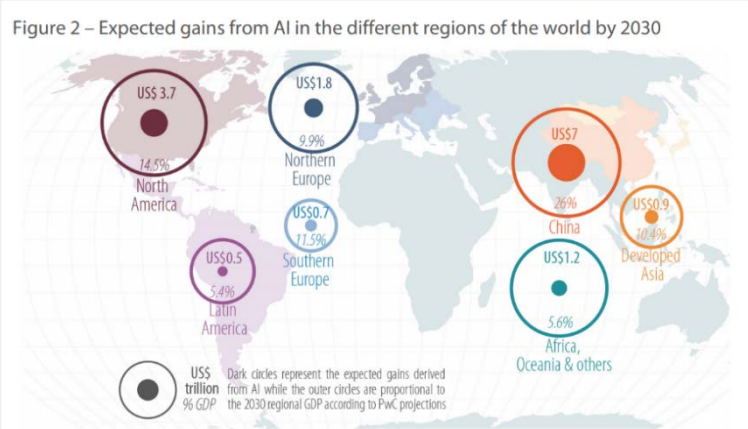
Responsable development of AI for industry and society

Francois TERRIER – francois.terrier@cea.fr

AI is ~~coming...~~ THERE!

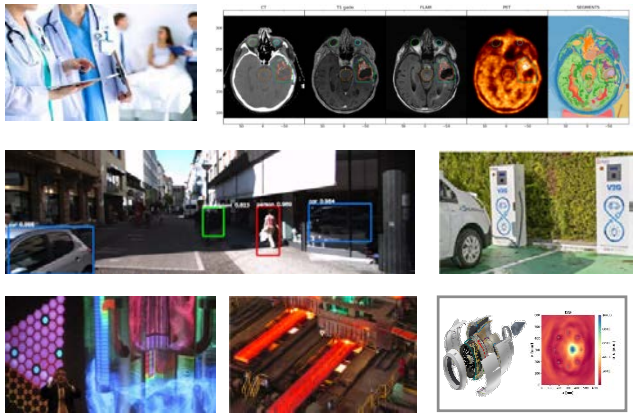
AI should represent in 2030
~10% of EU activity (US$ 1.8 trillions)



Source: The macroeconomic impact of artificial intelligence, PwC 2018

In many objects, products, systems of current life & industry

A lot of new applications in all public and private markets



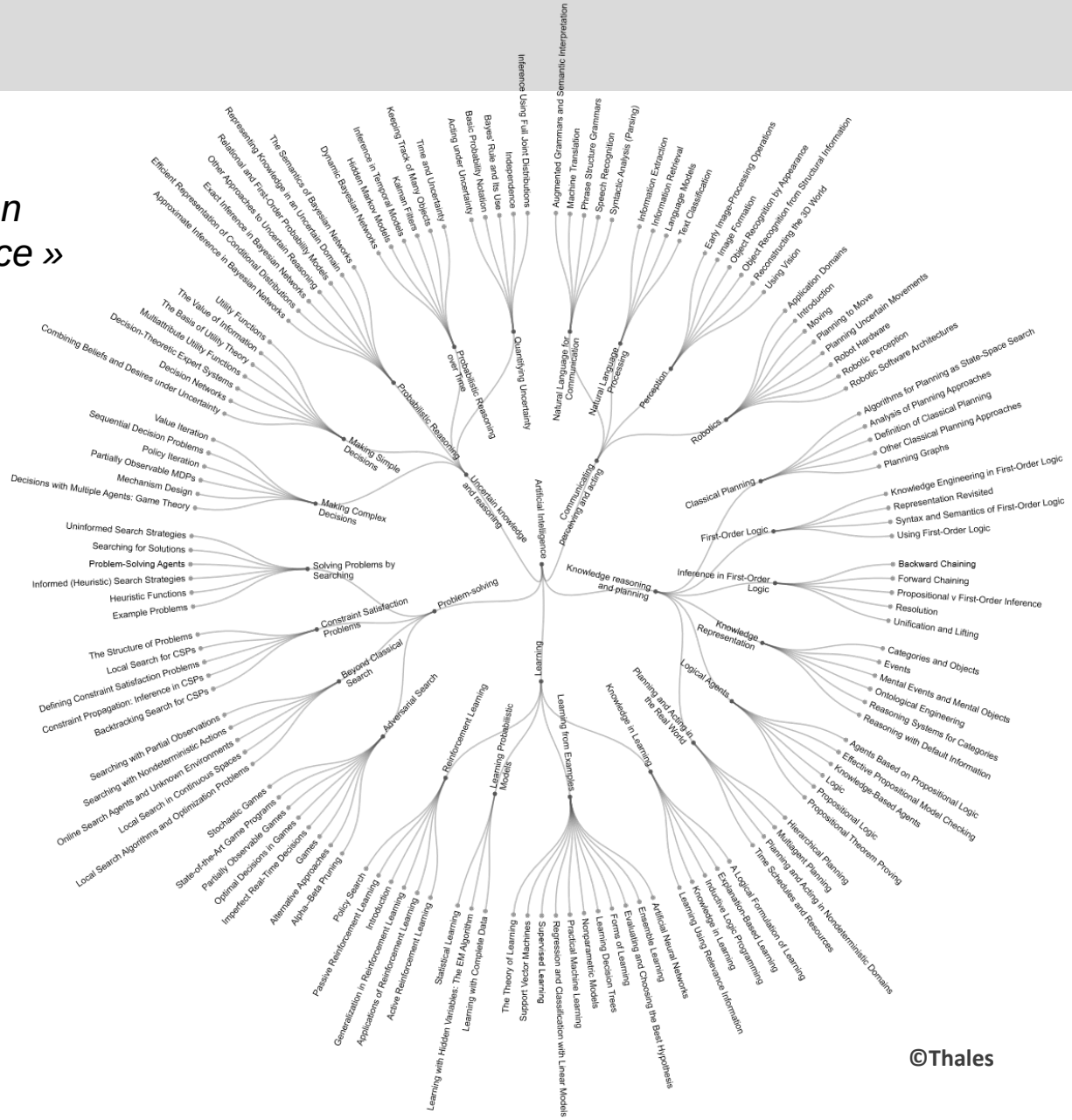
There is AI everywhere, everything is AI…

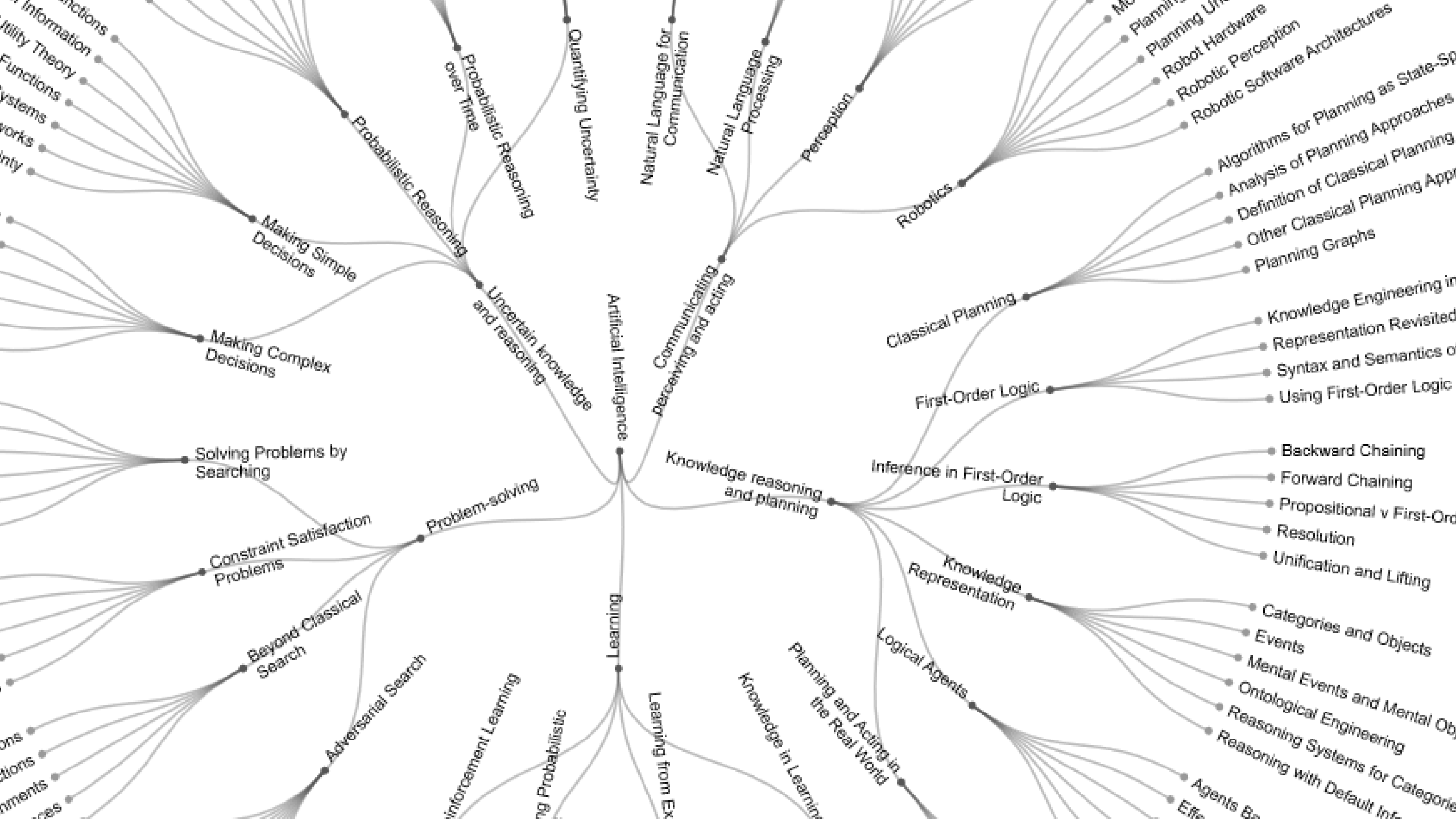
QU'EST-CE QUE L'IA ?

