

# **ia générative d'images**

**vers une mutation des pratiques**

P H - 21 NOVEMBRE 2024

# Introduction à l'ia générative

quels domaines sont concernés

- Images et Illustrations
- Textes et Littérature
- Musique
- Vidéo et animation
- Modèles 3D
- Voix et audio
- Code et logiciels
- Données et simulations
- Reconnaissance d'images et de sons

# Introduction à l'ia générative

quels domaines sont concernés

## IA Faible



### Purement réactive

- › Pas de mémoire, pas d'expérience, perception du monde à un instant donné.



### Mémoire limitée

- › Mémoire, expérience du passé, suivi de l'évolution de la représentation du monde.

## IA Forte



### Représentation de l'esprit

- › Comprend la représentation du monde et a conscience des objets s'y trouvant.



### Conscience propre

- › Construit sa propre représentation du monde. Conscience de soi.

# La réponse la plus probable

VOTRE DEMANDE → RÉPONSE

- *Il cherche la réponse la plus probable, pas la plus vraie.*
- *Il ne réfléchit pas, il associe les mots*  
= *prédiction*

De quelle couleur est le ciel ?

La couleur du ciel est bleue



**BLEU**









**Cette image dégage une atmosphère chaleureuse et familiale tout en capturant un moment de la vie quotidienne. Voici une description détaillée sous différents angles :**

**L'image présente une scène familiale en extérieur, probablement lors d'une promenade en pleine nature, dans une lumière douce et diffuse. La brume en arrière-plan donne une touche éthérée, presque magique, à l'image, créant une séparation naturelle entre le premier plan où se trouvent les personnes, et l'arrière-plan aux tons automnaux. Les couleurs dominantes sont les tons chauds et terreux – marron, jaune moutarde, et vert olive – qui se marient bien avec l'environnement automnal. La composition est équilibrée, avec le groupe de personnes occupant le centre de l'image, les lignes horizontales de l'horizon et du chemin guidant naturellement l'œil à travers la scène.**

- > Composition : L'image est bien équilibrée avec le groupe familial centré, mais légèrement décalé vers la droite, laissant un espace négatif sur la gauche qui renforce l'aspect naturel et spontané de la scène.**
- > Éclairage : La lumière semble naturelle, probablement un jour nuageux, ce qui crée une douceur sans ombre dure. La brume en arrière-plan diffuse davantage la lumière, renforçant cette atmosphère douce.**
- > Profondeur de champ : L'arrière-plan est légèrement flou, ce qui met davantage l'accent sur le groupe familial au premier plan, tout en conservant suffisamment de détails pour que l'on puisse identifier le paysage.**
- > Couleurs : Les tons chauds dominent, avec des touches de couleurs vives sur certains vêtements qui attirent le regard (par exemple, le jaune des pantalons et manteaux). Les teintes automnales des arbres en arrière-plan se fondent bien avec les vêtements des personnages, créant une harmonie visuelle.**
- > Mouvement : Il y a une certaine dynamique dans l'image, capturée par le mouvement des enfants, certains marchant, d'autres semblant interagir les uns avec les autres. Cela contraste avec la posture plus statique des adultes.**

**L'image exprime une certaine nostalgie et une tendresse particulière. On ressent une forte connexion familiale, avec les adultes (peut-être les grands-parents) entourant et veillant sur les enfants. Leurs expressions sont douces, et même si les visages ne sont pas entièrement visibles, on peut percevoir une certaine sérénité et bonheur dans ce moment partagé. Les enfants, par leur mouvement et leur posture, apportent une touche de vivacité et d'innocence à la scène. La brume et les couleurs automnales renforcent ce sentiment de moment privilégié, comme un souvenir chéri capturé dans le temps.**







# Historique

l'ai, domaine récent ?







**1952**

**Arthur Samuel**

est le pionnier de l'apprentissage automatique (**machine learning**), il est reconnu pour avoir créé le premier programme informatique permettant de jouer aux dames de niveau championnat.



**1957**

**Frank Rosenblatt**

un psychologue connu pour ses travaux sur le machine learning, développe le **perceptron**, l'un des premiers algorithmes à utiliser des réseaux de neurones artificiels. L'objectif du **Perceptron** était d'apprendre des données en ajustant ses paramètres jusqu'à atteindre une solution optimale.



**1970 > 1990**

**L'hiver de l'IA**

Pendant cette période, le financement de la recherche s'est tari et de nombreux projets ont été abandonnés en raison de leur manque de succès.



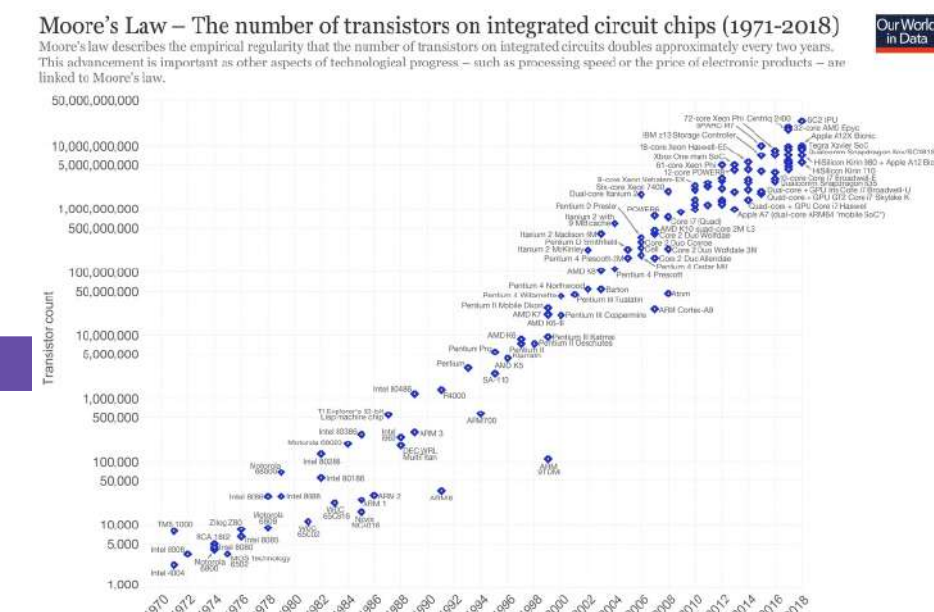
**1997**

le supercalculateur **IBM Deep Blue** bat le grand maître d'échecs **Garry Kasparov** lors d'un match.

C'était la première fois qu'une machine battait un joueur d'échecs expert. Il s'agit d'un événement marquant car il a montré que les systèmes d'IA pouvaient dépasser la compréhension humaine dans des tâches complexes.

