

L'intelligence artificielle

Définition(s) et idées reçues

Jean-Gabriel Ganascia

Équipe ACASA - LIP6 – CNRS (UMR 7606)

Sorbonne Université – Labex OBVIL

Institut Universitaire de France

Président du comité d'éthique du CNRS

4, place Jussieu, 75252 Paris Cedex 05, FRANCE

Jean-Gabriel.Ganascia@lip6.fr

Synoptique

1. Définition(s) de l'IA
2. Empan
 - i. Perception
 - ii. Mémoire et apprentissage
 - i. Apprentissage profond
 - iii. Raisonnement
 - iv. Fonctions exécutives
 - v. Fonctions communicatives
3. Perspectives actuelles
4. IA et défense
5. Nouvelles vulnérabilités
6. Éthique: jusqu'où?



Une machine peut-elle penser?



Number of Operation	Nature of Operation	Variables acted upon	Result
1	x	$1V_2 \times 1V_3$	$1V_4, 1V_5$
2	-	$1V_4 - 1V_1$	$2V_4$
3	+	$1V_5 + 1V_1$	$3V_4$
4	+	$+2V_4$	$4V_4$
5	+	$1V_5 - 1V_3$	$5V_4$
6	-	$0V_{10}$	$6V_4$
7	-	$1V_5 - 1V_1$	$7V_4$
8	+	$1V_5 - 0V_4$	$8V_4$
9	+	$1V_6 + 1V_7$	$9V_4$
10	x	$1V_{21} \times 2V_{11}$	$10V_4$
11	+	$1V_{12} + 1V_{13}$	$11V_4$

VOL. LIX. No. 236.]

[October, 1950]

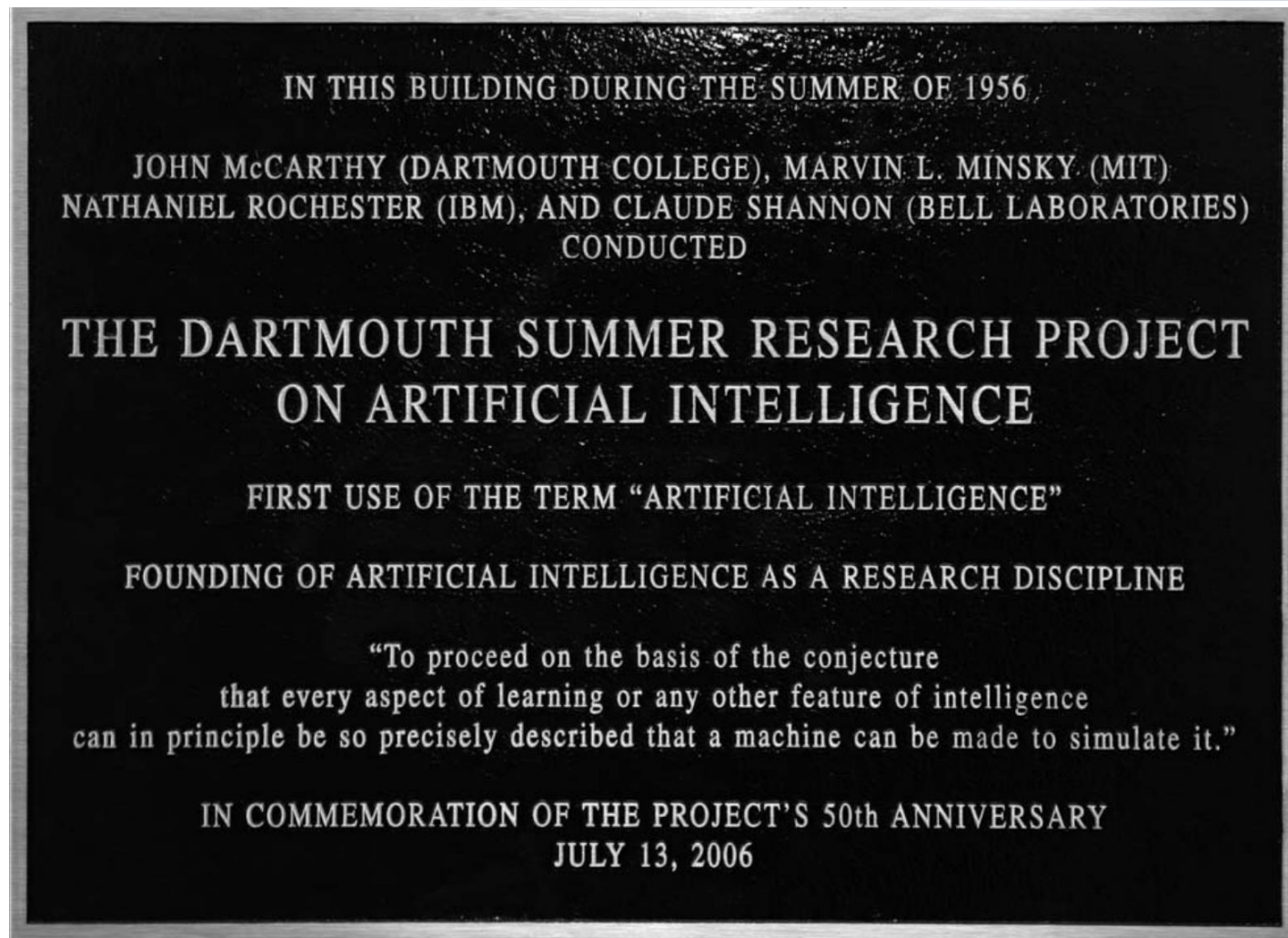
MIND

A QUARTERLY REVIEW
OF
PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

By A. M. TURING

Intelligence artificielle



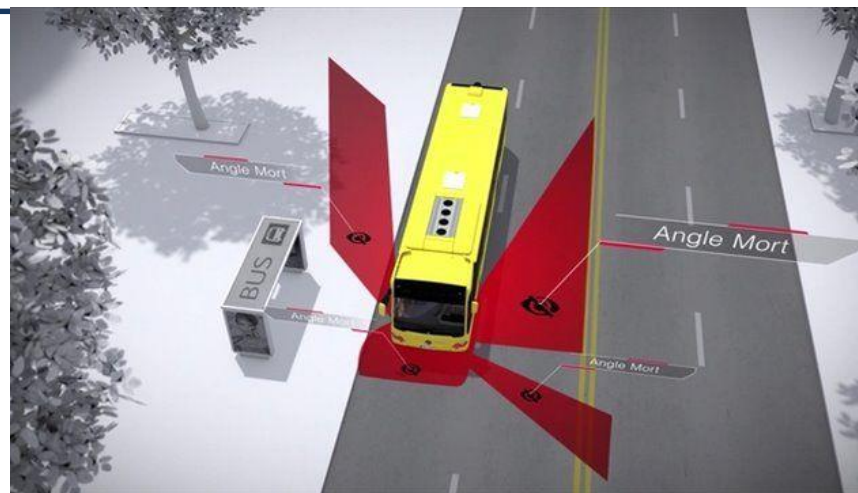
En 60 ans, l'IA a transformé le monde

- Réseaux
 - Darpanet, transpac, ...
 - Internet (1972)
- **Hypertexte**
 - Ted Nelson(1965)
- **Web** – couplage Internet/**Hypertext**
 - Tim Berners Lee 1990
 - **Moteurs de recherche** - Google
- **Web sémantique** – Tim Berners Lee 2004



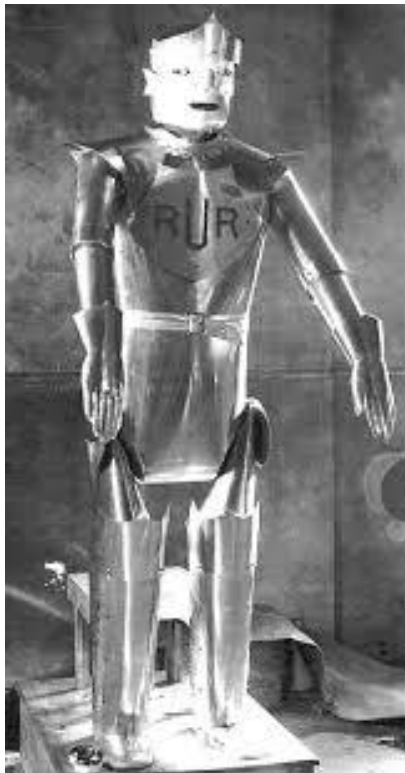
L'IA a changé le monde

- Robots (et bots)
- Véhicules autonomes
- Biométrie
- Vision
- Reconnaissance de la parole
- Traitement et Compréhension du langage naturel
- ...



Malentendu: confusion avec mythes

**Le but de l'intelligence artificielle
n'est pas de recréer un être
intelligent!**



Synoptique

1. Définition(s) de l'IA
2. **Empan**
 - i. Perception
 - ii. Mémoire et apprentissage
 - i. Apprentissage profond
 - iii. Raisonnement
 - iv. Fonctions exécutives
 - v. Fonctions communicatives
3. Perspectives actuelles
4. IA et défense
5. Nouvelles vulnérabilités
6. Éthique: jusqu'où?



