

INF8175 - Intelligence artificielle

Méthodes et algorithmes

**Module 10: utilisation en industrie,
éthique, et philosophie**



POLYTECHNIQUE
MONTRÉAL

Quentin Cappart

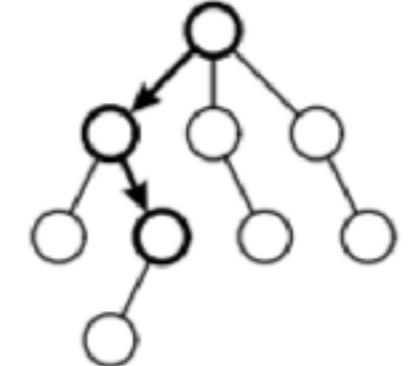
Contenu du cours

Raisonnement par recherche (essais-erreurs avec de l'intuition)

Module 1: Stratégies de recherche

Module 2: Recherche en présence d'adversaires

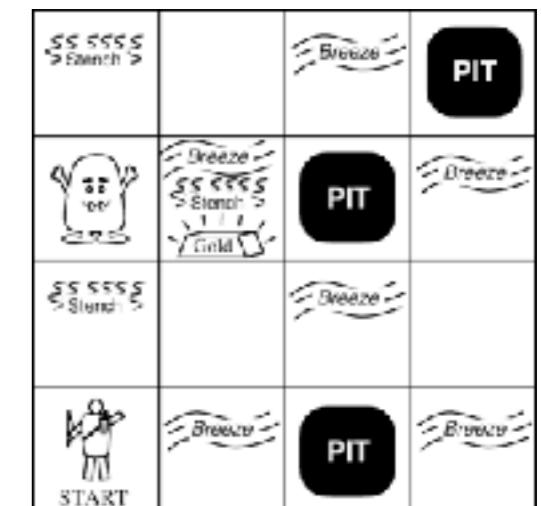
Module 3: Recherche locale



Raisonnement logique

Module 4: Programmation par contraintes

Module 5: Agents logiques



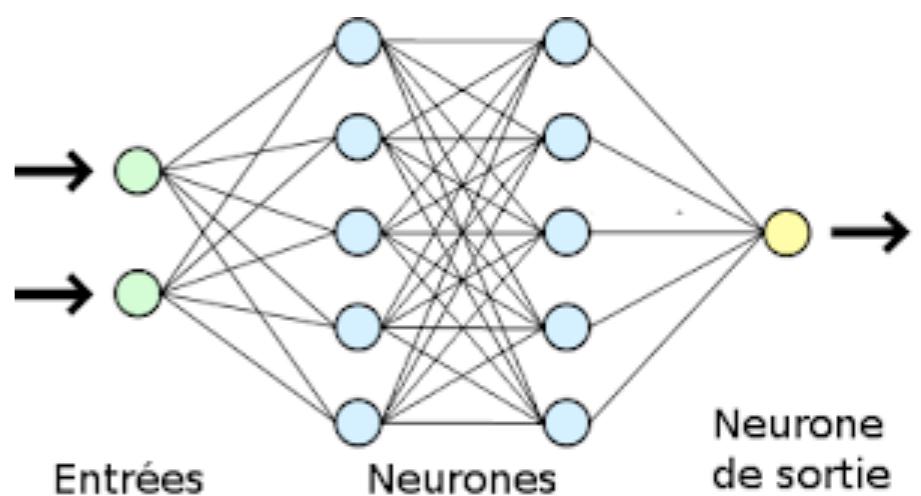
Raisonnement par apprentissage

Module 6: Apprentissage supervisé

Module 7: Réseaux de neurones et apprentissage profond

Module 8: Apprentissage non-supervisé

Module 9: Apprentissage par renforcement



Considérations pratiques et sociétales

Module 10: Utilisation en industrie, éthique, et philosophie

Conclusion du cours



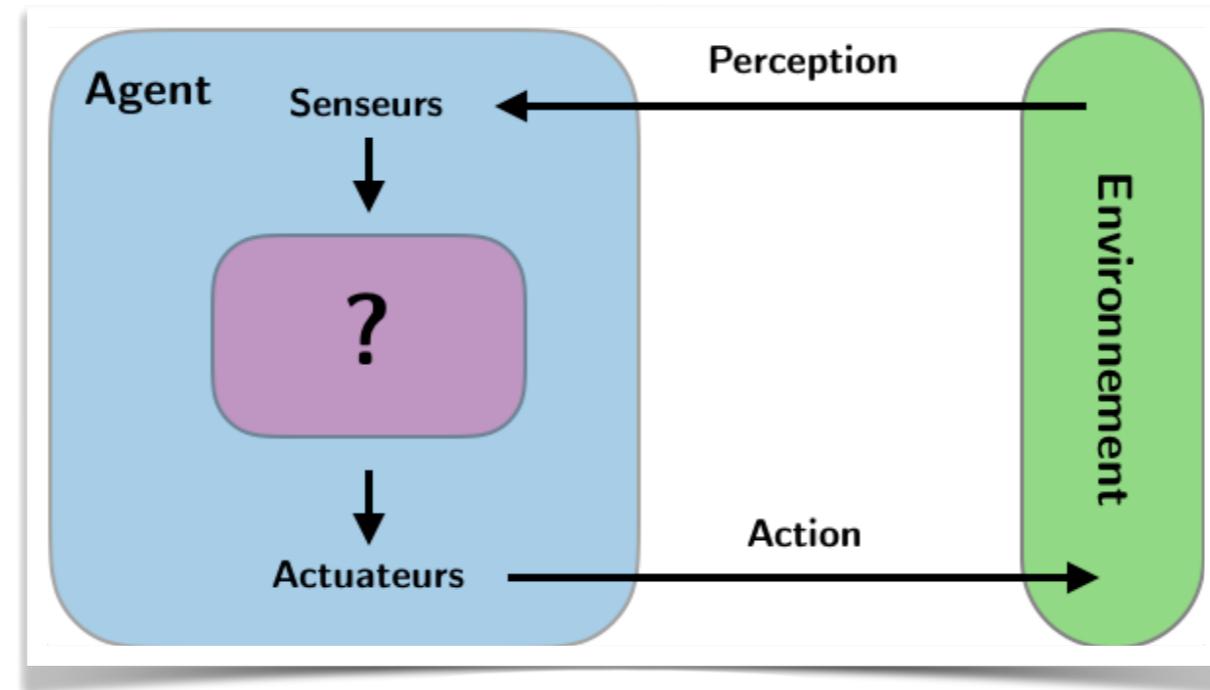
Définition moderne de l'intelligence artificielle

Une intelligence artificielle est une entité capable d'agir rationnellement



Décision rationnelle

Décision qui va maximiser les chances d'accomplir un objectif prédéfini



Objectifs du cours

Connaître, comprendre, et savoir utiliser judicieusement
les principaux types de raisonnement en intelligence artificielle

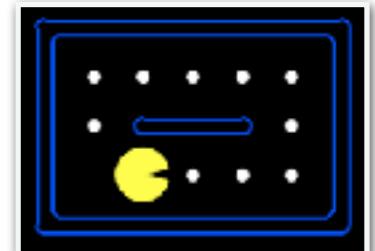
Module 1: stratégies de recherche

Problèmes de recherche

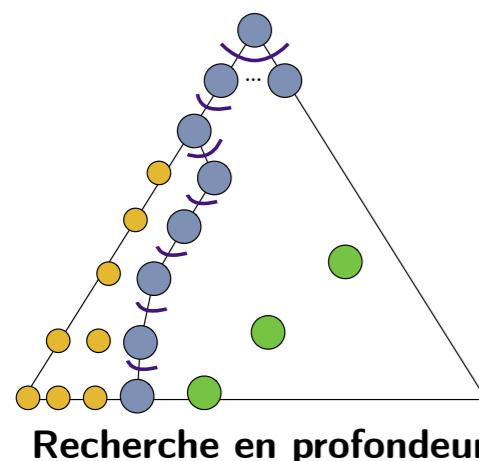
Motivation: planifier d'une séquence d'actions pour aller d'un état initial à un état final

Objectif idéal: trouver la séquence engendrant le moins de coûts

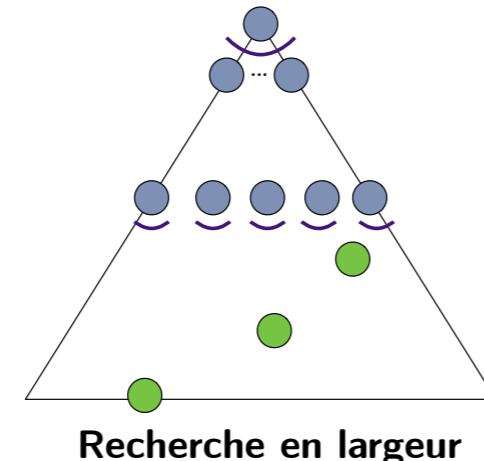
Patron de résolution générique: recherche en graphe ou en arbre



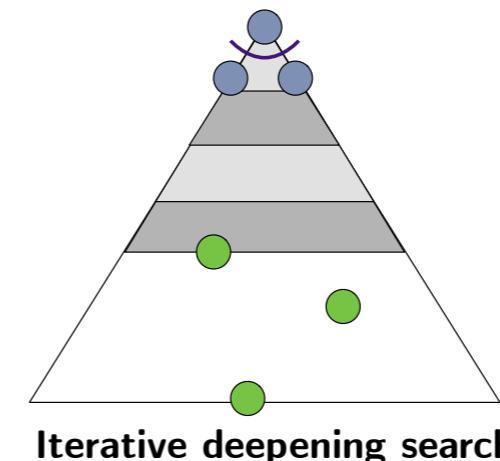
Stratégies sans information



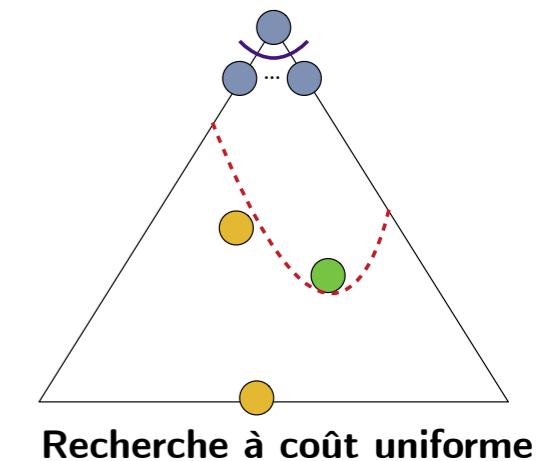
Recherche en profondeur



Recherche en largeur

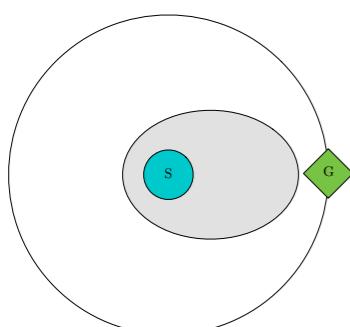


Iterative deepening search



Recherche à coût uniforme

Stratégies avec information



Inclusion d'une heuristique (information du problème) pour accélérer la recherche

Recherche gloutonne: rapide mais sans garantie

Recherche A*: combinaison de la recherche gloutonne et de coût uniforme

Heuristique admissible: A* avec une recherche en arbre est optimal

Heuristique consistante: A* avec une recherche en arbre/graphe est optimal

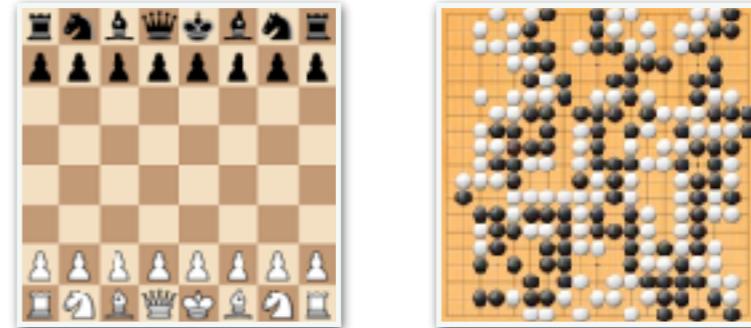
Conception d'heuristiques: relaxation, connaissances expertes, etc.

Module 2: recherche adversarielle

Recherche adversarielle

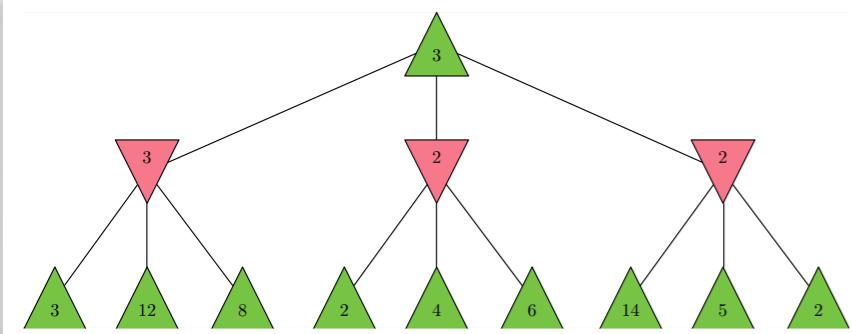
Contexte: l'agent évolue dans un environnement compétitif

Environnement: caractérisé par la présence d'un ou de plusieurs adversaires

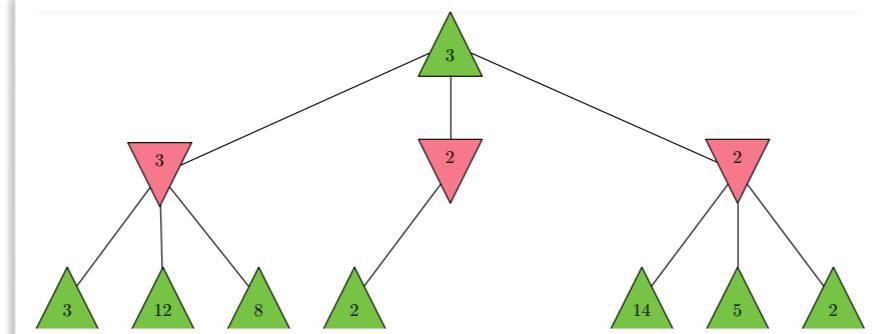


Cadre de référence: deux joueurs, déterministe, tour par tour, à information parfaite, et à somme nulle

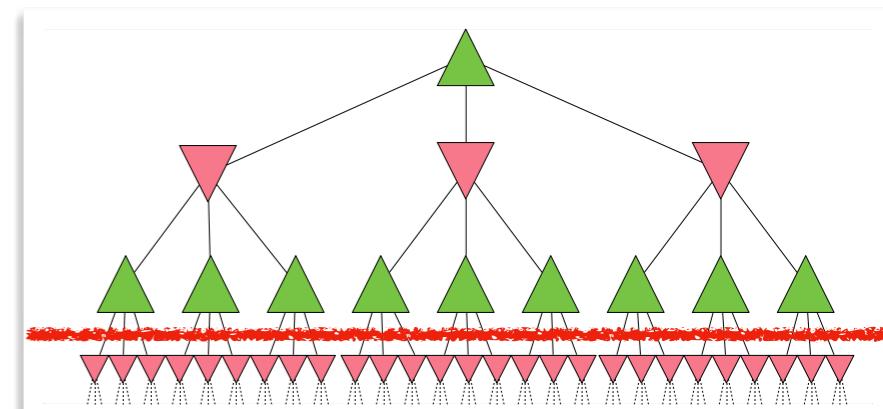
Stratégie de recherche Minimax



Minimax classique



Alpha-beta pruning



Avec des heuristiques

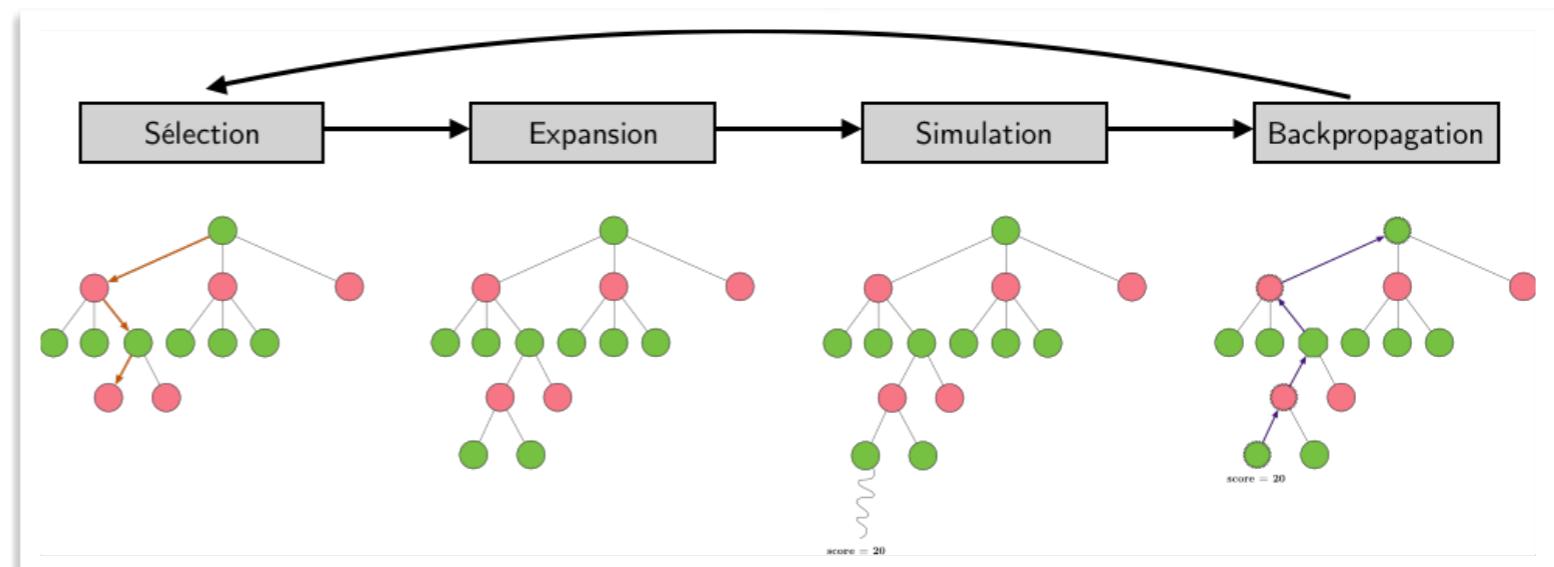
Autres améliorations: tables de transposition, méthodes hybrides, suppression de branches, etc.

