

Introduction à l'Intelligence Artificielle

Fabrice JUMEL
CPE LYON, CITI LAB. INRIA/CHROMA, ELESIA

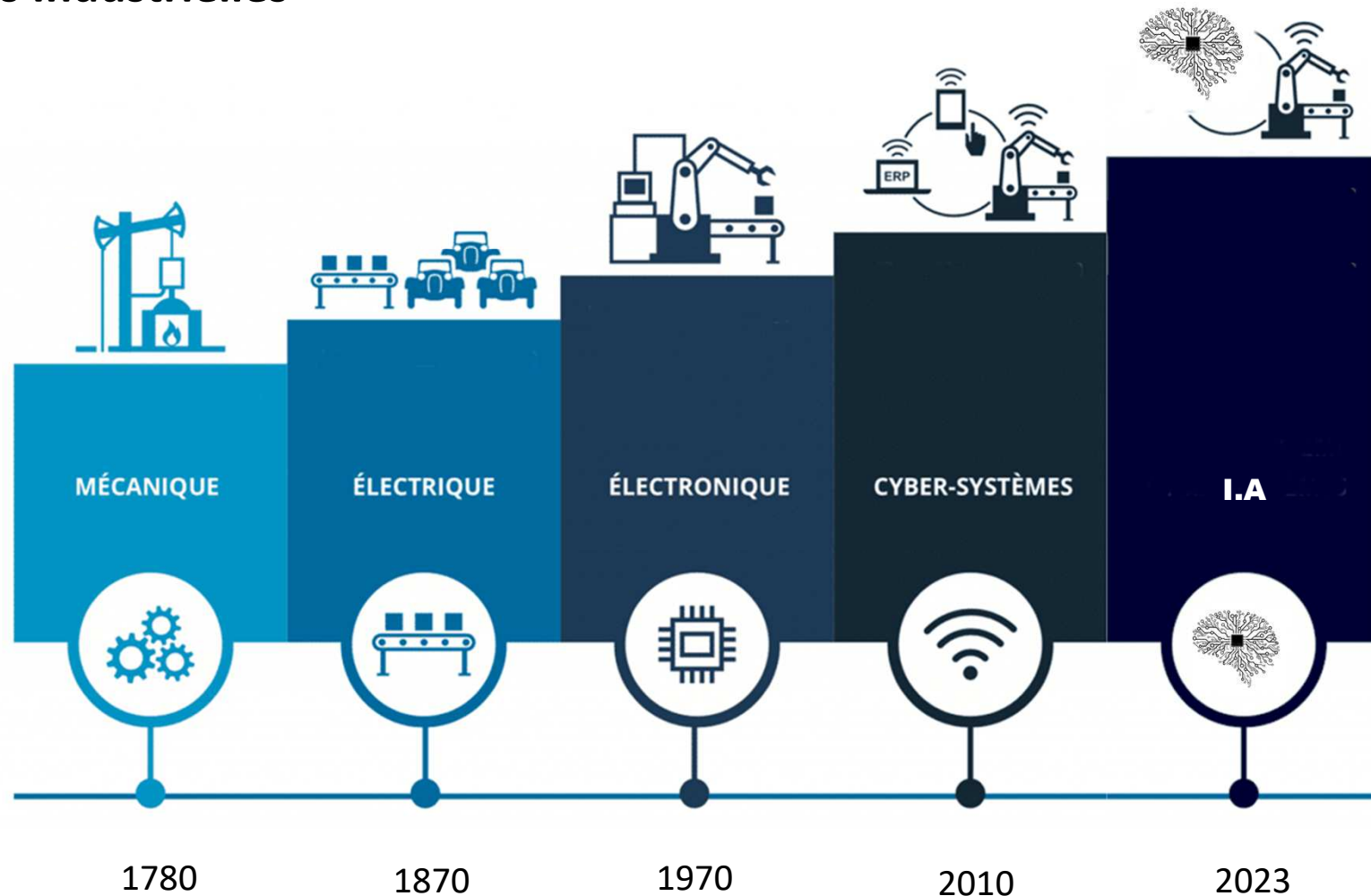


Attribution-NonCommercial-ShareAlike
4.0 International (CC BY-NC-SA 4.0)

établissement
associé



Révolutions Industrielles



<https://www.groupehyperforme.com/la-4e-revolution-industrielle-est-commencee/>

Lancement de la Roadmap

« Computing Machinery and Intelligence », MIND 1950, A. M. Turing

« Concernant le développement de l'IA deux approches doivent être menées de front.

L'une, est de s'attaquer à jouer à des activités abstraites comme les échecs.

L'autre, de faire une machine pouvant apprendre à nos cotés comme nous apprenons à nos enfants en montrant du doigt et en désignant les objets de notre quotidien en langage naturel »

460 A. M. TURING : COMPUTING MACHINERY AND INTELLIGENCE

of the different genetical combinations that had been tried, so as to avoid trying them again ?



We can only see a short distance ahead, but we can see plenty there that needs to be done.

BIBLIOGRAPHY

- Samuel Butler, Erewhon, London, 1865. Chapters 23, 24, 25, *The Book of the Machines*.
Alonzo Church, "An Unsolvable Problem of Elementary Number Theory", *American J. of Math.*, 58 (1936), 345-363.
K. Gödel, "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I", *Monatshefte für Math. und Phys.*, (1931), 173-189.

456

A. M. TURING :

Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain. Presumably the child-brain is something like a note-book as one buys it from the stationers. Rather little mechanism, and lots of blank sheets. (Mechanism and writing are from our point of view almost synonymous.) Our hope is that there is so little mechanism in the child-brain that something like it can be easily programmed. The amount of work in the education we can assume, as a first approximation, to be much the same as for the human child.

We have thus divided our problem into two parts. The child-programme and the education process. These two remain very closely connected. We cannot expect to find a good child-machine at the first attempt. One must experiment with teaching one such machine and see how well it learns. One can then try another and see if it is better or worse. There is an obvious connection between this process and evolution, by the identifications

Structure of the child machine	=	Hereditary material
Changes	" "	= Mutations
Natural selection		= Judgment of the experimenter

One may hope, however, that this process will be more expeditious than evolution. The survival of the fittest is a slow method for measuring advantages. The experimenter, by the exercise of intelligence, should be able to speed it up. Equally important is the fact that he is not restricted to random mutations. If he can trace a cause for some weakness he can probably think of the kind of mutation which will improve it.

It will not be possible to apply exactly the same teaching process to the machine as to a normal child. It will not, for instance, be provided with legs, so that it could not be asked to go out and fill the coal scuttle. Possibly it might not have eyes. But however well these deficiencies might be overcome by clever engineering, one could not send the creature to school without the other children making excessive fun of it. It must be given some tuition. We need not be too concerned about the legs, eyes, etc. The example of Miss *Helen Keller* shows that education can take place provided that communication in both directions between teacher and pupil can take place by some means or other.

Dualité de l'intelligence Artificielle

L'IA connexionniste : on part ici du principe que la machine n'est pas intelligente. Elle perçoit à travers les données des relations qui lui permettent de produire toute une série de perceptions.

L'IA symbolique : une manière de permettre à l'ordinateur de raisonner et de reproduire un raisonnement humain.

L'IA général : vise à créer des systèmes capables de comprendre, apprendre, raisonner et agir comme le ferait un être humain dans divers domaines, sans se limiter à des tâches spécifiques.

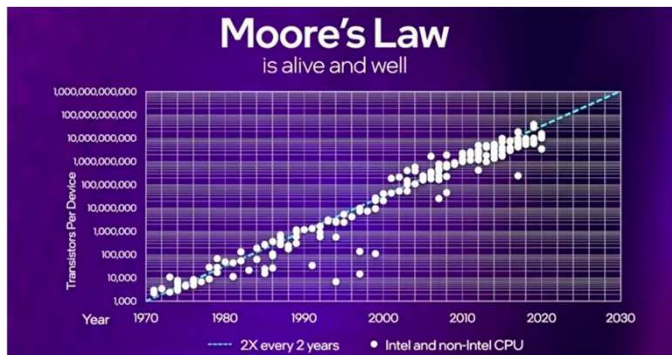
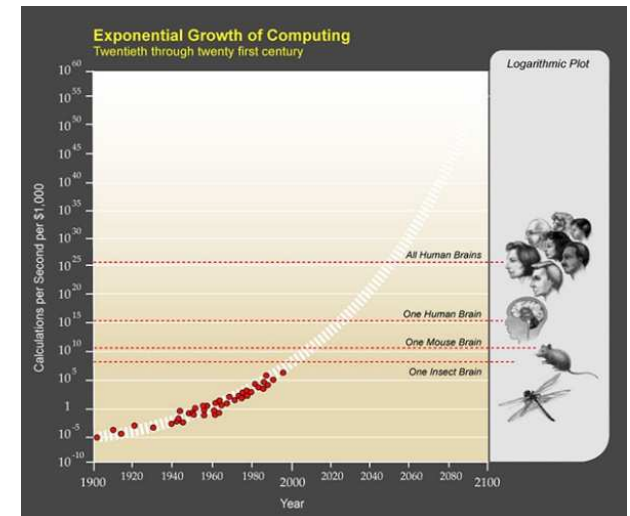
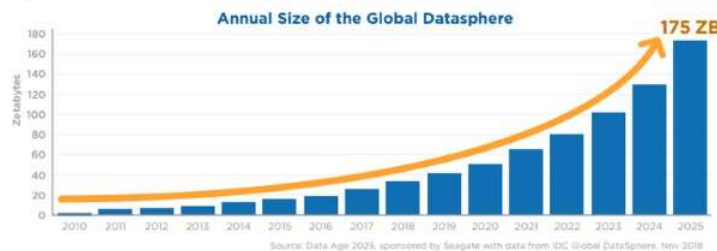
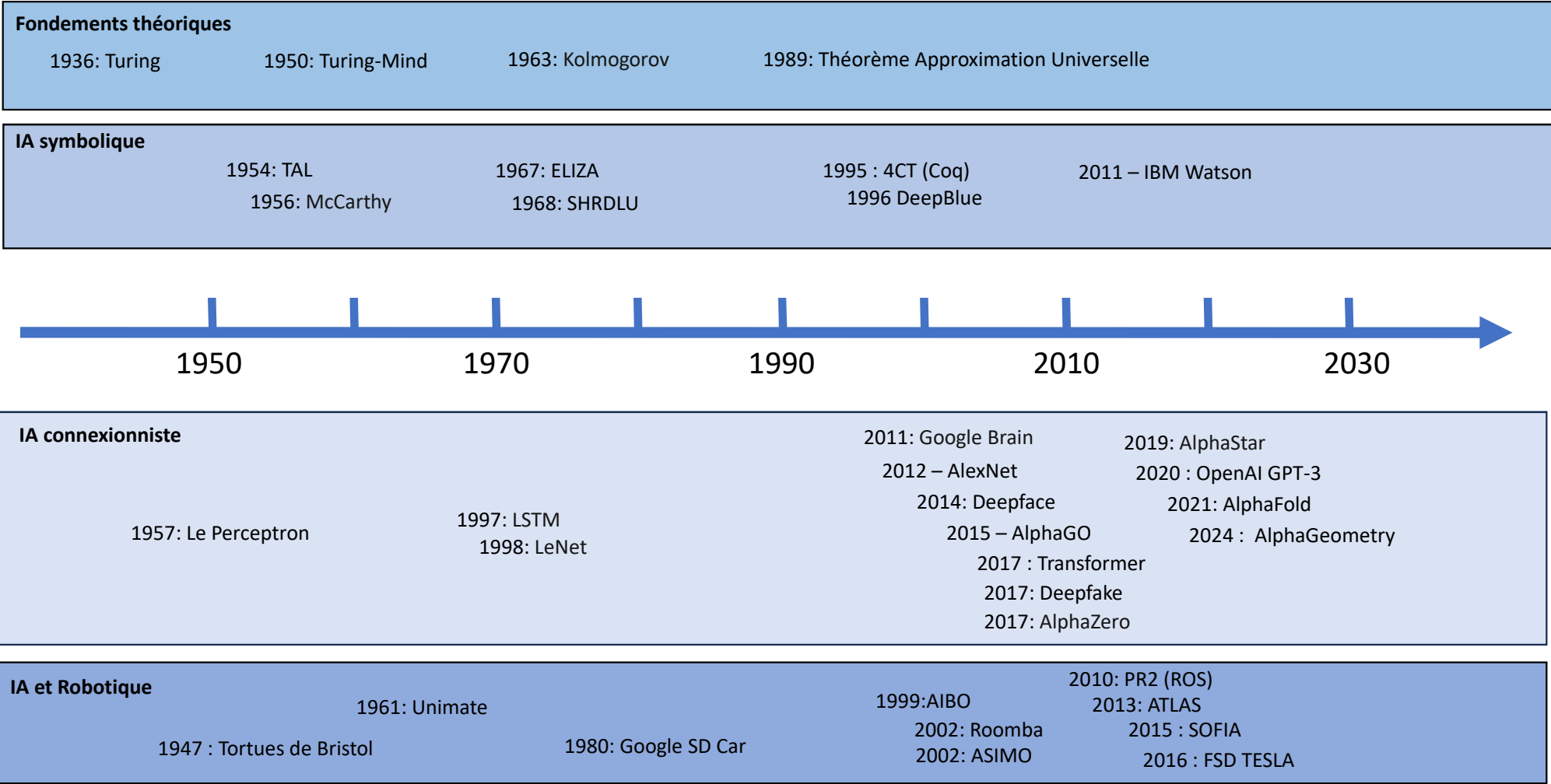