Panorama général (« IA @ large »)

Modélisation – 5 fonctions cognitives

1. Les *fonctions réceptives* : elles autorisent l'acquisition, le traitement, la classification et l'intégration de l'information.
2. La *mémoire* et l'*apprentissage* permettant le stockage et le rappel de l'information.
3. Le *raisonnement*, la *pensée*. Cela concerne aussi l'organisation et la réorganisation mentale de l'information ainsi que son utilisation.
4. Les *fonctions exécutives* de prise de décision et d'action.
5. Les *fonctions expressives* qui rendent possible la communication.

1- fonctions réceptives

- Interprétation automatique d'images
 - Reconnaissance des formes, reconnaissance des visages, identification d'objets, etc.
 - Accès par le contenu dans des bases d'images
- Interprétation automatique vidéo
 - Exemple: détection d'intentions ou de vortex en météorologie.
- Traitement de la parole
 - Reconnaissance pour la dictée automatique
 - Reconnaissance de mots ou de thèmes
 - Identification du locuteur, des accents et des langues
- Analyse et fusion d'informations issues de capteurs
 - Multi-modalités, *ex. désambiguïsation de l'image par le son.*
- Perception: construction d'une représentation

Quantité de données - performances

DeepFace: 4.4 millions d'images représentant 1000 images pour 4030 personnes.

Facenet: 200 millions d'images et 8 millions d'identités uniques

- FaceFirst,
- Face-Six,
- DeepFace (Facebook) 97,25%,
- FaceNet (Google) 99,63%!

Exemples d'images pour 6 identités



2- Mémoire et apprentissage

- Représentation des connaissances
- Ontologies – Logiques de description (inspiration psychologique).
- Web – modèle mémoire (hypertexte)
 - Réseaux sociaux → mémoires collectives et dynamiques.
 - Systèmes multi-agents et intelligence artificielle distribuée.
- Extraction de connaissances
 - Algorithmes d'apprentissage supervisé
 - Fouille de données complexes : *séquences, recherche de motifs dans les arbres et dans des graphes (communautés).*

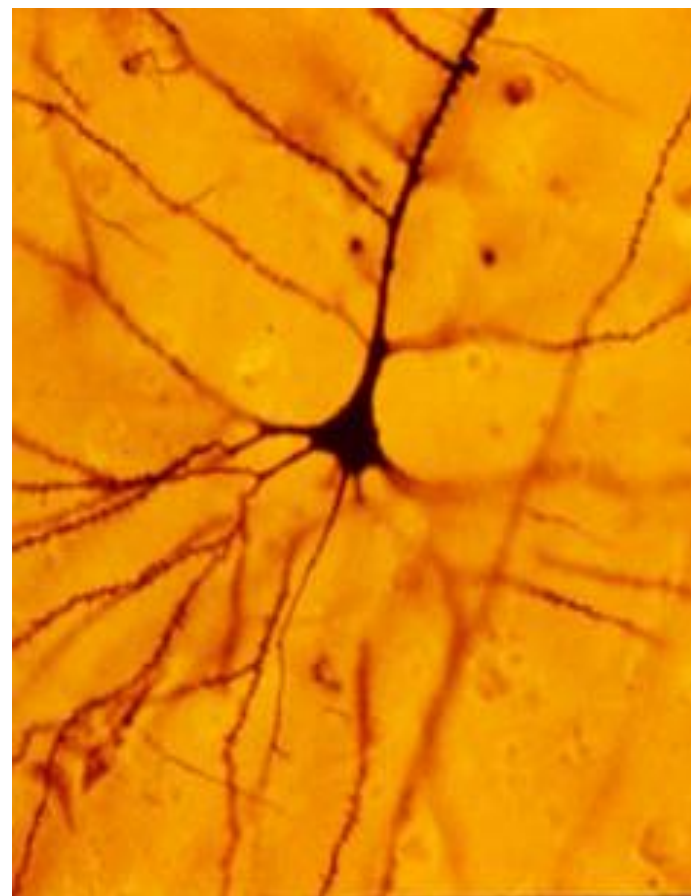
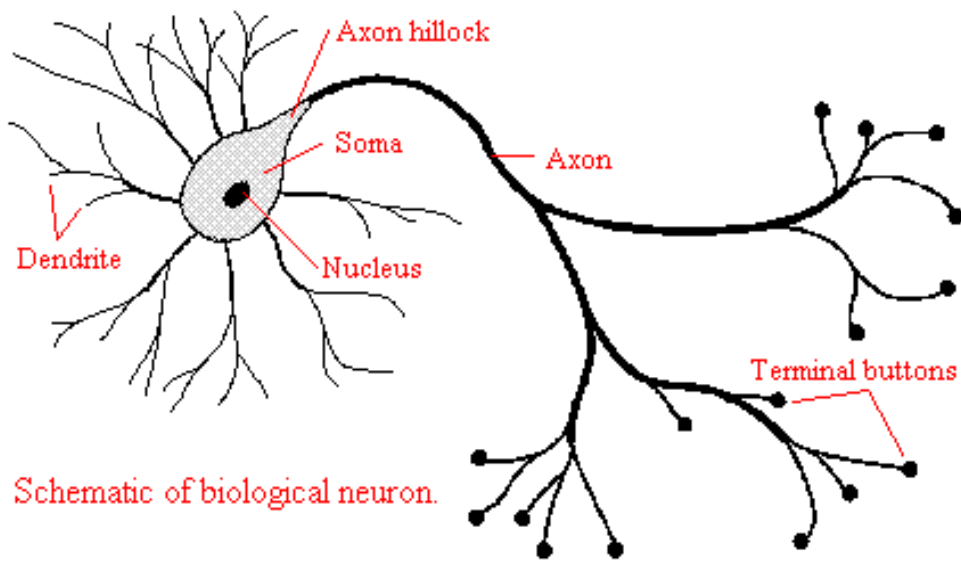
Synoptique

1. Définition(s) de l'IA
2. Empan
 - i. Perception
 - ii. Mémoire et apprentissage
 - i. **Apprentissage profond**
 - iii. Raisonnement
 - iv. Fonctions exécutives
 - v. Fonctions communicatives
3. Perspectives actuelles
4. IA et défense
5. Nouvelles vulnérabilités
6. Éthique: jusqu'où?