Arbre de recherche MC

Basé sur des simulations de parties

Dépendance moindre à une heuristique

Méthode assez générique



Module 3: recherche locale

Problèmes combinatoires

Satisfaction (CSP): trouver une solution satisfaisant un ensemble de contraintes

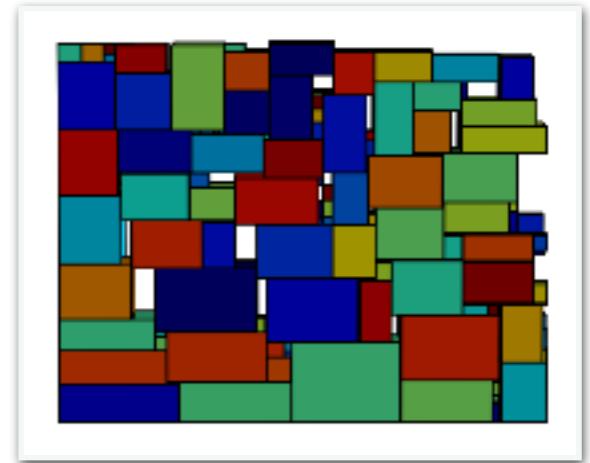
Optimisation (COP): CSP où on souhaite également optimiser une fonction objectif

Une solution est un état, qui nous est inconnu, et non une séquence d'action

5	3		7					
6		1	9	5				
	9	8				6		
8			6				3	
4		8	3				1	
7			2				6	
	6			2	8			
		4	1	9			5	
		8		7	9			

Investissement	Coût (\$)	Revenu espéré dans 10 ans (\$)
A	200	1 000
B	200	1 000
C	200	1 000
D	500	10 000
E	500	10 000
F	800	13 000
G	300	7 000

Budget maximal de 1000 \$



Recherche locale

Principe: résolution en se déplaçant de solutions en solutions via des mouvements locaux

Hill climbing: sélection systématique du meilleur voisin

Risque: être bloqué dans un minimum local

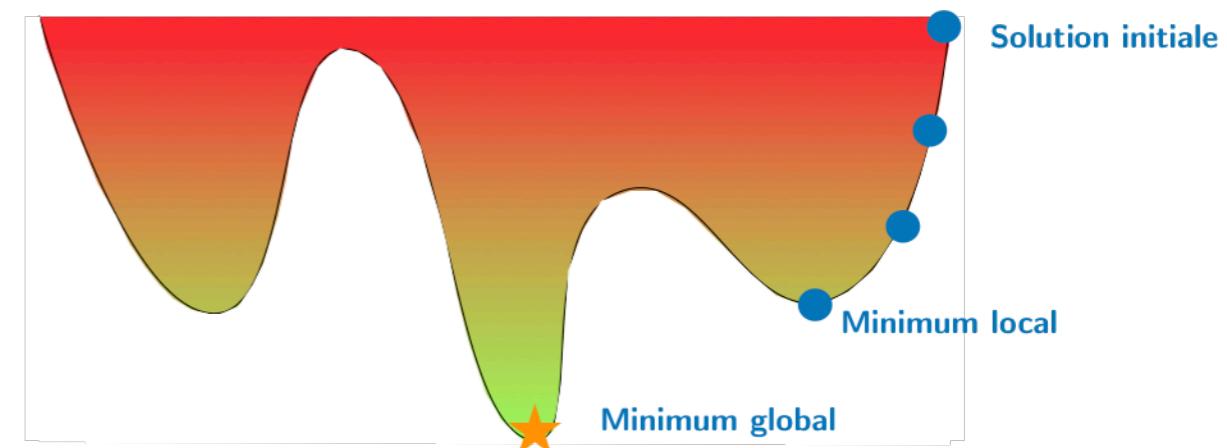
Mécanismes d'amélioration

Idée 1: utiliser un voisinage connecté

Idée 2: agrandir le voisinage

Idée 3: redémarrer aléatoirement la recherche

Idée 4: accepter de dégrader sa solution (p.e., *simulated annealing*)



Module 4: programmation par contraintes

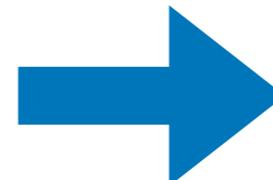
Formalisation des problèmes combinatoires (CSP et COP)

Variables

Domaines

Contraintes

Fonction objectif



Satisfy

Subject to $x_u \neq x_v$ $\forall (u, v) \in E$
 $x_v \in \{\text{red, green, blue}\}$ $\forall v \in V$

Programmation par contraintes = modèle + recherche + propagation

```
int: nc = 3;

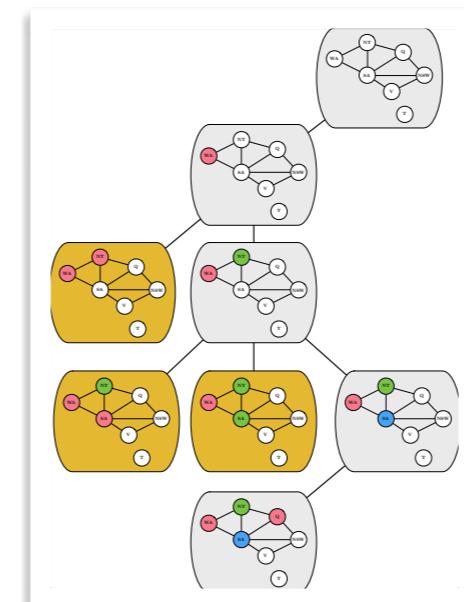
var 1..nc: wa;  var 1..nc: nt;  var 1..nc: sa;  var 1..nc: q;
var 1..nc: nsw;  var 1..nc: v;  var 1..nc: t;

constraint wa != nt;
constraint wa != sa;
constraint nt != sa;
constraint nt != q;
constraint sa != q;
constraint sa != nsw;
constraint sa != v;
constraint q != nsw;
constraint nsw != v;

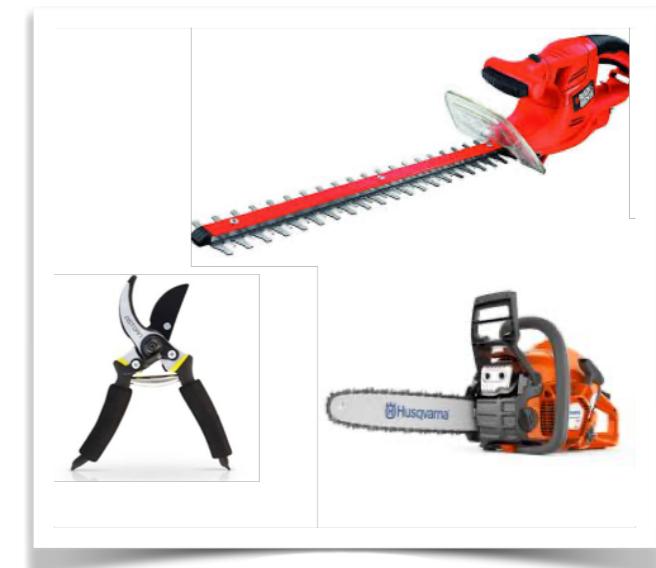
solve satisfy;
```



Modèle du problème



Recherche à
retours arrières



Propagateurs de contraintes
et algorithme point fixe

Module 5: programmation logique

Langage logique

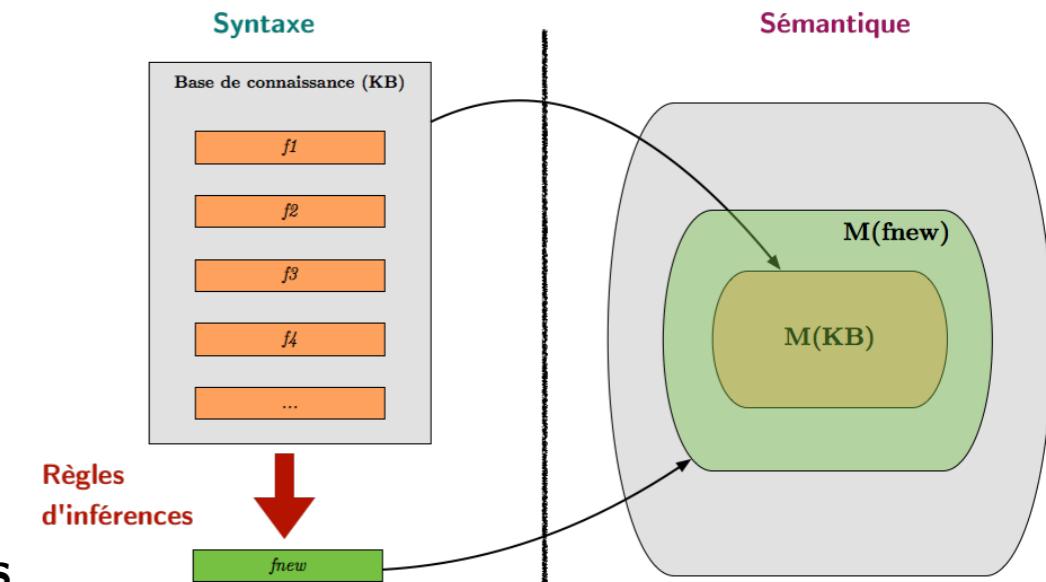
Objectif: raisonner sur base d'un ensemble de connaissances

Composé de 3 éléments fondamentaux

Syntaxe: définit un ensemble de formules valides

Sémantique: définit la signification d'une formule

Règles d'inférence: façon d'obtenir de nouvelles connaissances

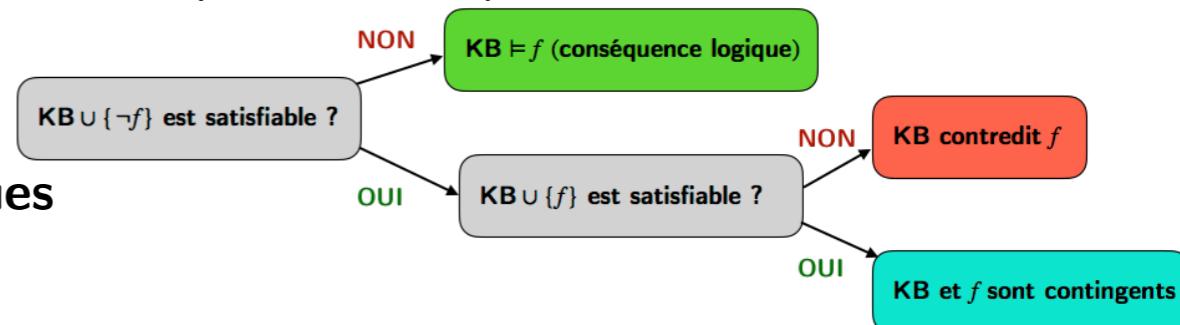


Vérification par model checking

Vérification de la satisfiabilité d'une formule basée sur les modèles (sémantique)

Résolution: via des méthodes de recherche

Limitation: ne tient pas compte qu'on a des formules logiques



Algorithmes d'inférences

Algorithme utilisé pour inférer des nouvelles connaissances en conséquence logique de KB

Cohérence: assure que l'algorithme ne génère que des formules vraies

Complétude: assure que l'algorithme génère toutes les formules vraies existantes

Algorithme de résolution: cohérent et complet dans la logique des propositions



Module 6: apprentissage supervisé

Apprentissage automatique



"A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E"

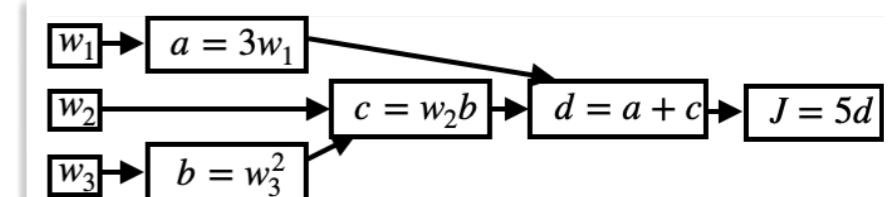
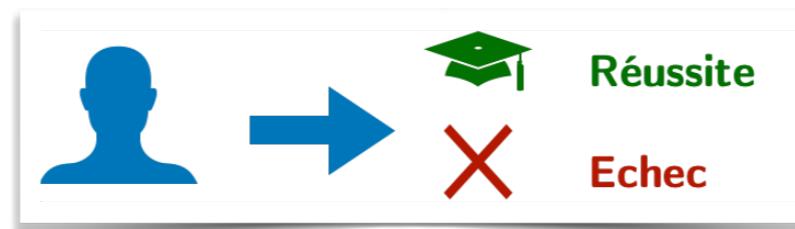
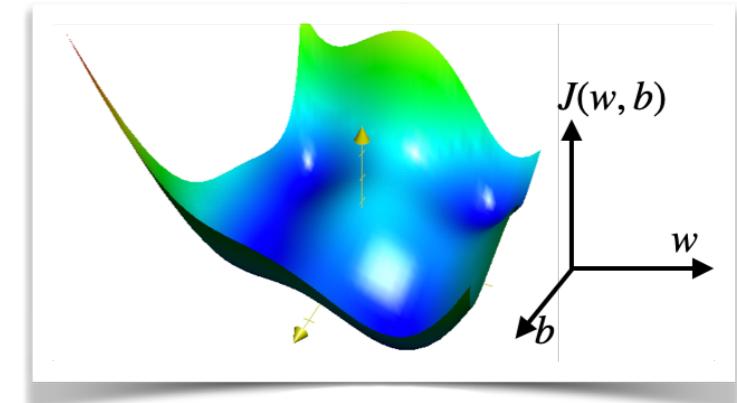
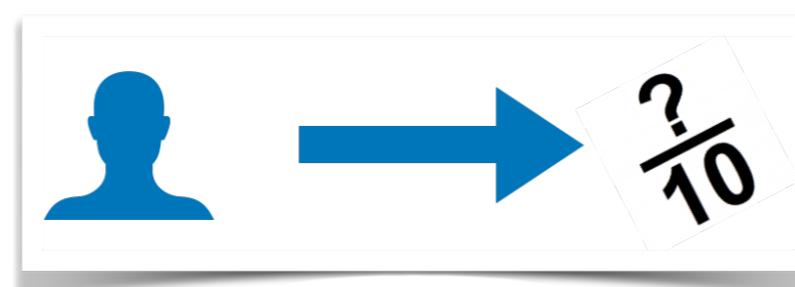
<http://www.cs.cmu.edu/afs/cs.cmu.edu/user/mitchell/ftp/mlbook.html>

- (1) L'apprentissage est dédié pour une tâche précise (T)
- (2) L'apprentissage a besoin d'expériences passées (E - données)
- (3) Besoin d'avoir un moyen d'évaluer les performances d'un modèle (P)

Apprentissage supervisé: minimiser d'un coût entre une prédiction et les données connues

Plusieurs modèles possibles:

- (1) Régression linéaire simple
- (2) Régression linéaire multiple
- (3) Régression logistique
- (4) Et d'autres !



