



HAL
open science

Pour une réflexion critique de la régulation éthique et juridique de l'intelligence artificielle

Raphaël Déchaux

► **To cite this version:**

Raphaël Déchaux. Pour une réflexion critique de la régulation éthique et juridique de l'intelligence artificielle. Doctorat. Séminaire critique de l'intelligence artificielle, France. 2022. hal-03815962


HAL Id: hal-03815962

<https://amu.hal.science/hal-03815962>

Submitted on 15 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Pour une réflexion
critique de la
régulation éthique
et juridique de l'IA

Raphaël Déchaux
Maître de conférences de droit public
Institut Louis Favoreu – GERJC
(CNRS UMR 7318)
Université Aix-Marseille



Introduction

Aspect révolutionnaire de l'IA :

- ADM : *automated decision-making*
- En droit : ne fonctionne pas.
- En médecine (par ex.) : fonctionne.

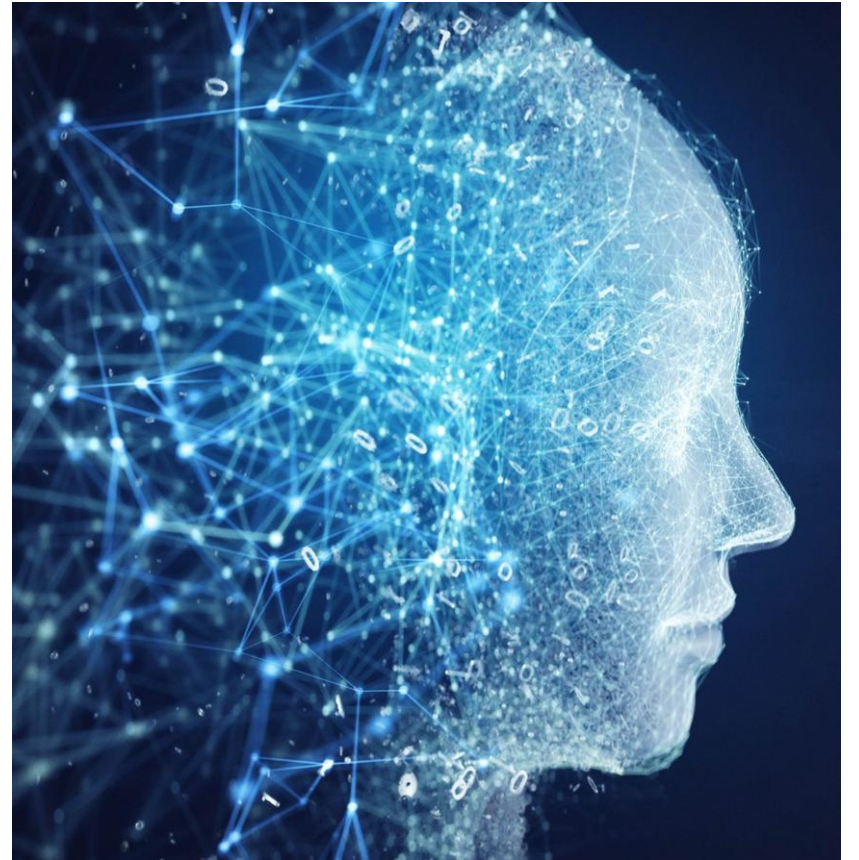


Si une IA ne peut juger mieux qu'un humain, elle peut diagnostiquer mieux qu'un humain

Introduction

Pourquoi est-il nécessaire de réguler l'IA ?

Quels sont les principes éthiques ou juridiques qui permettront de nourrir cette régulation ?