« L'ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence »

Encyclopédie Larousse - Cité par dalloz-actualite.fr

« Si son ambition initiale était d'imiter les processus cognitifs de l'être humain, ses objectifs actuels visent plutôt à mettre au point des ***automates qui résolvent certains problèmes bien mieux que les humains, par tous les moyens disponibles****. Ainsi l'IA vient au carrefour de plusieurs disciplines : informatique, mathématique (logique, optimisation, analyse, probabilités, algèbre linéaire), sciences cognitives... sans oublier les connaissances spécialisées des domaines auxquelles on souhaite l'appliquer. »*




Artificial Intelligence
Problem-solving
Solving Problems by Searching
Constraint Satisfaction Problems
Beyond Classical Search
Adversarial Search
Uncertain knowledge and reasoning
Probabilistic Reasoning over Time
Quantifying Uncertainty
Making Simple Decisions
Making Complex Decisions
Learning
Reinforcement Learning
Learning from Examples
Knowledge in Learning
Communicating, perceiving and acting
Natural Language Processing
Natural Language for Communication
Perception
Robotics
Robot Hardware
Robotic Perception
Robotic Software Architectures
Planning Graphs
Knowledge reasoning and planning
Classical Planning
Algorithms for Planning as State-Sp
Analysis of Planning Approaches
Definition of Classical Planning
Other Classical Planning Appr
First-Order Logic
Knowledge Engineering in
Representation Revisited
Syntax and Semantics o
Using First-Order Logic
Inference in First-Order Logic
Backward Chaining
Forward Chaining
Propositional v First-Orc
Resolution
Unification and Lifting
Knowledge Representation
Categories and Objects
Events
Mental Events and Mental Ob
Ontological Engineering
Reasoning Systems for Categorie
Reasoning with Default Inf
Logical Agents
Planning and Acting in the Real World
Agents Ba
Eff

QU'EST-CE QUE L'IA ?

Odorat ?

Toucher ?

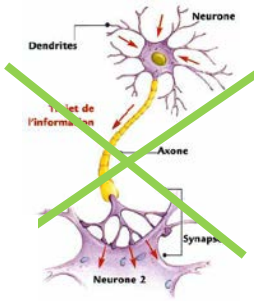
Perception - *Mimétisme*

Vision | Parole | Traitement des langues | Robotique | Controle | Diagnostic | Planification

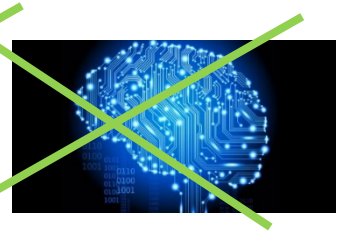
Systèmes à apprentissage → **Apprentissage machine, Réseaux de neurones**

Systèmes symboliques → **Raisonnement automatique**

Analyse de données



Anthropomorphisme !



Maîtriser l'effet buzz

… "biomimétisme"…

… "Bio-inspiré"… pourquoi pas si on analyse avec rigueur l'apport de la créativité/intuition !

QU'EST-CE QUE L'IA ?

Odorat ? *Toucher ?*

Perception - *Mimétisme*

Vision | **Parole** | **Traitement des langues** | **Robotique** | **Contrôle** | **Diagnostic** | **Planification**

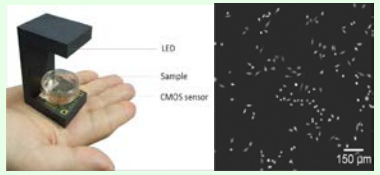
Systèmes à apprentissage → **Apprentissage machine, Réseaux de neurones**

Systèmes symboliques → **Raisonnement automatique**

Analyse de données

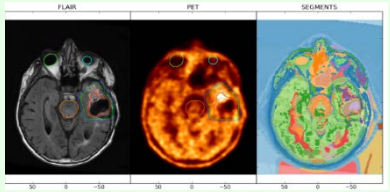
... vs réalité...


Automatic report analysis



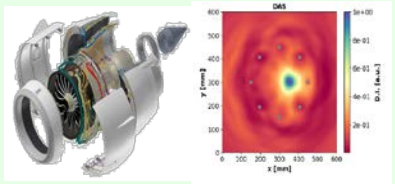
Lens-free microscopy


Virtual assistant



Medical diagnosis


Energy contract management

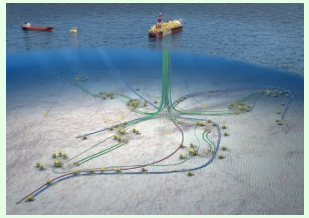

Structure Health Monitoring


Individuals identification

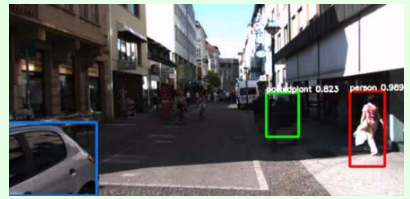

Autonomous shuttle


Medical prescription


Predictive diagnosis


Failure detection


Contrôle qualité



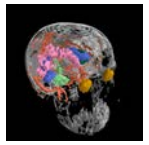
Object identification & localisation

Deux ambitions « complémentaires »

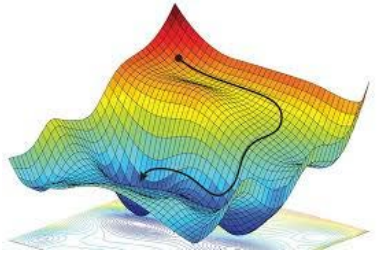
- **Modéliser et simuler les fonctions humaines**

Un rêve ?

Un mythe ?



Modéliser et simuler des fonctions complexes

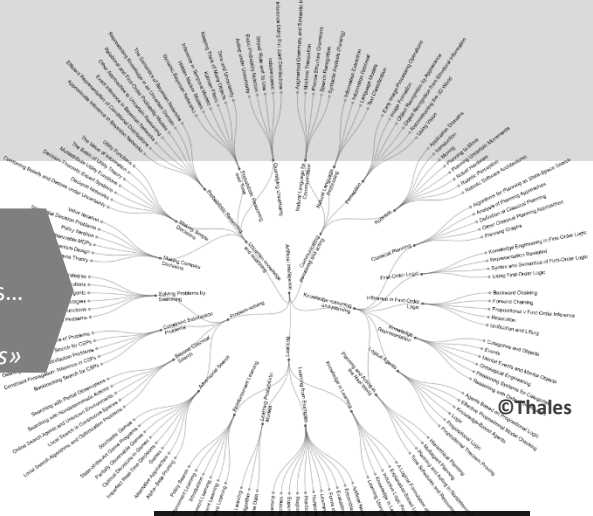


WHAT IS AI?

AI FOR HUMANITY
L'INTELLIGENCE ARTIFICIELLE AU SERVICE DE L'HUMAIN.

