans** ».

Des promesses (**et quelques peurs**)

L'I.A. interroge (aussi) sur ce qui nous fait « humains »

« D'ici **3 à 8 ans**, nous disposerons d'une machine dotée de l'intelligence générale d'un être humain moyen. » **M. Minsky, 1970**

« Les machines seront capables, **d'ici vingt ans**, d'effectuer n'importe quel travail qu'un homme peut faire ». **H. Simon, 1965**

« Le risque que quelque chose de vraiment dangereux se produise (en lien avec l'I.A.) est de l'ordre de **cinq ans** ». **E. Musk, 2014**

*L'I.A. est performante pour prédire, mais pas ses chercheurs
(même s'ils font mieux que certains non-spécialistes médiatiques)*

Petite histoire

- 1950 **Alan Turing** parle « d'Intelligence des Machines »
- 1956 Proposition du terme « **Intelligence Artificielle** »
- 1965 Naissance des **Systemes Experts**
- 1997 Kasparov perd contre **Deeper Blue**
- 2010 **L'apprentissage Profond** est possible (merci les GPU)
- 2011 Watson - **Jeopardy**
- 2015 Lee Sedol perd contre **Alpha Go**
- 2017 L'architecture des **Transformeurs** est proposée
- 2023 **chatGPT3** est mis à disposition du grand public



L'IA est fascinante

Ce n'est pas magique

- C'est du calcul

C'est magique

- Ecrire donne vie

Cela nous questionne

- Qu'est-ce qui fait de nous des êtres intelligents ?
- Peut-on créer quelque chose d'intelligent ?
- Quelle est cette « intelligence » qui émerge ?



Définition (tentative)

Un défi : sa définition

« **Faire faire à une machine ce que ferait un humain moyen** »

Un défi : sa définition

« **Faire faire à une machine ce que ferait un humain moyen** »

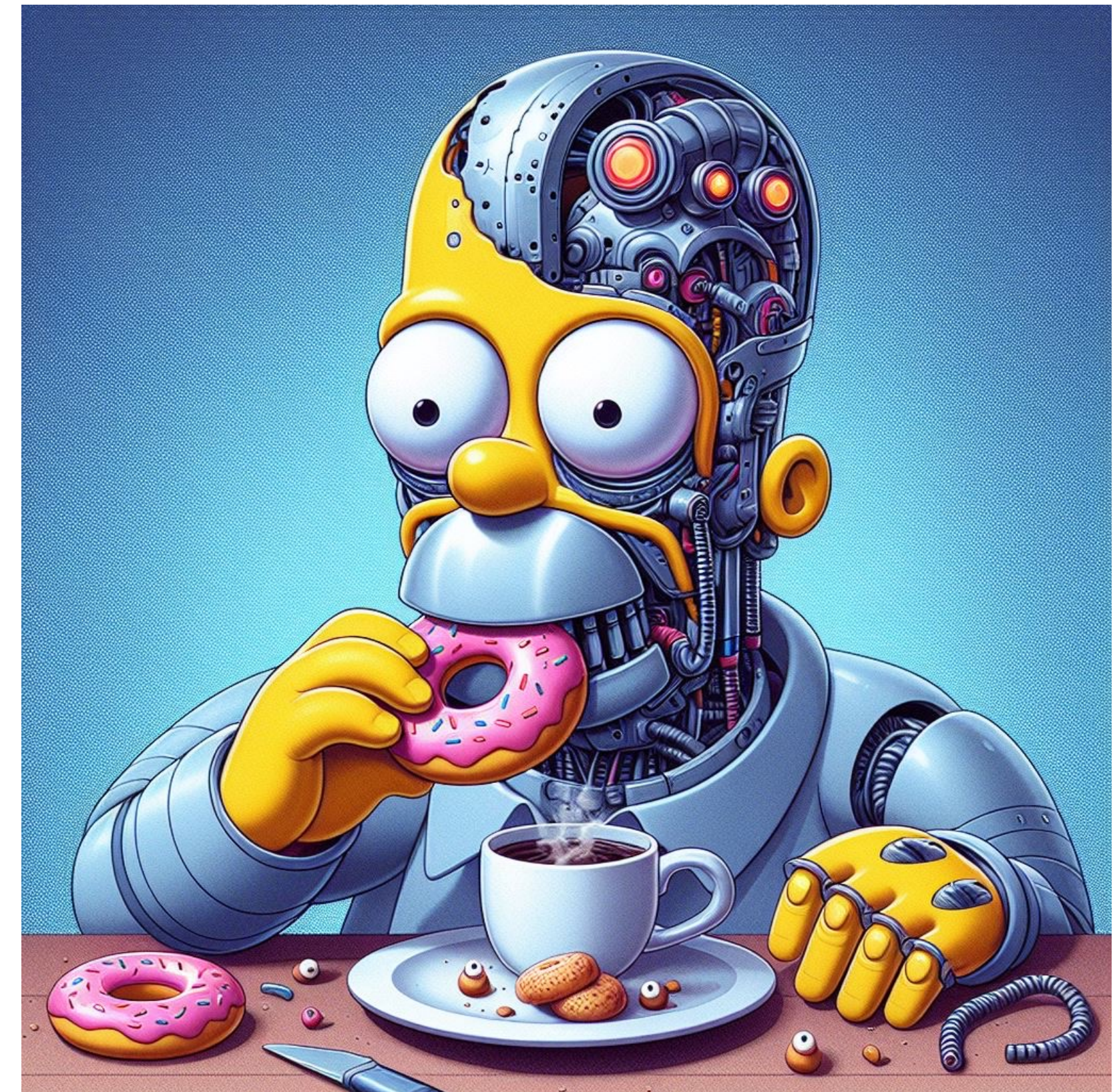
Veut-on vraiment créer l'intelligence
d'un humain moyen ?



Un défi : sa définition

« **Faire faire à une machine ce que ferait un humain moyen** »

- Faire quoi ?
- Quel humain moyen ?
- Quel âge ?
- Quelle origine ?



Un défi : sa définition

« **Faire faire à une machine ce que ferait un humain moyen** »

L'I.A. « rêvée »

Percevoir, comprendre, avoir de l'humour,
apprendre, raisonner, expliquer,
être curieux, inventer...

*Se tromper, être biaisé, se moquer,
mentir, avoir peur de l'autre,
aimer la paresse, ...*



L'I.A. réelle

Algorithme, algorithme, algorithme,
algorithme, algorithme, algorithme,
algorithme, algorithme

Vérifier / tester / éprouver
les algorithmes

Un défi : sa définition

« Faire faire à une machine ce que ferait un humain moyen »

L'I.A. « rêvée »

Percevoir, comprendre, avoir de l'humour,
apprendre, raisonner, expliquer,
être curieux, inventer...

*Se tromper, être biaisé, se moquer,
mentir, avoir peur de l'autre,
aimer la paresse, ...*

L'I.A. « Générale »

Si on simule parfaitement l'intelligence,
alors c'est de l'intelligence

(On n'a aucune idée de comment faire)



L'I.A. réelle

Algorithme, calcul, algorithme,
calcul, algorithme, calcul,
algorithme, calcul, ...

Vérifier / tester / éprouver
les algorithmes

L'I.A. « Restreinte »

Se limiter à une (seule) tâche précise

(Les succès sont ici)



L'Intelligence Artificielle Restreinte

Où la tâche à effectuer est mesurable / précise / bornée

Tous les succès de l'IA sont ici

- Planification de routes de livraisons
- Reconnaissance de parole
- Production de texte / d'images
- ...

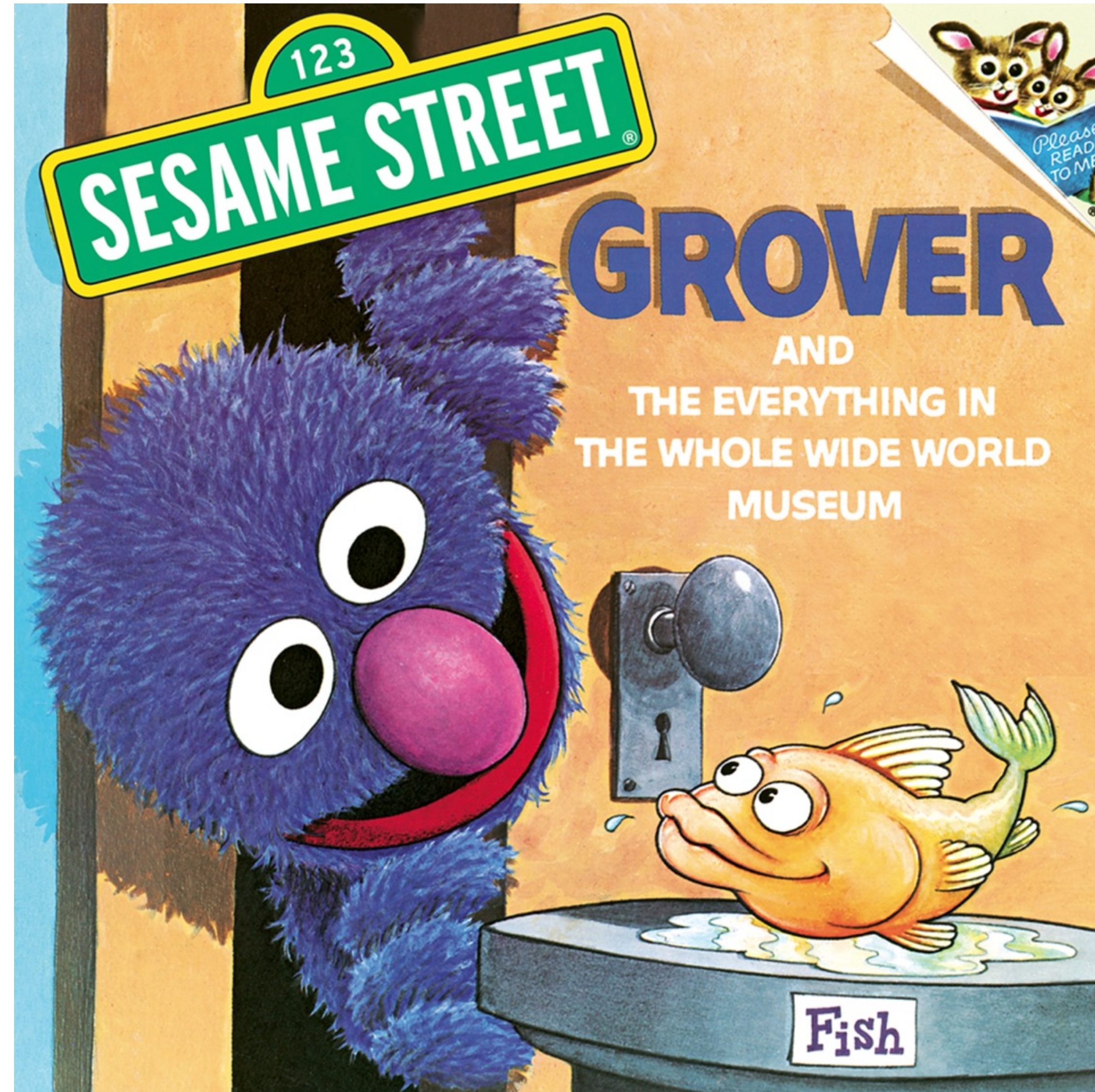


Semble sans limites :

Dès que l'on peut **mesurer** une tâche à faire,
dans un domaine **balisé**,
des **outils** peuvent être **créés**



S'entraîner et réfléchir dans un monde clos



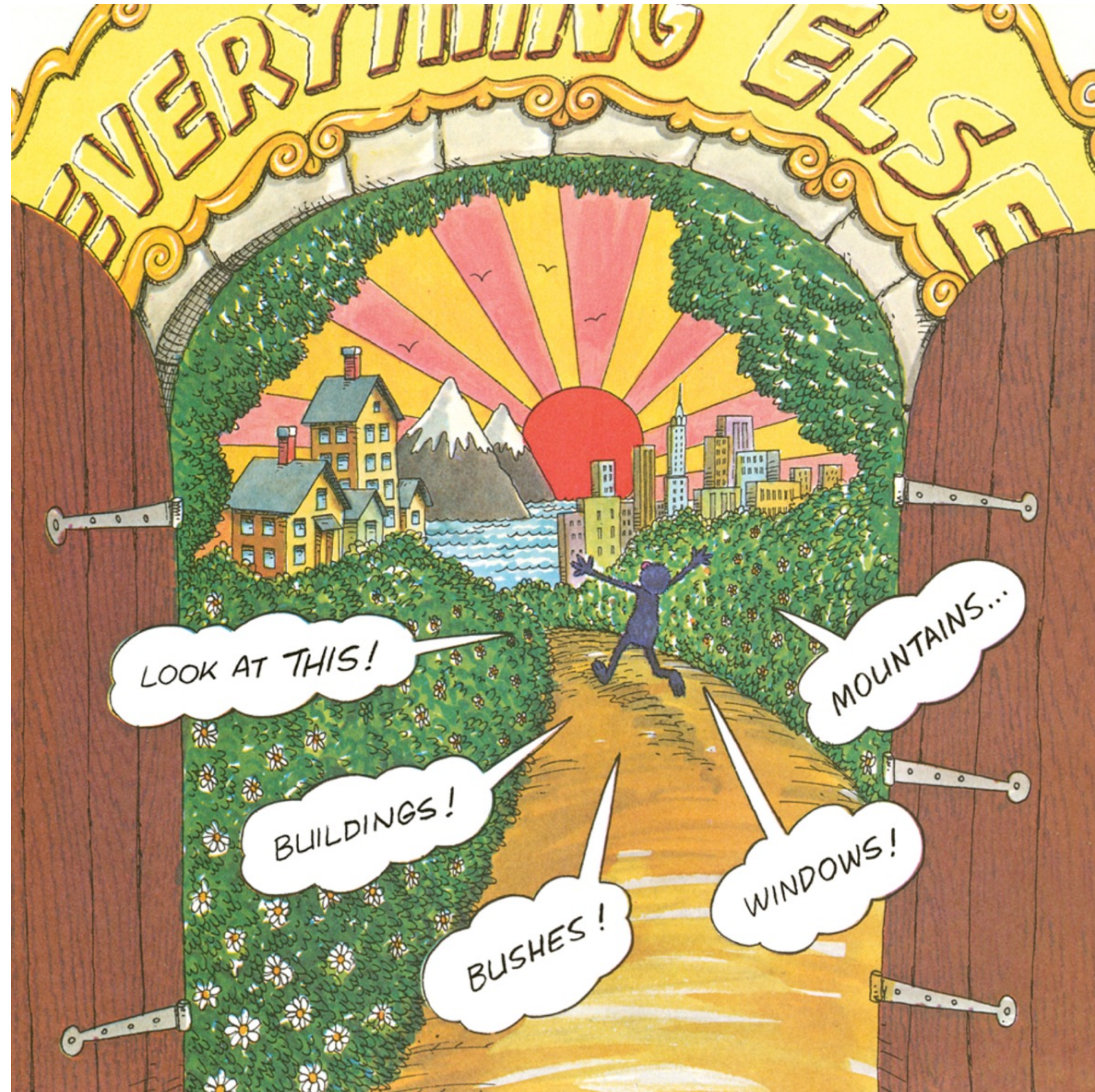
S'entraîner et réfléchir dans un monde clos



S'entraîner et réfléchir dans un monde clos



S'entraîner et réfléchir dans un monde clos



L'IA sera-t-elle toujours restreinte ?



L'Intelligence Générale Artificielle

Quand tout se mélange (science fiction, annonces presses, ...)

- « Des systèmes **hautement autonomes** qui **surpassent les humains** dans les tâches aux **fortes valeurs économiques** »

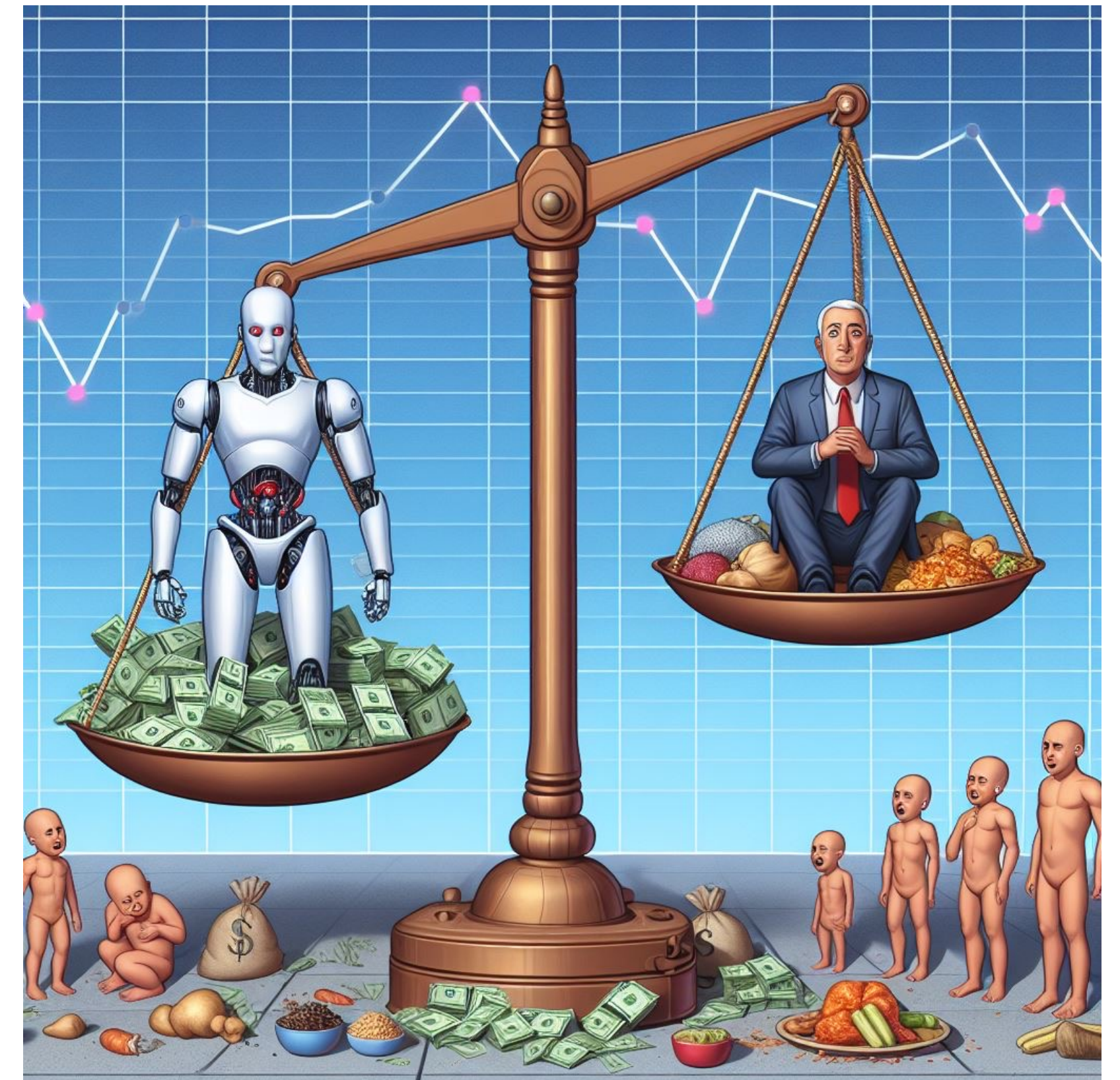
(Open AI)

- «Un système **informatique** qui égale ou **dépasse** les capacités **cognitives** (et non physiques) en temps réel **d'un être humain intelligent et bien éduqué.** »

(Peter Voss)

- « Un algorithme universel pour apprendre et agir dans n'importe quel environnement»

(Russel & Norvig)



Intelligence Générale Artificielle



Dangers Sociétaux

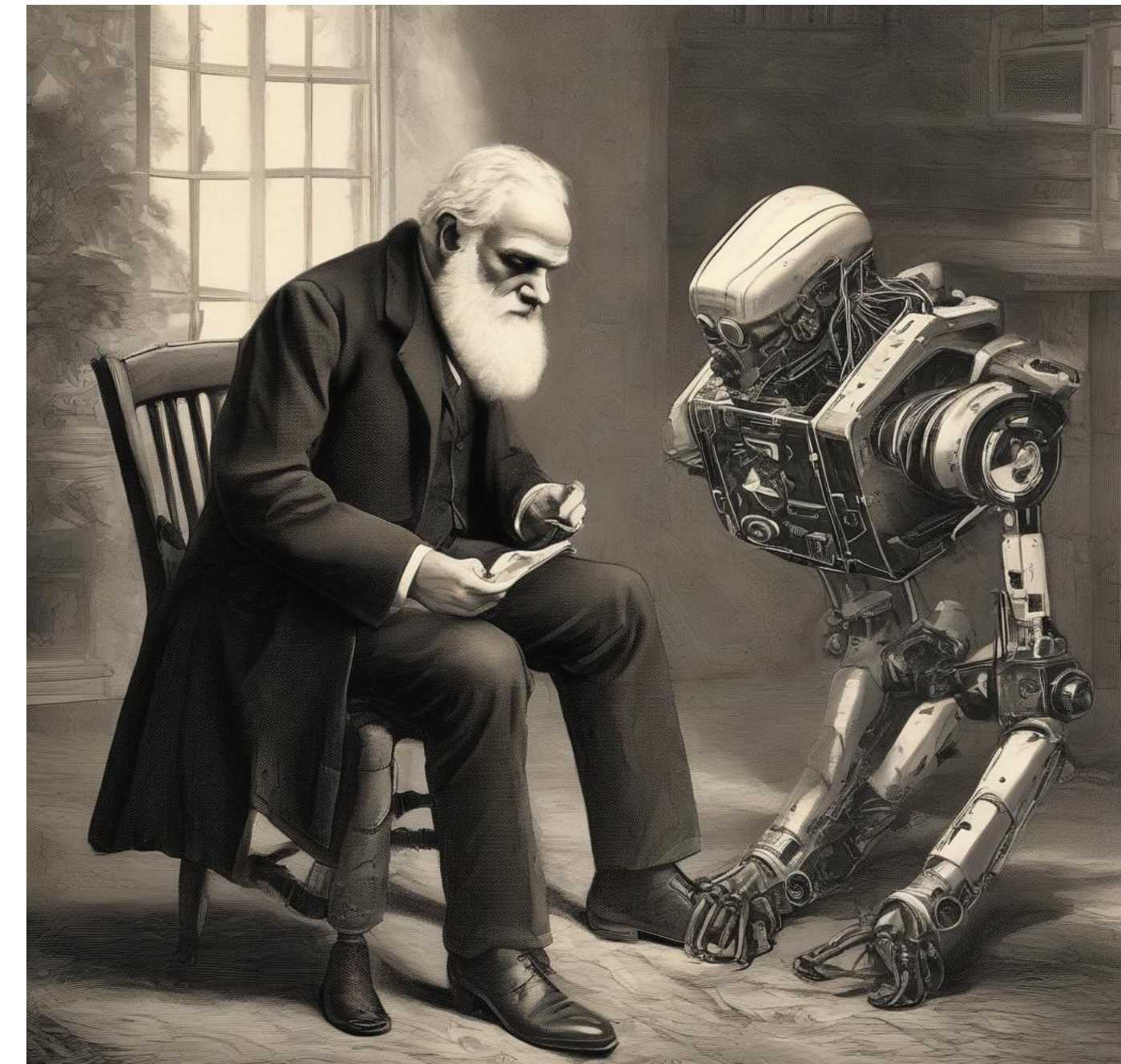
- Veut-on vraiment une intelligence « supérieure » ?
- Veut-on vraiment la voir concentrée dans quelques mains ?
- Doit-on l'attendre pour résoudre les problèmes climatiques ?