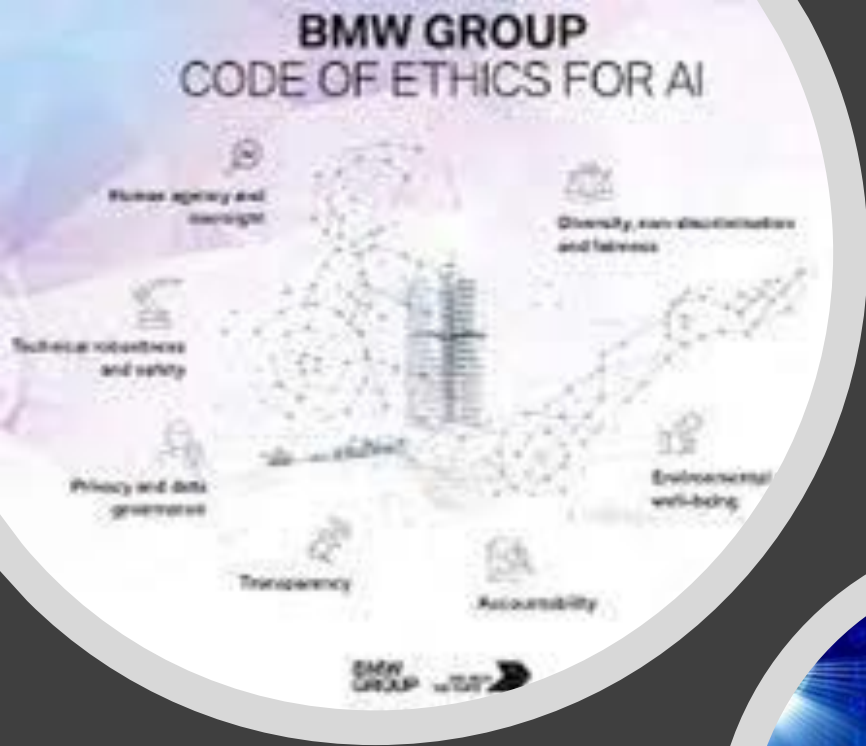


I La nécessité de réguler l'IA



Les
différentes
catégories de
régulation :
données
quantitatives

<https://lestempselectriques.net/index.php/2020/05/06/ia-algorithmes-big-data-data-science-inventaire-des-cadres-ethiques-et-politiques/>



Le développement des textes « éthiques »

- juin 2018 : Google => 7 grands principes sur l'IA.
- janvier 2019 : Facebook => centre de recherche consacré à l'éthique de l'intelligence artificielle
- octobre 2020 : BMW => code d'éthique pour l'utilisation de l'intelligence artificielle

Les différentes catégories de régulation

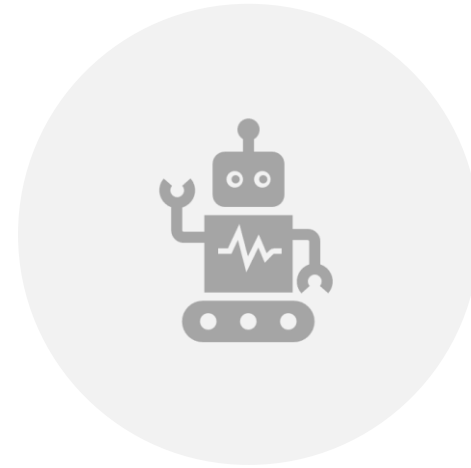
3 types de règles :

- Éthique : produites par des autorités non gouvernementales (morale professionnelle)
- Droit souple : produites par des autorités nationales ou internationales (non contraignantes)
- Droit : produites par des autorités nationales ou internationales (contraignantes)