Algorithme d'apprentissage: par descente de gradient via un graphe de dépendance

Module 7: apprentissage profond

Réseaux de neurones

Définition: ensemble de neurones organisés en un réseau

Intérêt: chaque neurone a ses paramètres et contribue à la prédiction

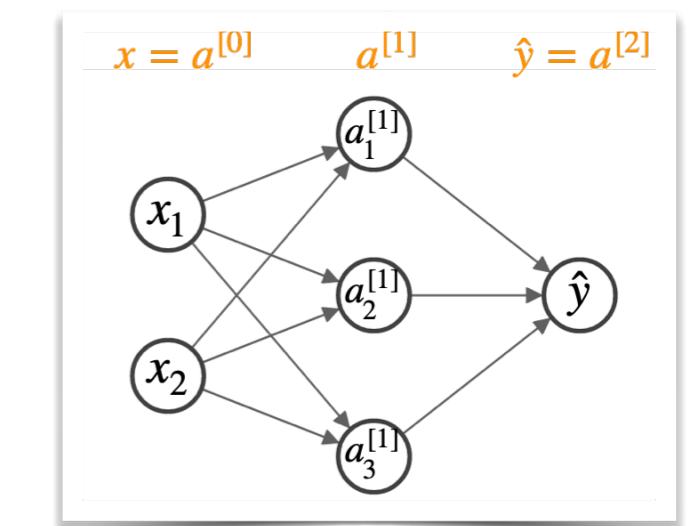
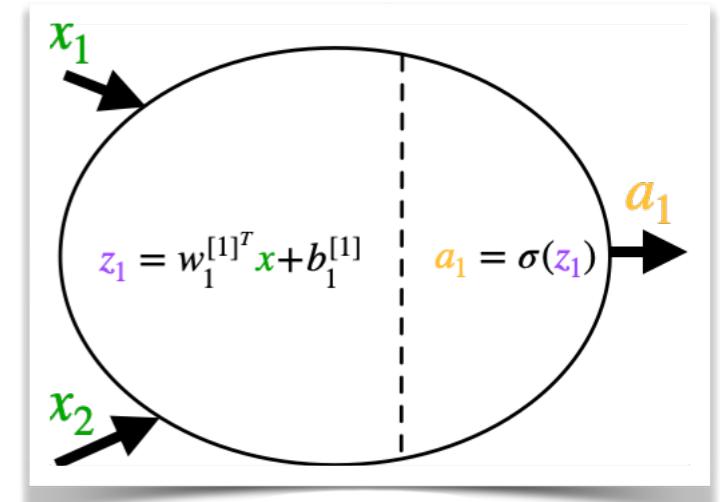
Besoin: activations non-linéaires

Approximateur universel: capacité d'approximer n'importe quelle fonction

Entraînement: descente de gradient et passe en arrière

Difficulté: la fonction de coût n'est plus convexe

Conséquence: utilisation de mécanismes pour faciliter l'apprentissage



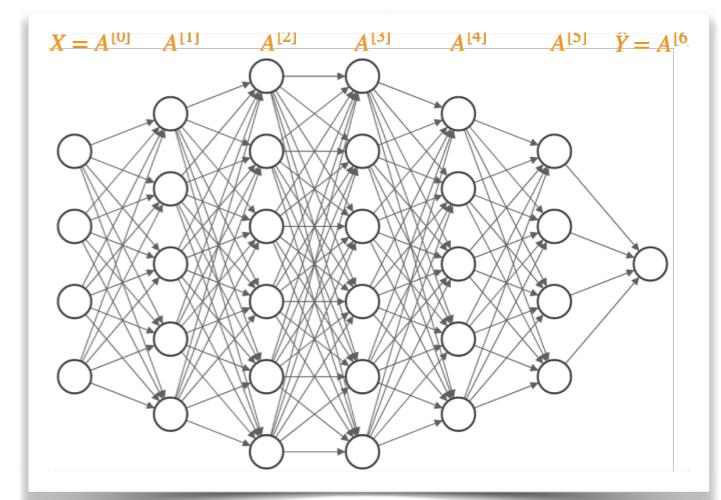
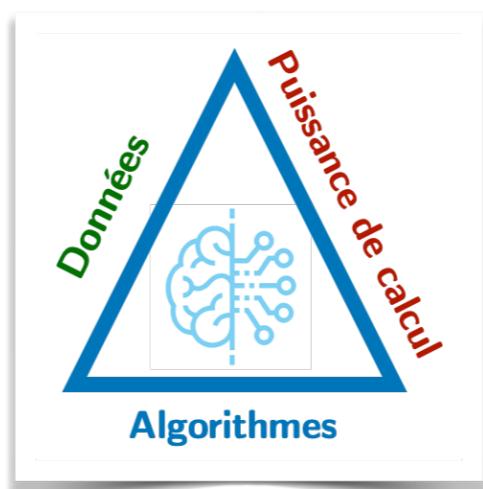
Apprentissage profond

Définition: réseau de neurones ayant plus d'une couche cachée

Intérêt: permet de réaliser des approximations plus précises

Trois ingrédients essentiels:

- (1) Une méthode d'apprentissage
- (2) Un nombre conséquent de données labellisées
- (3) De la puissance de calcul



Module 8: apprentissage non-supervisé

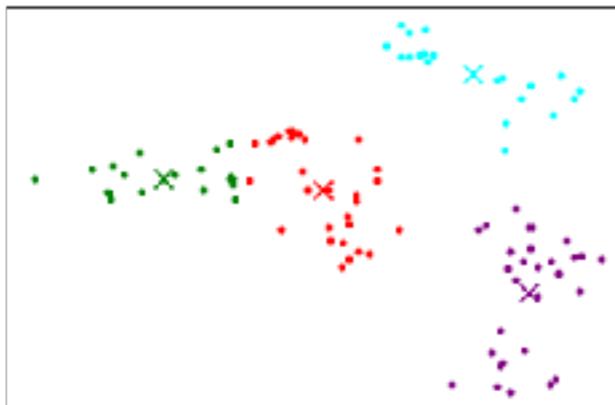
Apprentissage non-supervisé

$$D : \left\{ x^{(1)}, x^{(2)}, \dots, x^{(m)} \right\}$$

Définition: apprentissage qui se fait sur base de données non-labellisées

Principe: exploitation de la similarité entre les données pour l'apprentissage

Problème de regroupement



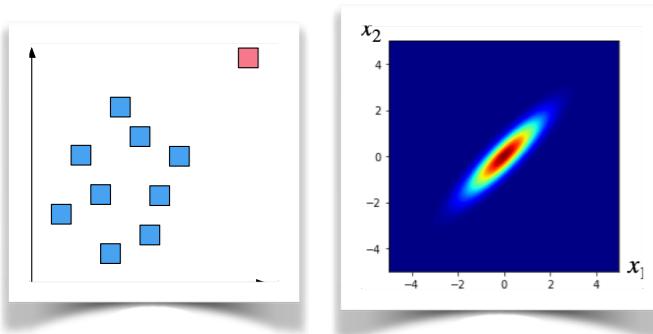
Objectif: regrouper les données similaires en groupes (clusters) cohérents

Méthode de résolution: algorithme *K-means*

Principe: recherche locale pour minimiser une fonction de coût

Améliorations possibles: redémarrages, choix des positions initiales, etc.

Problème de détection d'anomalies

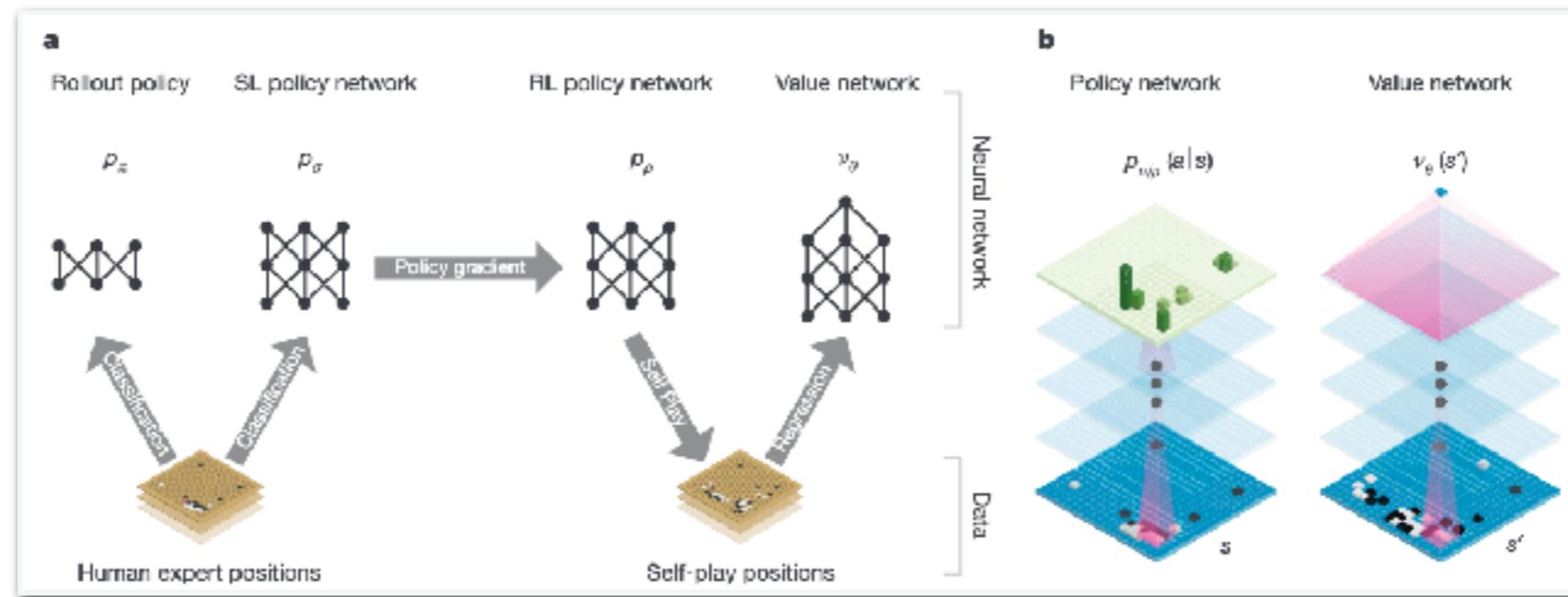


Objectif: prédire la probabilité qu'une nouvelle donnée soit régulière

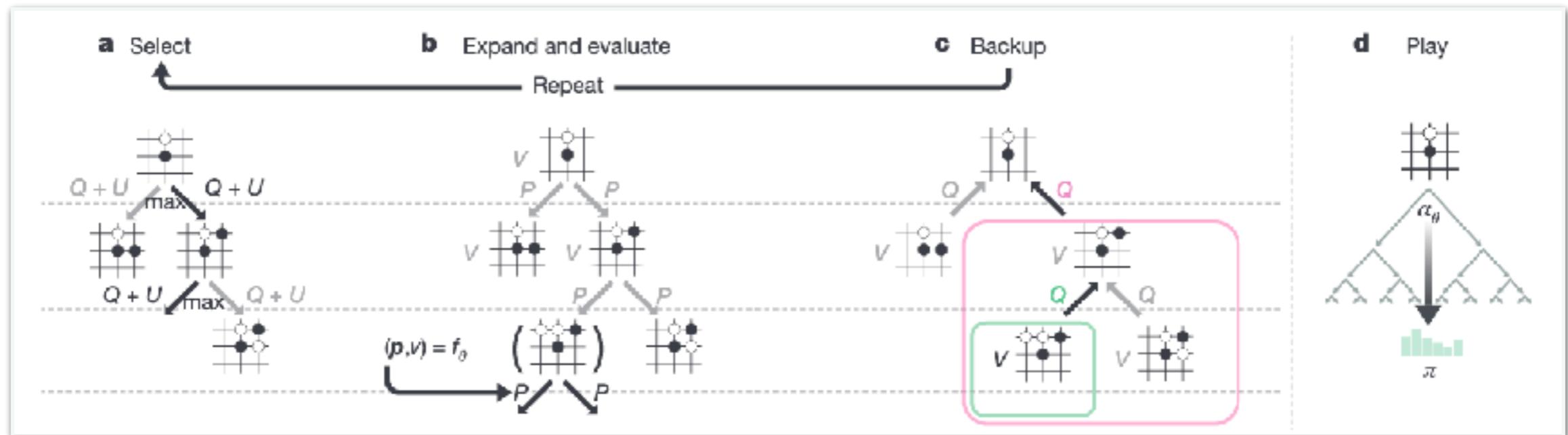
Principe: maximiser une fonction de vraisemblance

Modélisation: produit de lois normales univariées ou une loi normale multivariée

Module 9: apprentissage par renforcement



AlphaGo: combine MCTS, apprentissage supervisé, réseaux de neurones et apprentissage par renforcement



AlphaGo Zero: simplification de AlphaGo et n'utilise plus d'apprentissage supervisé

Apprentissage par vous-même: lecture et compréhension des deux articles (sur Moodle)

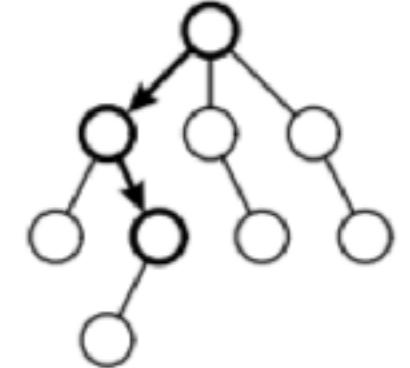
Contenu du cours

Raisonnement par recherche (essais-erreurs avec de l'intuition)

Module 1: Stratégies de recherche

Module 2: Recherche en présence d'adversaires

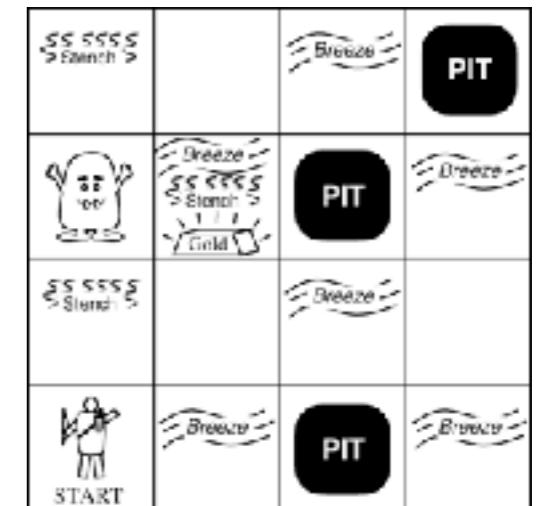
Module 3: Recherche locale



Raisonnement logique

Module 4: Programmation par contraintes

Module 5: Agents logiques



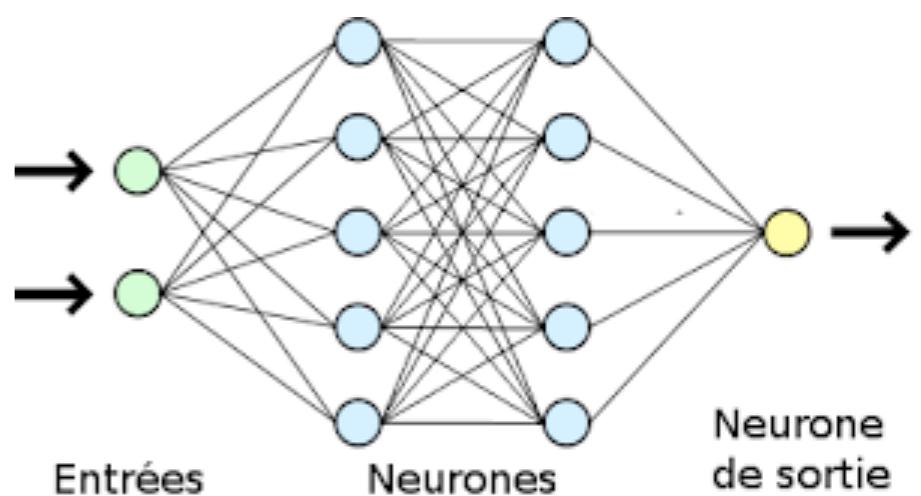
Raisonnement par apprentissage

Module 6: Apprentissage supervisé

Module 7: Réseaux de neurones et apprentissage profond

Module 8: Apprentissage non-supervisé

Module 9: Apprentissage par renforcement



Considérations pratiques et sociétales

Module 10: Utilisation en industrie, éthique, et philosophie

L'intelligence artificielle en industrie



Intervenant: Alexandre Audette Génier (Ubisoft - 2020)