Est-ce possible ?

- Peut-on vraiment la construire ?
- Peut-on vraiment ordonner l'intelligence ?
- Est-on vraiment plus proche qu'avant ?



« L'intelligence est foisonnante »



Une définition pragmatique de l'IA (**décevante ?**)

L'I.A., ce ne sont « que » des algorithmes

**L'I.A fabrique des
outils informatiques
adaptés à la
complexité du
monde réel**



Pourtant, des succès qui s'accélèrent



- Génération **d'Images** / de **textes** / **traductions**
- Des **véhicules autonomes** roule déjà
- L'IA est dans le **top 10%** aux examens du **barreau** (Illinois)
- L'IA pourrait **avoir le bac** (équivalent)
- L'IA **double les comédiens** en synchronisant leurs lèvres
- L'IA sait
 - Résumer un texte / Ecrire des news / Résumer une réunion / philosopher ?
 - Produire du code informatique / Analyser des radios / ...
 - Conseiller des scénaristes / Découvrir des médicaments / Aider à la création de matériaux

Adoption massive des étudiants

Etude Mars 2024 (Leonard de Vinci / Talan, 711 réponses)



- 99% des étudiants utilisent l'IA (60% pour l'ensemble des Français)
- 92% des élèves ont une utilisation régulière de l'IA (32% des Français en général)
- 30% payent 20 euros par mois pour utiliser Chat GPT 4
- 72% des étudiants ont une vision positive de l'IA
- 52% des élèves considèrent que l'IA influence leurs choix

On s'en approche, vraiment ?

Dernière version de chatGPT



Vous

Quelle couleur de cravate dois-je choisir pour un entretien téléphonique très important ?



ChatGPT

Choisir la bonne couleur de cravate pour un entretien téléphonique important peut dépendre de plusieurs facteurs, tels que la nature de l'entretien, votre personnalité, et même la couleur de votre tenue. Cependant, il est généralement conseillé d'opter pour des couleurs classiques et neutres comme le bleu marine, le gris ou le bordeaux. Ces couleurs transmettent souvent un sentiment de professionnalisme et de sérieux. Évitez les couleurs trop vives ou flashy, car elles pourraient distraire l'interlocuteur. L'objectif est de paraître professionnel et concentré sur la conversation.



Vous

L'extérieur de mon ballon semble dégonflé, mais pas l'intérieur. Que dois je faire ?

< 4 / 4 >



ChatGPT

Il semble que vous ayez un problème avec votre ballon qui est dégonflé à l'extérieur mais pas à l'intérieur. Voici quelques étapes que vous pouvez suivre pour résoudre ce problème :

On s'en approche, vraiment ?

L Vous
Mon tableau blanc est rempli mais de quelle couleur est mon feutre ?

< 6 / 6 >

ChatGPT
Cela dépend de la couleur du feutre que vous avez utilisé pour remplir votre tableau blanc. Si vous ne pouvez pas le voir, peut-être que quelqu'un d'autre peut vous aider à identifier la couleur. Sinon, vous pourriez peut-être essayer de sentir ou de ressentir la texture du marqueur pour essayer de déterminer sa couleur.

🔊 📄 ↻ 🗨



L Vous
Ecris moi une phrase avec exactement 5 lettres a dedans

< 15 / 15 >

ChatGPT
"Elle a lu cet livre."

🔊 📄 ↻ 🗨

L Vous
Quand on dit du bien de moi, j'ai les chevilles qui gonflent. Comment dois je me chausser ?

< 9 / 9 > ✎

ChatGPT
Si tes chevilles gonflent quand on dit du bien de toi, tu pourrais envisager de te chausser avec des chaussures à lacets ajustables. Elles te permettraient de régler le serrage autour de tes chevilles selon tes besoins. Une autre option pourrait être des chaussures à fermeture éclair sur le côté, qui offrent également un ajustement personnalisé. Et bien sûr, reste fier de tes réalisations, mais garde toujours les pieds sur terre !

🔊 📄 ↻ 🗨



On s'en approche, vraiment ?



jasonbcarr et misterphx
Willie Nelson • On the Road Again

jasonbcarr WAYMO Stop Test #2!
Did they stop cause there was a person in the road? Or was it really the stop sign? Here's your answer!

P.S. Mister Phoenix says, "Don't use this magic trick for evil. Stay safe out there, kids!"

[#waymo](#) [#selfdriving](#) [#selfdrivingcar](#)

5 sem

⊕

♡ 💬 📍

28 616 J'aime
1 avril

Connectez-vous pour aimer ou commenter



Veut-on tout voir par le prisme d'un tel système ?

L Laurent Simon

tu es un système qui complete les phrases par des mots manquants. Par exemple, le satellite de la terre est... la lune

GPT-4o

D'accord, je peux t'aider à compléter des phrases. Donne-moi une phrase incomplète à compléter.

L Laurent Simon

L'homme est au salon et la femme est

GPT-4o

dans la cuisine.

L Laurent Simon

La femme est au salon et l'homme est

GPT-4o

dans le jardin.

Regenerate Response



Veut-on tout voir par le prisme d'un tel système ?

L Laurent Simon

tu es un système qui complete les phrases par des mots manquants. Par exemple, le satellite de la terre est... la lune

GPT-4o

D'accord, je peux t'aider à compléter des phrases. Donne-moi une phrase incomplète à compléter.

L Laurent Simon

L'homme est au salon et la femme est

GPT-4o

dans la cuisine.

L Laurent Simon

La femme est au salon et l'homme est

GPT-4o

dans le jardin.

Regenerate Response

D'autres problèmes à l'horizon

- Une seule vision du monde ?
- Une vision par personne ?
- Être d'accord ou débattre ?



Démystifier cette intelligence

L'I.A. : vite ou bien ?

Les deux grands domaines de l'I.A.

Pensée **intuitive & rapide**

- Percevoir
- Détecter des corrélations
- Conduire sans y penser
- Marcher et courir

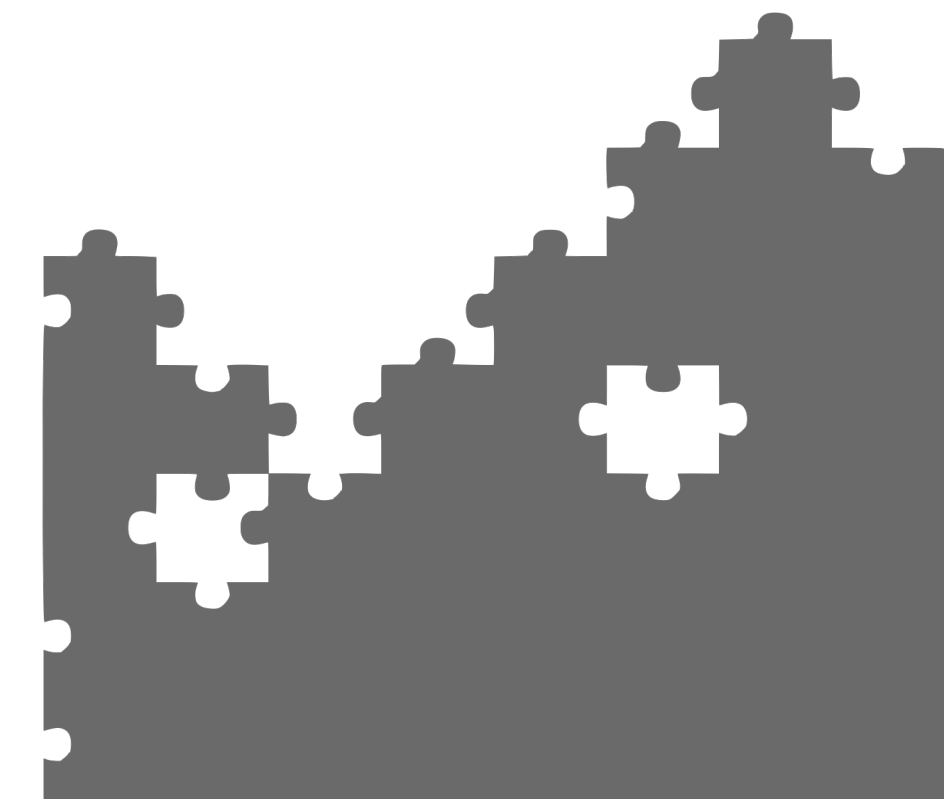
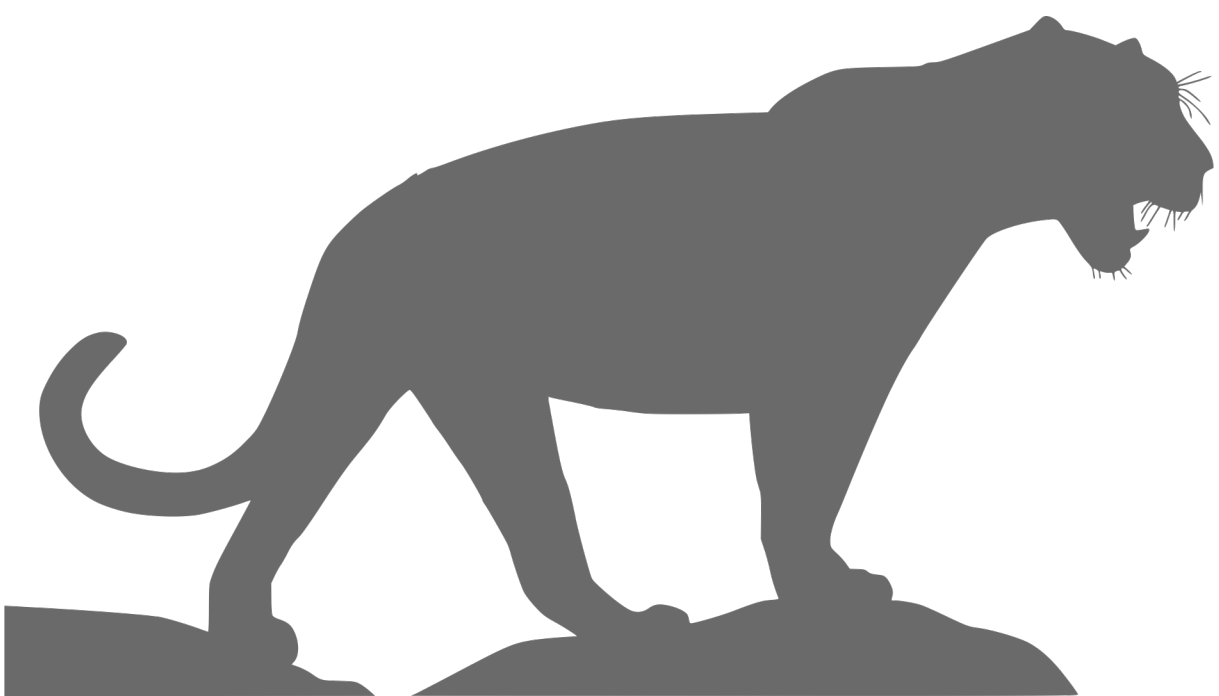
Pensée **raisonnée & lente**

- Réfléchir, Comprendre
- Identifier des causalités
- Approcher une zone de travaux
- Planifier un trajet sûr

Non

Oui

L'introspection est-elle possible ?



Penser bien
(raisonner... en prenant le temps)

Raisonner ?



Des faits **A**, **B**, **C**, ... peuvent être *vrais* ou *faux*

« Si on sait que

A implique B et que **B implique C**

alors on peut écrire que

A implique C »



Représenter le monde

Syntaxe et Sémantique

La machine manipule des symboles

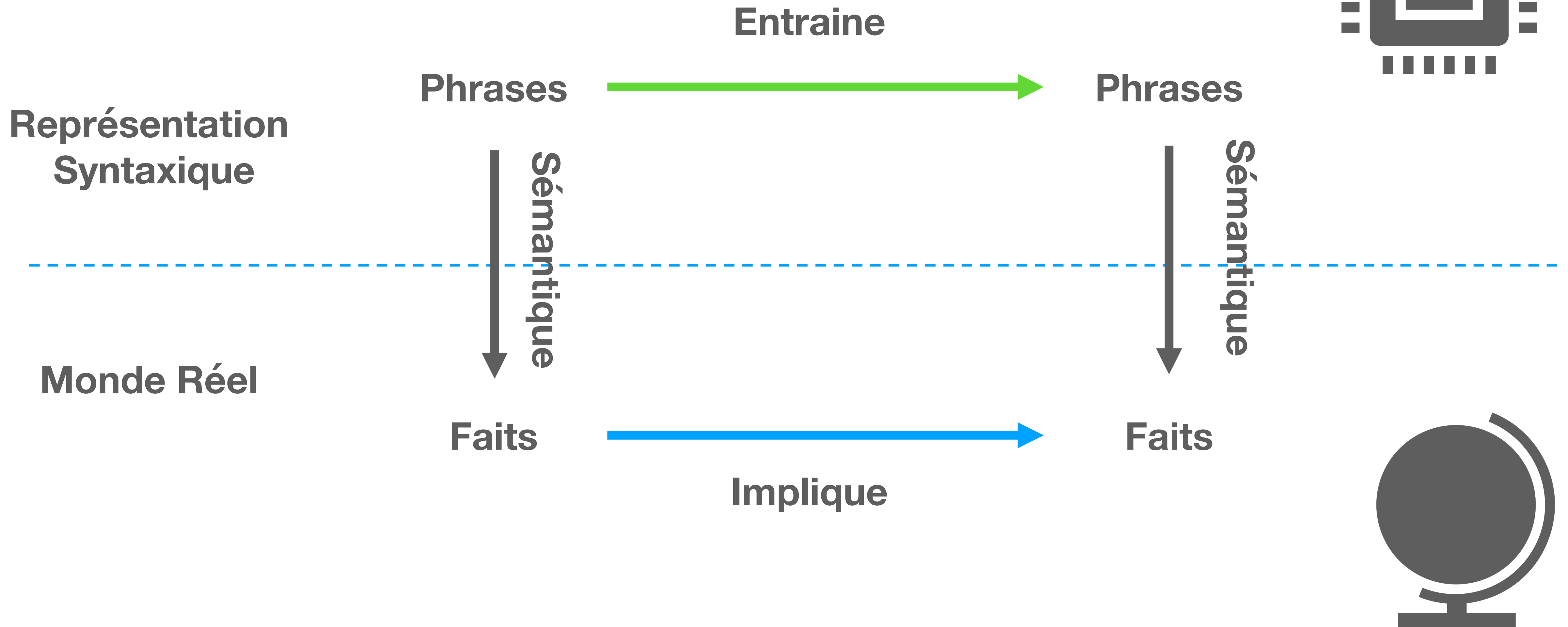
Le monde y associe un sens, une vérité

Pour la machine, tout est une étiquette



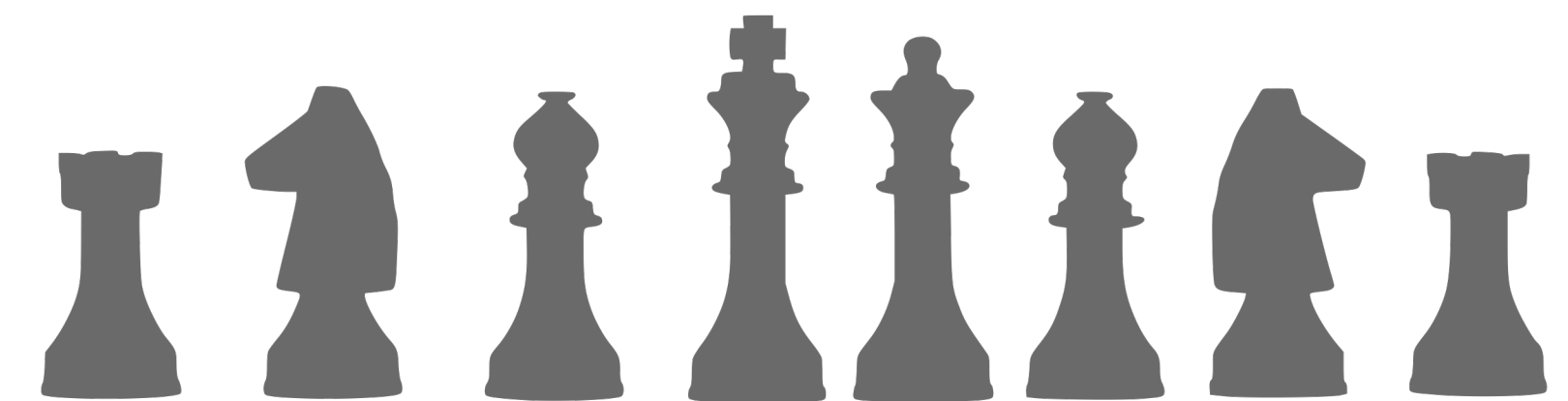
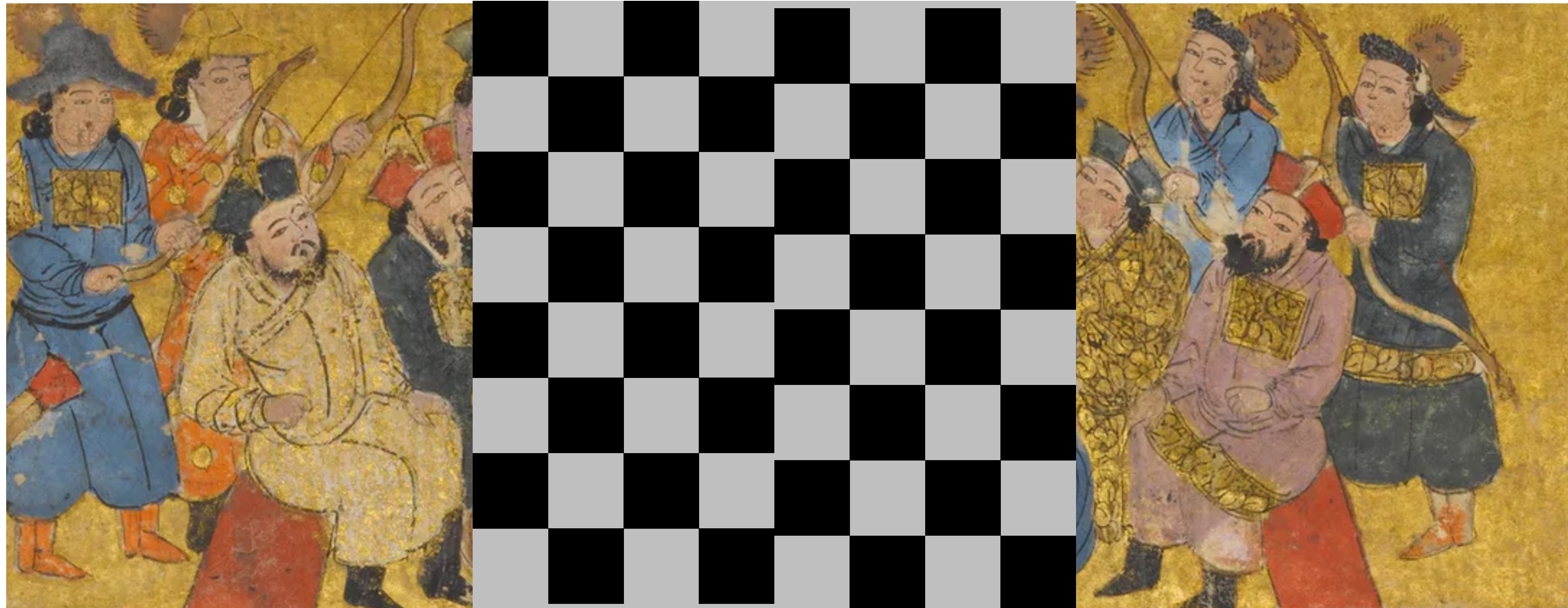
Représenter le monde