L'insuffisance des textes éthiques

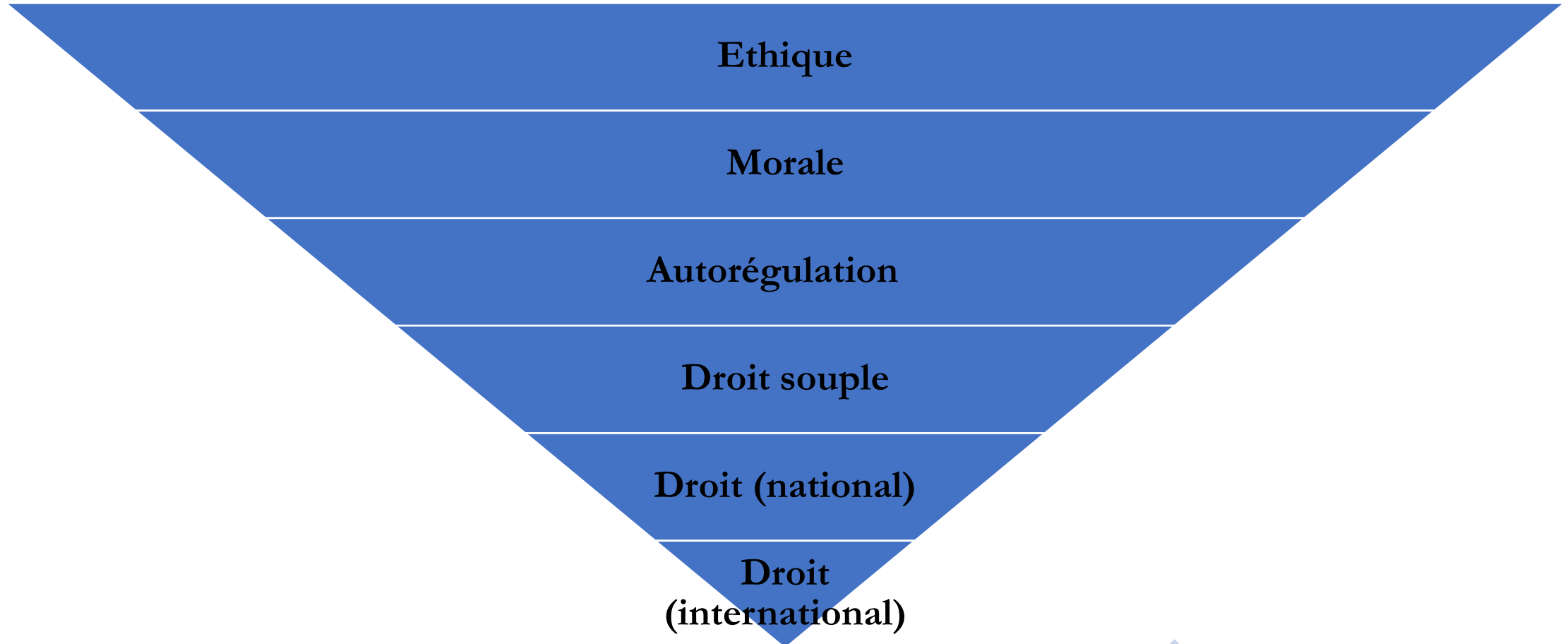


AUTOREGULATION
PAS DE GARANTIE



COMMUNICATION
ETHIC WASHING

Les différentes catégories de régulation de l'IA

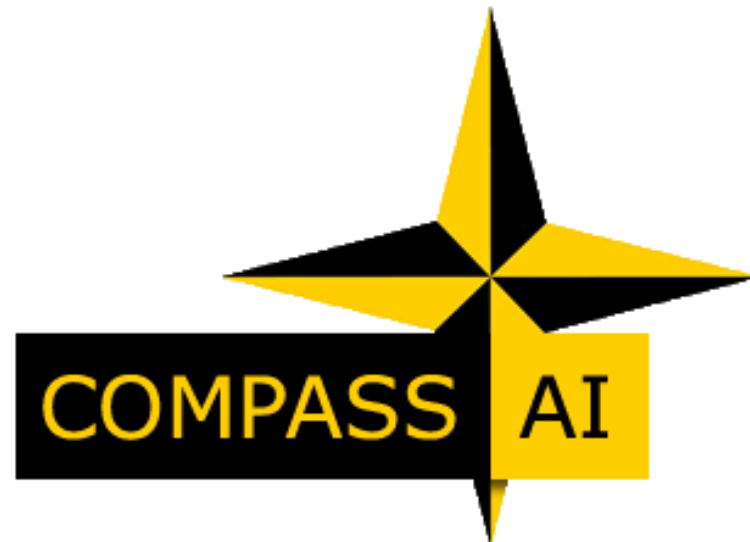


Les risques de l'IA

- Malentendu sur ce que sont les IA
- Biais causés par l'IA



Le risque de discrimination

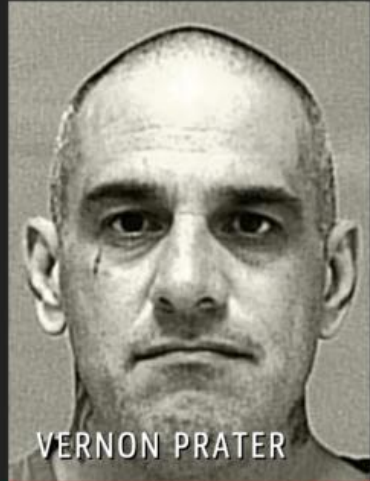





Résultats opérés par COMPAS




Résultats confrontés à la réalité de la dangerosité

 <p>DYLAN FUGETT</p>	 <p>BERNARD PARKER</p>
LOW RISK 3	HIGH RISK 10

 <p>VERNON PRATER</p>	 <p>BRISHA BORDEN</p>
LOW RISK 3	HIGH RISK 8

 <p>DYLAN FUGETT</p> <hr/> <p>Prior Offense 1 attempted burglary</p> <hr/> <p>Subsequent Offenses 3 drug possessions</p>	 <p>BERNARD PARKER</p> <hr/> <p>Prior Offense 1 resisting arrest without violence</p> <hr/> <p>Subsequent Offenses None</p>
LOW RISK 3	HIGH RISK 10

 <p>VERNON PRATER</p> <hr/> <p>Prior Offenses 2 armed robberies, 1 attempted armed robbery</p> <hr/> <p>Subsequent Offenses 1 grand theft</p>	 <p>BRISHA BORDEN</p> <hr/> <p>Prior Offenses 4 juvenile misdemeanors</p> <hr/> <p>Subsequent Offenses None</p>