Titre: des agents intelligents aux yeux du joueur

Résumé: la difficulté n'est pas de faire des IAs efficaces,
mais des **IAs crédibles et amusantes pour le joueur**



Intervenant: Raymond Li (ServiceNow Research - 2020)

Titre: application de l'apprentissage automatique pour le NLP

Résumé: construire des modèles plus volumineux n'est pas le seul moyen
d'améliorer les performances (p.e., mécanisme d'attention)



Intervenant: Guillaume Vergnolle (AlayaCare - 2021)

Titre: apprentissage automatique dans le domaine de la santé

Résumé: toutes les erreurs de prédictions n'ont pas le même impact
(faux négatif vs faux positif)



Intervenant: Olivier Blais (MoovAI - 2022)

Titre: il est temps de commencer à livrer des modèles de haute qualité

Résumé: applications diverses du ML en contexte industriel

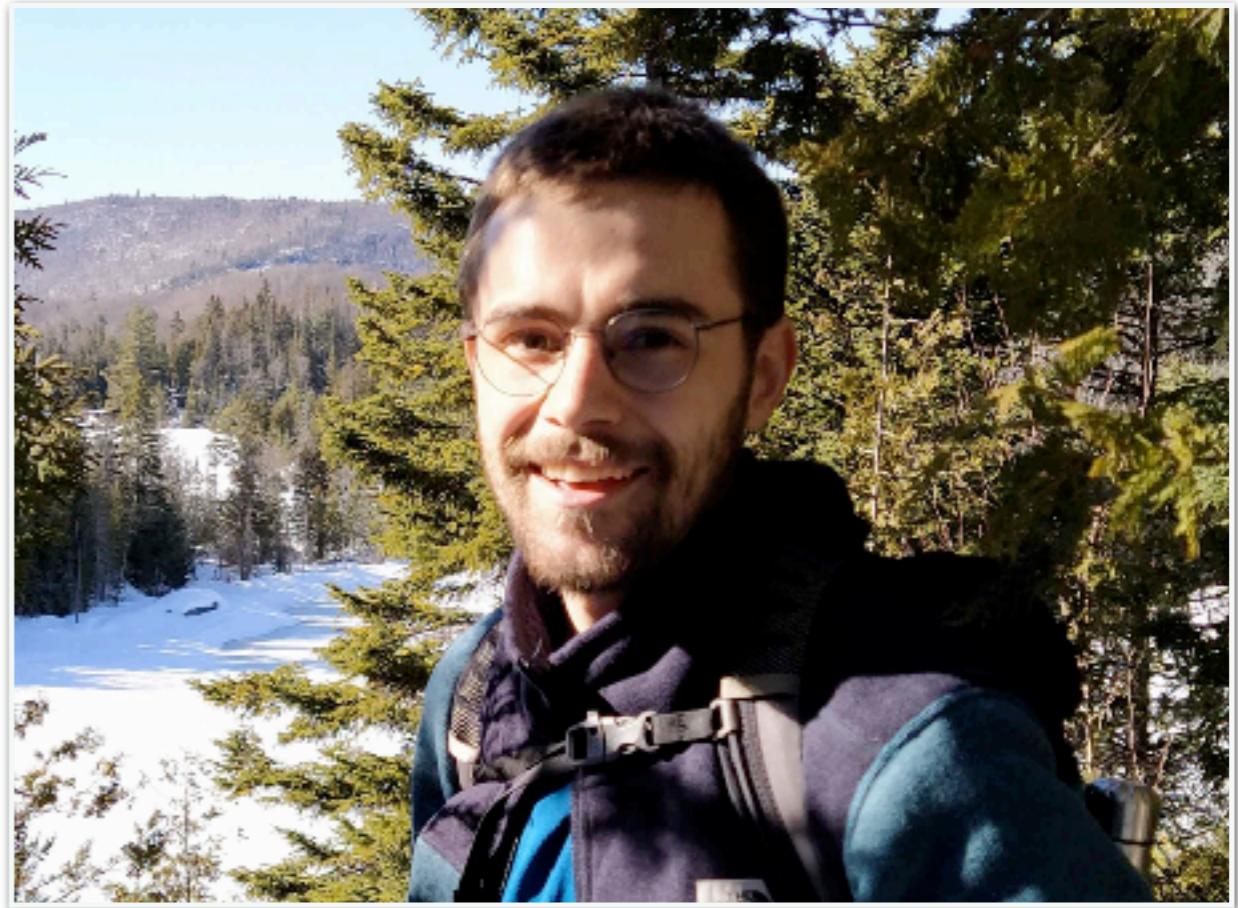
Intervenant de 2023



Maxime Gasse (scientifique de recherche senior à ServiceNow Research)

Déploiement d'agents LLMs en industrie

Intervenant de cette année



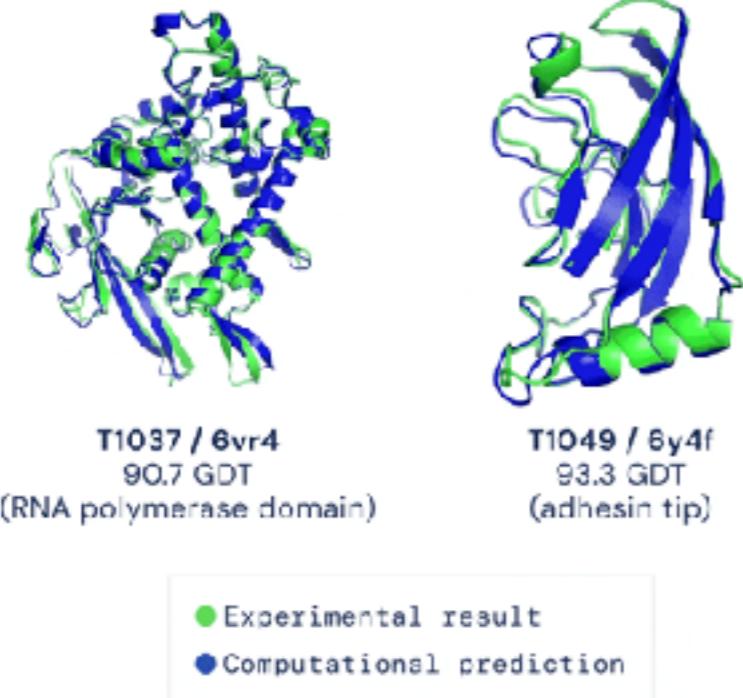
Augustin Parjadis (scientifique de recherche à IVADO Labs)

Combiner apprentissage et optimisation pour le transport électrique



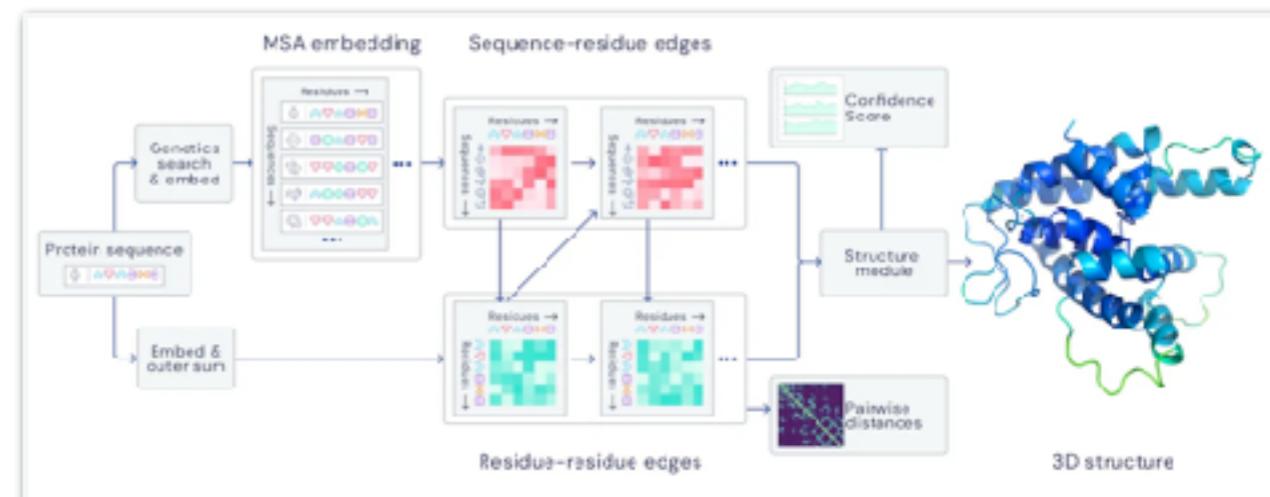
Quels bienfaits l'intelligence artificielle
peut apporter à notre société ?

Enjeux sanitaires



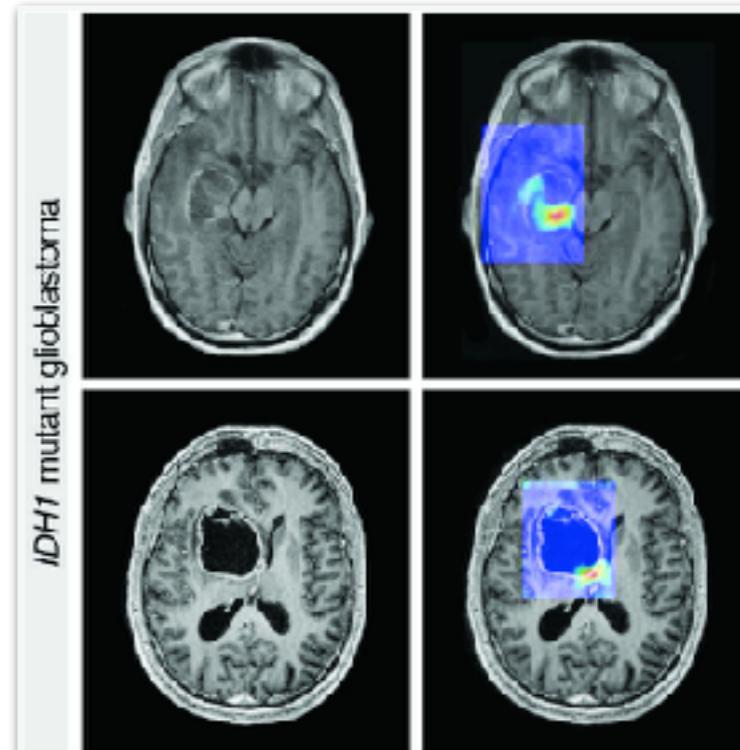
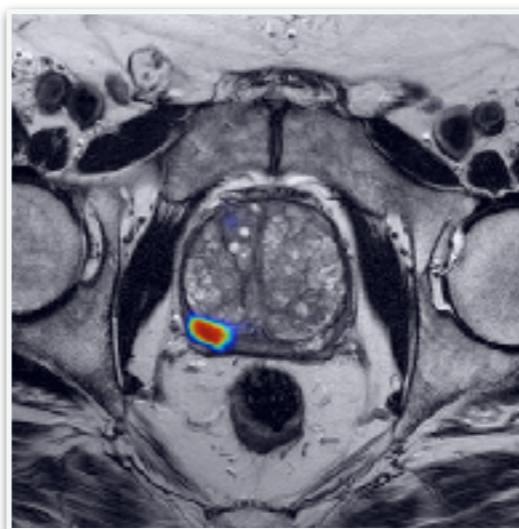
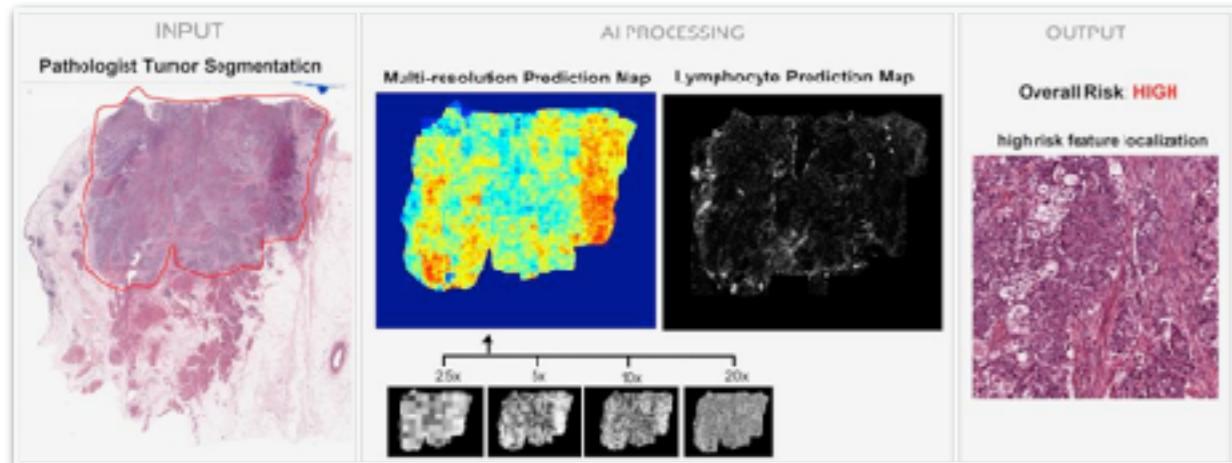
Application: découverte de composants biologiques