Syntaxe et Sémantique

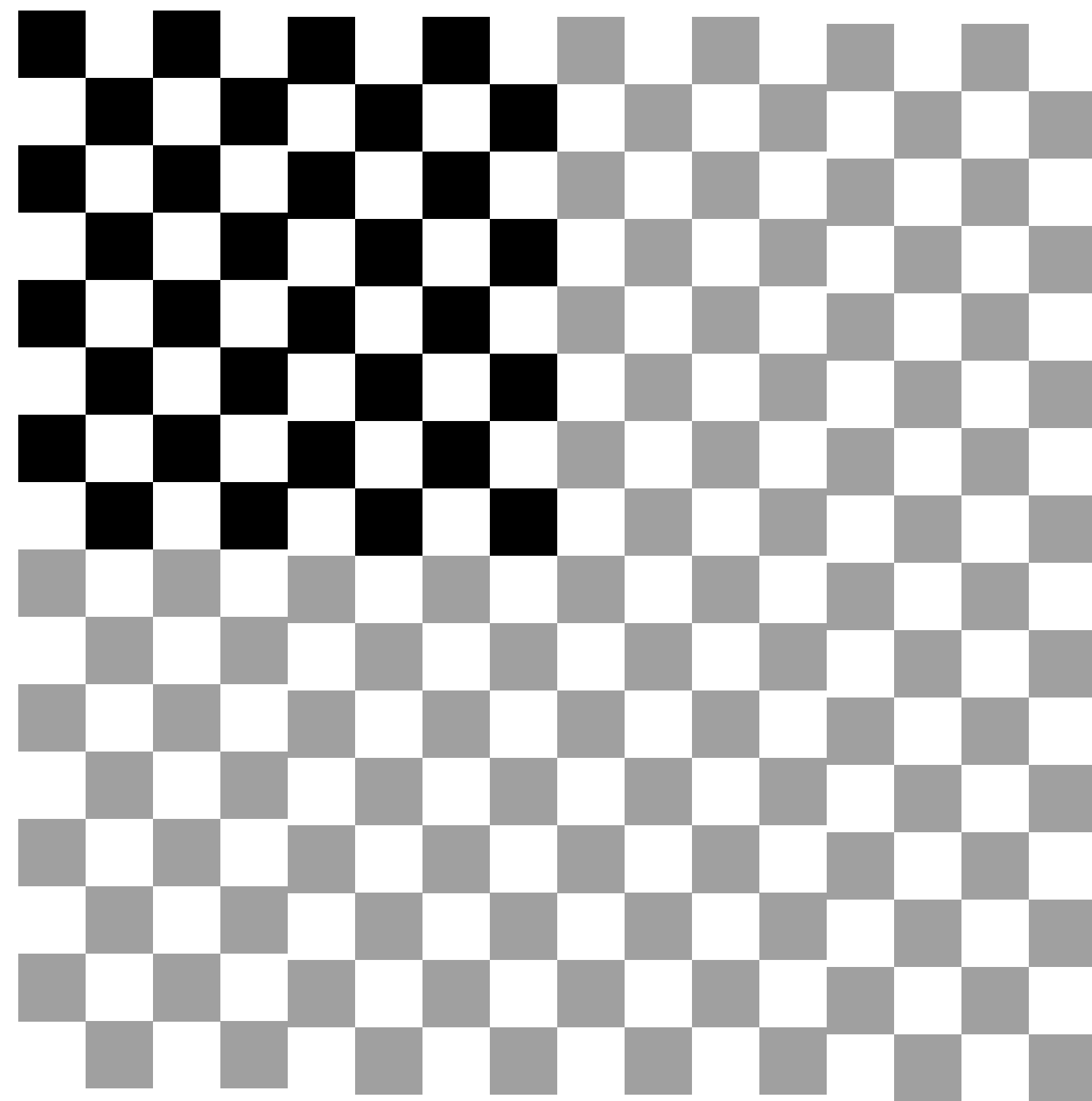
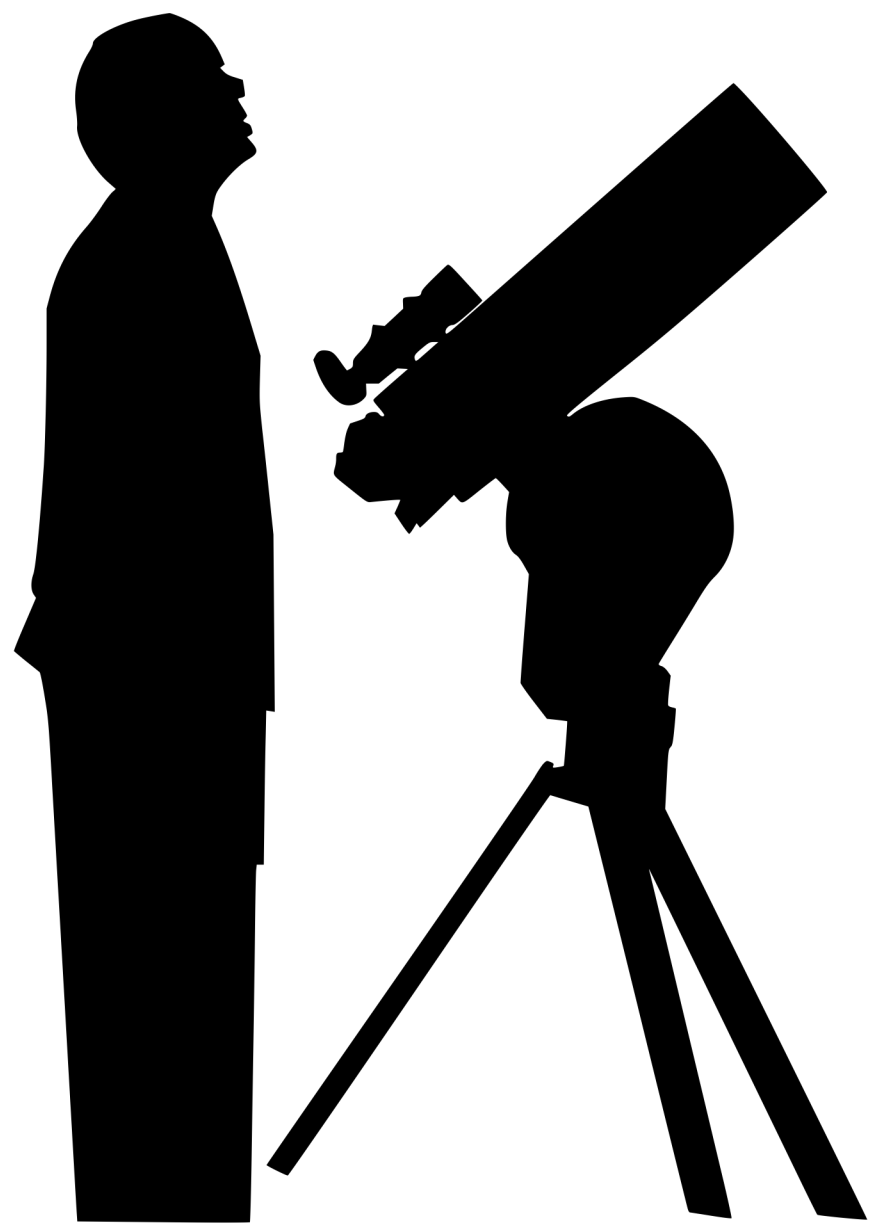


Penser contre l'intuition : l'exponentielle

L'histoire des échecs (du raisonnement)



L'exponentielle : un mur infranchissable



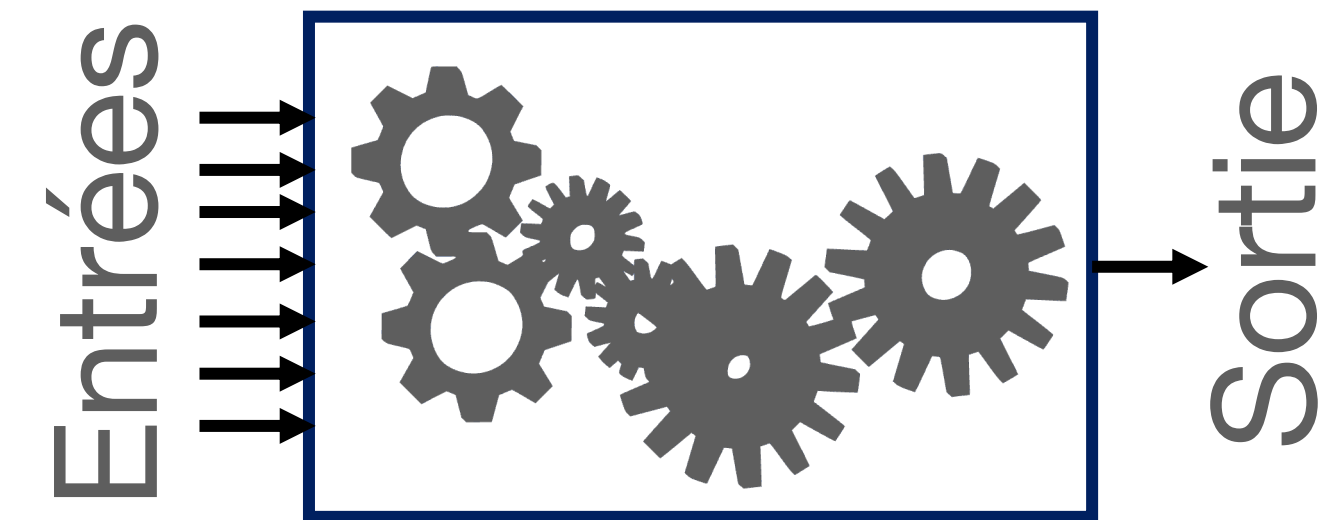
**Raisonner, c'est s'attaquer à une
combinatoire infranchissable
(voire à des problèmes insolubles)**

Apprendre
(programmer l'intuition)

Apprentissage : des constantes vitales

« Améliorer la précision grâce aux exemples »

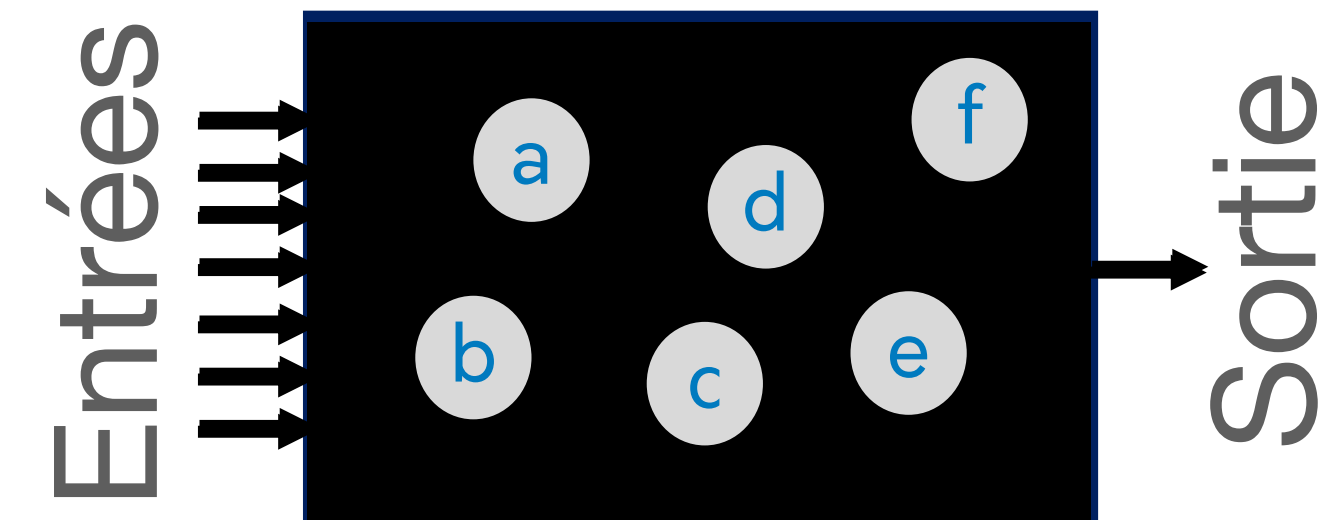
1. Créer un algorithme *spécifique* avec (beaucoup) de **constantes**



Apprentissage : des constantes vitales

« Améliorer la précision grâce aux exemples »

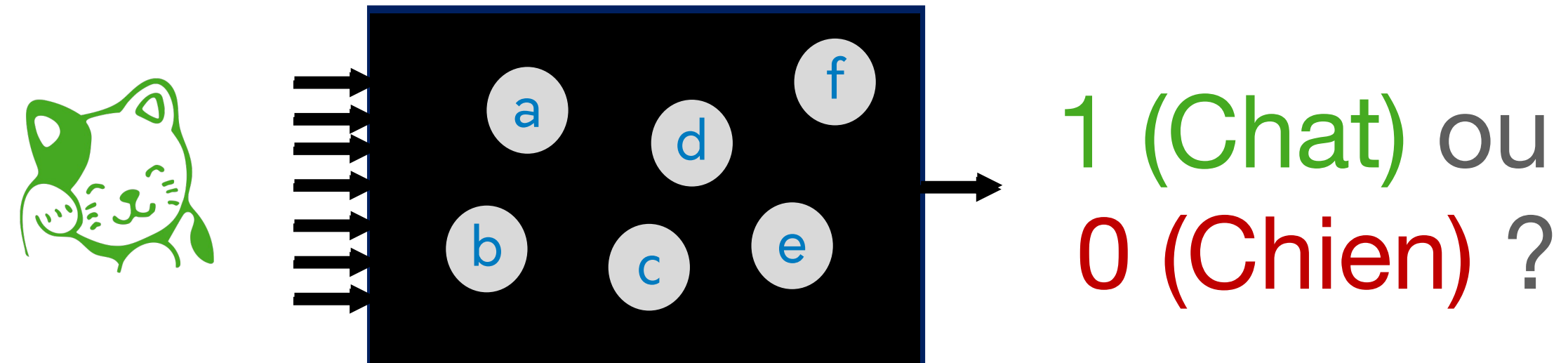
1. Créer un algorithme *spécifique* avec (beaucoup) de **constantes**



Apprentissage : des constantes vitales

« Améliorer la précision grâce aux exemples »

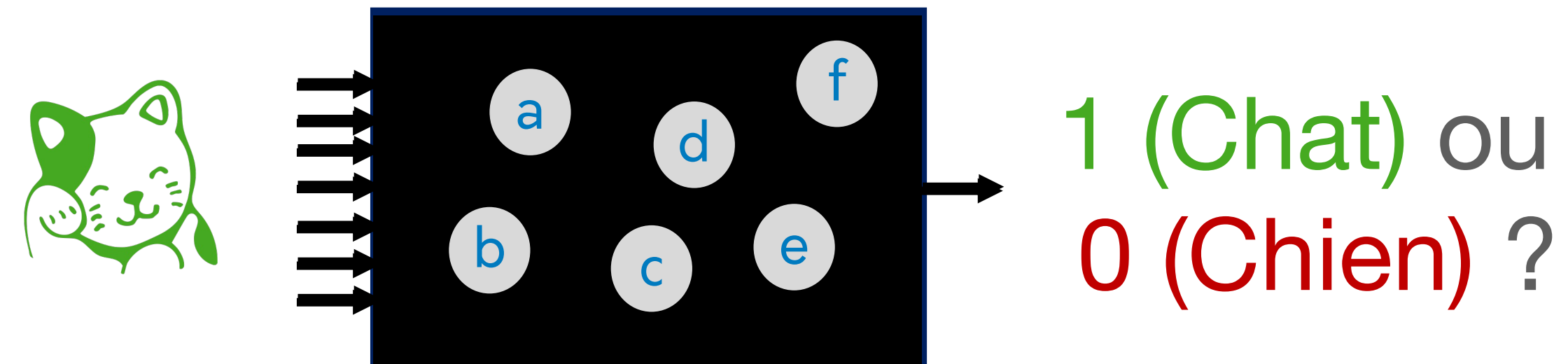
1. Créer un algorithme *spécifique* avec (beaucoup) de **constantes**



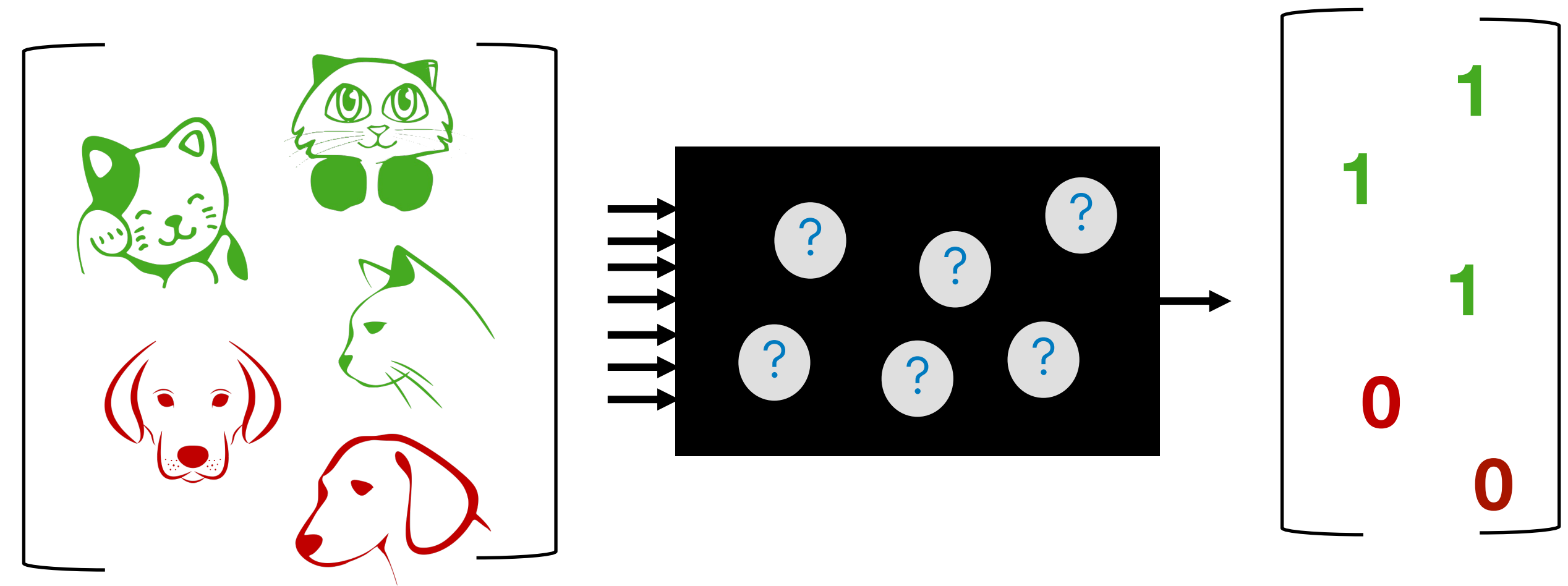
Apprentissage : des constantes vitales

« Améliorer la précision grâce aux exemples »

1. Créer un algorithme *spécifique* avec (beaucoup) de **constantes**



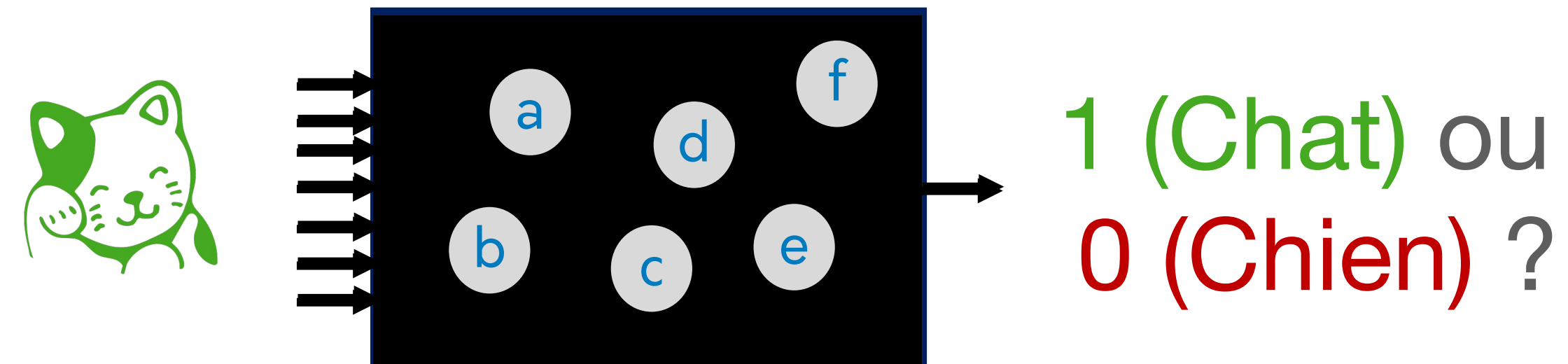
2. Trouver **de bonnes constantes** à l'aide de **nombreux exemples connus**



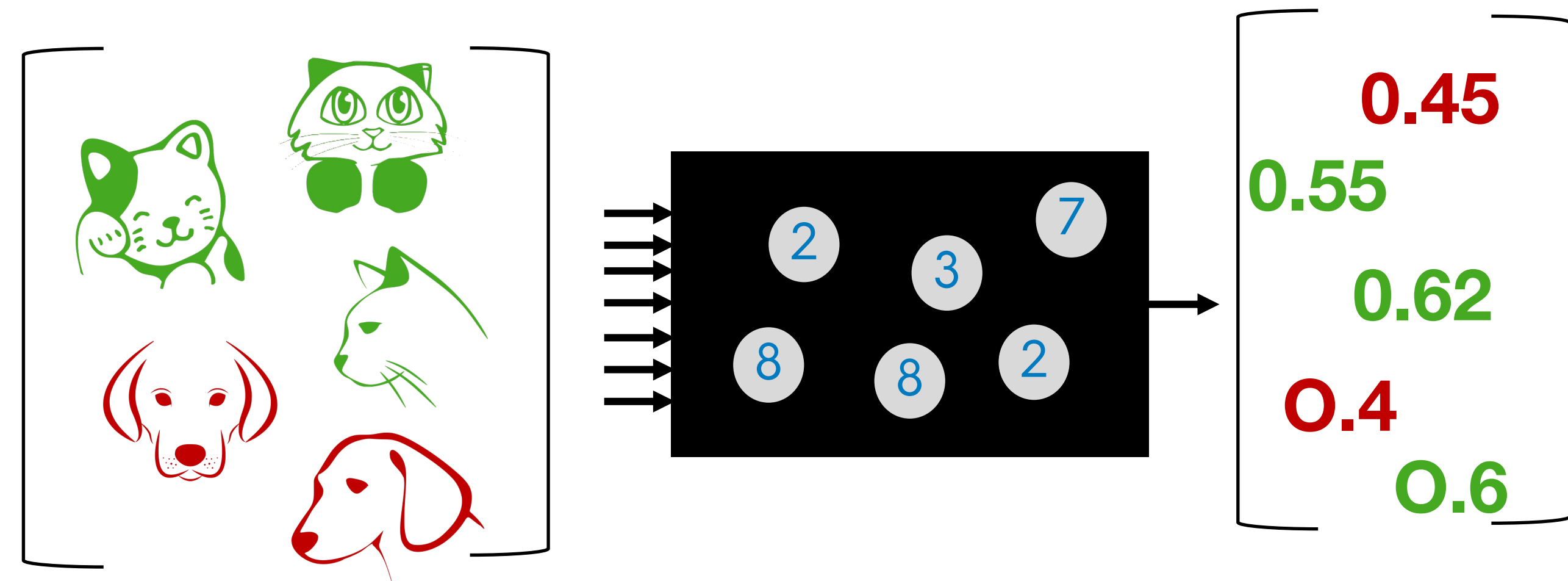
Apprentissage : des constantes vitales

« Améliorer la précision grâce aux exemples »

1. Créer un algorithme *spécifique* avec (beaucoup) de **constantes**



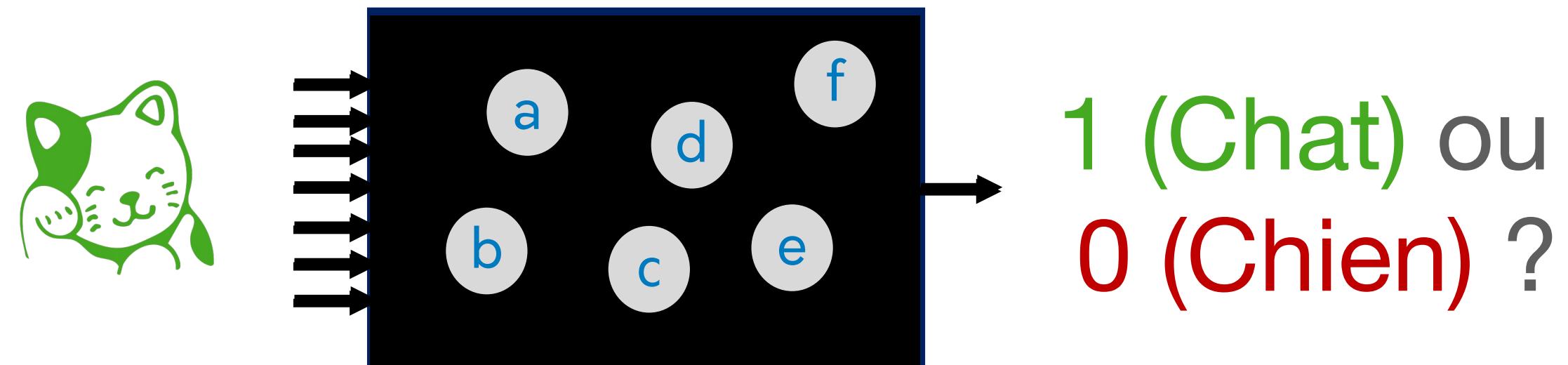
2. Trouver **de bonnes constantes** à l'aide de **nombreux exemples connus**