LOW RISK 3	HIGH RISK 8



**Pourquoi la
régulation
internationale
serait-elle
“meilleure” que
celle nationale ?**

Economie de
l'IA est
mondialisée

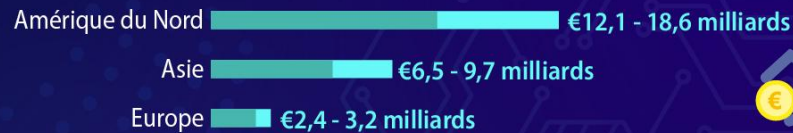
Contexte de la
compétition
internationale

La nécessité de maintenir une compétitivité européenne

Intelligence artificielle

Faits et chiffres

Investissements en IA*



€6 500 - €12 000 milliards/an

Impact économique de l'automatisation du travail intellectuel, des robots et des véhicules autonomes d'ici 2025**

*en 2016 | **estimation

Sources : Commission européenne (2019), IPOL (2020)



europarl.eu

Demandes de brevet IA*



Augmentation de **400%** des demandes de brevet IA publiées durant la dernière décennie

Part de l'IA dans la transition numérique et verte



60 millions de nouveaux emplois

pourraient être générés à travers le monde par l'IA et la robotique d'ici 2025

L'UE permet de créer des avantages pour les citoyens et les entreprises en finançant des projets IA par le biais de la Stratégie européenne en matière de compétences, le Fonds pour une transition juste, Horizon, etc.

