

Master IA2S (Intelligence Artificielle, Science des données et Systèmes Cyber-Physiques)

Apprentissage Automatique

Ferhat ATTAL

ferhat.attal@u-pec.fr

Université Paris Est Créteil (UPEC)

Octobre 2021

Cours 1 : Introduction à l'Intelligence Artificielle

Sommaire

- Un peu d'histoire
 - Dès l'Antiquité
 - A partir du 17e siècle
 - Début du 20e siècle
- La naissance de l'IA
- Les sous-domaines de l'IA
- Science de données
 - C'est quoi les données?
 - Les principaux types de données
 - Malédiction de la dimensionnalité

L'intelligence artificielle est une technologie liée à de nombreux domaines dont: **l'informatique**, les **mathématiques**, les **sciences cognitives**, les **statistiques**, la **psychologie** ou encore la **neuro-biologie**, etc.



Dès l'Antiquité

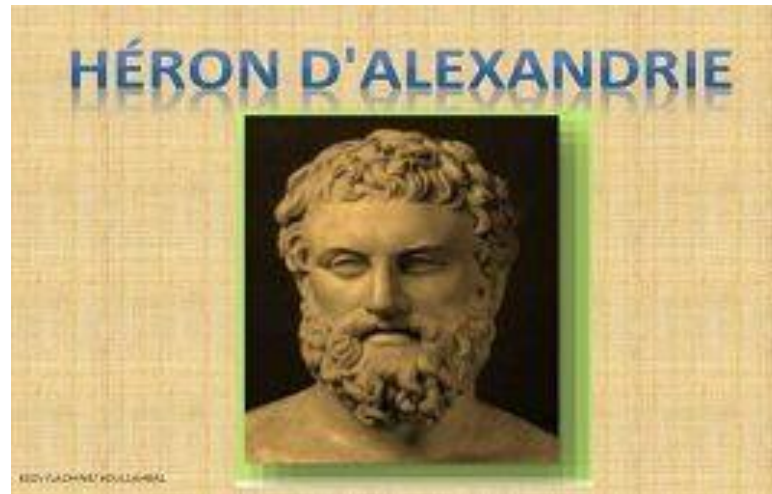
- Les mythes de Pygmalion et d'Héphaïstos



- La machine d'Anticythère, appelée également mécanisme d'Anticythère.

Plusieurs travaux et inventions antérieurs comme:

- La définition du syllogisme par Aristote en 384 av. J.-C.,
- La création par Héron d'Alexandrie des premiers automates au 1er siècle ap. J.-C.
- L'invention de l'*Ars Magna* de Ramon Lulle en 1275.



HIERONYMI CAR
DANI, PRÆSTANTISSIMI MATHE
MATICI, PHILOSOPHI, AC MEDICI,
ARTIS MAGNÆ,
SIVE DE REGVLIS ALGEBRAICIS,
Lib. unus. Qui & totius operis de Arithmetica, quod
OPVS PERFECTVM
inscribit, est in ordine Decimus.



HAbes in hoc libro, studiose Lector, Regulas Algebraicas (Itali, de la Col
fa vocant) nouis adinventionibus, ac demonstrationibus ab Authore ita
locupletatas, ut pro pauculis antea vulgò tritis, iam septuaginta euaferint. Nes
q; solum, ubi unus numerus alteri, aut duo uni, uerum etiam, ubi duo duobus,
aut tres uni quales fuerint, nodum explicant. Hunc aut librum ideo seors
sim edere placuit, ut hoc abstrusissimo, & planè inexhausto totius Arithmeti
cæ thesauro in lucem eruto, & quasi in theatro quodam omnibus ad specian
dum exposito, Lectores incitarentur, ut reliquos Operis Perfecti libros, qui per
Tomos edentur, tanto auidius amplectantur, ac minore fastidio perdiscant.

A partir du 17^e siècle

- La théorie de Descartes selon laquelle les corps des animaux n'étaient rien de plus que des machines complexes
- La théorie de Blaise Pascal qui lui a permis de créer la première calculatrice digitale en 1642
- Les travaux de Gottfried Leibniz en ce qui concerne le système de numération binaire



Le début du 20e siècle

Les travaux de Bertrand Russell et d'Alfred North Whitehead sur la logique, publiés dans *Principia Mathematica*.