« An artificial intelligence is first and foremost a computer program that aims to perform tasks [requiring a certain level of intelligence] at least as well as humans. »

A lot of technologies, methods, theories, representations, calculus… *« The whole computer science and mathematics»*



Symbolic

- Rules, contraints, ontologies

➔ Problems of « logic », « exact » results

… A simple calculator that solves big problems

The change is the declarative, empirical approach

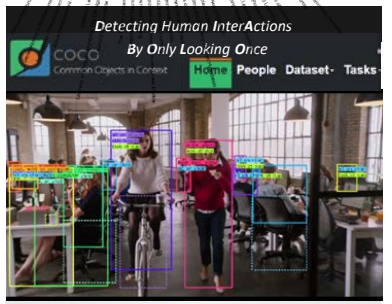
Learning based - connexionnism

- Learning, artificial neural networks…

➔ « Non linear interpolation among reference samples »

… How many pictures does a network need to recognise a cat?

… How many cats does a child see to recognise them?



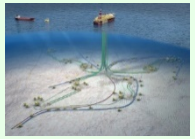
Top-1 Microsoft V-COCO, T-2 HICO-DET data set

Hybride

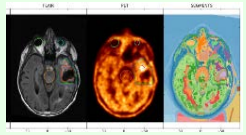

Prescription médicale


Diagnostic prédictif

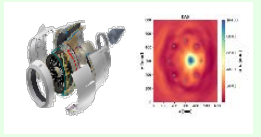

Optimisation de charge

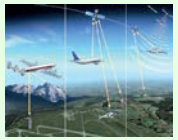

Détection de panne

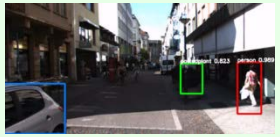

Contrôle qualité


Détection de tumeurs


Microscopie sans lentille


Contrôle non destructif


Optimisation trajectoire


Mobilité, identification d'objets

QU'EST-CE QUE L'IA ?

▶ *"L'IA c'est pas que les données"…*

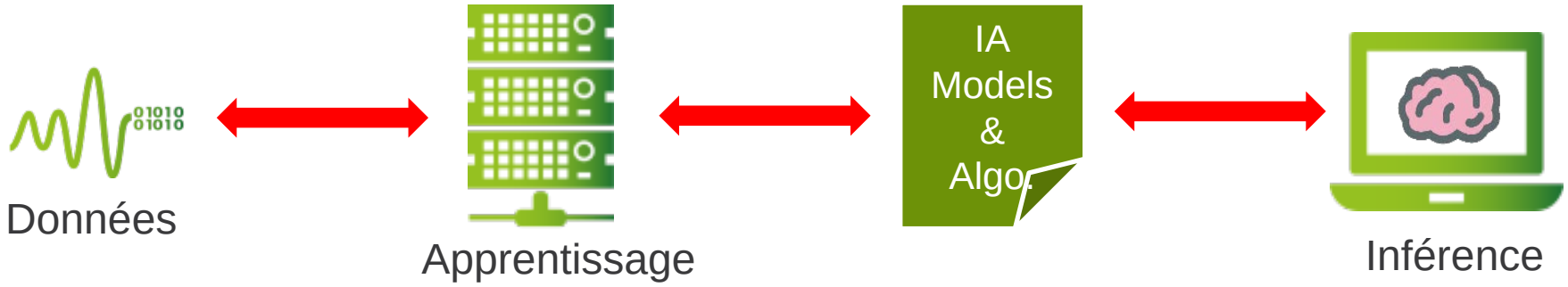
Ex : Véhicule autonome → *Perception par réseaux de neuronnes*
→ *Décision par IA "symbolique" = IA à base de connaissances structurées*

▶ ➔ **Oui, mais sans elles (les données) et sans le deep learning on ne parlerait pas d'IA…**

➔ **Rôle majeur des données** **(ne pas se leurrer sur l'enjeu de l'IA « classique »)**

La rupture actuelle vient des données et de l'apprentissage profond

- **Distinguer apprentissage et inférence :**

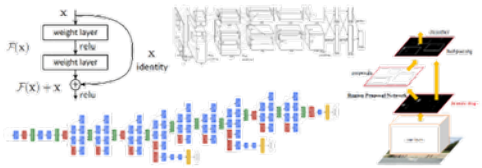


- Faire la part en ce qui peut / doit être fait en local sur chaque objet utilisateur
- Partage des apprentissages entre plusieurs utilisateurs (fédératif, centralisé vs distribué)

ARTIFICIAL INTELLIGENCE

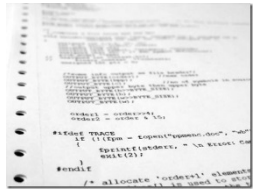
Data and knowledge

+

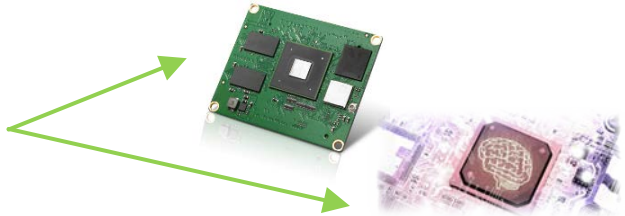


An engine / algorithm / program

→



A software
A function



Hardware, standard / dedicated

The breakthrough,
The « *Buz* »

Efficient!!!

On elementary task

- *Perception*
- *Reasonning*

and more and more on complex tasks

Big Data

DeepLearning

Transformers

Mega Models

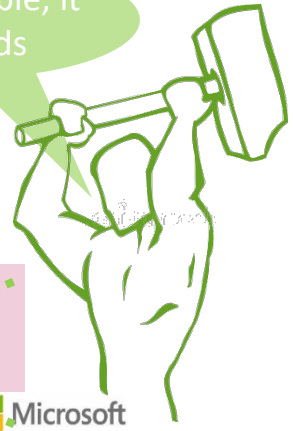
"AI does better than the human"

Ex. Vision, Speech

"... nous avons atteint le niveau humain :
- la transcription de la voix en texte,
- la traduction automatique,
- les réponses aux questions courantes,
- la compréhension globale d'un texte.
- l'ajout des légendes aux images..."

Microsoft

Its debatable, it depends



But...

Efficient and impressing!!!

on elementary task as

- *Perception*
- *Raisonnement*

But…

No common sense

« Chicken » or « Pedestrian »

« Nothing (recognized) behind? »

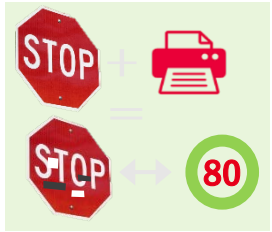
« Known known »

« Known unknown »

« Unknown unknown »

fragile

attackable



Or ***mis-used***

Report on Tesla first accident – Recommendation

Incorporate system safeguards that limit the use of automated vehicle control systems to those conditions for which they were designed. (H-17-41)

« At the best AI learns a result. It does not acquire a process of recognition, whereas the process is a key element of human intelligence »

… LES PROCESSUS DE DÉVELOPPEMENT NE SONT PAS SOUS CONTRÔLE !

Vrai aussi bien pour l'IA à base de données que de connaissances

… avec une très forte pression sur les réseaux de neurones

Approches formalisée, tracée, rationnelle

- Exigences
  - ↳ *Fonctions*
    - ↳ *Sous-fonctions*
      - ↳ *Actions*

Chaque instruction, chaque valeur est déduite et justifiée

vs

Une approche empirique

- Des exigences informelles, "par l'exemple"
  - ↳ **Essais / erreurs**

Très pauvres justifications, explication du résultat

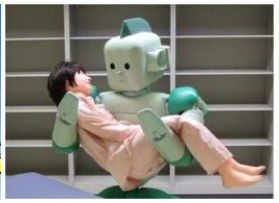
Exigences informelles avec peu / pas de structuration
Evolution dynamique de la définition des systèmes

Brise tous les principes et processus d'assurance de la conformité

DEVELOPMENT PROCESSUS ARE NOT (yet totally) UNDER CONTROL…

TRUST: the need is there





Ali Rahimi
(Google)

Machine learning becomes alchemy

Yann LeCun
(facebook)

Engineering artefacts preceded the understanding of the theory

The technology arrives with a poor (industrial) maturity and evident weakness on the definition of the *usages, specifications, design methods, robustness metrics, quality processes…*

breaks all principle of safety certification processes

LES DÉCEPTIONS DE L'IA

ESSORT DU GÉNIE LOGICIEL
Intégration des techniques matures de l'IA

Modélisation (UML), Méthodes formelles

IA – raisonnement automatique, début des « neurones »

- Trop de promesses, puissance de calcul faible

Systèmes experts
+ déclaratif, intuitif
- complexité, maintenance

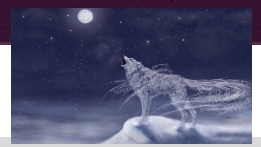
Deep Learning / Données

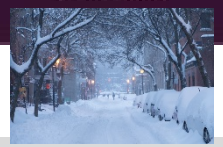
50 60 70 80 90 00 10

1956

1974 - 1980

1987 - 1993





… LES PROCESSUS DE DEVELOPPEMENT NE SONT PAS (*encore*) SOUS CONTROLE !