*Pour la période 1960 - 2018

Sources : Commission européenne (2019), IPOL (2020)



europarl.eu





La régulation existante

Les textes de droit souple :

- OCDE *Recommendation on Artificial Intelligence* 22 May 2019
 - UNESCO *Recommendation on the Ethics of Artificial Intelligence* 24 November 2021.
-

EUROPEAN COMMISSION FOR THE EFFICIENCY OF JUSTICE (CEPEJ)

European ethical Charter
on the use of Artificial Intelligence in
judicial systems and their environment



Adopted at the 31st plenary meeting
of the CEPEJ (Strasbourg, 3-4 December 2018)

le Conseil de l'Europe

1. Principle of respect for fundamental rights
2. Principle of non-discrimination
3. Principle of quality and security
4. Principle of transparency, impartiality and fairness
5. Principle “under user control”

La régulation en projet



Des textes de droit
contraignant

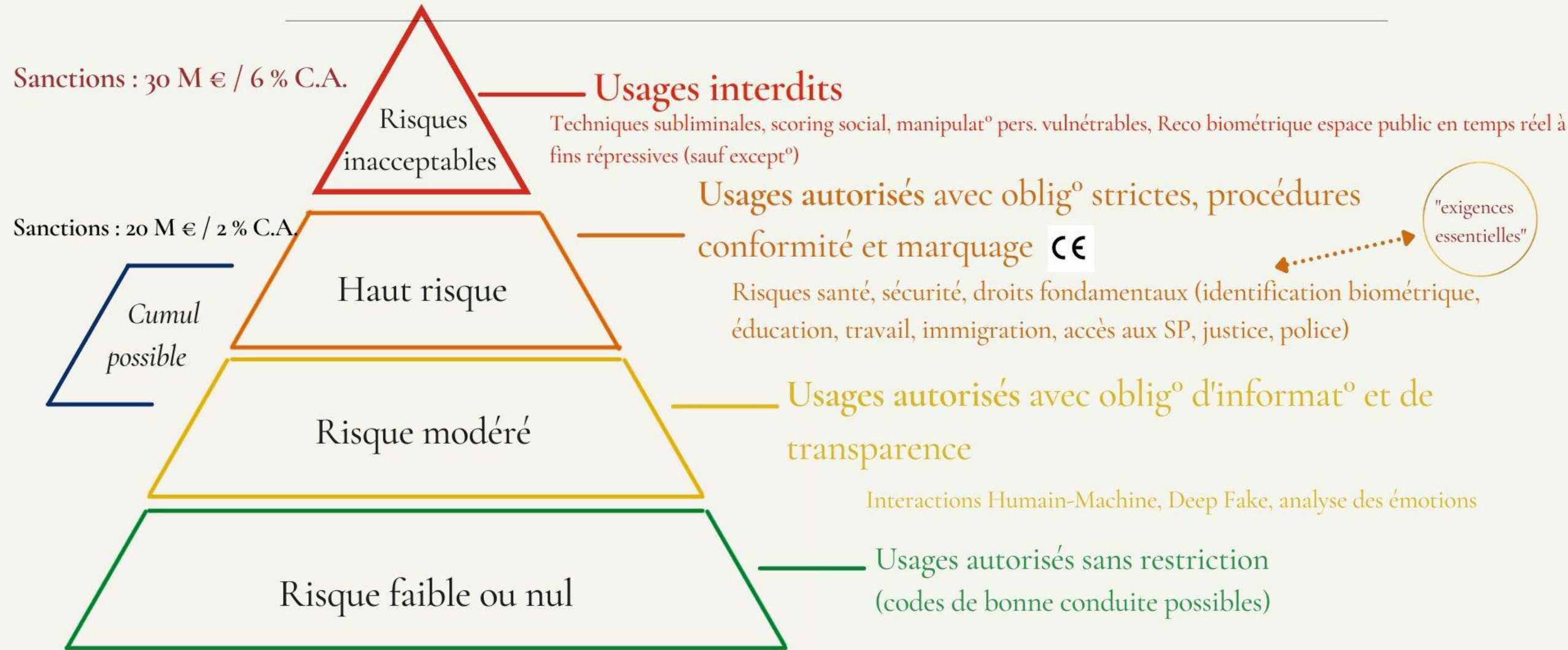
- **Conseil de l'Europe:**
Convention sur l'IA
- **Union européenne**
: *IA Act* (en cours de négociation)

