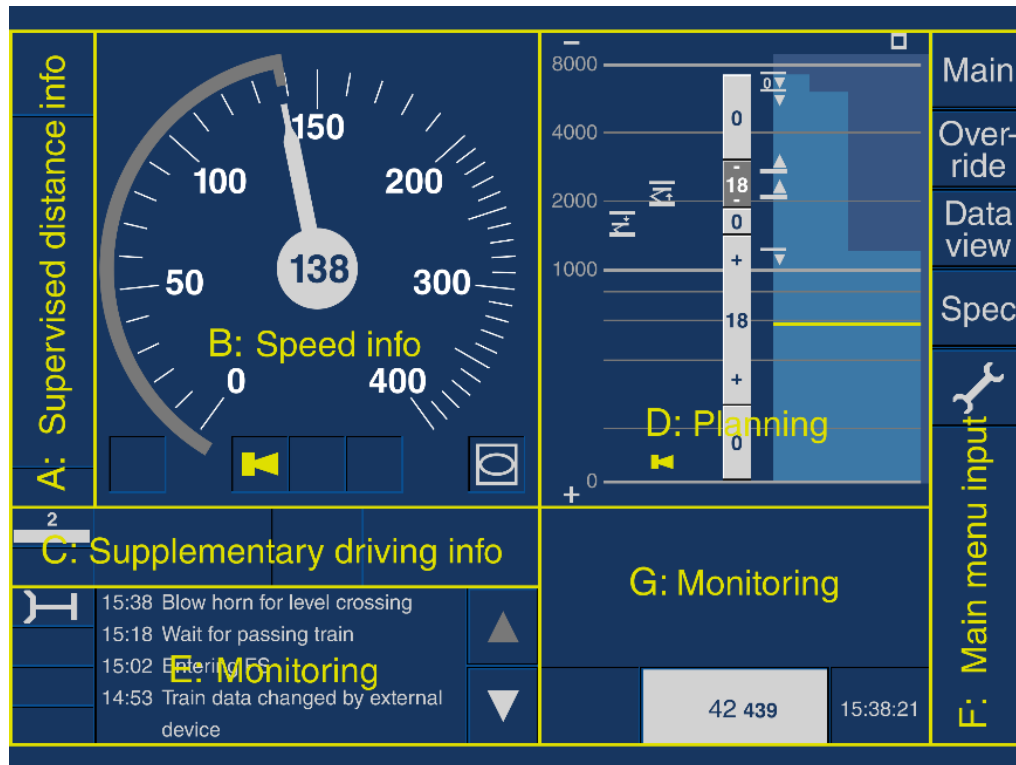


L'IA au secours de l'IS

lgm / leading great  
movements

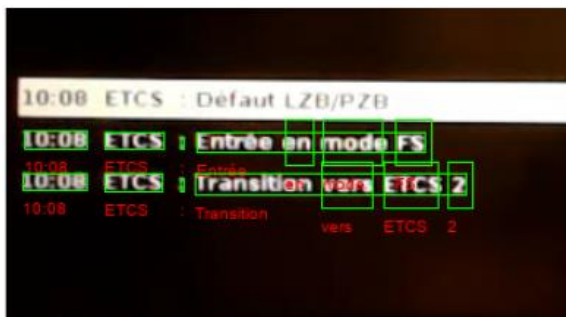




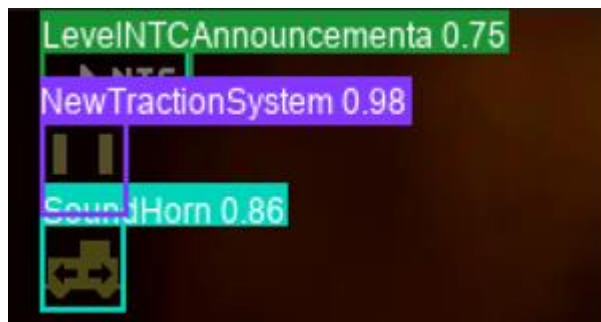
# ▶ PLUSIEURS MODALITÉS

Igm/

Reconnaissance  
texte



Reconnaissance  
pictogrammes



Analyse cadran