Aide: compréhension de certaines maladies

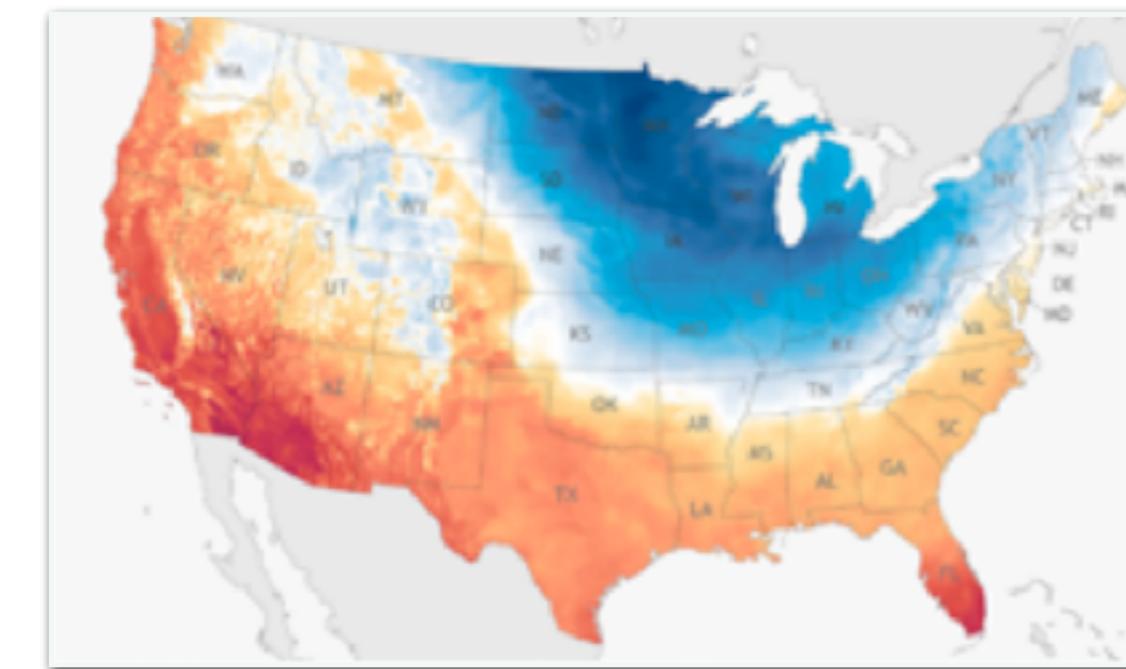


Application: diagnostic de maladies

Aide: complément à l'analyse d'un médecin



Enjeux climatiques

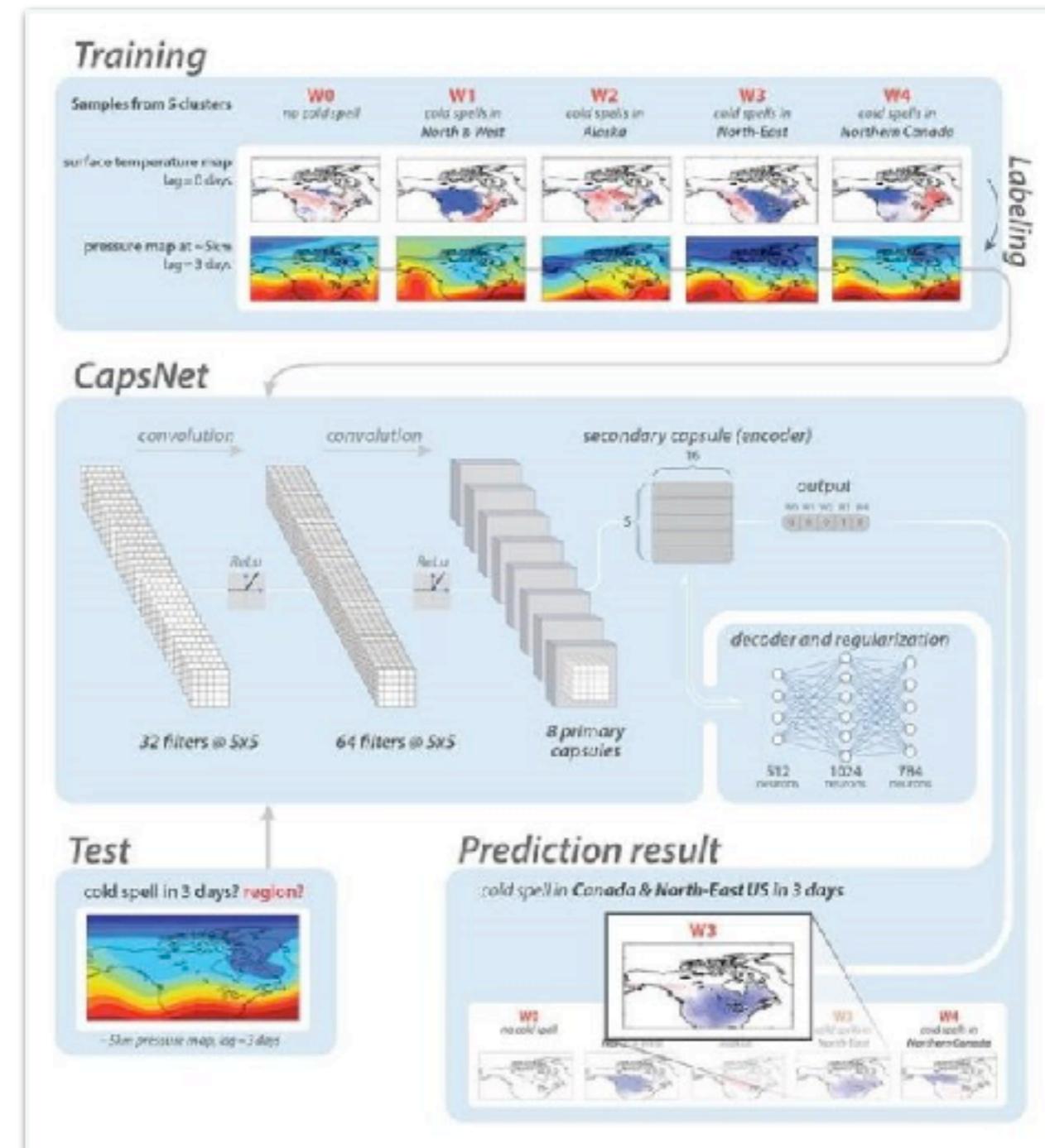


Application: prédition d'évènements météorologiques extrêmes (tornades, tremblements de terre, etc.)

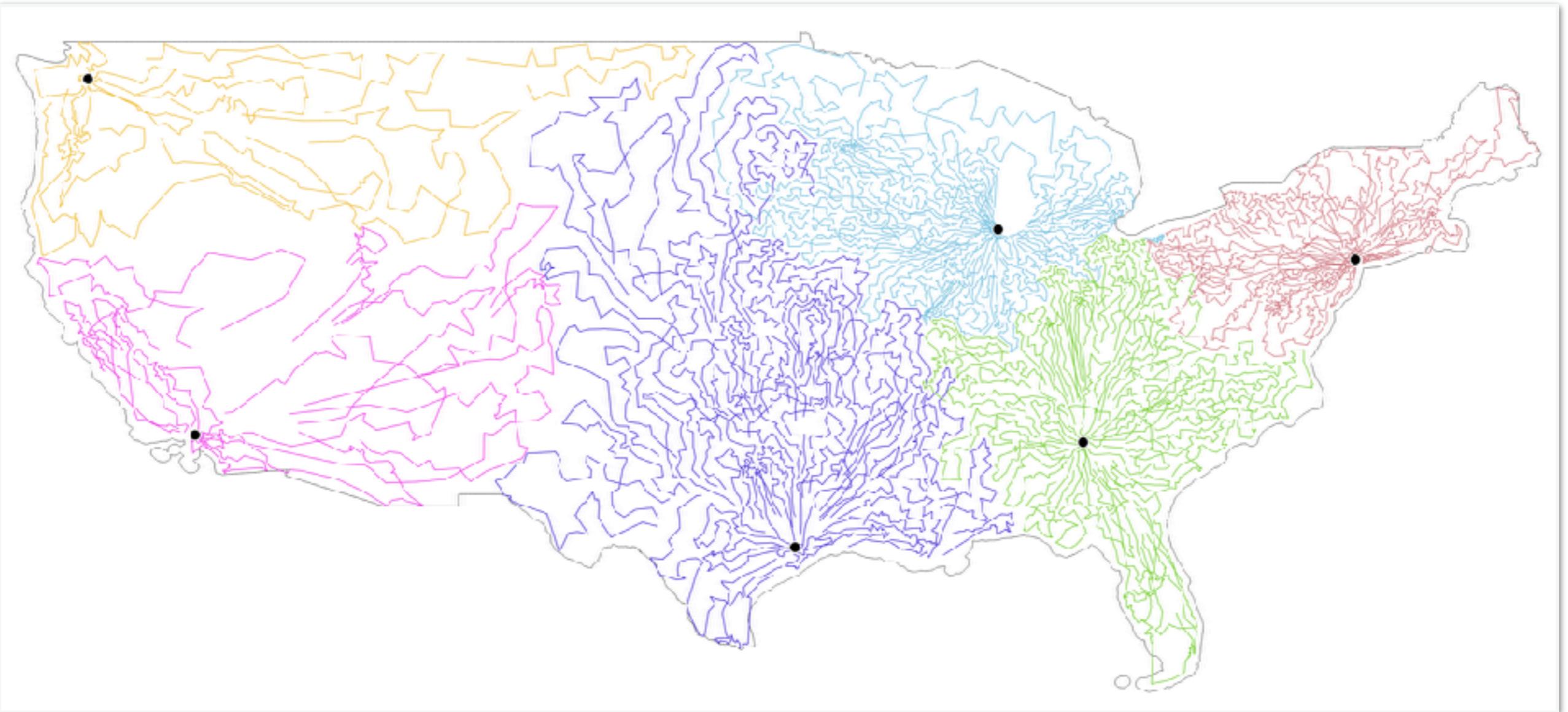
Application: à des fins de sensibilisation

<https://mila.quebec/en/ai-a-tool-to-better-respond-to-environmental-challenges/>

Quentin Cappart <https://developer.nvidia.com/blog/deep-learning-accurately-forecasts-extreme-weather-events/>



Réduction de la consommation des ressources

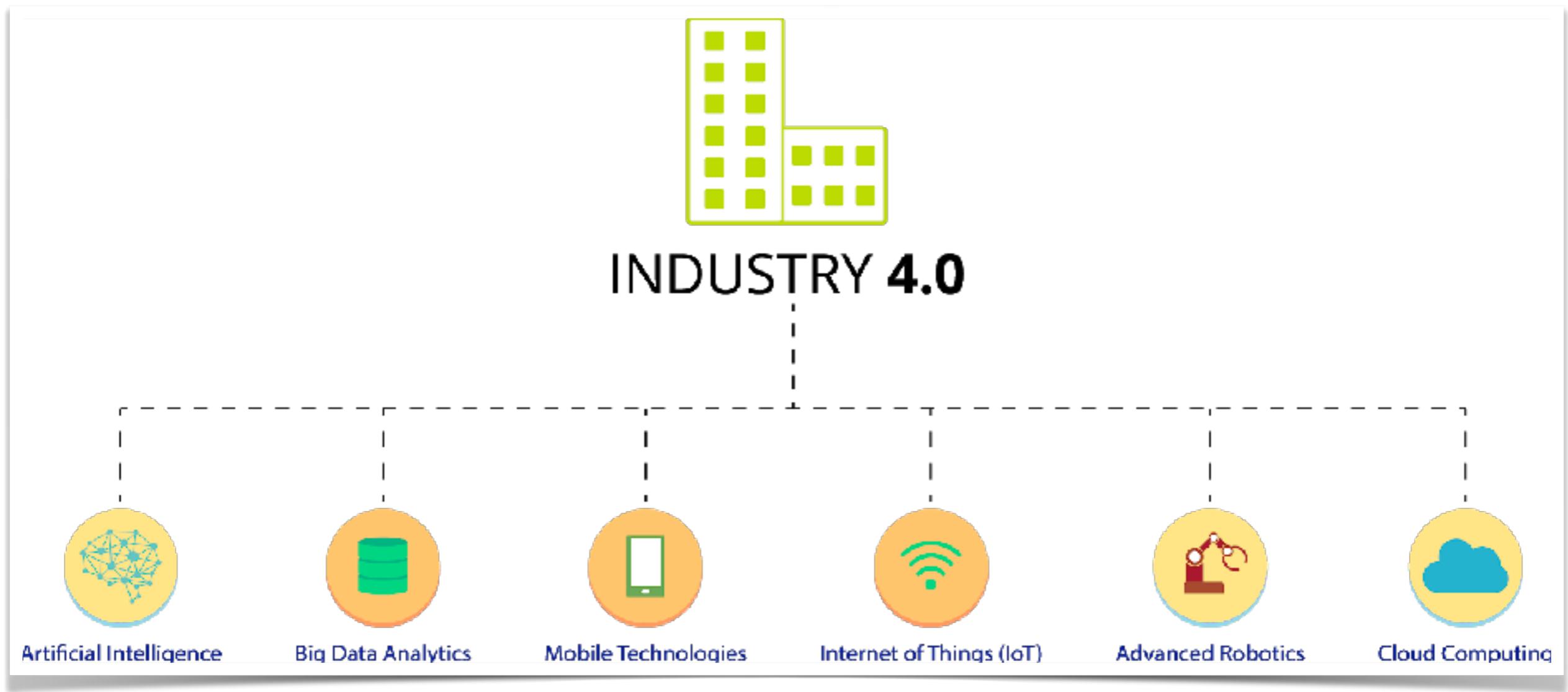


Application: toutes sortes de problèmes de transport ou de production (problème d'optimisation à résoudre)

Aide: réduire la consommation de diverses ressources (carburant, matières premières de construction, etc.)

Aide: rendre ces méthodes accessibles pour des petites entreprises ou pour des particuliers

Enjeux industriels et sociaux



Industrie 4.0: quatrième révolution industrielle basé sur le numérique (vapeur, électricité, informatique)

Effet positif: amélioration de la productivité et automatisation de certaines opérations

Effet positif: création de nouveaux emplois

Sur base de ces exemples, l'intelligence artificielle a le pouvoir de bonifier notre société

Quels dangers les découvertes scientifiques peuvent apporter à notre société ?



DALLE: *An image of a dark dystopia, with orange sky and ruins*

Observation historique - nucléaire

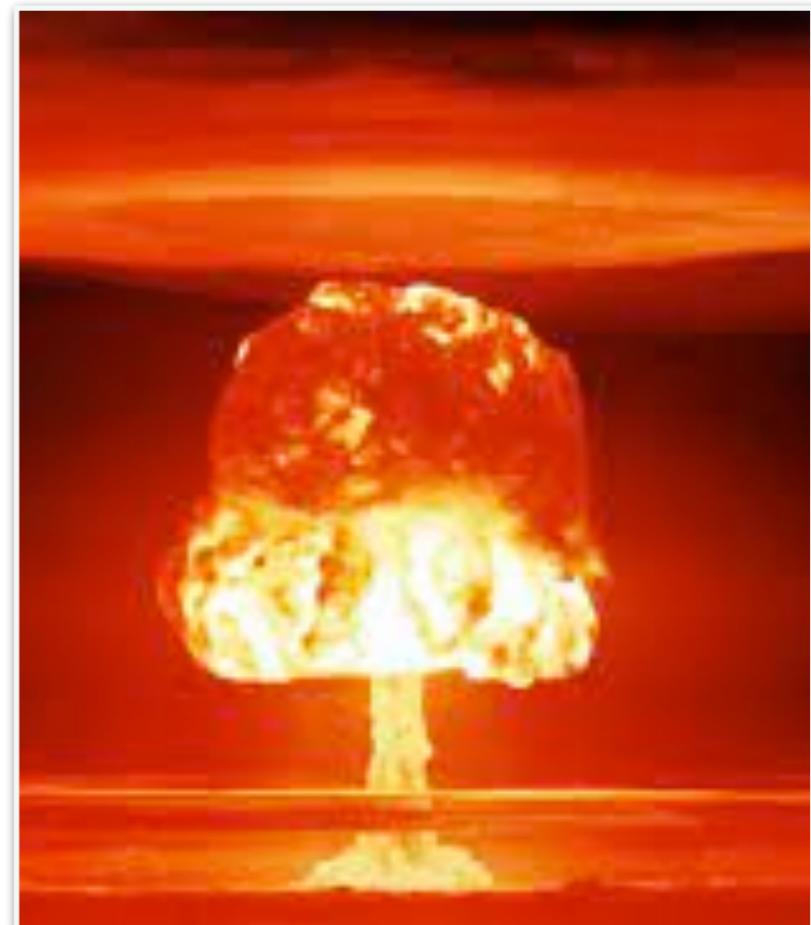
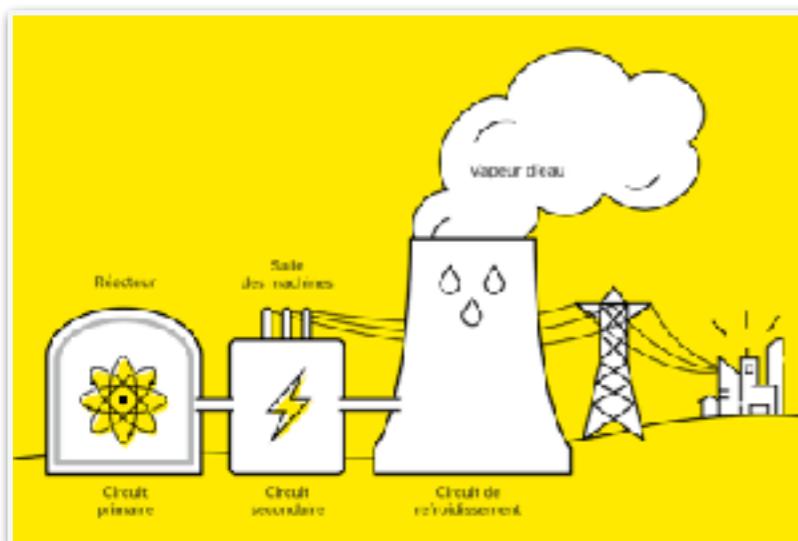
**Chaque invention, de base positive,
amène ses dangers, et ses dérives**

Invention: exploitation de l'énergie nucléaire (fission/fusion)

Utilisation bénéfique: source importante d'énergie

Dangers possibles: accidents nucléaires (Tchernobyl, Fukushima), gestion des déchets

Utilisation meurtrière: création d'une arme (bombe nucléaire)



Observation historique - synthèse de l'ammoniac



Invention: synthèse de l'ammoniac

Utilisation bénéfique: fabrication d'engrais et aide à l'agriculture

Utilisation meurtrière: création d'une arme chimique (gaz moutarde)

A YouTube thumbnail for a video titled "Fritz Haber - Sauver l'humanité en cherchant à la détruire - LPPV.04 - e-penser". The thumbnail features a portrait of Fritz Haber and a gas mask. The video has 1.6M views and was uploaded 3 years ago by e-penser 2.0. The description reads: "Fritz Haber est un chimiste incroyable - en cherchant à tuer en masse, il a sauvé l'humanité, et en cherchant à améliorer la ...". The video duration is 16:20.