Figure 1 - Annual Size of the Global Datasphere



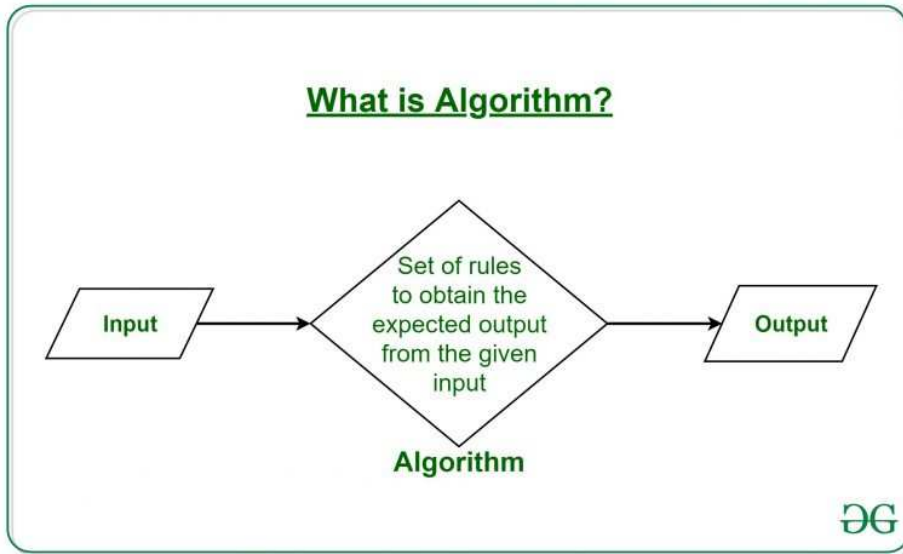
Quelques dates clefs de l'intelligence artificielle



Quelques définitions de l'Intelligence Artificielle

- L'ensemble des technologies informatiques, capable de résoudre des problèmes à forte complexité algorithmique,
- Algorithmes basés sur le biomimétisme et/ou l'imitation des fonctions cognitives humaines
- Algorithmes dont le fonctionnement n'est pas bien compris par les non experts.
- Algorithmes nécessitant une phase d'apprentissage sur des données

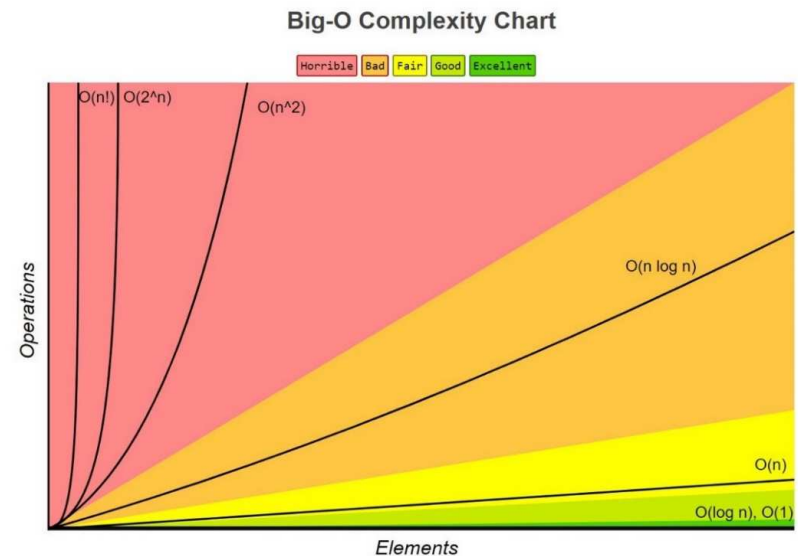
Rappel sur les algorithmes



Un algorithme à une certaine taille, nécessite du temps de calcul et des ressources mémoires.

La complexité du problème peut être liée à la taille de l'entrée et de la sortie) et/ou la variabilité associée.

Certains problèmes deviennent impossible à traiter si les temps de calcul évolue en fonction de l'entrée



Mathématique et algorithmes

Résoudre des problèmes complexes nécessite des outils mathématiques

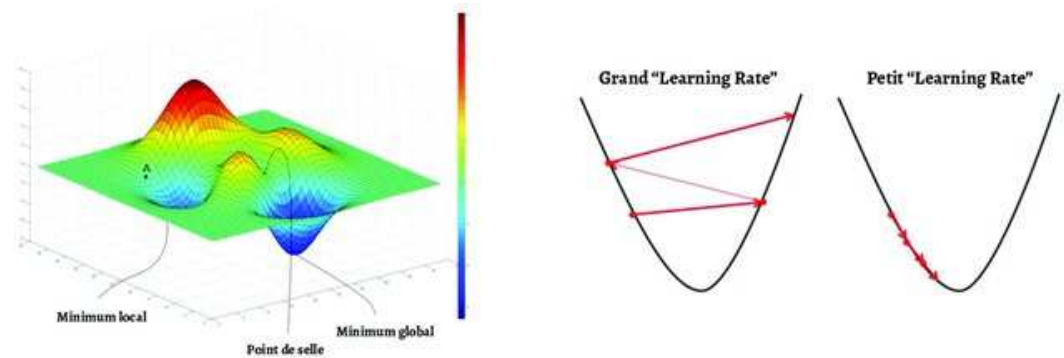
Optimisation
Classification
Parcours de graphe
Théorie des jeux
Preuves automatisés
Bases de connaissances
...

Importance de l'aléatoire (pour modéliser ou résoudre)
Random tree/forest
Algorithme génétique
Méthodes de montée carlo
Approche agent/particule

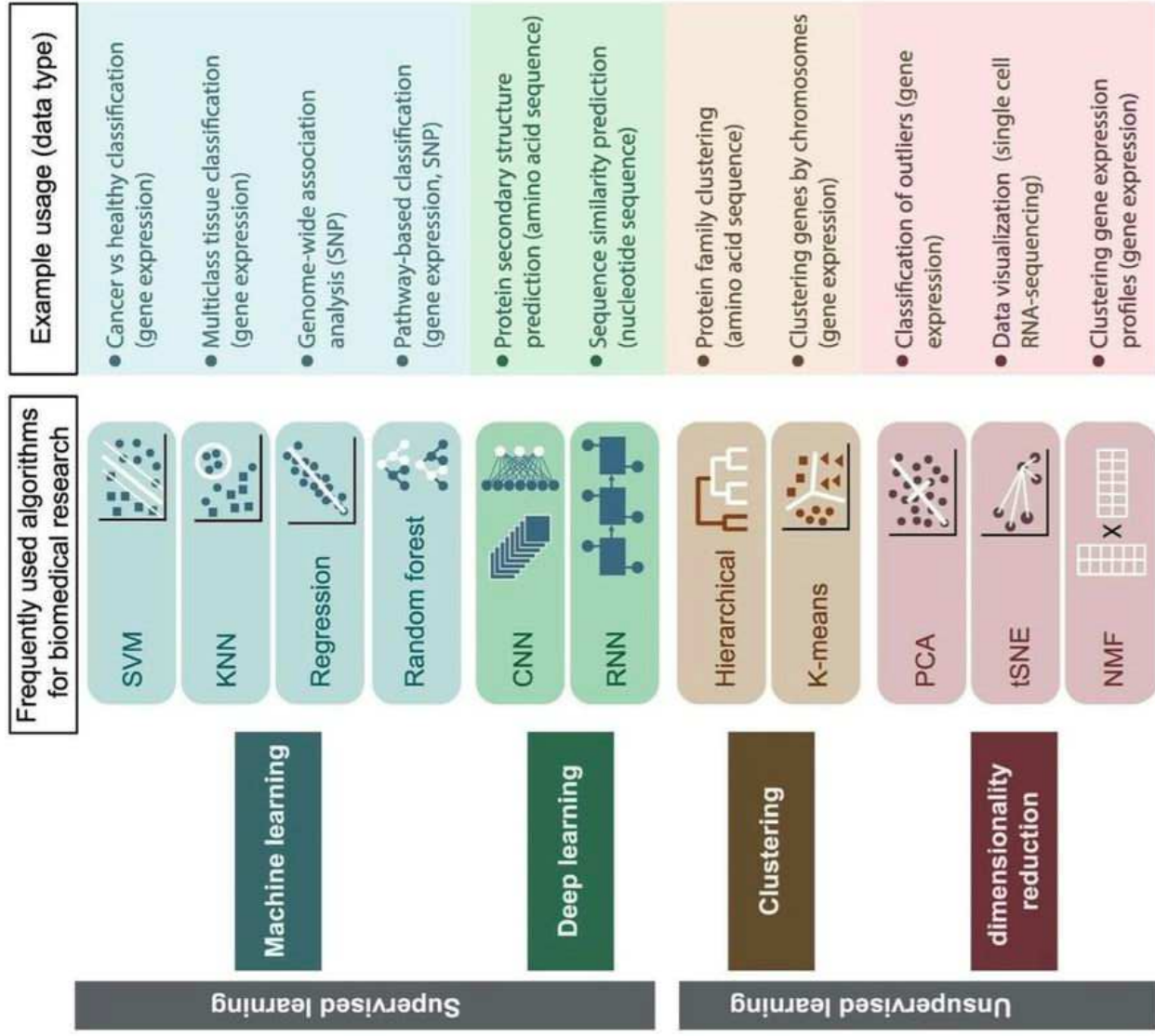
Outils de résolution mathématique => impose de modéliser le pb d'une façon précise (pas facile et peut changer la couverture du problème, on obtient par exemple une solution la meilleure avec cet outil mais peut être pas la meilleure en soit.)

- Beaucoup de méthodes nécessitent de paramétrer des coefficients (détails dans les algorithmes) pour s'adapter. Besoin de tuner ces algorithmes , approche dichotomique par exemple
- Problème plus complexe sur des multi paramètres , recherche de minimum dans des espaces multidimensionnelles. Approche de descente de gradient

- Convergence de l'algorithme
- Phase d'apprentissage
- Phase d'exploitation



- Besoin de beaucoup de données / temps d'apprentissage
- Certains algorithmes ont toujours besoin d'apprendre



Modes d'apprentissage

supervisé

classification

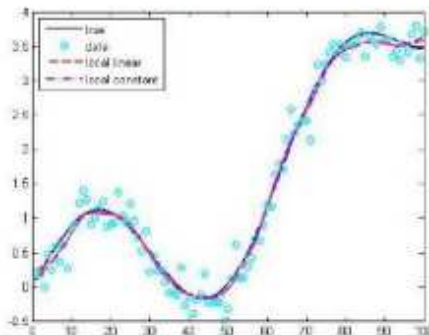
données avec label
(pixels) \rightarrow (label)



labelliser des images
identifier un virus
classifications binaires

régression

données chiffrées
prévoir (y) en fonction de (x)

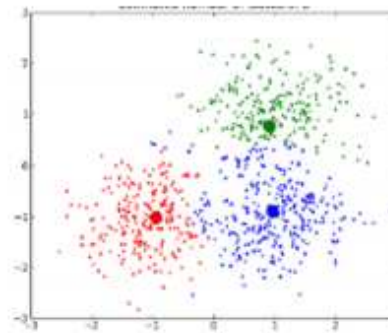


prévisions quantifiées
consommation électrique
ventes saisonnières

non supervisé

clustering

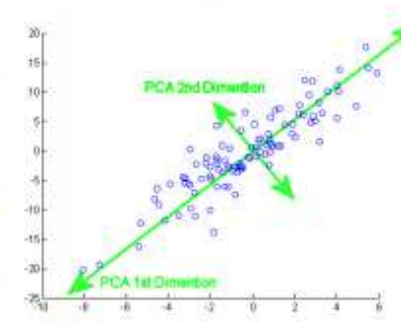
données sans label
(x, y, z, ...)



identification segments clients
regrouper des élèves
bizarrerie quelconque

réduction dimensions

données sans label
(x, y, z, ...)