AI is HUGE and EVERYWHERE! But …

No common sense



« Known known »
« Known unknown »
« Unknown unknown »


Les artéfacts d'ingénierie ont précédés la compréhension théorique

Yann LeCun
(facebook)

Fragile

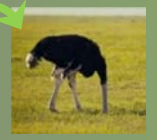

Scholar bus

+

=

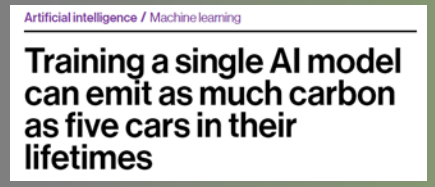

Ostrich

?



Quality
« How?»

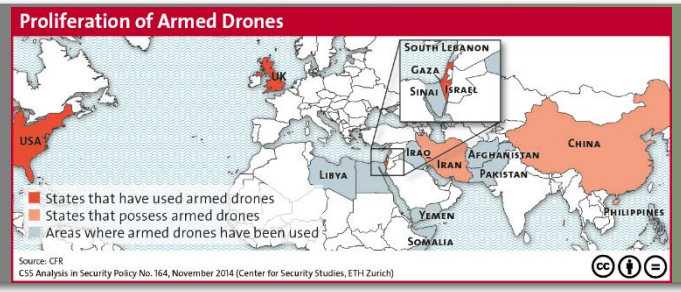
Heavily polluting



« Edge »
Embedded
« Where »

Application issues

European Parliament



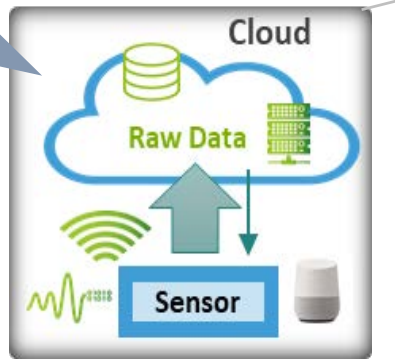
Usage
« Why? »

AI is coming… *HUGE and EVERYWHERE* !

From cloud to edge: AI closer to the user

Data today:
- 80 % in the cloud
- 20 % at edge
But within 5 years:
- 20 % in the cloud
- 80% everywhere

Source: Thierry Breton commissioner, 2020



Learning and Inference in the Cloud

Learning and Inference at the edge

Low Latency *(shortest paths)*

Reliable operations
(even with intermittent connectivity)

Data Privacy *(remain at the edge)*

Energy Efficiency *(less data exchange)*

Massive deployment of AI in industry

More and more expectations on trust and frugality

A lot of new applications

Usage challenge

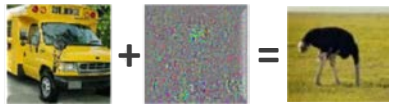
Automation

A lot of data, computation

Trust challenge

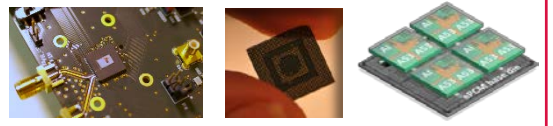
No common sense

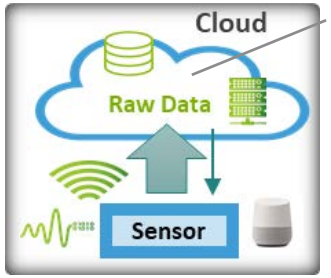
Fragile



Frugality challenge

Embedded challenge





Learning and Inference in the Cloud

Learning and Inference at the edge

From cloud to edge: toward embedded AI closer to the data and the users

AI DEPLOYMENT A GREAT OPPORTUNITY AND A STRATEGIC ISSUE

A European strategy

WHITE PAPER

On Artificial Intelligence – A European approach to excellence and trust

REGULATION OF THE PARLIAMENT AND OF THE COUNCIL

LAYING DOWN HARMONIZED RULES ON ARTIFICIAL INTELLIGENCE

A French strategy

An industrial strategy

GOUVERNEMENT

November 8, 2021

Innovation

Research

Embedded AI, Frugal AI

Trustworthy AI

AI foundations

Micro-electronics

Component & architecture

Algorithm for frugality and Model optimization

European approach to ethics and regulation

2018 — Expert group analyse the subject of TRUST

2019: refinement for key sectors

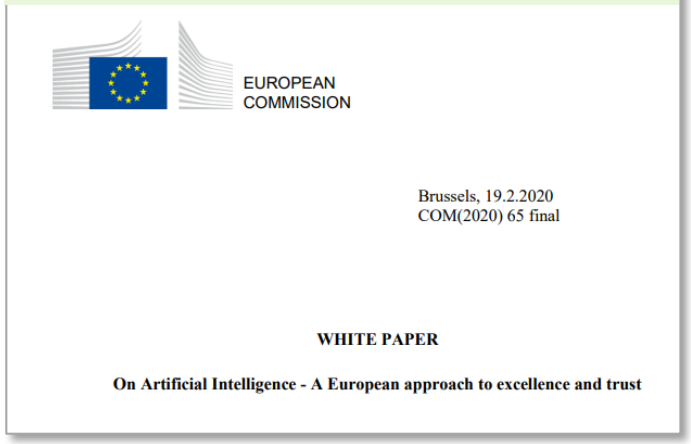
- Manufacturing
- Health
- Justice
- E-Government

Trustworthy AI should be:
(1) lawful - respecting all applicable laws and regulations
(2) ethical - respecting ethical principles and values
(3) robust - both from a technical perspective while taking into account its social environment

7 key requirements:
- Human agency and oversight
- Technical Robustness and safety.
- Privacy and data governance.
- Transparency.
- Societal and environmental well-being.
- Accountability.

European commission vision

2020: Approach for excellence and Trust


EUROPEAN COMMISSION

Brussels, 19.2.2020
COM(2020) 65 final

WHITE PAPER

On Artificial Intelligence - A European approach to excellence and trust

Toward a european regulation

2021: Proposal to the parliament


EUROPEAN COMMISSION

Brussels, 21.4.2021
COM(2021) 206 final

2021/0106 (COD)

Proposal for a

REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS

Human-centric AI:
- **AI system builder is responsible → robustness, safety, privacy, transparency…**
- **Human right must be respected and not subject to automated decision only**

Toward an European regulation for AI deployment respecting the european values

European Parliament presentation: www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2021)698792

The act (108 pg) : https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206

2021: Proposal to the parliament

EUROPEAN COMMISSION

Brussels, 21.4.2021
COM(2021) 206 final

2021/0106 (COD)

Proposal for a

REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS

A strong European legislation, i.e applicable as it is in each EU countries

- **Subject matter**
- **List of prohibited AI**
- **Rules for high risk AI systems**
- **Transparency obligations**
- **Support to innovation**

Outside of Europe: still at stage of recommandations...

Analysis by Future of Life Institute: https://artificialintelligenceact.eu/

NIST National Institute of Standards and Technology U.S. Department of Commerce

www.nist.gov/system/files/documents/2022/03/17/AI-RMF-1stdraft.pdf

The AI RMF is intended for voluntary use in addressing risks in the design, development, use, and evaluation of AI products, services, and systems.

OECD BETTER POLICIES FOR BETTER LIVES

https://oecd.ai/en/ai-principles

… innovative and trustworthy and that respects human rights and democratic values. (May 2019)

THE AI ACT

Centered on the usage and risk analysis

- **Forbidden usages:**
  - AI systems that deploy harmful manipulative 'subliminal techniques';
  - AI systems that exploit specific vulnerable groups (physical or mental disability);
  - AI systems used by public authorities, or on their behalf, for social scoring purposes;
  - 'Real-time' remote biometric identification systems in publicly accessible spaces for law enforcement purposes, except in a limited number of cases.