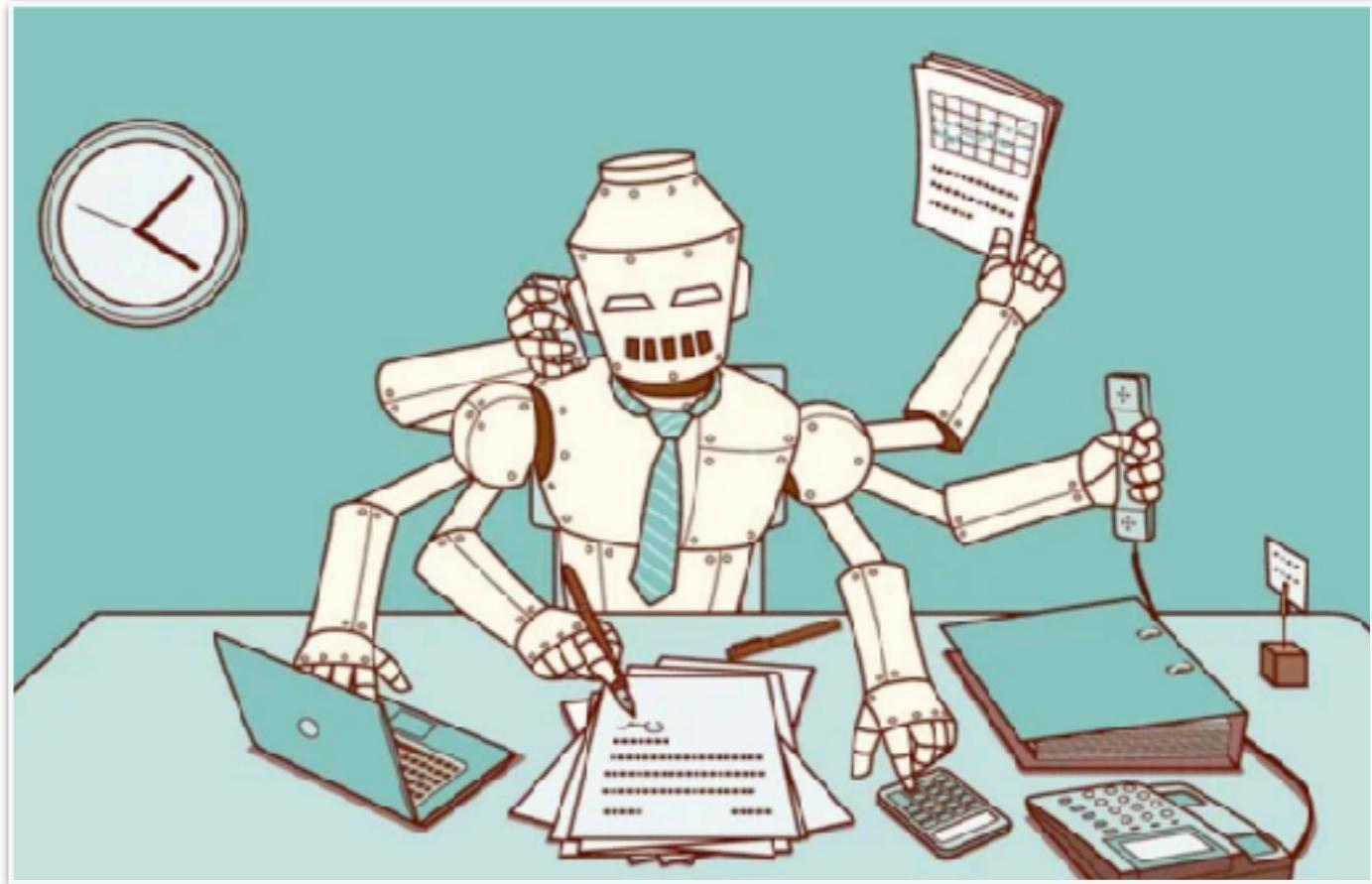
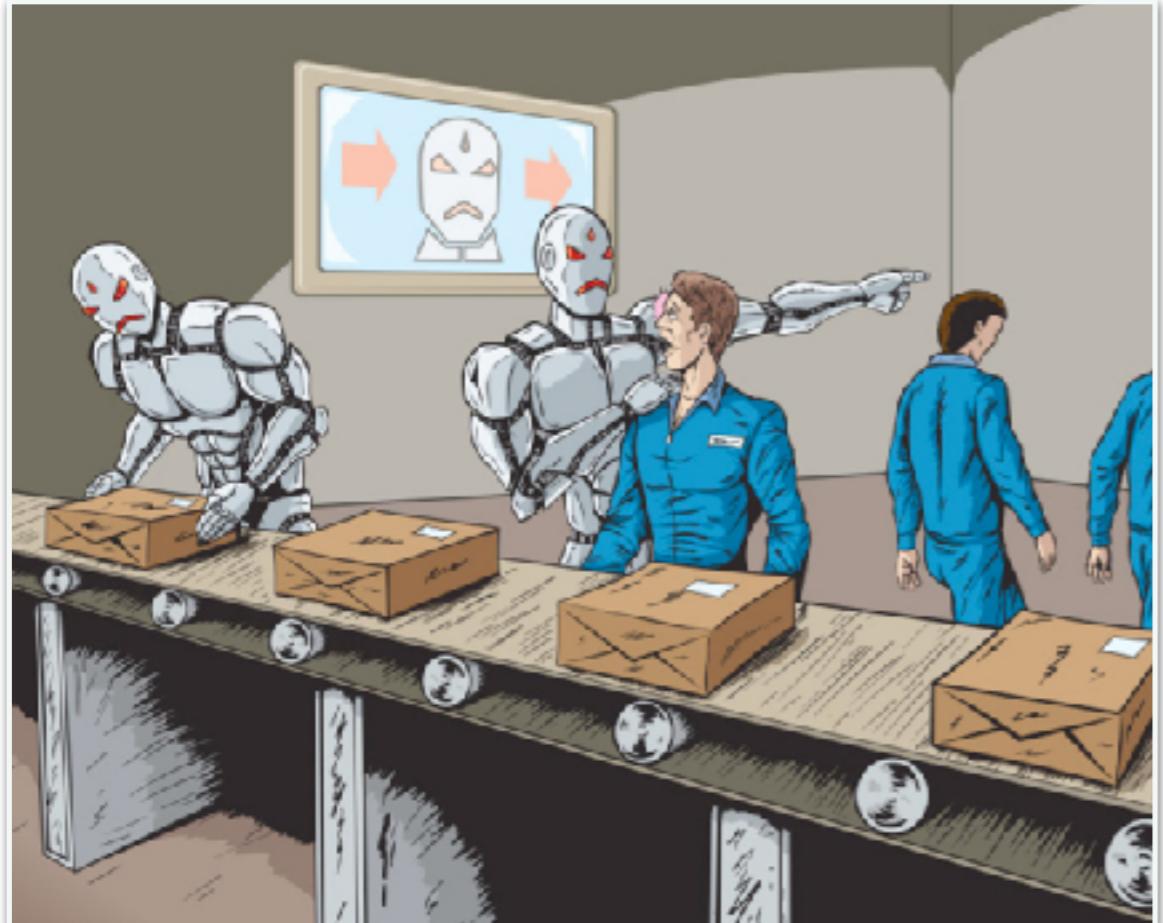
Quels dangers les découvertes scientifiques peuvent apporter à notre société ?

Ces dangers s'appliquent aussi pour l'intelligence artificielle !



Avez-vous en tête des dérives ou dangers potentiels dans la conception et l'utilisation d'intelligences artificielles ?

Difficulté possible - futur du monde du travail



Réalité: chaque progrès technique bouleverse le monde industriel

Conséquence: certains jobs risquent de disparaître progressivement

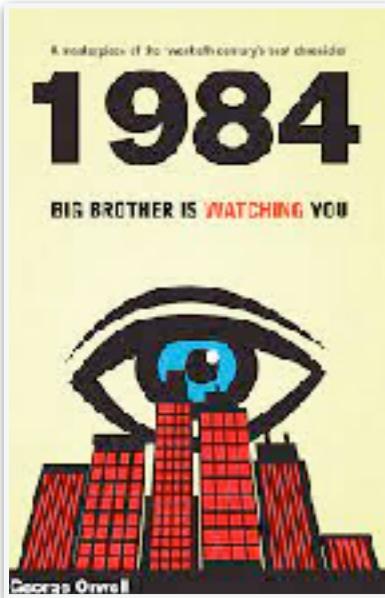
Exemple: Github Copilote peut automatiser une partie de la gestion de code

Attention: tous les outils imparfaits actuellement ne font faire que s'améliorer

Conseil personnel: ayez un profil unique, combinant beaucoup d'aptitudes

Clef du succès futur: savoir s'adapter à de nouvelles situations et apprendre continuellement

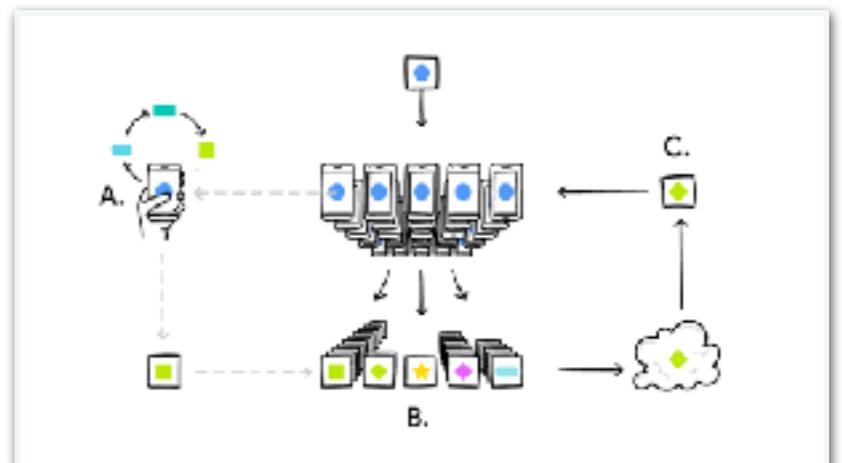
Dérive possible - surveillance, sécurité, et vie privée



Surveillance de la population: identification, trackage, et analyse des personnes

Cybersécurité: augmentation et facilitation des attaques (*phising*, fraude, *ransomware*, etc.)

Vie privée: de plus en plus de données personnelles sont collectées et fuite de données sensible



Exemple: concours Netflix pour prédire des recommandation sur base de données historiques anonymes

Difficulté: des chercheurs ont réussi à retrouver l'identité de personnes

Piste prometteuse: *federated learning* ou l'apprentissage n'est pas centralisé (chacun garde ses données)

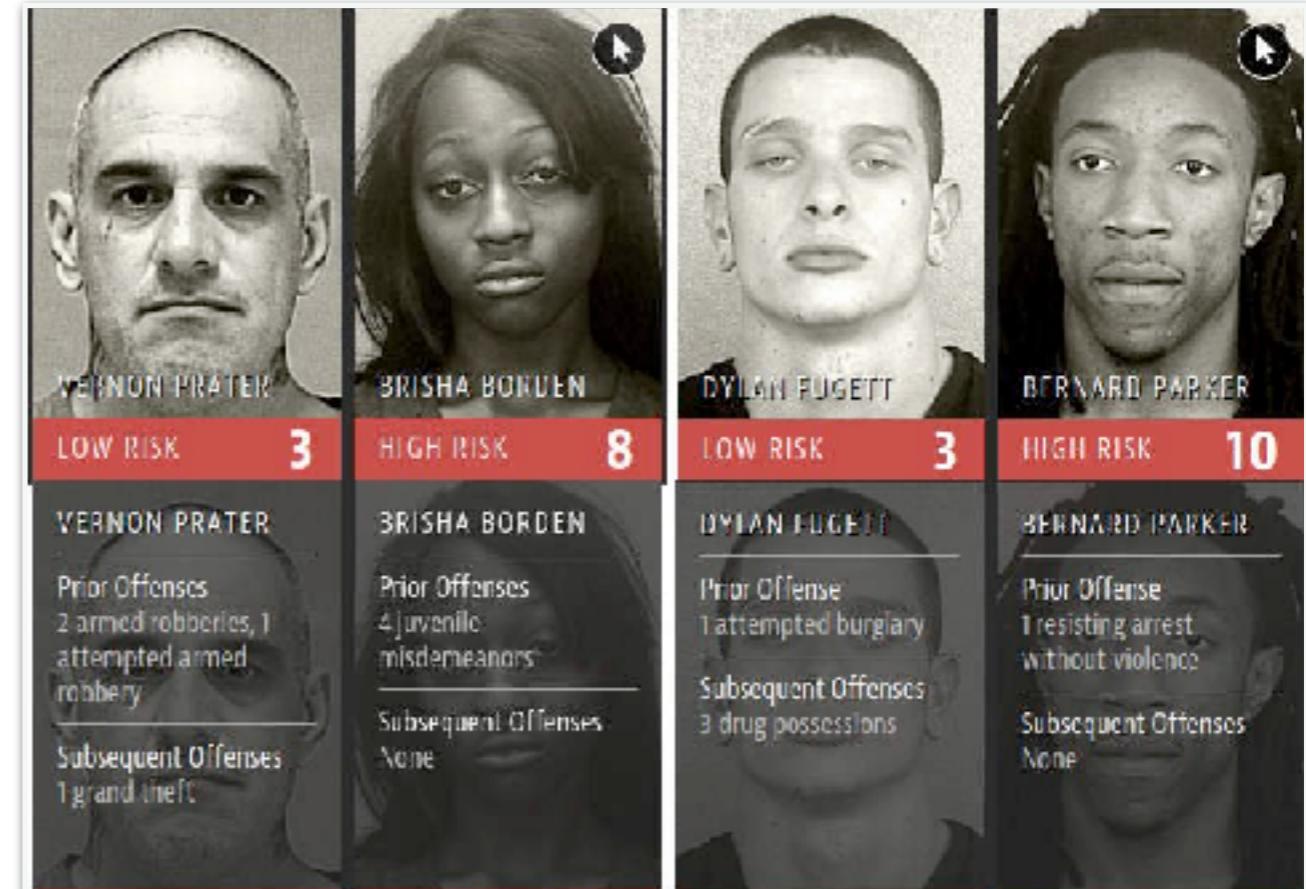
Dérive possible - discrimination automatisée

ARTIFICIAL INTELLIGENCE

LinkedIn's job-matching AI was biased. The company's solution? More AI.

ZipRecruiter, CareerBuilder, LinkedIn—most of the world's biggest job search sites use AI to match people with job openings. But the algorithms don't always play fair.