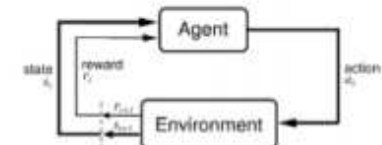


identifier des corrélations
entre des données
simplifier les modèles

par renforcement

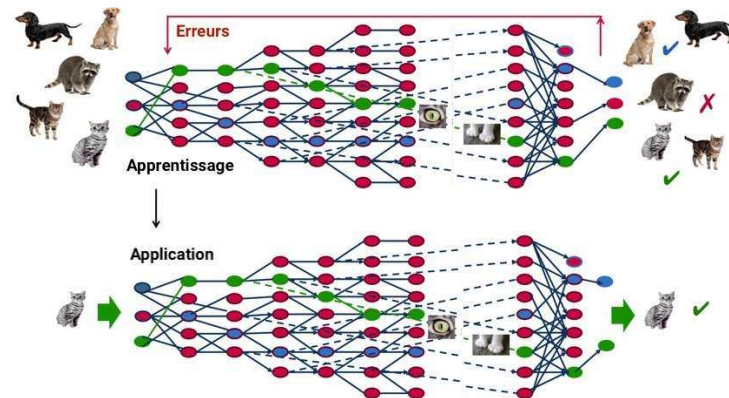
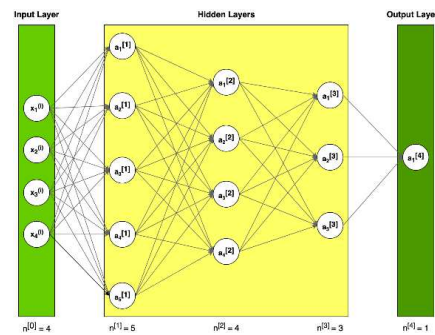
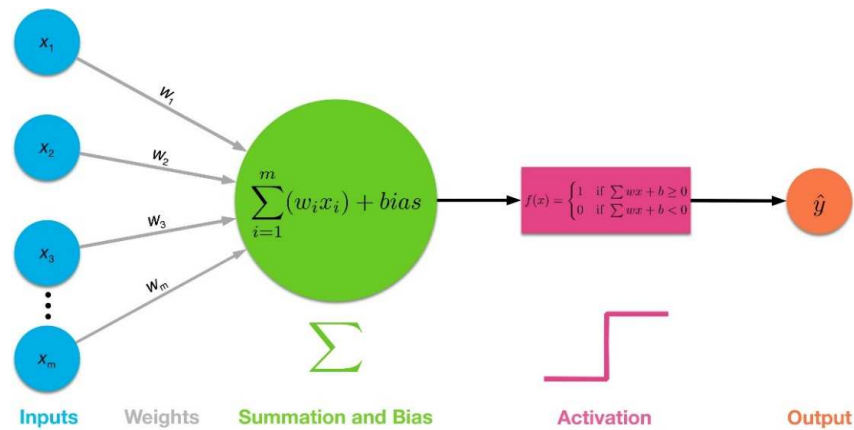
ajustement

toutes données
modèle entraîné

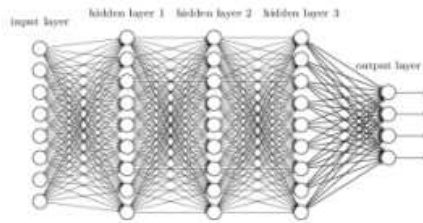


améliorer chatbot
apprentissage de robot

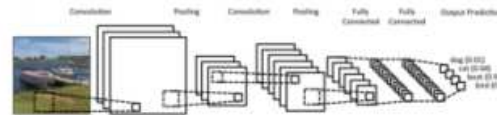
Réseau de Neurones



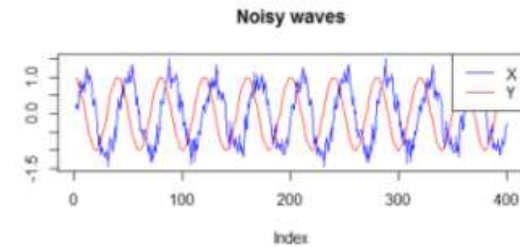
Quelques Architectures de Deeplearning



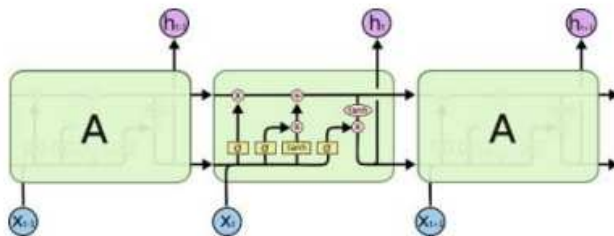
fully connected
classification
et prédictions



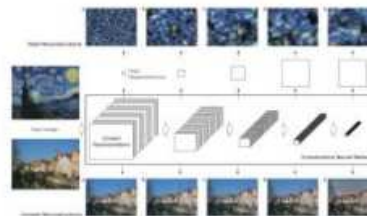
convolutionnels
spatial
reconnaissance images



récurrents
temporels
ECG, finance, bruit



LSTM
contexte - bidirectionnel
traduction, dialogue, recherche



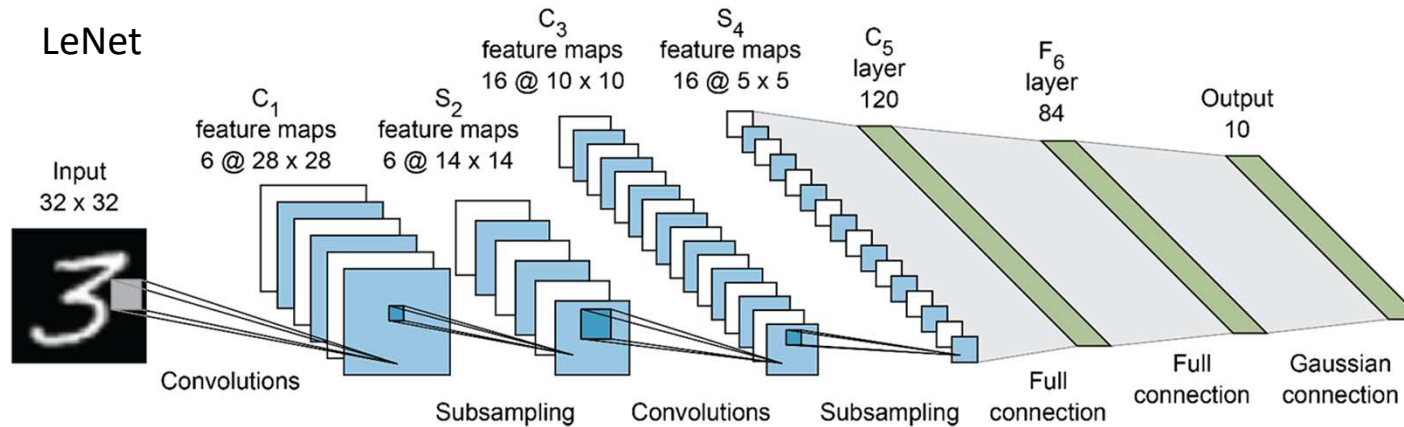
transfer networks
apprentissage incrémental
changement de domaine



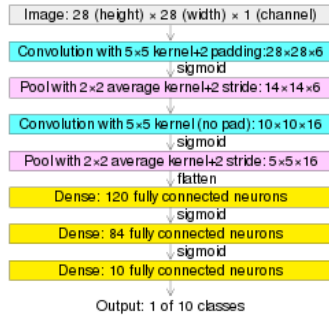
génératifs
variations – augmentation
modification d'images et de textes

Deeplearning, cas du CNN (Convolutional Neural Network)

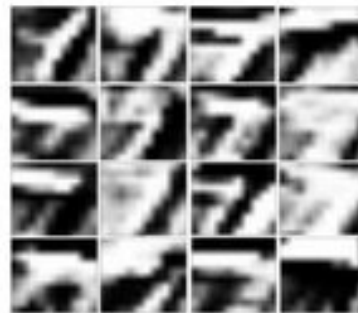
LeNet



LeNet



Convolutional layer C1
 $6 \times 28 \times 28$



Convolutional layer C3
 $16 \times 10 \times 10$

Démo : <https://cs.stanford.edu/people/karpathy/convnetjs/demo/mnist.html>

ConvNetJS MNIST demo

Description

This demo trains a Convolutional Neural Network on the [MNIST digits dataset](#) in your browser, with nothing but Javascript. The dataset is fairly easy and one should expect to get somewhere around 99% accuracy within few minutes. I used [this python script](#) to parse the [original files](#) into batches of images that can be easily loaded into page DOM with img tags.

This network takes a 28×28 MNIST image and crops a random 24×24 window before training on it (this technique is called data augmentation and improves generalization). Similarly to do prediction, 4 random crops are sampled and the probabilities across all crops are averaged to produce final predictions. The network runs at about 5ms for both forward and backward pass on my reasonably decent Ubuntu+Chrome machine.

By default, in this demo we're using Adadelta which is one of per-parameter adaptive step size methods, so we don't have to worry about changing learning rates or momentum over time. However, I still included the text fields for changing these if you'd like to play around with SGD+Momentum trainer.

Report questions/bugs/suggestions to [@karpathy](#)

Training Stats

Forward time per example: 2ms
Backprop time per example: 3ms
Classification loss: 2.20016
L2 Weight decay loss: 0.00085
Training accuracy: 0.22
Validation accuracy: 0.25714
Examples seen: 252
Learning rate: 0.01
Momentum: 0.9
Batch size: 32
Weight decay: 0.001

Instantiate a Network and Trainer

```
layer_defs = [
    layer_def.push({type:'input', out_sz:24, out_sz2:24, out_depth:1}),
    layer_def.push({type:'conv', szx:5, filterx:5, stride:1, pad:1, activation:'relu'}),
    layer_def.push({type:'pool', szx:2, stride:2}),
    layer_def.push({type:'conv', szx:5, filterx:5, stride:1, pad:1, activation:'relu'}),
    layer_def.push({type:'pool', szx:2, stride:2}),
    layer_def.push({type:'softmax', num_classes:10})
]
net = new convnetjs.Net();
net.makeLayers(layer_defs);
trainer = new convnetjs.SGDTrainer(net, {method:'adadelta', batch_size:32, l2_decay:0.001});
```

Network Visualization

Input ($24 \times 24 \times 1$)
max activation: 0.96607, min: 0
max gradient: 0.10002, min: -0.09581

Activations:

Activation Gradients:

conv ($24 \times 24 \times 8$)
filter size 5×5 , stride 1
max activation: 1.84505, min: -1.90654
max gradient: 0.07614, min: -0.07648
parameters: $8 \times 5 \times 5 \times 1 + 8 = 208$

Activations:

Activation Gradients:

Weights:

Weight Gradients:

relu ($24 \times 24 \times 8$)
max activation: 1.64565, min: 0
max gradient: 0.07614, min: -0.07648

Activations:

Activation Gradients:

pool ($12 \times 12 \times 8$)
pooling size 2×2 , stride 2
max activation: 1.84505, min: 0
max gradient: 0.07614, min: -0.07648

Activations:

Activation Gradients:

conv ($12 \times 12 \times 16$)
filter size 5×5 , stride 1
max activation: 1.38018, min: -1.91901
max gradient: 0.15682, min: -0.14014
parameters: $16 \times 5 \times 5 \times 1 + 16 = 3216$

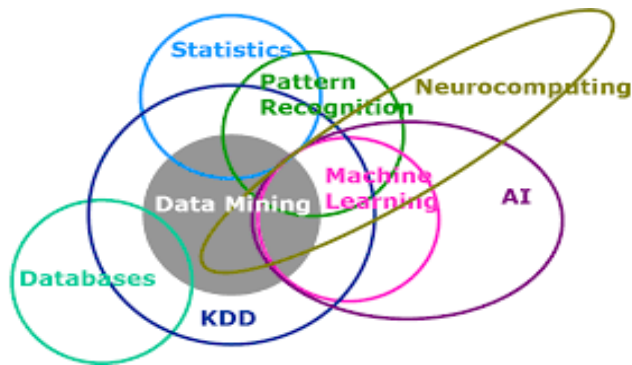
Activations:

Activation Gradients:

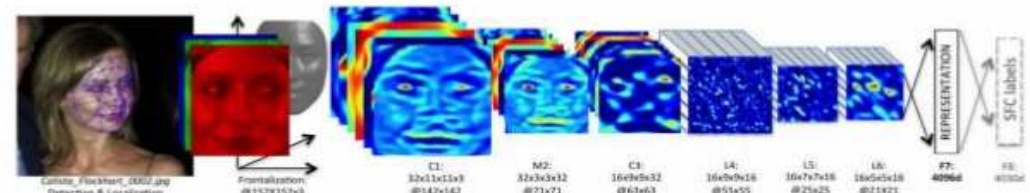
Weights:

Weight Gradients:

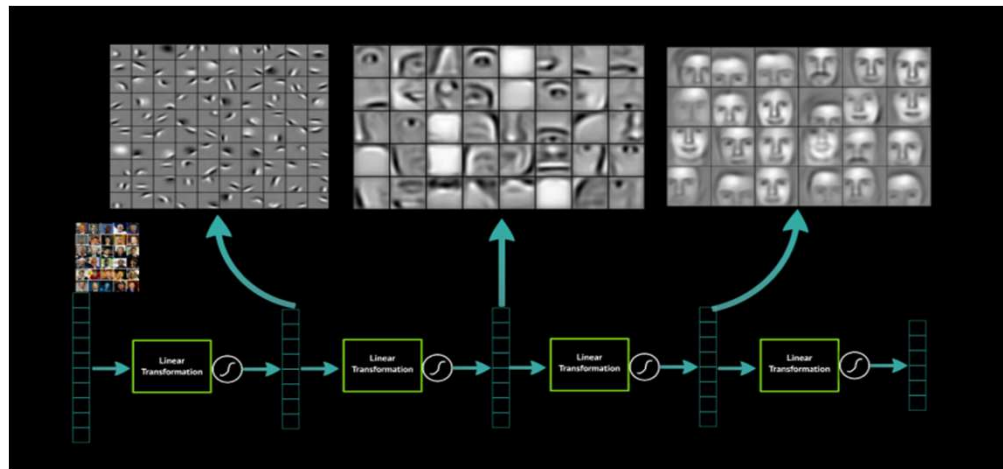
Extraction des Caractéristiques (Features)



DeepFace Architecture

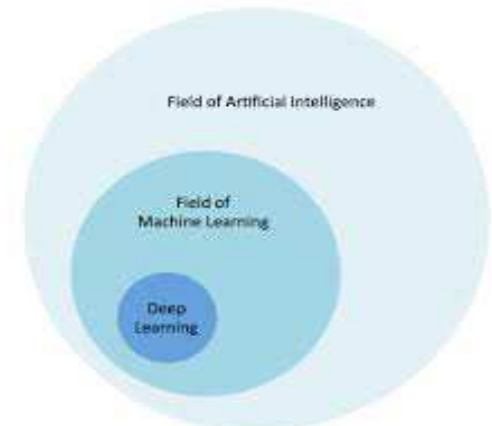


Yaniv Taigman, etc (Facebook) . [DeepFace: Closing the Gap to Human-Level Performance in Face Verification](#), CVPR 2014

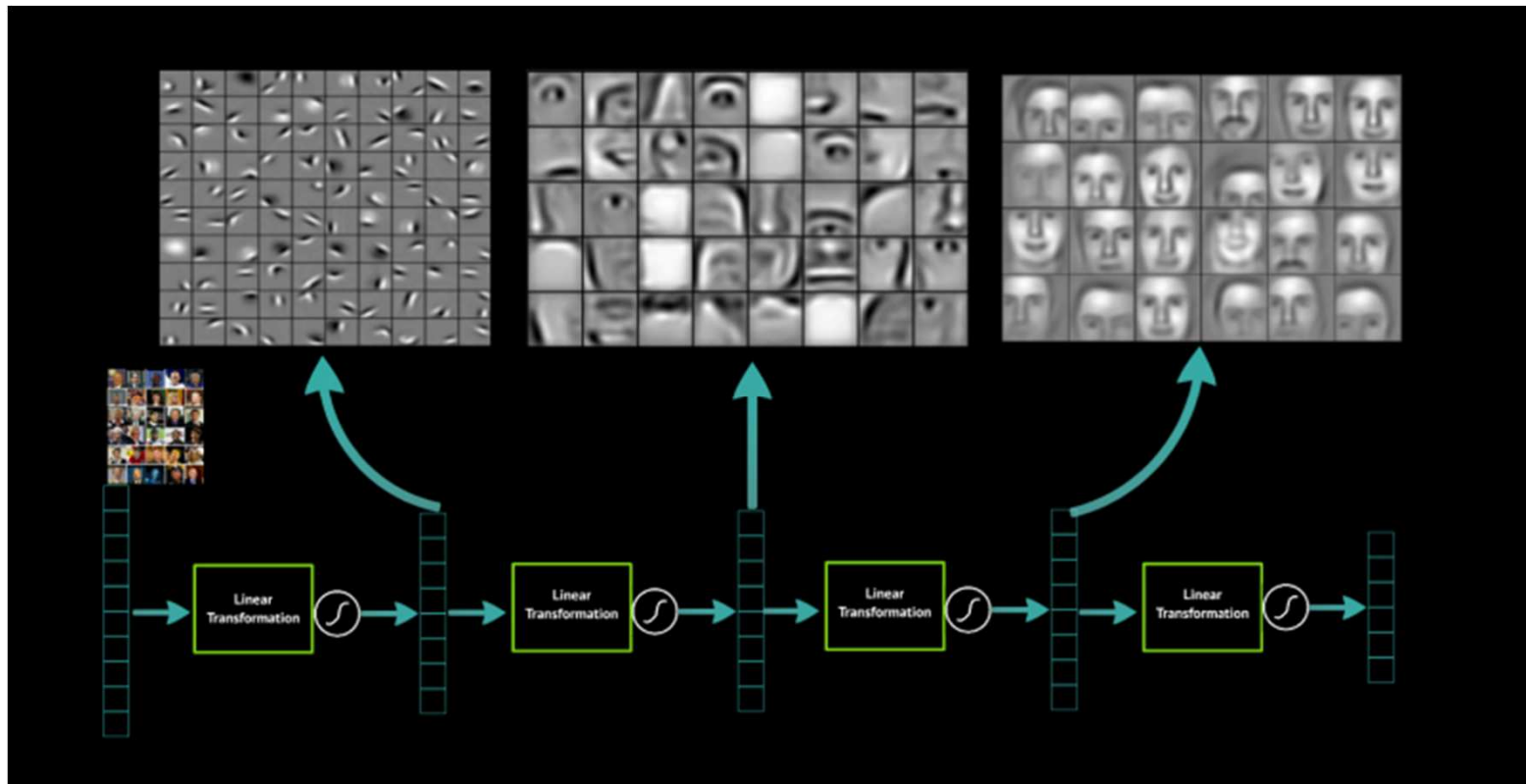


<https://www.datarobot.com/blog/a-primer-on-deep-learning/>

$$\begin{aligned} -\frac{\partial \log p(v)}{\partial W_{ij}} &= E_v[p(h_i|v) \cdot v_j] - v_j^{(i)} \cdot \text{sigm}(W_i \cdot v^{(i)} + c_i) \\ -\frac{\partial \log p(v)}{\partial c_i} &= E_v[p(h_i|v)] - \text{sigm}(W_i \cdot v^{(i)}) \\ -\frac{\partial \log p(v)}{\partial b_j} &= E_v[p(v_j|h)] - v_j^{(i)} \end{aligned}$$



Extraction des Features (Zoom)



<https://www.datarobot.com/blog/a-primer-on-deep-learning/>

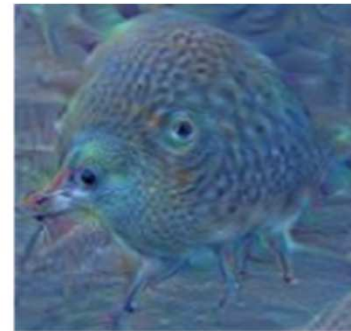
Intuition en action



"Admiral Dog!"



"The Pig-Snail"



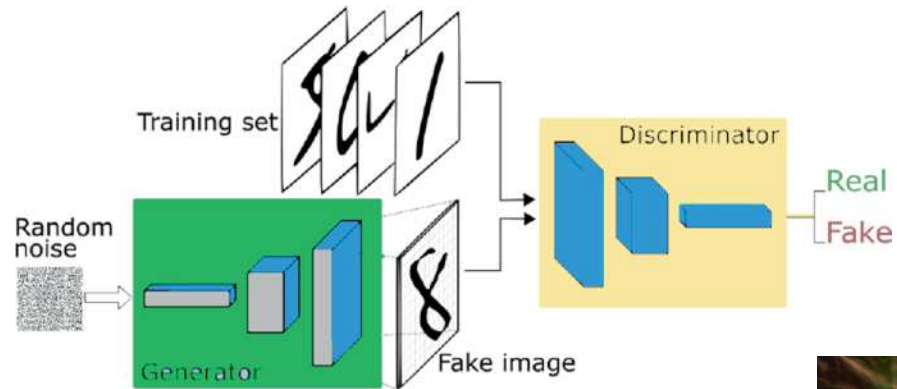
"The Camel-Bird"



"The Dog-Fish"

Source Inceptionism: Going Deeper into Neural Networks, google AI blog

IA générative, Exemple de GAN (Generative Adversarial Network)