3 levels of risks depending of the usage domain

- **High risk systems: Regulated according to usual EU rules + 8 specific applications areas**
  (biometry; Education; Employment management; essential private and public services; Law enforcement; Migration; Administration of justice; democratic processes)
  → self-assessment for applications not already governed by European legislation

- **Limited risk: Transparency obligations**

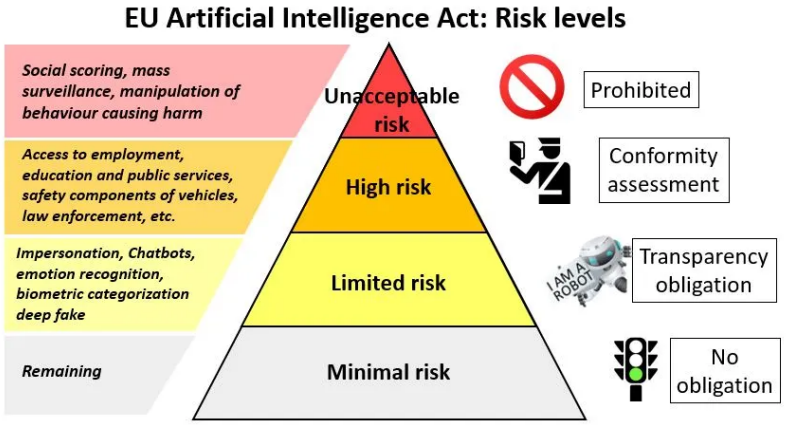
- **Low or minimal risk: No obligations**

« Confiance ➔ Certification ➔ Réglementation ➔ Normalisation »

Réglementation de l'IA : 1ière mondiale

Les autres pays en sont au stade « recommandations »

THE AI ACT



Influence les réglementations des autres pays
➔ ***« peur de l'effet RGPD »***

Octobre 2022, requête :
Production de standards support réglementation

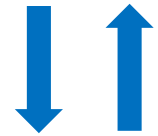
Décembre 2024 :
Normes harmonisées pour la conformité

Normalisation

Organisme européen

Commission Nationale IA

Coordination des contributions françaises



2018 « HLEG » - 2021 « AI Act » - 2022 Normalisation → 2024... !

AI is HUGE and EVERYWHERE! But …

TRUST
(Quality)
« How?»

FRUGALITY
(Embedded)
« Where? »

USAGE
« Why? »

Quality & Technical

Standards

ISO

cen CENELEC

+ domain Std (aero, auto, etc.)

THE AI ACT

AI liability

Data Act

Regulation

Ethics

Toward a GLOBAL POLICY

TRUST challenge: a set of characteristics

Trust challenge

No common sense

Fragile

Engineering view point

Safety community see AI components as (physical) systems

Certification, Quality, Safety, Reliability, Security, Privacy, Robustness, Accuracy, Traceability, Interpretability,

Usage view point

Ethics, Societal Impact, Accountability, Fairness, Transparency, Explainability

AI community see AI components as SW

What is the specification?
(formalized specification)

A STRUCTURING NATIONAL PROGRAM: CONFIANCE.AI

www.confiance.ai

Grand défi «Trusted AI method & tools»

Confiance ai

▶ 45 M€, 4 years

- ▶ Collective works
- ▶ Large industrials
- ▶ A rich set of use case

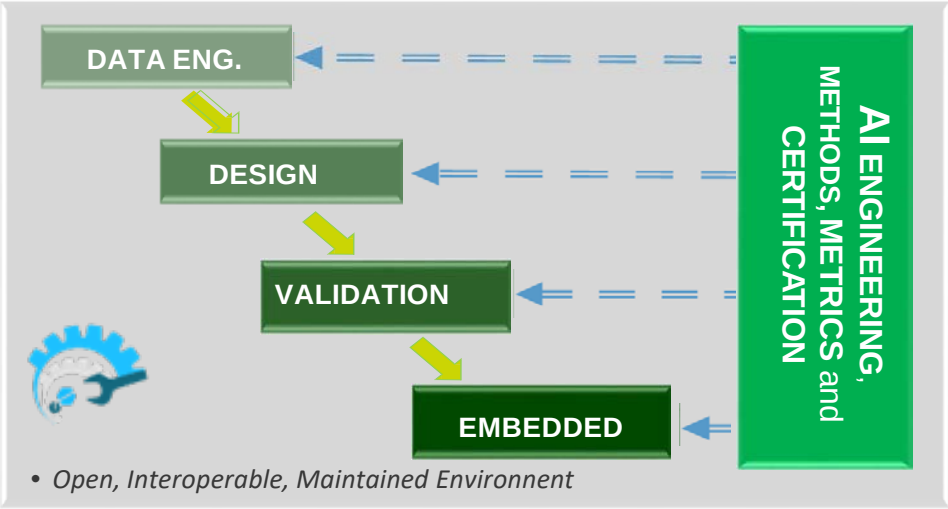
"From back office to embedded systems"

CEA involvement:
35 pers. + tools

STREAMER pixano N2D2 frama #AI Papyrus #AI

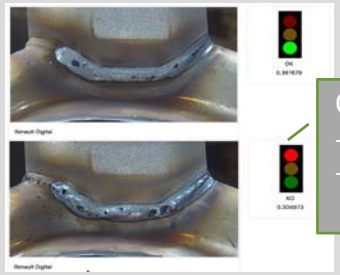
- ▶ Tooling components, safety engineering tools

"From data & knowledge collection to software deployment"



Real industrial use cases, Renault examples

DNN for Weld seam control



Qualification
- Process def.
- Robustness eval.

Intersection crossing

Exp. Sys. Safety valid.
- Completness
- Consistency

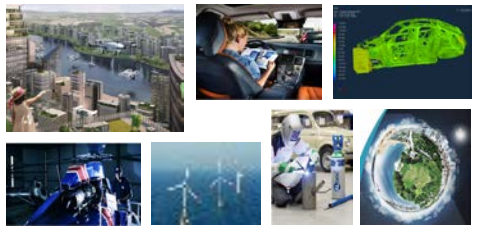
Opinion mining (NLP)

Qualification
- Robustness - Accuracy - Explainability

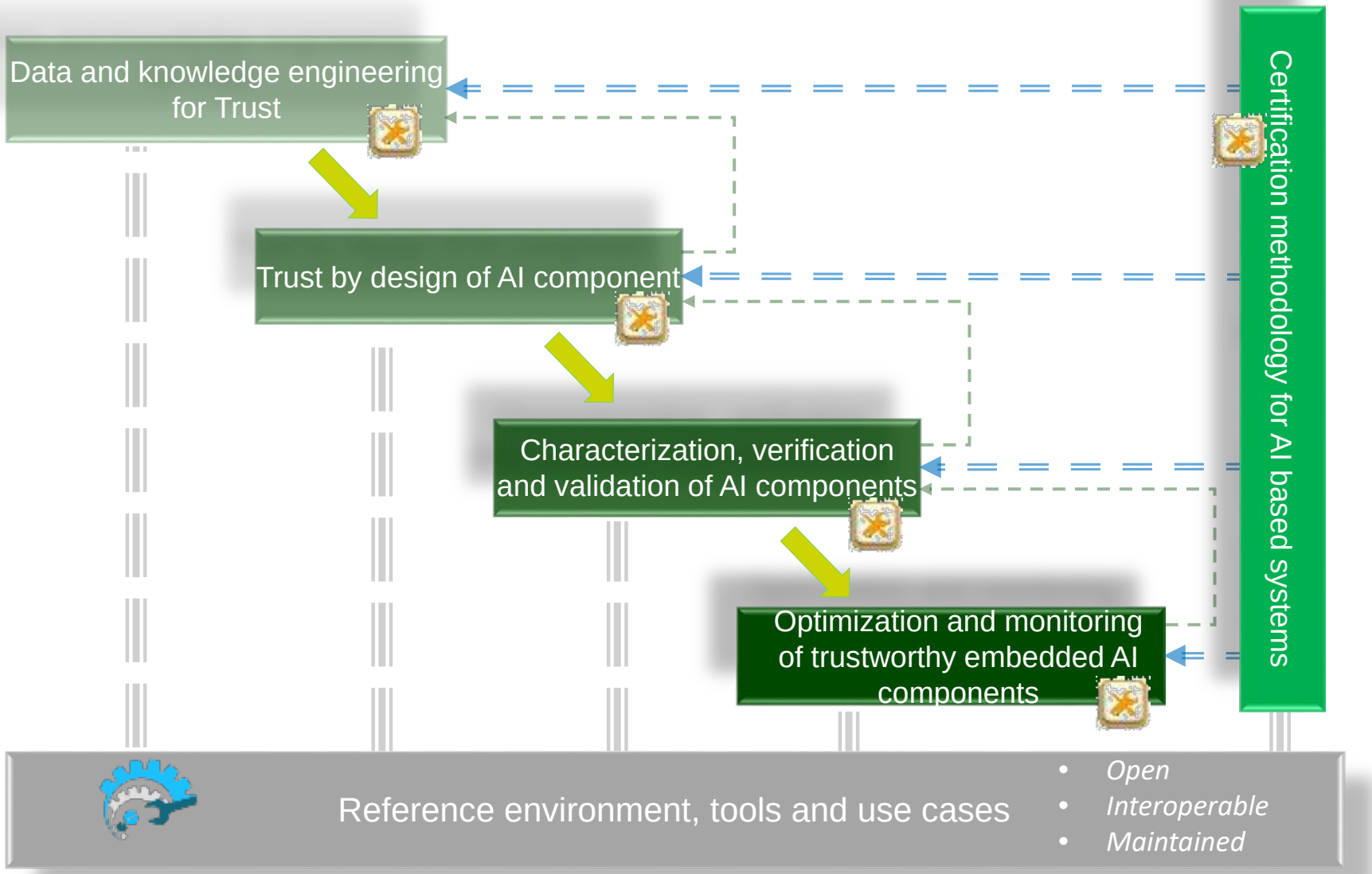
Building a methodology and tooling supporting the « four main stages of ML component development »