L'IS au secours de l'IA

lgm / leading great  
movements

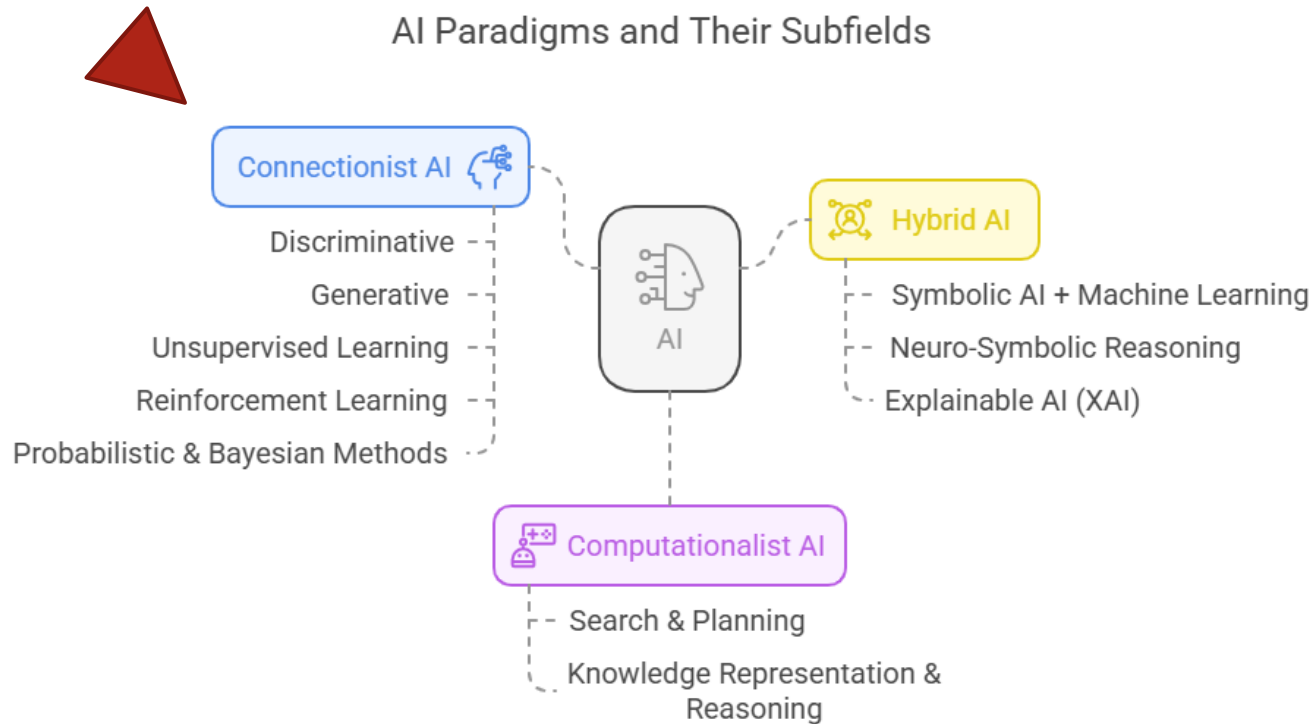






# /01

DE QUELLE IA PARLE T'ON ?



## Cas d'usages génériques

### IA Générative

Utilisé pour créer, analyser ou transformer des données, améliorant l'innovation et l'efficacité.