Les travaux Ludwig Wittgenstein et Rudolf Carnap sur l'analyse logique de la connaissance.

Les travaux d'Alonzo Church, qui a développé le système formel Lambda-calcul.

Des inventions ou découvertes du 20e siècle

- La machine et le test de Turing.
- La première définition d'un neurone formel de McCulloch et Pitts,
- Les travaux sur la cybernétique de Wiener ou encore ceux de von Neumann sur la théorie des jeux.
- L'invention des premiers ordinateurs en 1943

Sous le terme intelligence artificielle (IA) on regroupe l'ensemble des “*théories et des techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence.*”

les premières traces de l'IA remontent à **1950** dans un article d'Alan Turing intitulé “*Computing Machinery and Intelligence*” dans lequel le mathématicien explore le problème de définir si une machine est consciente ou non. De cet article découlera ce que l'on appelle aujourd'hui le Test de Turing qui permet d'évaluer la capacité d'une machine à tenir une conversation humaine.

L'officialisation de l'**intelligence artificielle** (nom donné par McCarthy) comme véritable domaine scientifique date de **1956** lors d'une conférence aux États-Unis qui s'est tenue au Dartmouth College (USA).

Une autre origine probable remonte à **1949** dans une publication de Warren Weaver avec un mémo sur la traduction automatique des langues qui émet l'idée qu'une machine pourrait très bien effectuer une tâche qui relève de l'intelligence humaine

Grâce au Logic Theorist proposé par Newell et Simon en **1955**, on peut résoudre les problèmes par le modèle d'arbre. Il a été conçu pour reproduire les compétences de résolution de problèmes d'un être humain et est considéré comme le premier programme d'intelligence artificiel. Il a été capable de prouver 38 des 52 théorèmes des “*Principia Mathematica*” de Whitehead et Russell.

En 1956, plusieurs des plus grands chercheurs de l'époque, dans ce domaine de recherche, se réunissent dans le New Hampshire lors d'une conférence d'été donnée sur le campus du Dartmouth Collège.

1956 Dartmouth Conference: The Founding Fathers of AI



John MacCarthy



Marvin Minsky



Claude Shannon



Ray Solomonoff



Alan Newell



Herbert Simon



Arthur Samuel



Oliver Selfridge



Nathaniel Rochester



Trenchard More

General Problem Solver, développé par
Allen Newell, Herbert Simon et Cliff Shaw
1957

Le langage LISP, créé
par McCarthy au MIT 1958

Eliza, un programme informatique
des chatbots actuels) créé par Joseph 1964 -
1966

Premier match entre programmes d'échecs
voit le jour, remporté par programme Kaissa
1966-1967

Robot virtuel SHRDLU développé par
Terry Winograd capable de dialoguer avec
l'utilisateur, 1971

Les systèmes experts,
- **DENDRAL** créé en 1965, par les
Edward Feigenbaum, Bruce Buchanan,

- **MYCIN** de Joshua Lederberg
pour assister les médecins dans
leur diagnostic, 1974

- Problèmes de logiciels.
- *Manque de mémoire et la puissance de calcul*
Beaucoup d'experts ne réussissent pas à faire

- Le rapport Lighthill (pronostic très pessimiste)
Arrêt des investissements

Période des grandes espoirs
L'âge d'or 1956–1974

Premières déceptions
**La première hibernation de l'intelligence
artificielle 1974–1980**

En 1980, un système expert appelé *XCON* a été réalisé à la **Carnegie Mellon University** pour la Digital Equipment Corporation

Créée l'Association américaine pour l'Intelligence Artificielle (AAAI, American Association for Artificial Intelligence) 1980

James F. Allen qui travaille sur la logique temporelle, introduit en 1983, l'algèbre des intervalles d'Allen, une méthode de calcul fondamentale pour le futur de l'intelligence artificielle

Le programme de dessin autonome *AARON*, créé par Harold Cohen et présenté à la conférence de l'AAAI en 1985, marque également cette décennie

Utilisation des méthodes de rétropropagation du gradient dans les réseaux de neurones, travaux de Rumelhart, Hinton et Williams en 1986