Lotissement

- **Fonction d'ingénierie, Briques outils**
  - ***De la collecte au déploiement***
- **Environnement fédérateur**
- **Cas d'usage**



  - ***du bureau à l'embarqué***



Défi Confiance.AI
AI Development Flow – Tooling examples

Methodologies and metrics for trustworthy AI

Data and knowledge engineering for Trust

- *Operational domain specification*
- *Building trustworthy data base*
- *Knowledge & data base charac.*
- *Augmentation of spare data base*

Trust by design of AI component

- *Iterative methods to build trustworthy data-driven AI*
- *Certified architecture for specific data-driven AI application*

Characterization, verification and validation of AI components

- *Quality assessment of a knowledge-based AI*
- *Performance metrics evaluation of data-driven AI*
- *Black box robustness characterization*
- *Metrology of explainability methods*

Optimization and monitoring of trustworthy embedded AI components

- *Automatic safe compression of neural networks*
- *Robust embedded code generation of data-driven AI*
- *Execution time optimization for neural networks*

- *AI risk spec. & analysis*
- *Assurance case based qualification support*
- *Incremental & evolutive qualification support*

Certification methodology for AI based systems

Reference environment, tools and use cases

AU-DELÀ DE LA CONFIANCE : 2 SUJETS CLÉS FORTEMENT ÉMERGEANTS AVEC DES VERROUS MAJEURS

- **Frugal, sobre, embarqué**

→ *Confusion des enjeux*

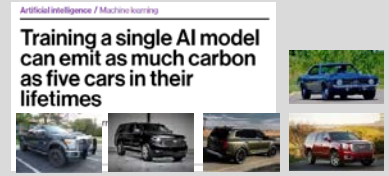
→ *Pas de vision de bout en bout*

→ *La confiance reste clé, mais…*

- ***Besoin de vision transverse de l'amont à l'industriel***

Données massives *Apprentissage profond* *Microélectronique*

Très grande performance *Mais…* → *Vorace*



IA FRUGALE et SOBRE de bout en bout pour un numérique de CONFIANCE

▶ **Distribué, sécurisé, efficace**

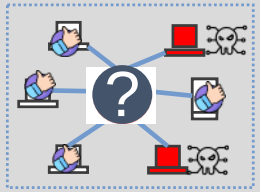
→ *Prise de conscience « aller au-delà du multi-agents »*

→ *Pas de vision technologiques des algorithmes aux systèmes*

▶ ***Enjeu clé pour développer le marché de la décision SoS et l'apprentissage embarqué***

Smart Mobility
Smart Production
Circular Economy

Très grands systèmes, très évolutifs, agiles *Mais…* → *Incontrôlable*

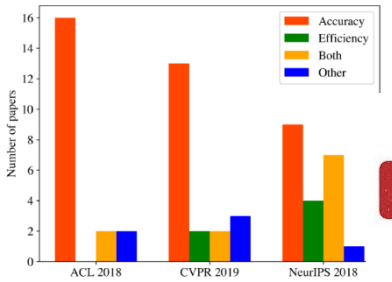


IA DÉCENTRALISÉE, résiliente, efficace pour systèmes distribués de CONFIANCE

« *Winter is coming. . .* »

➔ TRUST is a challenge, but sustainability too!

Trusted AI



Green AI

Green AI

Roy Schwartz* Jesse Dodge* Noah A. Smith Oren Etzioni

Allen Institute for AI, Seattle, Washington, USA
Carnegie Mellon University, Pittsburgh, Pennsylvania, USA
University of Washington, Seattle, Washington, USA

July 2019

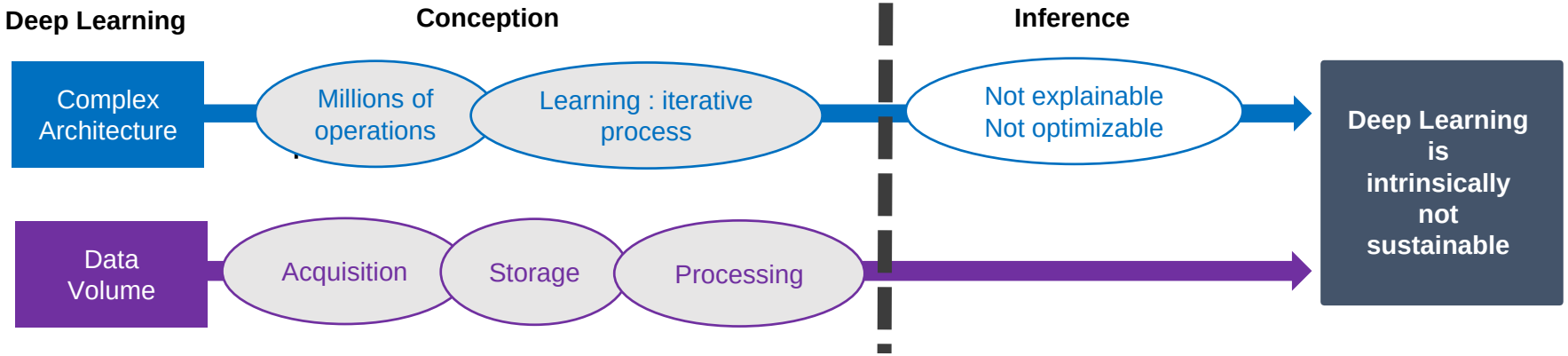

THE DAILY NEWSLETTER
Sign up to our daily email newsletter

NewScientist

News Technology Space Physics Health Environment Mind Crosswords Video

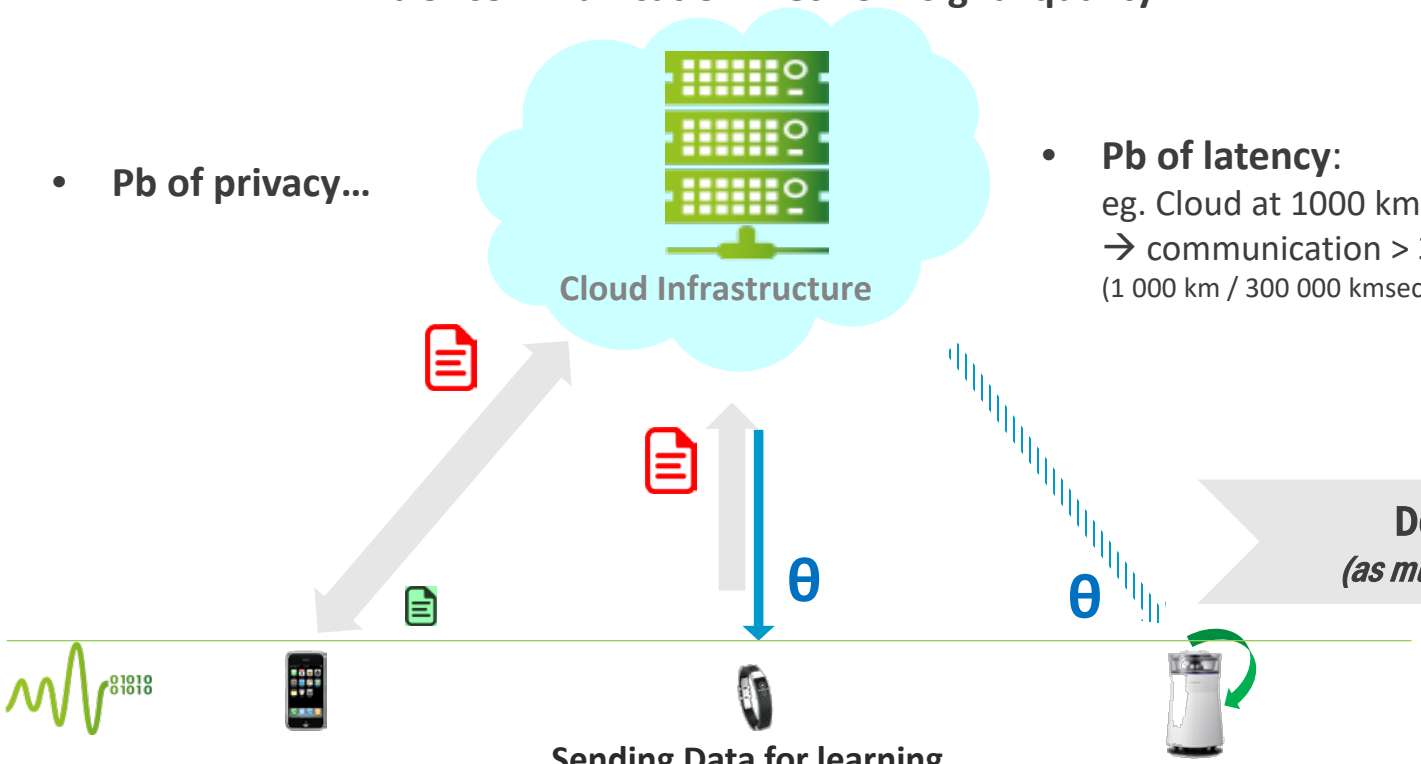
Creating an AI can be five times worse for the planet than a car