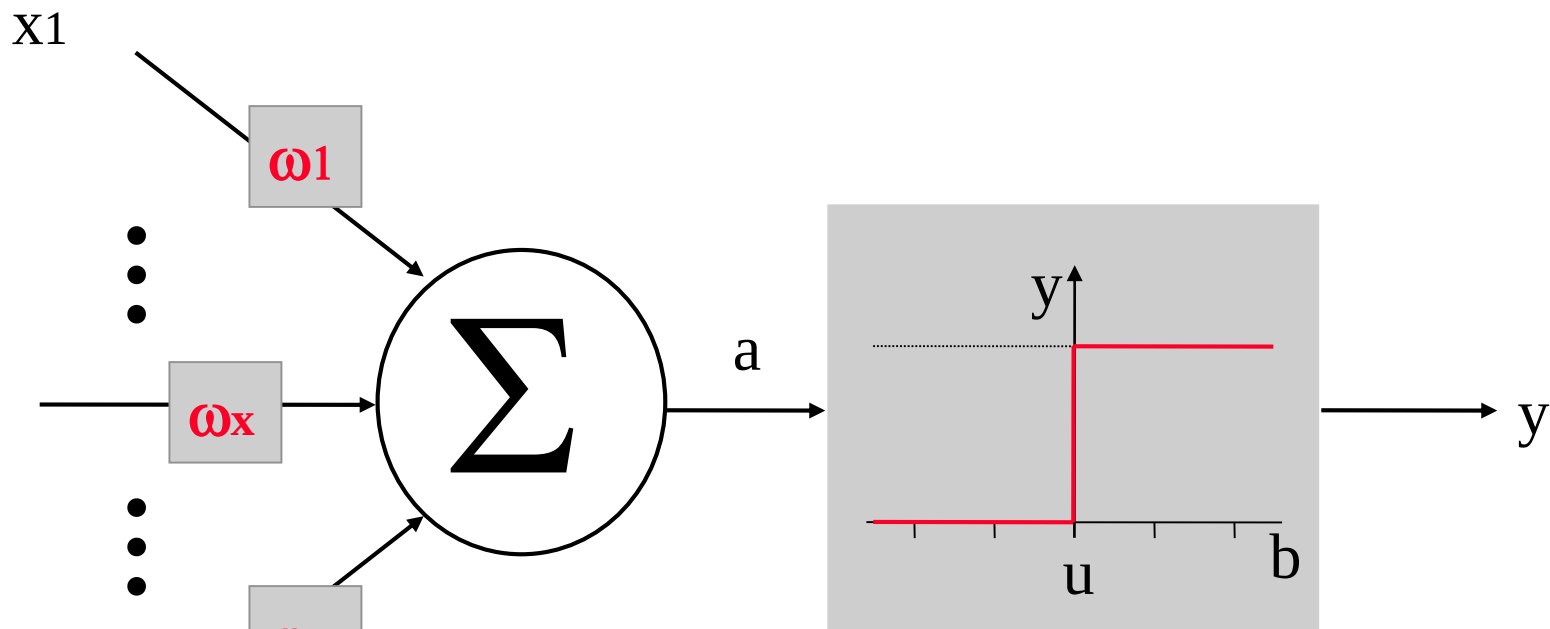
Réseaux de neurones

- Prix Nobel 1906
 - Camilo Golgi
 - Santiago Ramón Y Cajal



Réseaux de neurones

(Mc Culloch and Pitt, 1943) - Cybernétique



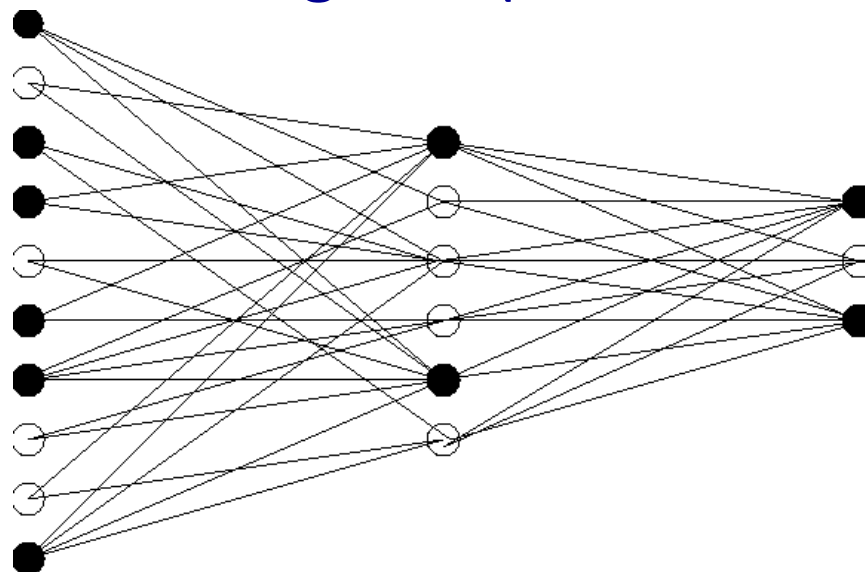
$$y(t+1) = f\left(\sum_{k=1}^n w_k x_k(t) - u\right) = f\left(\sum_{k=0}^n w_k x_k(t)\right)$$

Fonction logique

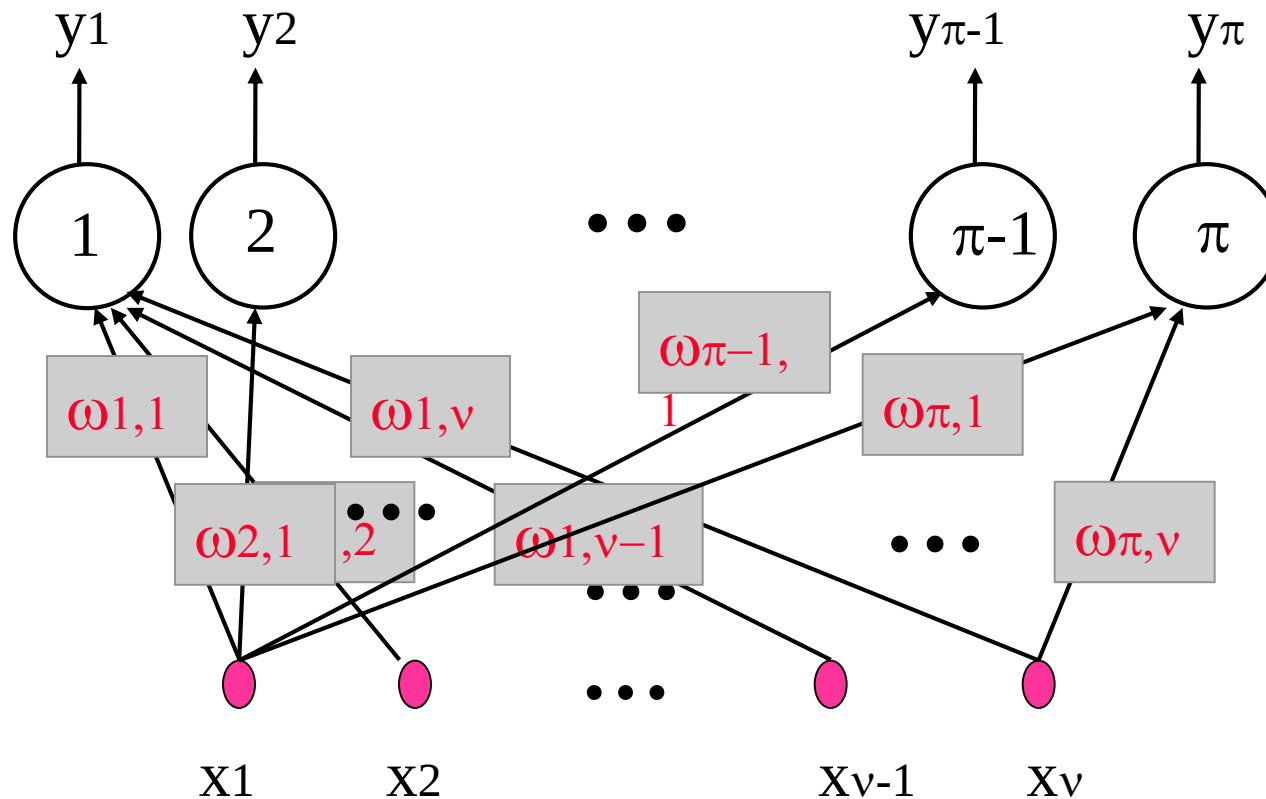
Propriété d'universalité

- Un réseau de neurones à deux couches avec une fonction d'activation en escalier peut réaliser **n'importe quelle fonction booléenne**, à condition que le nombre de neurones de la couche cachée H soit suffisamment grand (*Mc Culloch and Pitts, 1943*)

.