Alan Turing :  
Test - Imitation game

Marvin Minsky &  
John Mc Carthy  
(1956) - colloque  
aux USA



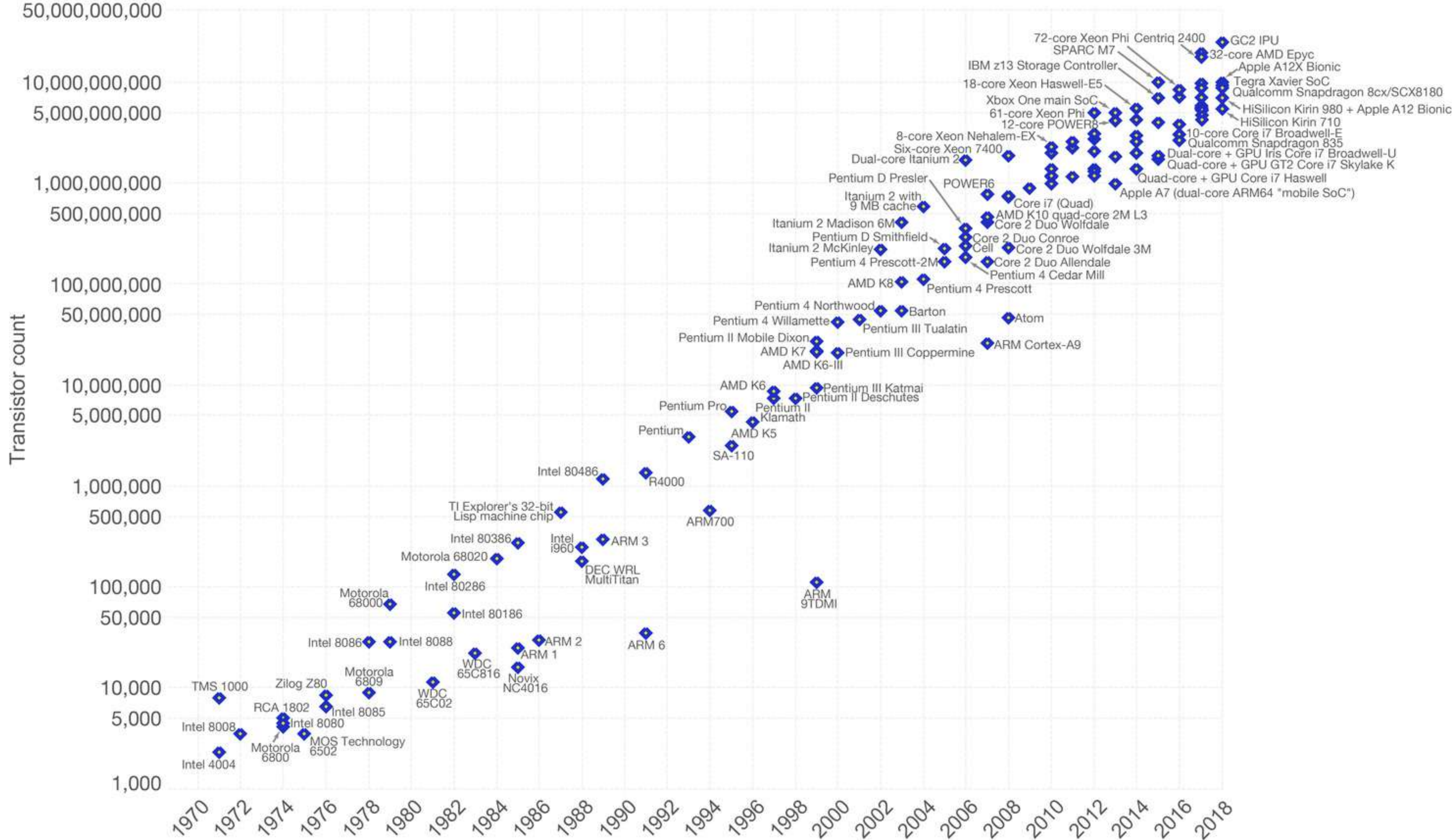
Manque de données



# Moore's Law – The number of transistors on integrated circuit chips (1971-2018)



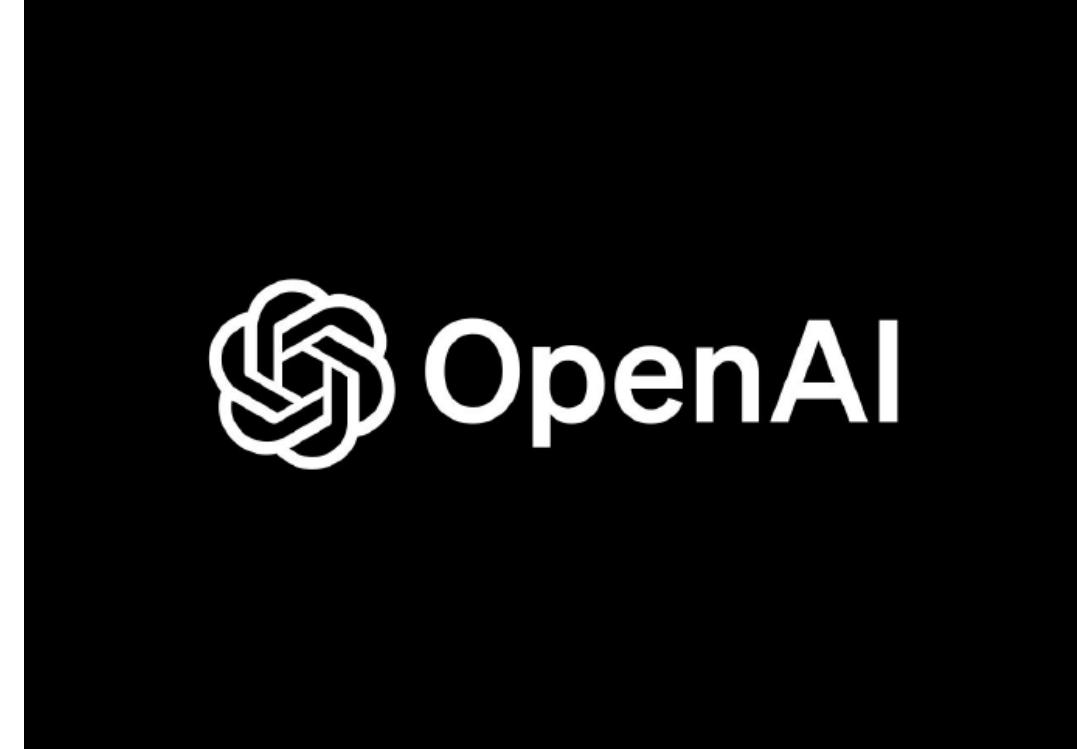
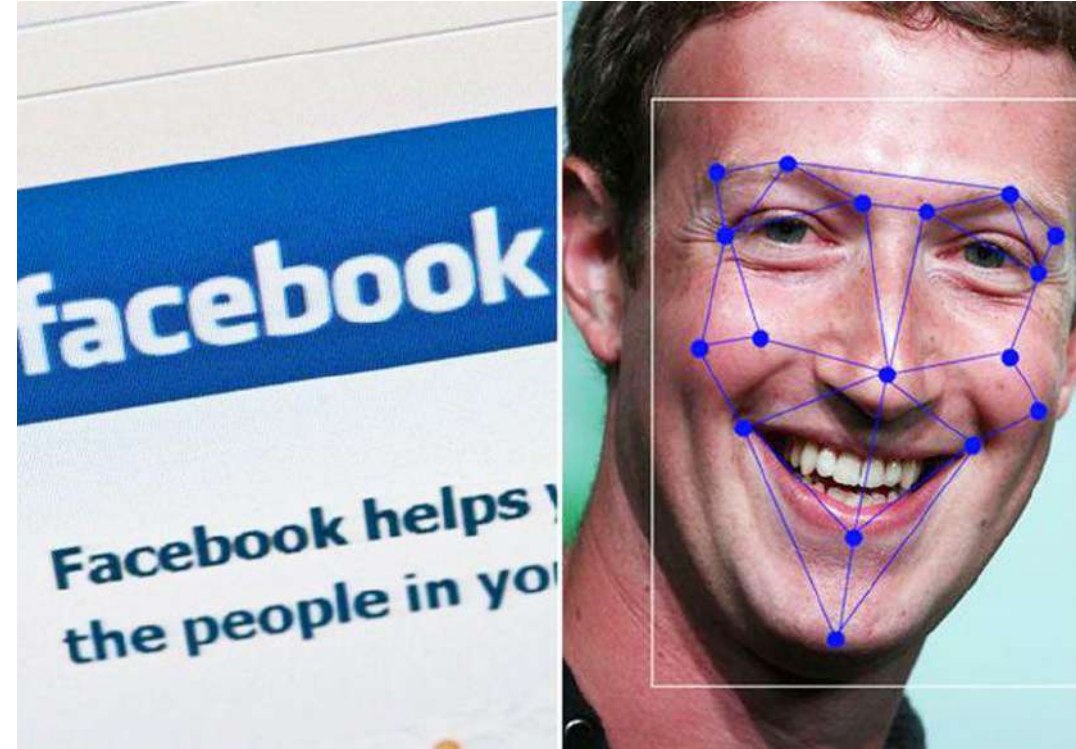
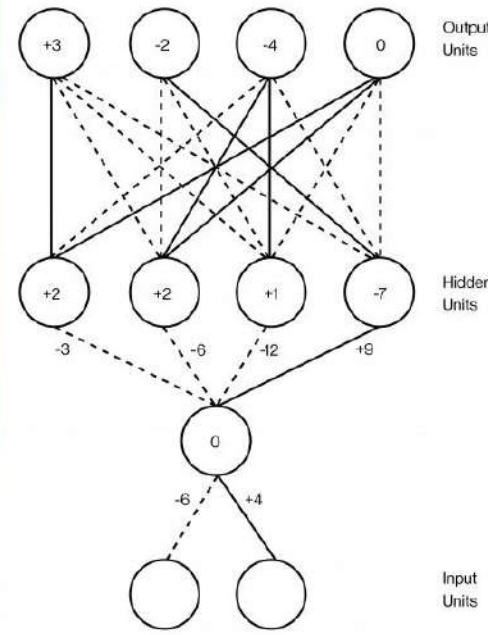
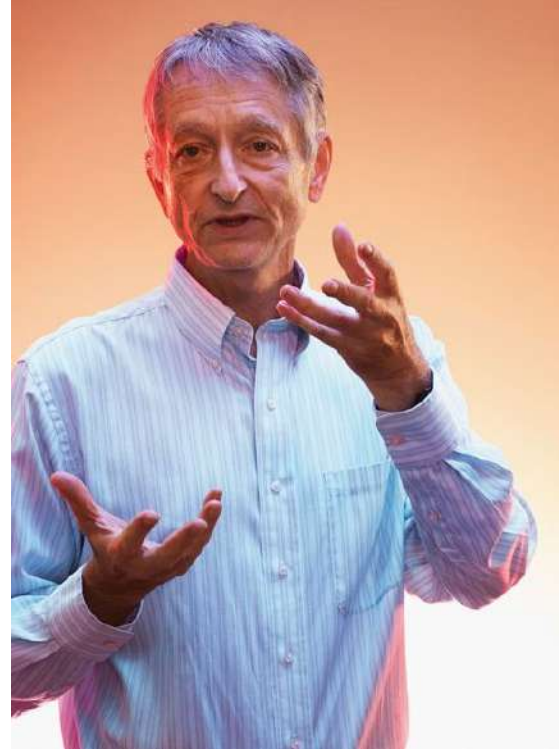
Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important as other aspects of technological progress – such as processing speed or the price of electronic products – are linked to Moore's law.