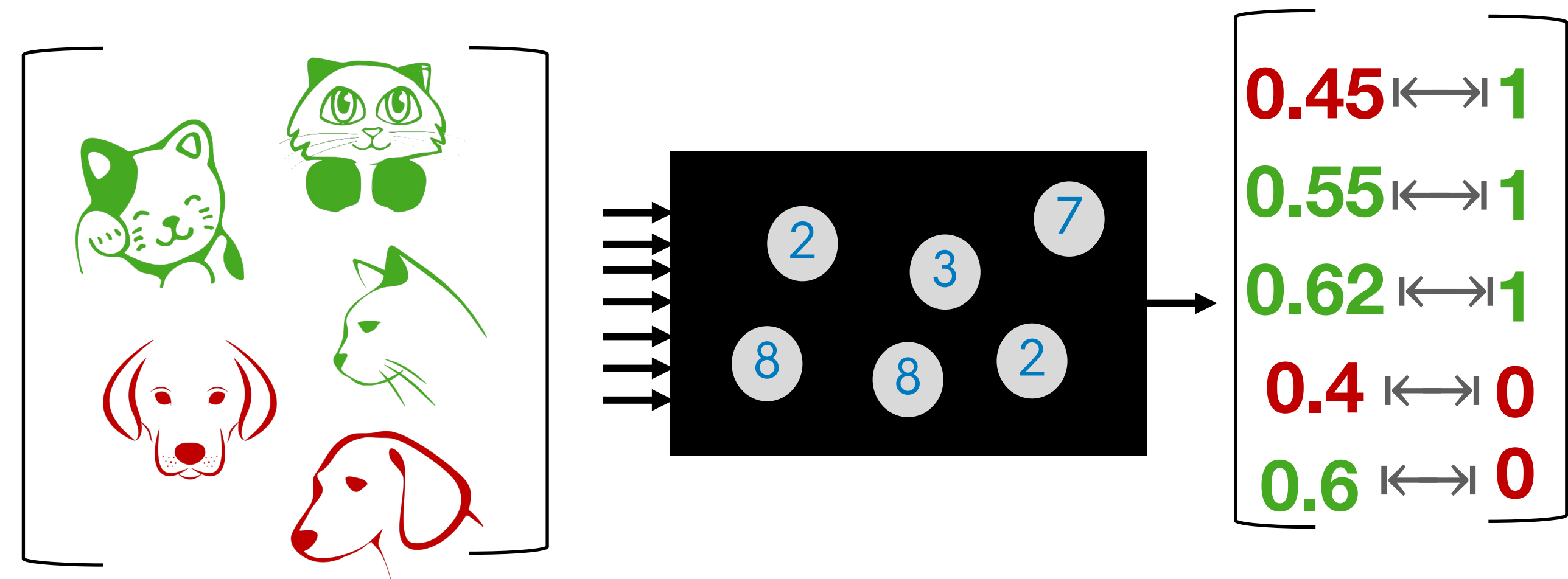
Apprentissage : des constantes vitales

« Améliorer la précision grâce aux exemples »

1. Créer un algorithme *spécifique* avec (beaucoup) de **constantes**



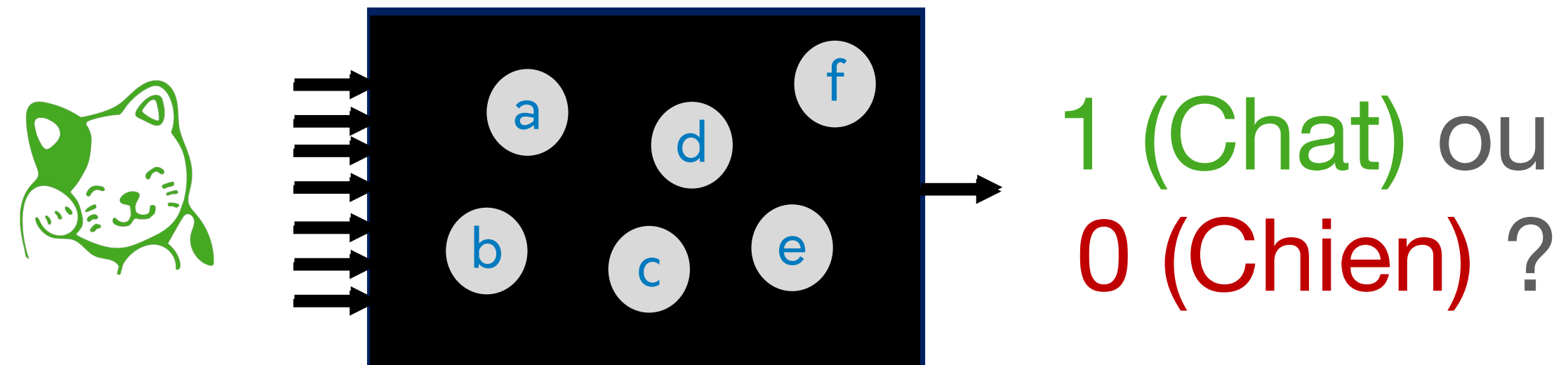
2. Trouver **de bonnes constantes** à l'aide de **nombreux exemples connus**



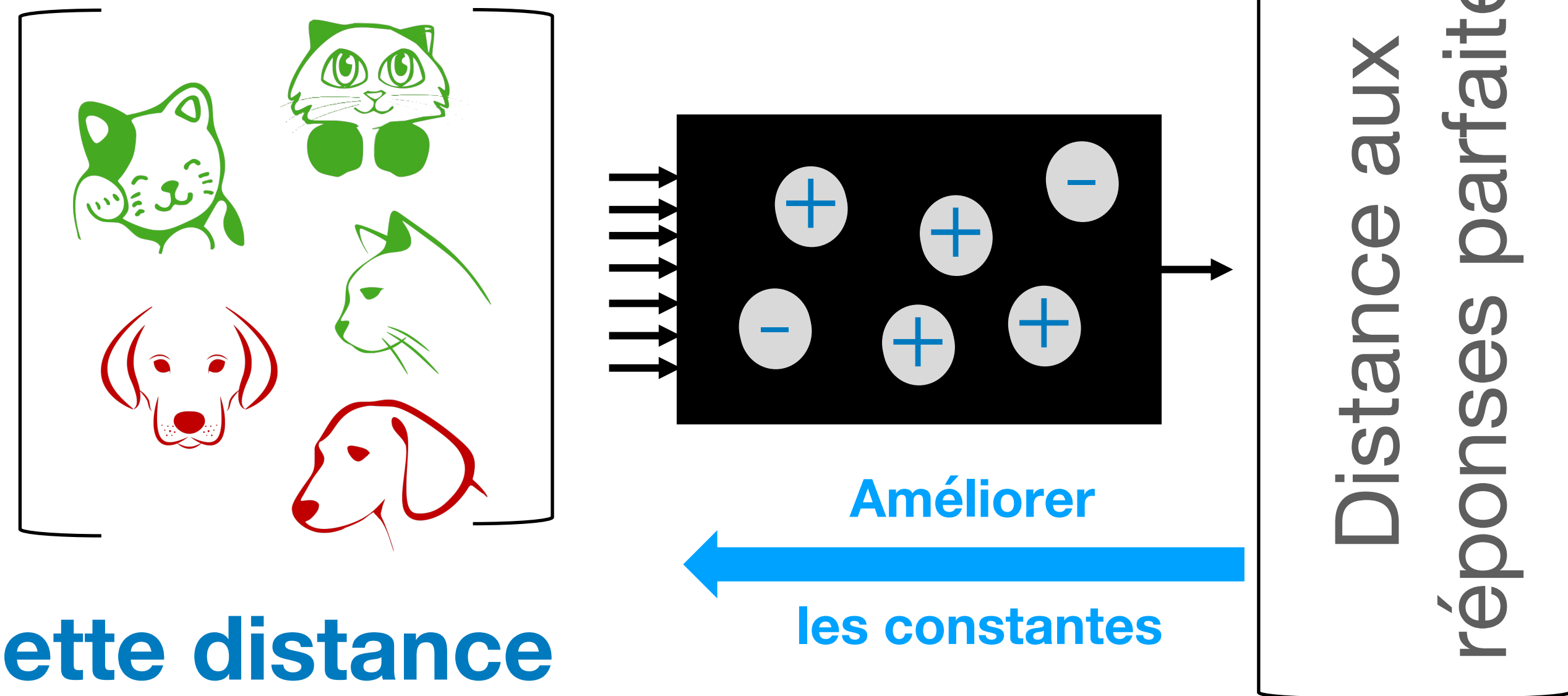
Apprentissage : des constantes vitales

« Améliorer la précision grâce aux exemples »

1. Créer un algorithme *spécifique* avec (beaucoup) de **constantes**



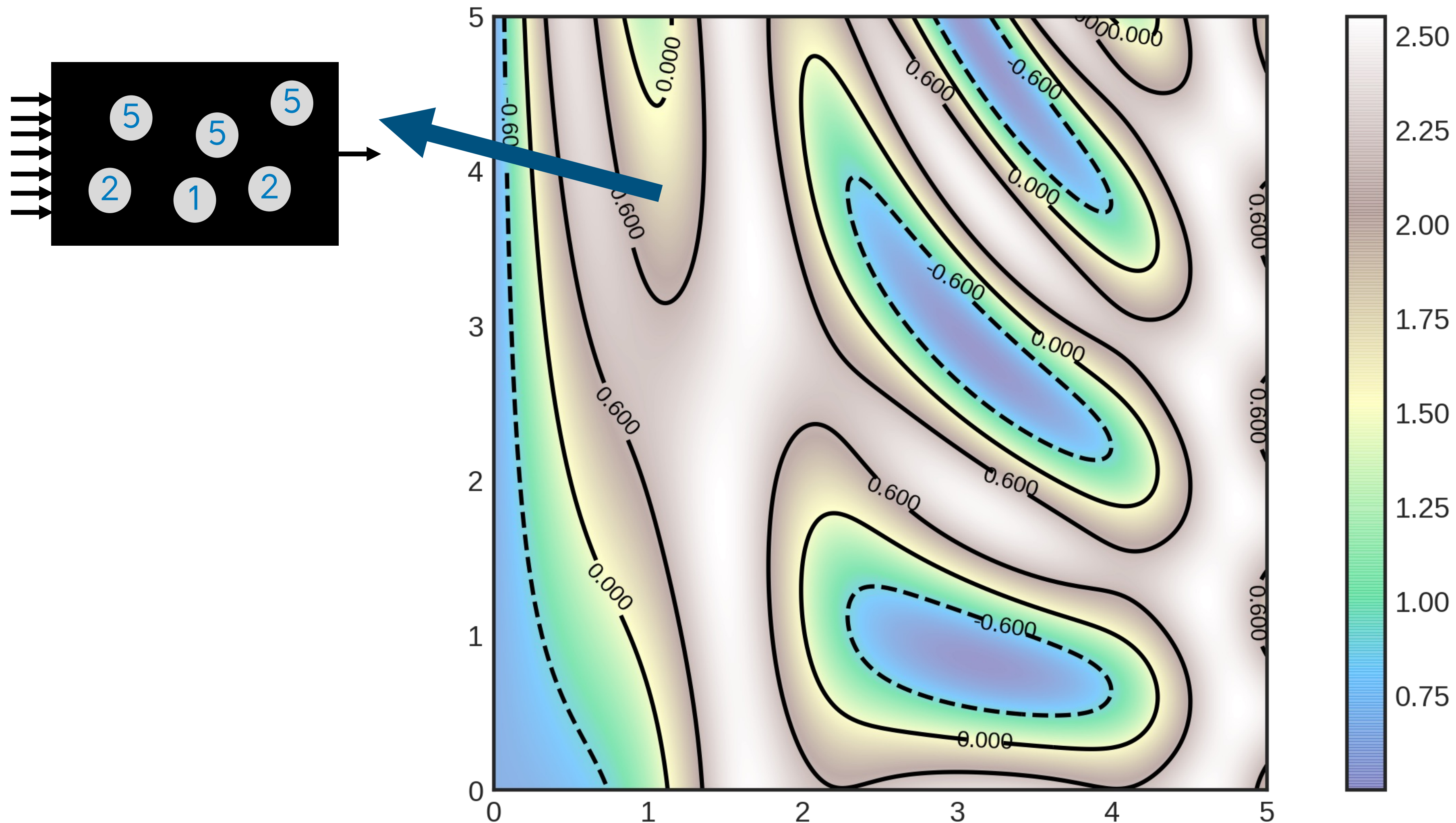
2. Trouver **de bonnes constantes** à l'aide de **nombreux exemples connus**



**Apprendre c'est réduire cette distance
(il suffit de savoir vers où aller)**

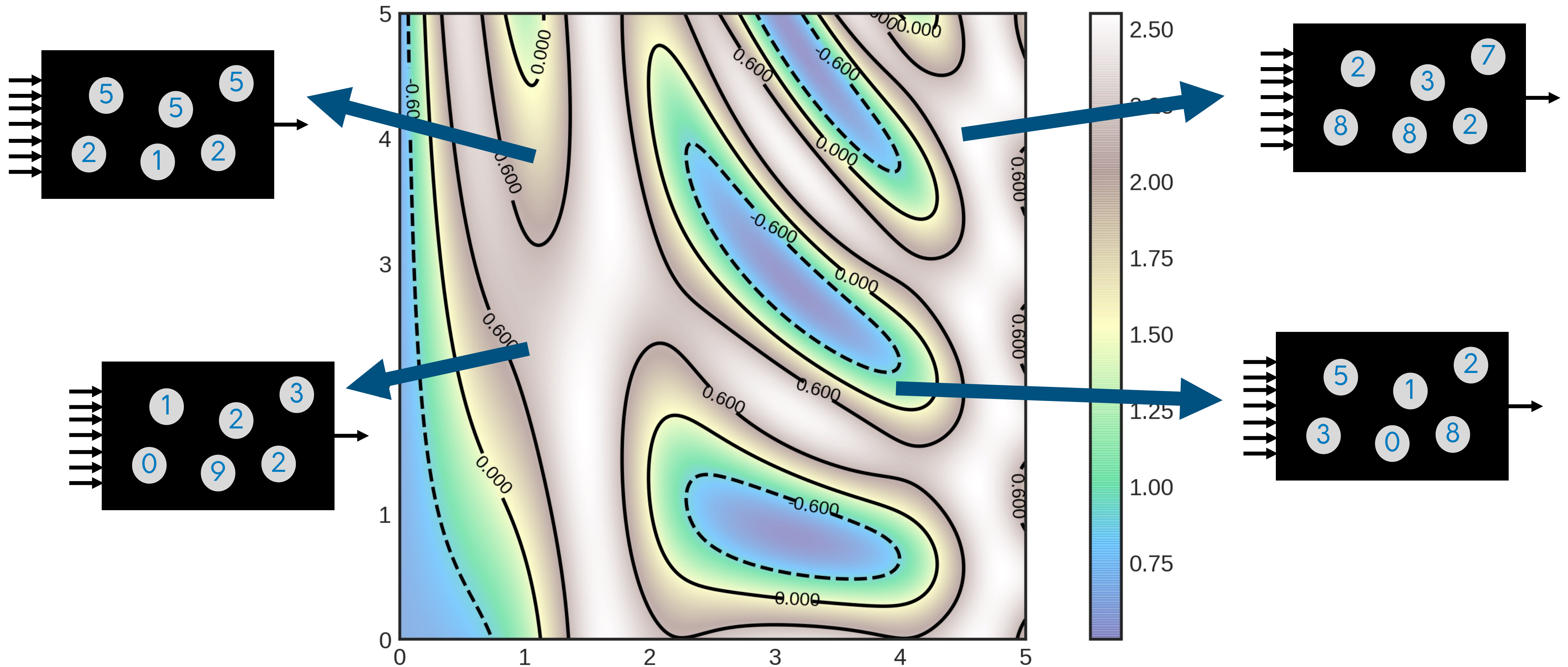
Apprendre, c'est descendre

On cherche (toujours) le minimum d'une fonction calculée



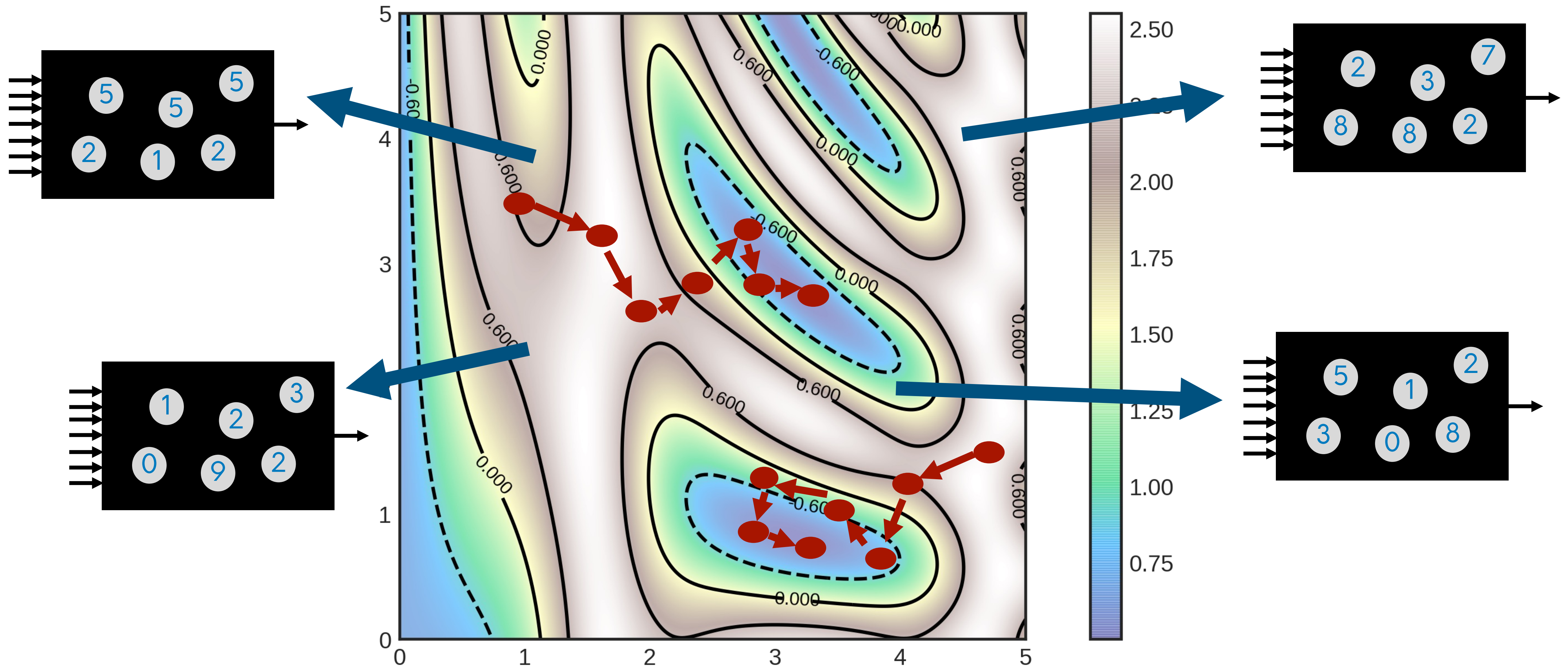
Apprendre, c'est descendre

On cherche (toujours) le minimum d'une fonction calculée



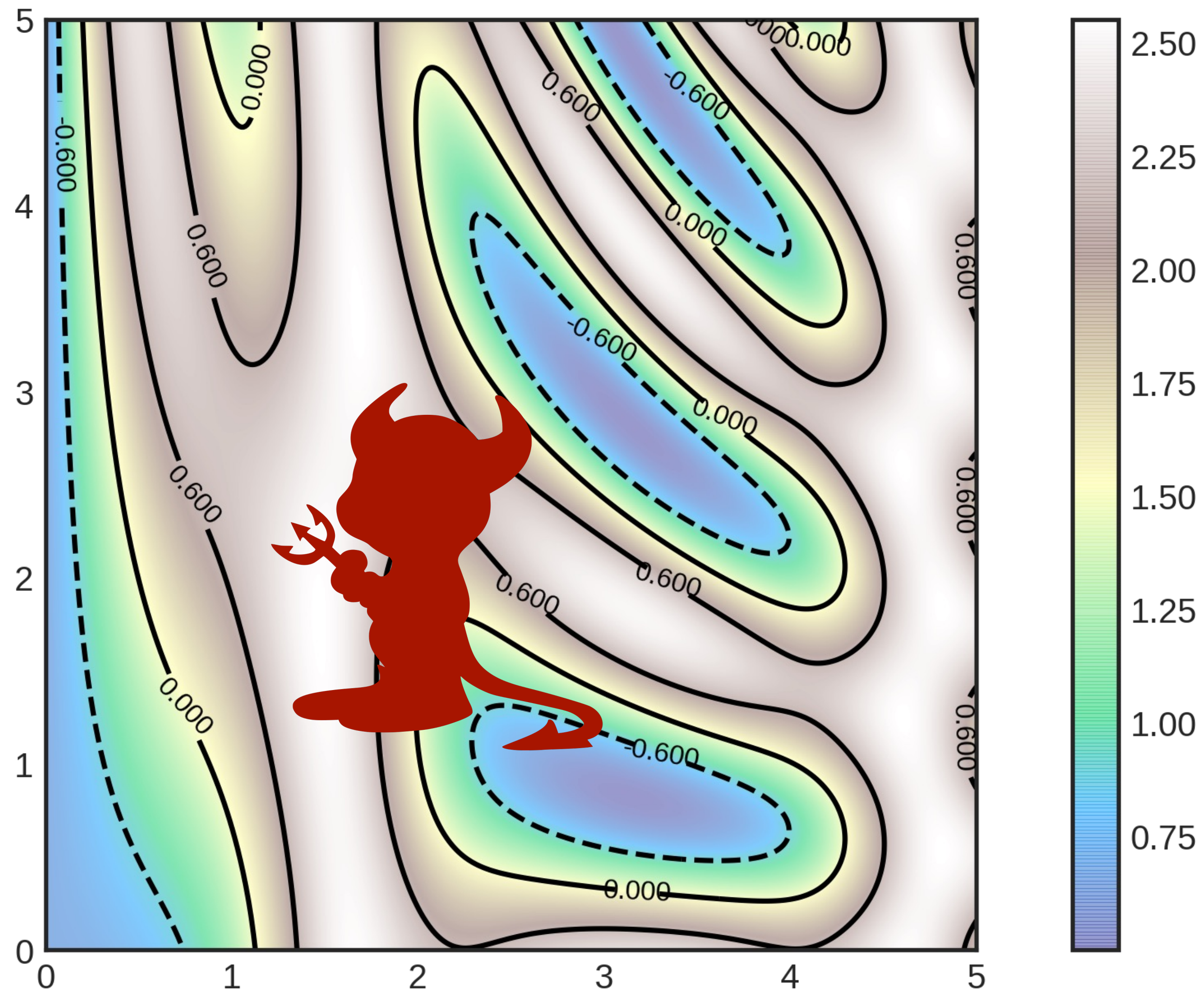
Apprendre, c'est descendre

On cherche (toujours) le minimum d'une fonction calculée

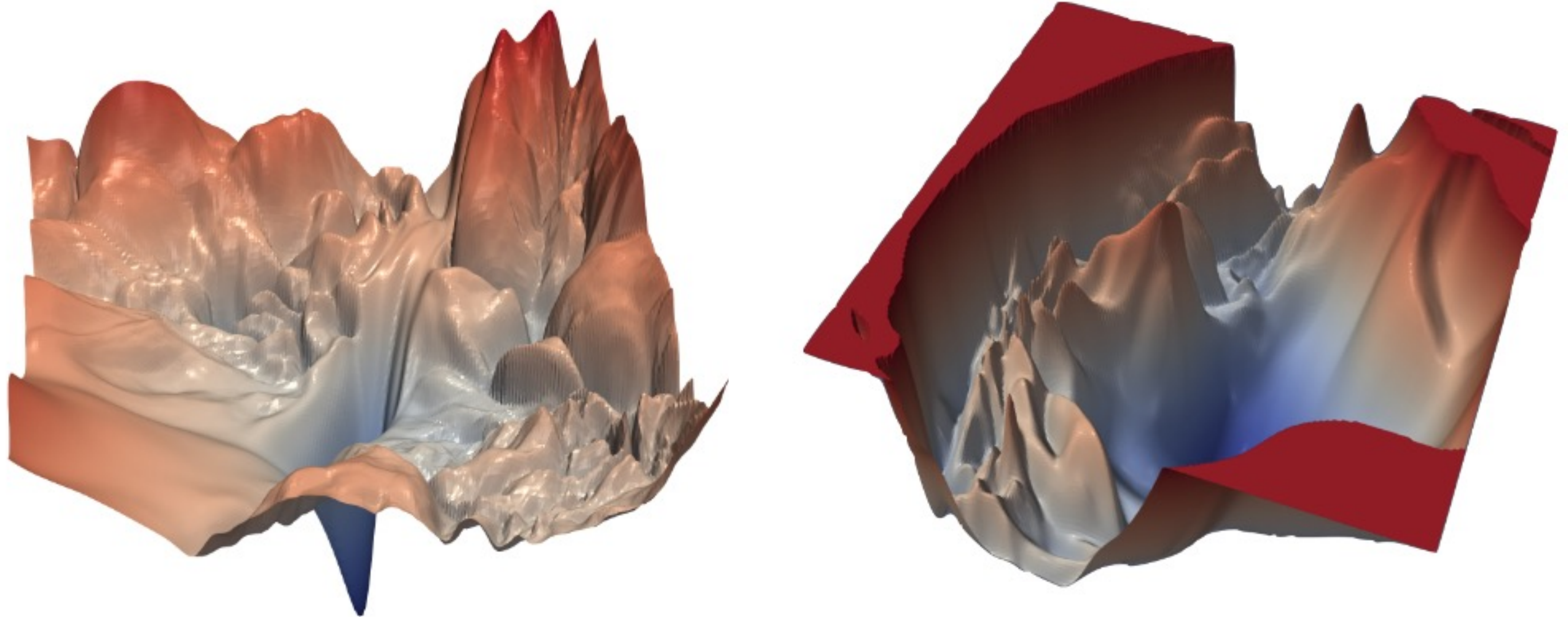


Apprendre, c'est descendre (aux enfers ?)