Stable Diffusion

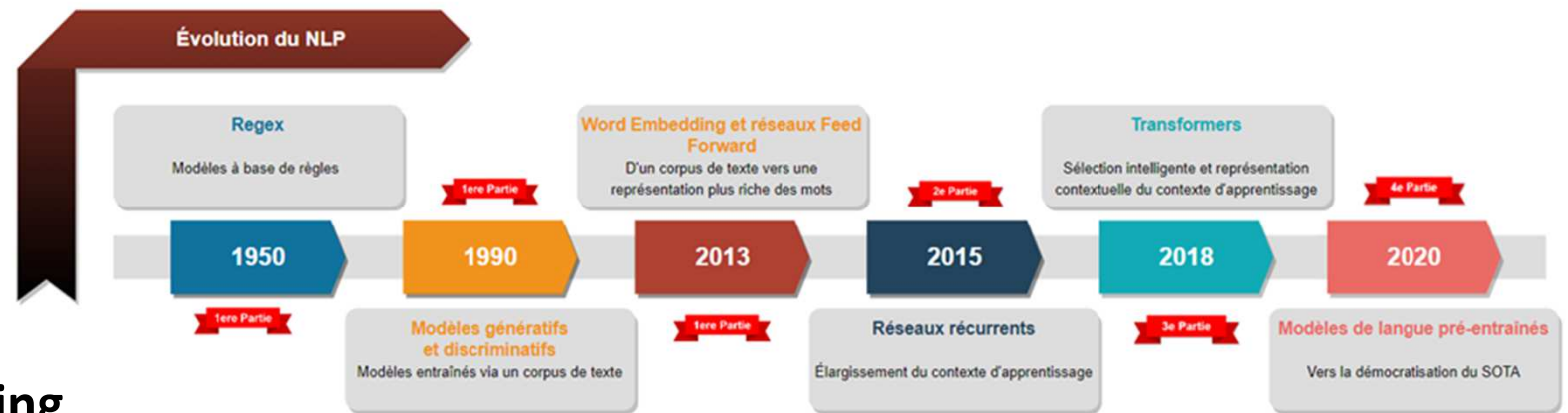
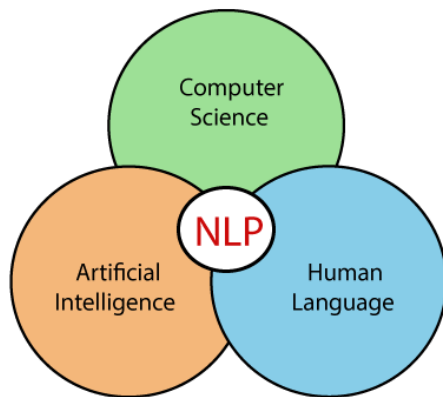
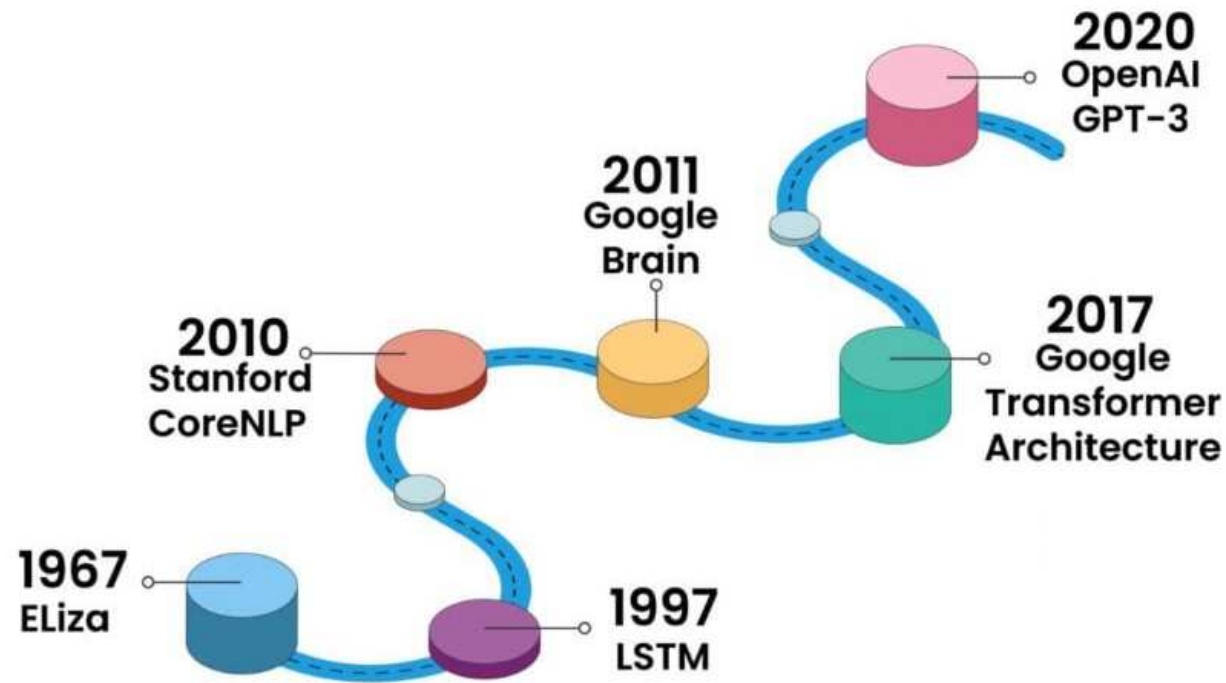


Sora



prompt: “A stylish woman walks down a Tokyo street filled with warm glowing neon and animated city signage. She wears a black leather jacket, a long red dress, and black boots, and carries a black purse. She wears sunglasses and red lipstick. She walks confidently and casually. The street is damp and reflective, creating a mirror effect of the colorful lights. Many pedestrians walk about.”

Chat GPT ?



Natural Language Processing

ChatGPT n'utilise de modèle de langages connu a priori, ce n'est pas une IA symbolique



<https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/>

Le cœur des approches est lié à la prédiction du mot suivant

Predict Next Word

The best thing about AI is its ability to

learn	4.5%
predict	3.5%
make	3.2%
understand	3.1%
do	2.9%

```
In[ ]:= NestList[StringJoin[#, model[#, "Decision"]]&,
"The best thing about AI is its ability to", 7]
```

```
Out[ ]:= {The best thing about AI is its ability to,
The best thing about AI is its ability to learn,
The best thing about AI is its ability to learn from,
The best thing about AI is its ability to learn from experience,
The best thing about AI is its ability to learn from experience.,
The best thing about AI is its ability to learn from experience. It,
The best thing about AI is its ability to learn from experience. It's,
The best thing about AI is its ability to learn from experience. It's not}
```

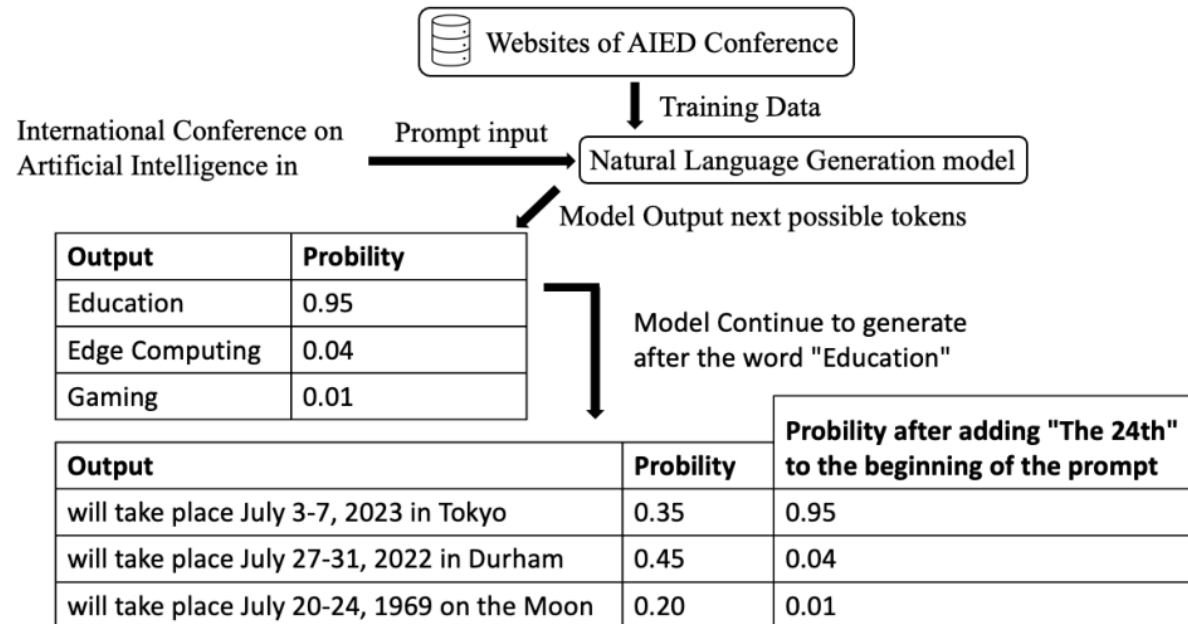
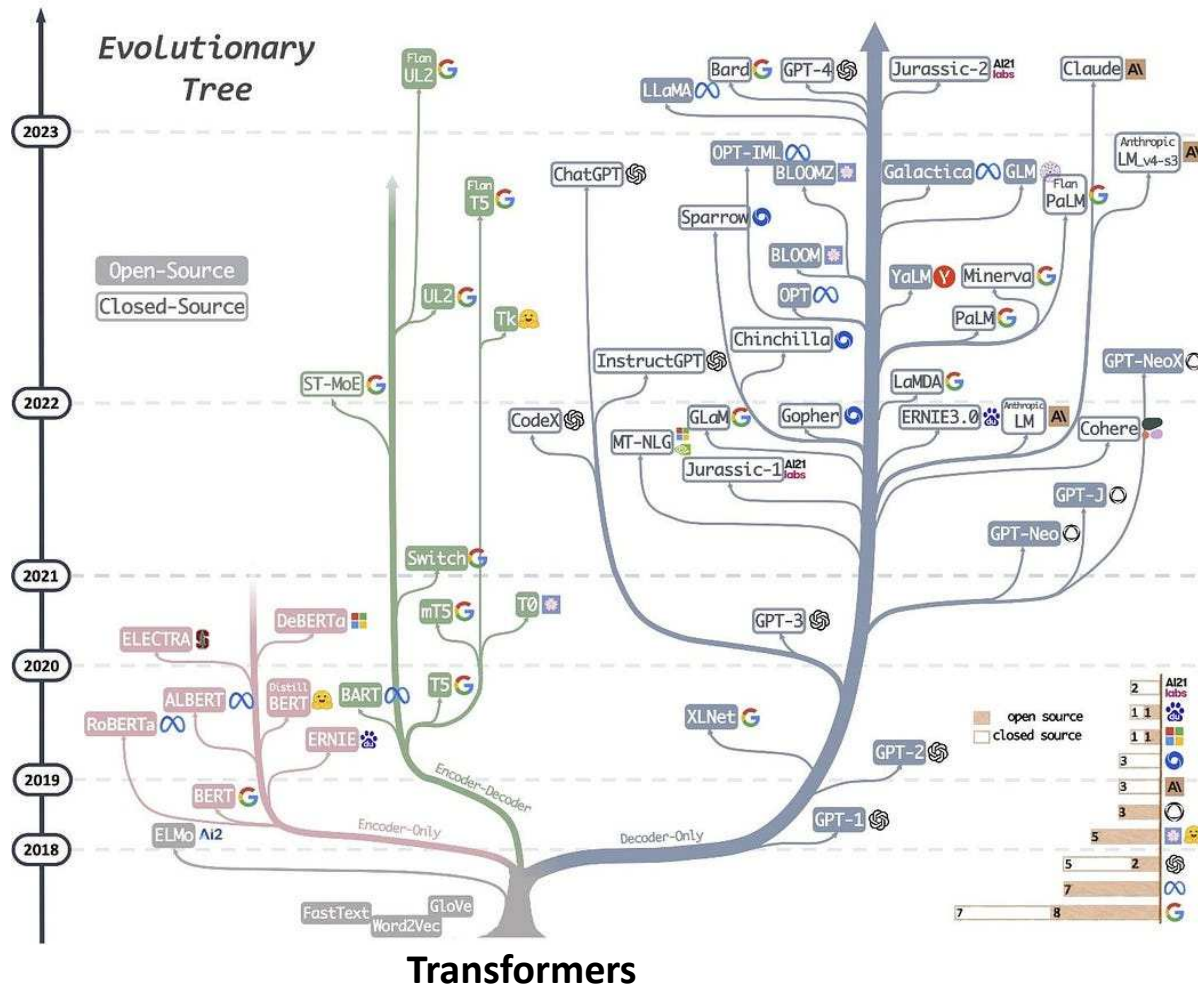


Fig. 1. Example of probability on generating different outputs during NLG process.

Attention is all you need.



