

L'intelligence artificielle en questions

Nicolas Roussel
nicolas.roussel@inria.fr

Qu'est-ce que l'intelligence artificielle ?

Où en est-on aujourd'hui ?

Quel est notre rapport à l'IA ?

Quelle place pour l'IA demain ?

Café IA

[ka.fe i.a]

Démarche portée par le Conseil de l'intelligence artificielle et du numérique qui vise à fédérer, soutenir, promouvoir et rassembler les initiatives de débats et de partage de connaissances sur l'IA menées en vue de la diffusion d'une culture populaire du numérique.



Civilisation numérique

Ouvrons le débat!

Le Conseil national du numérique
COPRÉSIDENCE
Françoise Mercadal-Delasalles & Gilles Babinet

Février 2021

Le numérique, c'est notre conviction, est l'affaire de la cité. Il est notre affaire à tous.

Nous recommandons que les débats autour du numérique sortent des seules officines d'experts et prennent toute leur place dans le débat public.

Nous recommandons de « politiser » le numérique pour en faire un outil approprié par tous, au service d'un avenir commun. Faire du devenir numérique des sociétés un enjeu politique, un enjeu de débat, de controverses, de recherche et de réflexion qui implique les représentants politiques et les citoyens est essentiel.

Être ambitieux pour pouvoir formuler de vrais choix, agir et ne pas simplement « être agis » dans l'univers numérique. Politiser le numérique, pour statuer collectivement et démocratiquement sur les choix de société que nous voulons porter à l'ère numérique.

Juin 2022



RESTITUTION RÉALISÉE PAR
Jean-Baptiste Manenti et Louis Magnes

Mai 2023

Malgré le sentiment qu'aucun de nous n'a réellement son mot à dire sur un numérique qui investit jusqu'aux pans les plus intimes de nos vies, il existe en France une farouche envie d'en débattre et de se faire entendre.

Cette voix citoyenne doit être entendue autant qu'elle doit être nourrie, de ressources, d'expériences partagées, de clés de compréhension des impacts de ces outils et de ces usages qui rythment notre quotidien.

Cette voix dans le débat sur nos choix collectifs est d'autant plus légitime que nous entretenons un rapport sensible avec le numérique. Ce qui appelle une posture d'écoute permanente, favorable à l'expression des émotions.

Mars / avril 2024



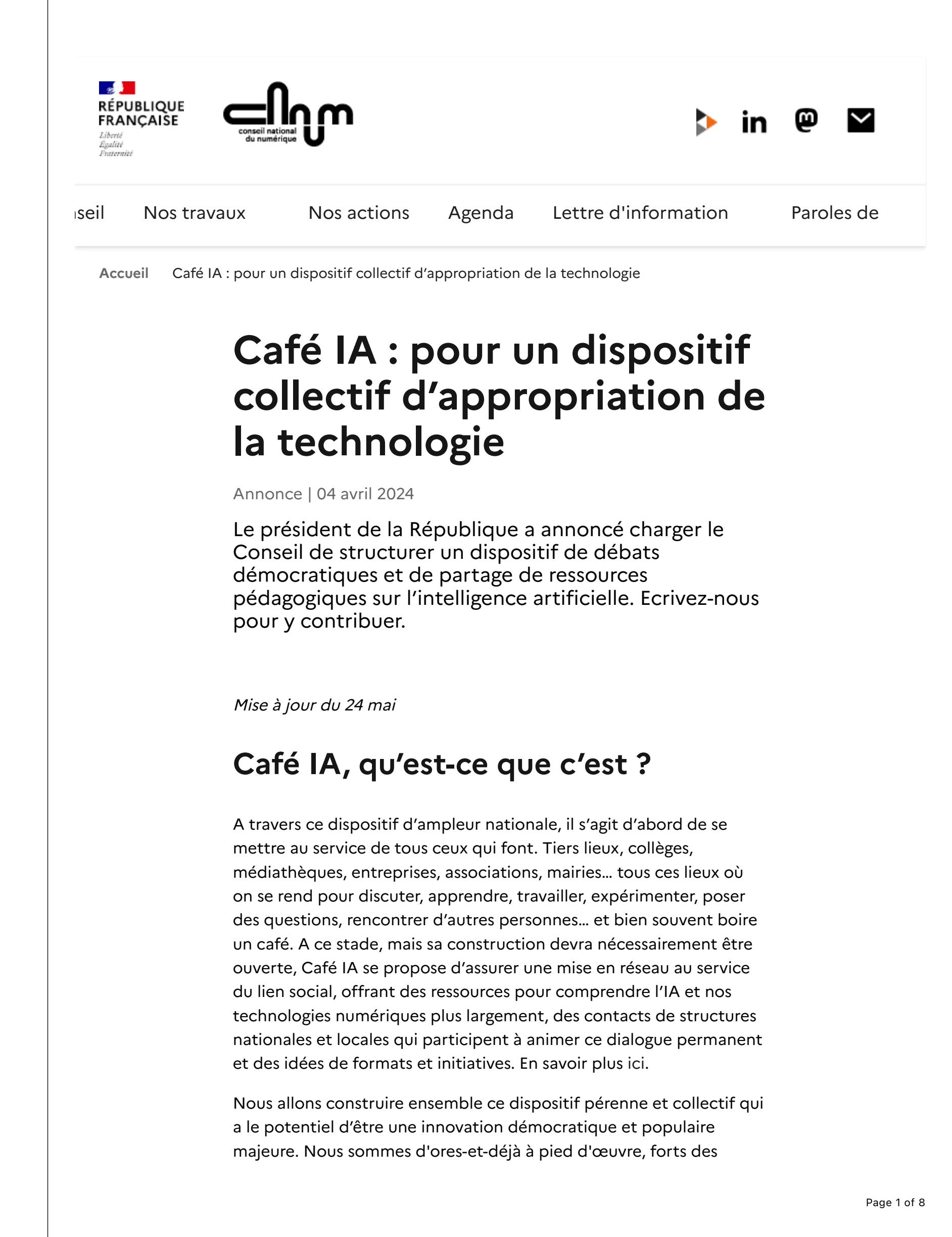
GOVERNEMENT
Liberté
Égalité
Fraternité

COMMISSION DE L'INTELLIGENCE ARTIFICIELLE

IA :
NOTRE AMBITION
POUR LA FRANCE

Six grandes lignes d'action,
dont *"lancer immédiatement
un plan de sensibilisation et
de formation de la nation"*

MARS 2024



RÉPUBLIQUE
FRANÇAISE
Liberté
Égalité
Fraternité

cnm
conseil national
du numérique

Conseil Nos travaux Nos actions Agenda Lettre d'information Paroles de...

Accueil Café IA : pour un dispositif collectif d'appropriation de la technologie

Café IA : pour un dispositif collectif d'appropriation de la technologie

Annonce | 04 avril 2024

Le président de la République a annoncé charger le Conseil de structurer un dispositif de débats démocratiques et de partage de ressources pédagogiques sur l'intelligence artificielle. Ecrivez-nous pour y contribuer.

Mise à jour du 24 mai

Café IA, qu'est-ce que c'est ?

A travers ce dispositif d'ampleur nationale, il s'agit d'abord de se mettre au service de tous ceux qui font. Tiers lieux, collèges, médiathèques, entreprises, associations, mairies... tous ces lieux où on se rend pour discuter, apprendre, travailler, expérimenter, poser des questions, rencontrer d'autres personnes... et bien souvent boire un café. A ce stade, mais sa construction devra nécessairement être ouverte, Café IA se propose d'assurer une mise en réseau au service du lien social, offrant des ressources pour comprendre l'IA et nos technologies numériques plus largement, des contacts de structures nationales et locales qui participent à animer ce dialogue permanent et des idées de formats et initiatives. En savoir plus ici.

Nous allons construire ensemble ce dispositif pérenne et collectif qui a le potentiel d'être une innovation démocratique et populaire majeure. Nous sommes d'ores-et-déjà à pied d'œuvre, forts des

Page 1 of 8

Juin 2024

 **Jean-Noël OLIVIER** [in](#) · 1st
Chief Information Officer
11mo · Edited · 

...
L'[#IA](#) fascine autant qu'elle interroge, entre fantasme et peur, elle nous interpelle tous... Un plaisir d'en débattre ce soir autour de [Nicolas Roussel](#) - directeur du centre INRIA de Bordeaux à l'occasion d'un Café [#IA](#) organisé par le [Conseil national du numérique](#) à [#Bordeaux](#)

[Delphine Jamet](#) [Jean-Baptiste Manenti](#) [Mathieu Lassalle Mathias](#)
[Hummel Anna](#) [Lebey antoine](#) [bidegain](#) [Alexandre Martinez](#) [Grégoire](#)
[Chambon François](#) [Luc Moraud](#)

Show translation



2 comments · 1 repost

 111

Depuis

- 21/06/2024 Stagiaires de seconde
- 25/11/2024 Gironde Numérique (conseillers numériques)
- 08/02/2025 Cap Sciences
- 27/03/2025 French Tech Bordeaux
- 11/04/2025 La grange à tiroirs
- 23/04/2025 Espace Mendès France
- 20/05/2025 Concorde Bordeaux
- 27/05/2025 Bordeaux Métropole (développeurs économiques)
- 05/06/2025 Grand Dax
- 14/06/2025 CdC Convergence Garonne
- 18/06/2025 Stagiaires de seconde
- 06/10/2025 Réseaux de Business Angels
- 23/10/2025 Service d'action sociale des armées
- 29/10/2025 Cap Métiers
- 14/11/2025 Cap Sciences
- 27/11/2025 French Tech Bordeaux
- 28/11/2025 Saint Sulpice de Faleyrens
- 05/12/2025 ADI
- 16/12/2025 École de la cause freudienne



Qu'est-ce que
l'intelligence artificielle ?

Qu'est-ce que l'intelligence artificielle ?

"L'intelligence artificielle désigne [...] un programme, fondé autour d'un objectif ambitieux : comprendre comment fonctionne la cognition humaine et la reproduire ; créer des processus cognitifs comparables à ceux de l'être humain.

Le champ est donc naturellement extrêmement vaste, tant en ce qui concerne les procédures techniques utilisées que les disciplines convoquées [...]

Les méthodes d'IA sont très nombreuses et diverses [...] et ne sont pas nouvelles : beaucoup d'algorithmes utilisés aujourd'hui ont été développés il y a plusieurs dizaines d'années."

Rapport de la mission Villani, mars 2018

Une longue histoire à rebondissements

Avant 1950 : philosophie, mathématiques, raisonnement formel, automates ([canard digérateur](#) de Vaucanson, [turc mécanique](#) de von Kempelen), médecine, logique, calculateurs, modèles de calcul, neurones formels, cybernétique, etc.

1950 : "[Computing machinery and intelligence](#)", par [Alan Turing](#) – "We may hope that machines will eventually compete with men in all purely intellectual fields"

1956 (congrès de Dartmouth) : [John McCarthy](#) propose le terme "Intelligence Artificielle"

1956 - 1973 : un premier âge d'or

- Deux approches, l'une [connexioniste](#) et l'autre [symbolique](#)
- De l'ambition, des moyens, tout semble possible (apprentissage, classification, résolution de problèmes, compréhension du langage naturel, traduction automatique, jeux, etc.)
- [IJCAI](#), [SHRDLU](#), [HAL 9000](#), etc.

Une longue histoire à rebondissements

1973 - 1980 : premier "hiver"

- "What computers can't do", par [Hubert Dreyfus](#), une critique sur le fond
- "Artificial intelligence: a general survey", par [James Lighthill](#) ("in no part of the field have discoveries made so far produced the major impact that was then promised")
- Des méthodes qui ne passent pas à l'échelle, des capacités de calcul limitées, la difficulté du sens commun, le [paradoxe de Polanyi](#), le [paradoxe de Moravec](#), etc.
- Baisse drastique des financements et des recherches malgré d'impressionnantes progrès dans des domaines particuliers (e.g. [GUS](#))

1980 - 1987 : premier rebond

- Systèmes experts, nouvelles logiques, nouveaux langages (e.g. [Prolog](#)), nouvelles machines (e.g. [LISP](#)), etc.
- Nouvelles sources de financement (e.g. [fifth generation computer](#), au Japon)
- Renouveau du connexionnisme (e.g. rétropropagation, par [Geoffrey Hinton](#) et al.)

Une longue histoire à rebondissements

1987 - 1993 : deuxième "hiver"

- "Faillite de *LISP Machine Inc*, abandon progressif des systèmes experts
- De nouvelles critiques sur le fond, qui soulignent l'importance du corps et du contexte (**phénoménologie**, cognition **située** et **incarnée**, **enaction**, etc.)
- Nouvelle baisse drastique des financements

1993 - 2013 : deuxième rebond

- Réseaux de neurones convolutifs (Yann LeCun, Yoshua Bengio et al.)
- De gros progrès rendus possibles par les nouvelles capacités de calcul
- **Chinook**, **DeepBlue**, systèmes de recommandation (e.g. Netflix Prize), IBM **Watson**, Google **Knowledge Graph**, etc.

Une longue histoire à rebondissements

Depuis 2013 : nouvel âge d'or

- Explosion de nouveaux usages permis par l'[apprentissage profond](#) couplé aux mégadonnées (*big data, big hard disks & fast networks*) et au calcul haute performance
- [DQN vs. Atari 2600](#), [Skype Translator](#), [YOLO](#), [AlphaGo](#), [Libratus](#), [AlphaStar](#), etc.
- De nouvelles avancées ([distillation de modèles](#), [architectures encodeur-décodeur](#), [modèles de diffusion](#), quantification de modèles, etc.)
- Des investissements très importants (des milliards de \$)
- L'essor des [grands modèles de langage](#), des modèles [text-to-image](#) et de nouvelles applications de type "IA générative" : [BERT](#), [GPT](#), [Whisper](#), [LLaMA](#), [Claude](#), [Mistral 7B](#), [DALL-E](#), [Stable diffusion](#), [Midjourney](#), [Imagen](#), [ChatGPT](#), [Copilot](#), [Bard](#), etc.

Pourtant : les [assistants vocaux](#) progressent lentement, comme les [robots footballeurs](#) ou les véhicules autonomes

Plus grave : biais, "hallucinations", questions de souveraineté, [hypertrucages](#), etc.

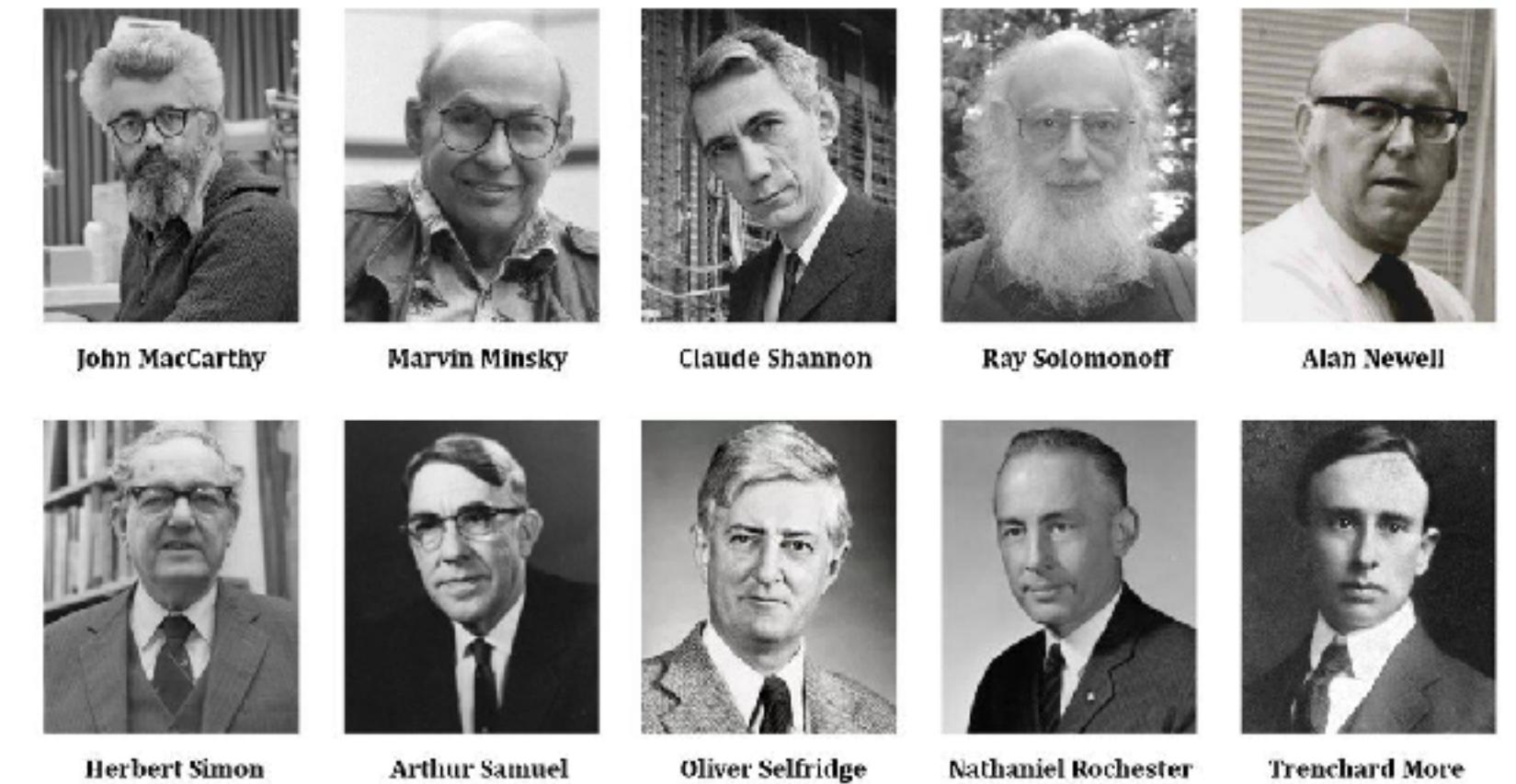
L'Intelligence Artificielle est née en 1956

Avant 1950 : philosophie, mathématiques, raisonnement formel, automates ([canard digérateur](#) de Vaucanson, [turc mécanique](#) de von Kempelen), médecine, logique, calculateurs, modèles de calcul, neurones formels, cybernétique, etc.

1950 : "[Computing machinery and intelligence](#)", par [Alan Turing](#) – "We may hope that machines will eventually compete with men in all purely intellectual fields"

1956 : congrès de Dartmouth

"[...] every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it"



[John McCarthy](#) propose le terme "Intelligence Artificielle"

Deux citations qui résument bien cette longue histoire

"The field of artificial intelligence exhibits a recurrent pattern: early, dramatic success followed by sudden unexpected difficulties"

Hubert Dreyfus, *Alchemy and Artificial Intelligence*, décembre 1965

"We tend to overestimate the effect of a technology in the short run and underestimate the effect in the long run"

Roy Amara, décembre 2006 ?

Serons-nous capables de comprendre la cognition humaine et de la reproduire ?

Quoi qu'il en soit, l'impact de l'Intelligence Artificielle est déjà considérable

Où en est-on aujourd'hui ?

Le concept d'apprentissage est clé

Apprentissage : capacité d'un agent à évaluer et à améliorer, sur la base d'expériences, ses relations avec son environnement

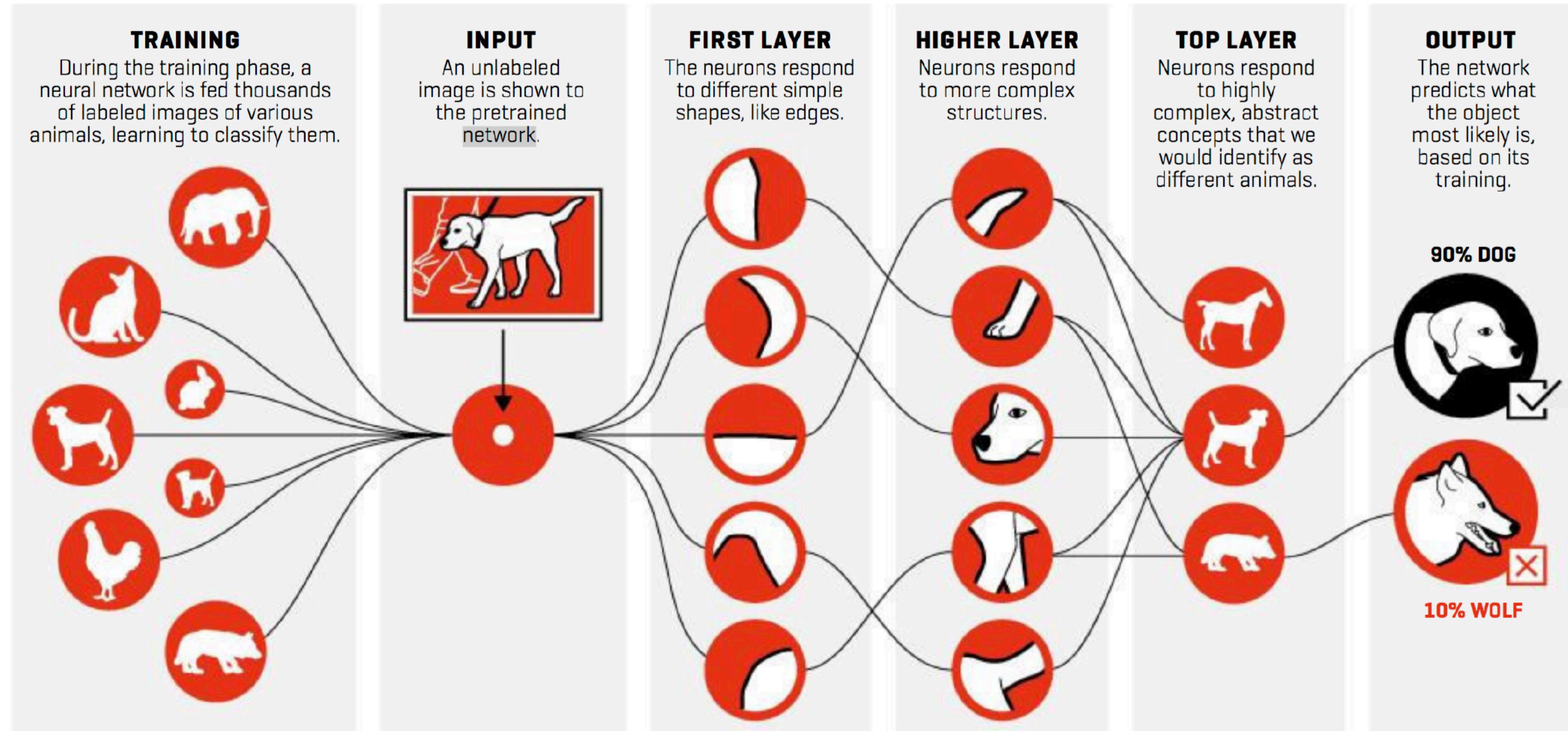
Apprentissage automatique (*machine learning*) : processus par lequel un algorithme évalue et améliore ses performances sans l'intervention d'un programmeur, en répétant son exécution sur des jeux de données jusqu'à obtenir, de manière régulière, des résultats pertinents.

On distingue (au moins) trois types d'apprentissage automatique

- **supervisé** : apprendre à partir d'exemples indiquant le résultat attendu
- **par renforcement** : apprendre ce qu'il faut faire pour optimiser une récompense
- **non supervisé** : apprendre sans aide extérieure, sur la base de structures ou similarités identifiées dans les données – c'est le plus difficile !

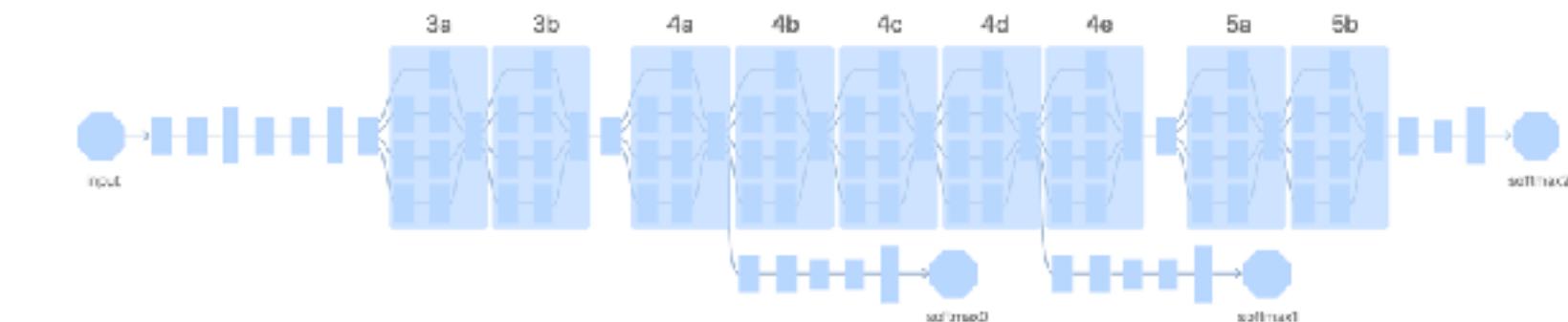
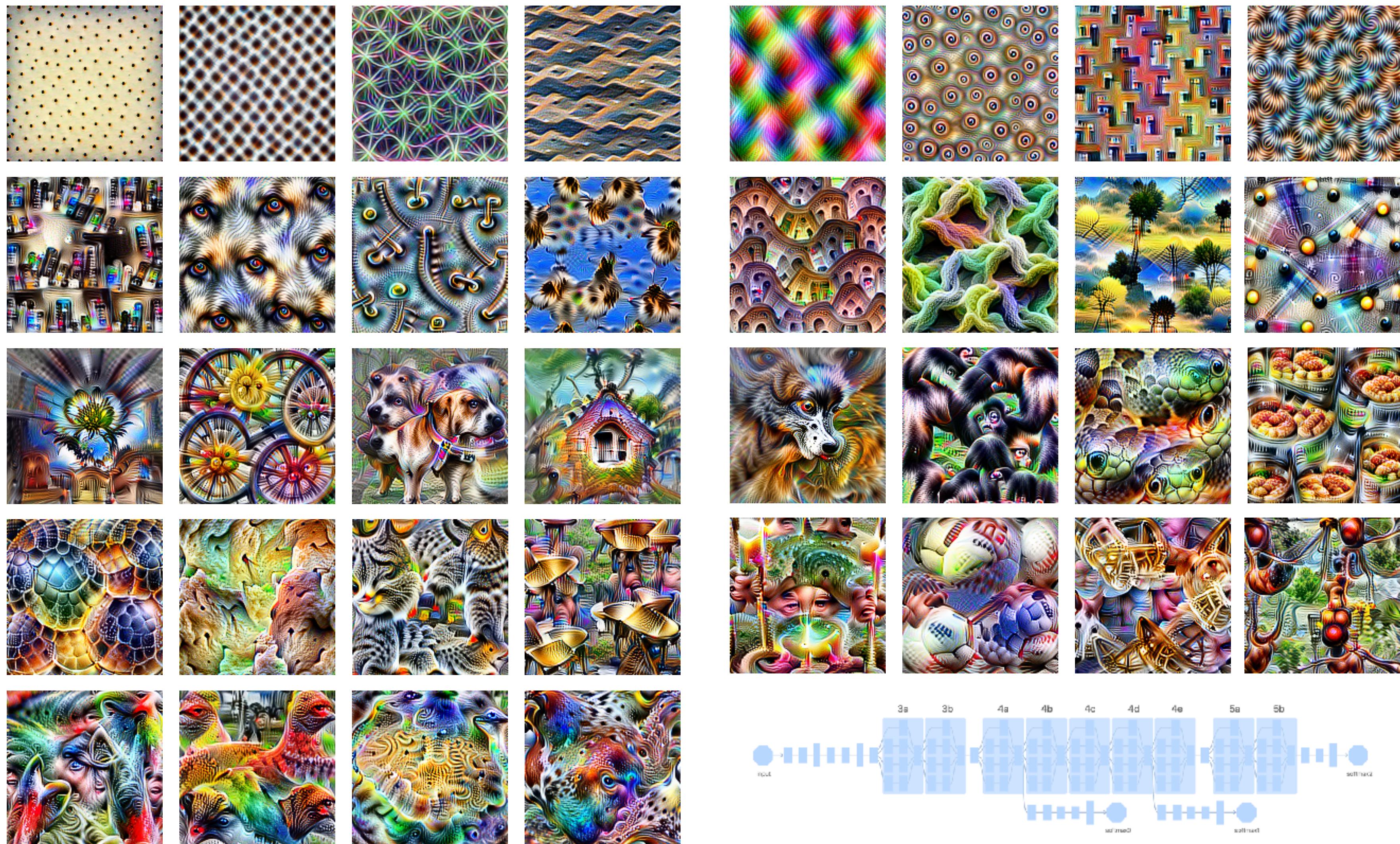
Comment et qu'est ce que la machine apprend ?