By Sheridan Wall & Hilke Schellmann June 23, 2021



Objectif de l'algorithme: faire des correspondances entre candidats et offres d'emplois

Nature de la discrimination: l'algorithme avait tendance à prioriser les hommes

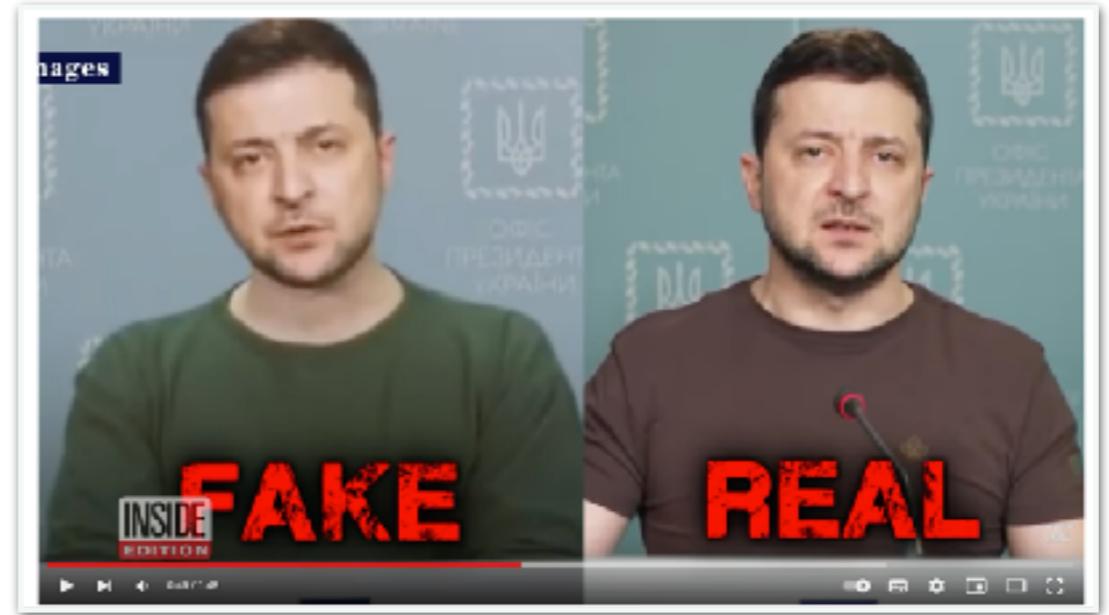
COMPAS: système donnant un score de risque qu'un prisonnier libéré récidive

Utilisation: aider à la décision de libérer ou non un prisonnier

Nature de la discrimination: risque supérieur et injustifié donné pour les personnes de couleur

Equité: assurer que la prédiction ne discrimine pas les personnes via des données sociales et humaines

Dérive possible - confiance et désinformation



DeepFake: vidéo très réaliste d'une personne, mais qui a été générée artificiellement

Danger 1: désinformation en utilisant des personnalités publiques

Danger 2: dans certains contextes, peut être utilisé comme une arme de guerre

Danger 3: diffamation en utilisant votre propre image

Dérive possible - armes mortelles automatiques



Définition de l'ONU: arme mortelle qui vise, sélectionne, et tue un humain sans supervision humaine

Exemple: Le missile Harop, capable de chercher pendant 6 heures une cible satisfaisant un certain critère

Danger alarmant: à l'heure actuelle, on a la technologie pour la conception de ce genre d'armes

Les progrès en AI ne font faire que faciliter ces développements

Vidéo de sensibilisation sur youtube: *Slaughterbots*

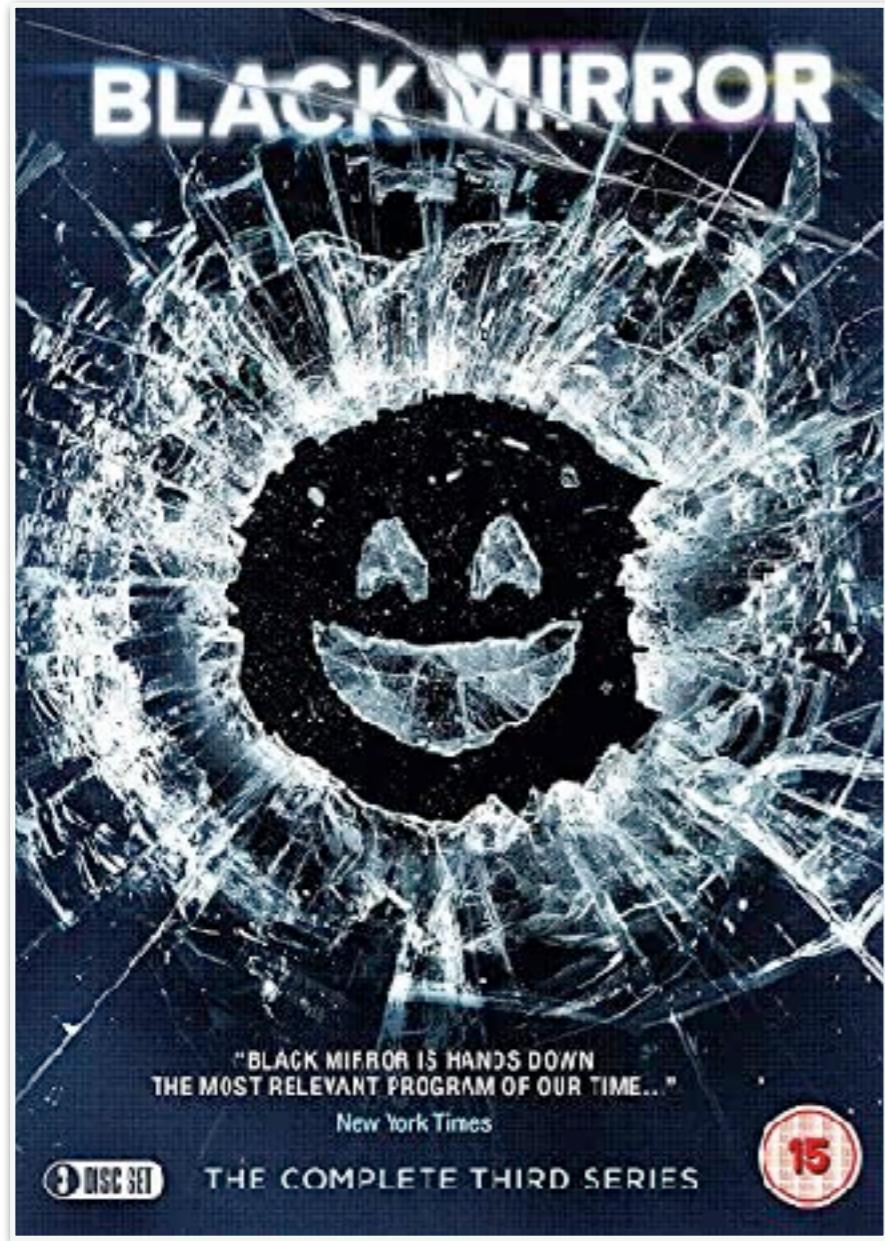
ONU: la majorité des pays membres de l'ONU (mais pas tous) sont favorables à un ban de ces armes



Antonio Guterres (secrétaire général de l'ONU)

« **Autonomous machines with the power and discretion to select targets and take lives without human involvement are politically unacceptable, morally repugnant and should be prohibited by international law.** » (2019)

Dérive possible - autres exemples

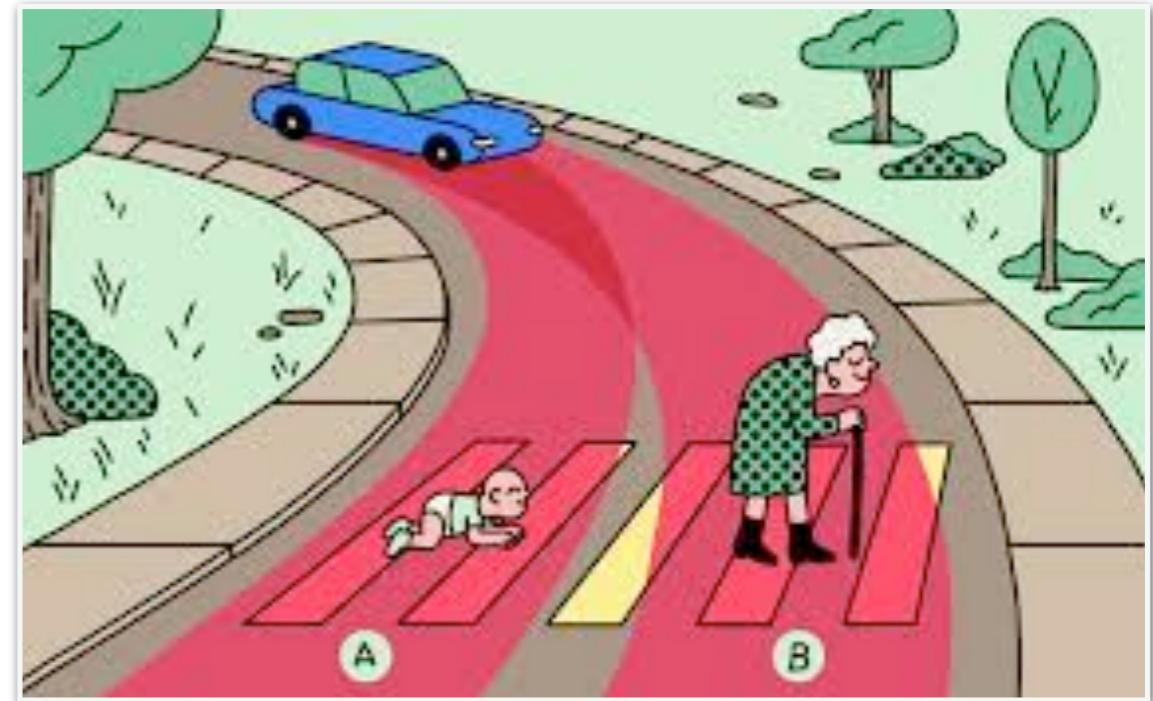
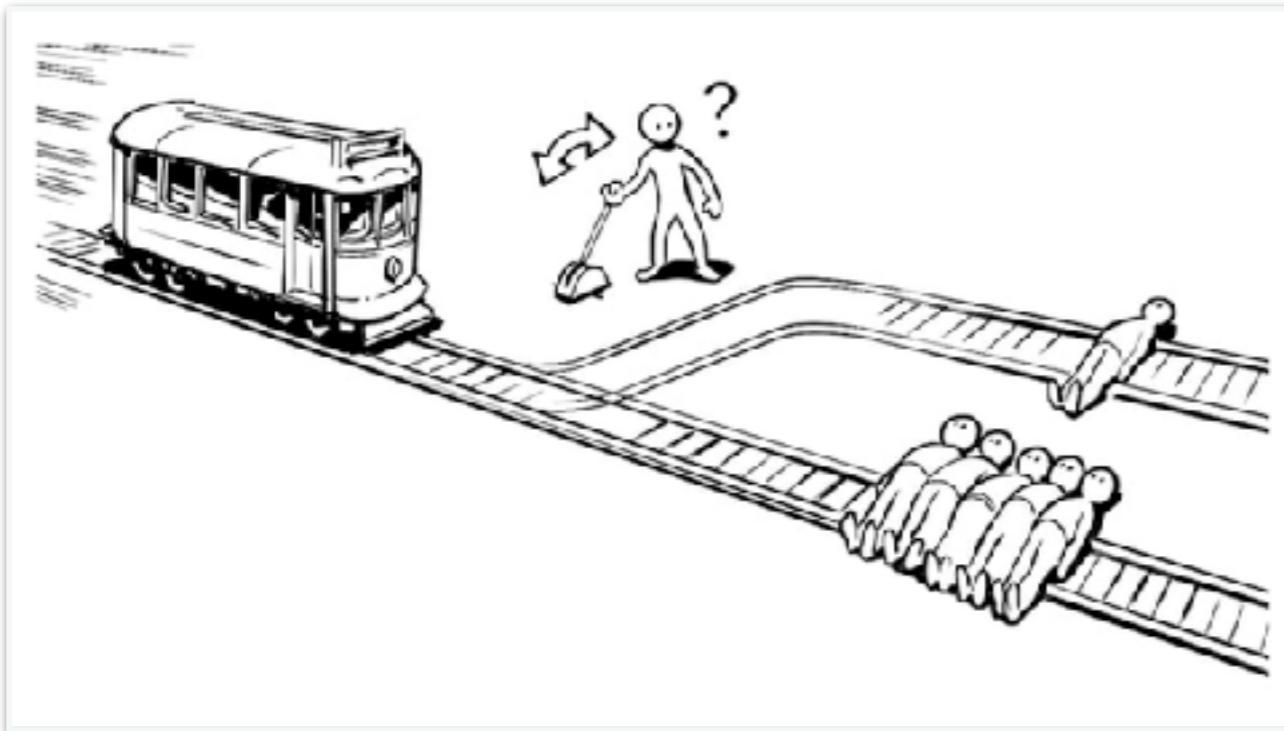


Black Mirror: série montrant différents futurs dystopiques (souvent lié à l'abus de technologies)

Certains épisodes: on a théoriquement toutes les technologies pouvant amener à ces situations !

Upload: comédie liée à l'utilisation de l'IA pour communiquer avec les personnes mortes

Dérive possible - où est le mal et le bien ?

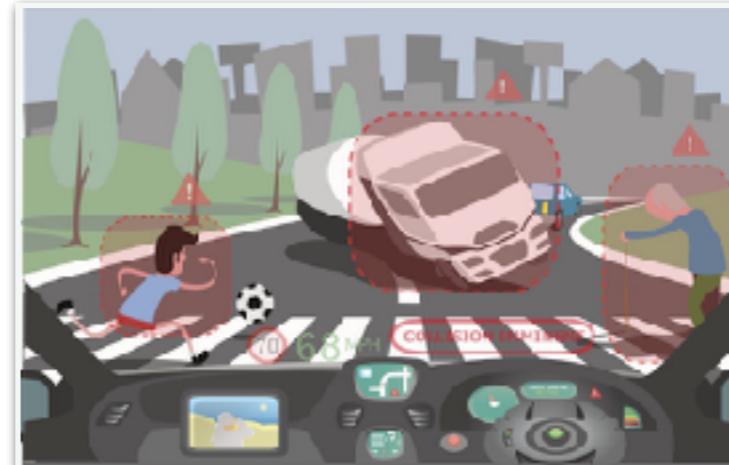
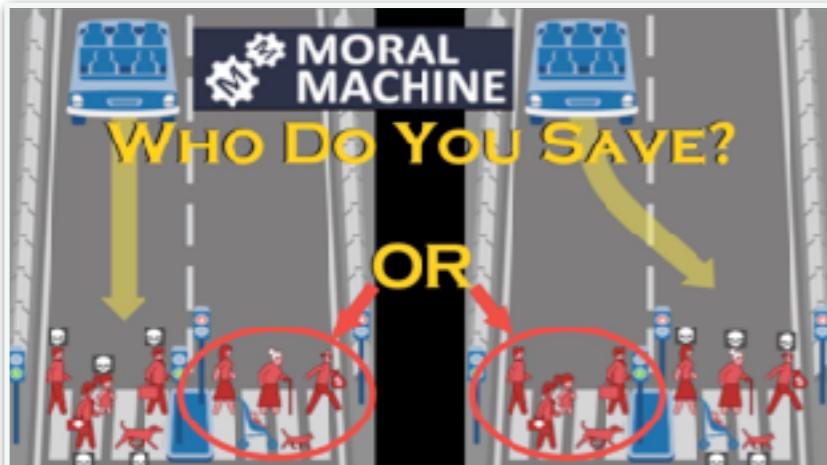


Dilemme du tramway: est-il correct de tuer une personne pour en sauver d'autres ?

Véhicules autonomes: entre deux accidents inévitables, quelle issue l'IA doit-elle prendre ?