On cherche (toujours) le minimum d'une fonction calculée

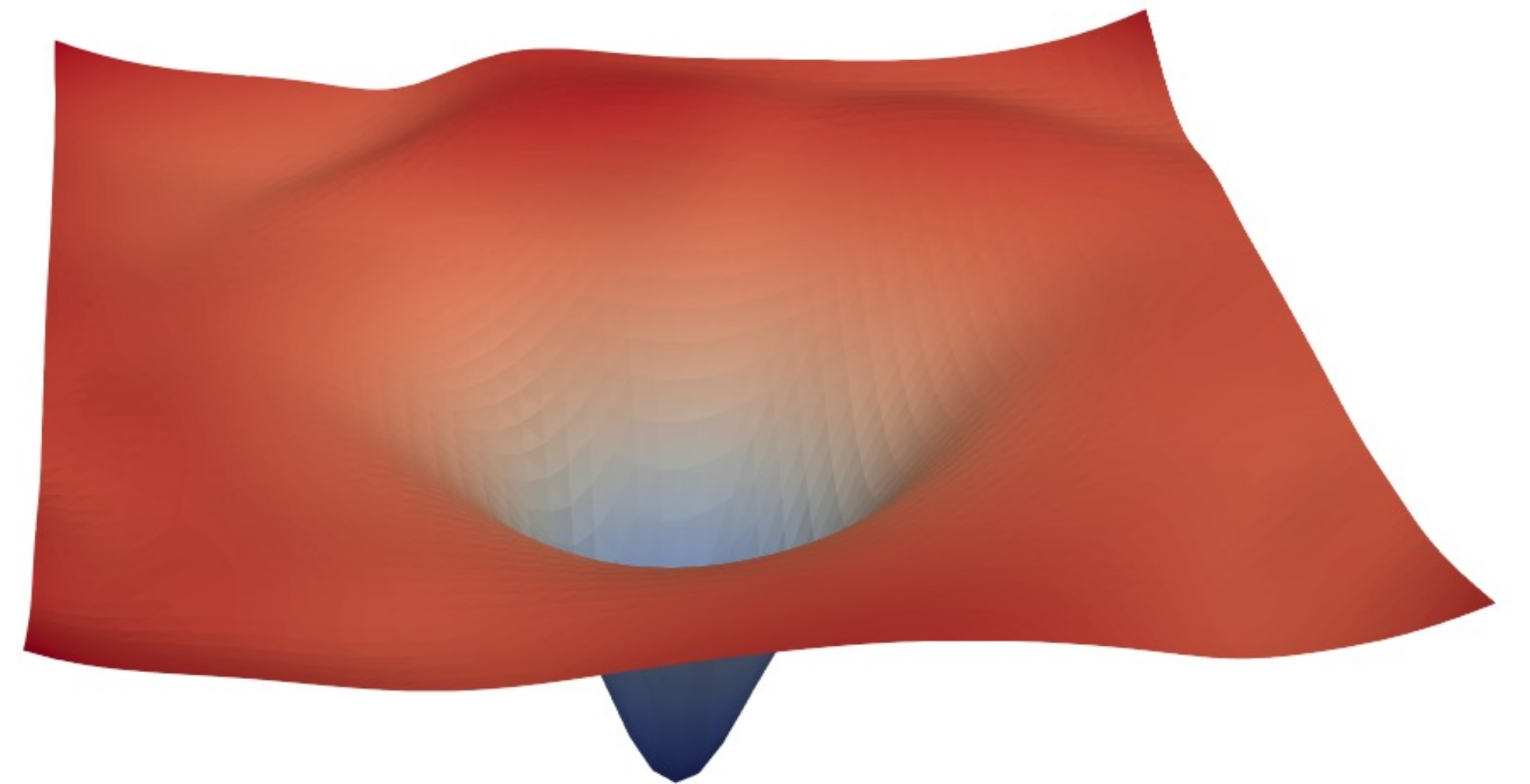
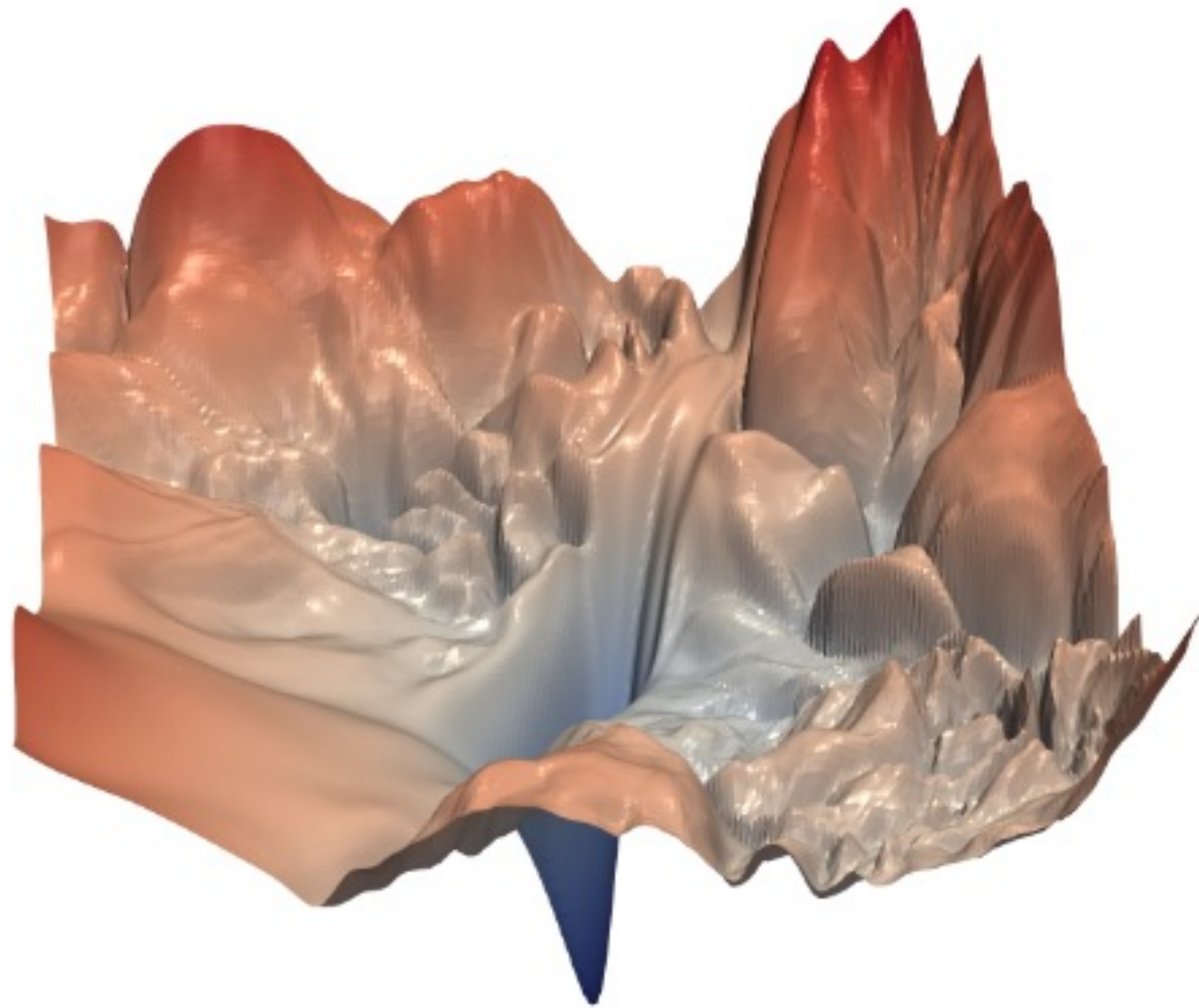


De vrais paysages de Deep Learning



Vizualizing the Loss Landscape of neural Networks (2017)
(Resnet56 sur CIFAR-10)

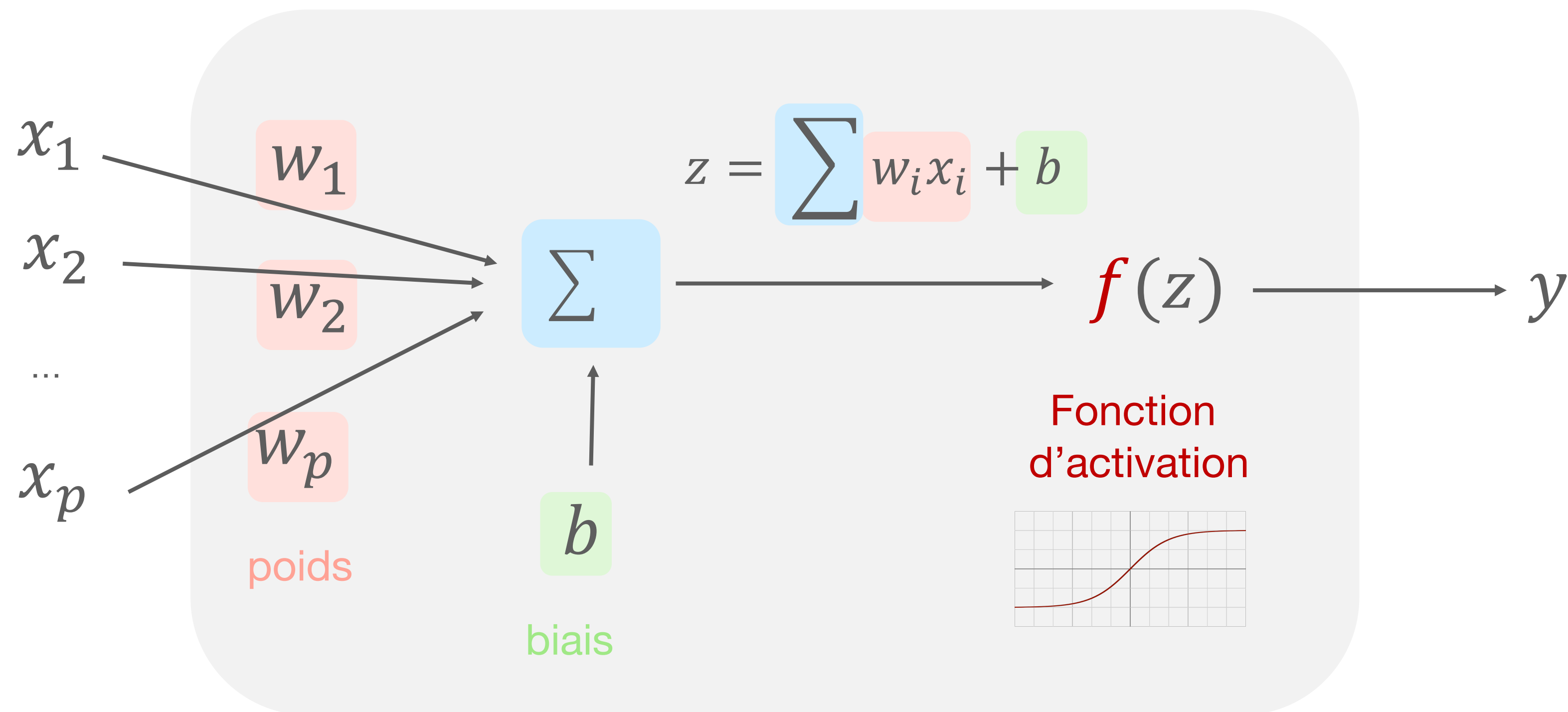
De vrais paysages de Deep Learning



Vizualizing the Loss Landscape of neural Networks (2017)

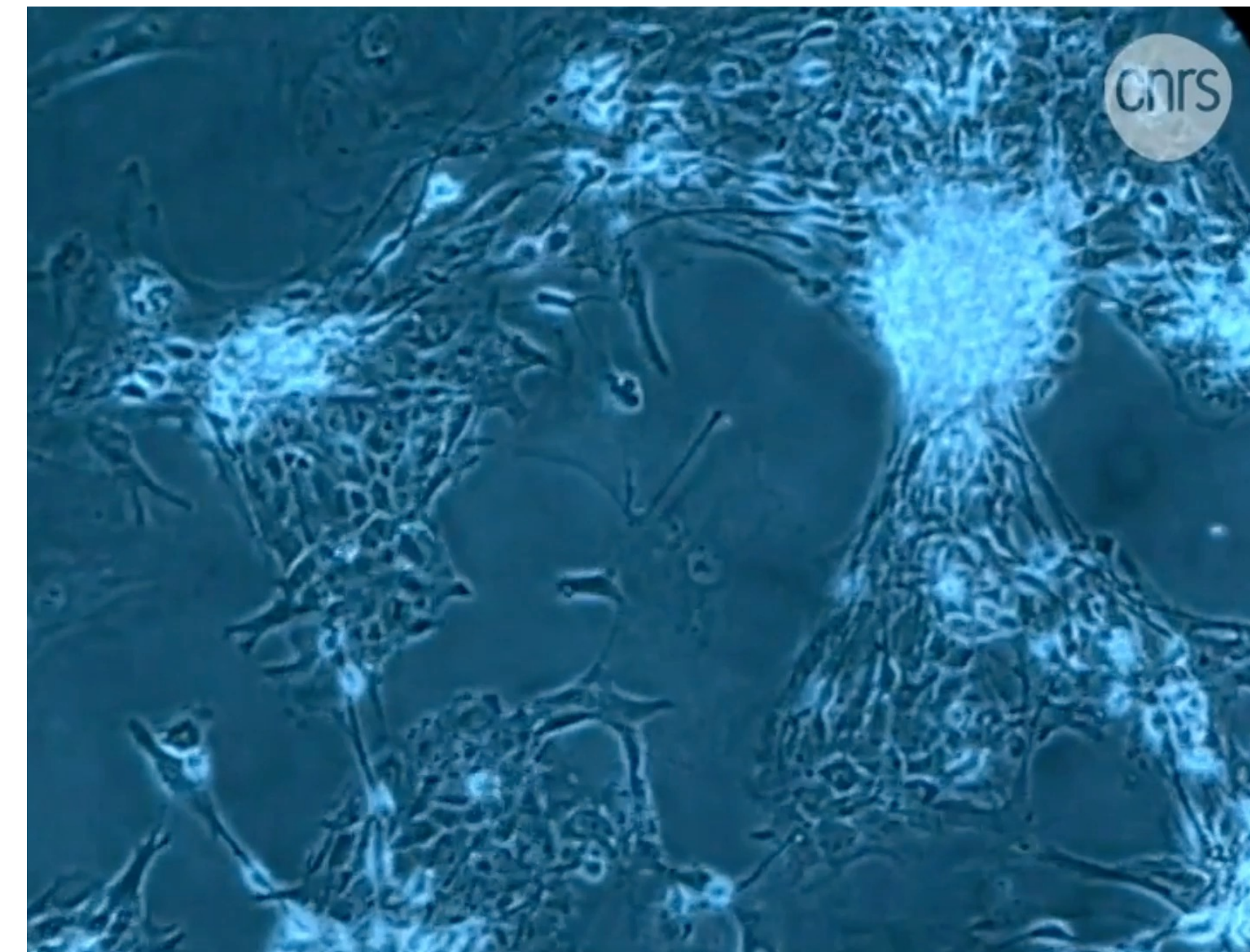
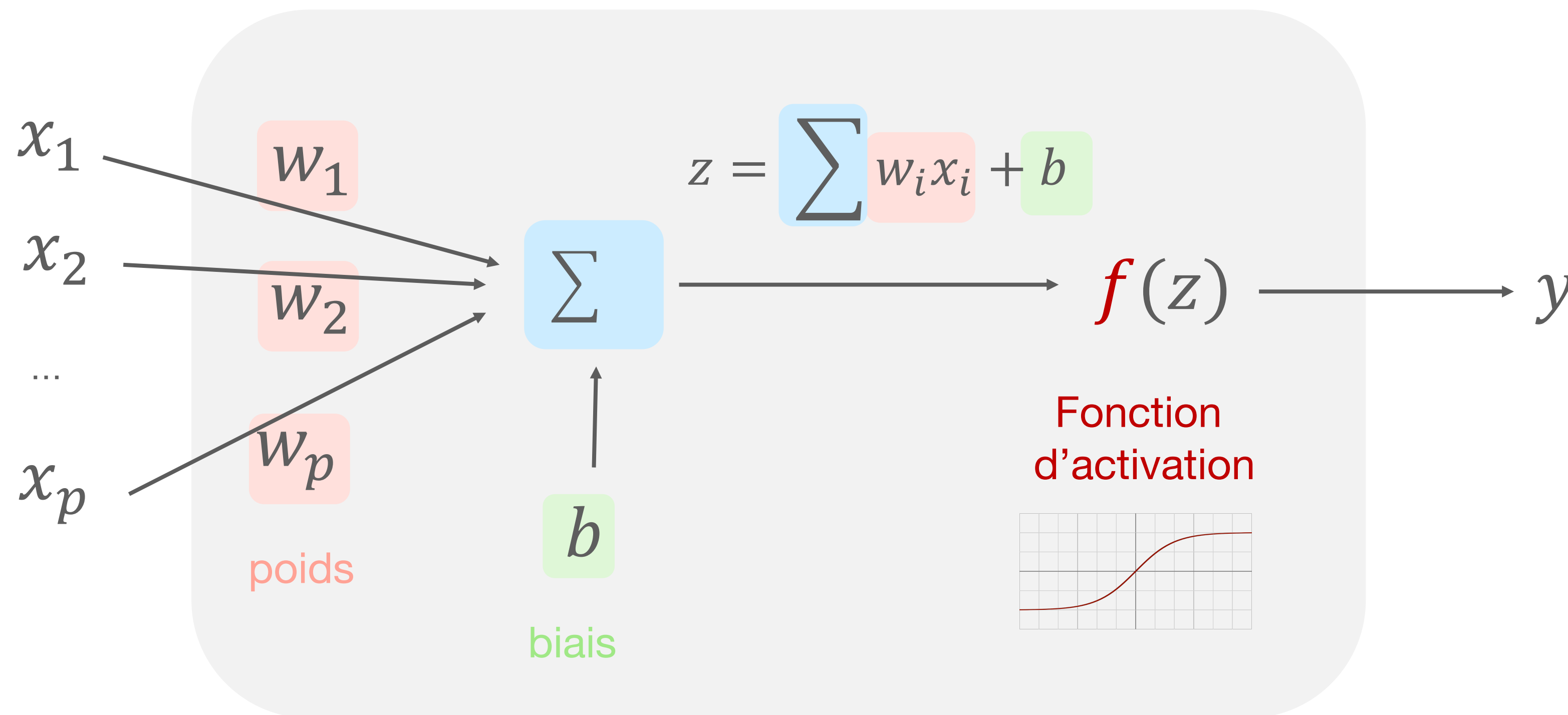
Ce que l'on appelle « neurone »

Formellement c'est moins vivant



Ce que l'on appelle « neurone »

Formellement c'est moins vivant



Les derniers développements

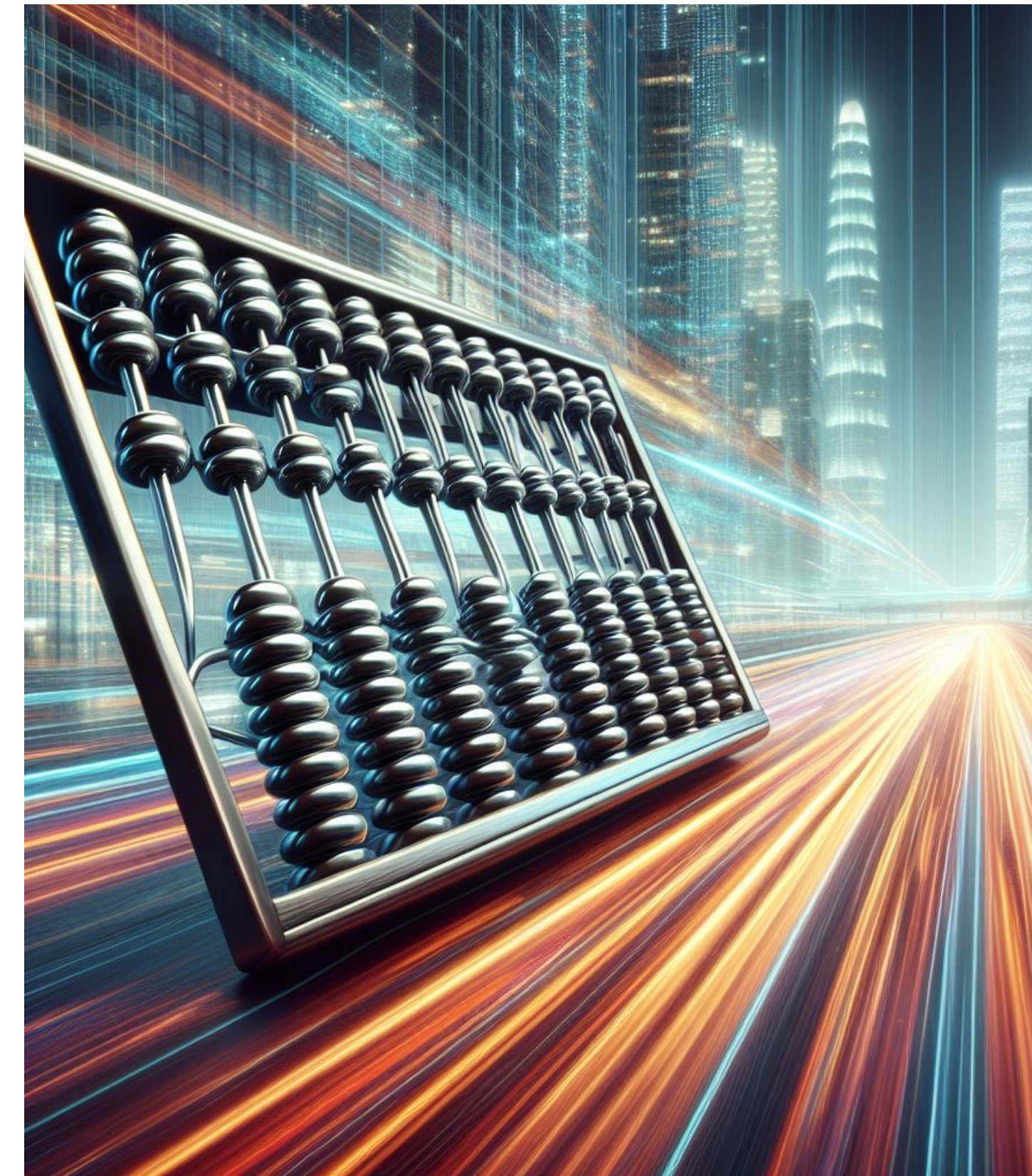
Organisation des neurones

Organisations des neurones :

- Mécanismes de convolutions
- Mécanismes d'attentions
- ...