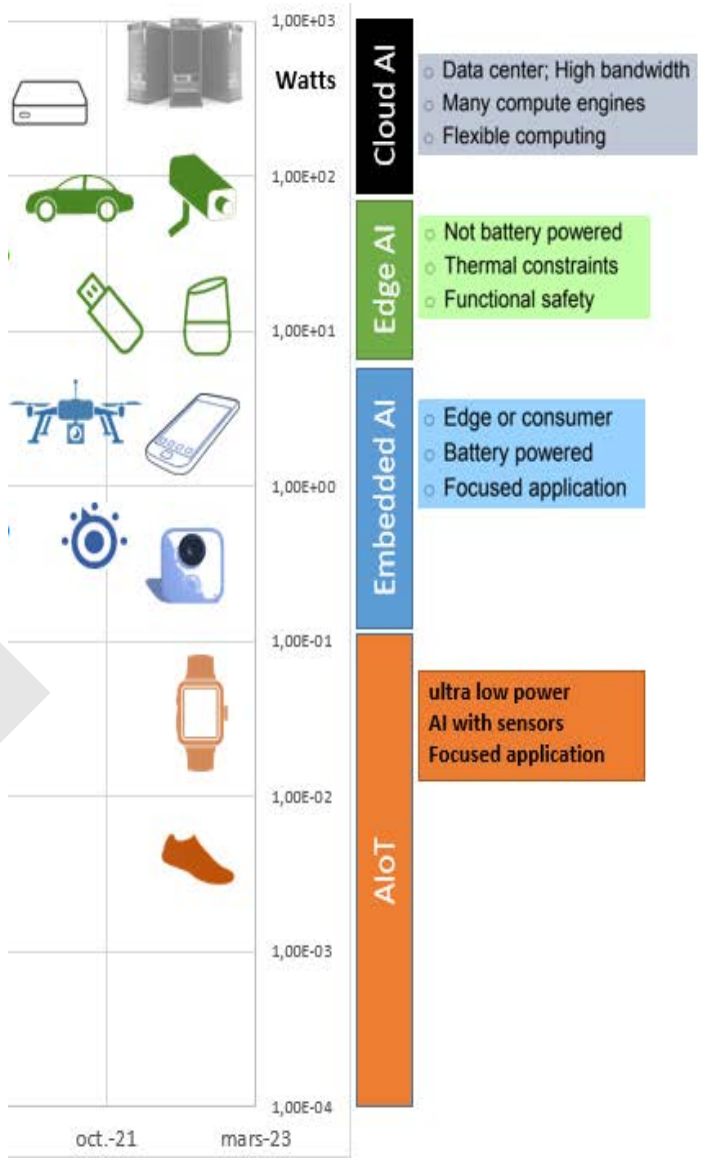
A RECENT ISSUE, BUT A DEEP ONE…



Cloud based learning: very costly in energy, communication…

- **Pb of communication: network signal quality…**
- **Pb of privacy…**
- **Pb of latency**:
  eg. Cloud at 1000 km
  → communication > 3 msec…
  (1 000 km / 300 000 kmsec-1)





FEDERATED LEARNING:

Estimation of vehicle speed from accelerometer data (vehicle vibrations)

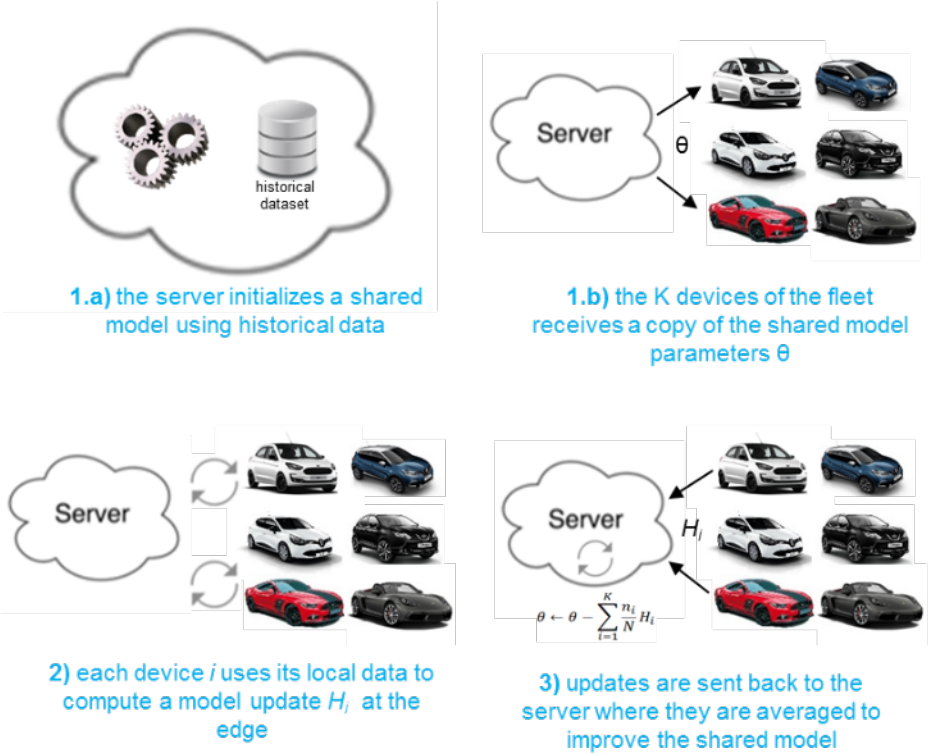
OBJECTIVES

Use federated learning to train a model on the local data of many users without the need to ever upload these data to a central server.

- Inputs: data from a fleet of cars equipped with 3-axis accelerometers measuring local car's vibrations.
- Outputs: estimation of the speed of vehicles.

RESULTS

Performance evaluation: 'model retraining' improves the accuracy of the model for each car while 'model aggregating' improves the robustness of the shared model without transmitting any raw data



1.a) the server initializes a shared model using historical data

1.b) the K devices of the fleet receives a copy of the shared model parameters θ

2) each device *i* uses its local data to compute a model update H_i at the edge

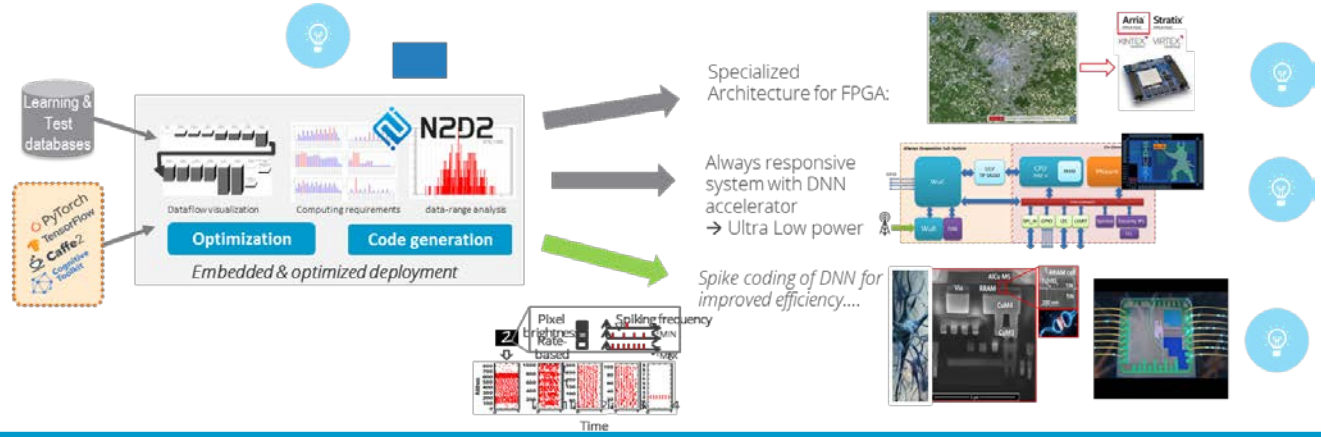
3) updates are sent back to the server where they are averaged to improve the shared model