Allan Newell, John Laird et Paul Rosenbloom développent l'architecture cognitive symbolique *Soar*, présentée en 1987

Période de renaissance **1980–1987**

La deuxième hibernation **1987-1993**

Victoire de Deep Blue d'IBM aux échecs contre le champion du monde Garry Kasparov, 1997

Succès des robots de Stanford et du CMU au Darpa Urban Challenge en 2005 et 2007

Victoires de Watson d'IBM au Jeopardy en 2011

Eugène Goostman a réussi à passer le test de Turing

Victoires d'AlphaGo de Google DeepMind au jeu de Go en 2015 et 2016 contre des champions professionnels

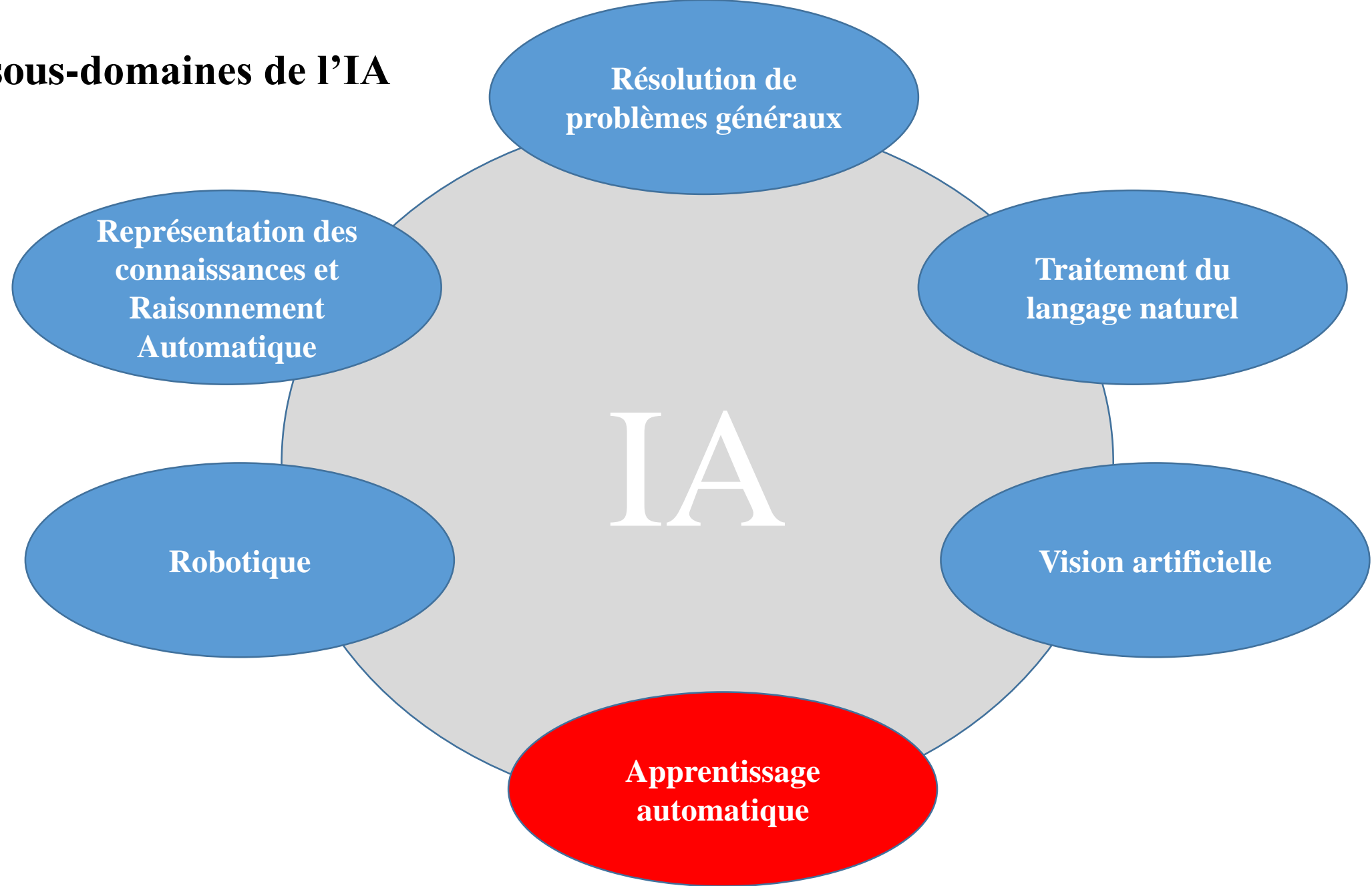
Victoires Libratus, contre quatre joueurs de poker Texas Hold'em, 2017