On parle de dilemme éthique (situation où la distinction entre bien/mal de chaque décision est floue)

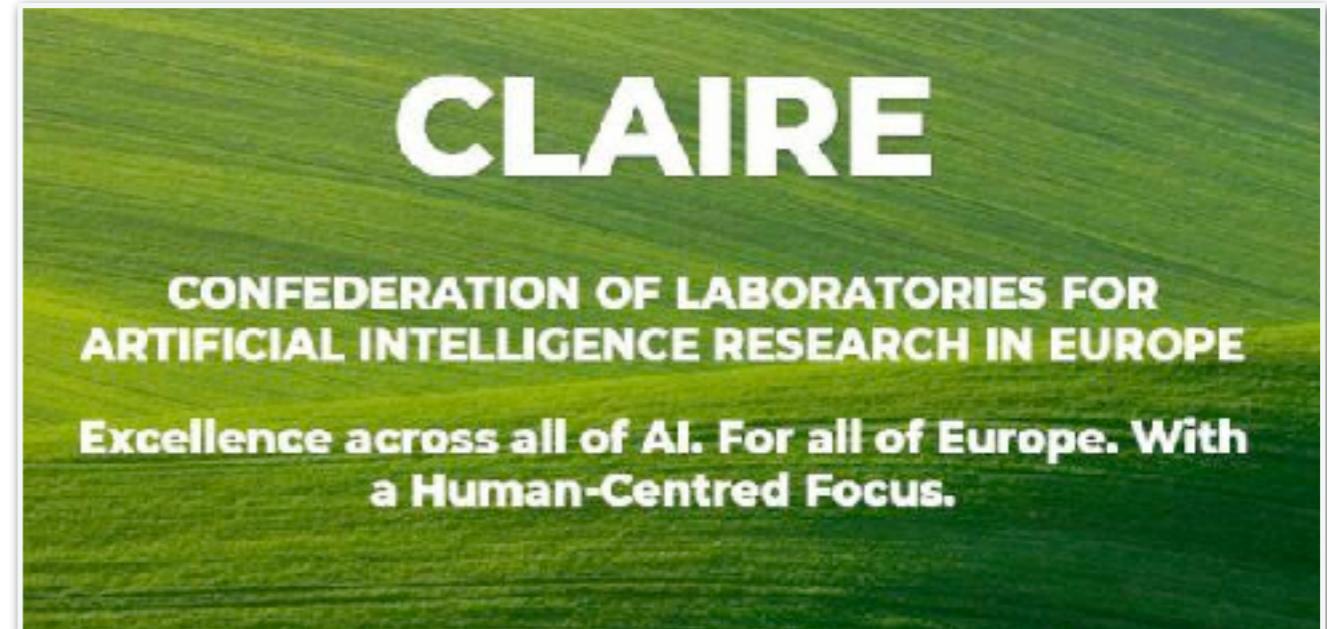
Conception d'IAs: comment gérer ces situations, et à qui revient la responsabilité d'un acte ?



Que peut-on faire actuellement pour éviter un futur sombre ?



Actions et résolutions groupées



De nombreuses actions collectives et mondiales sont menées à ce sujet

Objectif: notamment de codifier les pratiques éthiques de l'intelligence artificielle

Difficulté: tous les pays ne s'entendent pas forcément sur les principes à suivre

Actions que vous pouvez faire



(1) Être conscient des risques et des dangers qui peuvent découler de vos conceptions

Même si vous êtes bien intentionnés, cela ne va pas empêcher les personnes mal-intentionnées

(2) Agir selon votre conscience sur ce qui est bon ou mauvais

(3) Refuser de participer à des projets qui vont à l'encontre de vos principes

Le monde de demain, c'est en partie vous qui le construirez

INF8175 - Intelligence artificielle

Méthodes et algorithmes

**Utilisation en industrie,
éthique, et philosophie: FIN**



POLYTECHNIQUE
MONTRÉAL

Quentin Cappart

Suite de vos aventures...

Objectifs du cours

Objectif 1: comprendre les différentes méthodes et types de raisonnement en intelligence artificielle

Objectif 2: étudier rigoureusement ces méthodes, et vous rendre capable de les utiliser en pratique

Objectif 3: nous munir d'un bagage suffisant pour approfondir les champs selon vos intérêts

Suite possible de votre cursus à Polytechnique en IA

Apprentissage automatique

INF8953DE (Sarah Chandar)

INF8111 (Daniel Aloise)

INF8245E (Sarah Chandar)

Algorithmes de recherche

INF6101 (Gilles Pesant)

INF6102 (Quentin Cappart)

INF8775 (Gilles Pesant)

Problèmes combinatoires

MTH8414 (L.-M. Rousseau)

MTH6404 (Guy Desaulniers)



Applications

LOG6308 (Michel Desmarais)

INF8460 (Amal Zouaq)

LOG8235 (Olivier Gendreau)

INF6804 (G.-A. Bilodeau)

INF6805E (Giovanni Beltrame)

Agents logiques et probabilistes

INF8410 (non donné actuellement)

INF8225 (Christopher Pal)

Aspects logiciels de l'IA

LOG6309E (Heng Li)

Quelques statistiques et remerciements...



Achievement unlocked
LEARN 32 ALGORITHMS



Achievement unlocked
DIGEST ROUGHLY 600 SLIDES



Achievement unlocked
DISCOVER 9 ASPECTS/MODES OF AI



Achievement unlocked
SURVIVE TO 40 HOURS OF RECORDING



Remerciements

(1) Chargés de laboratoires pour leur immense travail réalisé

Temps de permanence, correction rapide, organisation du concours, etc.

(2) Plusieurs références et inspiration pour le contenu du cours



<http://ai.berkeley.edu/home.html>



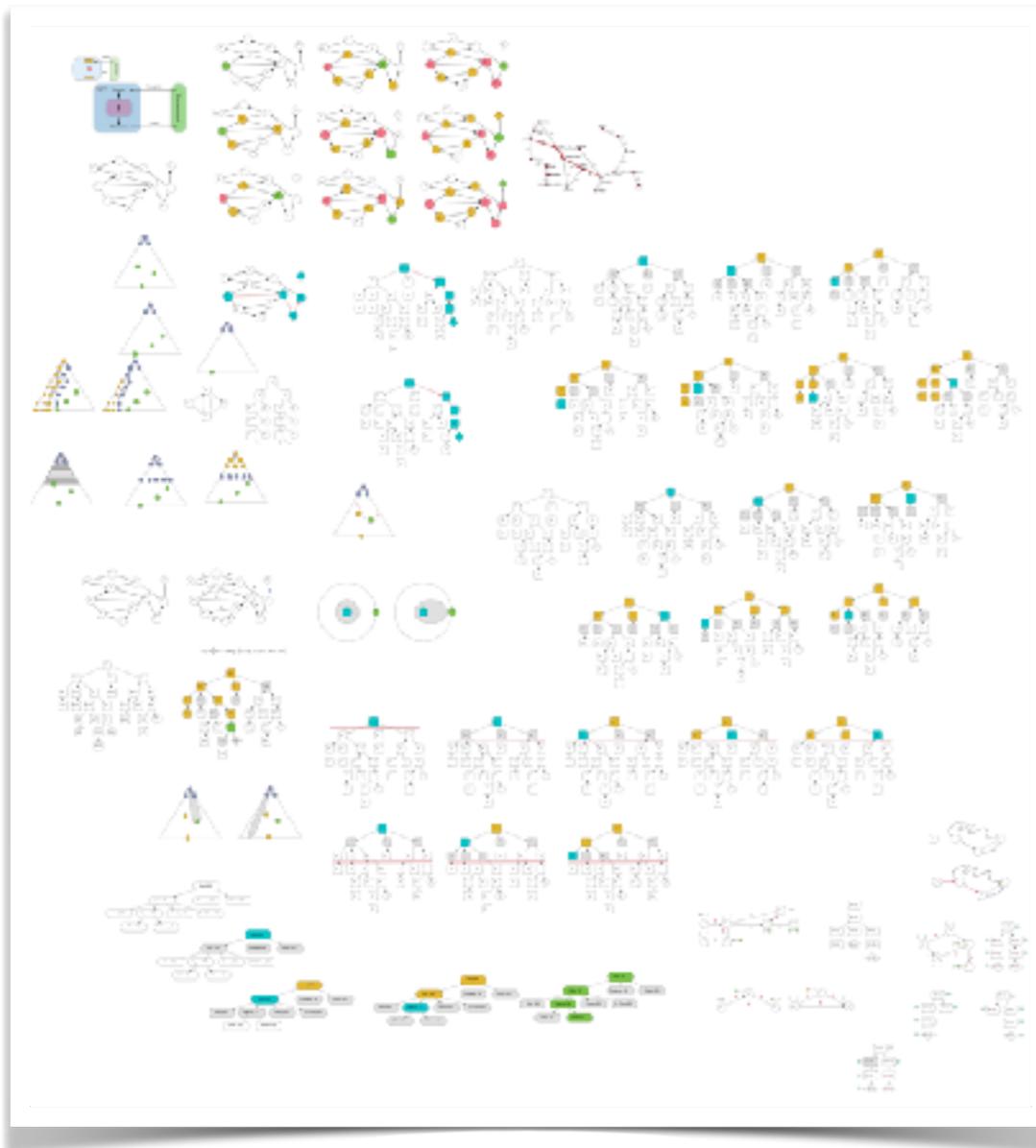
CS221

(3) A toute la classe: pour votre intérêt, motivation, respect et bonne humeur!

(4) Aux anciens chargés et responsable du cours: la construction d'un cours se fait de manière itérative

(5) Meme-generator: pour la centaine de mèmes des slides :-)

Génération de mes schémas

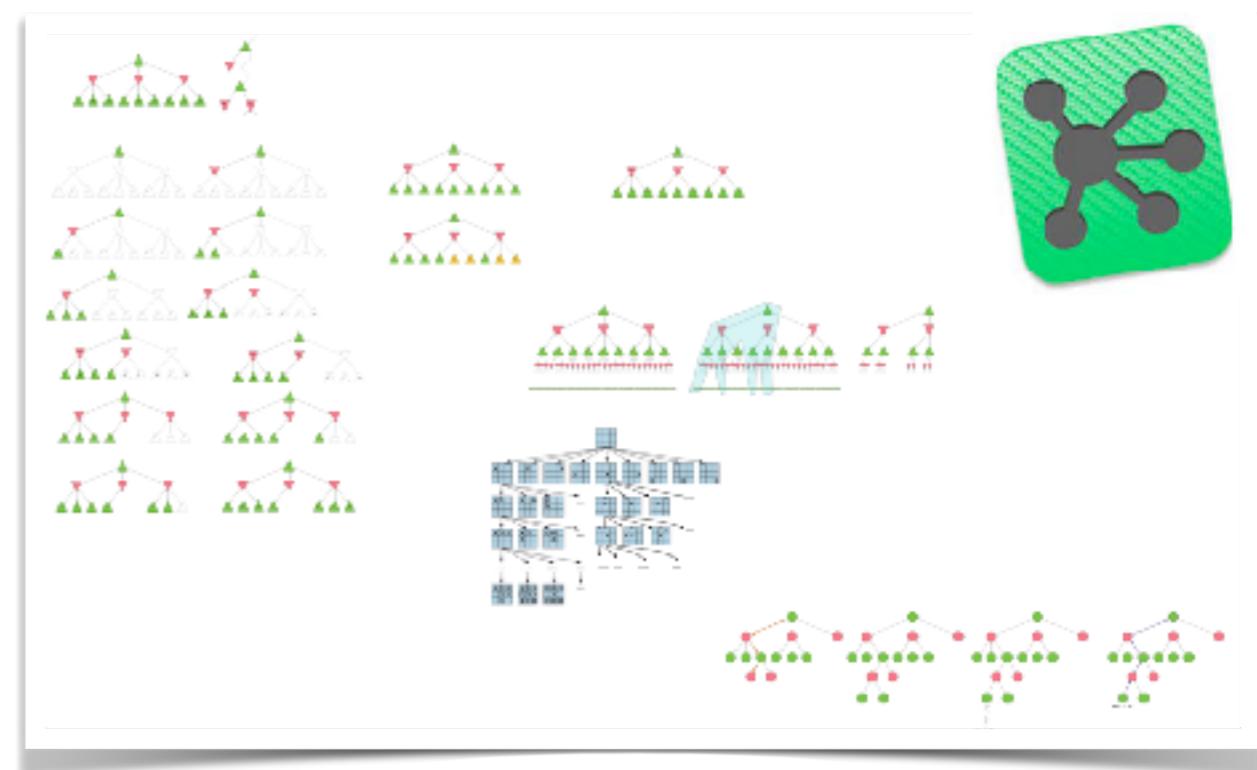
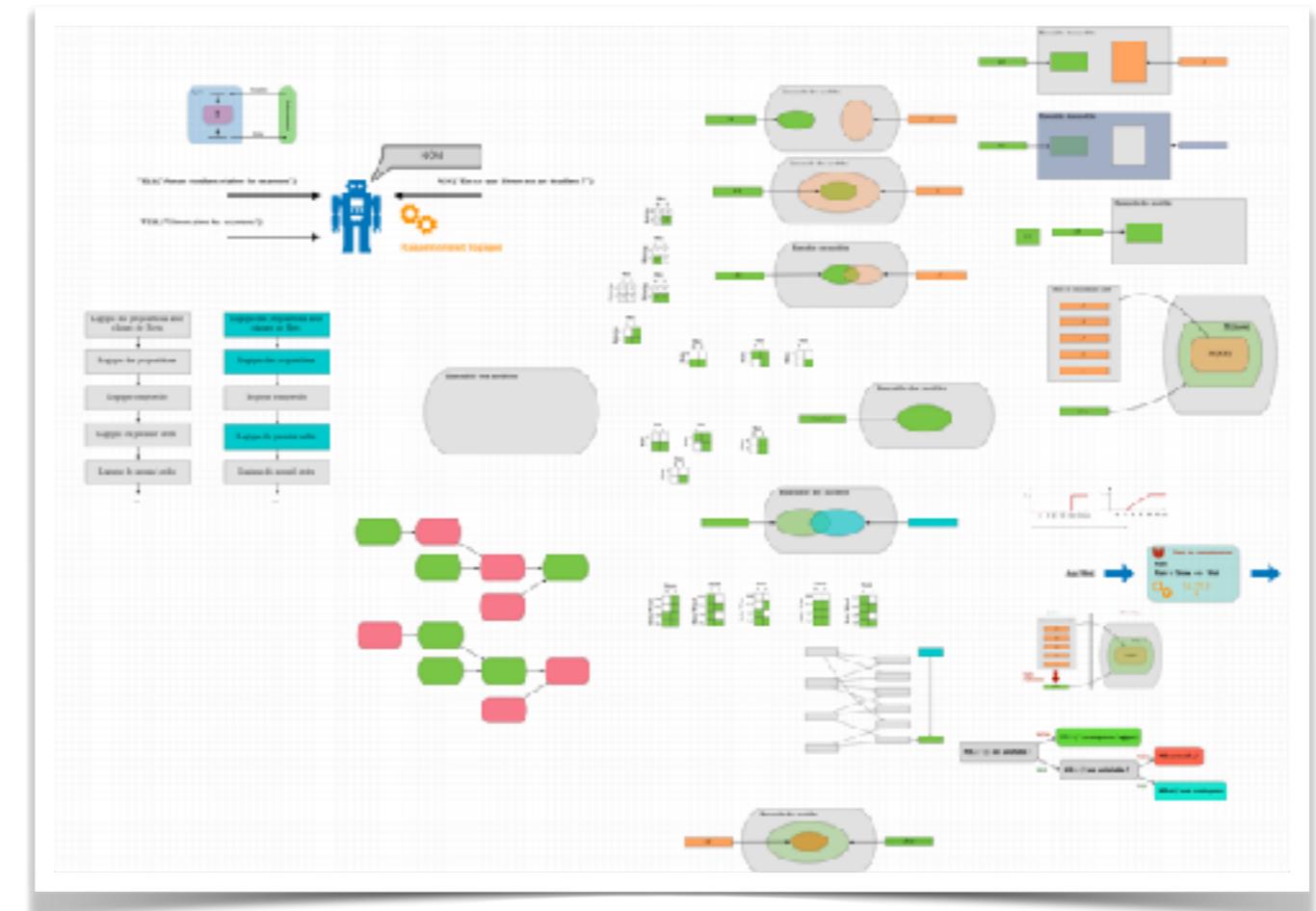


Logiciel: **Omnigraffe**

Avantage: génération aisée de figures

Inconvénient 1: fonctionne que sur Mac OS

Inconvénient 2: malheureusement payant (et cher)





THE END