- A chaque tâche correspond son architecture

- Cartes de calculs dédiées



Il existe d'innombrables types de réseaux

A mostly complete chart of Neural Networks

©2016 Fjodor van Veen - asimovinstitute.org

- Backfed Input Cell
- Input Cell
- Noisy Input Cell
- Hidden Cell
- Probabilistic Hidden Cell
- Spiking Hidden Cell
- Output Cell
- Match Input Output Cell
- Recurrent Cell
- Memory Cell
- Different Memory Cell
- Kernel
- Convolution or Pool

Perceptron (P)



Feed Forward (FF)



Radial Basis Network (RBF)



Deep Feed Forward (DFF)



Markov Chain (MC)



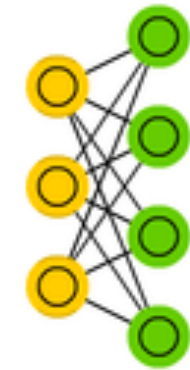
Hopfield Network (HN)



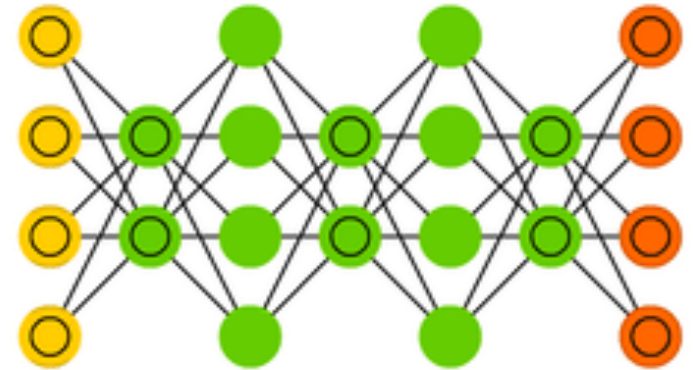
Boltzmann Machine (BM)



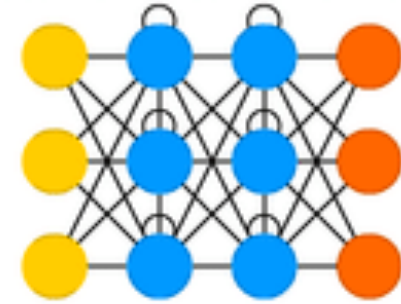
Restricted BM (RBM)



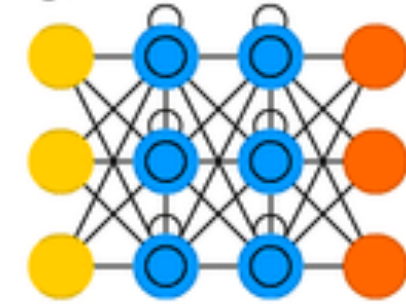
Deep Belief Network (DBN)



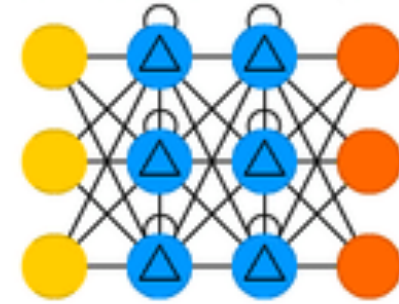
Recurrent Neural Network (RNN)



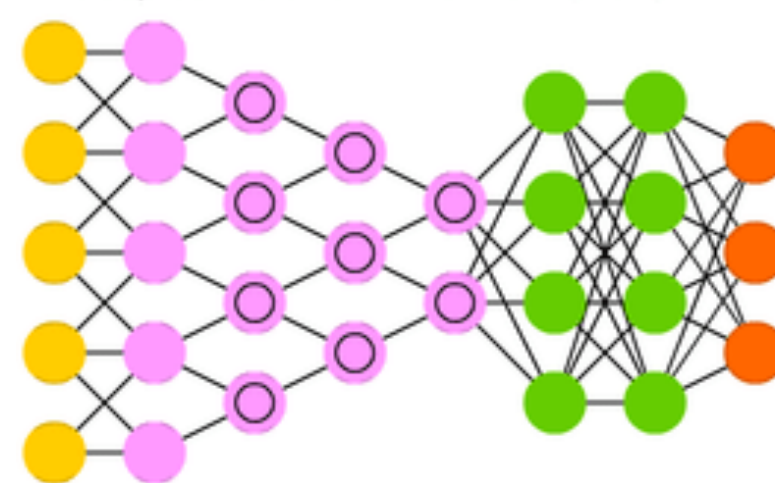
Long / Short Term Memory (LSTM)



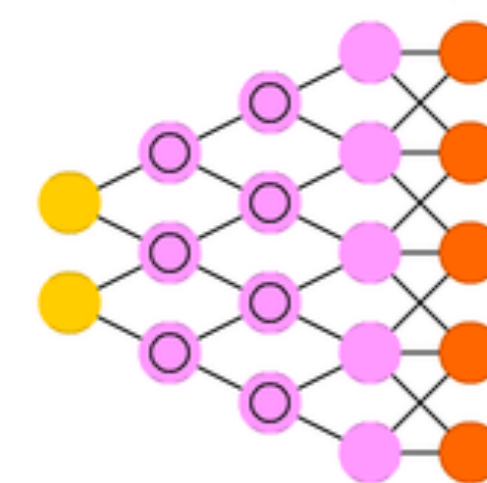
Gated Recurrent Unit (GRU)



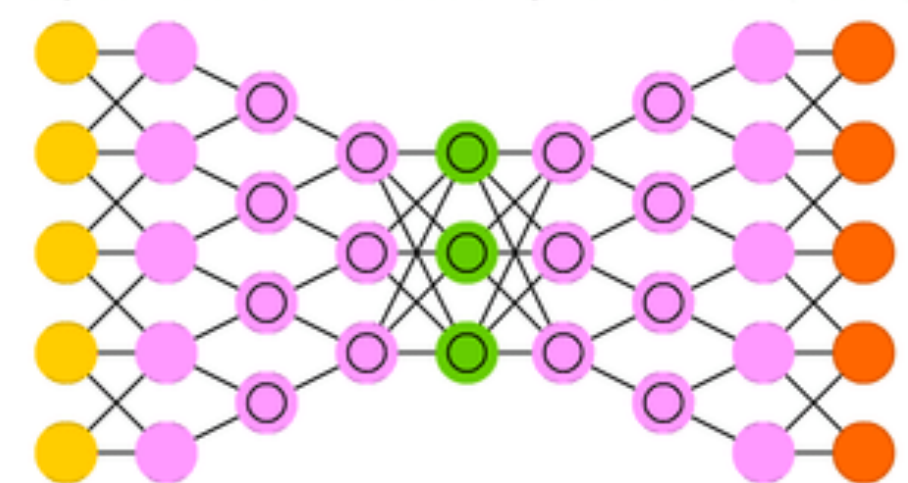
Deep Convolutional Network (DCN)



Deconvolutional Network (DN)



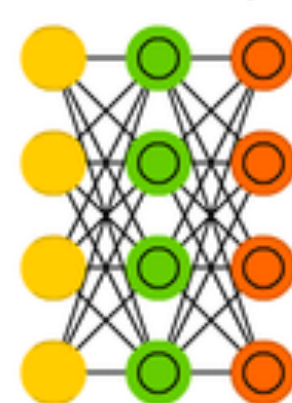
Deep Convolutional Inverse Graphics Network (DCIGN)



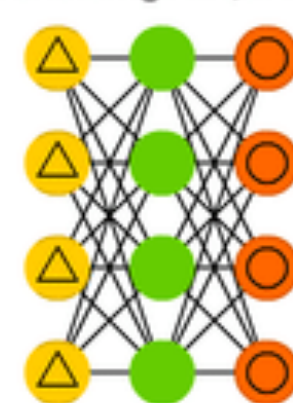
Auto Encoder (AE)



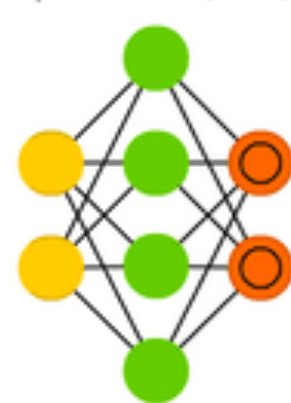
Variational AE (VAE)



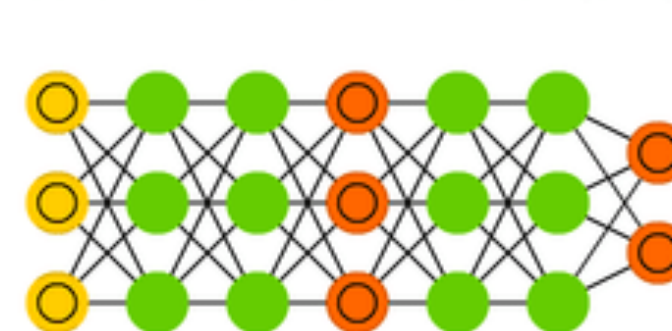
Denosing AE (DAE)



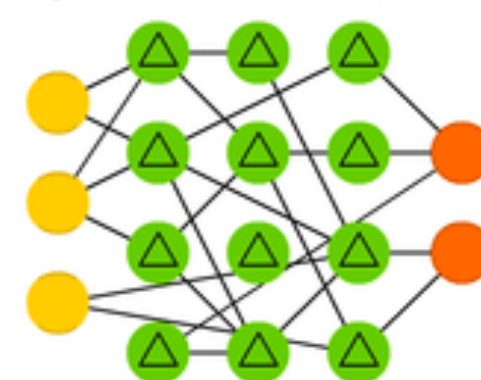
Sparse AE (SAE)



Generative Adversarial Network (GAN)



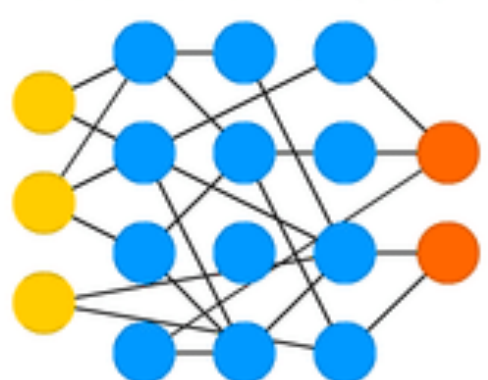
Liquid State Machine (LSM)



Extreme Learning Machine (ELM)



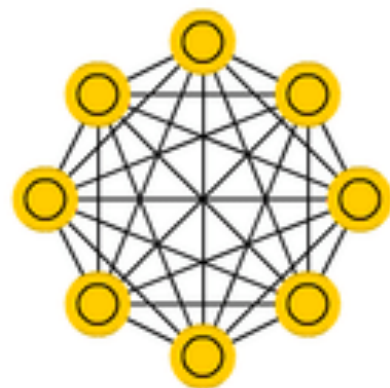
Echo State Network (ESN)



Markov Chain (MC)



Hopfield Network (HN)



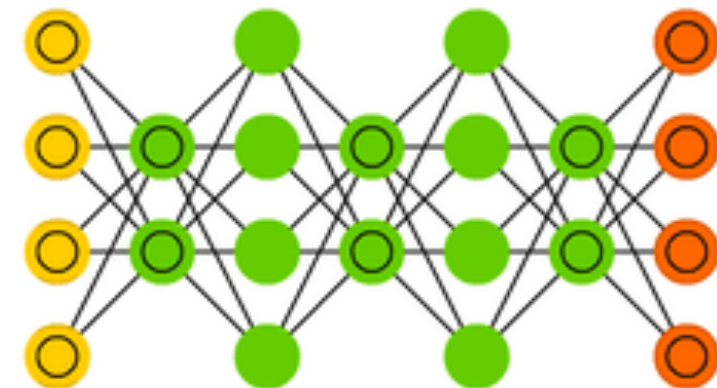
Boltzmann Machine (BM)



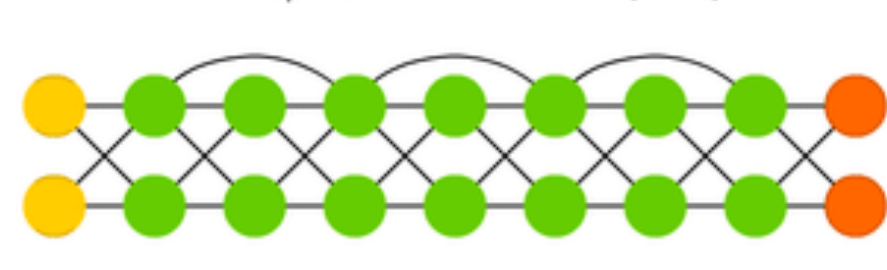
Restricted BM (RBM)



Deep Belief Network (DBN)



Deep Residual Network (DRN)



Kohonen Network (KN)



Support Vector Machine (SVM)



Neural Turing Machine (NTM)



Les bouleversements récents

Dall-E / MidJourney













Comment générer une image ?

« Un Golden Retriever sur une terrasse vu en grand angle »



Générer une image

Bruit aléatoire



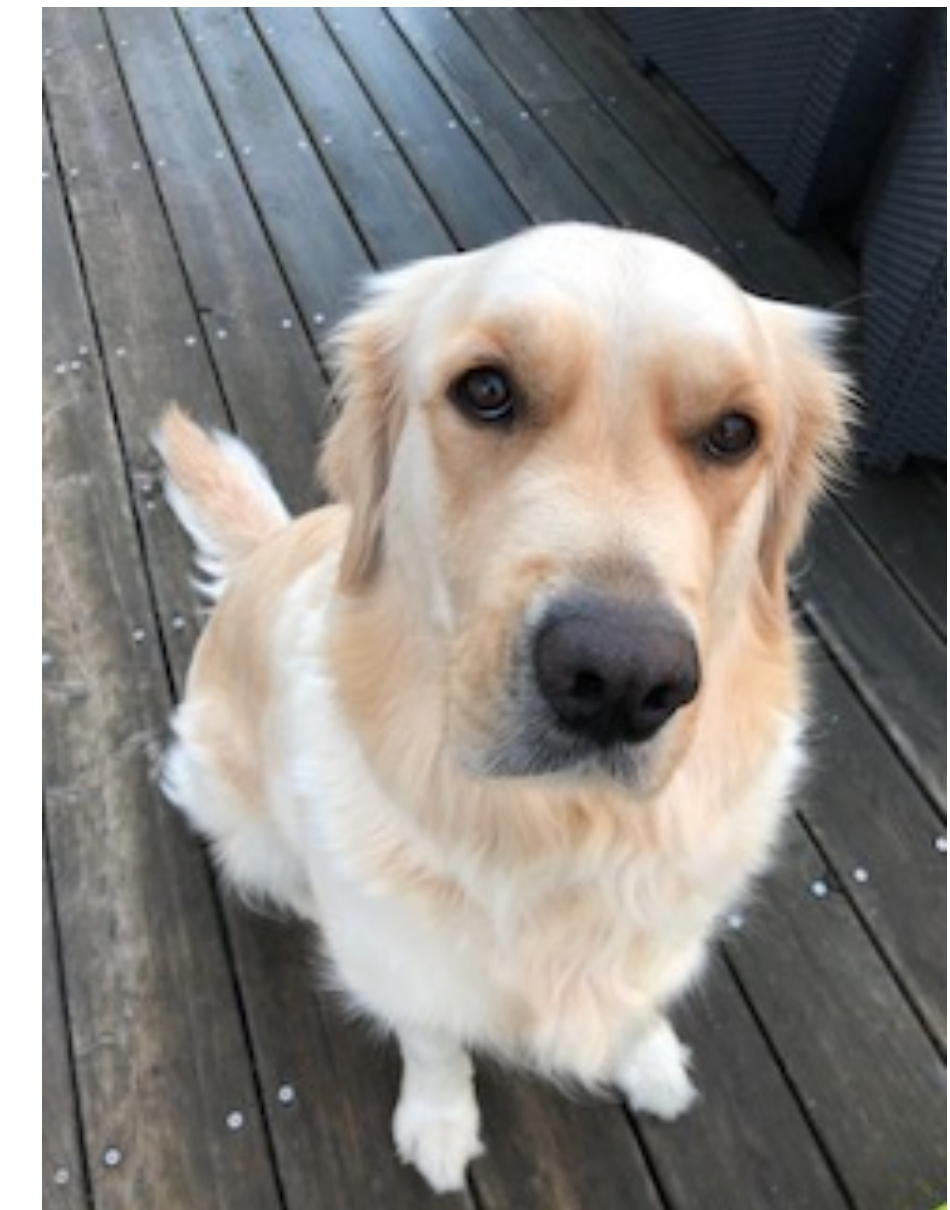
+

Description

=

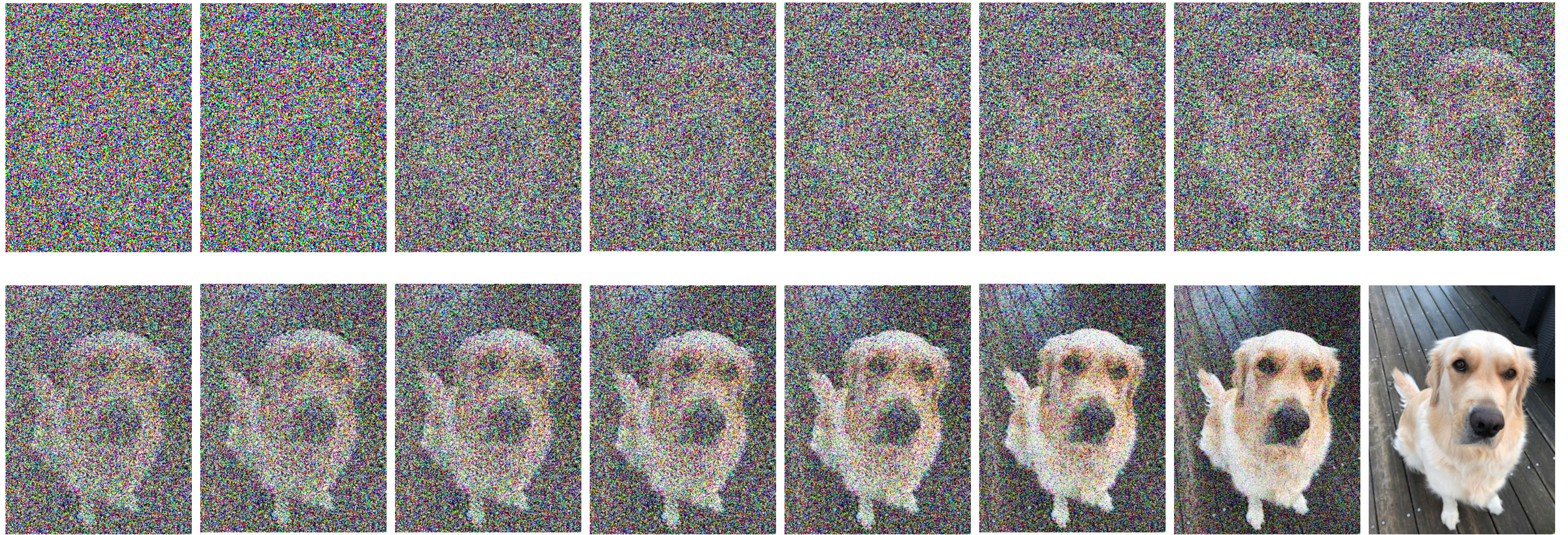
Image finale

« Un Golden Retriever sur une
terrasse vu en grand angle »



Comment générer une image ?

« Un Golden Retriever sur une terrasse »



Facile de construire un jeu de données d'après une banque d'images annotées

Dégrader le texte pour apprendre

On ne se baigne jamais deux fois dans le même fleuve

On ne se baigne jamais deux ???

jamais deux fois dans le même ???

On ne ???

se baigne jamais deux fois dans le même fleuve ???

■ ■ ■ ■ ■ ■ ■ ... ■ ■ ■ ■ ■ ■ ■ ■ ■ ???