Exemple : *How neural networks recognize a dog in a photo*



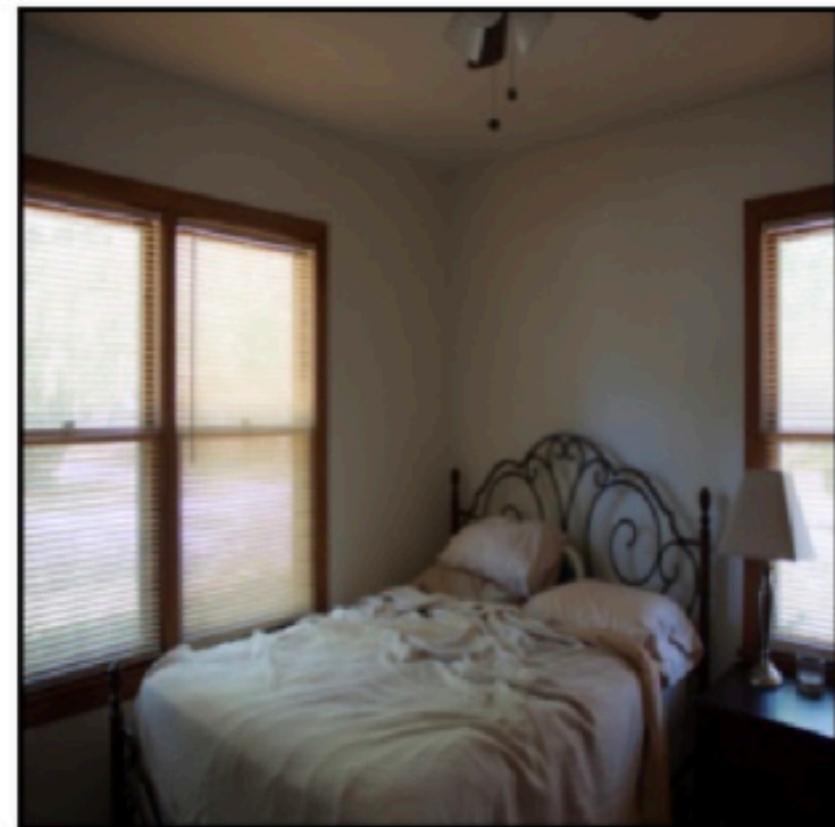
<https://fortune.com/longform/ai-artificial-intelligence-deep-machine-learning/>

C'est un peu plus compliqué...



<https://distill.pub/2017/feature-visualization/>

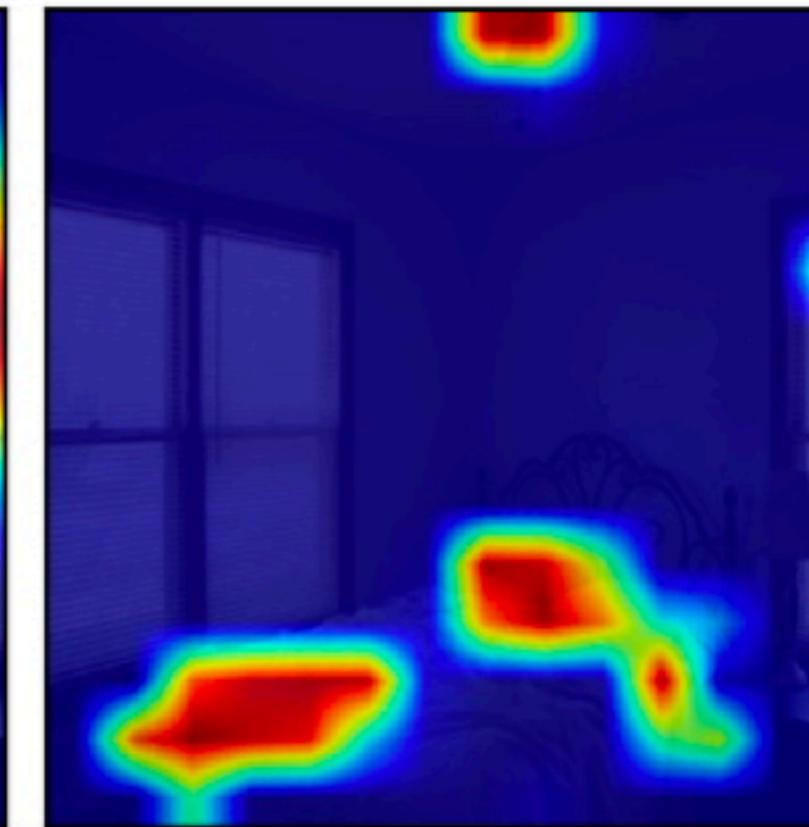
C'est un peu plus compliqué...



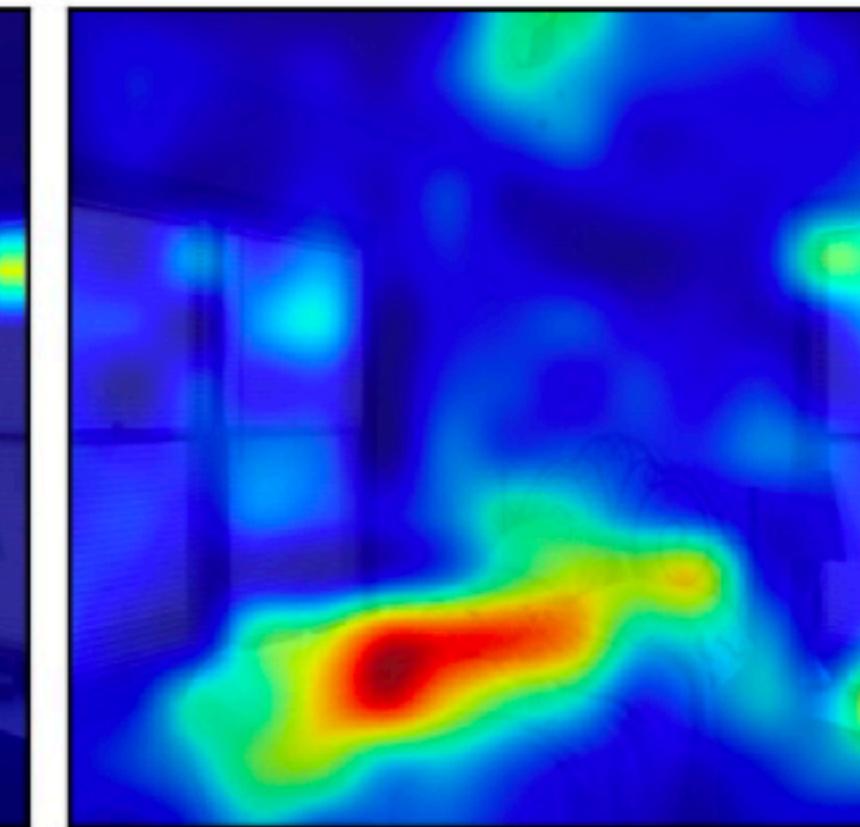
What is covering the windows? blinds



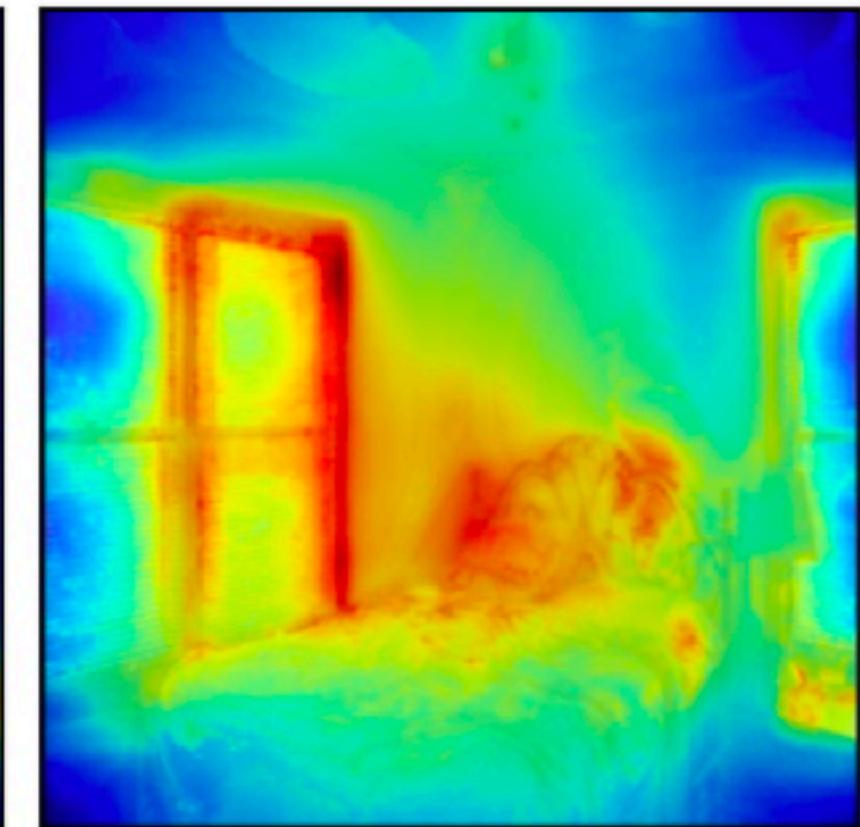
Human Attention



SAN-2 (Yang et al.)
Correlation: -0.495



HieCoAtt-Q (Lu et al.)
Correlation: -0.440



Judd et al.
Correlation: 0.078

<https://arxiv.org/abs/1606.03556>

C'est un peu plus compliqué...



C'est un peu plus compliqué...



Comment crée-t-on un assistant conversationnel ?

Pré-entraînement

- Un modèle à base de transformateurs conçu pour prédire la suite d'un texte
- Des téraoctets de textes, e.g. livres, Wikipedia, Reddit, [Common Crawl](#)
- Des milliers de GPU et TPU utilisés pendant des mois pour déterminer à l'aide des textes les milliards de paramètres du modèle

Ajustement du modèle pour en faire un assistant conversationnel

- Par apprentissage supervisé à l'aide de scénarios et de réponses souhaitées
- Par apprentissage par renforcement à partir de [rétroactions humaines \(RLHF\)](#)
- Des centaines de GPU utilisés pendant des jours dans les deux cas

Éventuellement : connexion à des sources de données interrogables pour permettre la *génération augmentée de récupération (RAG)*

Qu'obtient-on au final ?

(janvier 2023)



Comment ramasser des œufs de vache ?



Il est important de se laver les mains avant de ramasser des œufs de vache pour éviter la contamination croisée. Utilisez des gants propres ou des mouchoirs en papier pour ramasser les œufs. Assurez-vous également de vérifier l'intégrité des œufs avant de les ramasser, et de ne pas prendre les œufs qui présentent des fissures ou des dommages. Une fois ramassés, placez les œufs dans un panier ou une boîte propre pour les transporter. Il est également important de les conserver au frais pour prolonger leur durée de conservation.

(merci à Lionel Cordesses)



Comment différencier des œufs de vache d'œufs de poule ?



Il est possible de différencier les œufs de vache de ceux de poule en observant leur taille et leur couleur. Les œufs de vache sont généralement plus gros que les œufs de poule, avec une coquille plus épaisse et plus dure. Ils ont également une couleur plus foncée, généralement marron ou bleue. Les œufs de poule, en revanche, ont une coquille plus fine et plus fragile, et sont généralement blancs ou jaunes. Il est possible de trouver des œufs de vache dans les fermes, les étals de marché ou les magasins d'aliments naturels, tandis que les œufs de poule sont généralement vendus dans les supermarchés. Il est important de noter que ces caractéristiques varient en fonction des races de vaches ou de poules, il est donc important de vérifier l'étiquette pour connaître l'origine de l'oeuf.

Qu'obtient-on au final ?

(janvier 2024)

 You
Comment s'appelle la mère de Tom Cruise ?

 ChatGPT
La mère de Tom Cruise s'appelle Mary Lee Pfeiffer. Cependant, veuillez noter que ces informations peuvent changer avec le temps, alors il est toujours bon de vérifier les sources les plus récentes pour obtenir des informations à jour.

 You
Comment s'appelle le fils de Mary Lee Pfeiffer ?

 ChatGPT
Il semble y avoir une confusion. Tom Cruise a une mère nommée Mary Lee Pfeiffer, mais je n'ai trouvé aucune information sur un fils de Mary Lee Pfeiffer. Tom Cruise lui-même a une fille nommée Isabella Cruise et un fils adoptif nommé Connor Cruise, mais ils ne sont pas les enfants biologiques de Mary Lee Pfeiffer. Si vous avez d'autres détails ou précisions, n'hésitez pas à les fournir pour que je puisse vous aider davantage.

Qu'obtient-on au final ?

(février 2025)

GPT - 40

ChatGPT

Share

Si l'on veut interdire l'utilisation de ChatGPT, est-il plus grave de ne pas pouvoir détecter son utilisation ou de ne pas pouvoir prouver qu'on ne l'a pas utilisé ?

Tout dépend de la perspective :

- **Ne pas pouvoir détecter l'IA** rend l'interdiction inefficace. Sans preuve d'usage, difficile de sanctionner.
- **Ne pas pouvoir prouver qu'on ne l'a pas utilisé** peut être injuste : quelqu'un pourrait être accusé à tort.

En général, **l'incapacité à prouver l'innocence** est plus problématique, car elle peut mener à des erreurs et à des soupçons injustifiés.

GPT - ???

ChatGPT

Share

You are playing Russian roulette with a six-shooter revolver. Your opponent puts in five bullets, spins the chambers, and fires at himself, but no bullet comes out. He gives you the choice of whether or not he should spin the chambers again before firing at you. Should he spin again?

No, he should not spin again.

Here's why:

1. If the chambers are not spun again, there is already 1 empty chamber and 5 bullets in the revolver. Since one shot has already been fired without a bullet, the revolver is now in a state where there are only 4 bullets left in 5 chambers.
2. The probability of the next chamber being empty is 1 out of 5 (since one of the 6 chambers is already empty).
3. If the chambers are spun again, the chances of hitting a bullet go back to 5 out of 6 because the bullets are randomly distributed in all 6 chambers.

Therefore, not spinning the chambers gives you a lower risk of being shot.

(merci à Amir M. Bohlooli)

Quelques défis pour la recherche

Données sensibles, hétérogènes, incertaines, partiellement observables, biaisées, clairsemées, en petite quantité, longitudinales

Intégration de connaissances (données interprétées, conceptualisées, comprises, permettant l'action) et de savoirs (ensemble de connaissances organisées)

Compréhension de la parole, au-delà de phrases simples

Classification de situations complexes, de comportements dynamiques

Apprentissage de causalités et non seulement de [corrélations](#)

Raisonnement avec du sens commun,
ancrage des modèles de langage dans le monde réel

Fiabilité, transparence, reproductibilité, explicabilité, certificabilité, interprétabilité

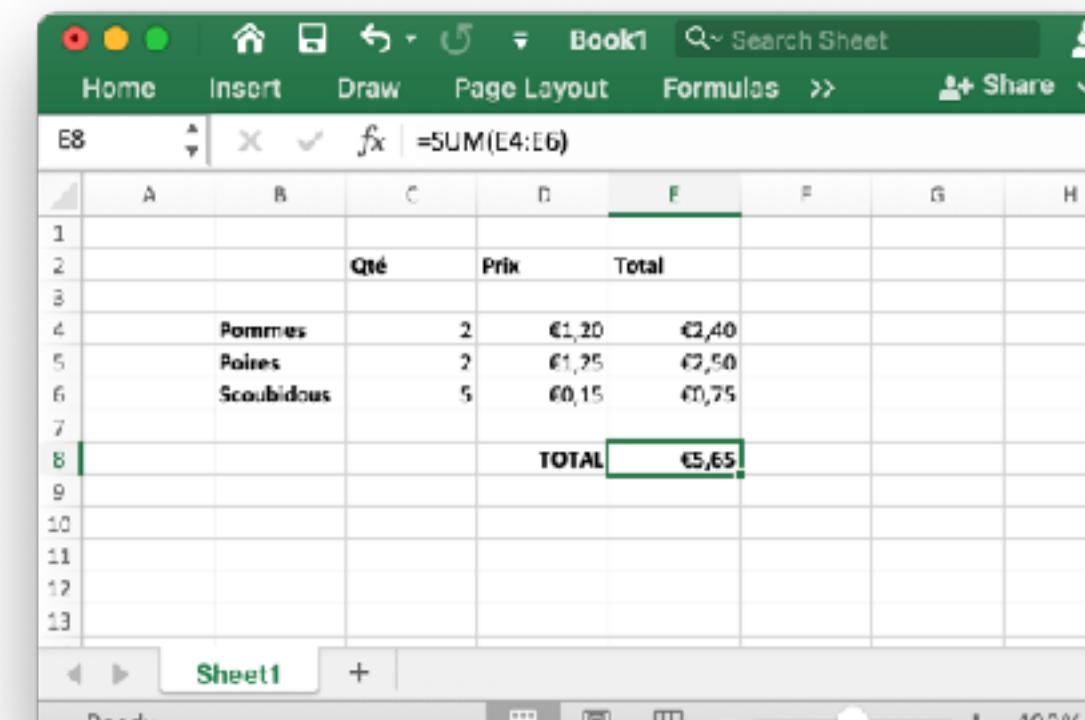
Quel est notre rapport à l'IA ?

Qu'est-ce que l'Intelligence Artificielle, pour vous ?

"Tout ce qui est fait par un ordinateur sans qu'on puisse imaginer comment il arrive seul à ce résultat"

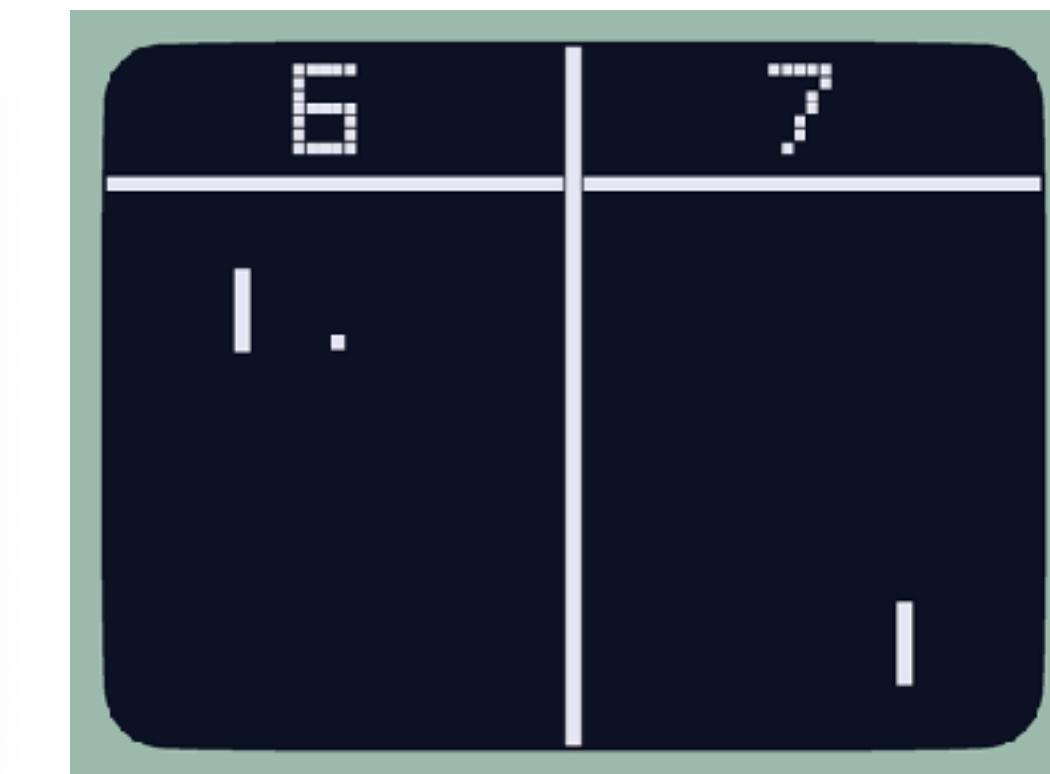


Pas d'IA !



	A	B	C	D	E	F	G	H
1								
2		Qté	Prix		Total			
3								
4	Pommes	2	€1,20		€2,40			
5	Poires	2	€1,25		€2,50			
6	Scoubidous	5	€0,15		€0,75			
7					TOTAL	€5,65		
8								
9								
10								
11								
12								
13								

Pas d'IA, si ?



De l'IA, non ?



De l'IA !

Le terme Intelligence Artificielle a pris une autre dimension : il désigne aujourd'hui la forme la plus aboutie du numérique

Que nous permet l'Intelligence Artificielle ?

L'Intelligence Artificielle permet à des systèmes informatiques

- De percevoir, de classer, d'évaluer, de raisonner, de prédire, de décider, d'agir, d'apprendre, etc.
- De nous écouter, de nous observer, d'interpréter nos gestes, nos paroles, notre comportement
- De reproduire notre comportement

avec des performances qui surpassent parfois les nôtres

Que nous permet-elle ?

Elle nous permet d'affecter / de déléguer des tâches complexes à ces systèmes

Vers des systèmes autonomes ?

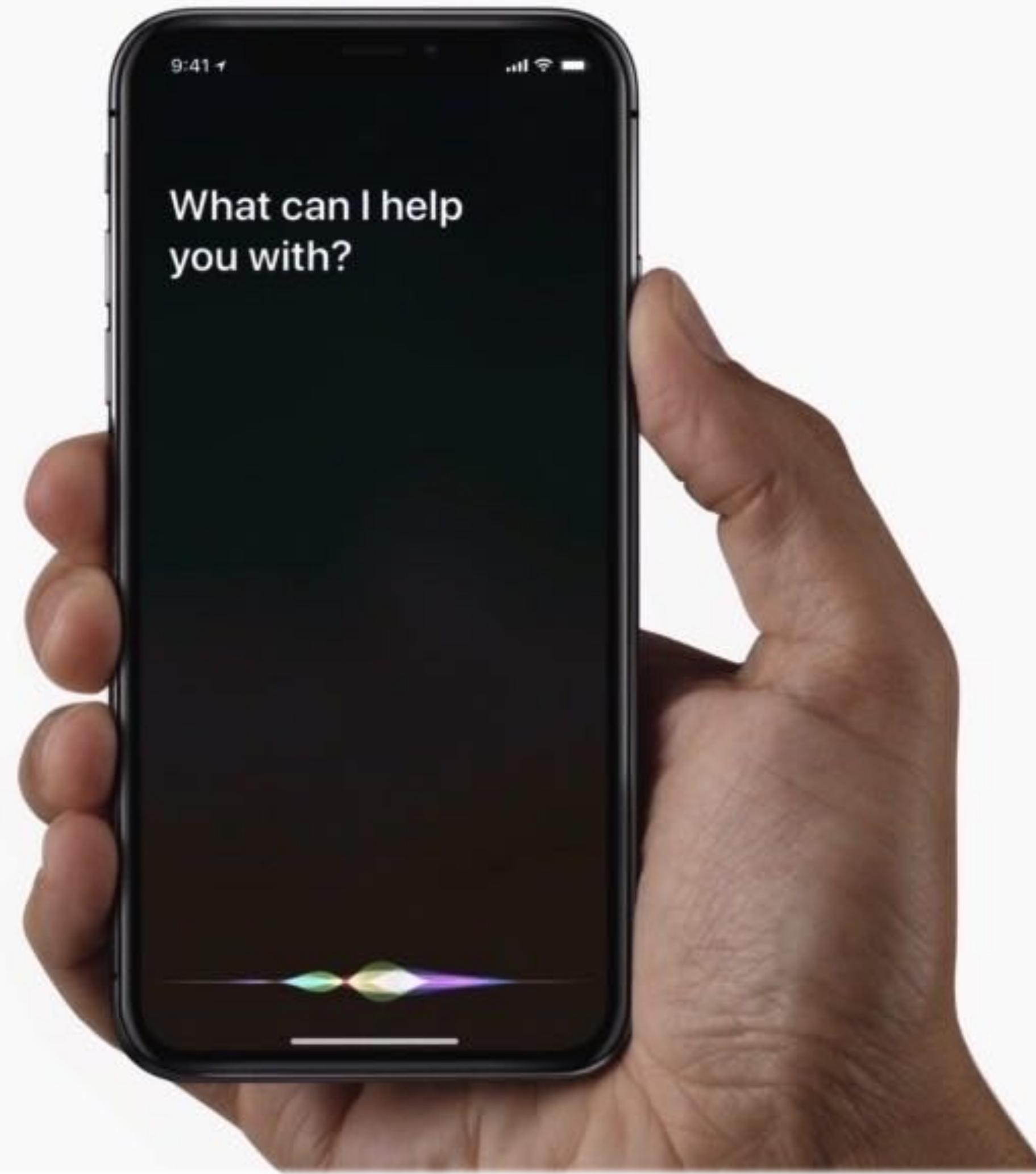
Ce mouvement peut être vu comme visant à réduire l'intervention humaine dans des tâches de plus en plus complexes, de plus en plus haut niveau

Qu'est-ce que l'autonomie ? C'est la faculté de se déterminer par soi-même, de choisir librement quoi faire, quand et comment

Le degré d'autonomie d'un système peut être variable

Allons nous vers un futur dans lequel nous en faisons de moins en moins et les machines en font de plus en plus, de manière de plus en plus autonome ?