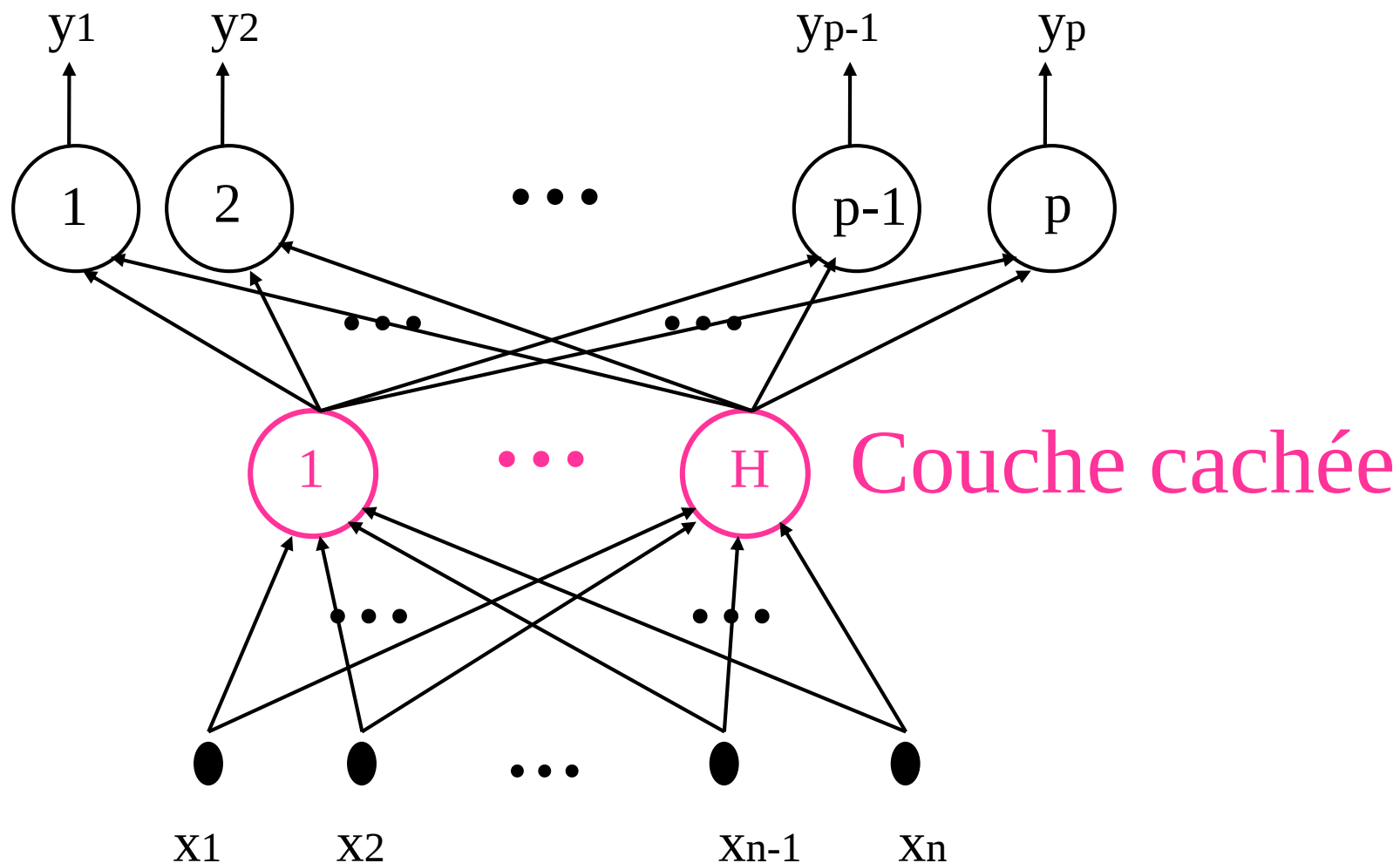
Perceptron – Rosenblatt 1958



$$y_i(t+1) = f\left(\sum_{k=0}^n w_{ik} x_k(t)\right) \quad i = 1, 2, \dots, p$$

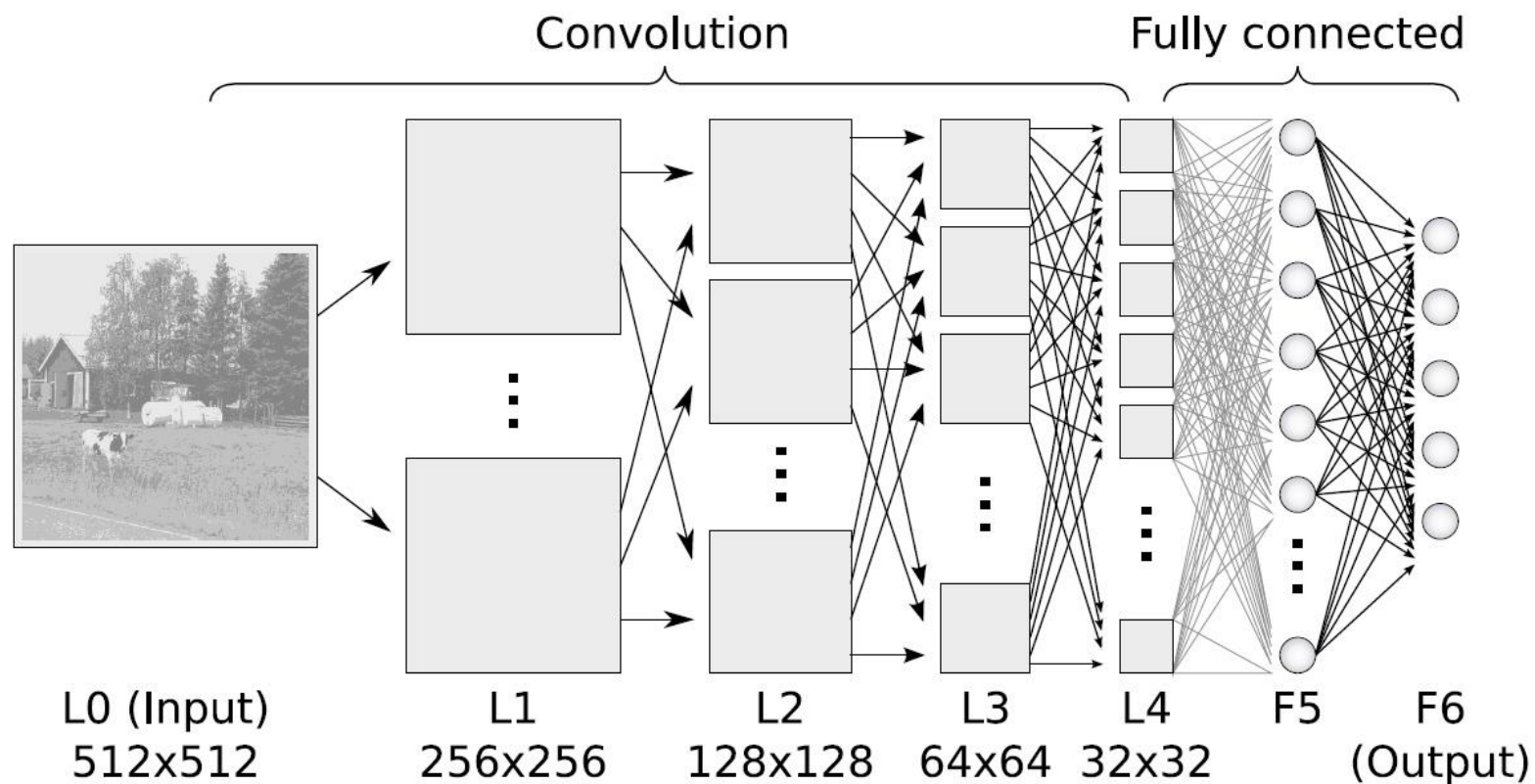
Perceptrons multicouches

Rumelhart 1985



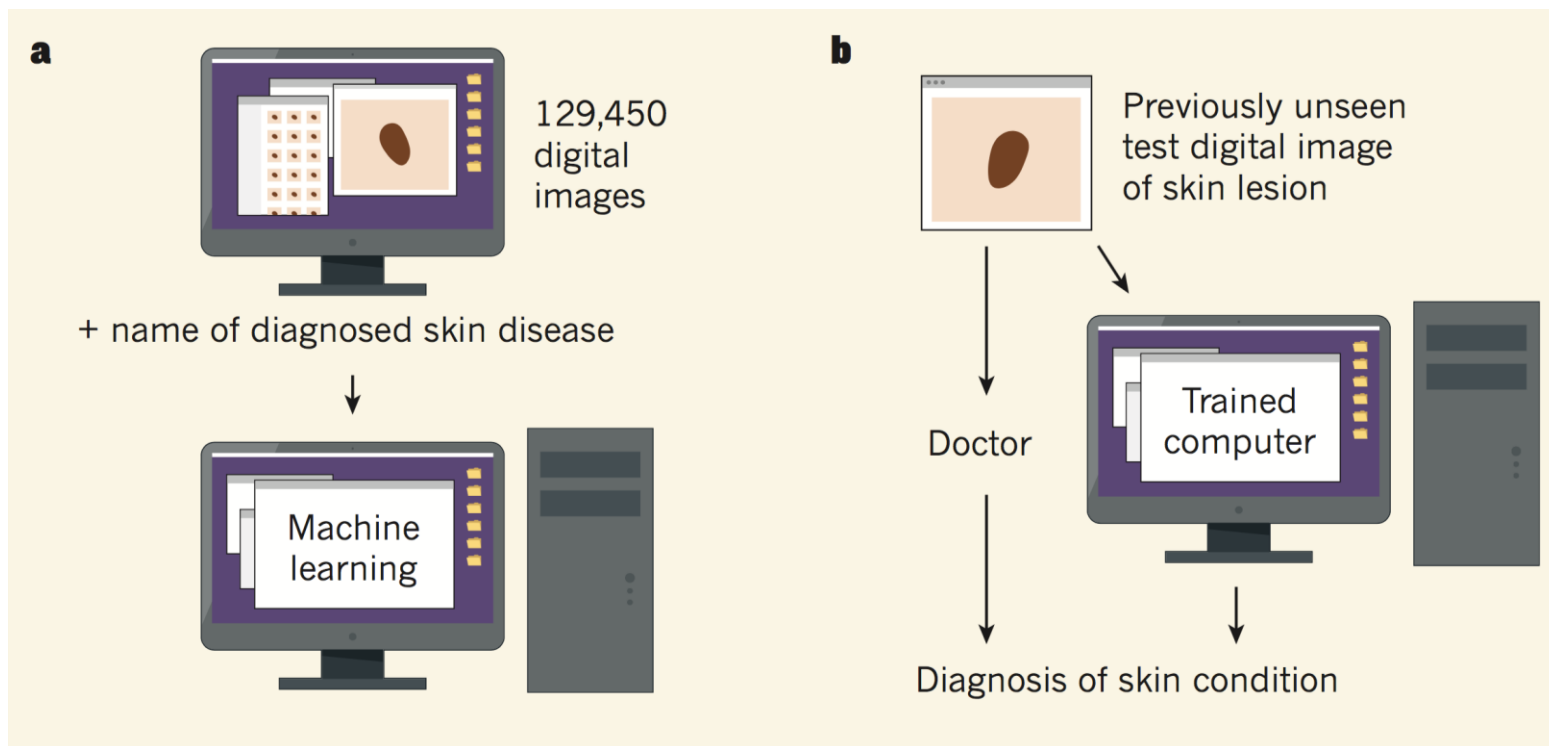
2010 – 3^e renaissance...