Data source: Wikipedia ([https://en.wikipedia.org/wiki/Transistor\\_count](https://en.wikipedia.org/wiki/Transistor_count))  
 The data visualization is available at [OurWorldinData.org](https://www.ourworldindata.org). There you find more visualizations and research on this topic.

Licensed under CC-BY-SA by the author Max Roser.





### 2006

#### Geoffrey Hinton

publie "A Fast Learning Algorithm for Deep Belief Nets." . Cet article décrit le premier algorithme de **Deep learning** capable d'atteindre des performances de niveau humain sur des tâches de reconnaissance de formes difficiles et complexes.

spécialiste des neurones artificiels

### 2014

#### DeepFace

permet à Facebook de "reconnaître" les visages sur les photos téléchargées par les utilisateurs. Ce système de reconnaissance faciale est tellement avancé qu'il peut identifier des visages avec une précision presque comparable à celle d'un humain.

Eugène : première machine à réussir le test de Turing

### 2017

un cap est franchi lors de **L'ImageNet Challenge**, un concours de vision par ordinateur organisé depuis 2010 : 29 équipes sur 38 ont atteint une précision de 95 % avec leurs modèles de vision par ordinateur. L'amélioration de la reconnaissance d'images est immense.

2015 : Elon Musk + Sam Altman (open ai)

### 2018

**OpenAI** présente **GPT-1** (Generative Pre-trained Transformer) entraîné sur 40 Gigabytes de données et consistant en 117 millions de paramètres, GPT ouvre la voie aux prochains modèles de langages de grandes tailles (LLM) en génération de contenu, chatbots et traductions.

2010 : ère du Big data les données sont disponibles







"A 1980 photograph of a smiling Chinese crested dog wearing sunglasses and drinking cocktails by the pool of a motel in Las Vegas, a slice of birthday cake next to him"

Emi • DALL-E  
Human GAI

## Stable Diffusion Public Release



2021

### DALL-E

développée par OpenAI, cette IA permet de générer des images à partir de **prompts**.

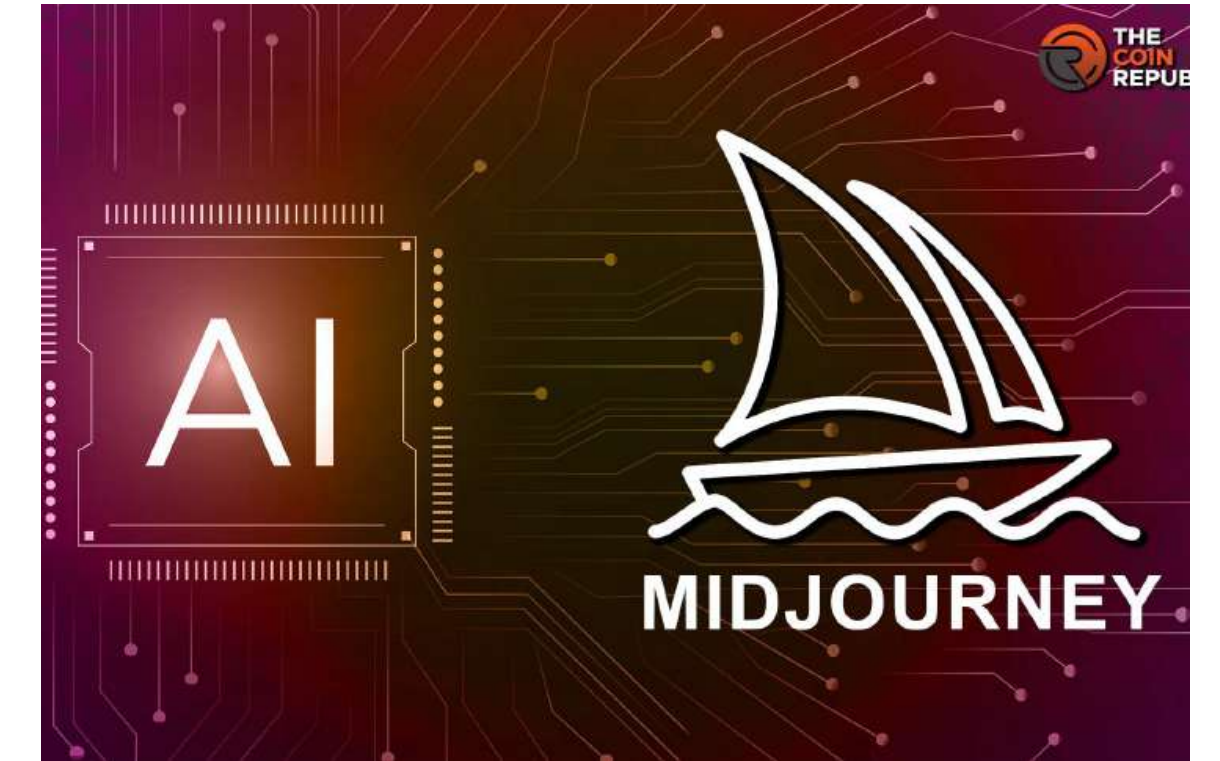
2022

**Stable Diffusion** développée par des chercheurs de Runway research, Stability AI et CompVis LMU, cette IA open source permet de générer des images à partir de prompts.



### Chat GPT

OpenAI attire 100 millions d'utilisateurs en 2 mois avec son chat basé sur la version 3.5 de son GPT LLM (Generative Pre-trained Transformer Large language model)