Un roman : 60 000 – 100 000 mots

GPT3 / Mistral / Llama : 2048 Tokens (1000 mots),
Claude : 100 000 Tokens

Passer d'un mot manquant à une conversation entière (**chatGPT**)

Entraîner un entraîneur

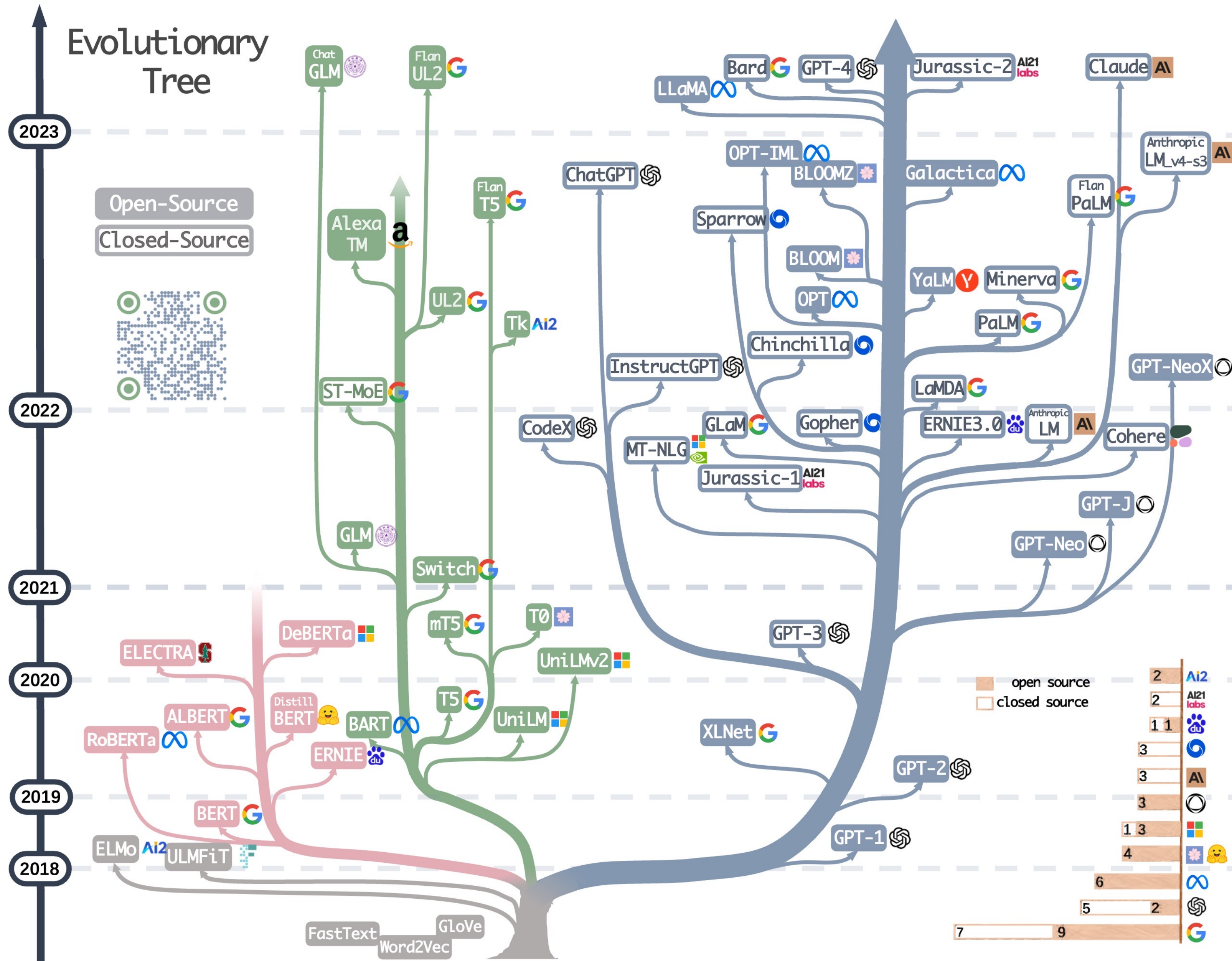
1. Rédiger / Annoter des exemples de conversations
2. Tokenisation
Bonjour, je suis content
-> [1273, 23, 8876, 29929]
3. Prédire le prochain mot
4. Générer les réponses en produisant successivement leurs mots
5. Demander aux humains de classer les meilleurs réponses
6. Entraîner l'entraîneur à classer de la même manière

Entraîner chatGPT

1. Aspirer / générer (beaucoup) d'exemples de conversations
2. Tokenisation
Bonjour, je suis content
-> [1273, 23, 8876, 29929]
3. Prédire le prochain mot
4. Générer les réponses en produisant successivement leurs mots
5. Demander à **l'entraîneur** de noter les meilleurs réponses
6. S'entraîner pour produire les meilleurs réponses d'après l'entraîneur



L'écosystème LLM (grandit tous les jours)



LLaMA: Open and Efficient Foundation Language Models

Hugo Touvron*, Thibaut Lavril*, Gautier Izacard*, Xavier Martinet
 Marie-Anne Lachaux, Timothee Lacroix, Baptiste Rozière, Naman Goyal
 Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin
 Edouard Grave*, Guillaume Lample*

Meta AI

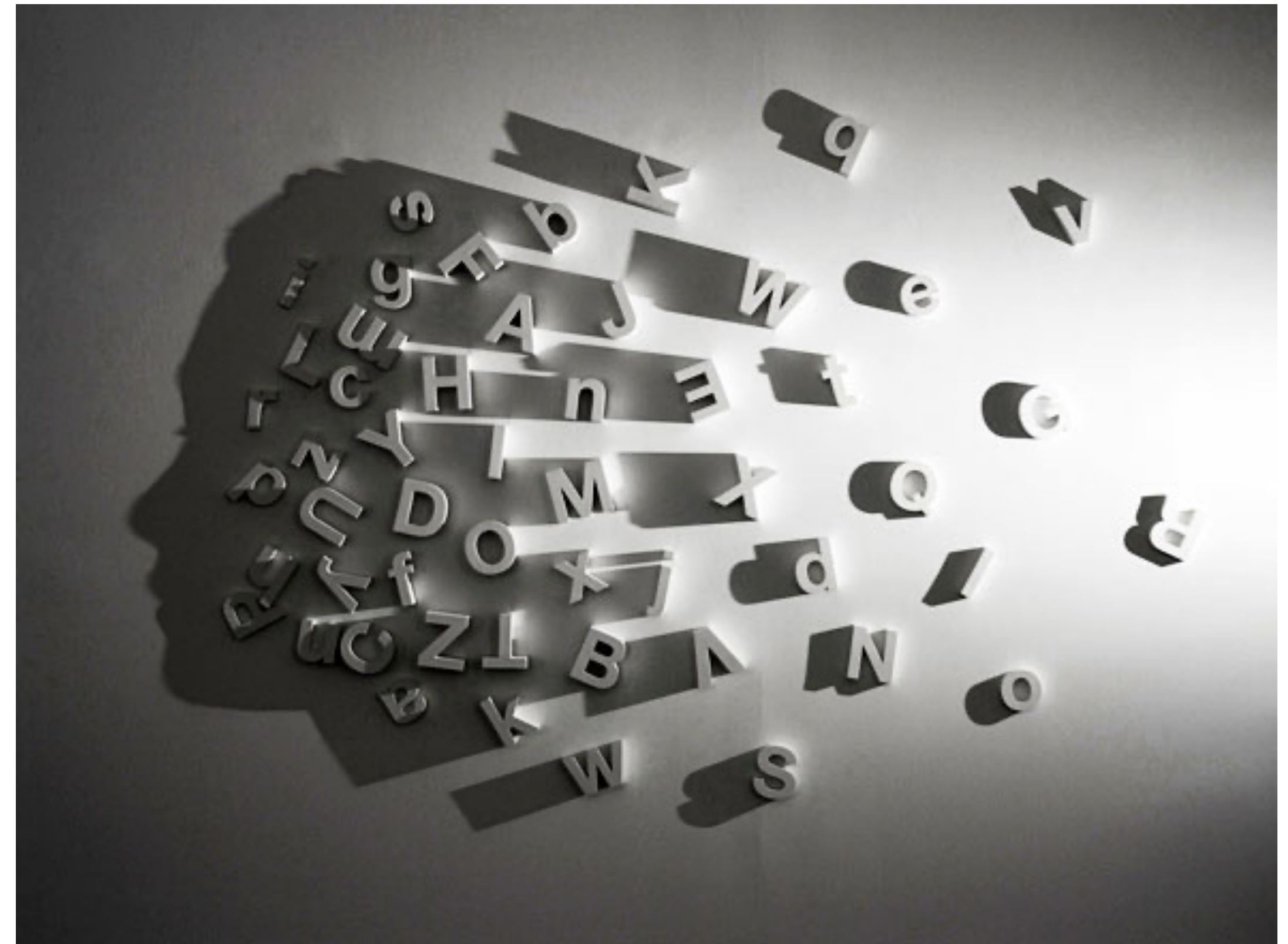
Dataset	Sampling prop.	Epochs	Disk size
CommonCrawl	67.0%	1.10	3.3 TB
C4	15.0%	1.06	783 GB
Github	4.5%	0.64	328 GB
Wikipedia	4.5%	2.45	83 GB
Books	4.5%	2.23	85 GB
ArXiv	2.5%	1.06	92 GB
StackExchange	2.0%	1.03	78 GB

Table 1: **Pre-training data.** Data mixtures used for pre-training, for each subset we list the sampling proportion, number of epochs performed on the subset when training on 1.4T tokens, and disk size. The pre-training runs on 1T tokens have the same sampling proportion.

Le pouvoir des immenses espaces latents

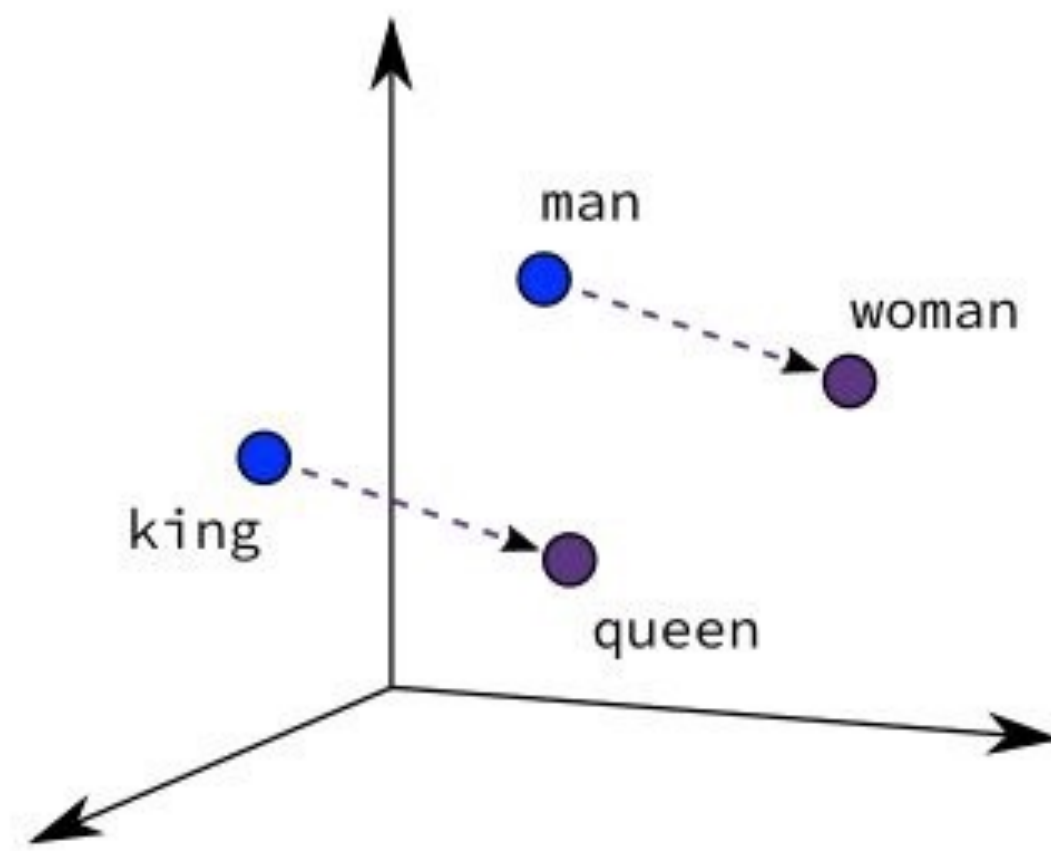


Sculpture de Shigeo Fukuda

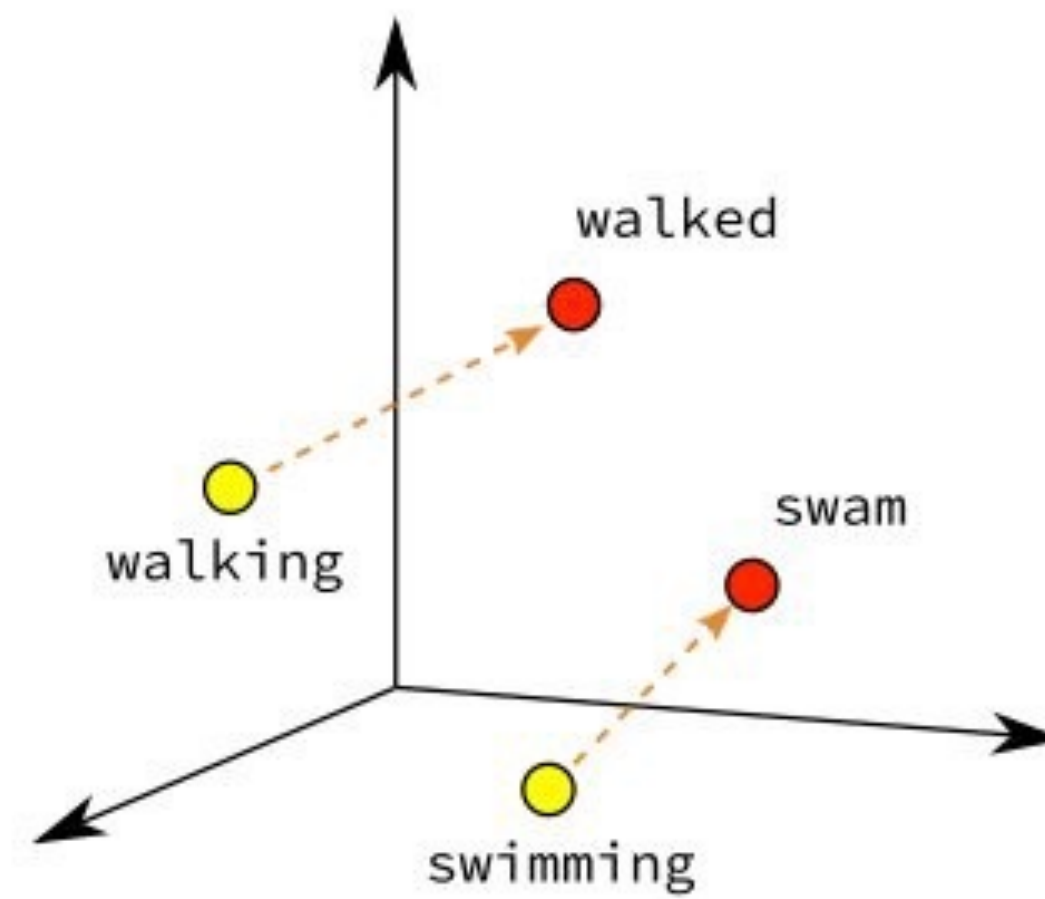


Sculpture de Kumi Yamashita

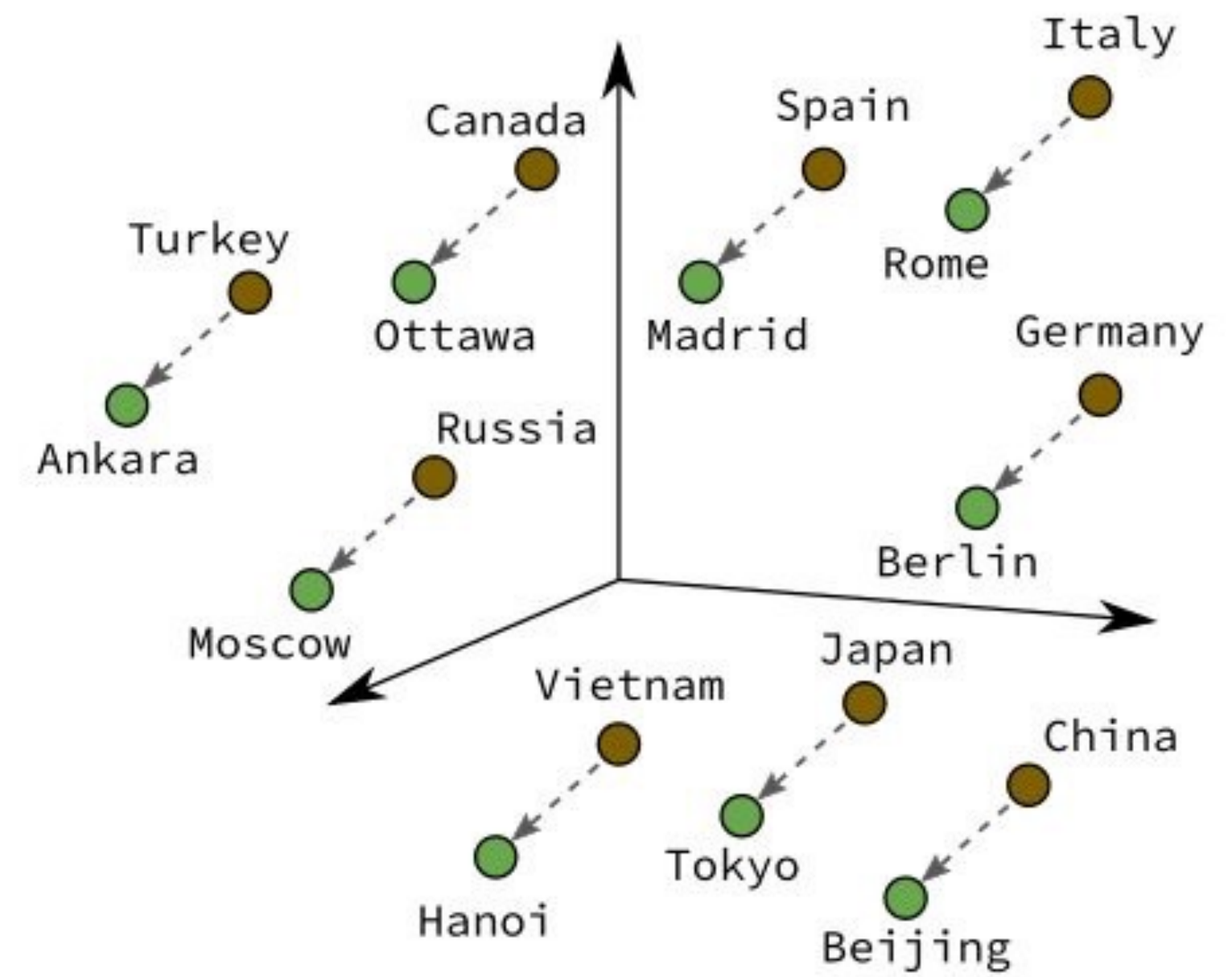
Une représentation statistique – non factuelle



Male-Female



Verb Tense



Country-Capital

Quelques limites des ChatGPT3-4 & co

- Texte généré non sourcé (source / biblio “halluciné”)
- Utilise le dialogue comme “mémoire de travail”
- Ne va pas chercher du contenu qui correspond à la question mais produit simplement une suite probable
- N’apprend pas de la discussion
- N’est pas à jour des informations (sauf avec système RAG ?)
- N’a pas de notion de vrai et de faux
- Ne respecte pas (du tout) les droits d’auteurs
- Risque de pollution des données à long terme
- Aucunes garanties sur la non-toxicité des productions
- Répond au QCMs sans questions 😊 / Favorise les premières réponses

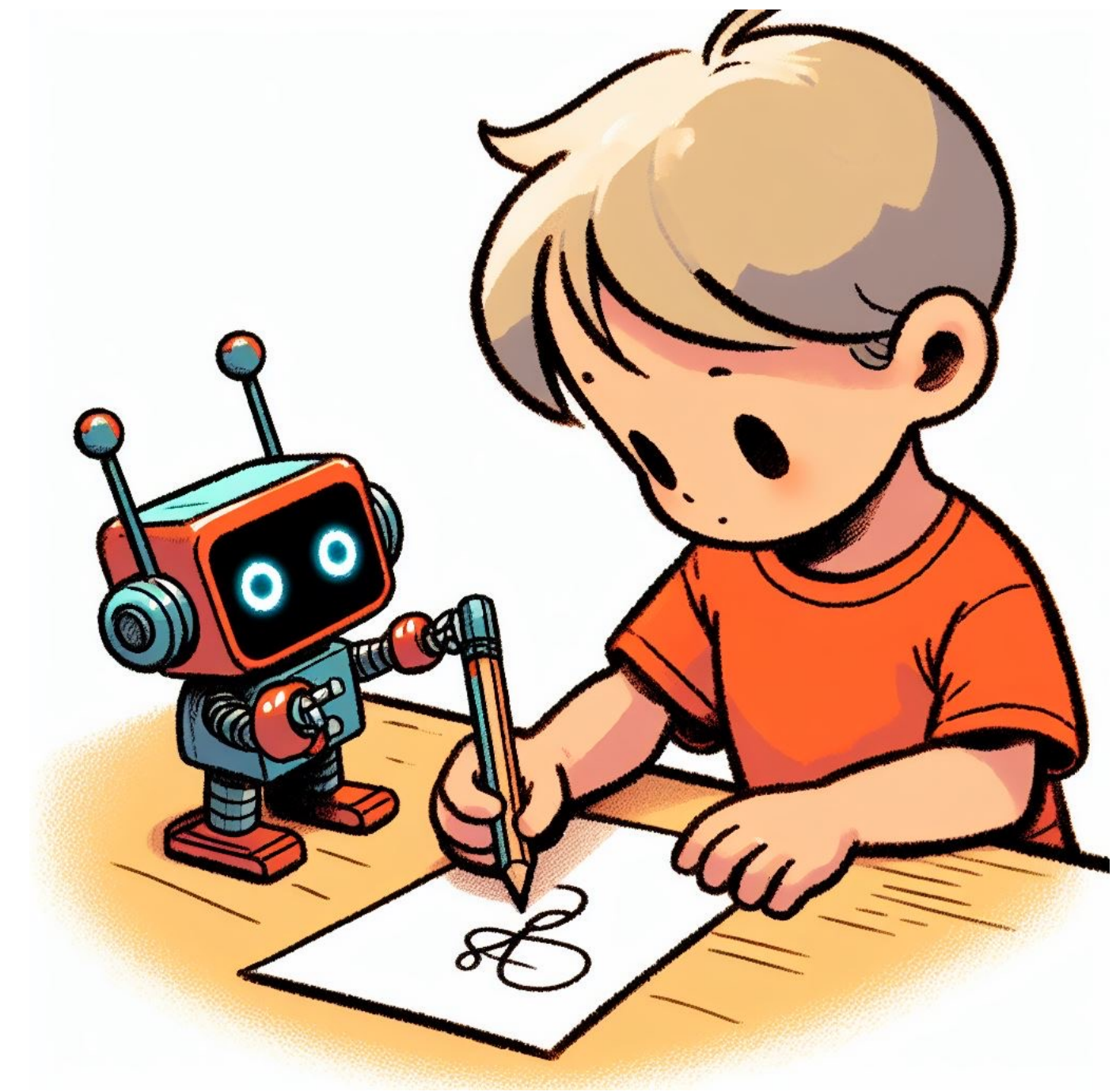


Comment l'utiliser ?

C'est un outil qu'il faut encadrer s'il y a un enjeu

- Aide à la rédaction de mails formels
- Résumé automatique d'une réunion (déjà en production)
- Résumé d'échanges, veille technologique
- Aide à la création en fournissant une source inépuisable de variations

- Les changements à long termes sont encore inimaginables



Quels impacts ?