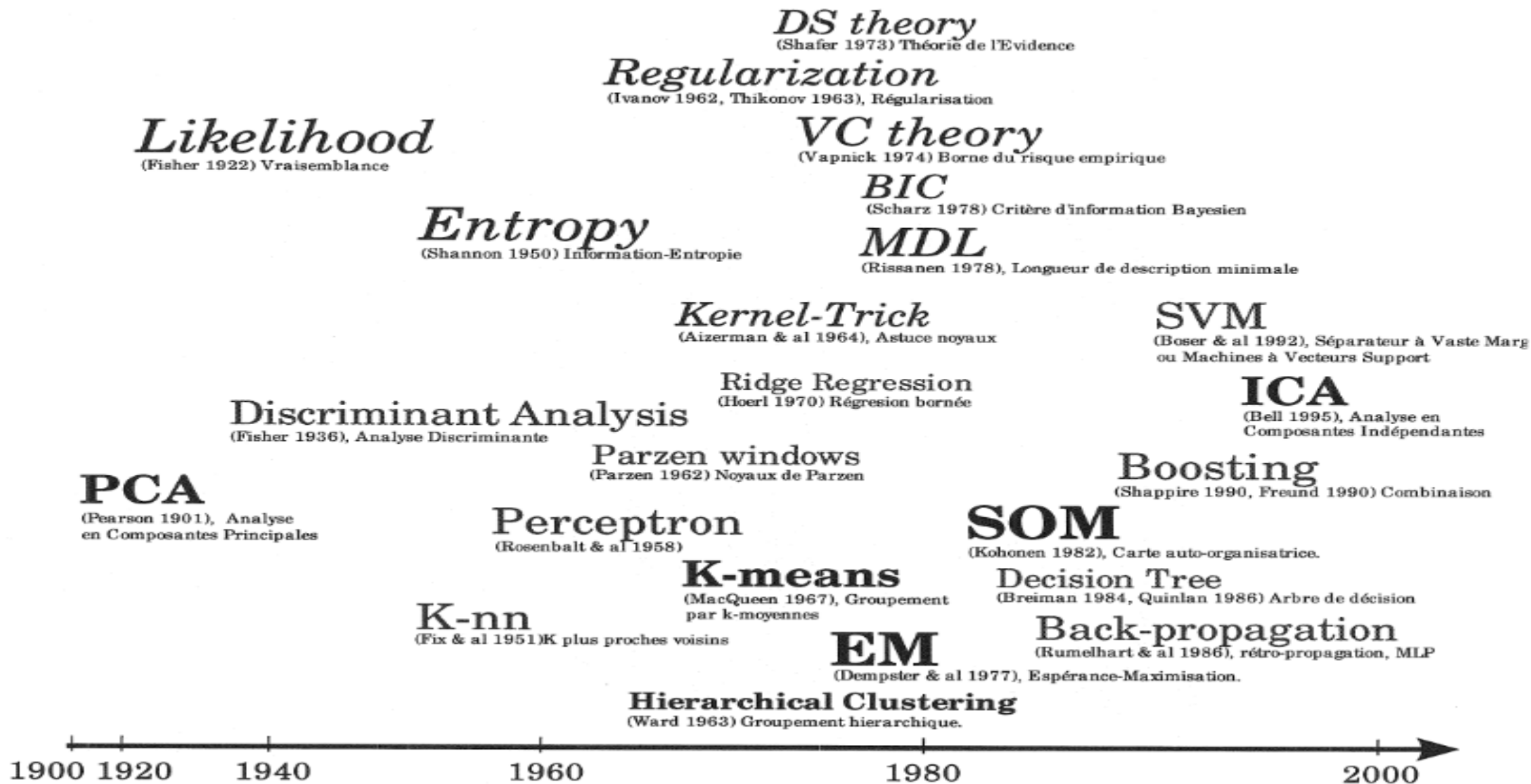
Lancement du projet Project Debater , IBM 2018

Période de renaissance **1993–2011**

L'IA sans limites **2011- ...**

Les sous-domaines de l'IA





Un siècle d'apprentissage automatique

Science de données (data science)

La *data science* est une démarche empirique qui se base sur des données pour apporter une réponse à des problèmes.