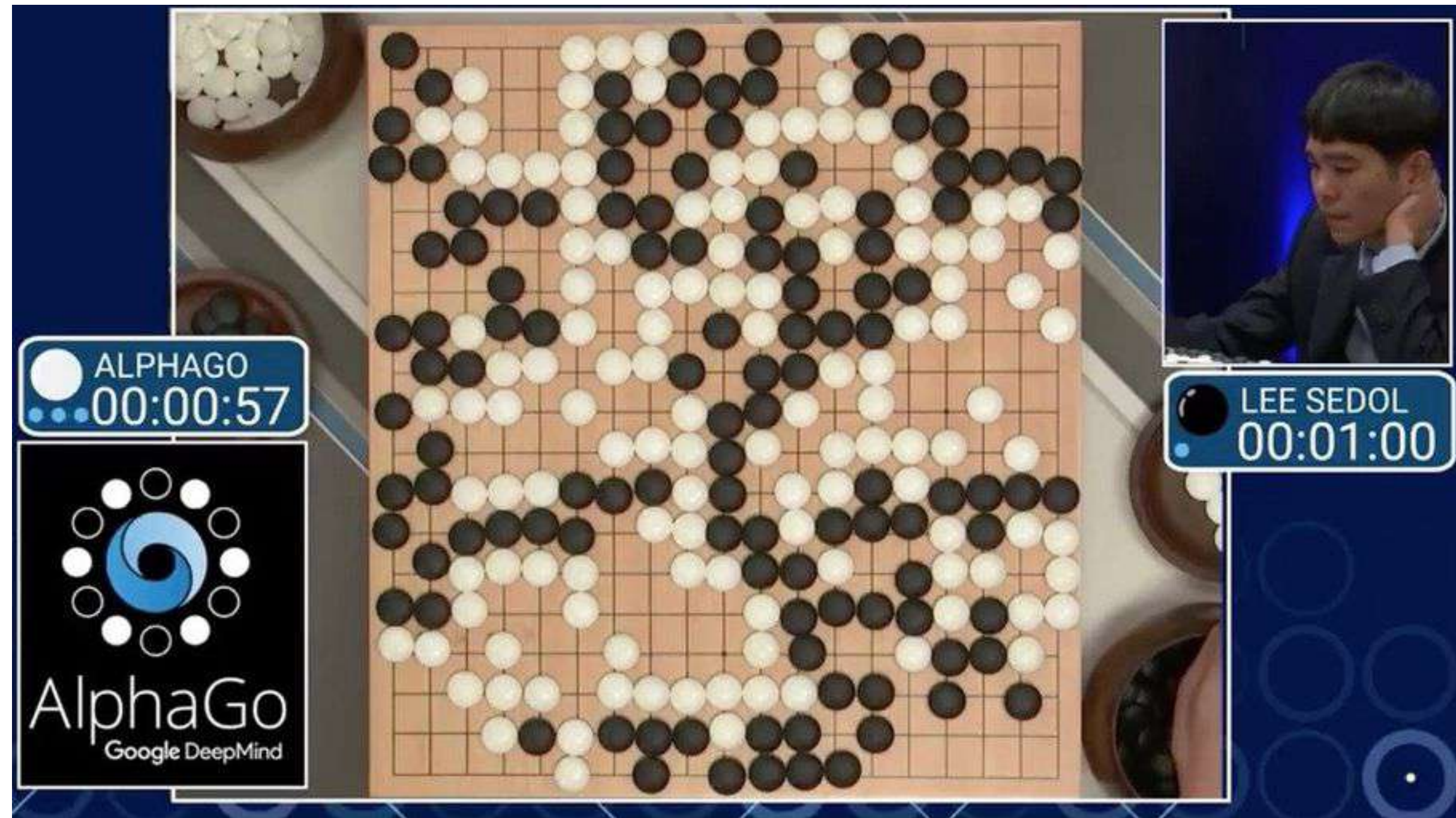


### Midjourney

sort en version bêta limitée le 21 mars 2022, et bêta ouverte **le 12 juillet 2022**. En juillet 2022, le laboratoire de recherche a commencé à tester la troisième itération (V3) de ses algorithmes de génération d'images..







# Et le jeu de GO ?

AlphaGo, une créativité algorithmique ?



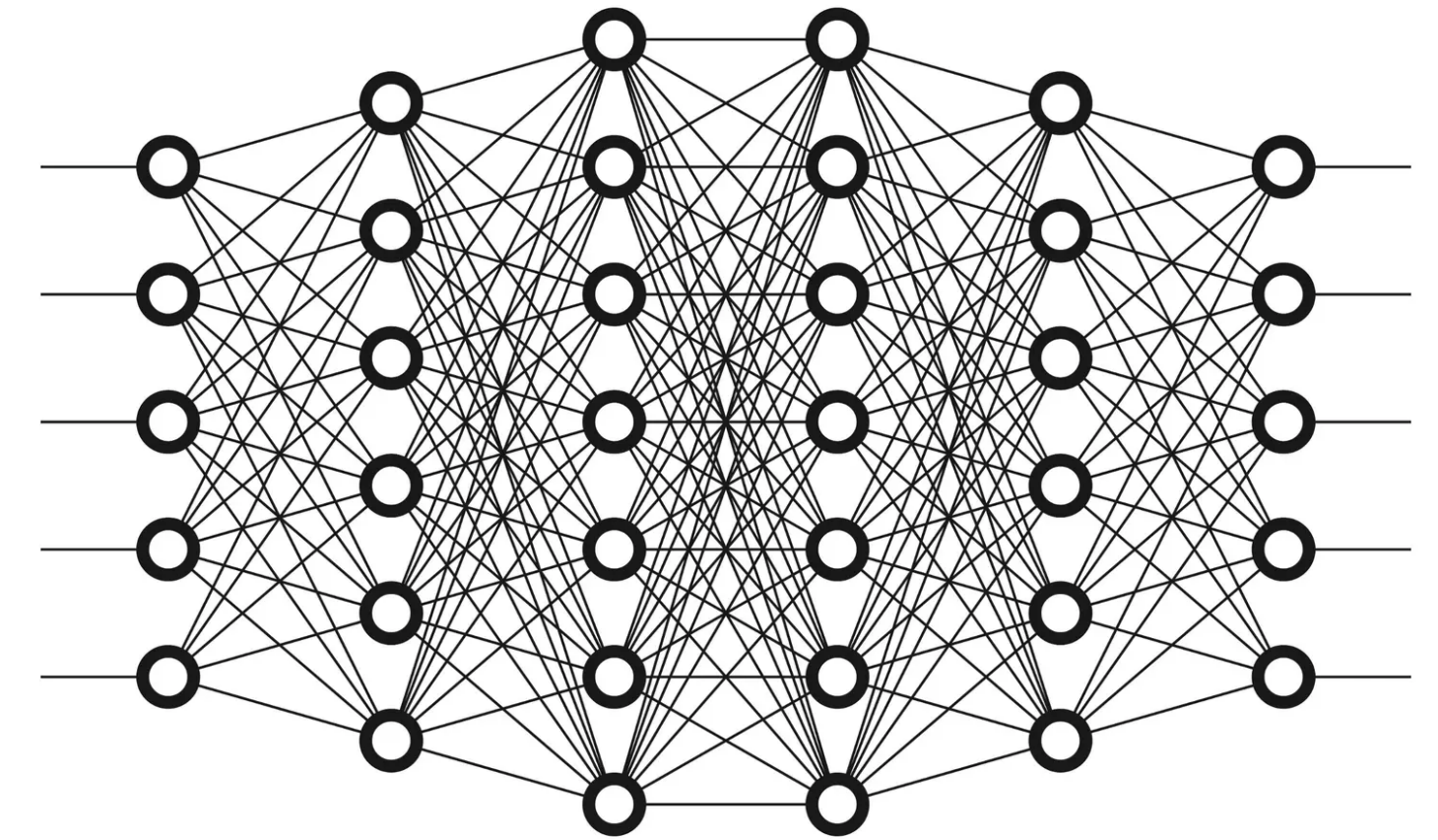
# Quelques concepts clefs

sinon on s'y perd

ai faible ou forte

Réseaux de neurones artificiels (ANN)

Machine learning



- Apprentissage profond (deep learning+modèles de données)
- Apprentissage supervisé (tâches spécifiques)
- Apprentissage non supervisé (analyse exploratoire)



# IA génératives d'images

quels outils

DALL·E - MidJourney - Stable Diffusion - Firefly - Craiyon -  
DeepArt - NightCafe Creator

Nom de l'Outil	Éditeur	Modèle d'IA	Payant
DALL·E 3	OpenAI	Transformer (GPT-3 basé)	Payant (avec options gratuites)
Midjourney	Midjourney, Inc.	Modèle propriétaire	Payant (avec essai gratuit limité)
Stable Diffusion	Stability AI	Diffusion latente	Gratuit (open-source)
Adobe Firefly	Adobe	Modèle propriétaire	Payant (avec options gratuites)
Craiyon	Boris Dayma	Modèle propriétaire	Gratuit
DeepArt	DeepArt AG	Réseaux convolutifs	Payant (avec options gratuites)
NightCafe Creator	NightCafe Studio	Diffusion latente, VQGAN+CLIP	Payant (avec options gratuites)