### Classification/ Clustering

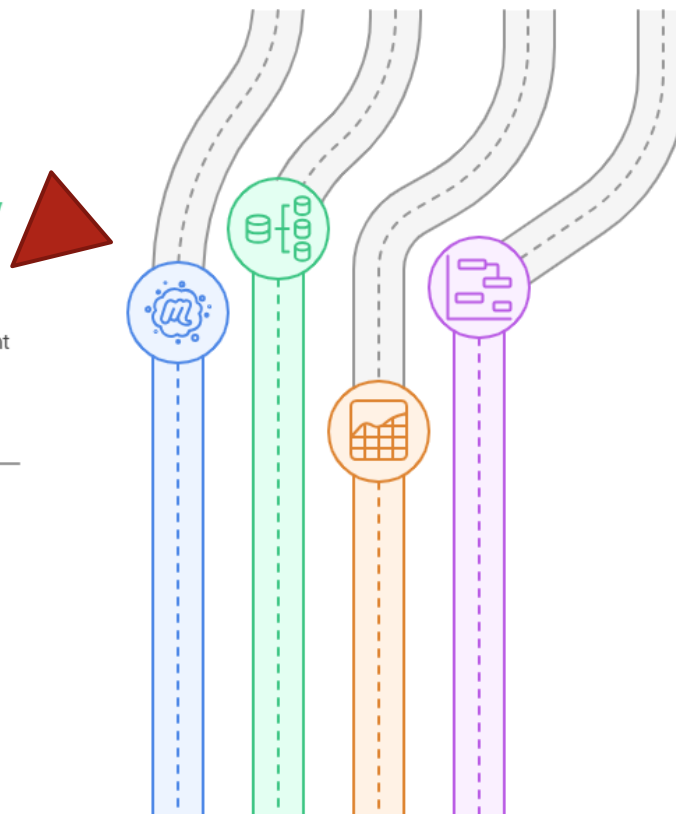
Classifie les données en catégories, facilitant l'organisation et l'analyse des données.

### Régression

Prédit les résultats basés sur des données historiques, soutenant la prise de décision proactive.

### Planification

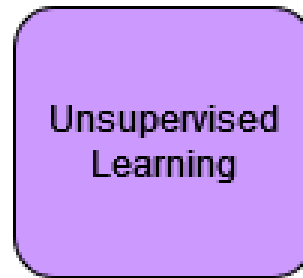
Optimise les parcours ou les ressources sous contraintes, améliorant la gestion des ressources.



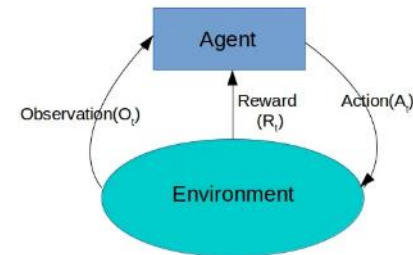
- Les tâches d'apprentissage automatique sont généralement regroupées en trois grandes catégories.



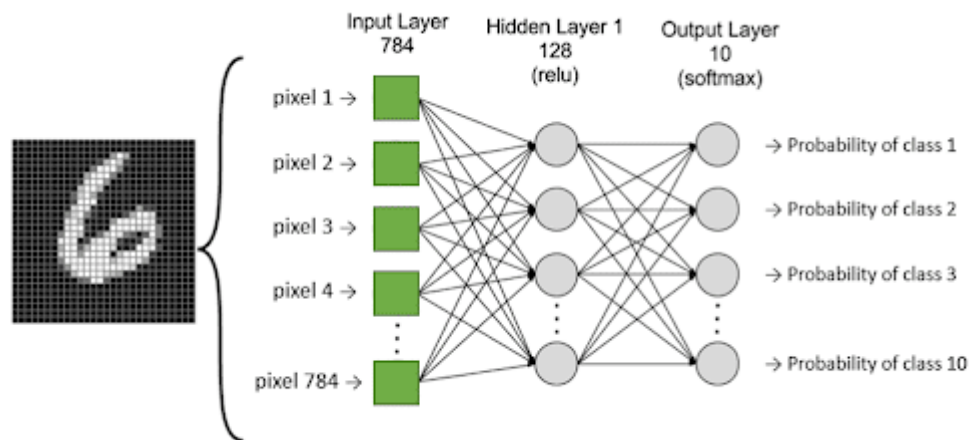
7210414959  
0690159784  
9665407401  
3134727121  
1742351244  
1 7 4 2 3 5 1 2 4 4