Toward distributed and decentralized learning

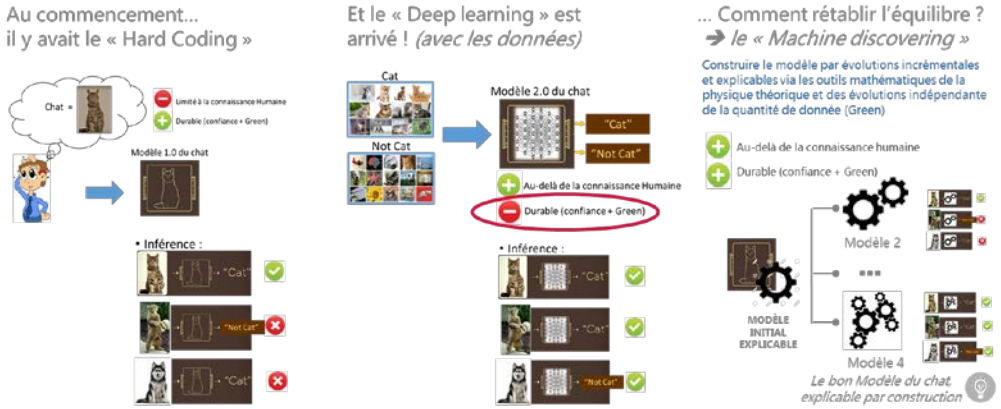
- Efficient, resilient and trusted consensus (multi-agent decision)
- Federated distributed decentralized learning (no more cloud)

La FRUGALITE :
une condition au déploiement massif de l'IA dans l'industrie

Par la technologie, les composants et les architectures électroniques

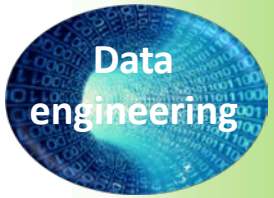


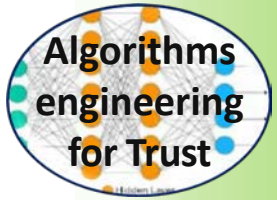
Par l'algorithmie, l'optimisation et les nouveaux paradigmes

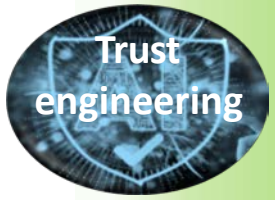


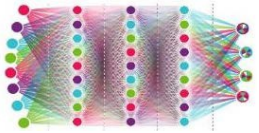
Main research axes for trustworthy AI

Risk analysis (ODD)









Data quality management

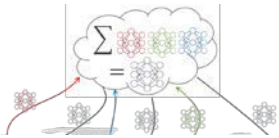
Automated annotation

Frugal & interpretable ML

Physics mathematics (PCA)

Robustness to env. noise

Learning with noise profile

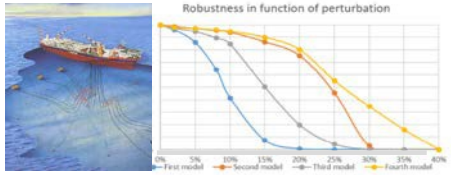


Federative learning & privacy

Keep data on edge, privacy

Robustness to perturbations

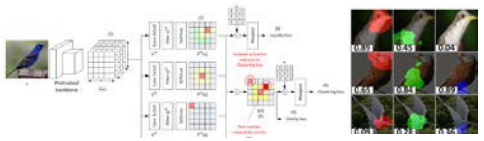
Exhaustive analysis, formal methods



Properties verification

Model interpretation

Automated identification of properties

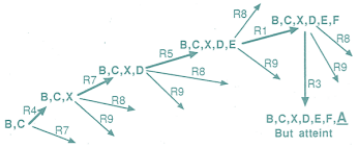


Real time risk monitoring

Associate risk graph with Bayesian models

Risk based certification

Risk and process modeling

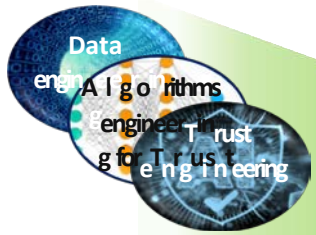


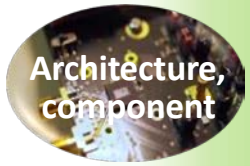
Explainability and rule learning

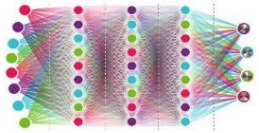
ExpressIF®

Proved inference engine

Main research axes for embedded AI





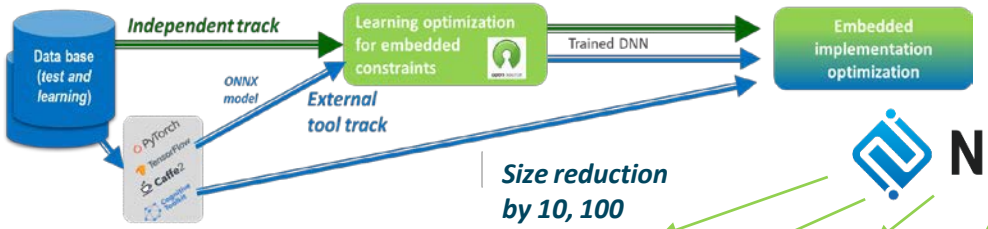


Quantization

Reduce model size, data accuracy

HW oriented code gen.

Target agnostic platform



Reliability to HW fault

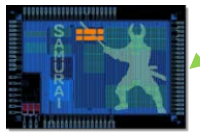
CNN memory protection

RAPTOR

Dedicated accelerator

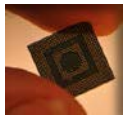
FPGA (DNeuro), IP (Pneuro), architectures

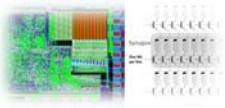
Pneuro

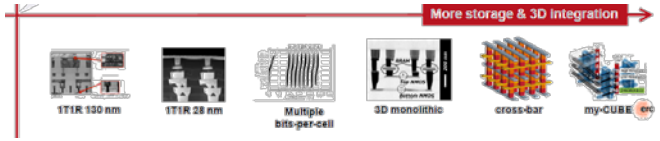


HW technologies

RRAM, IMC, 3D







C'est aussi des enjeux sociétaux

Réferentiels normatifs, légaux

Laws!
« No law, no pb »

Ugo Pagallo, University of Turin

From Automation to Autonomous Systems: A Legal Phenomenology with Problems of Accountability

Acceptabilité et éthique

Ethics!
« No autonomous weapon »

Economie, emploi, SOUVERAINETÉ

Destruction !

Artificielle : Au lieu De Supprimer Des Emplois, L'IA en Crée

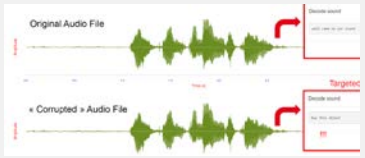
Moteur de croissance !

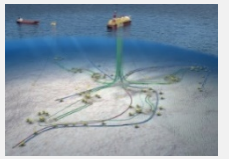
France stratégie – Intelligence artificielle et travail,
https://www.strategie.gouv.fr/publications/intelligence-artificielle-travail

« Juste » une des composantes du numérique… C'est d'abord un enjeu global de transition numérique

Le jumeau numérique
Couplage vues
Physique / Structure / Fonctionnel

Les systèmes distribués ouverts
Protocoles
consensus, équité





Thanks!

Imagine new usages & understand real needs

Applications

Design, V&V Certification

Deployment technologies HW/SW

New tools New process

Performance Cost

Responsible AI:

TRUST & FRUGALITY

will make the difference

francois.terrier@cea.fr