# Le prompt engineering

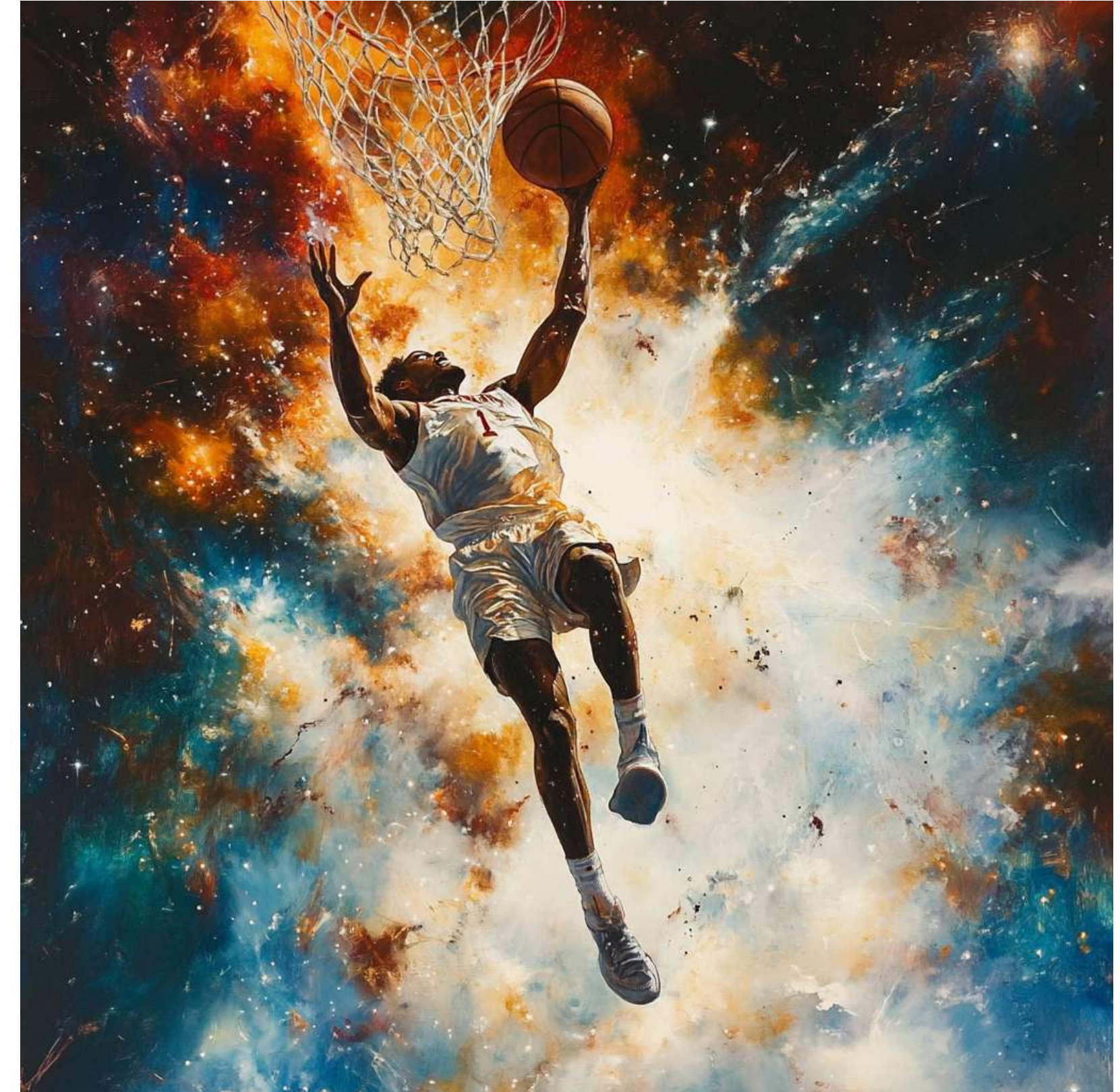
des prompts identiques, des images différentes



Firefly



Dall·e



MidJourney

An expressive oil painting of basketball player dunking, depicted as an explosion of a nebula



# Le prompt engineering

des prompts identiques, des images différentes DANS LE MEME OUTIL (FF)



An expressive oil painting of basketball player dunking, depicted as an explosion of a nebula



# Le prompt engineering

des prompts identiques, des images différentes DANS LE MEME OUTIL (D•E)



An expressive oil painting of basketball player dunking, depicted as an explosion of a nebula



# Le prompt engineering

des prompts identiques, des images différentes DANS LE MEME OUTIL (MJ)



An expressive oil painting of basketball player dunking, depicted as an explosion of a nebula



# Le prompt engineering

des prompts identiques, des images différentes et des versions différentes (MJ)

**V1**



*Feb 2022*

**V2**



*April 2022*

**V3**



*July 2022*

**V4**



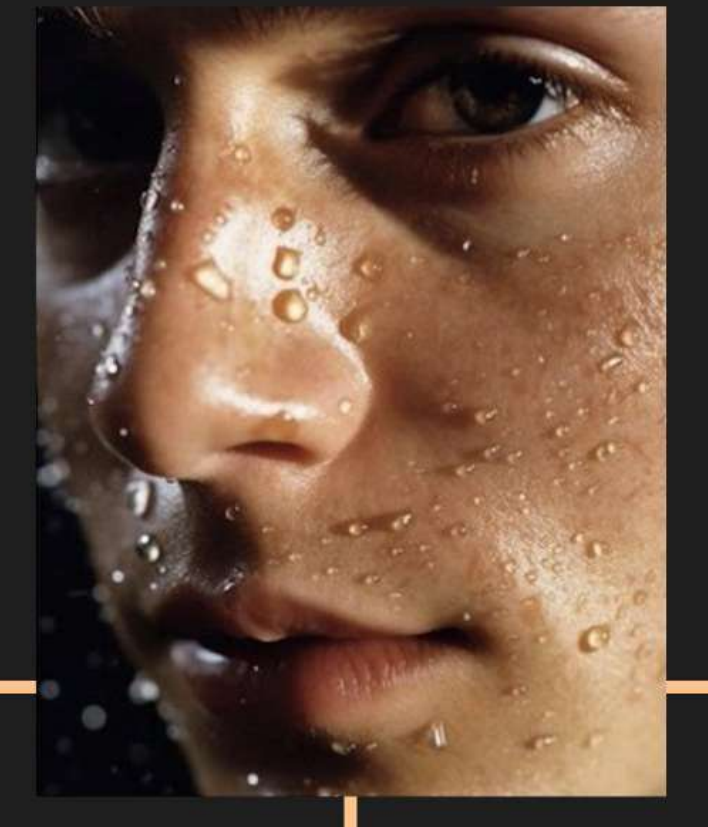
*Nov 2022*

**V5**



*March 2023*

**V5.1**



*May 2023*



# Le prompt engineering

## conseils

- Style d'image et/ou références (photo, dessin, gravure, peinture, cyberpunk...) -
  - État émotionnel, ambiance, atmosphère, climat général -
- Angle de vue, perspective et cadrage (contre plongée, plan moyen...) -
  - Objectif, appareil photo -
- Description du sujet et de l'arrière plan (mots clefs spécifiques) -
  - Ajouter du mouvement -
  - Choix de la lumière et/ou couleurs -
- Concept (limiter le nombre, trop complexe à gérer) -
  - Ponctuation et syntaxe -
  - Références ou style -

**TOUJOURS DEMANDER CE QU'ON VEUT (et pas ce que l'on ne veut pas)**



# Test sur MidJourney

prompts spécifiques

1 - un ordre spécifique

(sujet, médium, environnement, éclairage, couleur, humeur, composition)

2 - mots distincts et précis

3 - la ponctuation est primordiale

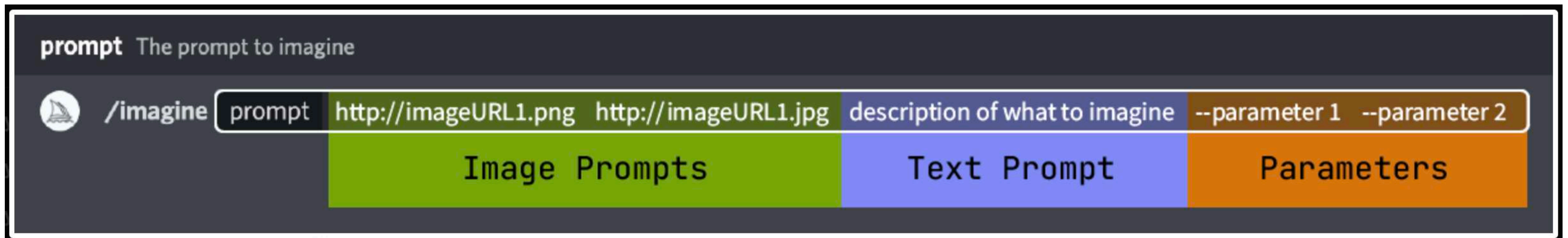
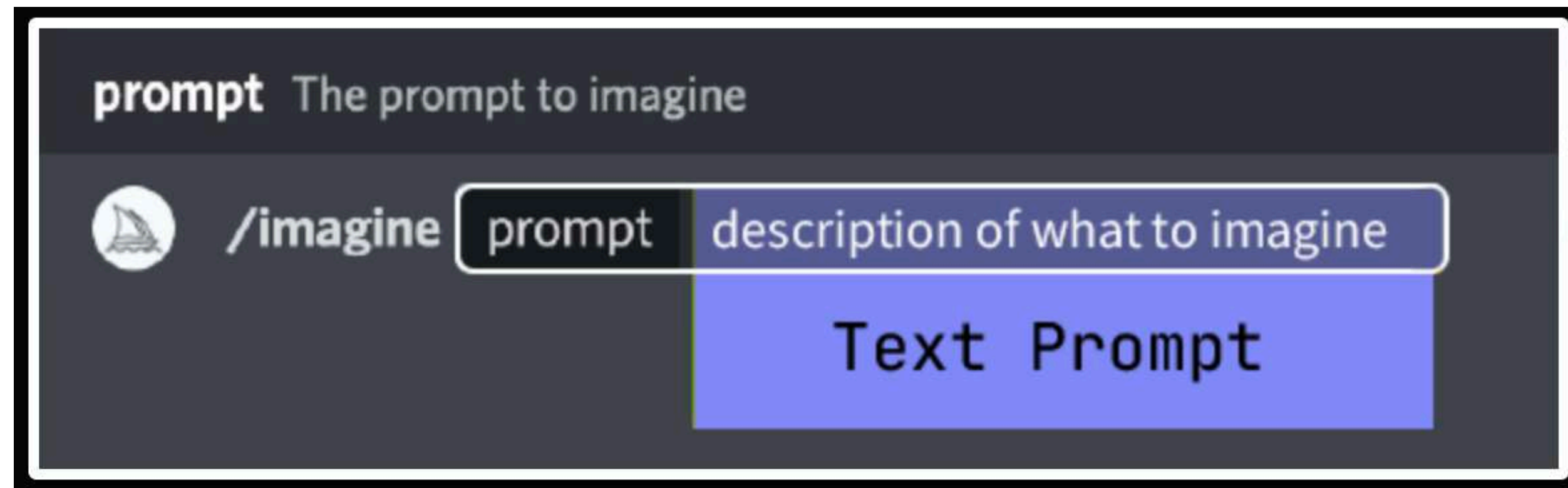
4 - les majuscules sont inutiles