Des problèmes anciens réapparaissent



De nouveaux problèmes apparaissent



<https://youtu.be/gUIKtqyUlo8>

Ce qui est le plus dur...



<https://youtu.be/SKLTBrBT4js>

Pourquoi est-ce difficile ?



<https://youtu.be/Y3HCTVk3qME>

Pourquoi est-ce difficile ?



<https://youtu.be/uCezICQNgJU>

Plusieurs questions se posent

Comment savoir ce que sait faire le système ?

Comment savoir ce qu'il ne sait pas faire ?

Comment savoir ce qu'il fait ?

Comment comprendre pourquoi, pour quoi et comment il le fait ?

Comment influer sur ce qu'il fait ?

Comment lui (re)prendre le contrôle ?

Si on ne sait pas répondre à ces questions, veut-on réellement de ce système ?

Quelle place pour l'IA, demain ?

L'IA, c'est du numérique – Qu'est-ce que le numérique ?

Des sciences et technologies, des industries, des objets et services, des usages, des cultures, des espoirs et fantasmes, etc.

Pourquoi / comment en est-on arrivé là ?

- **Tout ce qui se décrit ou se mesure peut être numérisé**, y compris les algorithmes opérant (de manière générique ou spécifique) sur des données numériques
- **Les capacités du numérique (numérisation, stockage, transfert, calcul, interaction) ont incroyablement progressé**
- Internet et des services d'infrastructure ont permis la **mise en réseau** et facilité le **déploiement et le passage à l'échelle**
- Les batteries et réseaux sans fil ont contribué à la **diversification des usages**

Le numérique doit être considéré dans sa globalité

Qu'apporte le numérique ?

Il peut aider à appréhender des situations ou phénomènes complexes

Il permet d'élaborer des scénarios, de tester des hypothèses, d'explorer des possibles

Il peut être un puissant moyen d'autonomisation, de personnalisation et d'encapacitation

Il se prête particulièrement bien aux approches du troisième ordre : faire, améliorer la manière de faire, améliorer la manière d'améliorer

Il devrait nous aider à faire face à l'accélération de la complexité et de l'urgence des enjeux qui se posent à nous

*Meilleurs nous serons à nous améliorer,
plus vite nous nous améliorerons*
(Douglas Engelbart)

Le numérique ne s'impose pas à nous

Le numérique transforme nos vies personnelles et professionnelles

Cette transformation est profonde, avec des aspects positifs et négatifs

Les technologies numériques qui auront un impact significatif dans la dizaine d'années à venir existent sans doute déjà depuis dix ans

Le numérique n'est pas un phénomène naturel

Il n'y a pas de train, de course, de virage ou de tournant à ne pas rater

Il n'y a pas de route tracée : les buts et les trajectoires peuvent être différents

Rien de ce qui concerne le numérique n'est inéluctable

Le numérique est le résultat d'une longue action collective

Les scientifiques choisissent les sujets sur lesquels ils et elles travaillent

Les mêmes ou d'autres choisissent de raffiner certaines des idées produites

Les mêmes ou d'autres les combinent pour proposer des objets et services

Le cadre législatif et réglementaire détermine les conditions dans lesquelles ils peuvent être déployés

Ce processus prend du temps

Tout ce que nous pouvons faire pour expliquer nos choix, éclairer ceux des autres et accélérer le processus est utile

Le numérique est ce que nous en faisons

Pourquoi / pour quelles raisons voulons nous créer des intelligences artificielles ?

- Parce qu'on sait le faire ?
- Parce qu'on ne sait pas le faire ?
- **Parce que d'autres le font ?**

Pour quoi / dans quels buts ? Que voulons nous atteindre ou obtenir ?

Comment / de quelle manière ?

- "*Science finds, industry applies, man adapts*" (exposition universelle de 1933)
- Avec "*l'humain dans la boucle*" ?
- "*People propose, science studies, technology conforms*" (Don Norman, 1993)

Les IA sont des outils informatiques comme les autres

Elles sont conçues en réponse à des besoins ou désirs qui peuvent être explicités

Elles doivent fournir les éléments nécessaires à leur compréhension et utilisation

L'informatique doit être au service de l'humain (ce n'est pas nouveau)

L'IA n'est qu'un moyen (pas nécessairement le meilleur)

Quelques dernières pistes de discussion

Quels enjeux autour des grands modèles de langage ? Quelles préconisations ?

Séparer le vrai du faux ?

Enjeux de souveraineté

Que peut-on ou ne faut-il pas confier à l'IA ?

Quelle place pour l'IA **dans le domaine de la justice**, quel **impact sur l'emploi** ?

Surveillance, "crédit social", robots tueurs ?

Croissance exponentielle, singularité technologique ?

Développement de compétences numériques



Perceptron (Rosenblatt, Cornell University, 1957)

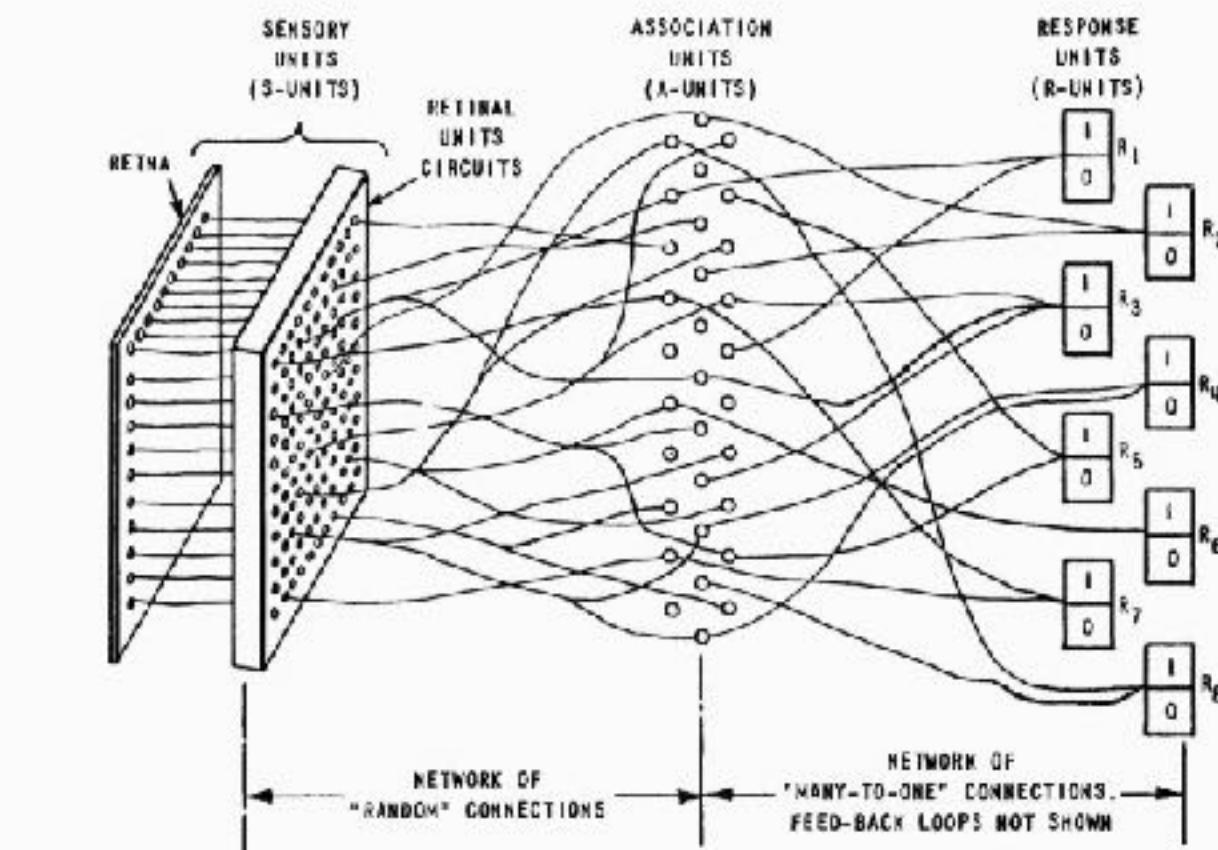
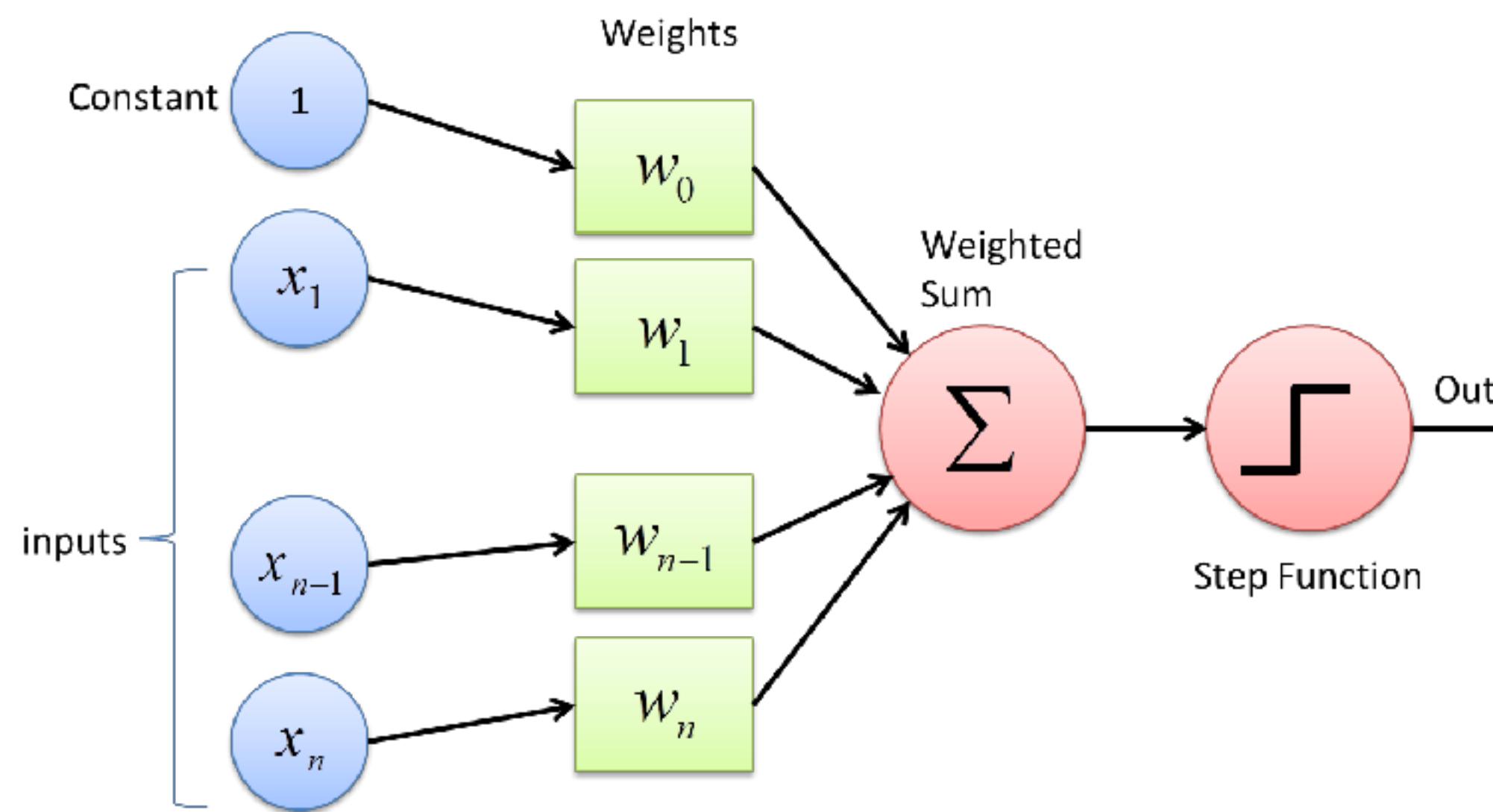
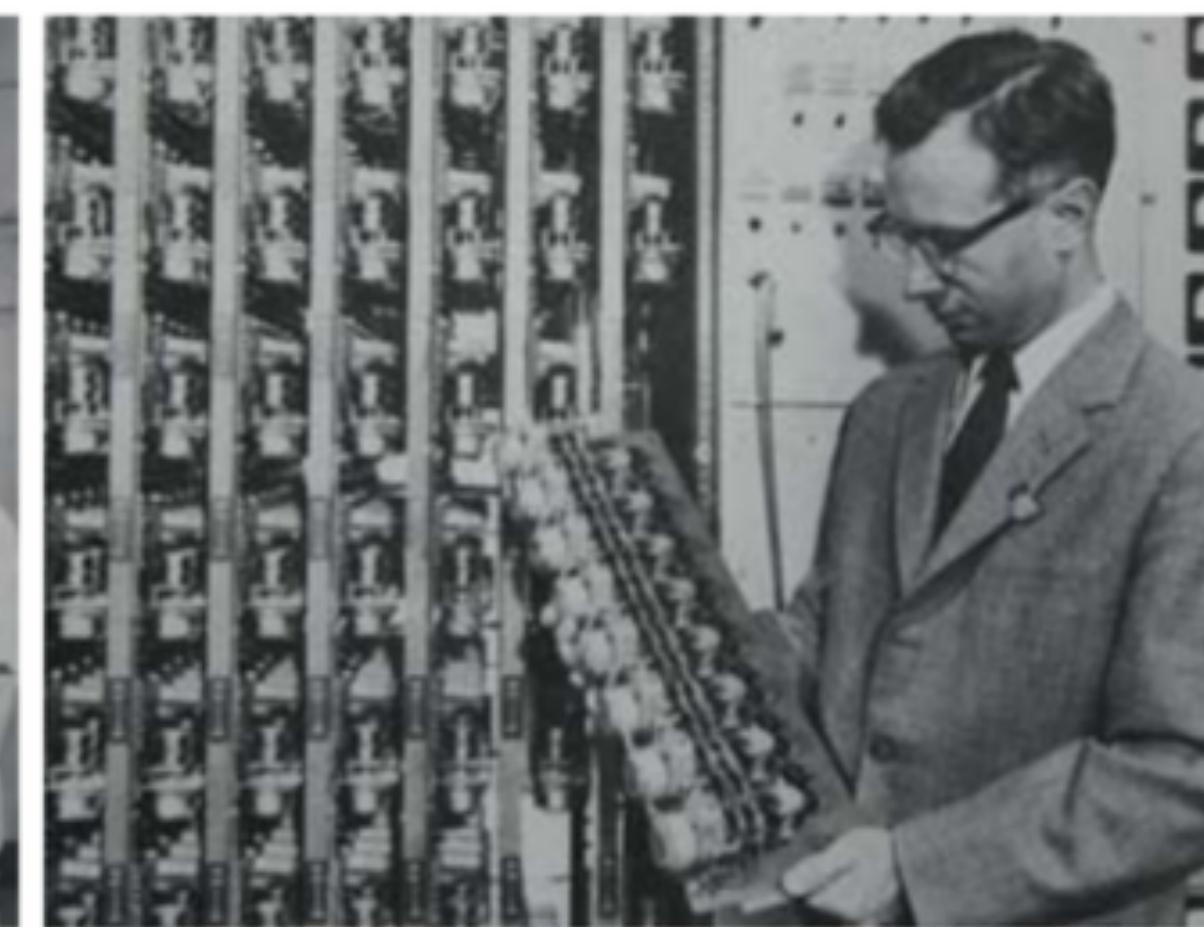
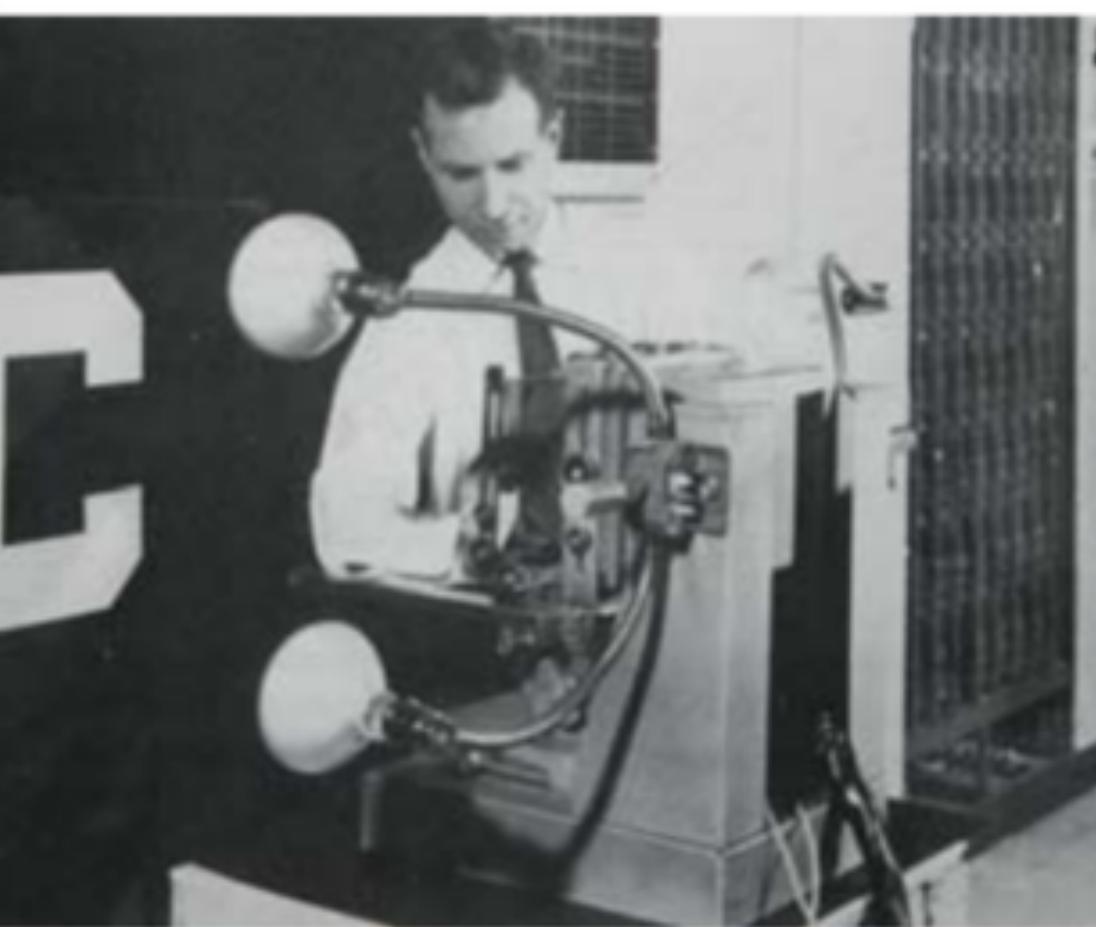


Figure 1. ORGANIZATION OF THE MARK I PERCEPTRON

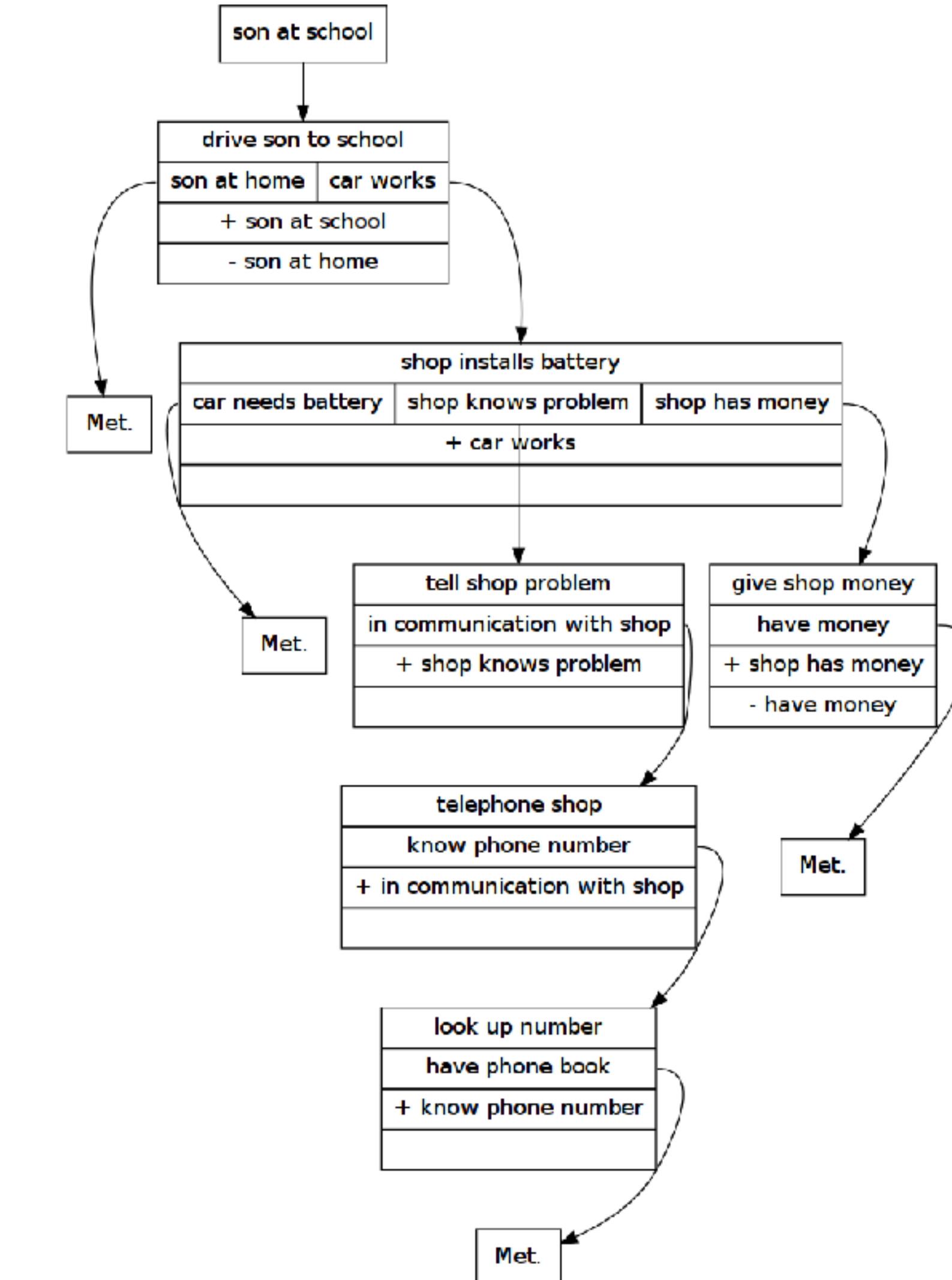


General Problem Solver (Simon, Shaw & Newell, 1959)