C'est quoi les données?

Une données c'est le résultat d'une observation faite sur une population ou sur un échantillon

Une donnée peut être un nombre, une caractéristique, qui m'apporte une information sur un individu, un objet ou une observation

Généralement, on lie les données à des variables parce que le nombre/la caractéristique varie si on observe plusieurs objets/individus/observations

Les principaux types de données

On distingue généralement les données **quantitatives** des données **qualitatives**

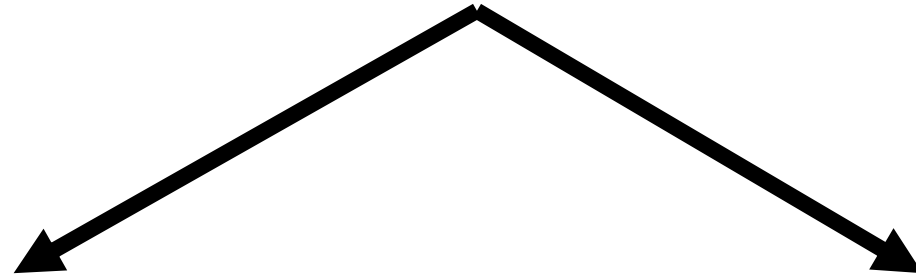
Les différents niveaux de structuration des données

Niveau de structuration	Modèle de données	Exemples	Facilité de traitement
Structuré	Système de données relationnel objet/colonne	Base de données d'entreprise...	Facile (indexé)
Semi-structuré	XML, JSON, CSV, logs	API Google, API Twitter, web, logs...	Facile (non indexé)
Non structuré	Texte, image, vidéo	web, e-mails, documents...	Complexe

Variable quantitatives

Variable qualitative :

Une variable statistique est qualitative si ses valeurs, ou **modalités**, s'expriment de façon littérale ou par un codage sur lequel les opérations arithmétiques telles que moyenne, somme, ... , n'ont pas de sens.



Variable qualitative nominale :

C'est une variable qualitative dont les modalités ne sont pas ordonnées.

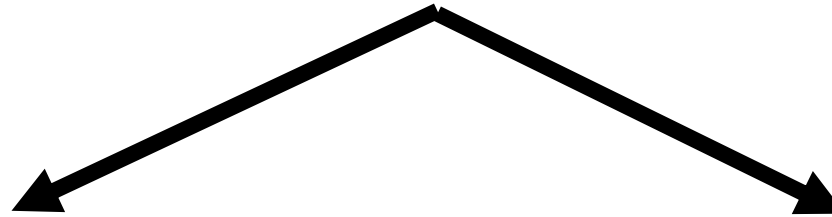
Variable qualitative ordinale :

C'est une variable qualitative dont les modalités sont naturellement ordonnées

Variable qualitatives

Variable quantitative :

Une variable statistique est quantitative si ses valeurs sont des nombres exprimant une quantité, sur lesquels les opérations arithmétiques (somme, etc...) ont un sens.



Variable quantitative discrète:

Une variable quantitative est discrète si elle ne peut prendre que des valeurs isolées, généralement entières.

Variable quantitative continue:

Une variable quantitative est continue si ses valeurs peuvent être n'importe lesquelles d'un intervalle réel.