<https://medium.com/@paul.k.pallaghy/a-fascinating-tree-of-gpts-and-llms-reveals-whats-been-going-on-4d4235f2a2b1>

Attention is all you need (Transformer)

- Les Transformers sont une architecture révolutionnaire en intelligence artificielle, utilisée historiquement dans le traitement du langage naturel (Il faut maîtriser les RNN et les LSTM pour les comprendre)
- Modélise efficacement des séquences de longueurs variables grâce à des mécanismes d'attention en fait une méthode puissante.
- Performant dans une large gamme de tâches NLP, telles que la classification de texte, la traduction automatique, et la génération de texte.
- Adaptés au pré-entraînement massif et au fine-tuning, ils peuvent apprendre des représentations linguistiques générales à partir de grandes quantités de données non étiquetées.
- **Haut degré de parallélisme les rend efficaces pour l'entraînement rapide sur de vastes ensembles de données**

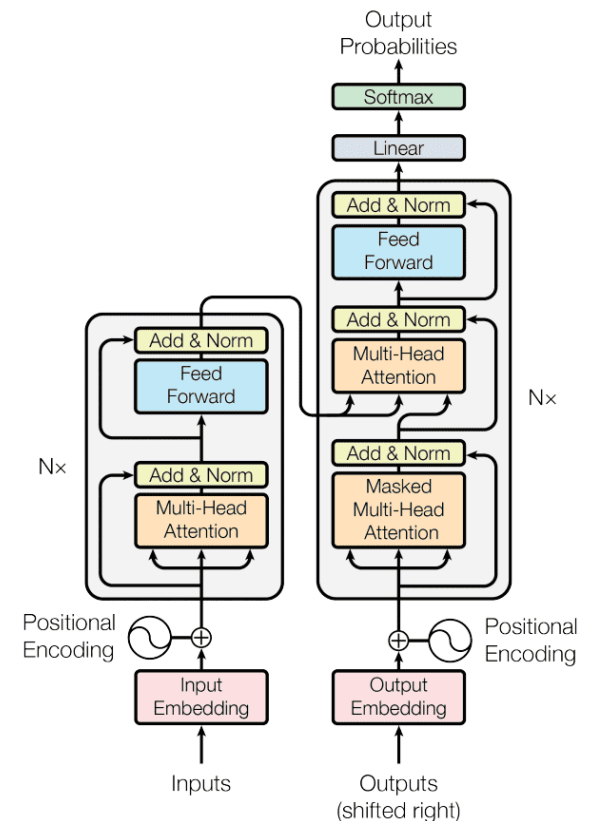
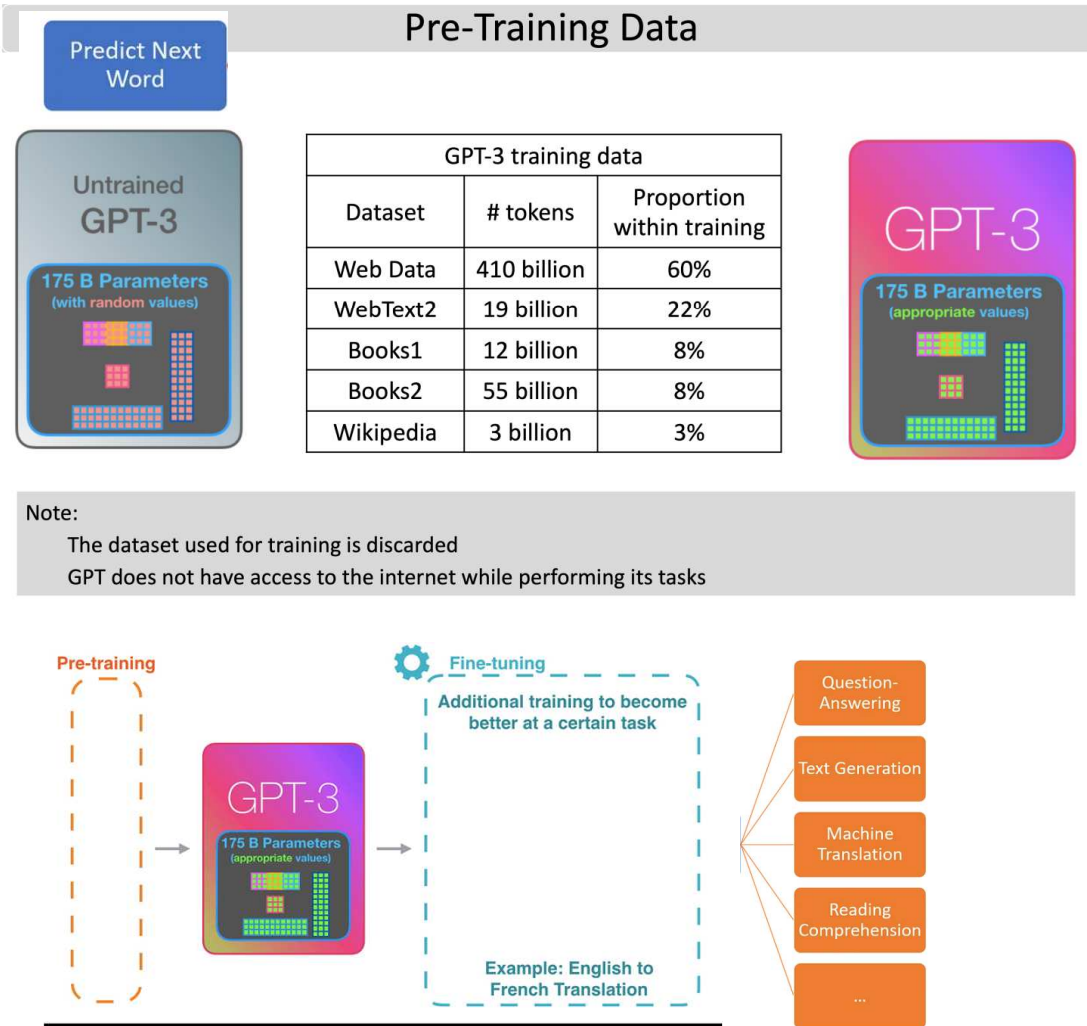


Figure 1: The Transformer - model architecture.

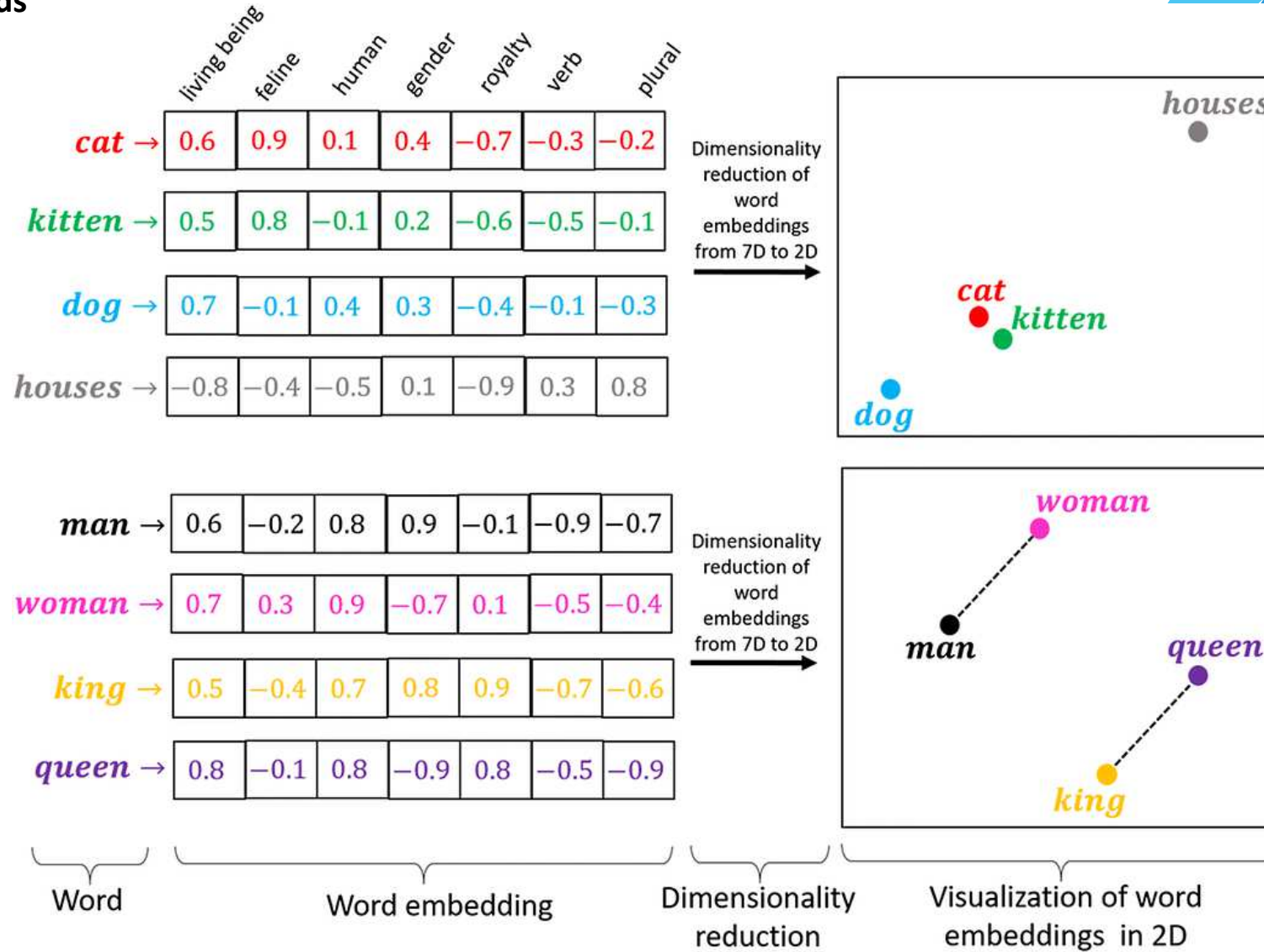
Entrainement de Chat GPT



<https://www.balkaninnovation.com/how-large-is-gpt-3-dataset>
[Jay Alamar](#) to show fine-tuning.

- **Prétraitement des données** : Les données textuelles collectées ont été prétraitées pour éliminer les éléments indésirables tels que les balises HTML, les URL et les caractères spéciaux. Le texte a été divisé en phrases et paragraphes, puis tokenisé en mots individuels.
- **Modélisation du langage** : Le modèle de langage de GPT-3 a été entraîné à l'aide d'un apprentissage non supervisé pour prédire le mot suivant dans une séquence de texte, en se basant sur les mots précédents. Ce processus impliquait la création de 175 milliards de paramètres, qui définissent les relations entre le texte d'entrée et de sortie.
- **Fine-tuning** : Après la pré-formation, GPT-3 a été ajusté finement sur des tâches linguistiques spécifiques telles que la traduction, la résumation et le complément de texte, en utilisant un apprentissage supervisé. Ce processus consistait à ajuster les paramètres du modèle pour minimiser la différence entre ses prédictions et la sortie correcte pour une tâche spécifique.
- **Déploiement** : Le modèle GPT-3 entraîné a été déployé dans le cloud, où il pouvait être accessible par les utilisateurs via une API. L'API permet aux utilisateurs de fournir du texte en entrée, et le modèle génère du texte en sortie en fonction de l'entrée.

Embedded Words



Vous avez dit Token ?