- Apprentissage profond – *Deep Learning*



Exemple d'application

Diagnostic des mélanomes



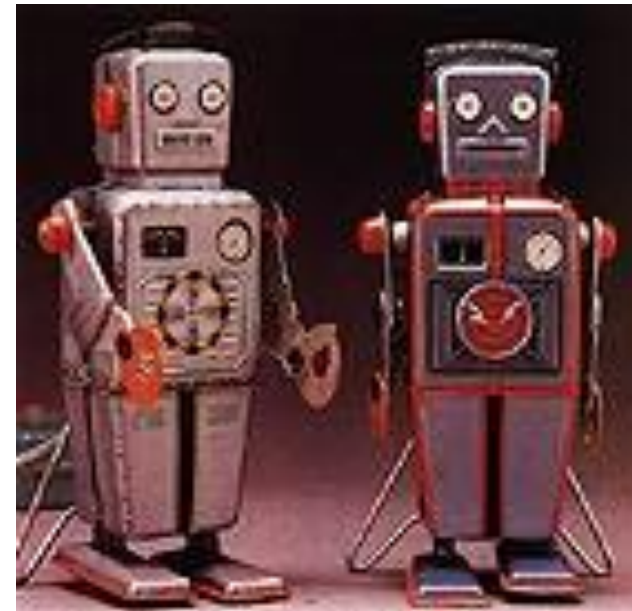
2 FEBRUARY 2017 | VOL 542 | NATURE | 37

3- Raisonnement – pensée

- Inférences
 - Démonstration automatique de théorème
 - Logiques non-monotones, modélisation sens commun
 - Raisonnement automatique.
 - Satisfaction de contraintes
- Logiques épistémiques – multi-agents
 - Intelligence artificielle distribuée
- Retour d'expérience (RETEX)
 - Traçabilité
 - Réflexion
 - Capitalisation des connaissances

4- fonctions exécutives

- Systèmes autonomes et/ou semi-autonomes
 - Autonomie de mouvement et/ou d'action
 - exemple: voiture autonome, armes autonomes, ...
- Aide à la décision
 - Conflit d'autorité
 - Décision assistée
- Robot – planification de tâches



5- Fonctions expressives

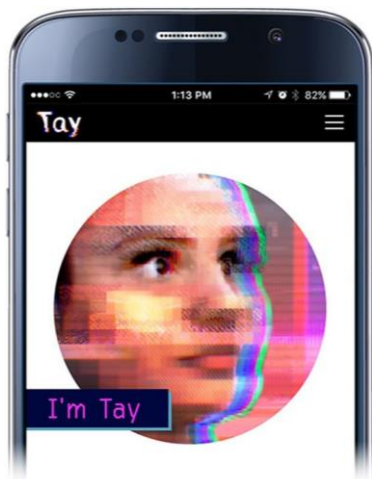
- Langue naturelle (écrit)
 - Traitement du langage
 - Compréhension du langage
 - Système question-réponse
 - *Apprentissage sur d'immenses masses de textes*
- Interfaces homme-machine
 - Dialogue homme-machine (robots bavards, etc.)
 - Informatique affective – avatars émotifs.
- Interfaces cerveau-ordinateur

Synoptique

1. Définition(s) de l'IA
2. Empan
 - i. Perception
 - ii. Mémoire et apprentissage
 - i. Apprentissage profond
 - iii. Raisonnement
 - iv. Fonctions exécutives
 - v. Fonctions communicatives
3. **Perspectives actuelles**
4. IA et défense
5. Nouvelles vulnérabilités
6. Éthique: jusqu'où?



Agents conversationnels - *Chatbots*



INTRODUCING
amazon echo

Always ready, connected,
and fast. **Just ask.**



Your Google Assistant,
always ready to help.



Industrie 4.0 – Hanovre 2011

Changements technologiques (capteurs + IA)

- Masses de données
- Apprentissage machine
- Anticipation
- Traitement en temps réel

exemple « logistique 4.0 »:

- calcul du stock en temps réel (sans attendre l'inventaire)
- commande en temps réel
- anticipation des évolutions du stock et dimensionnement des entrepôts et des transports



FinTech – livre blanc

pôle de compétitivité finance innovation

- Banque:
 - Risque
 - Relation clients
(traitement courriels, conseil en investissement, ...)
 - Recherche d'informations
(en particulier juridiques)
- Assurances:
 - Etablissement des primes (nécessité de disposer d'informations personnelles...)
 - Nouveaux risques et nouveaux objets à assurer



Synoptique

1. Définition(s) de l'IA
2. Empan
 - i. Perception
 - ii. Mémoire et apprentissage
 - i. Apprentissage profond
 - iii. Raisonnement
 - iv. Fonctions exécutives
 - v. Fonctions communicatives
3. Perspectives actuelles
4. **IA et défense**
5. Nouvelles vulnérabilités
6. Éthique: jusqu'où?



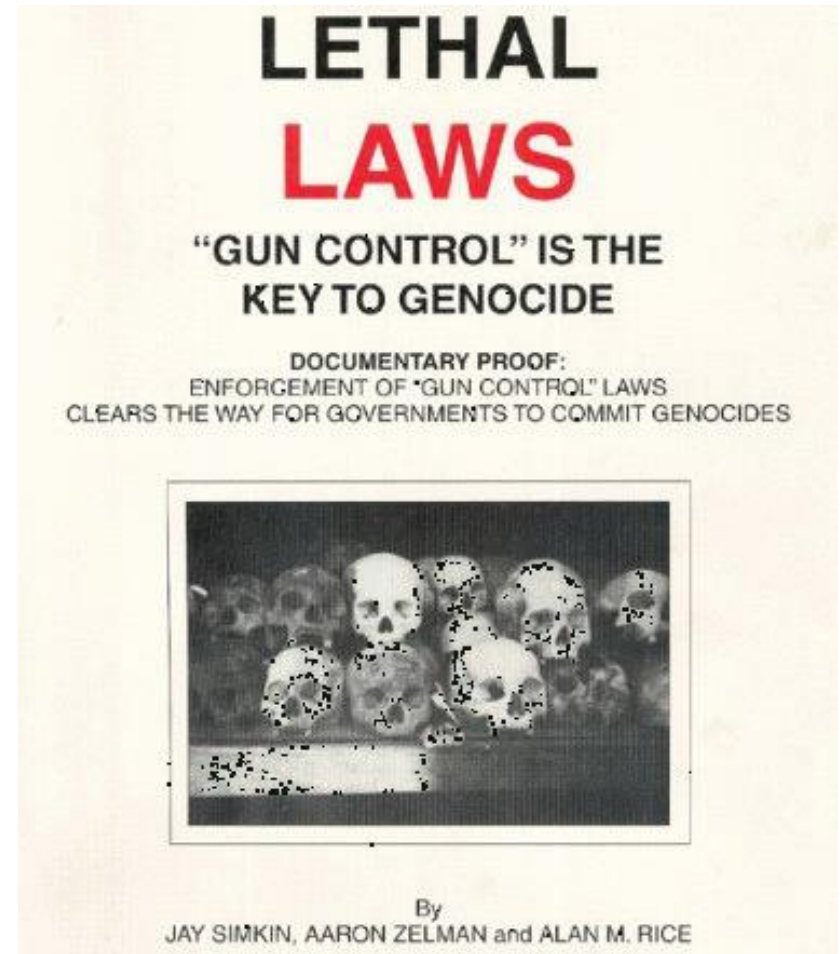
IA & défense: applications

- **Reconnaissance et sécurité** - biométrie (empreintes digitales, reconnaissance faciale et vocale, etc.)
- **Interaction homme-machine** – L'homme face à l'automatisation.
- **Automatisation d'armes et de systèmes de soutien** – drones, robots, SALAs, etc.



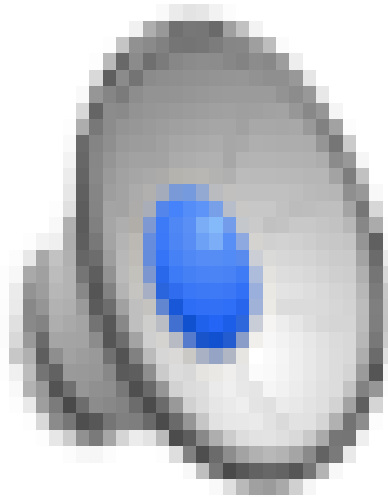
SALA – LAWS – ils nous tuent!

- Lethal Autonomous Weapon Systems (LAWS)
- Systèmes d'armes létales autonomes (SALA)
- Doit-on les encourager ou les condamner et imposer un moratoire?



Lethal Autonomous Weapon Systems

- Doit-on les encourager ou les condamner?



Moratoire sur les « robots tueurs »



[UK](#)
[world](#)
[sport](#)
[football](#)
[opinion](#)
[culture](#)
[business](#)
[lifestyle](#)
[fashion](#)
[environment](#)
[tech](#)
[travel](#)

[browse all sections](#)

[home](#) > [tech](#)

[Robots](#)

Elon Musk leads 116 experts calling for outright ban of killer robots

Open letter signed by Tesla chief and Alphabet's Mustafa Suleyman urges UN to block use of lethal autonomous weapons to prevent third age of war



A killer robot from the 2014 remake of RoboCop. The open letter read: 'lethal autonomous weapons will permit armed conflict to be fought at a scale greater than ever, and at timescales faster than humans can comprehend.'
Photograph: Allstar/Studio Canal/Sportsphoto Ltd./Allstar

[f](#) [t](#) [e](#) [...](#)

This article is 3 months old

< 53,095

[Samuel Gibbs](#)

Sunday 20 August 2017 15.01 BST

Most popular



Argentina: grief and anger after 'explosion' near missing submarine



Cologne 1-0 Arsenal: Europa League - as it happened



Zimbabwe grants Robert Mugabe immunity from prosecution



Trump hails 'invisible' plane in remarks to coast guard: 'The enemy cannot see it'

Robots soldats

- Enfin, des guerres propres!
- Contrôleur éthique:
 - encodage des règles de la guerre juste.
- Est-ce possible?