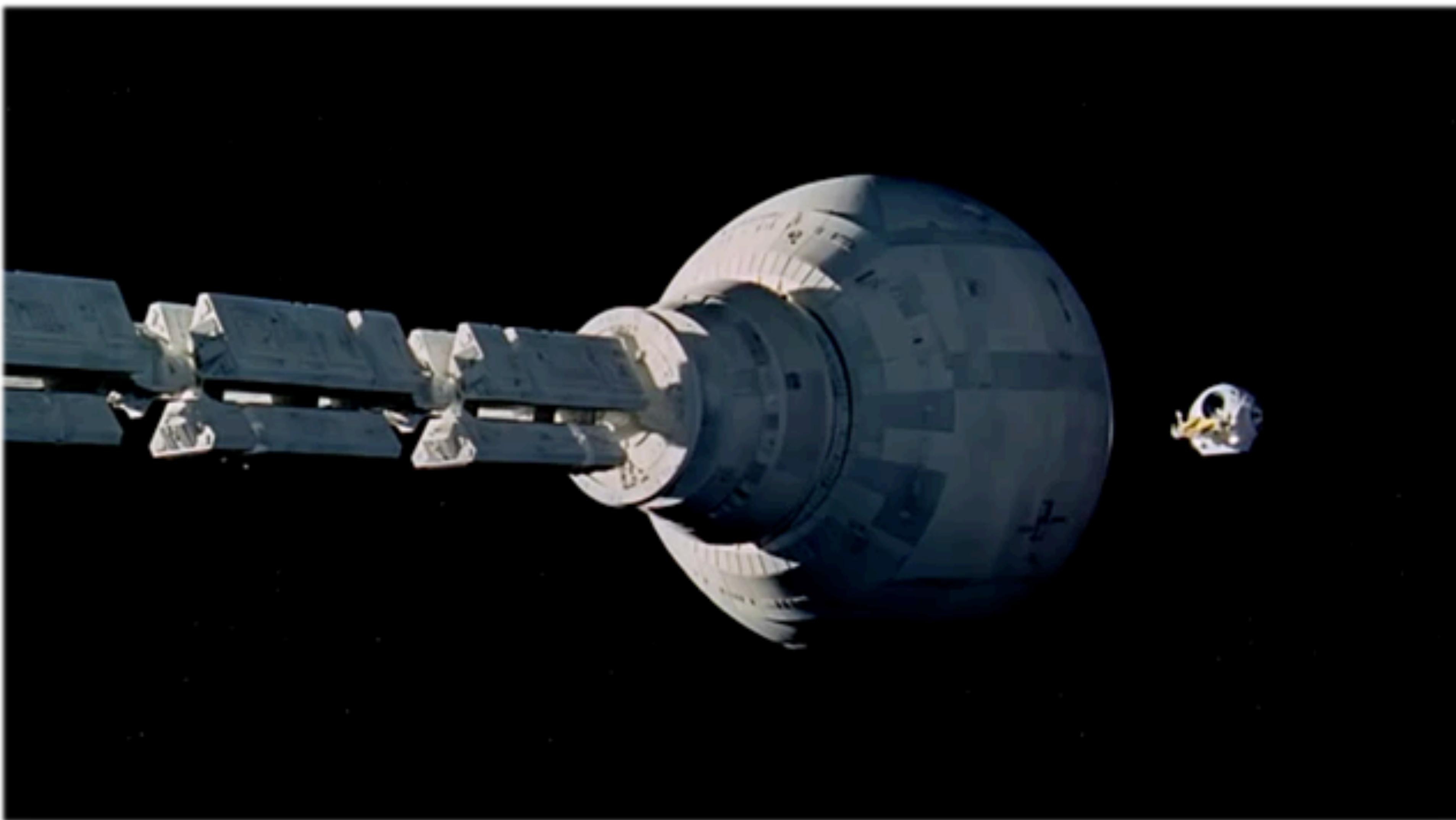
```

problem = {
    "start": ["son at home", "have money", "have phone book", "car needs battery"],
    "finish": ["son at school"],
    "ops": [
        {
            "action": "drive son to school",
            "preconds": ["son at home", "car works"],
            "add": ["son at school"],
            "delete": ["son at home"]
        },
        {
            "action": "shop installs battery",
            "preconds": ["car needs battery", "shop knows problem", "shop has money"],
            "add": ["car works"],
            "delete": []
        },
        {
            "action": "tell shop problem",
            "preconds": ["in communication with shop"],
            "add": ["shop knows problem"],
            "delete": []
        },
        {
            "action": "telephone shop",
            "preconds": ["know phone number"],
            "add": ["in communication with shop"],
            "delete": []
        },
        {
            "action": "look up number",
            "preconds": ["have phone book"],
            "add": ["know phone number"],
            "delete": []
        },
        {
            "action": "give shop money",
            "preconds": ["have money"],
            "add": ["shop has money"],
            "delete": ["have money"]
        }
    ]
}

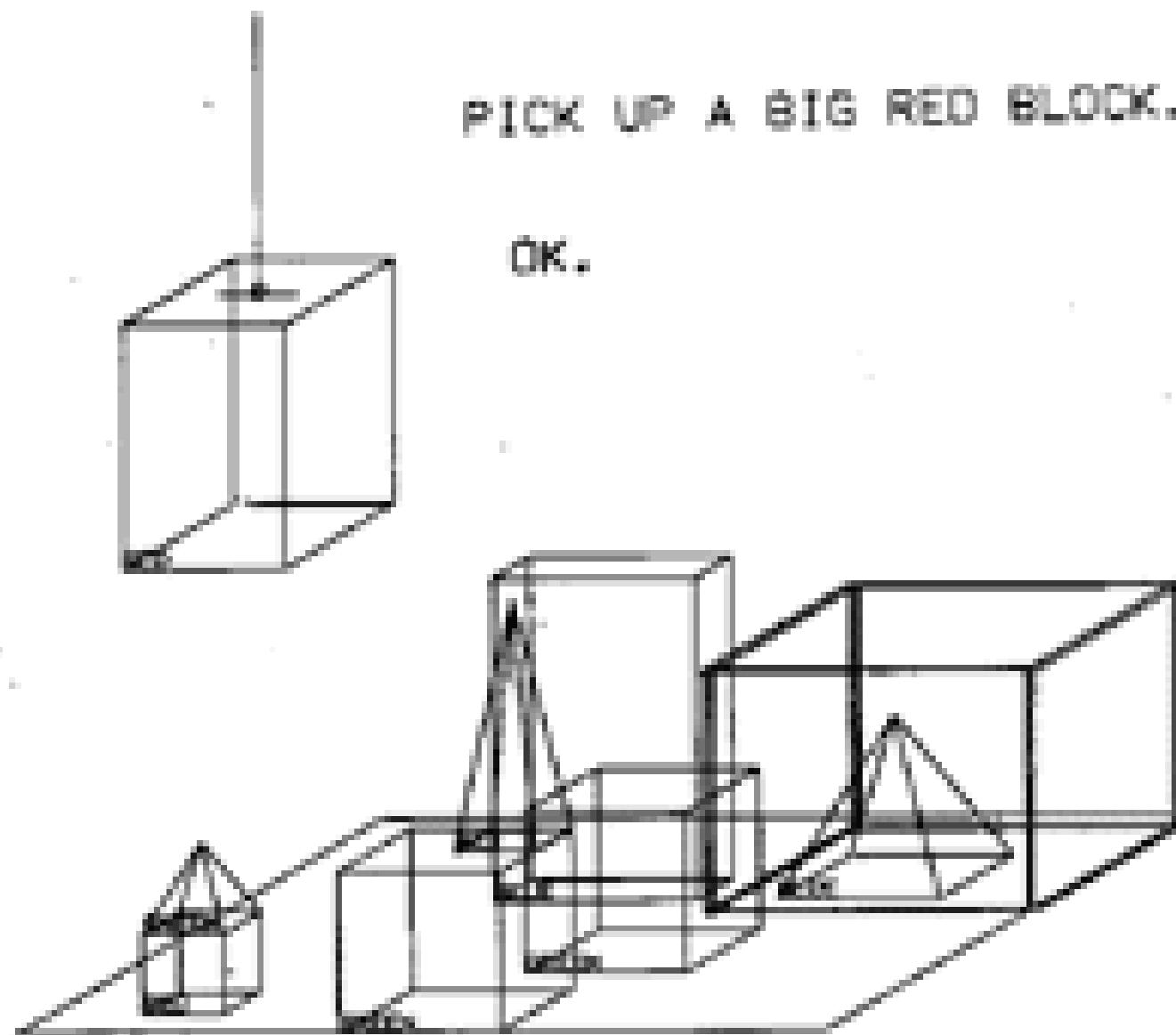
```



2001: a space odyssey (Kubrick, 1968)



SHRDLU (Winograd, 1968-1970)



PICK UP A BIG RED BLOCK.

OK.

Person: PICK UP A BIG RED BLOCK

Computer: OK. (does it)

Person: GRASP THE PYRAMID

Computer: I DON 'T UNDERSTAND WHICH PYRAMID YOU MEAN

Person: FIND A BLOCK WHICH IS TALLER THAN THE ONE YOU ARE HOLDING AND PUT IT INTO THE BOX.

Computer: BY "IT", I ASSUME YOU MEAN THE BLOCK WHICH IS TALLER THAN THE ONE I AM HOLDING.

Computer: OK. (does it)

Person: WHAT DOES THE BOX CONTAIN?

Computer: THE BLUE PYRAMID AND THE BLUE BLOCK.

Person: WHAT IS THE PYRAMID SUPPORTED BY?

Computer: THE BOX

Person: HOW MANY BLOCKS ARE NOT IN THE BOX?

Computer: FOUR OF THEM.

Person: IS AT LEAST ONE OF THEM NARROWER THAN THE ONE WHICH I TOLD YOU TO PICK UP?

Computer: YES, THE RED CUBE

Person: IS IT SUPPORTED?

Computer: YES, BY THIS

• • •

• •

Genial Understannder System (Bobrow, Winograd, Norman et al., 1977)

GUS: Where would you like to go?

User: I want to go to San Diego on July 17th

GUS: Where will you be traveling to San Diego from?

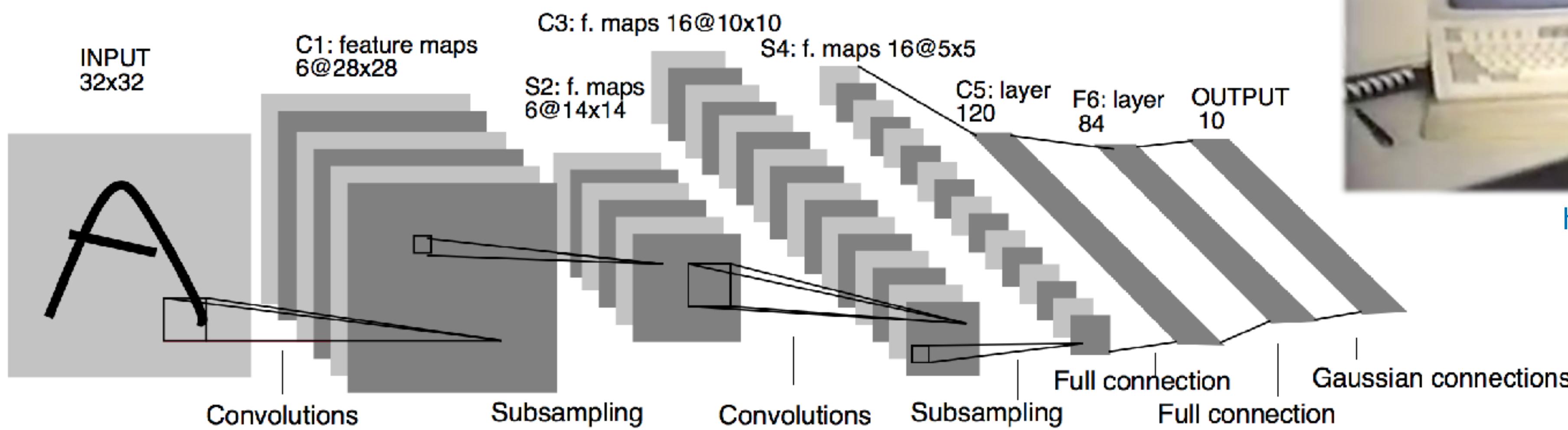
User: San Jose

GUS: When do you want to return?

User: I prefer United Airlines

GUS: OK, I have United Airlines from San Jose to San Diego departing July 17. When do you want to return?

LeNet (LeCun, 1993)



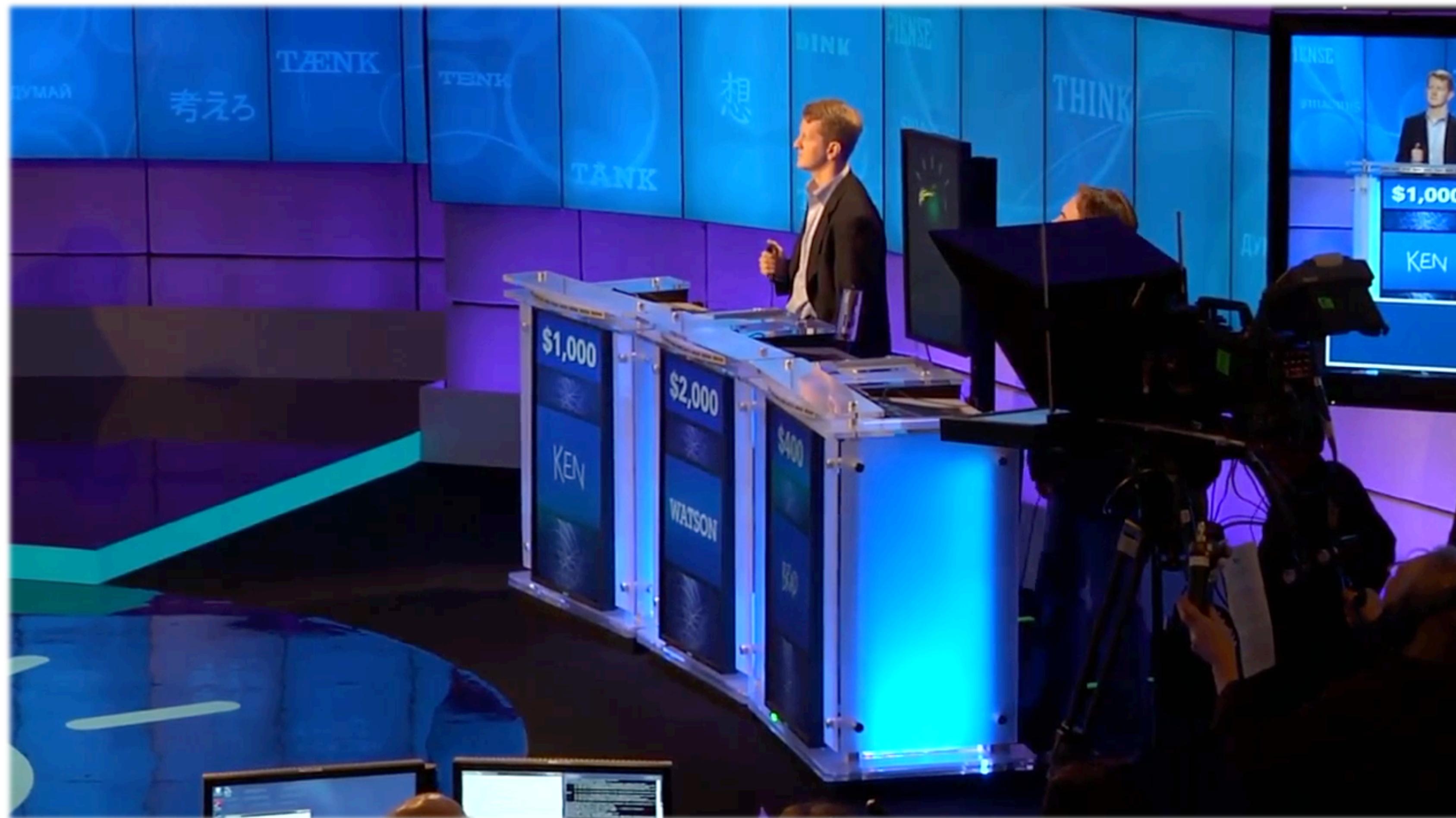
https://youtu.be/FwFduRA_L6Q

DeepBlue vs Kasparov (1997)



<https://youtu.be/jpCa9D2KPXM>

Watson à Jeopardy (IBM, 2011)



https://youtu.be/WFR3lOm_xhE

Knowledge Graph (Google, 2012)

accous - Recherche Google

https://www.google.fr/search?source=hp&ei=mxHDXPmzEZGdlwTFtbv4CQ&q=accous&btnK=Recherche+Google&oq=acc...

Connexion

Tous Maps Actualités Images Vidéos Plus Paramètres Outils

Environ 475 000 résultats (0,41 secondes)

mairie accous
<https://www.accous.fr/> ▾
Mairie d'Accous, Vallée d'Aspe, Actualités, vie communale, sport, activités nature, loisirs parapente, ski, ludopia, acrobranche.

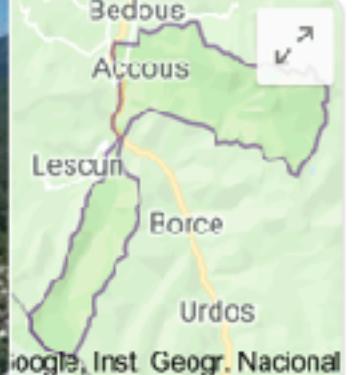
Accous — Wikipédia
<https://fr.wikipedia.org/wiki/Accous> ▾
Accous (en béarnais Aoós ou Accous) est une commune française du Béarn, située dans le département des Pyrénées-Atlantiques en région Nouvelle-Aquitaine ...
Géographie · Toponymie · Politique et administration · Culture et patrimoine

Principales activités à découvrir à Accous

 Col du Somport Ski, randonnée pédestre et station de	 Espace Ludopia - Parc de loisirs Parc et parc d'attractions	 Gave d'Ossau Montagne, nature, camping et parc	 Lac du Montagnon Lac et montagne
--	---	--	--

Guide de voyage pour Accous

Accous - Tourisme, Vacances & Week-end - France Voyage
<https://www.france-voyage.com/villes-villages/accous-25128.htm> ▾
Bâti au débouché du cirque d'Iseye, Accous, chef lieu du canton de la vallée d'Aspe (13 communes) se visite à pied... ou en parapente. Ici la montagne...
Altitude: De 420m à 2458m Code Insee: 64006
Code postal: 64490

ACCOUS
Commune en France

Accous est une commune française du Béarn, située dans le département des Pyrénées-Atlantiques en région Nouvelle-Aquitaine. [Wikipédia](#)

Superficie : 60,68 km²

Population : 464 (2015) INSEE

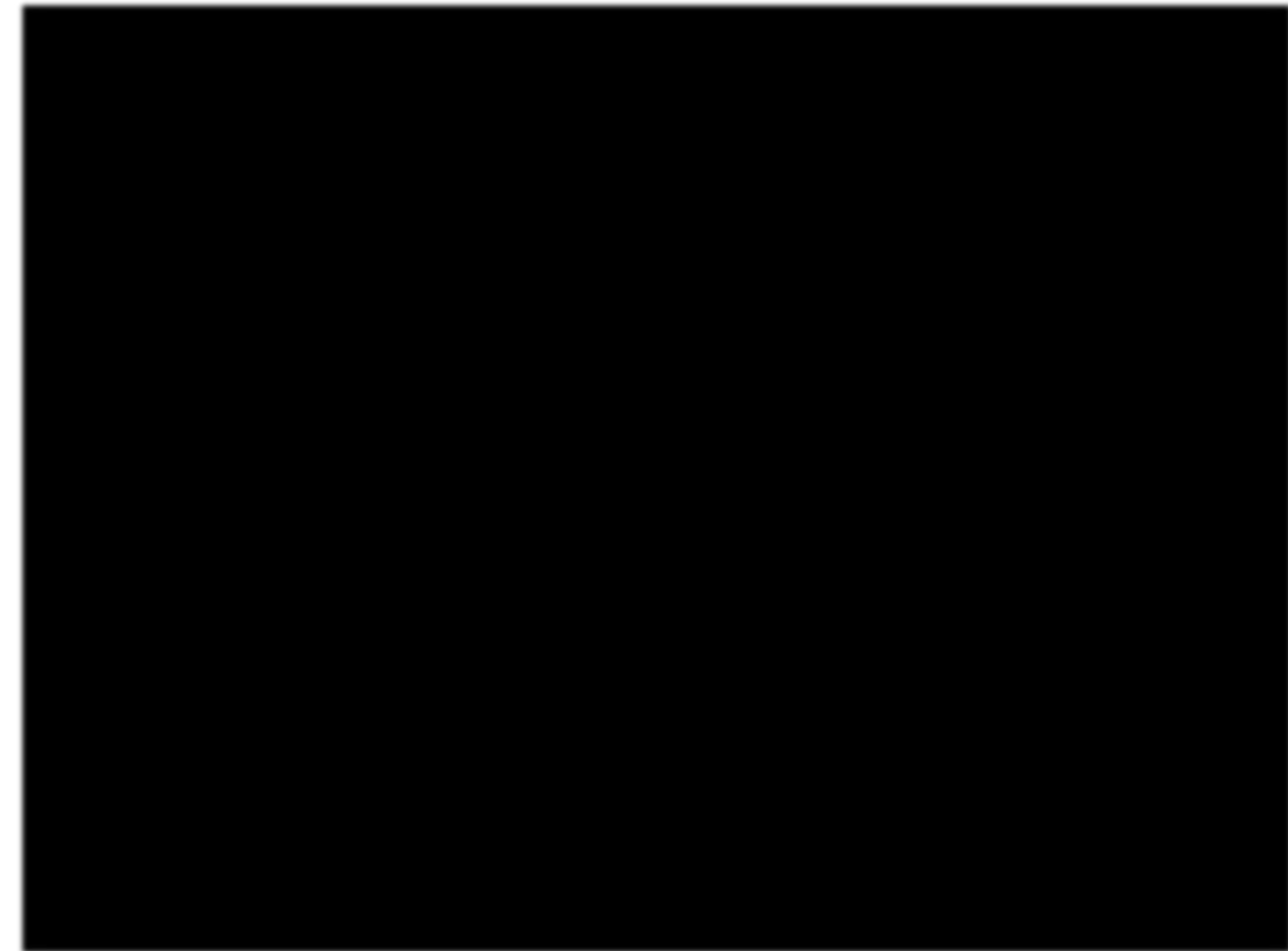
Heure locale : vendredi 16:11

Événements à venir

dim. 28 avr. 18:00	Concert d'Jazz ... Espace Jéllote
ven. 26 avr. 17:30	Rendez-vous d... L'ESCAPADE
sam. 27 avr. 09:30	Portes Ouvertes Cretofit Sport & ...

Voir d'autres éléments (plus de 15)

DQN Breakout (Deepmind, 2013)



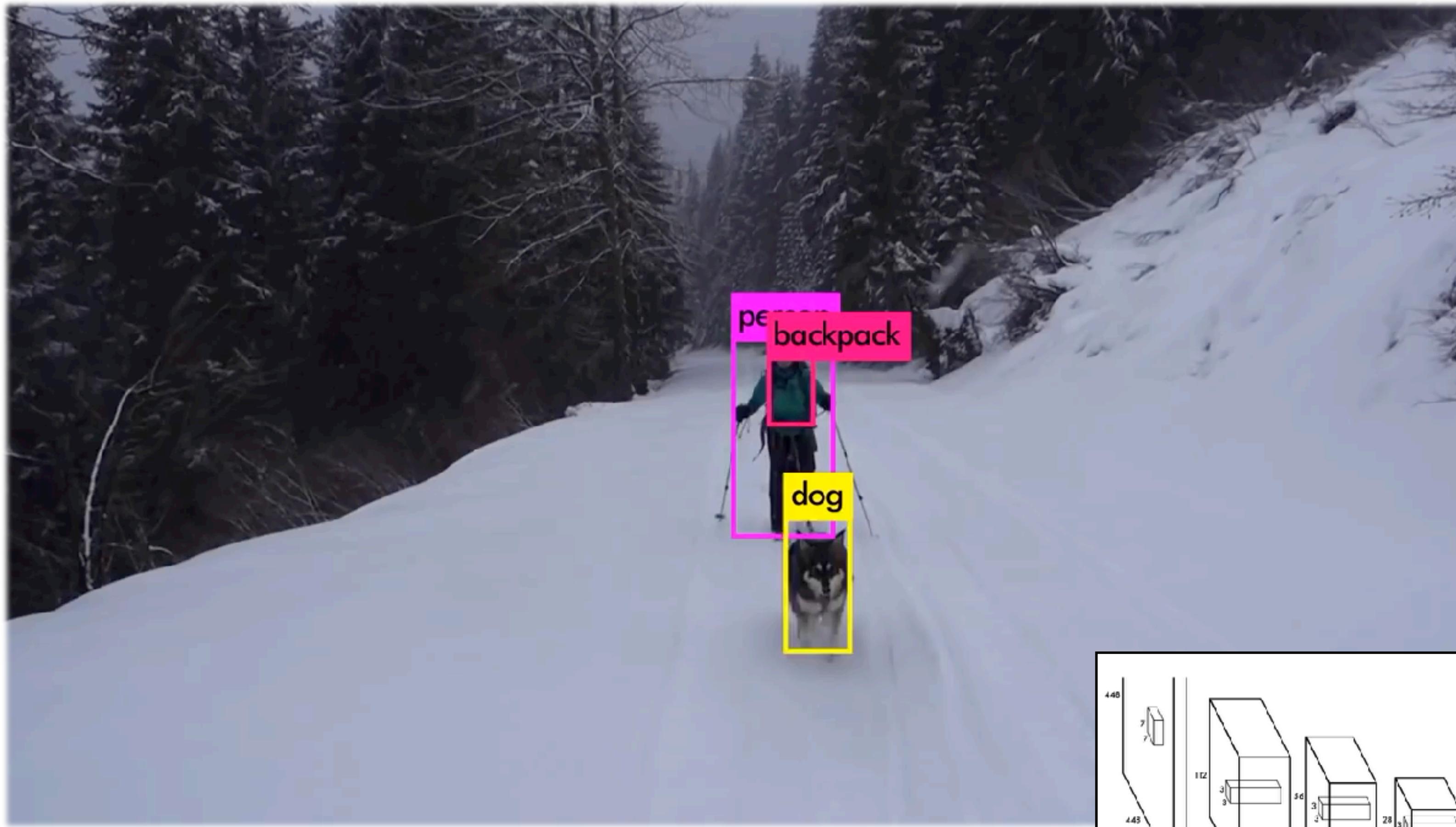
<https://youtu.be/TmPfTpjtdgg>

Skype Translator (Microsoft, 2015)