Si on observe n individus, et d variables . Le nuage de points peut se décomposer de deux manières,

- l’espace des individus,
- l’espace des variables,

On note x_{ij} l’observation de la j ème variable sur le i ème individu. On a donc le tableau de données

		Variables				
		<i>1</i>	<i>...</i>	<i>j</i>	<i>...</i>	<i>d</i>
Individus	<i>1</i>	x_{11}	$...$	x_{1j}	$...$	x_{1d}
	\vdots	\vdots		\vdots		\vdots
	<i>i</i>	x_{i1}	$...$	x_{ij}	$...$	x_{id}
	\vdots	\vdots		\vdots		\vdots
	<i>n</i>	x_{n1}	$...$	x_{nj}	$...$	x_{nd}

Malédiction de la dimensionnalité

Ou fléau de la dimensionnalité (curse of dimensionality), Richard Bellman

nombre d'échantillons nécessaires pour estimer une fonction arbitraire avec un niveau de précision donné augmente de façon exponentielle par rapport au nombre de variables d'entrée (c'est-à-dire, la dimensionnalité) de la fonction.

Etant donné un ensemble d'apprentissage $S = \{x_i \in \mathbb{R}^{n \times d}\}$

➤ **Cas 1) n est très grand, d est très grand**

Typiquement un problème de données massives BIG DATA ($n \nearrow, d \nearrow$) \Rightarrow solutions algorithmiques et architecturales.

➤ **Cas 2) n est petit et d est très grand**

Problème d'inférence (mauvaise capacité de généralisation)