7210414959  
0690159784  
9665407401  
3134727121  
1742351244





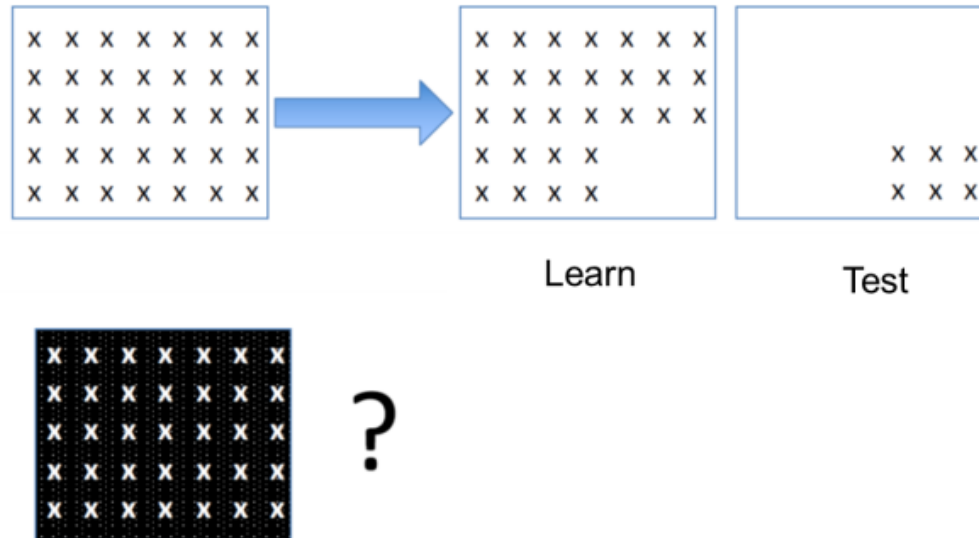




# /02

POURQUOI S'INQUIETER DE L'IA ?

# ▶ QUELQUES LIMITES





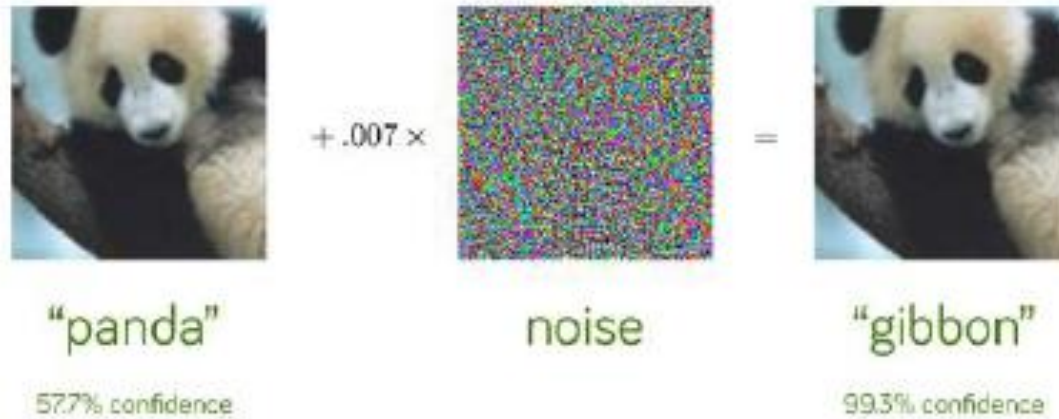
**Figure 11: Raw data and explanation of a bad model's prediction in the "Husky vs Wolf" task.**

	Before	After
Trusted the bad model	10 out of 27	3 out of 27
Snow as a potential feature	12 out of 27	25 out of 27