Critiques des projets de régulation européenne

- Régulation fondée sur le risqué
- Objectif : la confiance dans l'IA
- Pas de droits individuels



LA PYRAMIDE DES RISQUES SELON LA PROPOSITION IA ACT 21.04.21



Measure

if your organisation's AI is

trustworthy



ALTAI – Assessment List for Trustworthy Artificial Intelligence

Les textes de droit contraignant au niveau national

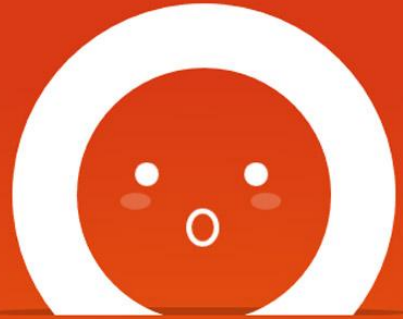
Loi relative à la bioéthique
du 2 août 2021

- vise les IA de *deep learning*
- consentement du patient
- liberté d'utilisation
- pas de « garantie humaine »





Quelle définition
juridique de l'IA ?



Hello, je suis **OUIbot**, ton assistant de voyage. Avec moi, **une phrase** suffit pour trouver ton train !

C'est parti !



GPT-3

Catégorisation juridique de l'IA

algorithmes opérant

- dans un *système fermé composé de règles objectives et quantifiables* ≠
- Dans un *système ouvert subjectif nécessitant une interprétation ou une contextualisation*

algorithmes

- à *prétention descriptive* ≠
- à *prétention prescriptive*

Définitions européennes de l'IA

Recommandation de l'OCDE : « Un système d'IA est un système automatisé qui, pour un ensemble donné d'objectifs définis par l'homme, est en mesure d'établir des prévisions, de formuler des recommandations, ou de prendre des décisions influant sur des environnements réels ou virtuels. Les systèmes d'IA sont conçus pour fonctionner à des degrés d'autonomie divers ».

IA Act définit un système d'IA comme « un logiciel développé à l'aide d'une ou plusieurs des techniques et approches énumérées à l'annexe I et capable, pour un ensemble donné d'objectifs définis par l'homme, de générer des résultats tels que du contenu, des prédictions, des recommandations ou des décisions influençant les environnements avec lesquels ils interagissent ».

Faut-il interdire par principe certaines intelligence artificielle ?

Interdictions

- les IA de reconnaissance émotionnels
- les IA de notation sociales
- les IA judiciaires
- IA militaires complètement autonome (robots tueurs)

Interdictions relatives

- les IA éducatives
- Les IA psychographiques
- les IA d'augmentation des capacités sensorielles
- Toutes les IA qui visent à reproduire la cognition humaine.



II LES PRINCIPES PERMETTANT D'ENCADRER L'IA

A) Responsabilité et transparence

Qui est le responsable ?

Implique le droit à la transparence de l'IA

- Secret des affaires
- IA d'apprentissage automatique

Implique de réfléchir à l'étendu du préjudice subi

- Identité cognitive
- Quel consentement éclairé ?
- Relation IA – cerveau humain

B) Dignité de la personne humaine, neurotechnologies et transhumanisme

Questions classique et moderne : intégrité du corps humain

Domaine thérapeutique : traitement des maladies graves

Domaine « amélioratif » : capacités morales, atténuer l'agressivité ou augmenter les capacités d'apprentissage

- Normalité subjective
- Réification (chosification) de l'être humain : eugénisme

Domaine commercial : neuromarketing



Fin