Ron Arkin



Just War Theory

Jus ad bellum:

- Règles d'engagement (*proportionalité*)

Jus in bello:

- *discrimination* entre civiles et militaire (pb. guerres asymétriques)
- Exceptions à la discrimination

Jus post bellum



IA & défense: applications

- **Reconnaissance et sécurité** - biométrie (empreintes digitales, reconnaissance faciale et vocale, etc.)
- **Interaction homme-machine** – L'homme face à l'automatisation.
- **Automatisation d'armes et de systèmes de soutien** – drones, robots, SALAs, etc.
- **Aide au commandement opérationnel et à l'entraînement** – visualisation
- **Cyber Défense et renseignement** – analyse automatique d'images, analyse de signaux, analyse de textes, analyse vocale ; recherche sur de très grands réseaux.



Synoptique

1. Définition(s) de l'IA
2. Empan
 - i. Perception
 - ii. Mémoire et apprentissage
 - i. Apprentissage profond
 - iii. Raisonnement
 - iv. Fonctions exécutives
 - v. Fonctions communicatives
3. Perspectives actuelles
4. IA et défense
5. **Nouvelles vulnérabilités**
6. Éthique: jusqu'où?



Données et information – vulnérabilités

- Données-information comme **ressources**

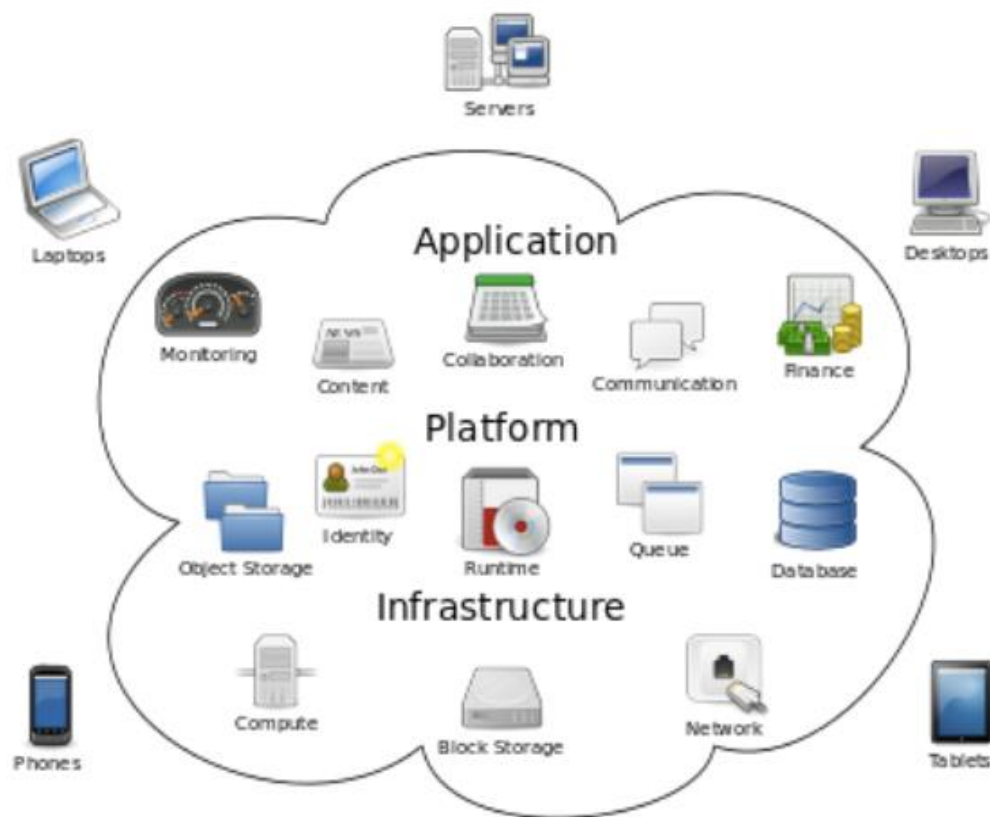
- Disponibilité,
- Correction,
- Accessibilité
- Fracture numérique
- « Infobésité »
- Confiance



- **Manipulation et diffusion** des données et de l'information:

- Diffamation
- Faux témoignages
- Plagiat, réutilisation
- Publicité
- Propagande
- Désinformation
- Calomnie
- Tromperie
- Mensonge
- ...

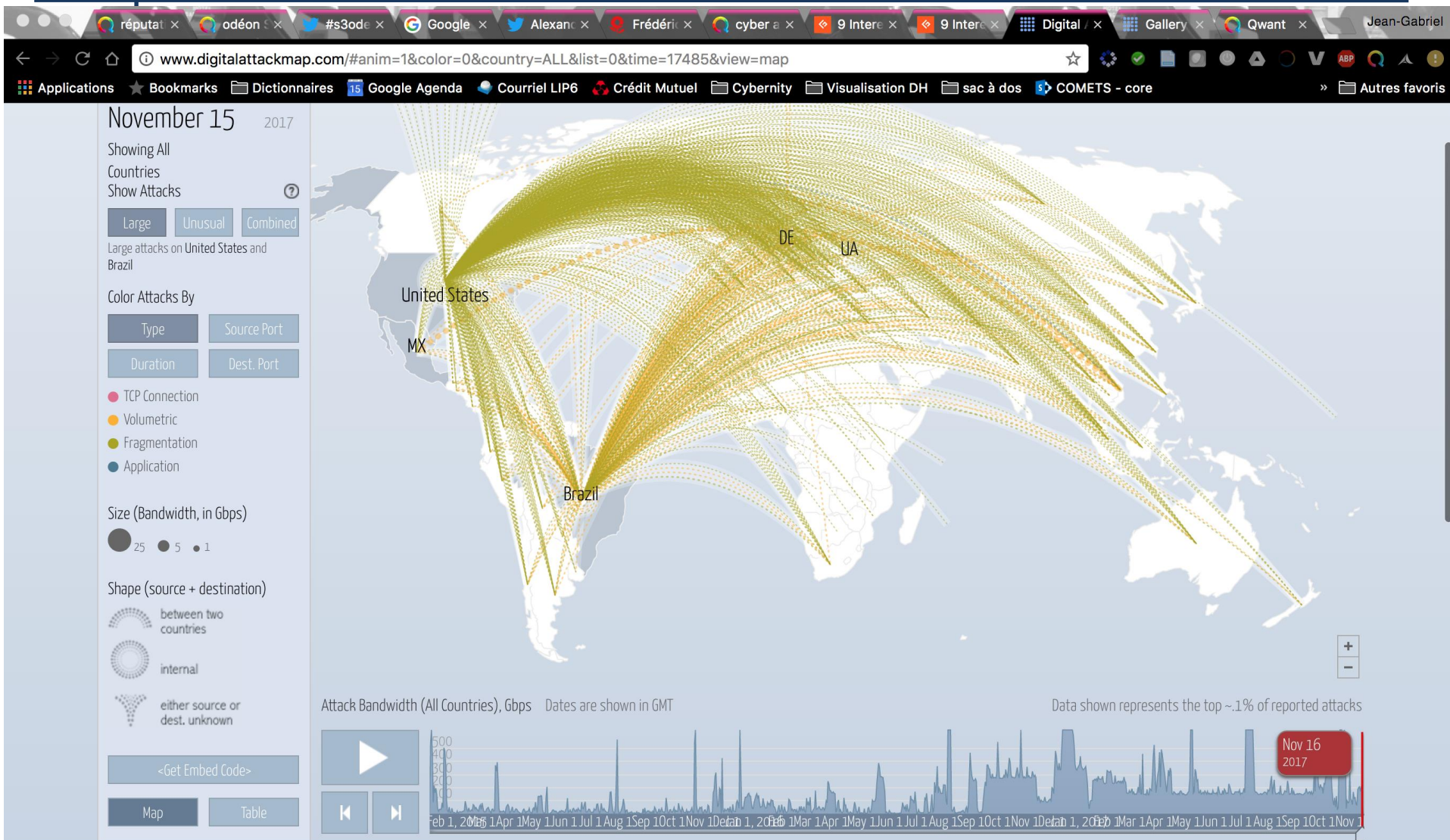
« Cloud computing » Objets connectés



- Distribution de l'information et du calcul
- Ressources partagées
- Réseaux haut débit