**Il NE comprend PAS la grammaire et les structures de phrases**



# Test sur MidJourney

les commandes spéciales et paramètres





# Test sur MidJourney

## les commandes spéciales et paramètres

/imagine	-- no = éviter	-- iw (poids de l'image 0>2)
/blend (mélange)	-- aspect = ratio 4/3	-- ar (4:3 - 16:9)
/describe (décrire)	-- chaos 0>100= aléatoire	-- niji5 (manga)
/tune (différents styles)	-- seed = numéro d'image	-- w (weird - 0>3000)
/relax + /fast + /turbo	-- tile = motif	-- c (chaos)
/stealth + /public (privé-publique)	-- v (version 1 à 6.1)	:: (parties séparées)
	-- video = image par image	::2 (hiérarchie (poids positif ou négatif ::-2))



« A digital art of a large landscape, trees::100, mountains::1 »

Vous donnera une série d'image davantage orienté sur le terme « Trees »



« A digital art of a large landscape, trees::1, mountains::100 »

Vous donnera une série d'image davantage orienté sur le terme « Mountains »



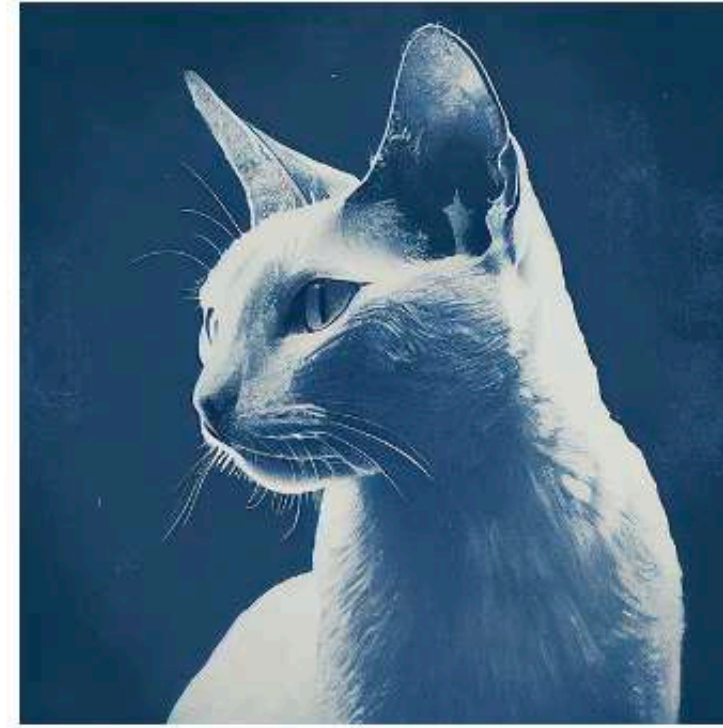