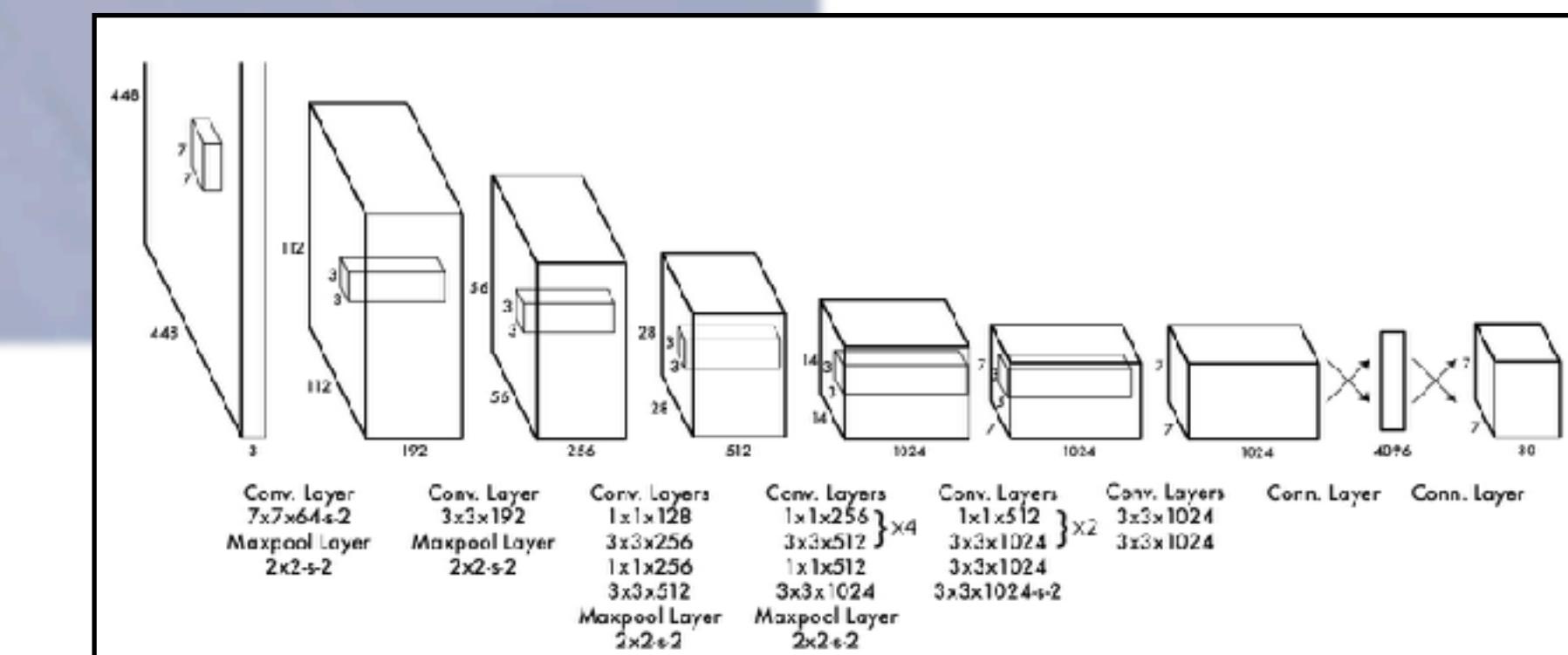


https://youtu.be/_VJ_C521an8

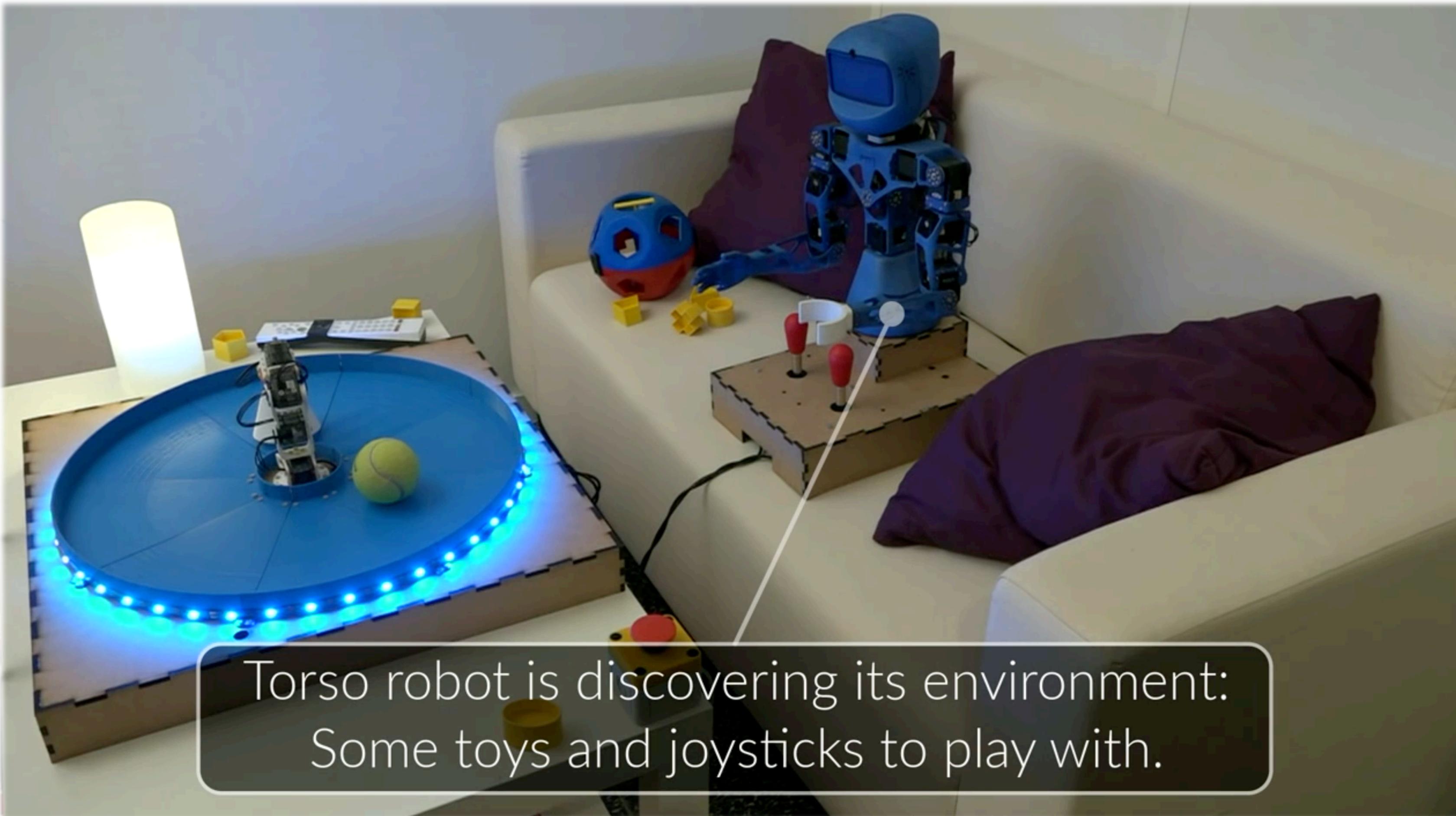
You Only Look Once (Joseph Redmon et al., 2016)



<https://youtu.be/MPU2HistivI>



Open-ended robot learning (Forestier et al., 2016)



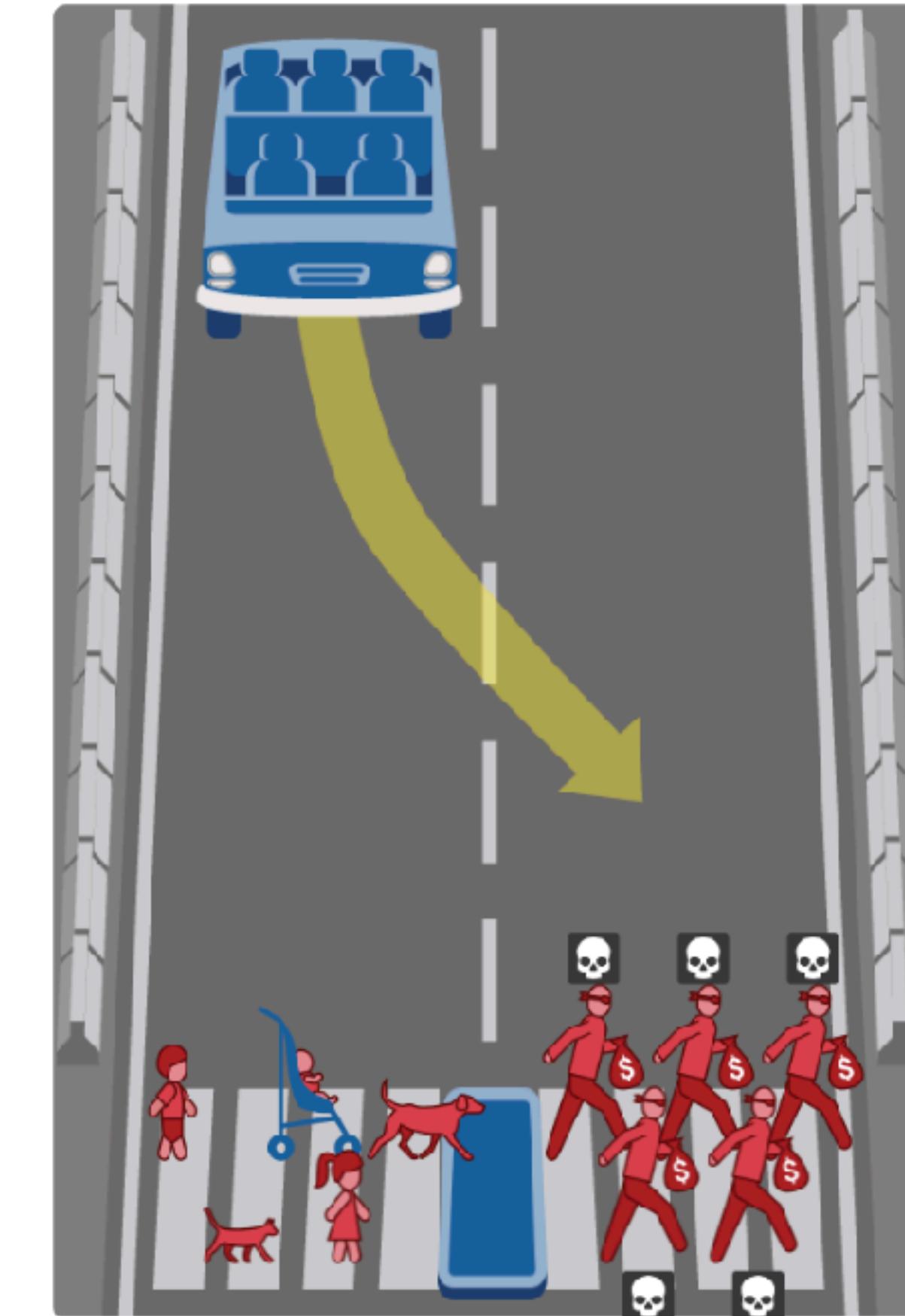
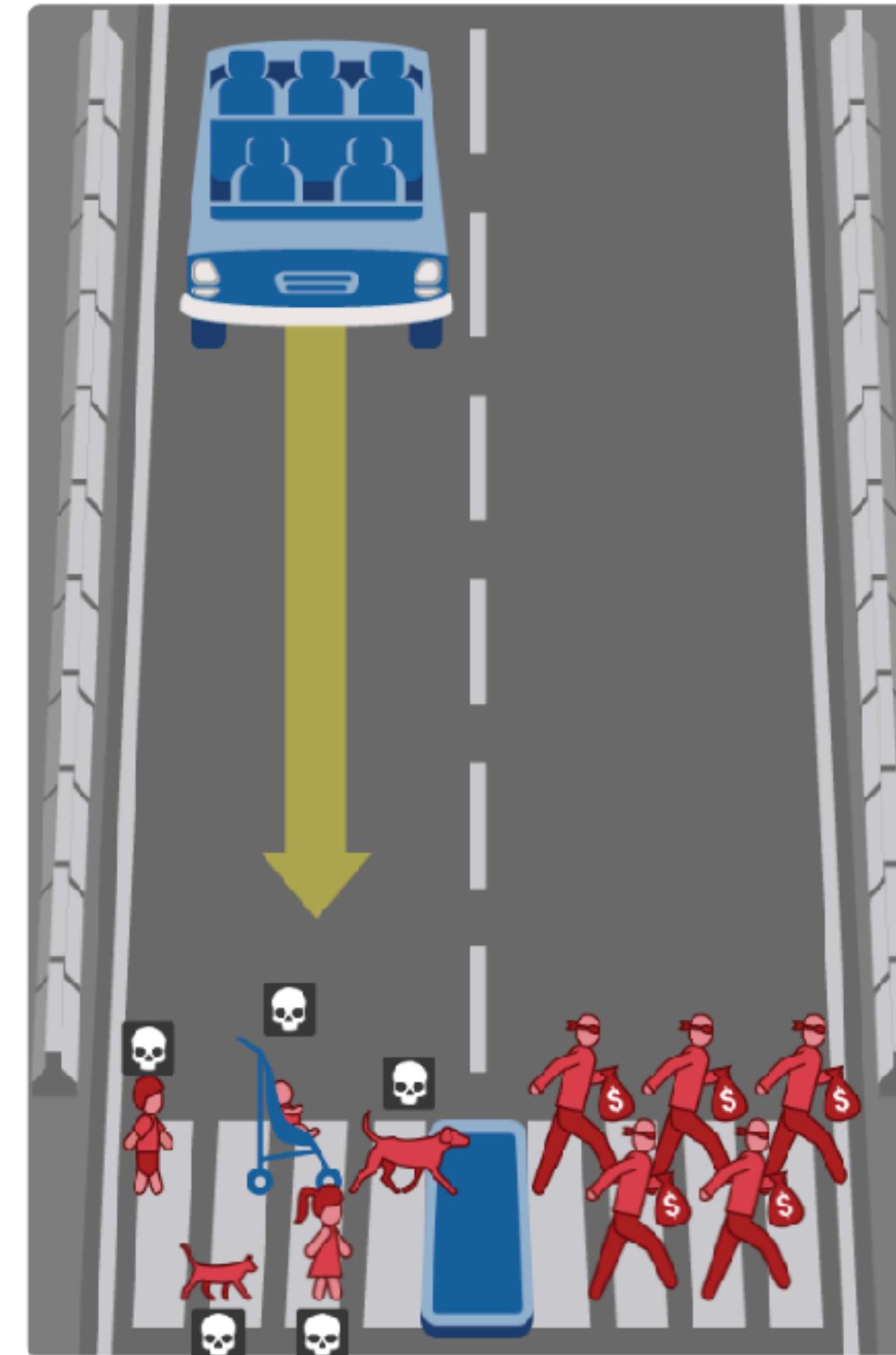
<https://youtu.be/NOLAwD4ZTWO>

Pepper learning bilboquet (Vollmer & Hemion, 2016)



https://youtu.be/jkaRO8J_1XI

Moral machine (MIT, 2016)



<http://moralmachine.mit.edu/>

CaptionBot (Microsoft, 2016)

Microsoft

English (US)

CaptionBot



I think it's a herd of sheep standing on top of a grass covered field.



How did I do?

★★★★★

Microsoft

English (US)

CaptionBot



I think it's a person sitting on top of a mountain.



How did I do?

★★★★★

Microsoft

English (US)

CaptionBot



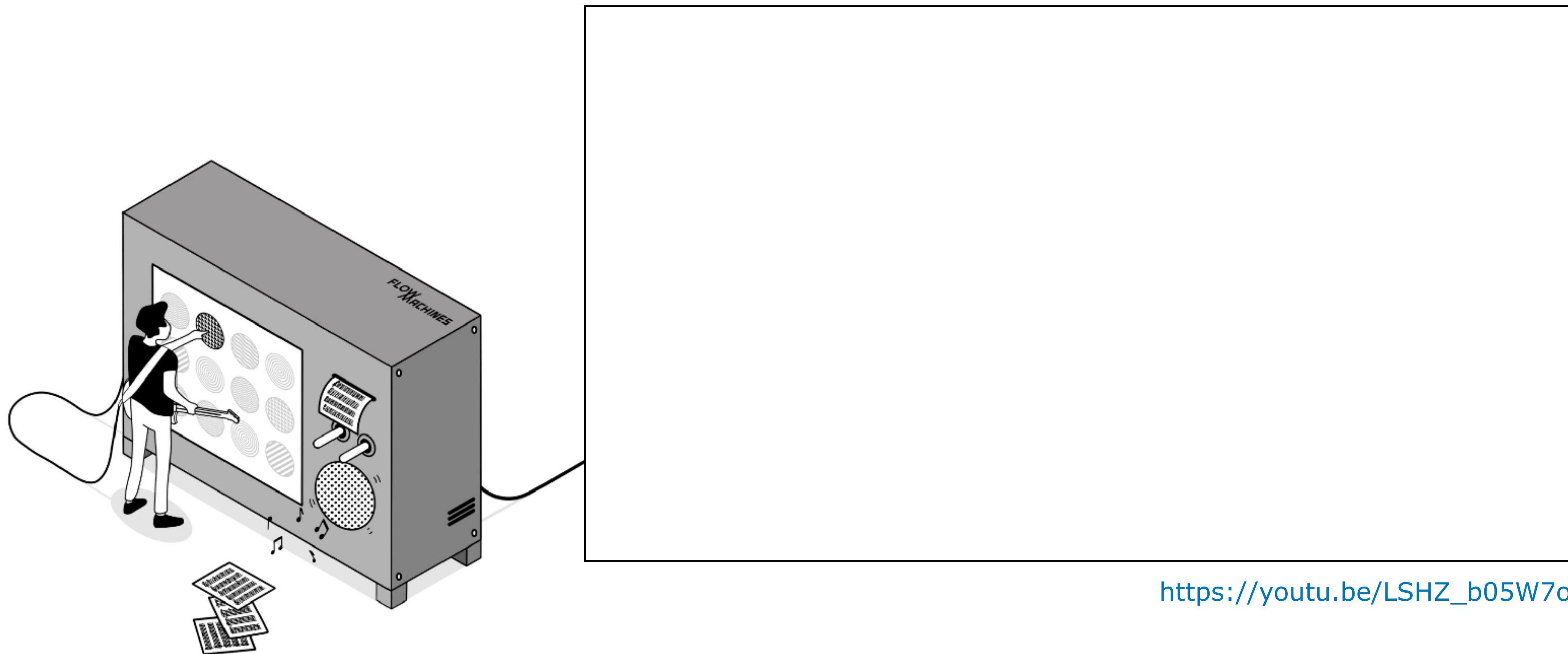
I think it's a view of a large mountain in the background.



How did I do?

★★★★★

Daddy's car (François Pachet & Benoit Carré, 2017)



https://youtu.be/LSHZ_b05W7o

Slaughterbots (Stop Autonomous Weapons, 2017)



<https://youtu.be/9CO6M2HsoIA>

Ke Jie à propos d'AlphaGo (2017)



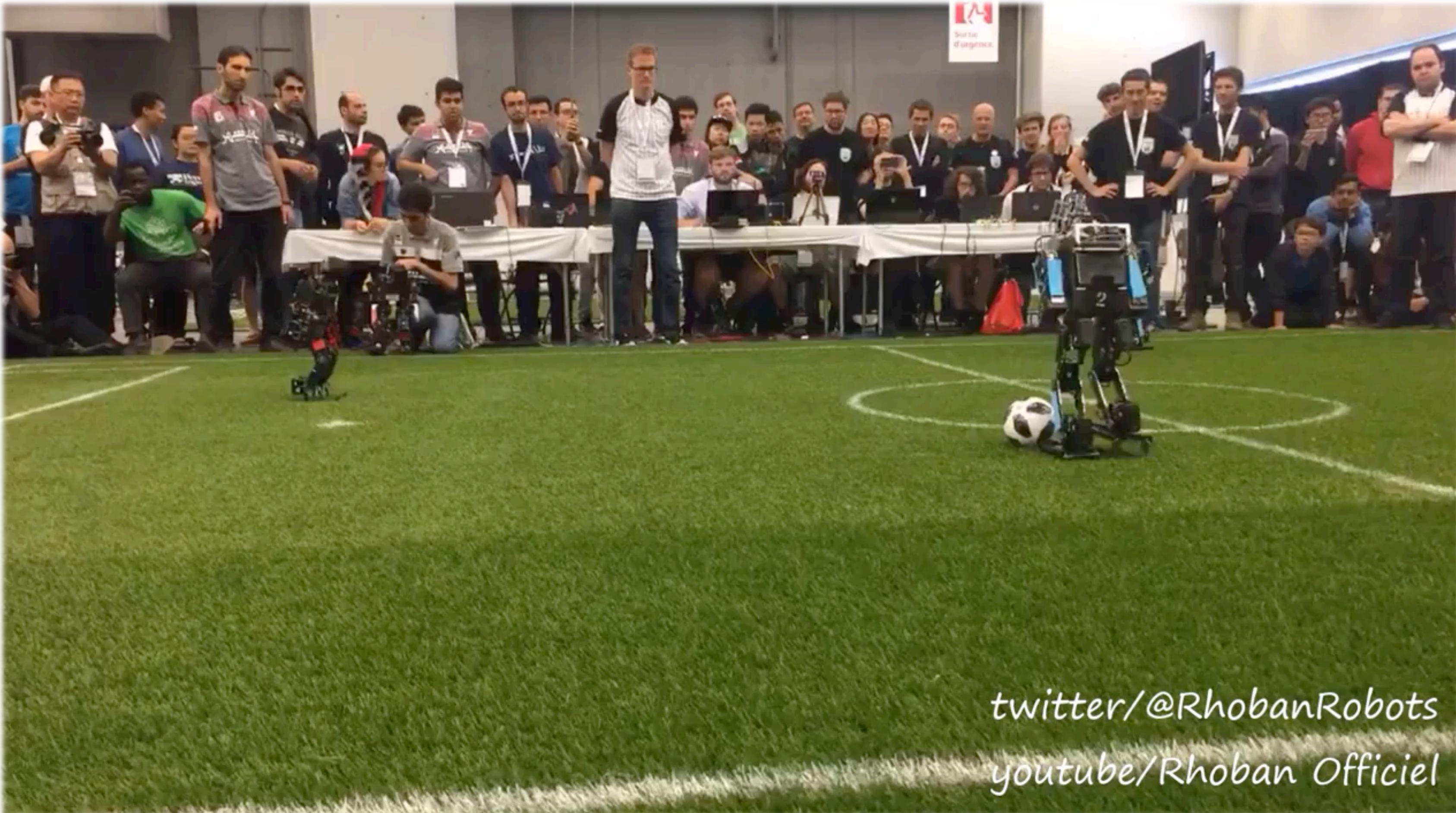
<https://youtu.be/UHU8ICctC-Q>

Libratus (2017)



<https://youtu.be/Jgau2BKTHbk>

Finale de la RoboCup 2018 : MRL vs. Rhoban



twitter/@RhobanRobots
youtube/Rhoban Officiel

<https://youtu.be/tmr3rVOS15Y>

Rhoban à la RoboCup 2019

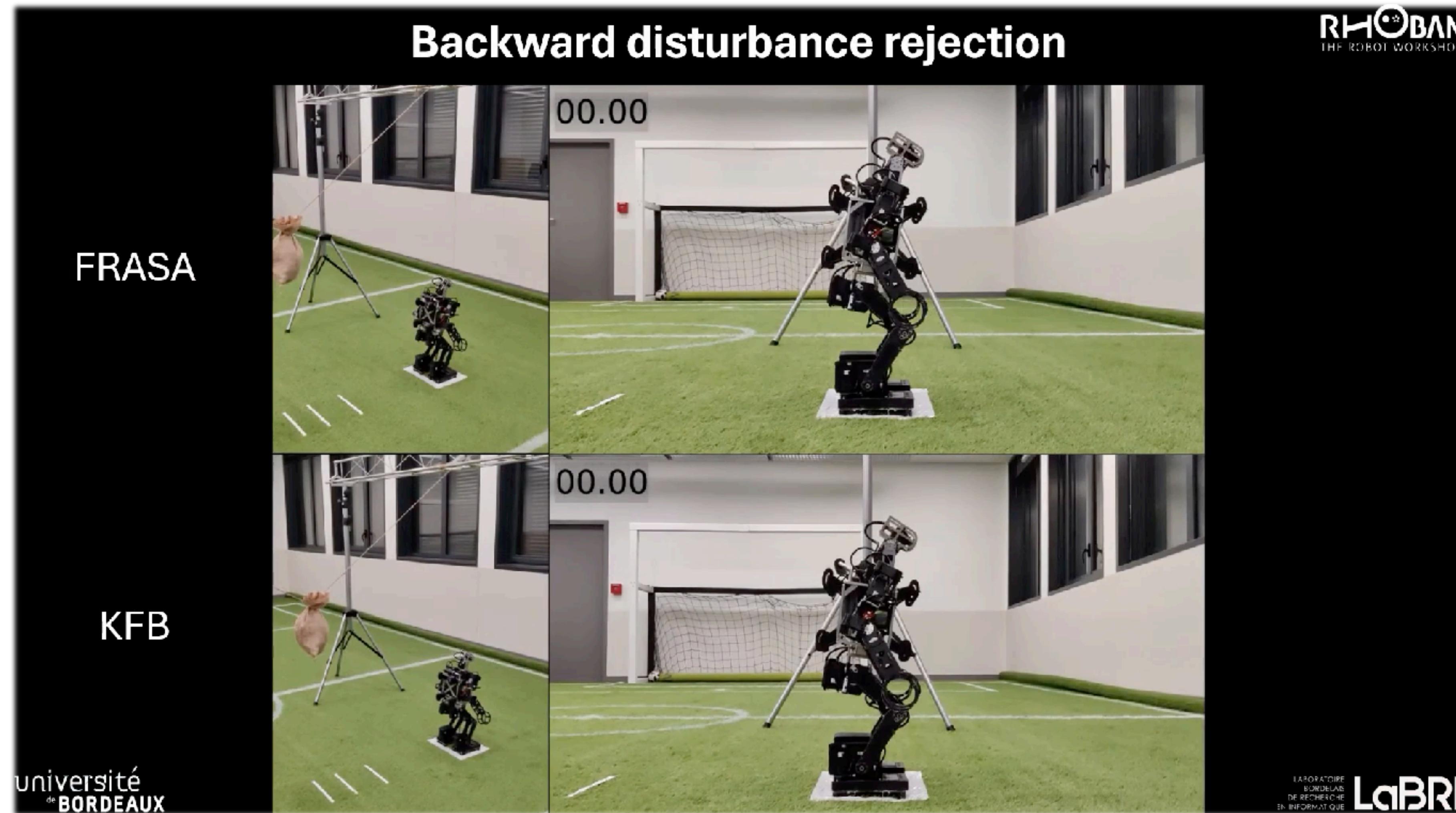


<https://twitter.com/RhobanRobots/status/1146262032796323841>



<https://twitter.com/LauRn964/status/1147733578211479553>

Rhoban en 2025



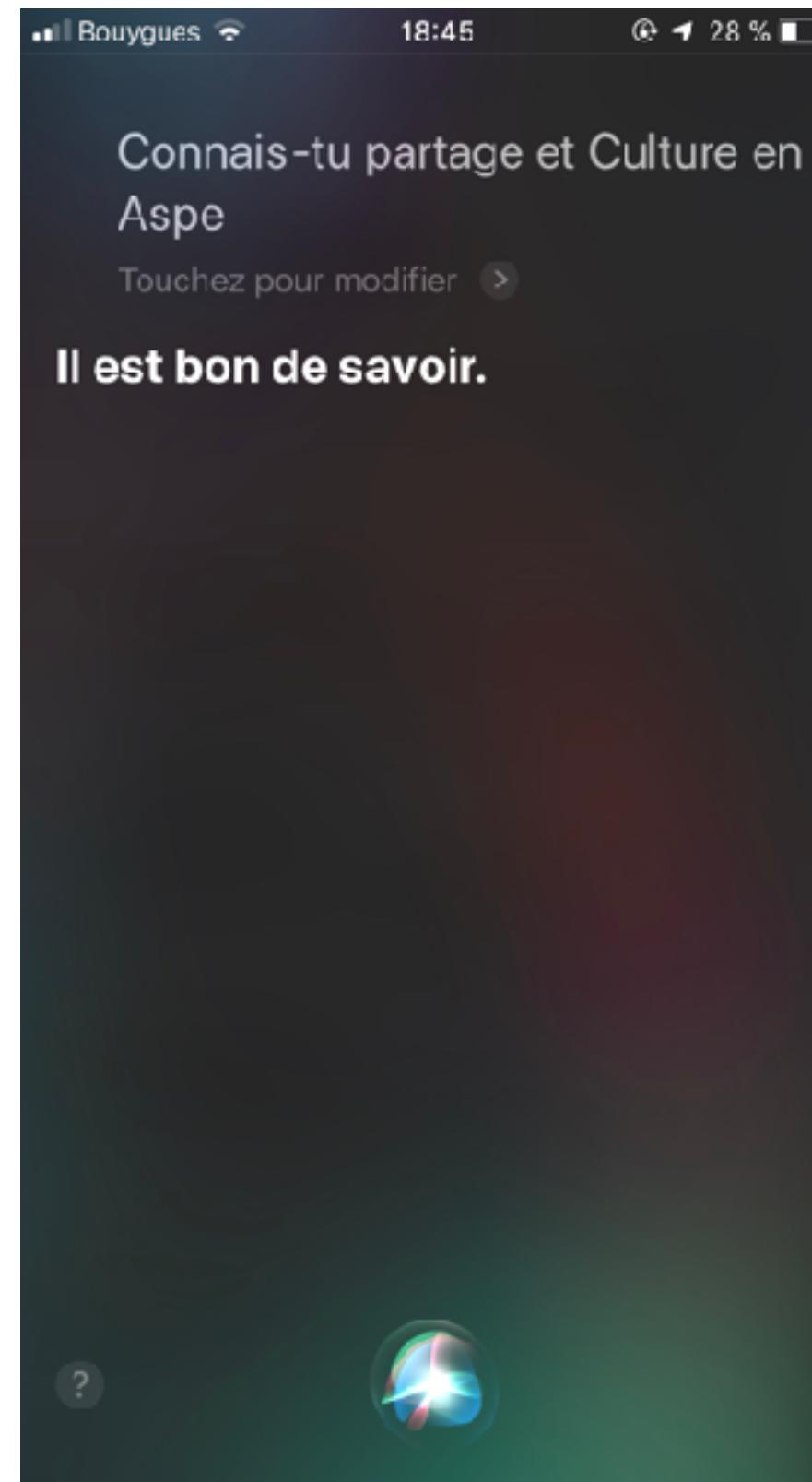
KFB = méthode *key frame based* utilisée lors de la 5e victoire à la RoboCup en 2023

AlphaStar vs "MaNa" (2019)



<https://youtu.be/zgIFoepzhIo>

Assistants vocaux : peuvent mieux faire ?