→ **Vulnérabilité accrue**

Carte des cyber-attaques



Sources des attaques

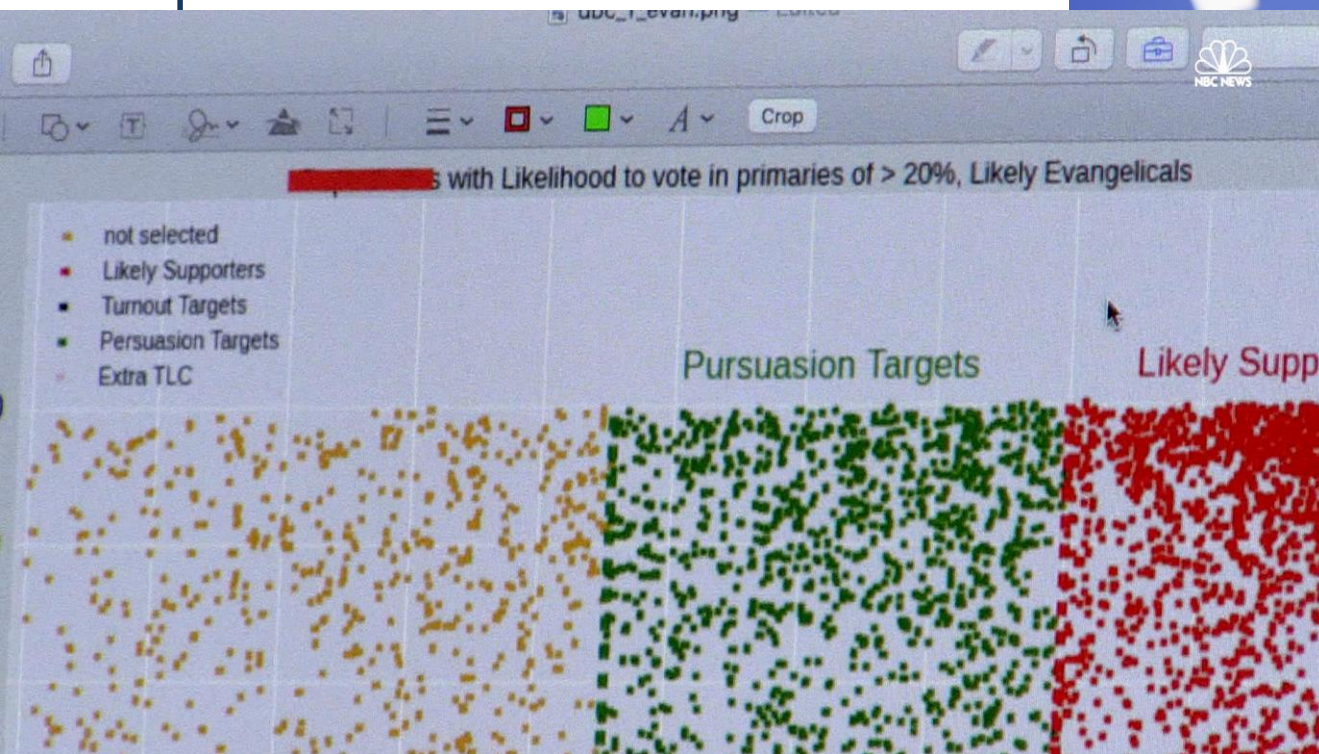
	Script Kiddies	Hacktivistes	Int. économ./maffias	Piratage d'états
Nombre/technicité	++++/+	+++/>++	++/>+++	+/>++++
Motivation	Prouesse, virtuosité	Message militant	Economique et financière	Politique et économique



IA et cybersécurité

- **Détection d'attaques**
 - Utilisation d'apprentissage machine.
 - Agrégation de signaux faibles (analyse et identification)
- **Réaction aux attaques**
 - Contre-attaques défensives et/ou offensives...
 - Questions éthiques et politiques
- **Création de failles de sécurité**
 - Cyber attaques utilisant du machine learning.
 - Détection de patterns et de motifs.
- **Leurres, biais et manipulations**
 - Prise de décision sur la base de signaux manipulés par des informations erronées...

Nouvelles vulnérabilités



Nouvelles vulnérabilités: deep fake!

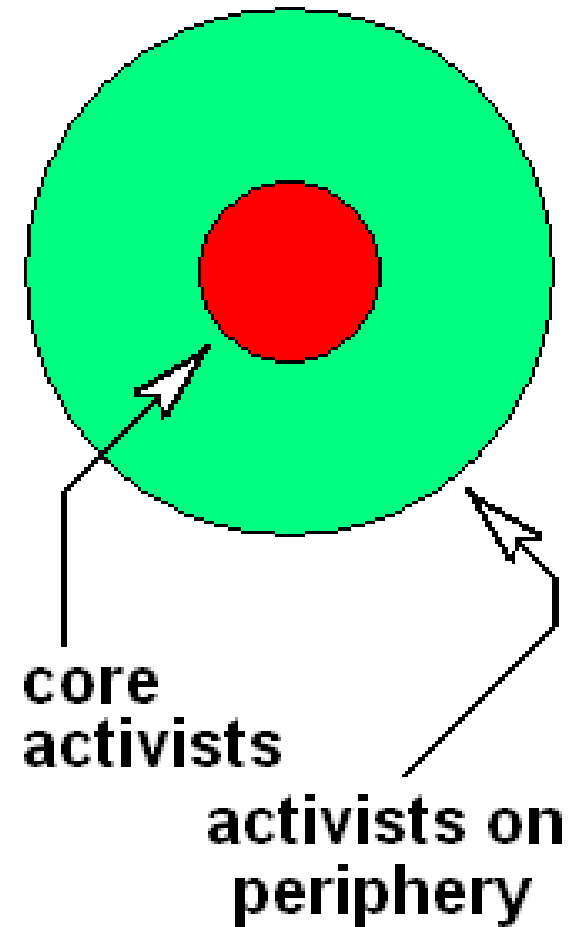


Manipulations
de paroles et
d'images...



Armes de communication massive!

- **hacktivisme** (militantisme sur les réseaux),
- **infox** (interférence avec le jeu démocratique),
- **cybercriminalité**,
- **cyberterrorisme** (attaques massives de site sensibles ou de ministères)
- **cyberespionnage**
- **cyberguerre**

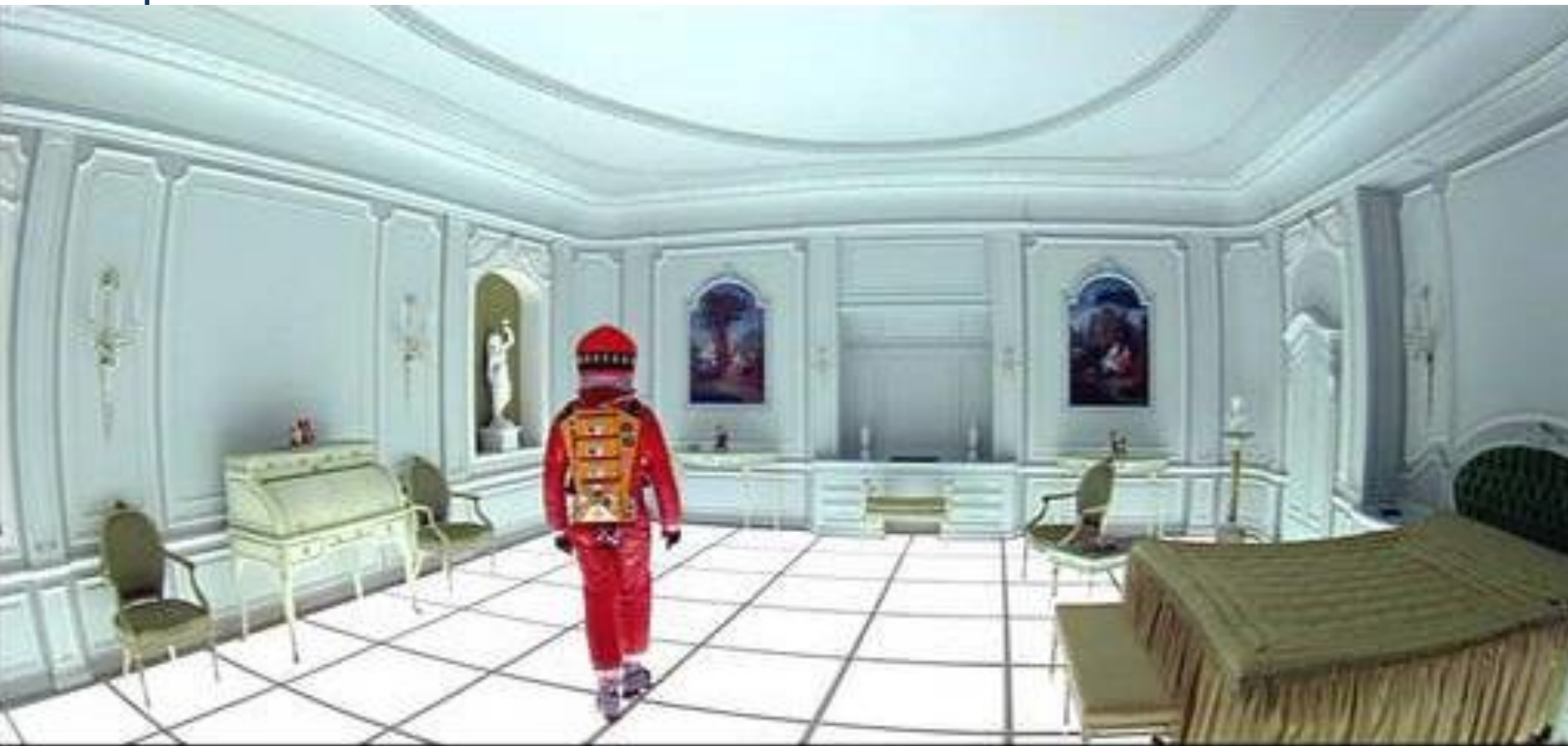


Synoptique

1. Définition(s) de l'IA
2. Empan
 - i. Perception
 - ii. Mémoire et apprentissage
 - i. Apprentissage profond
 - iii. Raisonnement
 - iv. Fonctions exécutives
 - v. Fonctions communicatives
3. Perspectives actuelles
4. IA et défense
5. Nouvelles vulnérabilités
6. **Éthique: jusqu'où?**



Jusqu'où?