**Block Print**



**Folk Art**



**Cyanotype**



**Graffiti**



**Paint-by-Numbers**



**Risograph**



**Ukiyo-e**



**Pencil Sketch**



**Watercolor**



**Pixel Art**



**Blacklight Painting**



**Cross Stitch**

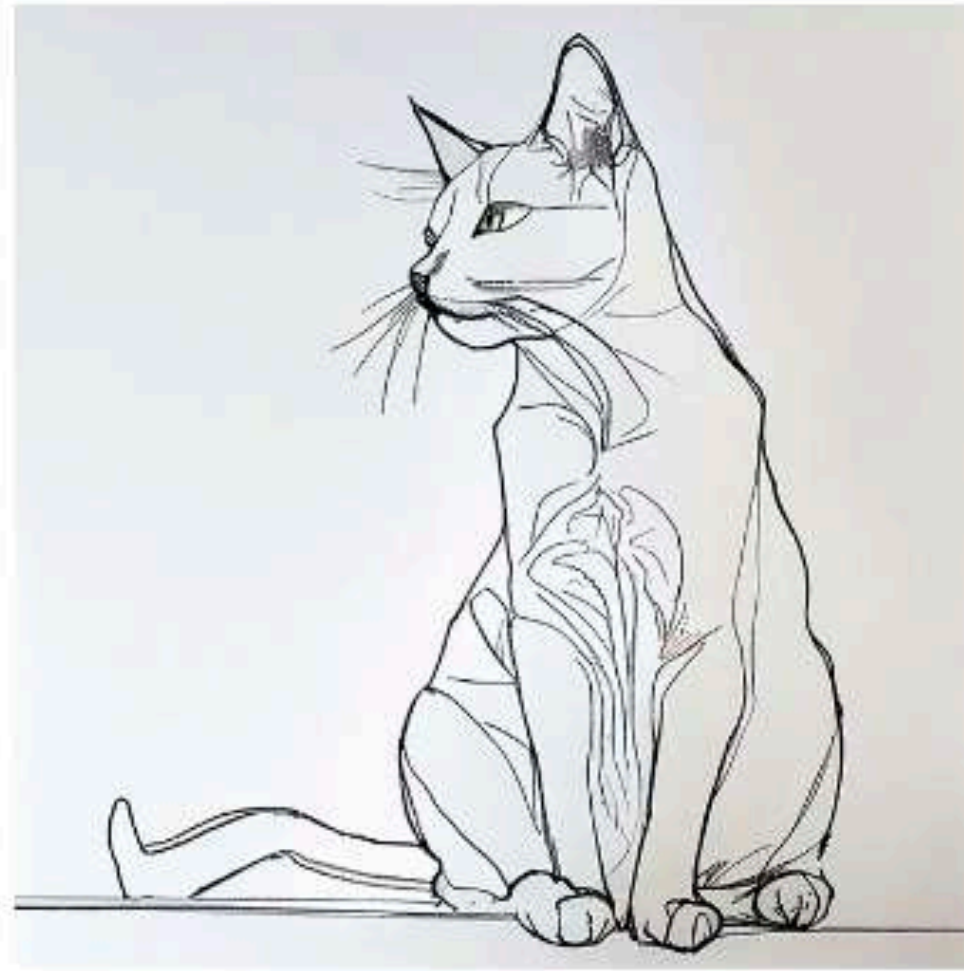




**Life Drawing**



**Continuous Line**



**Loose Gestural**



**Blind Contour**



**Value Study**



**Charcoal Sketch**





1700s



1800s



1900s



1910s



1920s



1930s



1940s



1950s



1960s



1970s



1980s



1990s





**Determined**



**Happy**



**Sleepy**



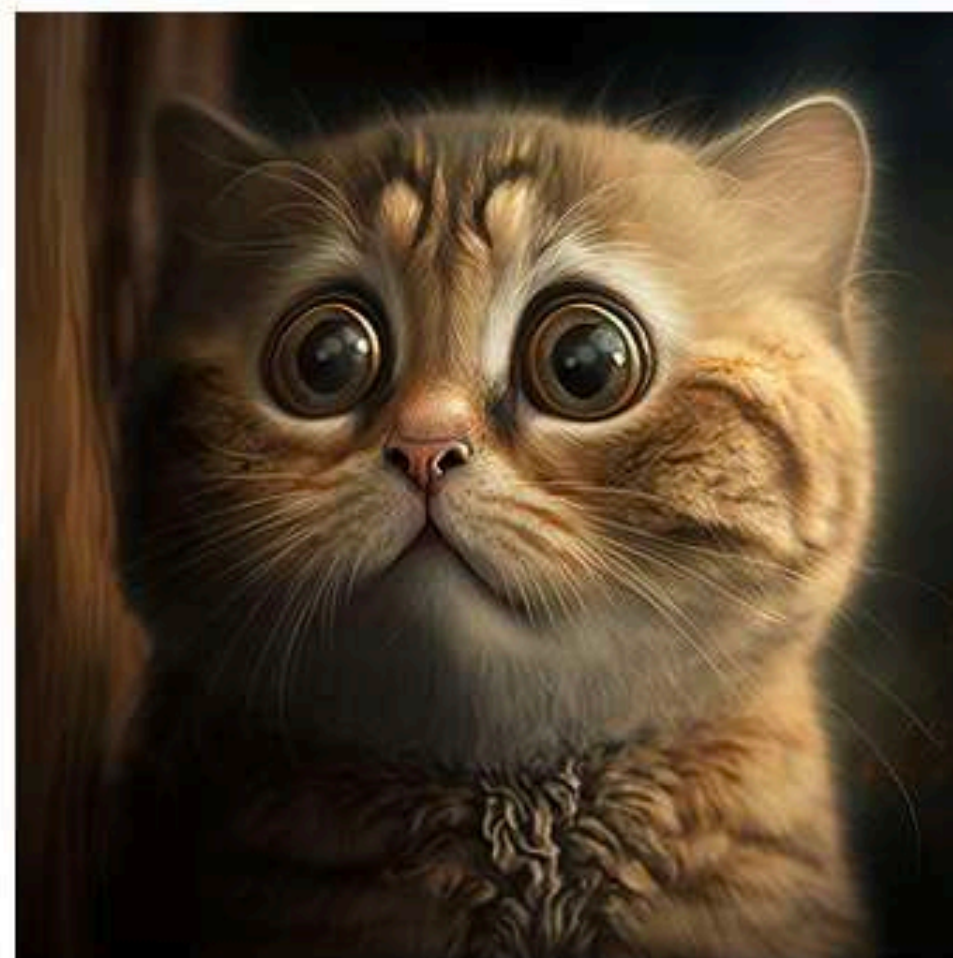
**Angry**



**Shy**

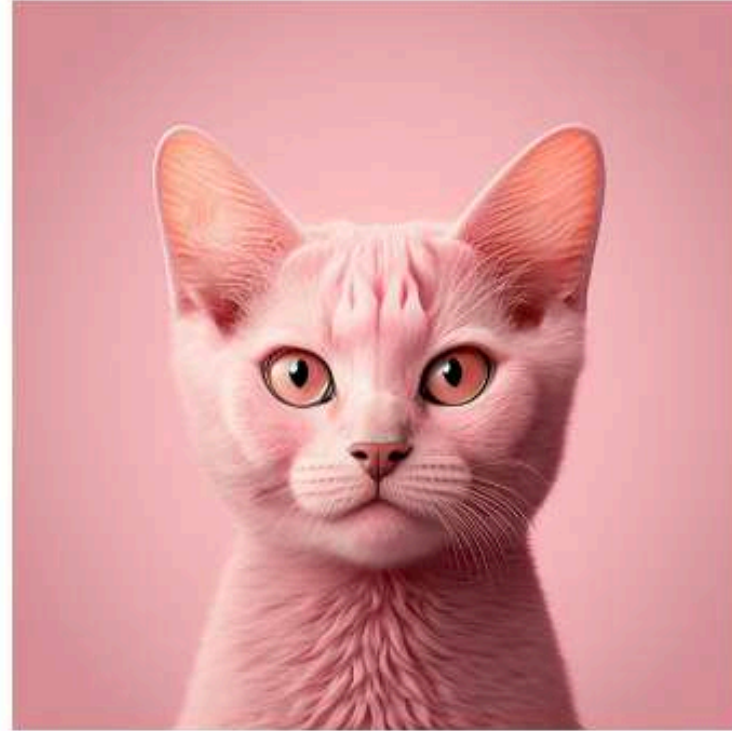


**Embarrassed**





**Millennial Pink**



**Acid Green**



**Desaturated**



**Canary Yellow**



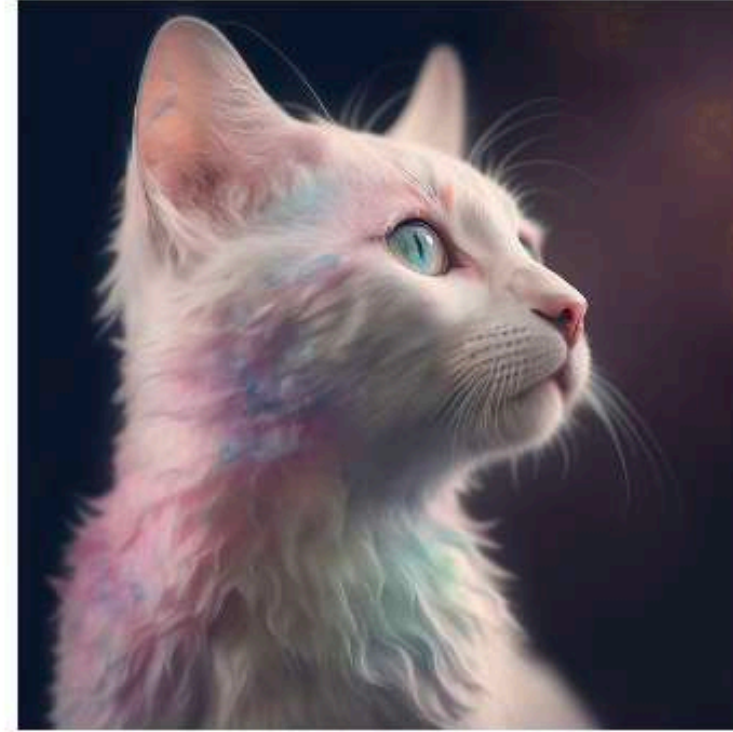
**Peach**



**Two Toned**



**Pastel**



**Mauve**



**Ebony**



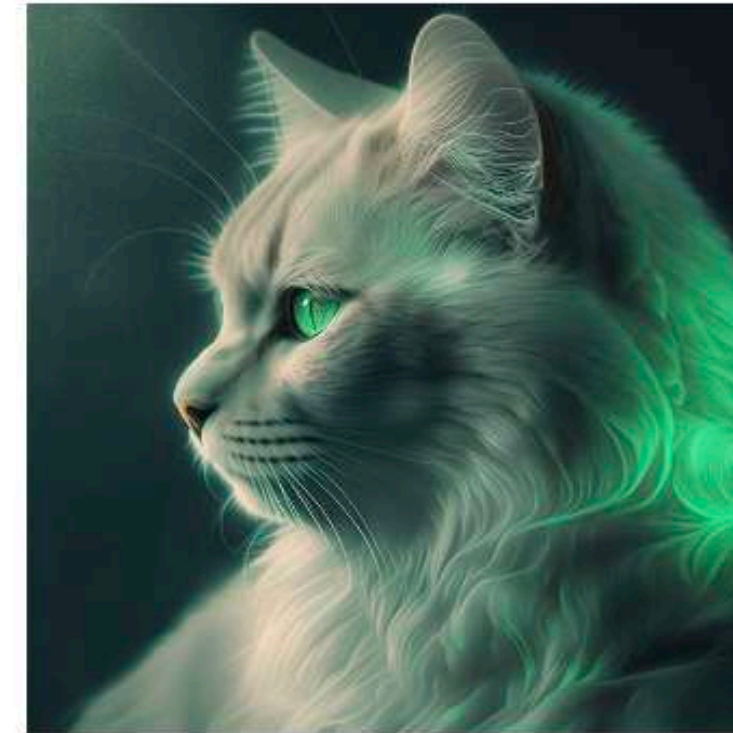
**Neutral**



**Day Glo**



**Green Tinted**





**Tundra**



**Salt Flat**



**Jungle**



**Desert**



**Mountain**



**Cloud Forest**





# Itération

An expressive oil painting of basketball player dunking, depicted as an explosion of a nebula