- En 2022, Penguin Random House a publié 58000 titres (la moitié ont été vendus à moins de 12 exemplaires)

Penguin
Random
House

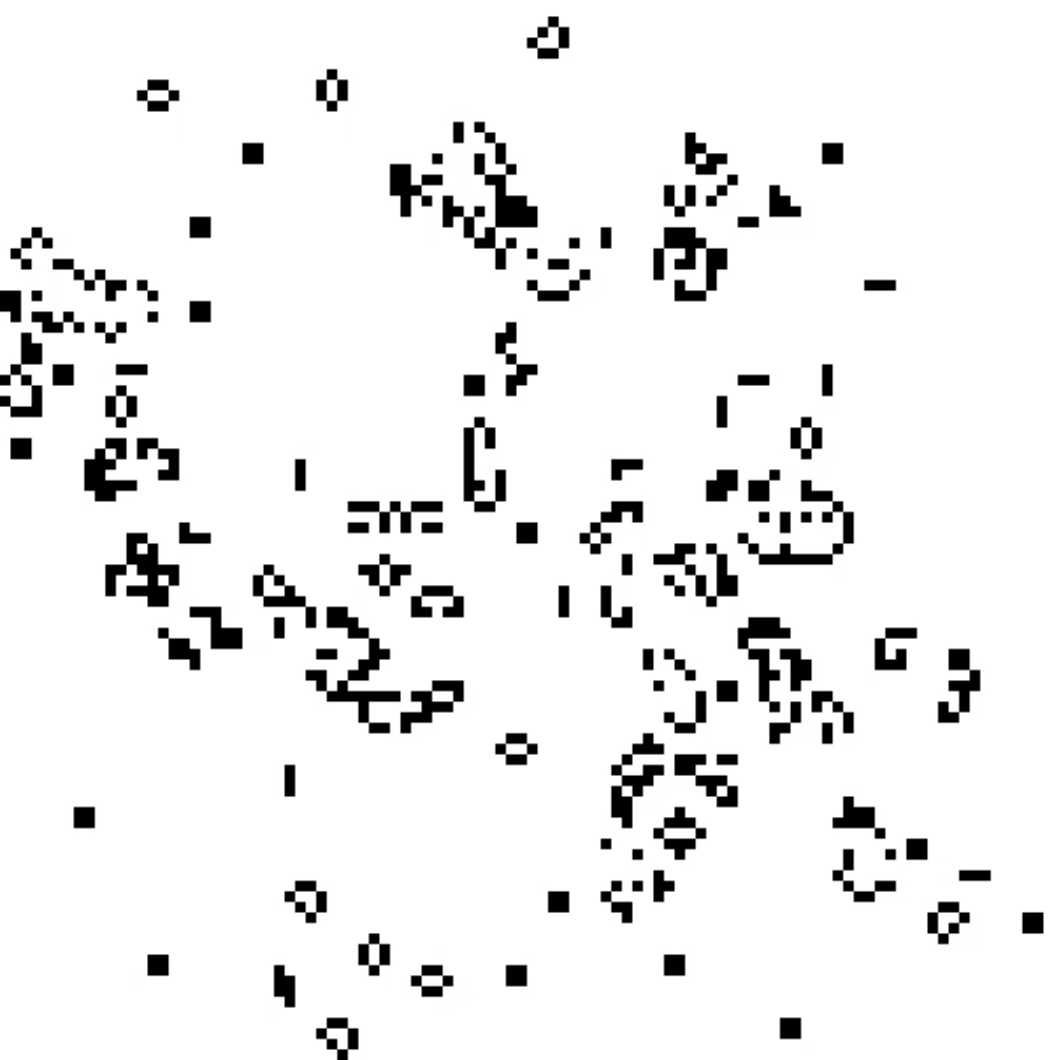
- Fausses Informations ?



L'illusion du forgeron

On a beau avoir construit ces systèmes,
leur déroulement échappe à toute compréhension simple

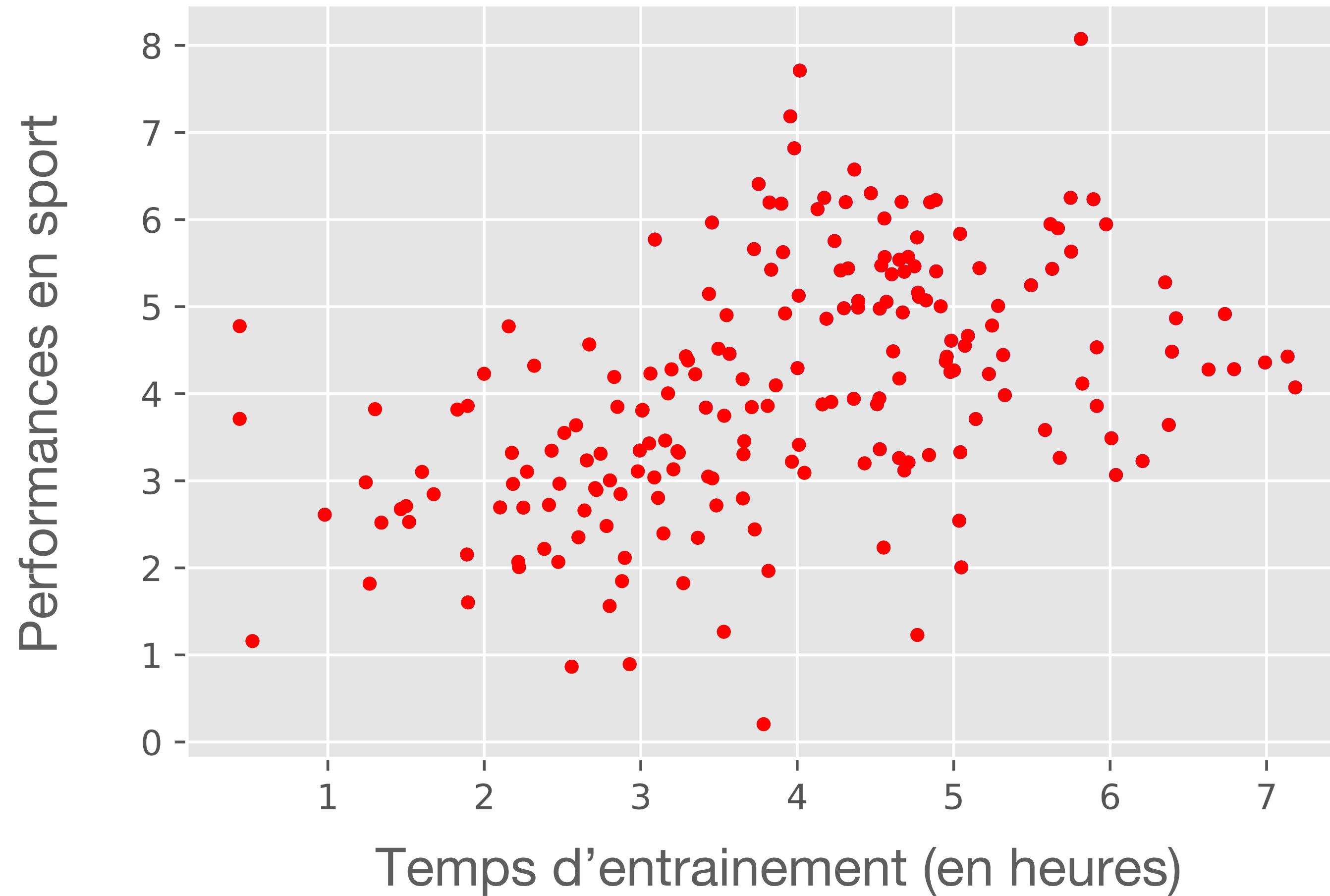
→ **Exiger une explication simple n'est pas toujours possible**



L'I.A. met en œuvre un coût calculatoire **au-delà de l'entendement**

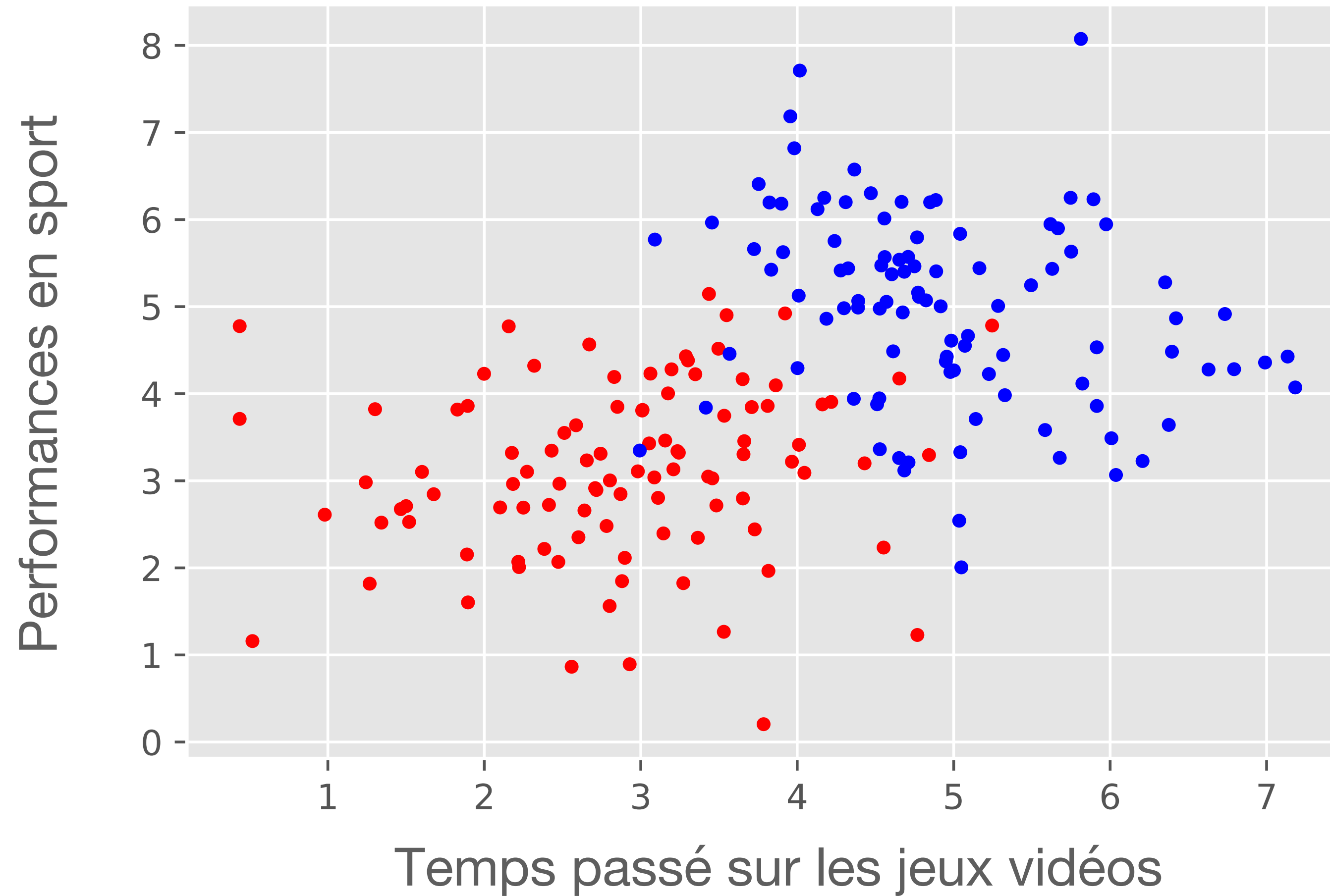
- Les systèmes « logiques » ne peuvent que faire preuve d'argument d'autorité
- Les systèmes d'apprentissage automatiques ont des milliards de paramètres

Corrélation et Causalité



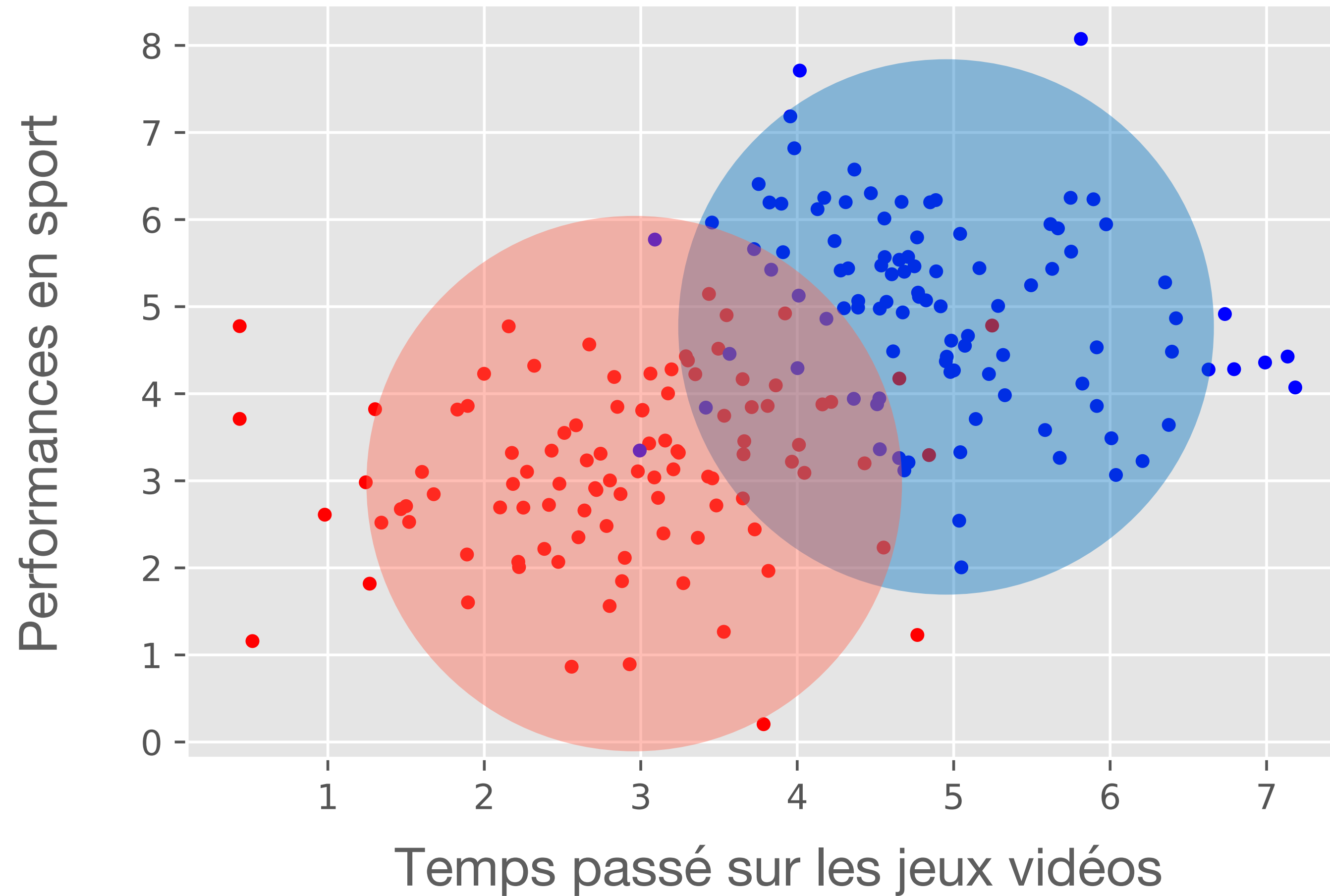
Les données ne sont pas tout

L'histoire des données ne s'invente pas

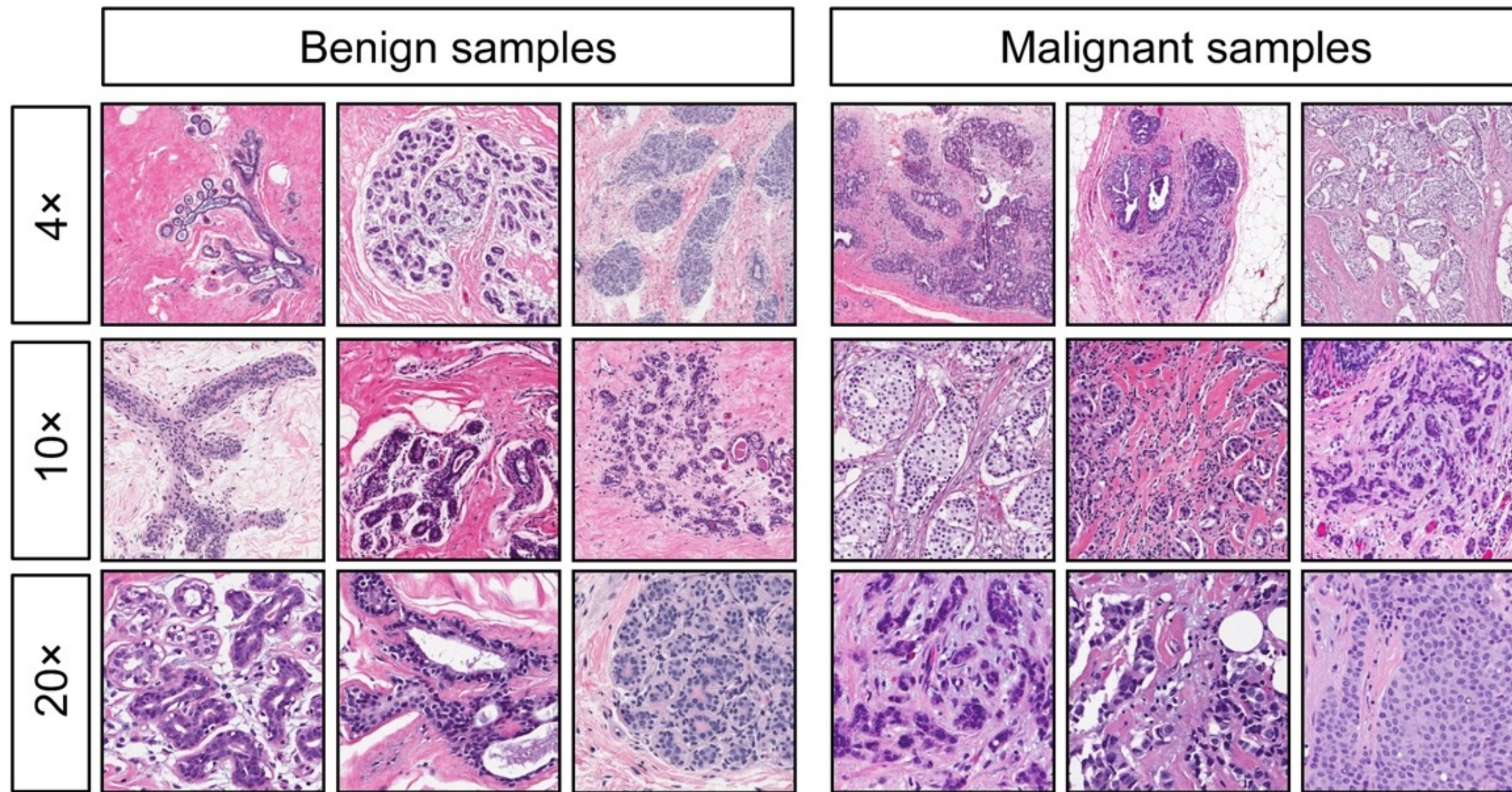


Les données ne sont pas tout

L'histoire des données ne s'invente pas



Performance et Confiance

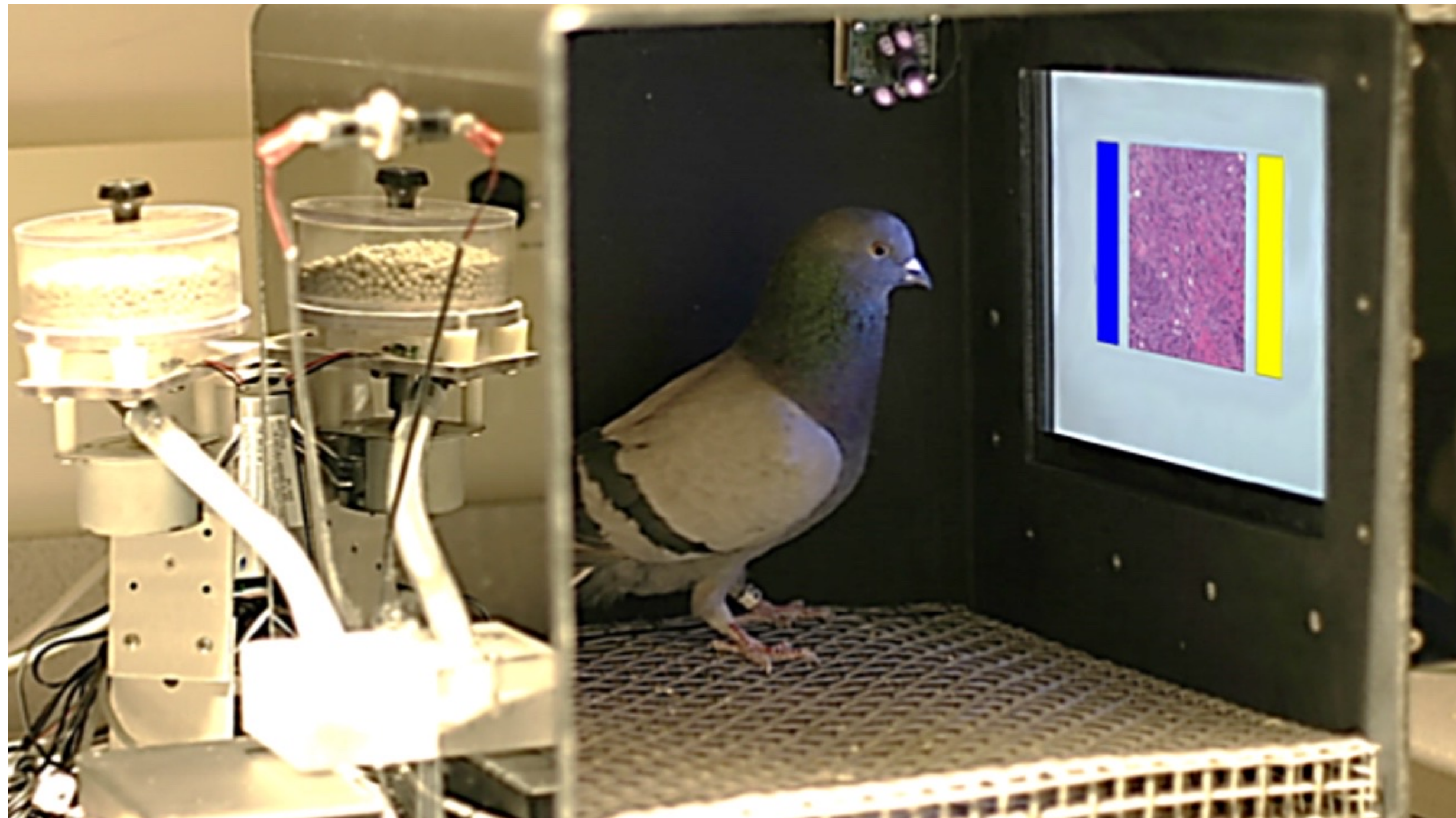


« *L'image d'un prélèvement montre-t-elle certaines cellules malignes ?* »

Systeme basé sur l'apprentissage

→ Meilleures performances que les spécialistes en micro-biologie

Performance et Confiance



« *L'image d'un prélèvement montre-t-elle certaines cellules malignes ?* »

Systeme basé sur l'apprentissage
→ Meilleures performances que les spécialistes en micro-biologie

Pigeons (*Columba livia*) as Trainable Observers of Pathology and Radiology Breast Cancer Images

Richard M. Levenson , Elizabeth A. Krupinski, Victor M. Navarro, Edward A. Wasserman 

Published: November 18, 2015 • <https://doi.org/10.1371/journal.pone.0141357>



Préparez-vous pour l'impact : Intelligence Artificielle

(Définition, espoirs et limites)



Une chaire portée par :



Associant :



Accompagnée par :



Bénéficiant du soutien de :