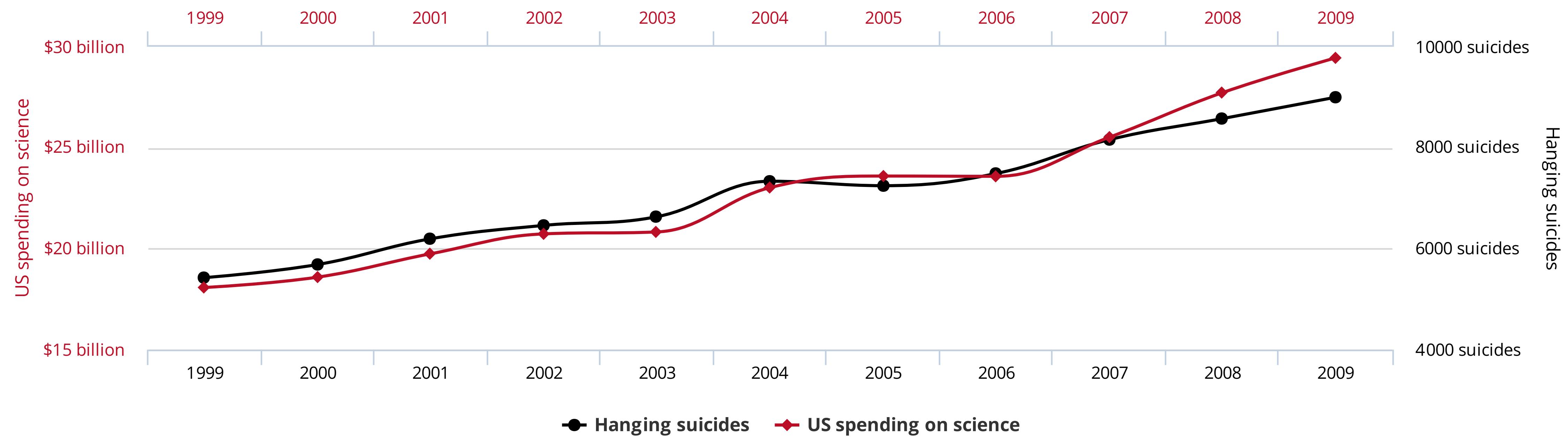
Amazon's Alexa started ordering people dollhouses after hearing its name on TV (2017)

Google's Super Bowl ad accidentally set off a lot of Google Homes (2017)

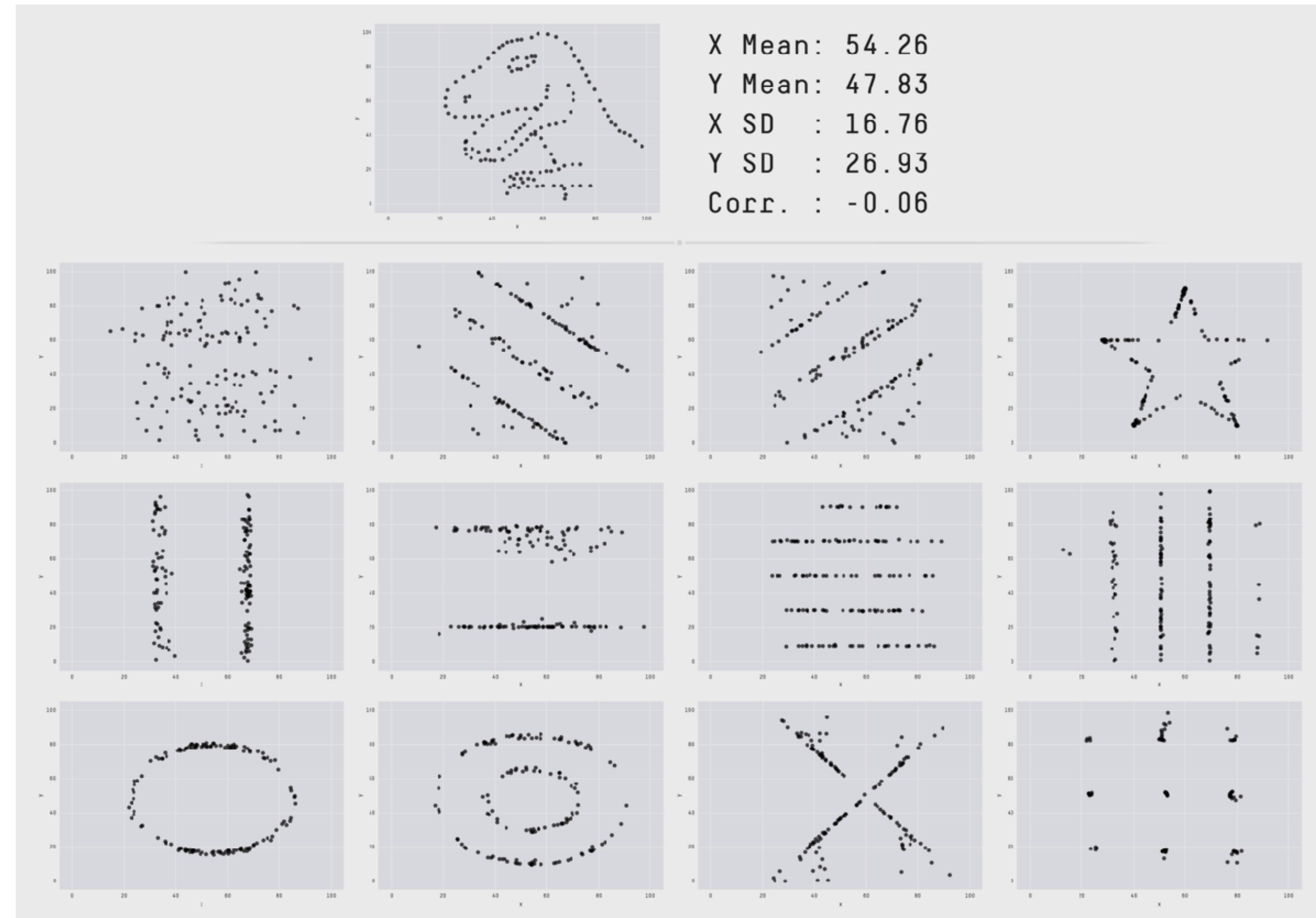
Amazon : les possesseurs d'Alexa sont-ils sur écoute ? (2019)

Corrélation n'est pas causalité

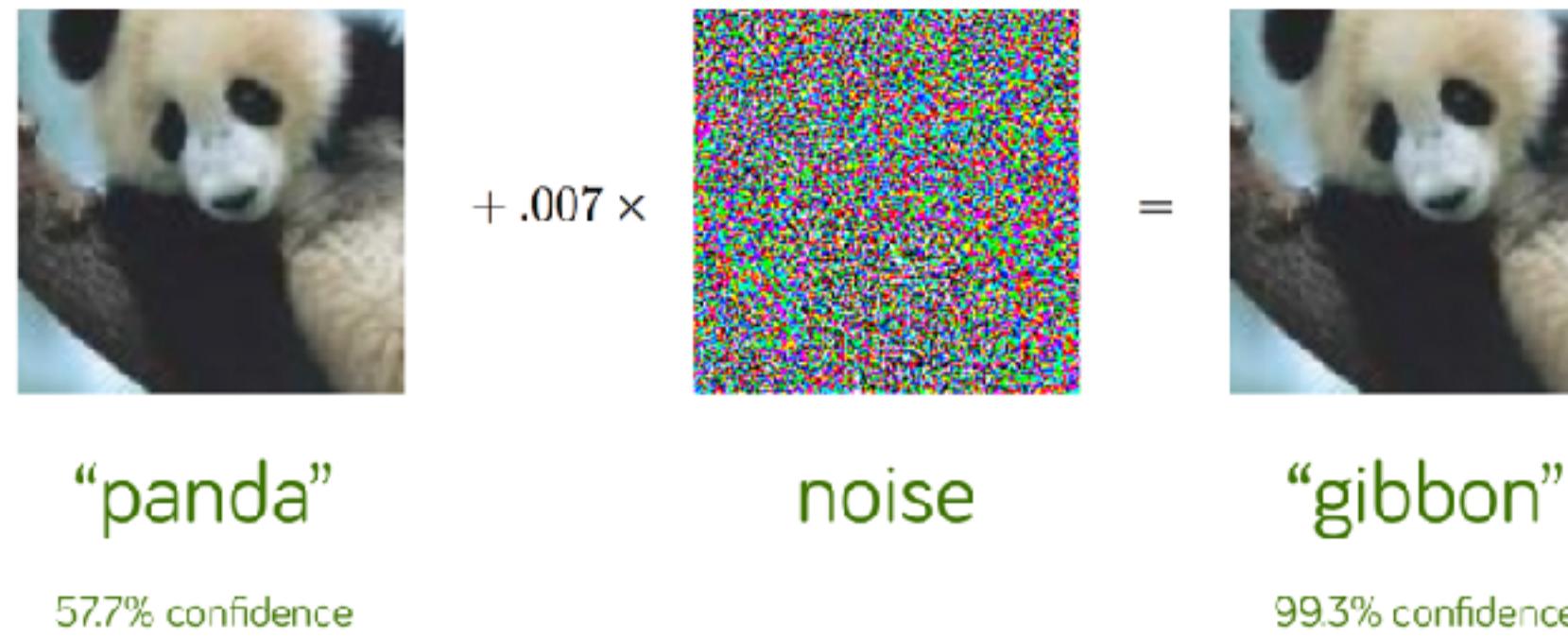
US spending on science, space, and technology
correlates with
Suicides by hanging, strangulation and suffocation



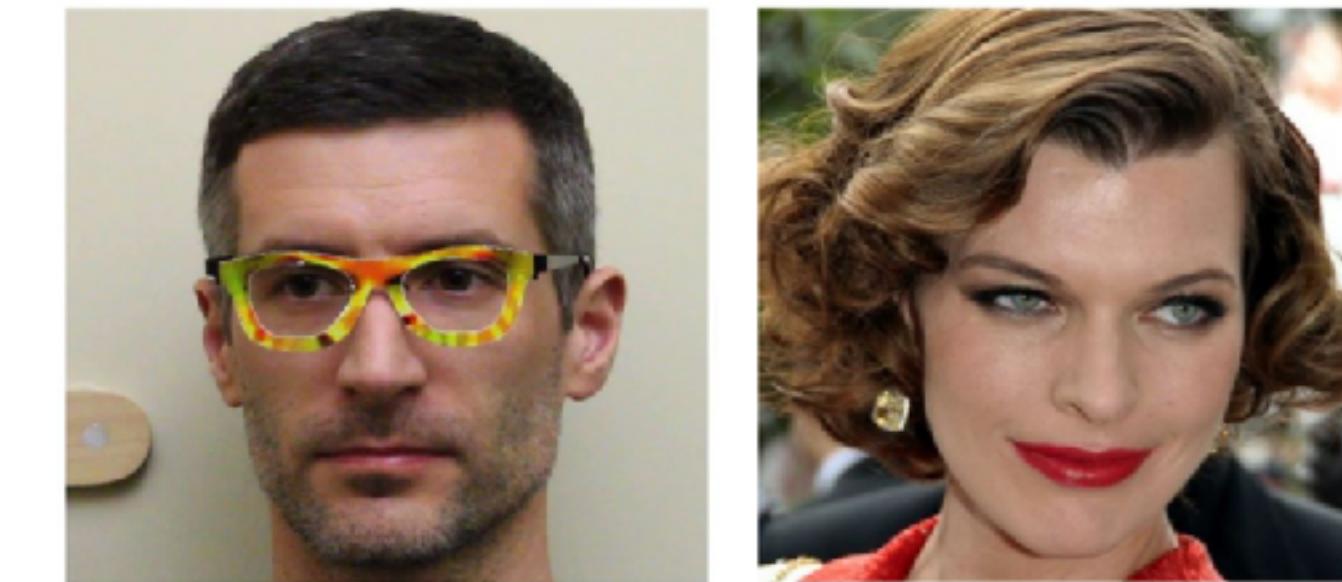
Same stats, different graphs (Matejka & Fitzmaurice, 2017)



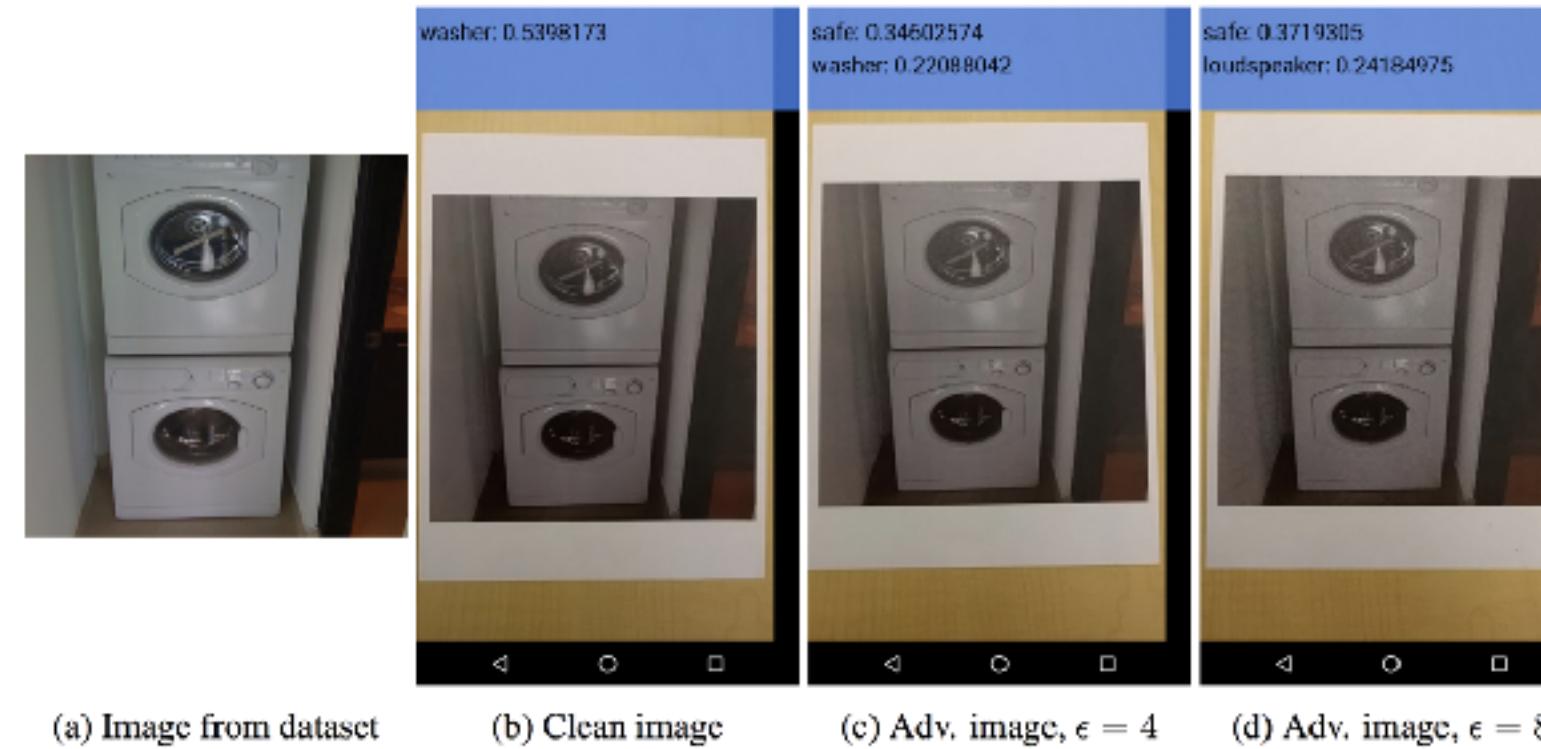
Adversarial examples



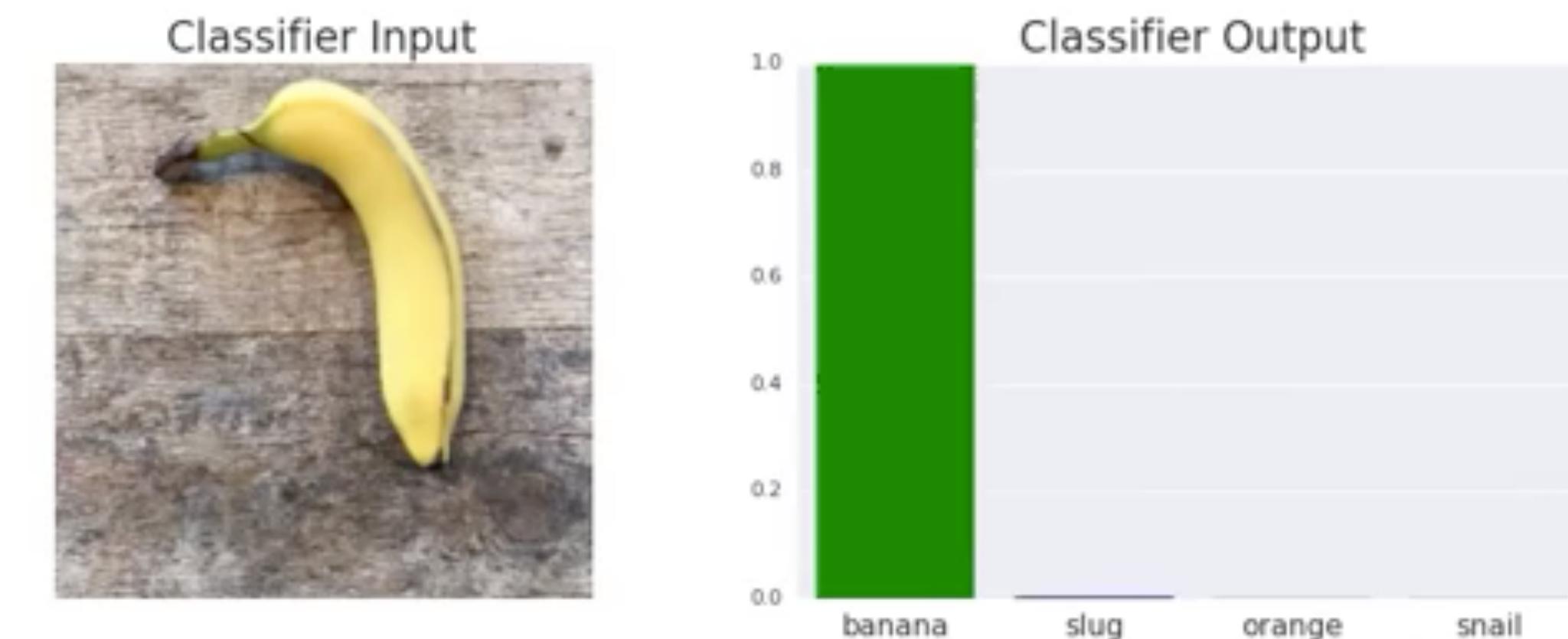
<https://arxiv.org/abs/1412.6572>



<https://www.cs.cmu.edu/~sbhagava/papers/face-rec-ccs16.pdf>



<https://arxiv.org/pdf/1607.02533.pdf>



<https://arxiv.org/pdf/1712.09665.pdf>

Time lapse of a baby playing with his toys



<https://youtu.be/8vNxjwt2AqY>

The men who stare at goats (Heslov, 2009)



So, they've started doing psi research because they thought we were doing psi research, when in fact we weren't doing psi research?

Yes, sir. But now that they're doing psi research, we're gonna have to do psi research, sir.

“L'informatique doit être au service de chaque citoyen”

"L'informatique doit être au service de chaque citoyen. [...] Elle ne doit porter atteinte ni à l'identité humaine, ni aux droits de l'homme, ni à la vie privée, ni aux libertés individuelles ou publiques. Toute personne dispose du droit de décider et de contrôler les usages qui sont faits des données à caractère personnel la concernant, dans les conditions fixées par la présente loi."

[Loi 78-17 du 6 janvier 1978](#)

Un taxi pour Tobrouk (de La Patellière, 1961)



"Je crois, docteur, que l'homme de Neandertal est en train de nous le mettre dans l'os. Deux intellectuels assis vont moins loin qu'une brute qui marche."

Intelligence artificielle et justice

Ce logiciel de justice prédictive ne nous apportait aucune plus-value

Après quatre mois de test [...], le Premier président de la cour d'appel de Rennes a mis fin à l'expérimentation, considérant que les résultats livrés étaient trop aléatoires. "C'est une opération complexe que la motivation intellectuelle d'un juge, une opération subtile qui s'articule à un raisonnement et non à une simple corrélation d'items factuels", explique Xavier Ronsin.

En Estonie, une intelligence artificielle va bientôt rendre la justice

Une intelligence artificielle développée par le gouvernement estonien pourra rendre des jugements autonomes dans des délits mineurs afin d'alléger la charge de travail des juges. Le recours aux algorithmes dans la sphère judiciaire soulève des questions éthiques.

Justice : faut-il faire confiance au verdict des algorithmes ?

L'expérience montre que les algorithmes sont source d'erreurs et de discriminations. Une première charte éthique énonce les règles du jeu.

Véhicules autonomes



https://youtu.be/4aGXEJ_9G3Y

5 Reasons Experts Think Autonomous Cars Are Many Years Away

In the world of autonomous vehicles, Pittsburgh, Phoenix and Silicon Valley are bustling hubs of development and testing. But ask those involved in self-driving vehicles when we might actually see them carrying passengers in every city, and you'll get an almost universal answer: Not anytime soon.

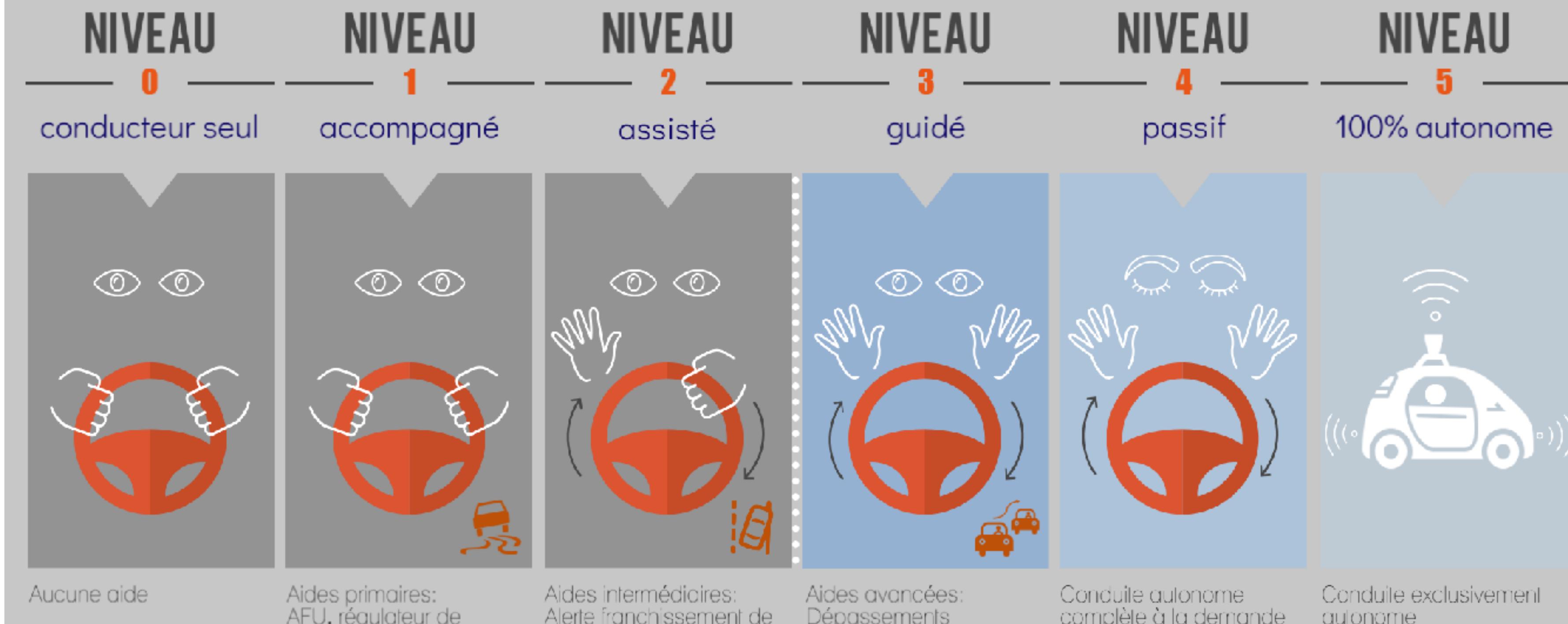
PSA revoit sa stratégie pour la voiture autonome

Le groupe PSA annonce ne pas vouloir développer de technologie de conduite 100 % autonome pour ses véhicules particuliers. Jugée trop chère pour le constructeur comme les clients, elle n'aura pas d'intérêt et risquerait de faire trop augmenter les tarifs.

Niveaux de conduite autonome

Les niveaux de conduite autonome

Groupe
Argus



Source: SAE International levels of automation

Quel impact sur l'emploi ?

Les cinq métiers les plus menacés par l'intelligence artificielle

Selon le think tank Institut Sapiens, cinq métiers semblent plus menacés que les autres par l'intelligence artificielle et l'automatisation. Ils pourraient disparaître dans le courant du XXI^e siècle.