L'IApocalypse... « révélation »



Stephen Hawking says A.I. could be 'worst event in the history of our civilization'

- Physicist Stephen Hawking said the emergence of artificial intelligence could be the "worst event in the history of our civilization."
- He urged creators of AI to "employ best practice and effective management."
- Hawking is among a number of voices including Elon Musk who have warned about the dangers of AI.

Arjun Kharpal | @ArjunKharpal

Published 2:11 PM ET Mon, 6 Nov 2017 | Updated 3:39 PM ET Mon, 6 Nov 2017



Elon Musk issues a stark warning about A.I., calls it a bigger threat than North Korea

- Tesla's billionaire CEO renewed his critique of artificial intelligence, saying that if you're not concerned, 'you should be.'
- He likened autonomous machines to North Korea, saying they were a bigger threat.

Javier E. David | @TeflonGeek

Published 10:21 PM ET Fri, 11 Aug 2017 | Updated 10:39 AM ET Mon, 14 Aug 2017

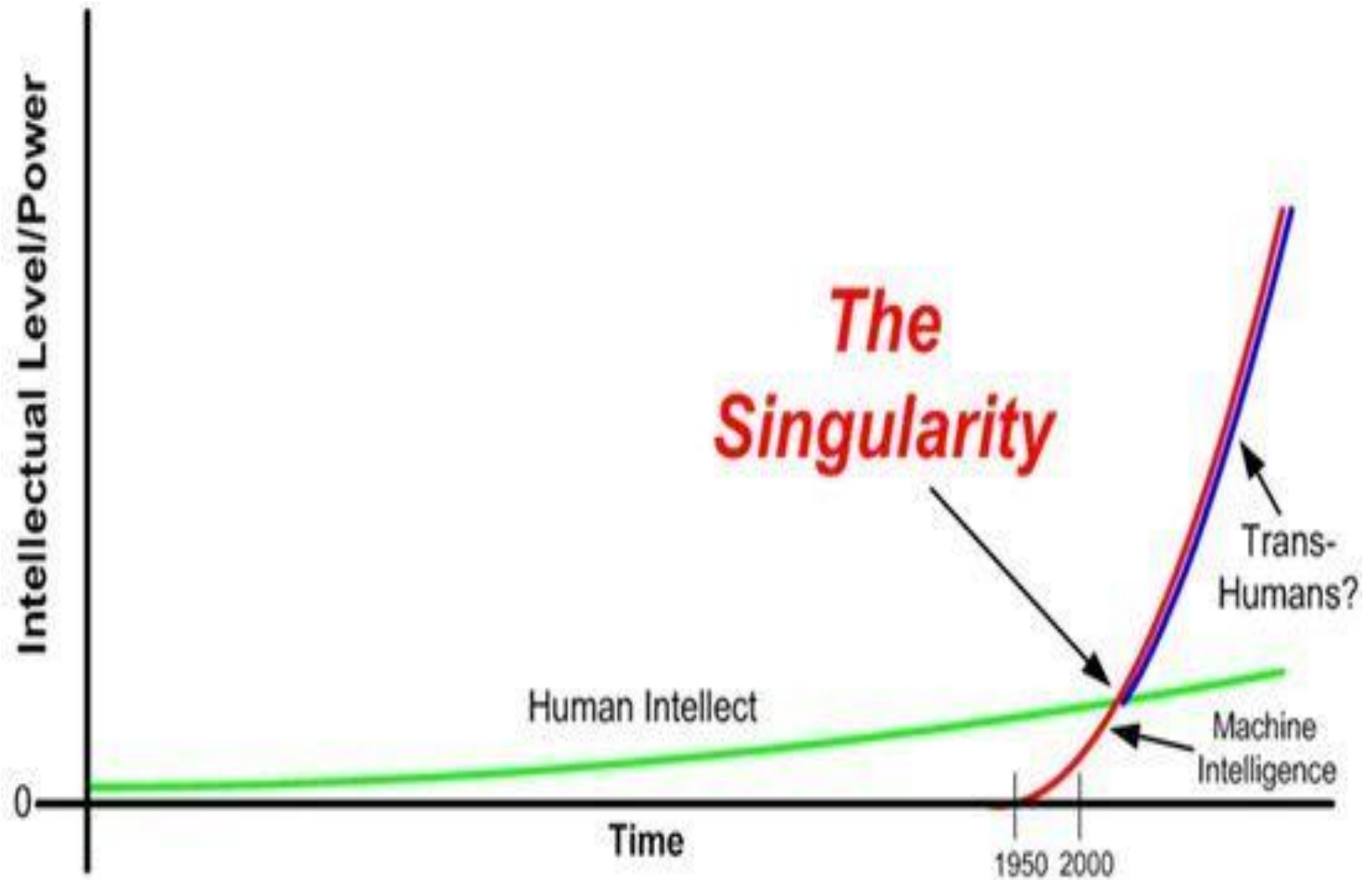
“AI is a fundamental risk to the existence of human civilization”

L'IApocalypse

- *“Whereas the short-term impact of AI depends on who controls it, the long-term impact depends on whether it can be controlled at all.” Stephen Hawking*
1er mai 2014



La Singularité technologique



« Mind Uploading »



De la Singularité au transhumanisme

2045.com

Applications DMKM Dictionnaires Google Agenda Page J-GG Master IAD Courriel LIP6 Crédit Mutuel LOCUPLETO Cybernity



Strategic Social Initiative

RUS / ENG

ENROLLMENT **37815** MEMBERS

+ Join us

Twitter Facebook

experts / ideology / faq
 / technology / interfaith dialog
 / discussion / press / rebrain / about us
 / make a donation

/ #3 Avatar Technology Digest / Bionic hand, BioMicroRobot, 3D-printed bionic ants etc.



Welcome to #3 Avatar Technology Digest. And here are the top stories of the last week. As always we start our Digest with incredible news on Technology, Medical Cybernetics and Artificial Intelligence.



Immortality Button

Click this button to start the development of your **personalized immortal avatar**



Global Future 2045
 New York, June 15-16, 2013

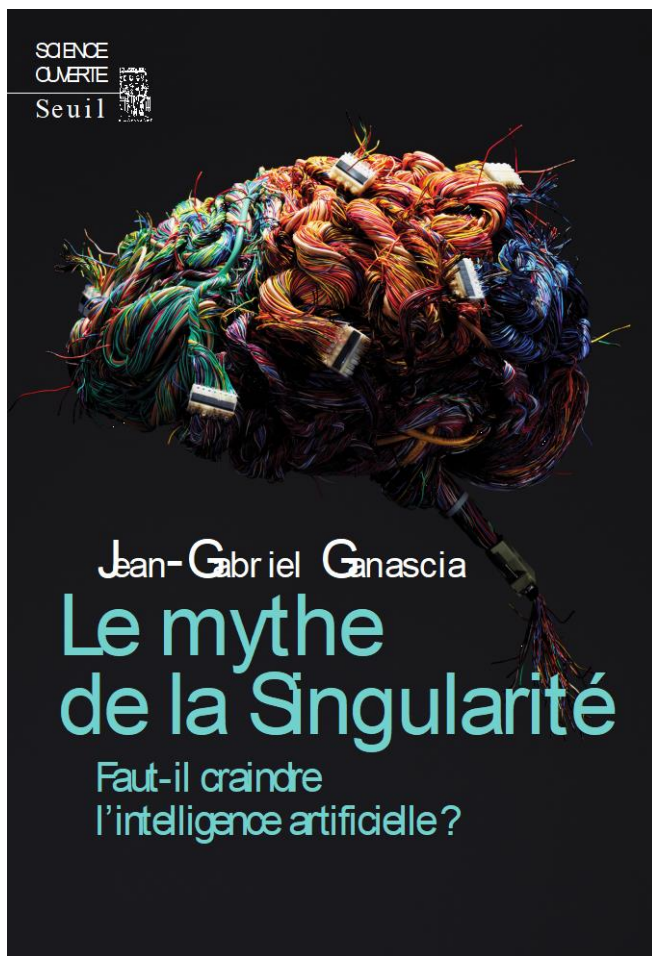
The second international Global Future 2045 congress focused on discussion of a new evolutionary strategy for humanity aimed at overcoming the 21st century's civilization challenges.

02.04.2015

The path to neo-humanity

2045 AVATAR PROJECT MILESTONES

L'Apocalypse



- Accélération exponentielle de la technologie (Loi de Moore)
 - Fin loi de Moore
 - Fréquence de calcul $\neq$ conscience
- Autonomie des machines: apprentissage et *big data*
 - Limité à apprentissage supervisé
 - Question des changements de paradigmes
- Les « pompiers-pyromanes »

Enjeux politiques: fin de la souveraineté des États?

Prérogatives régaliennes des Etats

- Défense
 - Cyber-sécurité (acquisition de protection et d'armes offensives par les grands groupes...)
- Sécurité intérieure
 - Reconnaissance faciale
 - État civil – réseaux sociaux (formulaire ESTA)
 - Traitement masses de données - Palantir
- Finance
 - Battre monnaie – Paypal, Bitcoin
 - Percevoir l'impôt: cadastre (Google)
- Justice
 - Utilisation des masses de données pour prévoir les récidives (US)
 - Site <https://crowdlaw.org/>: législation participative

Autre attributs de la souveraineté

- Santé
- Education
- ...