/imagine joueur de basket, homme, blanc, effectuant un dunk, peinture à l'huile, graffiti, nébuleuse, espace, couleurs vert et bleu, expressif, dynamique, plan moyen





# Itération

/imagineWhite male::4 basketball player dunking, oil painting, graffiti, nebula, space, green and blue colors, expressive, dynamic, medium shot, --no ball





# Bibliothèques de prompts

les communautés partagent

- MidJourney
- <https://midlibrary.io/> - bibliothèque de styles
- <https://promptlibrary.org/>
- [https://www.freemidjourneyprompt.com/#google\\_vignette](https://www.freemidjourneyprompt.com/#google_vignette)
- <https://promptlibrary.org/>
- DALL·E
- <https://www.promptpal.net/>
- <https://dallery.gallery/wp-content/uploads/2022/07/The-DALL·E-2-prompt-book-v1.02.pdf>
- <https://ai-art-tutorials.com/best-dalle-3-art-styles-ai-image-prompts-examples-12>



# Art, création et ia

ce qu'il faut retenir

1 : technique intrinsèque et technique extrinsèque

2 : l'incohérence (raison, émotion et intuition)

3 : le dialogue

4 : le fond et la forme

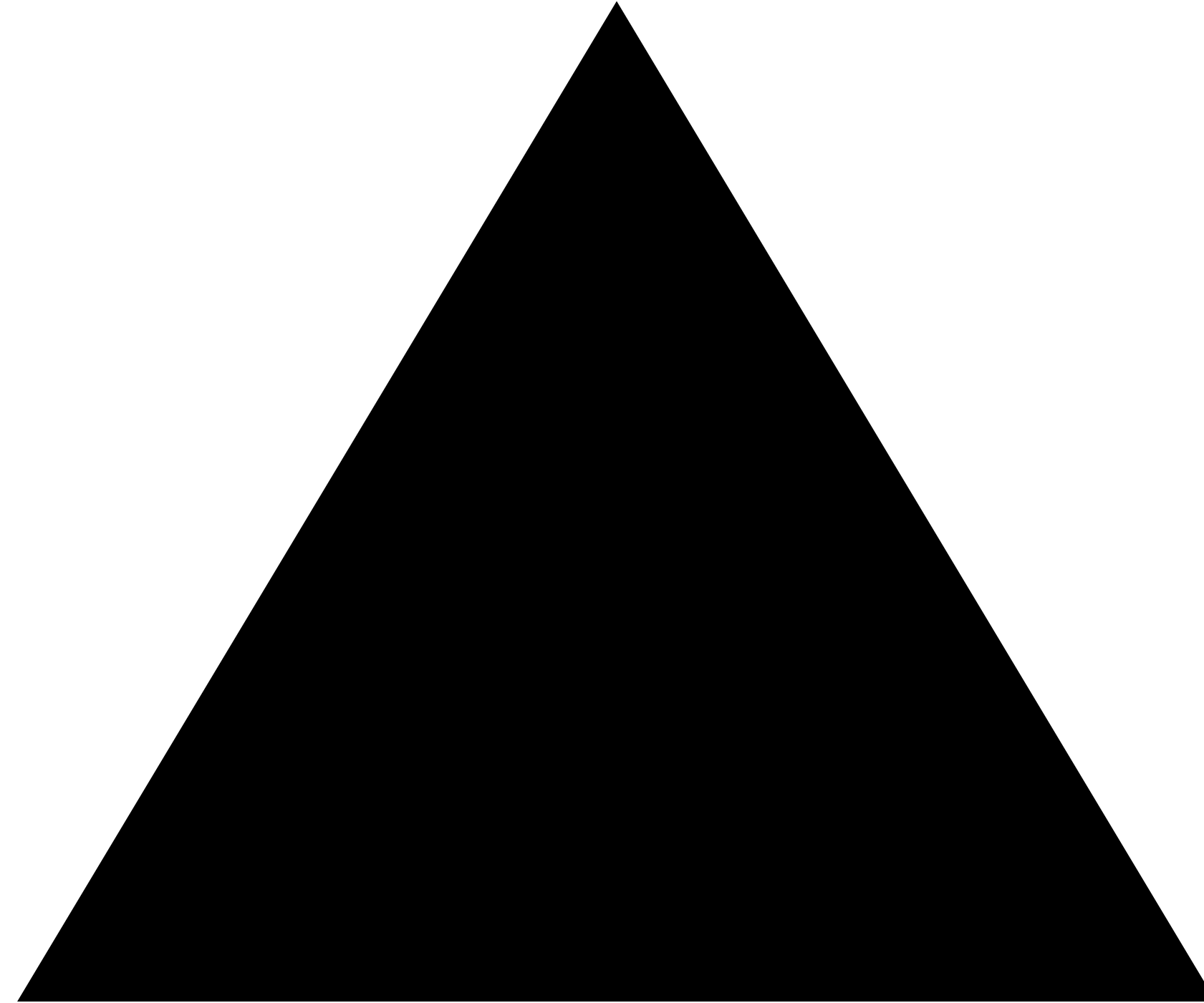
5 : le choix - la résilience

authenticité, sincérité, contexte, compréhension du monde,

ambiguïté, émotions, esprit critique...



**Technique**



**Narration  
Concept**



# Droits d'auteurs et plagiat

## un imbroglio juridique



<https://suno.com/song/a5d096df-98f7-4ba8-993b-b1696134e4c3>





# Droits d'auteurs et plagiat

## à la manière de

Art Movement	Major Artists	Midjourney Prompt
<b>Cubism</b>	['Pablo Picasso', 'Georges Braque', 'Juan Gris', 'Fernand Léger', 'Robert Delaunay']	Fragmented forms, geometric shapes, multiple perspectives, monochromatic palette, abstract objects, modern style, Cubism art, 20th century
<b>Surrealism</b>	['Salvador Dalí', 'René Magritte', 'Max Ernst', 'André Breton', 'Joan Miró']	Dreamlike scenes, surreal imagery, bizarre combinations, subconscious expression, illogical art, surrealism, abstract forms, 20th century
<b>Abstract Expressionism</b>	['Jackson Pollock', 'Mark Rothko', 'Willem de Kooning', 'Franz Kline', 'Barnett Newman']	Action painting, large canvas, emotional intensity, non-representational art, abstract expressionism, vivid colors, gestural brushstrokes, 20th century
<b>Dada</b>	['Marcel Duchamp', 'Hugo Ball', 'Hans Arp', 'Man Ray', 'Raoul Hausmann']	Anti-art, absurd, readymade objects, satirical humor, Dada movement, random compositions, avant-garde, 20th century
<b>Futurism</b>	['Umberto Boccioni', 'Giacomo Balla', 'Filippo Tommaso Marinetti', 'Carlo Carrà', 'Luigi Russolo']	Dynamic movement, speed, modernity, futurism, bold shapes, technological imagery, industrial colors, 20th century
<b>Pop Art</b>	['Andy Warhol', 'Roy Lichtenstein', 'Claes Oldenburg', 'Richard Hamilton', 'James Rosenquist']	Bright colors, commercial imagery, pop art, consumer culture, irony, mass production, 20th century style, playful designs
<b>Bauhaus</b>	['Walter Gropius', 'László Moholy-Nagy', 'Marcel Breuer', 'Paul Klee', 'Wassily Kandinsky']	Functional design, simplicity, Bauhaus movement, geometric forms, modern materials, integration of art, 20th century architecture
<b>Minimalism</b>	['Donald Judd', 'Frank Stella', 'Agnes Martin', 'Dan Flavin', 'Carl Andre']	Minimalist design, geometric abstraction, focus on form, non-representational art, minimalism, simplicity, monochromatic, 20th century
<b>Expressionism</b>	['Edvard Munch', 'Egon Schiele', 'Wassily Kandinsky', 'Ernst Ludwig Kirchner', 'Franz Marc']	Distorted figures, bold colors, expressionist art, emotional impact, subjective perspective, exaggerated lines, 20th century
<b>Op Art</b>	['Victor Vasarely', 'Bridget Riley', 'Richard Anuszkiewicz', 'Yaacov Agam', 'Julian Stanczak']	Optical illusions, geometric patterns, high contrast, Op Art, visual trickery, precise lines, vibrant colors, 20th century





# Education et méthodologie

un raz-de-marée

- brainstormer
- associer des idées
- rebondir sur les résultats
- jouer avec l'aléatoire
- mixer des idées
- tester différents rendus
- pousser la recherche
- moodboards
- rassembler des références
- dossier de recherches
- découvrir des styles
- associer des styles