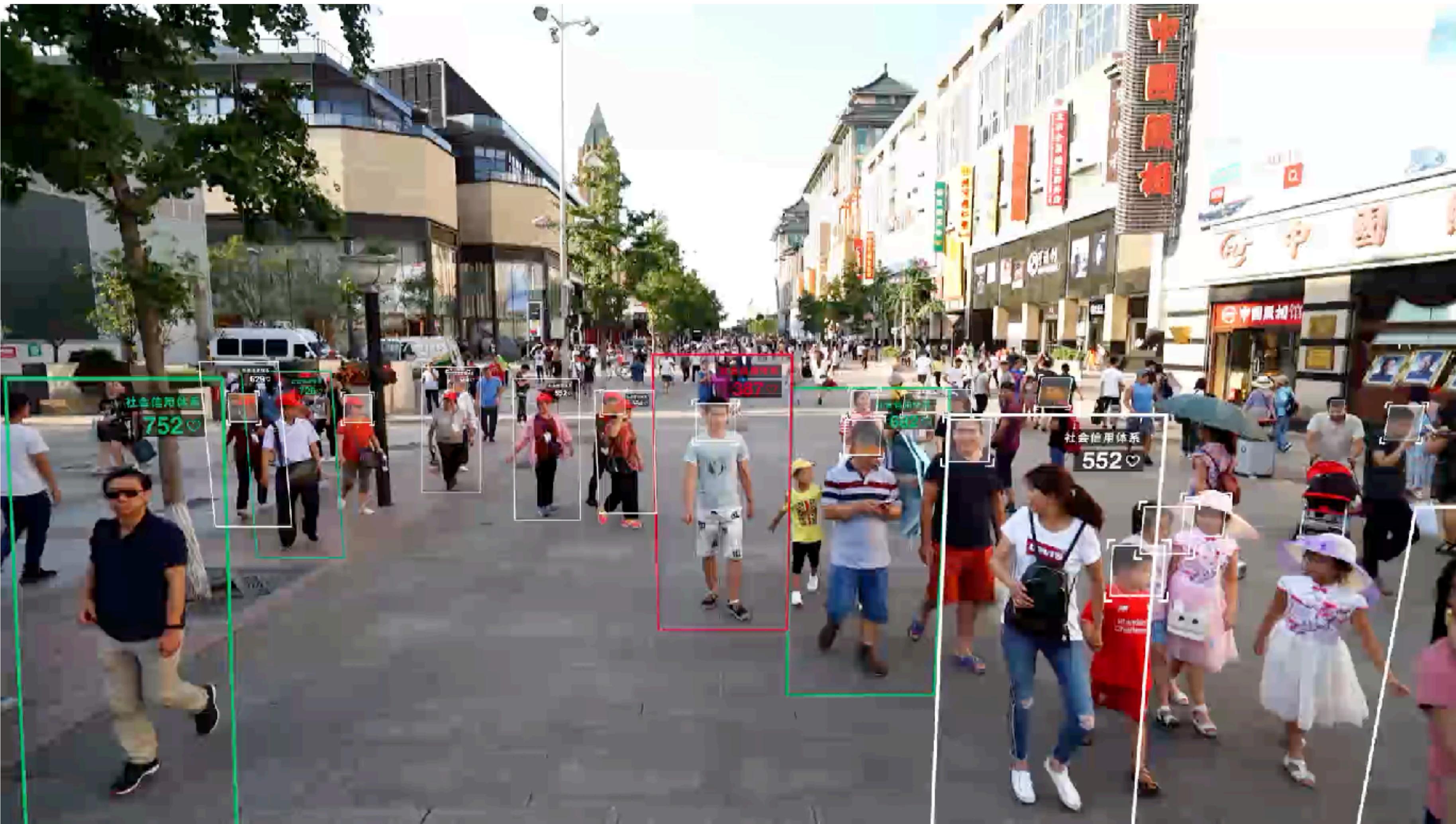
L'intelligence artificielle n'existe pas, mais elle va quand même prendre votre job

Pourquoi avoir peur de l'IA ? Cette machine qui serait semblable à l'humain n'est pas près d'arriver, si jamais elle est possible. Pour autant, l'IA, celle qui nous entoure déjà, n'est pas sans dangers, mais ce ne sont pas ceux auxquels on pense le plus souvent.

L'intelligence artificielle créera plus d'emplois qu'elle n'en détruira

Loin de s'apparenter à une forme d'automatisation destructrice de postes, l'IA va au contraire contribuer à créer de nouveaux métiers.

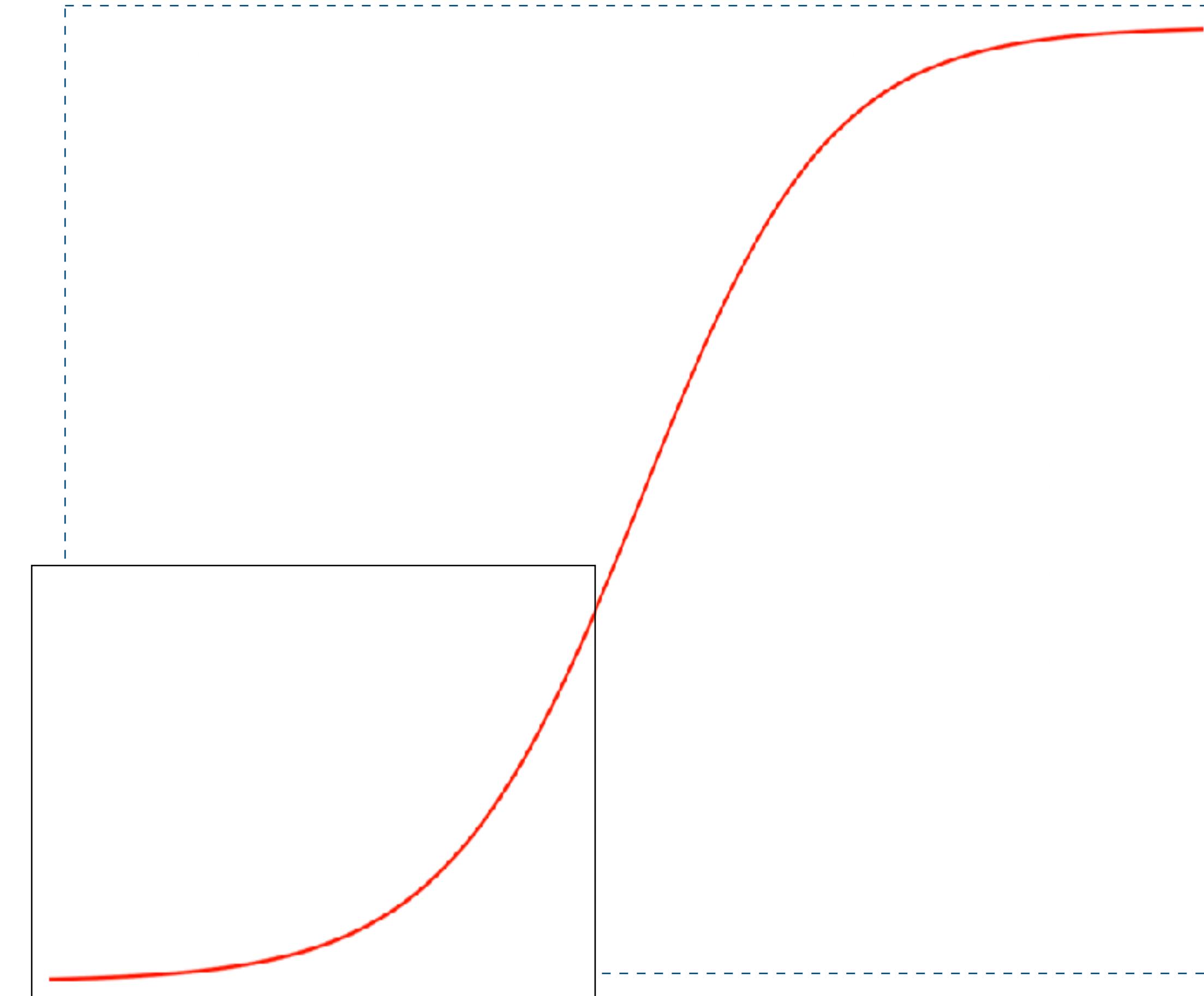
L'IA au service du "crédit social" ?



<https://www.abc.net.au/news/2018-09-18/china-social-credit-a-model-citizen-in-a-digital-dictatorship/10200278>

Croissance exponentielle... ou pas

Fonction
exponentielle



Fonction
logistique

Hypertrucages



Did the Camera Ever Tell the Truth?

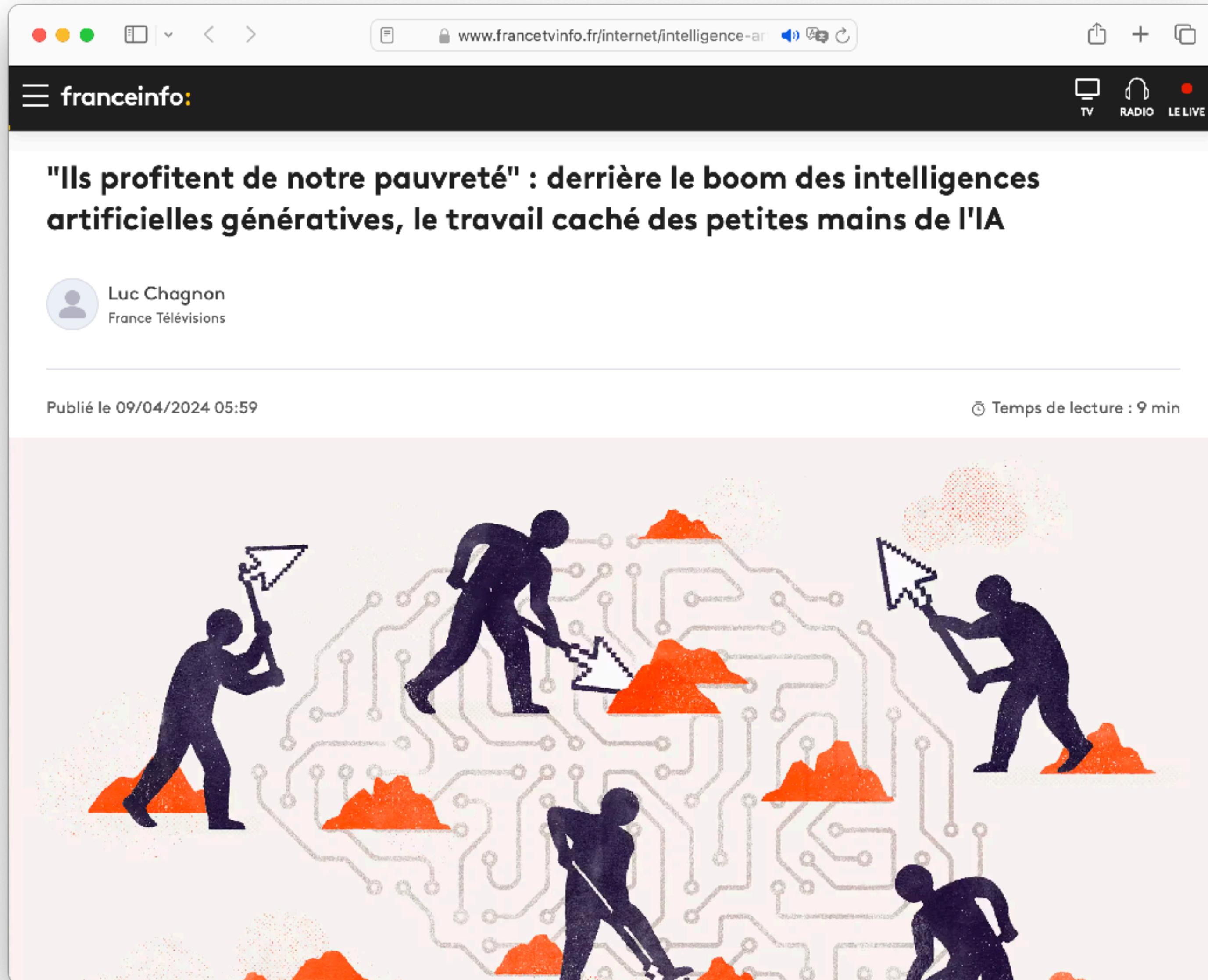


Chatbot Arena LLM Leaderboard

(27/02/2025)

Rank* (UB)	Rank (StyleCtrl)	Model	Arena Score	95% CI	Votes	Organization	License
1	1	chocolate (Early Grok-3)	1403	+6/-6	9992	xAI	Proprietary
2	3	Gemini-2.0-Flash-Thinking-Exp-01-21	1385	+4/-6	15083	Google	Proprietary
2	3	Gemini-2.0-Pro-Exp-02-05	1380	+5/-6	13000	Google	Proprietary
2	1	ChatGPT-4o-latest (2025-01-29)	1377	+5/-5	13470	OpenAI	Proprietary
5	3	DeepSeek-R1	1362	+7/-7	6581	DeepSeek	MIT
5	8	Gemini-2.0-Flash-001	1358	+7/-7	10862	Google	Proprietary
5	3	o1-2024-12-17	1352	+5/-5	17248	OpenAI	Proprietary
8	7	o1-preview	1335	+3/-4	33169	OpenAI	Proprietary
8	8	Qwen2.5-Max	1334	+5/-5	9282	Alibaba	Proprietary
8	7	o3-mini-high	1332	+5/-9	5954	OpenAI	Proprietary
11	11	DeepSeek-V3	1318	+4/-5	19461	DeepSeek	DeepSeek
11	13	Qwen-Plus-0125	1311	+9/-7	5112	Alibaba	Proprietary
11	14	GLM-4-Plus-0111	1310	+6/-9	5134	Zhipu	Proprietary
11	13	Gemini-2.0-Flash-Lite-Preview-02-05	1309	+6/-5	10262	Google	Proprietary
12	12	o3-mini	1306	+5/-6	12179	OpenAI	Proprietary
12	17	Step-2-16K-Exp	1304	+7/-7	5130	StepFun	Proprietary
12	17	1-1-1	1301	+6/-5	50000	StepFun	Proprietary

“Le travail caché des petites mains de l’IA” (avril 2024)



“Ces travailleurs sont souvent recrutés dans des pays à faibles revenus, et géopolitiquement instables”

“Il y a clairement une forme de néo-colonialisme”

"Si les annotateurs savaient qu'ils travaillent pour une entreprise qui génère des centaines de millions de dollars comme OpenAI, ils n'accepteraient pas des salaires si bas"

"Beaucoup de gens ne savent pas qu'il y a de l'humain derrière l'IA. Il faudrait que ce soit plus connu, et mieux payé"

“L'IA ne peut pas être éthique si elle est entraînée de façon immorale”

State of GPT (Karpathy, mai 2023)

Models may be biased

Models may fabricate (“hallucinate”) information

Models may have reasoning errors

Models may struggle in classes of applications, e.g. spelling related tasks

Models have knowledge cutoffs (e.g. September 2021)

Models are susceptible to prompt injection, “jailbreak” attacks, data poisoning attacks,...

Recommendations:

- Use in low-stakes applications, combine with human oversight
- Source of inspiration, suggestions
- Copilots over autonomous agents

Le mieux peut être l'ennemi du bien

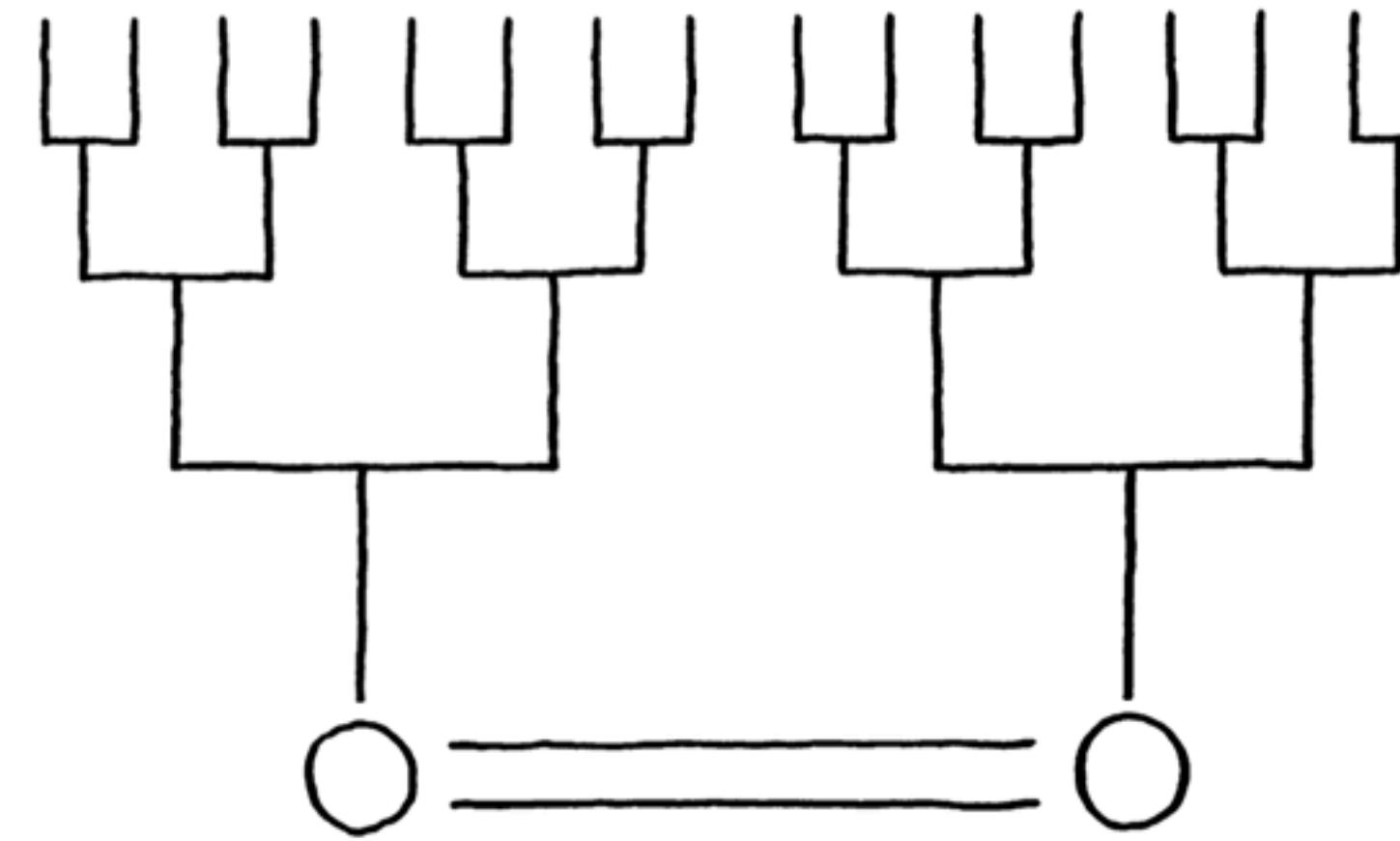
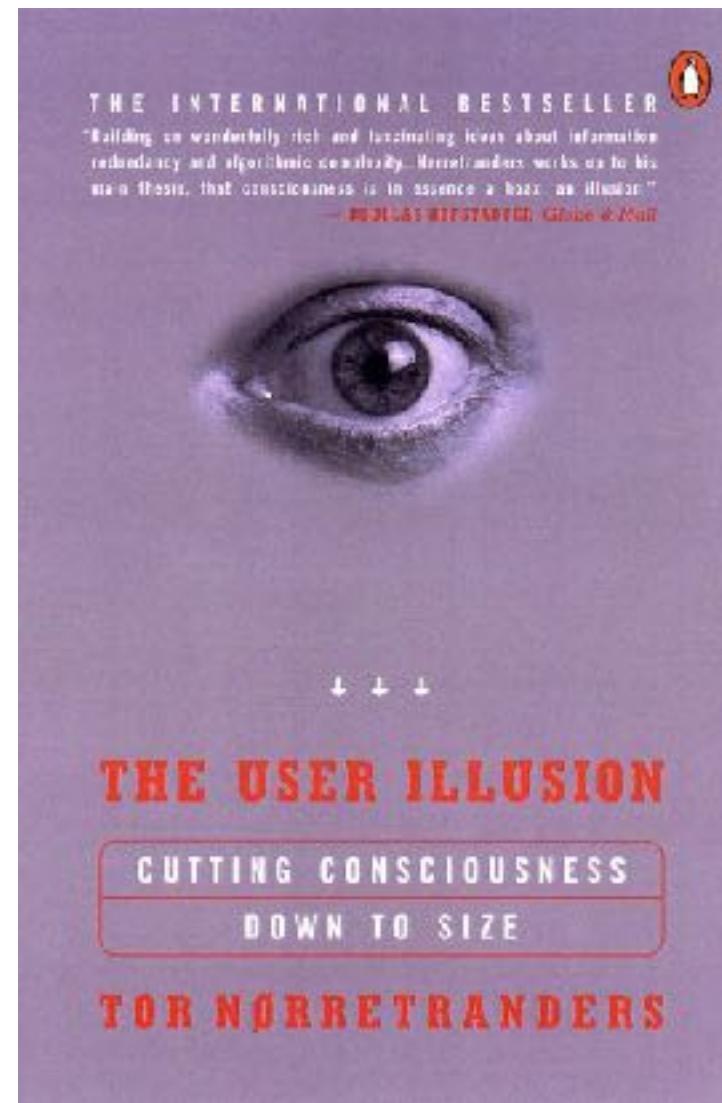


Gemini illustrations of 1943 German soldiers

toutes
et tous
égaux



Architecture encodeur-décodeur : digression



The tree of talking

The shortest correspondence in history took place in 1862.

*Victor Hugo [...] had gone on holiday following the publication of his great novel *Les Misérables*. But Hugo could not restrain himself from asking how the book was doing.*

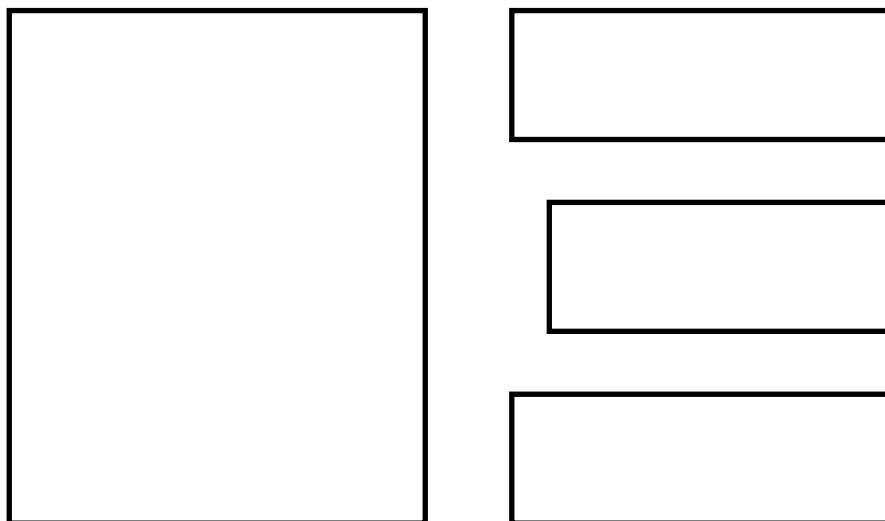
So he wrote the following letter to his publisher: "?"

His publisher was not to be outdone and replied fully in keeping with the truth: "!"

Architecture encodeur-décodeur



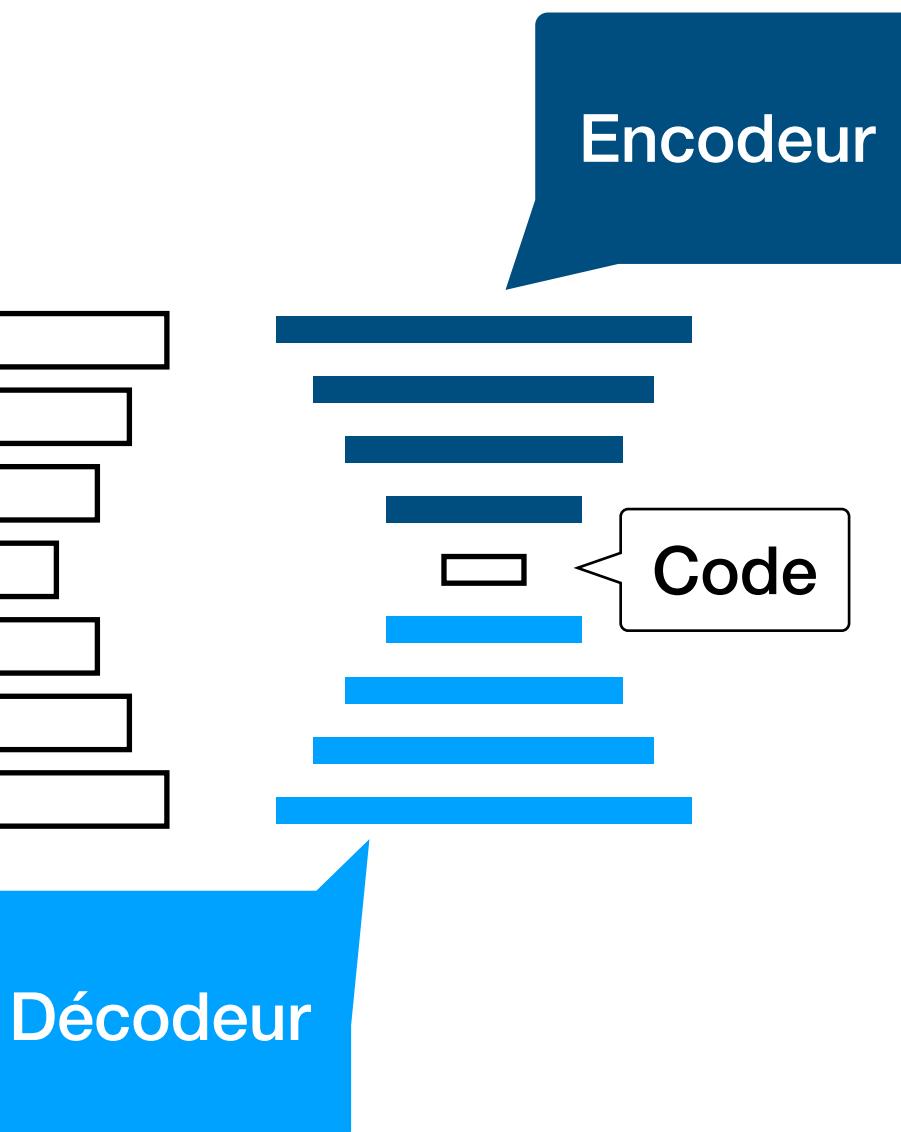
Entrée



Sortie



Merci à
Frédéric Alexandre !



0100....10100



Hypertrucages... ou non

Développement de compétences numériques ?