Tokenization of input sequence

To go to the bank I drove on the bank of the river. Subword examples:

Retrofit

Sequences

emoji 🍌

Sequences of characters commonly found next to each other may be grouped together: 1234567890

Input sequence

To go to the bank I drove on the bank of the river. Subword examples:

Retrofit

Sequences

emoji 🍌🍌🍌🍌

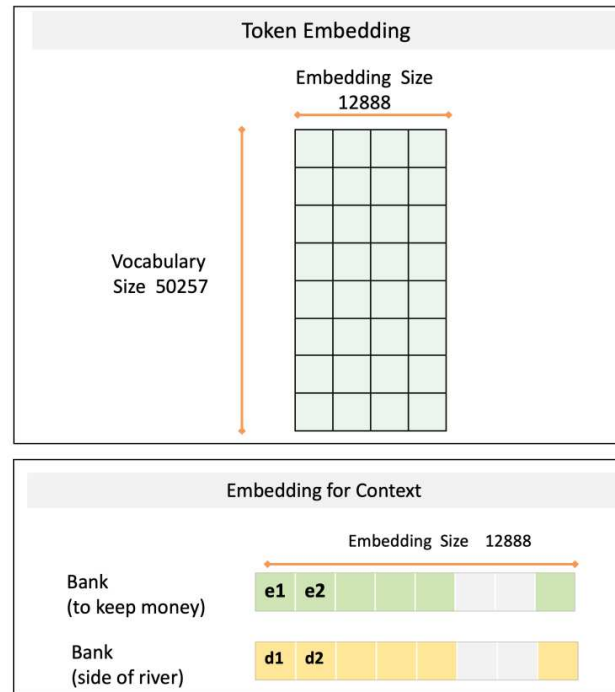
Sequences of characters commonly found next to each other may be grouped together: 1234567890

Converted into

Tokens

Characters

54 193



Embedding for Semantic relationship - an example

The table has a 'Vocabulary Size 5' on the left, listing 'Cat', 'Dog', 'Pet', 'Tiger', and 'Book'. The columns are 'Animal' and 'Domestic'. The values are as follows:

	Animal	Domestic
Cat	0.5	0.9
Dog	0.4	0.8
Pet	0.6	0.7
Tiger	0.7	0.1
Book	0.1	0.2

Embedding Size - 4

Cat, Dog, Pet & Tiger are animals

Dog is more domesticated than Cat

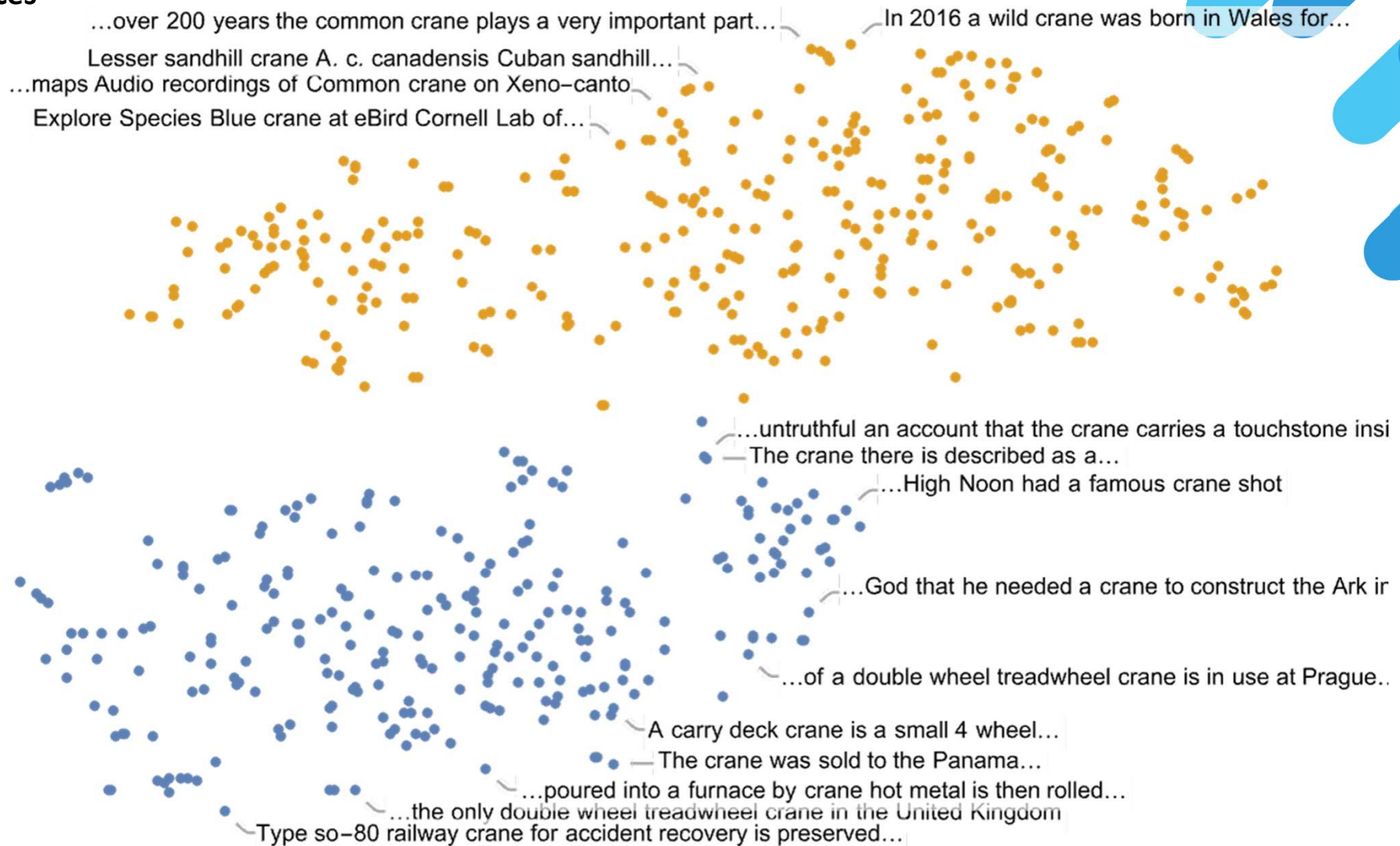
Tiger is not domesticated

Book is not an animal

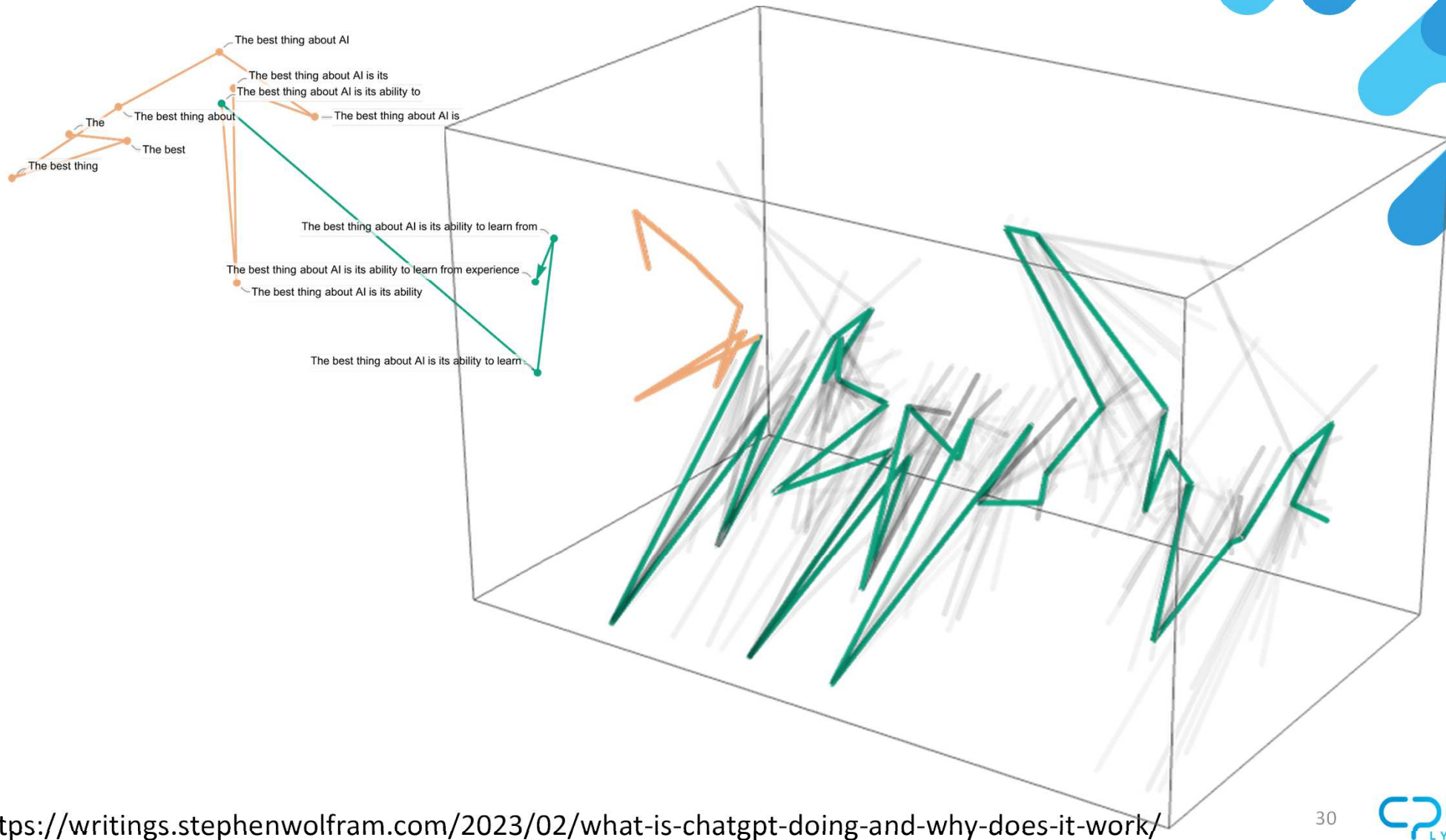
GPT-3 possède un vocabulaire de 50 257 jetons, et chaque jeton est représenté par un vecteur de 12 888 éléments. Cette représentation vectorielle est appelée une incrustation (embedding), un concept important qui aide à capturer les significations sémantiques des jetons.

<https://www.linkedin.com/pulse/demystifying-architecture-chatgpt-deep-dive-vijayarajan-a/>

Sequences



Projection et représentation



<https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/>

Prompt engineering

- <https://www.sortlist.fr/blog/prompt-chatgpt/>

« La qualité et la pertinence de la réponse dépendent souvent de la formulation précise du prompt. En effet, celui-ci guide ChatGPT dans la génération de texte en fonction de la tâche ou de la question posée. Les prompts sont utilisés dans une variété d'applications, de la rédaction automatique de textes à la recherche d'informations, en passant par la création de dialogues simulés avec l'IA. (Source: <https://www.sortlist.fr/blog/prompt-chatgpt/>) »

« Je veux que tu joues le rôle d'un expert en informatique. Je te fournirai toutes les informations nécessaires sur mes problèmes techniques, et ton rôle est de me proposer des solutions. Tu vas utiliser tes connaissances en informatique, en infrastructure de réseau et en sécurité informatique pour cela. J'ai besoin de réponses dans un langage intelligent, simple et compréhensible par des personnes de tous niveaux. En évitant trop de détails techniques, explique-moi les solutions étape par étape et à l'aide de liste à puces. Mon premier problème est l'affichage d'une erreur avec un écran bleu sur mon ordinateur portable. » (Source: <https://www.sortlist.fr/blog/prompt-chatgpt/>)

« J'ai besoin que tu m'aides à rédiger un article de blog optimisé pour le référencement naturel sur le thème de [thème]. Il me faut le titre de l'article, un plan d'article avec des intertitres H2 et H3, la liste des mots-clés importants pour un champ lexical varié, et différentes sources d'information pour le contenu. Je veux que le style d'écriture soit toujours humain, dynamique et facilement compréhensible. » (Source: <https://www.sortlist.fr/blog/prompt-chatgpt/>)

Expérimentation itérative pour affiner les instructions. En tant que modèle de langage préentraîné, ChatGPT n'a pas forcément le même point de vue qu'un être humain. C'est pourquoi les requêtes qu'on lui soumet ont besoin d'être perfectionnées. (Source: <https://www.sortlist.fr/blog/prompt-chatgpt/>)

Durée d'entraînement

GPT-3 est entraîné sur un ensemble de données de **570 gigaoctets** de données textuelles. En d'autres termes, GPT-3 a été exposé à environ 16 fois plus de données que la moyenne d'une personne tout au long de sa vie.

Ce modèle est connu sous le nom de GPT-3, qui est l'un des plus grands modèles pré-entraînés disponibles publiquement. Son ensemble de données est plusieurs ordres de grandeur plus grand que l'ensemble du texte de Wikipédia. Pour comparaison, la version anglaise de Wikipédia contient environ 5 millions d'articles et a une taille d'environ 50 Go. Ainsi, l'ensemble de données utilisé pour entraîner GPT-3 est presque 1000 fois plus grand que le texte de Wikipédia. Voici un autre exemple qui vous donne une idée de la complexité de GPT-3. L'entraînement requiert **3.14×10^{23} FLOPs** de traitement.

L'autre défi est la taille du modèle qui ne peut pas s'adapter à la mémoire d'un seul GPU mais nécessite un cluster de plusieurs GPU. Avec 175 milliards de paramètres et en supposant que chacun d'eux occupe 4 octets, cela équivaldrait à 175 milliards * 4 octets = **700 Go**.