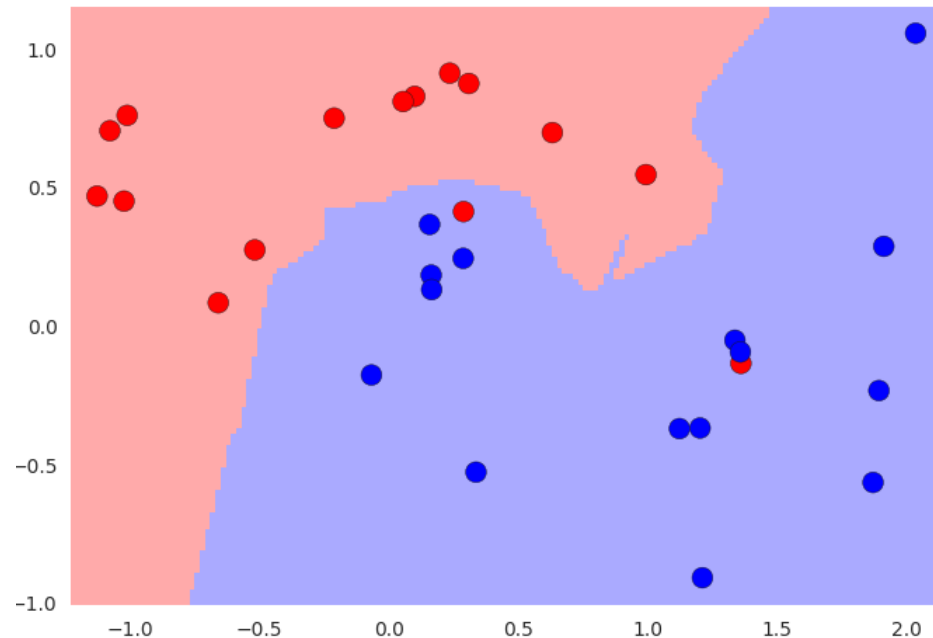


Malédiction de la dimensionnalité

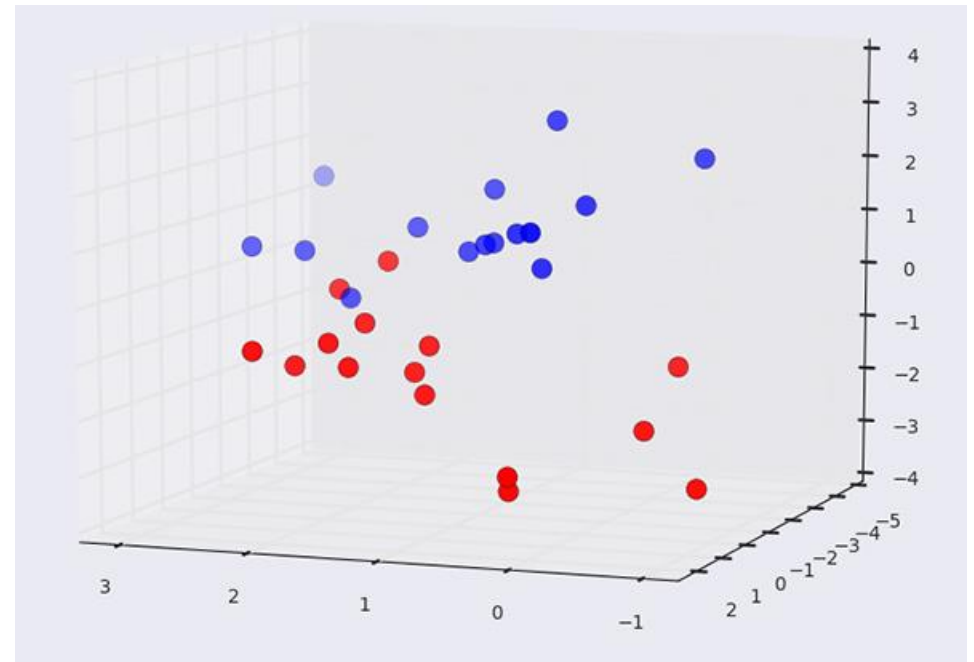
➤ Cas 2) $n \searrow, d \nearrow$

- Classification d'images
- Traitement de textes
- Analyse des séries temporelles

Généralement, dans ces domaines on dispose des bases d'apprentissage dont la taille est plus petite que la dimension.

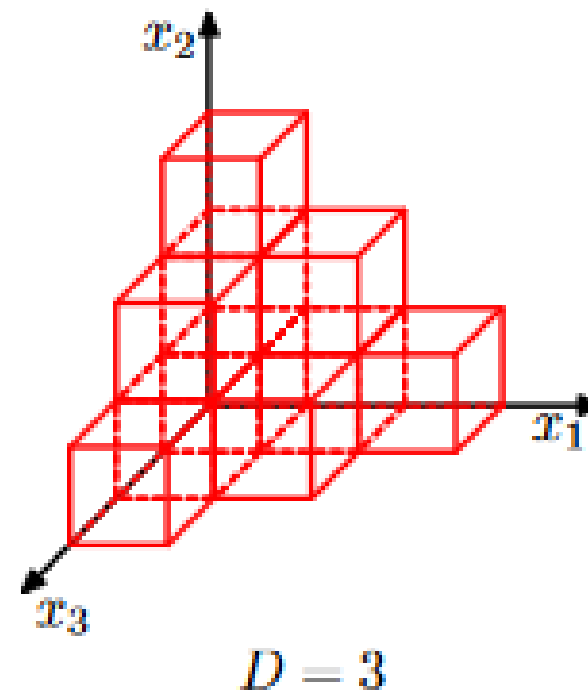
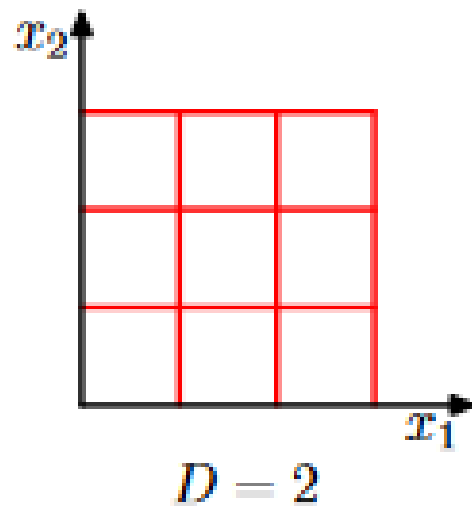


Jeu de données en 2 dimensions



Jeu de données en 3 dimensions

Malédiction de la dimensionnalité



Malédiction de la dimensionnalité

Conséquences

- Problème de la capacité de généralisation
- Des attributs non pertinents peuvent affecter l'optimalité des algorithmes.
- La redondance d'attributs nuit à la régularisation.

Solutions: Réduire la dimension

1. Sélectionne de caractéristiques (feature selection) : choisir p attributs les plus pertinents parmi les d attributs disponibles
 - Augmenter la capacité généralisation des algorithmes.
 - Réduire la complexité de le modèle sous-jacent.
2. Changement d'espace de représentation: créer p nouveaux attributs
 - Réduire la redondance
 - Augmenter la capacité discriminant des algorithmes