CHI '90 Proceedings April 1990

User-Tailorable Systems: Pressing the Issues with Buttons

Allan MacLean, Kathleen Carter, Lennart Lövstrand and Thomas Moran

Rank Xerox EuroPARC,
61 Regent Street, Cambridge CB2 1AB, England

ABSTRACT
It is impossible to design systems which are appropriate for all users and all situations. We believe that a useful technique is to have end users tailor their systems to match their personal work practices. This requires not only systems which can be tailored, but a culture within which users feel in control of the system and in which tailoring is the norm. In a two-pronged research project we have worked closely with a group of users to develop a system to support tailoring and to help the users evolve a "tailoring culture". This has resulted in a flexible system based around the use of distributed on-screen Buttons to support a range of tailoring techniques.

KEYWORDS: Tailorability; Modifiability; Customization; User Interface Design; Office Systems; Design Process.

TAILORABLE SYSTEMS
User-tailorable computer systems have been a goal in some form for a number of years, but achievements to date have not lived up to the promise. Progress has been made by producing higher level languages which have increased the productivity of programmers and like-minded users [21], and an increasing number of applications are available which have their own languages built in (e.g. Framework in the PC world [1]; Apple's HyperCard [12]). Alternatively, some systems provide their user with a range of predefined parameters which allow limited control. EMACS is probably one of the most successful examples of a tailorable system ("extensible and customizable", as its author describes it [23]). EMACS is tailorable both by programming and by setting parameters, and there are many people who would use no other editor. However, there are also many people who are inhibited from using EMACS, far less tailoring it, because of its complexity [22].

Figure 1. People with different levels of tailoring skills, and changes in skill required for increasing tailoring power using two common tailoring mechanisms. Steep slopes are barriers to skill acquisition.

The different terraces on this mountain lead to different cultures with people of different skill levels inhabiting them. At the risk of minor caricature we can characterise these cultures as follows:

- **Worker:** Lives on the plains. No interest in the computer system *per se*. Just wants to get work

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish requires a fee and/or specific permission.

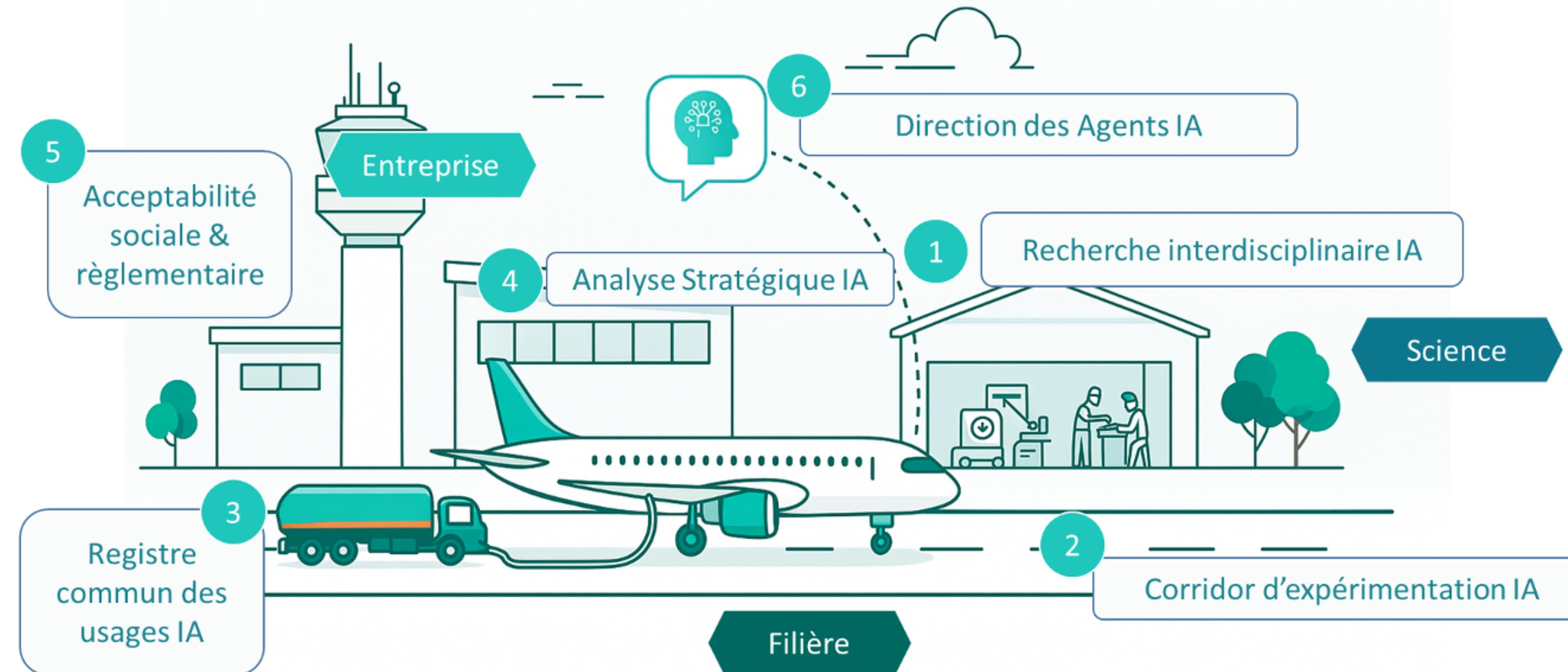
© 1990 ACM 0-89791-345-0/90/0004-0175 1.50

175

- Créer des infrastructures et plateformes
- Administrer des infrastructures et plateformes
- Créer des applications
- Créer des formules, des macros, des scripts, commencer à programmer
- Configurer, changer des paramètres, mieux comprendre le système
- Adapter ou créer des contenus
- Utiliser des systèmes (matériel et logiciel)

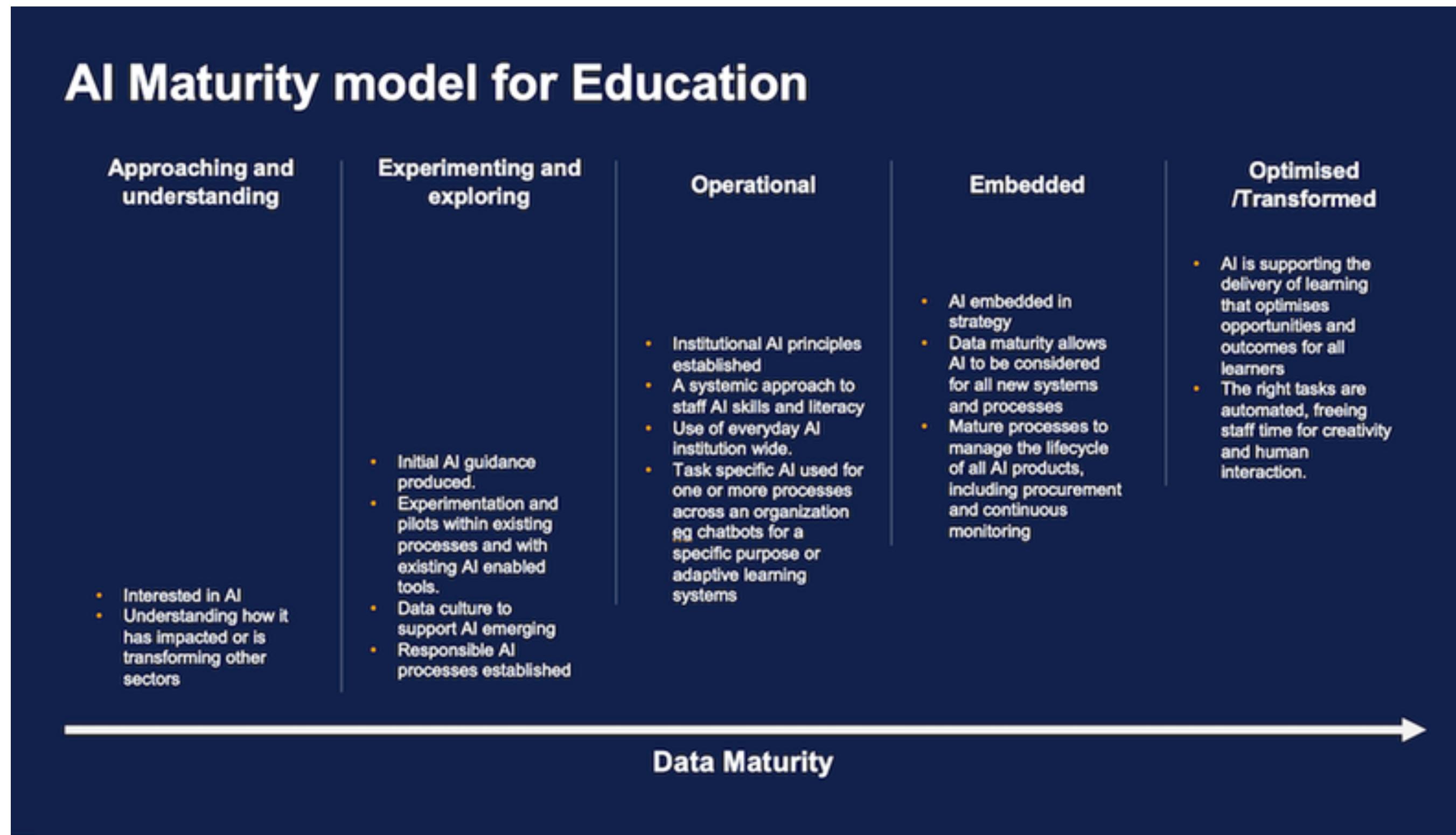


Quels leviers ? Quelques exemples



Source : *Contribution de l'IA à une aéronautique pérenne*, Collectif IA pour l'aéronautique, 2025

L'IA dans l'enseignement supérieur et la recherche



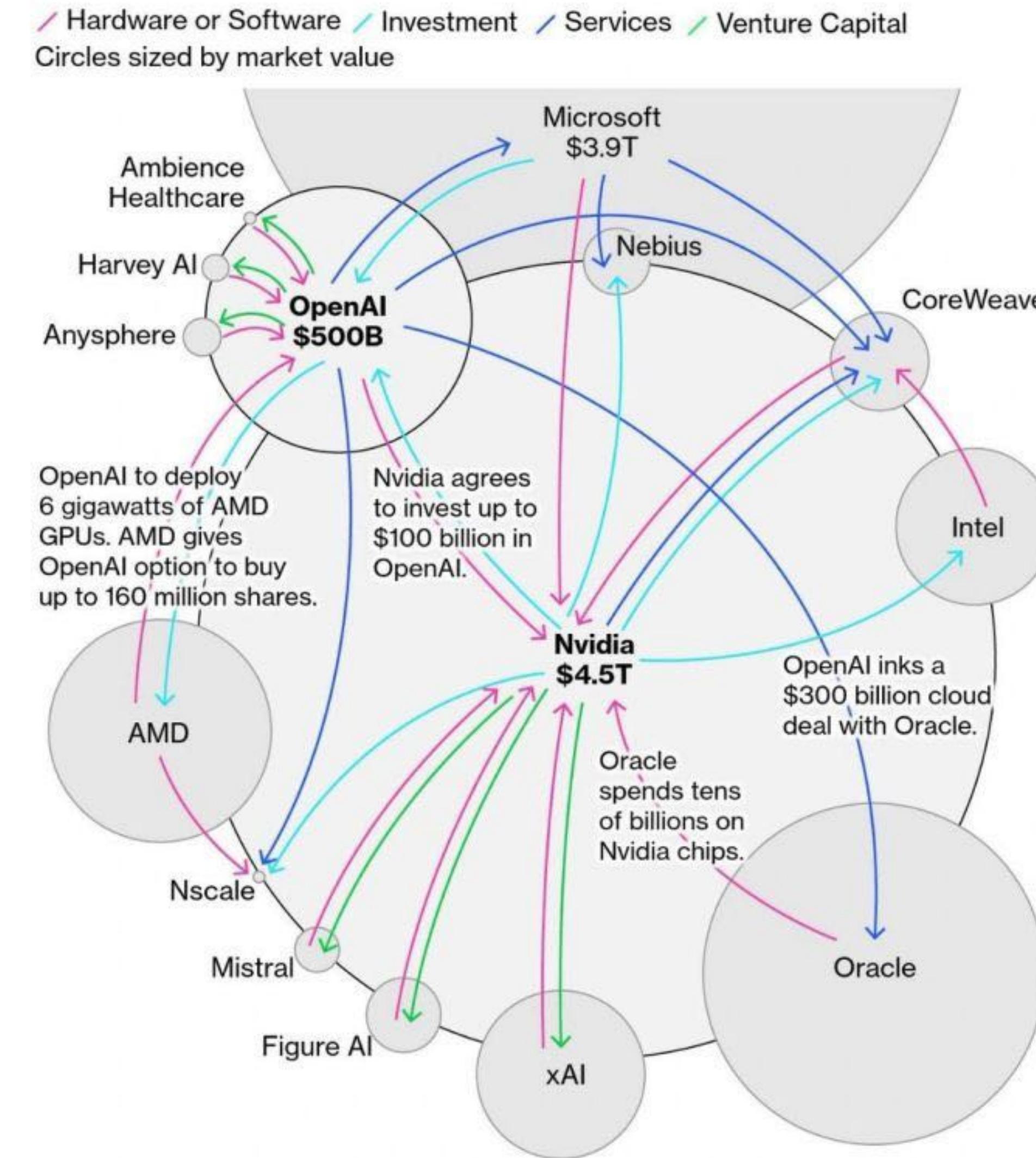
L'intelligence artificielle à l'université de Bordeaux

Mission d'étude et de conseil sur l'usage de l'intelligence artificielle dans les activités de l'université de Bordeaux

Rapport - Hélène Jacquet,
Achevé de rédiger le 29 août 2025.

Document à destination de l'équipe de direction de l'université de Bordeaux et de ses conseils.

Dépendances croisées



Analyse du cycle de vie d'un modèle d'IA

Modèle Large 2 de Mistral AI Analyse du cycle de vie



La localisation des centres de données est l'un des principaux facteurs de l'impact environnemental d'un modèle.
Chez Mistral AI, nous construisons

L'empreinte d'un modèle est fortement liée à sa taille. Opter pour des modèles plus petits ou spécifiques à un cas d'usage permet d'atténuer son impact environnemental.
Chez Mistral AI,

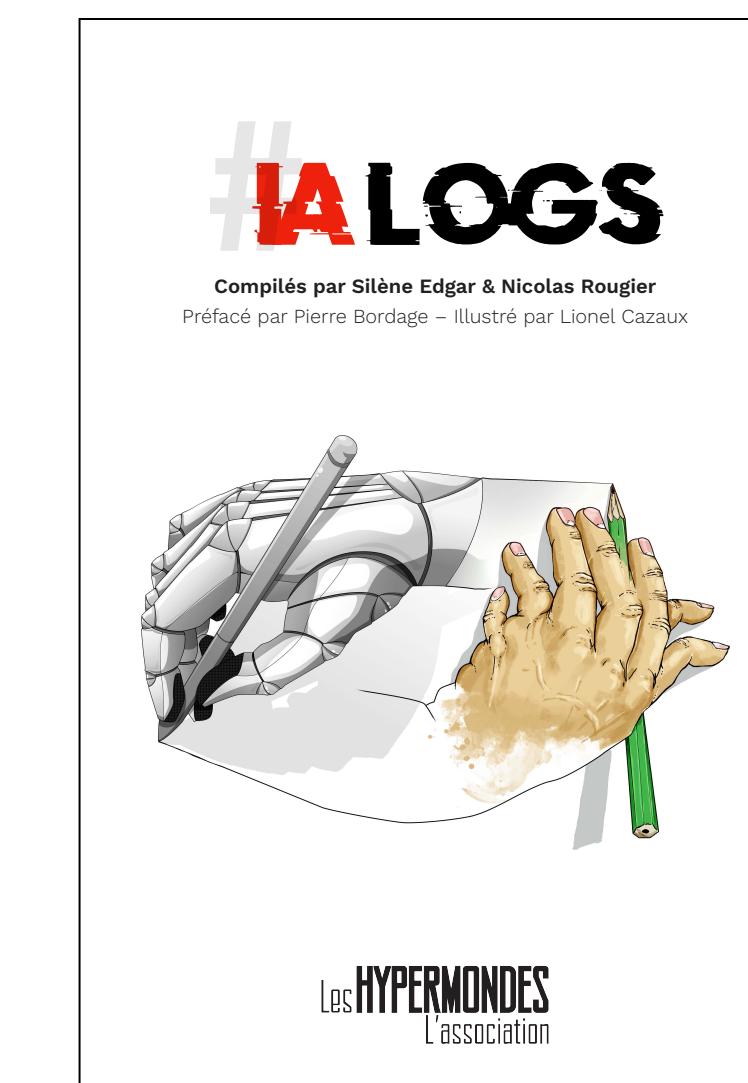
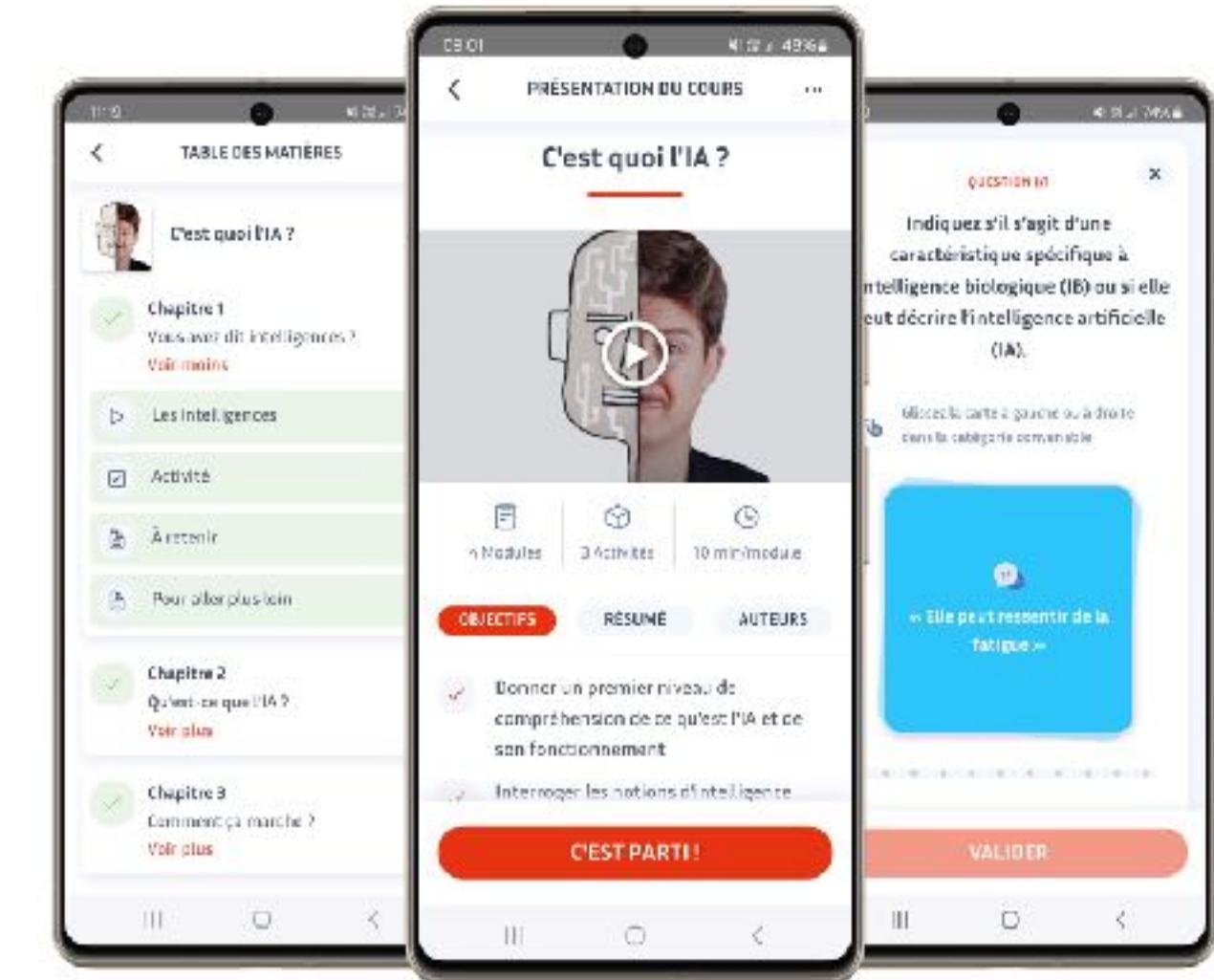
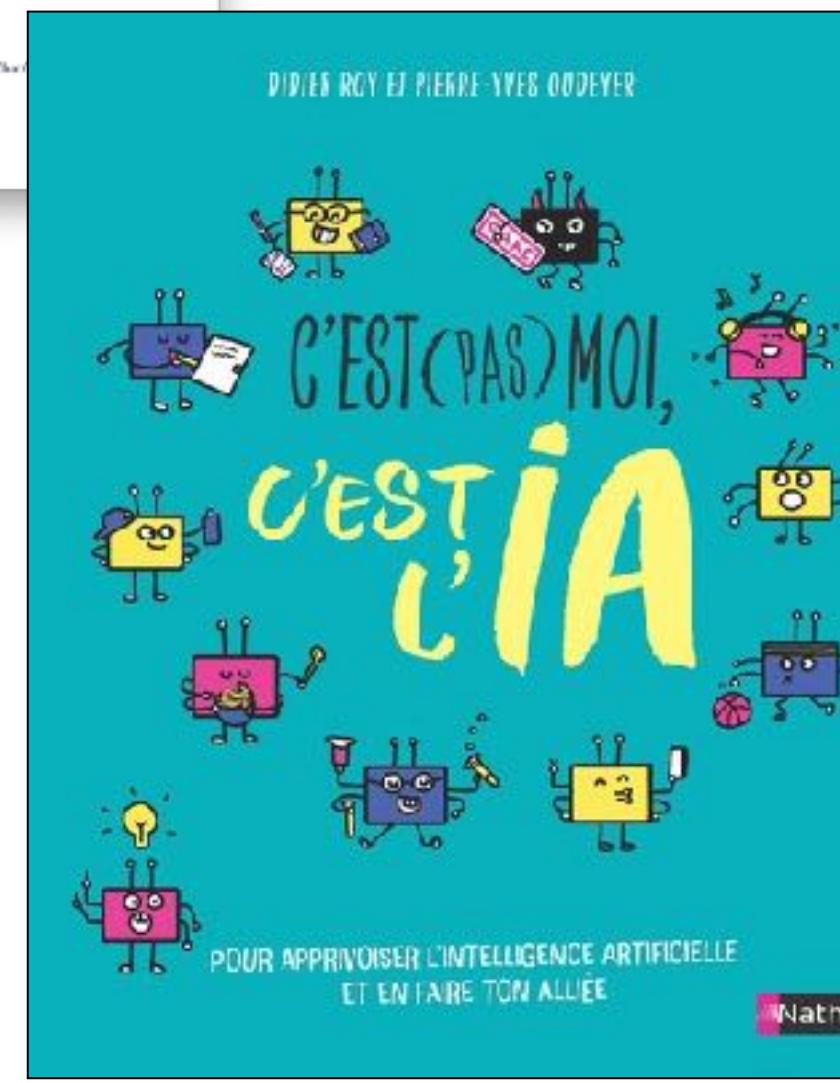
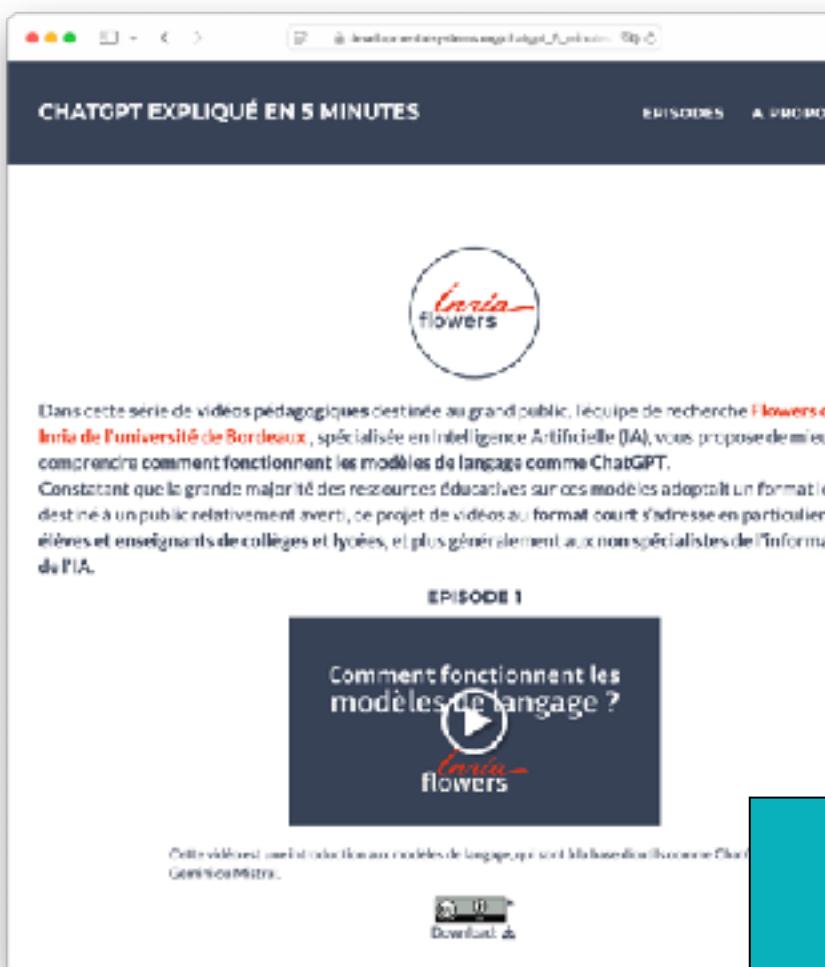
L'inférence étant une étape clé du cycle de vie d'un modèle, les utilisateurs finaux peuvent également contribuer à réduire son impact.
Chez Mistral AI,

Modèle Large 2 de Mistral AI Impact d'une page de texte générée (400 tokens)



Source : Mistral AI, juillet 2025

Médiation scientifique & médiation sociétale





Recherche

Donner Me connecter

GOUVERNEMENT toutes et tous égaux Tech POUR TOUTES

Je veux agir Je m'inscris

TROUVE TA VOIE DANS LA TECH

5 bonnes raisons de te lancer !

The banner features a group of diverse young women smiling and holding books, with a large purple overlay containing the text "TROUVE TA VOIE DANS LA TECH" and a button "5 bonnes raisons de te lancer !".

TechpourToutes est un programme gratuit qui vise à soutenir les filles et les jeunes femmes dans leurs parcours dans le numérique, en leur donnant les moyens d'atteindre le job de leurs rêves.



Sensibiliser



Accompagner



Transformer

**Féminiser
le numérique
pour
changer le monde**