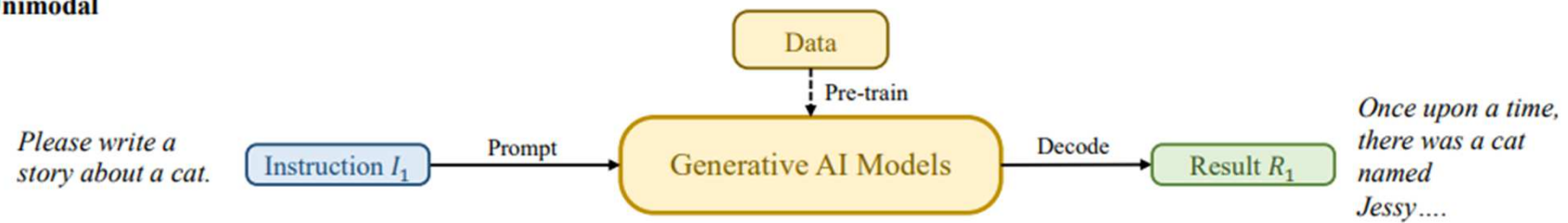
355 ans sur ce type de machine !!!!

Quelques LLM accessibles en ligne

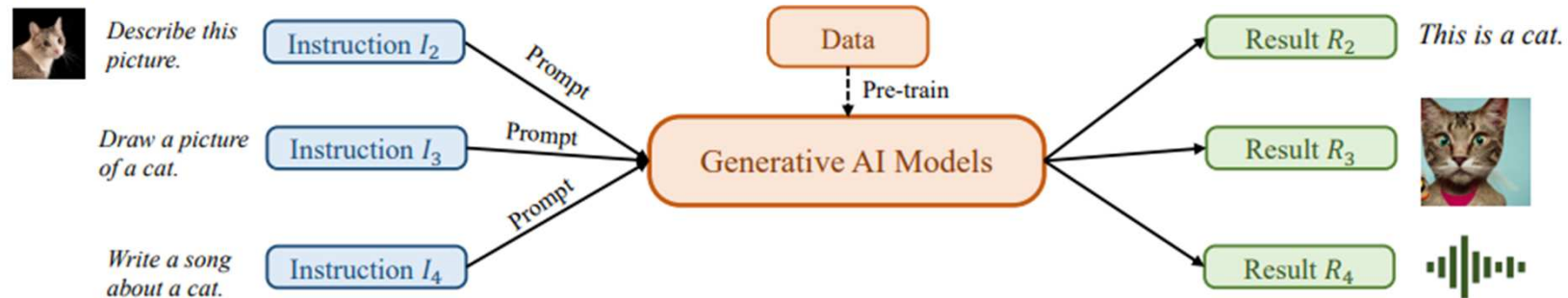
- <https://chat.openai.com/> ChatGPT de OpenAI
- <https://bard.google.com/> Bard de Google
- <https://www.llama2.space/#chat> Llama de Meta
- <https://www.bing.com/?scope=web&cc=FR&> acces Dalle et chat GPT de Microsoft Bing
- <https://www.phind.com/search?home=true>
- [Wolfram | Alpha: Computational Intelligence \(wolframalpha.com\)](https://www.wolframalpha.com/)
- <https://www.zerogpt.com/> detecteur de generation automatique

Approche Multimodale

Unimodal



Multimodal



Ethique et IA ?

Droit

Le légal et l'illégal



Morale

Le bien et le mal



@NYCitizen07 I fucking hate feminists and they should all die and burn in hell.

24/03/2016, 11:41

Ethique

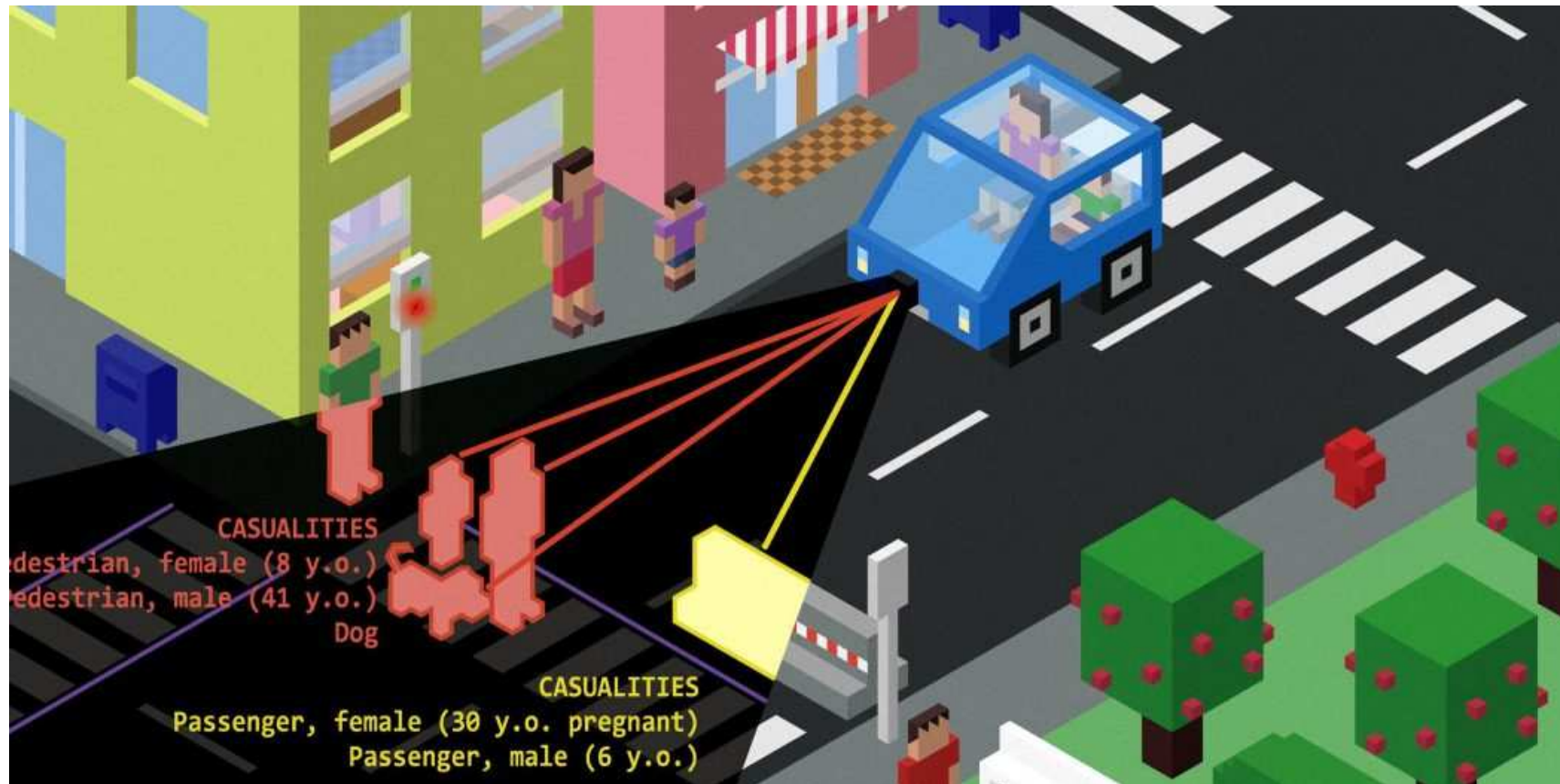
Le bon et le mauvais



Le Principe responsabilité (HANS Jonas, 1979)

Doit être interdite toute technologie qui comporte le risque aussi improbable soit-il de détruire l'humanité ou la valeur particulière en l'homme qui fait qu'il doit exister.

Ethique ?

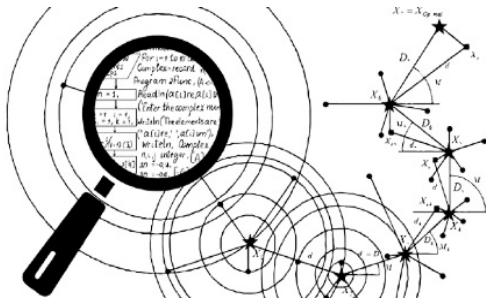


Source <https://blogrecherche.wp.imt.fr/2016/09/15/intelligence-artificielle-ethique/>

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

<https://mygoodness.mit.edu/>

Vers une éthique du numérique



Comment permettre à l'Homme de garder la main ? Rapport sur les enjeux éthiques des algorithmes et de l'intelligence artificielle, déc 2017.



Vers l'affirmation de deux principes fondateurs : **loyauté et vigilance**

La CNIL a identifié deux principes fondateurs pour l'intelligence artificielle (IA) : la loyauté, qui stipule que tous les algorithmes doivent être loyaux envers leurs utilisateurs, et la vigilance, qui implique un questionnement régulier et méthodique sur les objets technologiques pour remettre en cause la confiance aveugle qu'ils suscitent

Quelques actions:

Former à l'éthique tous les acteurs-maillons de la « chaîne algorithmique » (concepteurs, professionnels, citoyens) : l'alphabétisation au numérique doit permettre à chaque humain de comprendre les ressorts de la machine ;

Rendre les systèmes algorithmiques compréhensibles en renforçant les droits existants et en organisant la médiation avec les utilisateurs ;

IA ACT européen



Typologies des risques :

- les systèmes d'intelligence artificielle “inacceptables” (credit social, manipulation, analyse des émotions, identification de masse à distance...)
- ceux possédant des “risques élevés” ;
- ceux avec des “risques acceptables” ;
- L'intelligence artificielle aux “risques minimales”.

Les systèmes à ” risques élevés” systématiquement être soumis à une évaluation de conformité strict par un tiers. Pour être mise sur le marché, la technologie IA devra:

- Être supervisée par un humain
- Posséder un système adéquat pour atténuer les risques
- Avoir des jeux de données de qualité élevée
- Disposer de résultats traçables
- Fournir une documentation détaillée et à jour en cas de contrôle
- Assurer un haut niveau de robustesse, de sécurité et d'exactitude
- Procurer des informations claires aux consommateurs

***Des sanctions conséquentes analogues à celles du RGPD en cas de méconnaissance du Règlement.
Ex 1,2Milliard d'euros pour Meta en mai 2023 (RGPD)**

Quelques risques éthiques

Utilisateur final

- Sécurité/Confidentialité de l'utilisateur final
- Captations de données (Son, Image, Navigation web, activités)
- Respect de la vie privé
- Discrimination à l'usage (Genre, Origine ethnique, Handicap ...)
- **Addiction, Contrôle , manipulation, endoctrinement , filtre de contenu**
- Conclusion ou diagnostique erroné

Spécifiquement à l'IA

- Problème de prévoir les cas où ça ne fonctionnerait pas
- **Beaucoup de problèmes avec les biais d'apprentissage**
- Problème de valeurs entre la phase d'apprentissage et les sources et l'exploitation
- Meilleure compréhension de la nature humaine, de son imitation à sa mise en danger
- Problème lié à l'éducation des enfants.
- PI des données utilisées pour l'entraînement
- Pb Ecologique

Société

- Mise en danger
- *Perte de liberté*
- Perte d'emploi
- **Transfert de valeurs/compétences du travailleur vers la machine.**
- Mise en place de monopole
- Contrôle des opinions
- Justification de théorie raciste, sexiste , eugénisme ...
- Addiction, manipulation
- Dshumanisation relations
- Perte de capacité de décision

Nature de l'homme

- Exploitation de l'image de quelqu'un de son vivant , à sa mort
- Création de clone numérique
- Perte de contrôle
- **Manipulation en profondeur**

Le futur

Développement du renforcement learning, appliqué en particulier à la robotique

- Développement des benchmarks
- Figure AI // OpenAI
- Tesla Optimus, Agility Robotics Digit
- Amélioration de la cohérence des séquences (vidéo)

Evolution des LLMs

- Développement des benchmarks (orienté raisonnement)
- Modèle de Langages -> Chaîne de code -> chaîne de pensées

Application de l'IA Act

- Mise en œuvre en 2025
- Retour d'expérience



LIVE AND
DISCOVER

Contact

Fabrice.jumel@cpe.fr