**Table 2: "Husky vs Wolf" experiment results.**

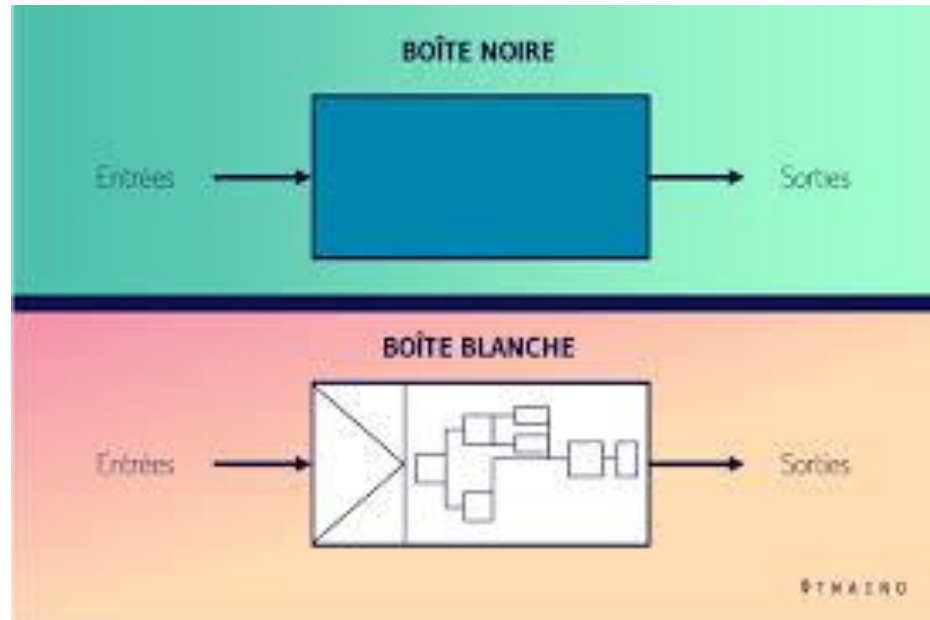
## ▶ QUELQUES LIMITES

Igm/





# ▶ QUELQUES LIMITES



# L'IA c'est du Logiciel !

Du Logiciel intégré dans des systèmes  
HW et SW « classiques »

Déploiement croissant de l'IA dans tous les  
domaines et secteurs

**Appliquer toutes les  
bonnes Pratiques de  
développement  
logiciel existantes**

# ▶ QUID DES NORMES LIÉES À LA SAFETY LOGICIELLE ?

- ✓ Déjà envisagé, avec l'ISO 26262 par exemple
  - Identifier les dangers
  - Fautes et modes de défaillances
  - Utilisation de sets d'apprentissage
  - Niveau de l'usage du Machine Learning
  - Techniques logicielles



/03

ELEMENTS DE REPONSES



# L'IA c'est du Logiciel !

Du Logiciel intégré dans des systèmes  
HW et SW « classiques »

Déploiement croissant de l'IA dans tous les  
domaines et secteurs

**Appliquer toutes les  
bonnes Pratiques de  
développement  
logiciel existantes**

**Puis une approche  
complémentaire pour  
les spécificités du  
Machine Learning**



## SPECIFICATIONS IMPRECISES



- Les spécifications des applications d'IA(ML) sont non explicites
- Les spécifications sont basées sur des exemples dont la qualité et la représentativité ne sont pas démontrées

## EFFET BOITE « NOIRE »



- La taille des modèles empêchent la compréhension de leur fonctionnement interne
- Les performances de l'IA(ML) ne sont pas formellement démontrables

## STANDARDS LOGICIELS INCOMPLETS



- Les standards logiciels (IEC61508,...) s'appliquent difficilement à l'IA,
- Les architectures d'IA/ML évoluent très rapidement (exemple avec les LLM)

## DONNEES D'APPRENTISSAGE



- La représentativité des données d'apprentissage par rapport à l'environnement réel est difficile à évaluer, en particulier pour les événements rares ou critiques
- Les conditions d'exploitation de l'IA peuvent évoluer dans le temps

## VALIDATION & ROBUSTESSE



- Les algorithmes d'IA adressent des problèmes non abordables par des algorithmes classiques, mais ne peuvent prétendre à une fiabilité de 100%
- La robustesse est difficile à vérifier
- Le processus de validation est complexe

## EXPLICABILITE



- Les algorithmes de ML produisent des résultats basés sur des processus difficilement explicables aux utilisateurs
- L'explicabilité est liée aux profils des utilisateurs

## SPECIFICATIONS IMPRECISES



Appliquer une démarche d'IS

## EFFET BOITE « NOIRE »



Maîtriser la conception et contrôler le comportement de l'IA

## STANDARDS LOGICIELS INCOMPLETS



Etablir de nouveaux standards dans le contexte de l'AI ACT



## DONNEES D'APPRENTISSAGE



Maîtriser le contexte d'utilisation (ODD)

## VALIDATION & ROBUSTESSE



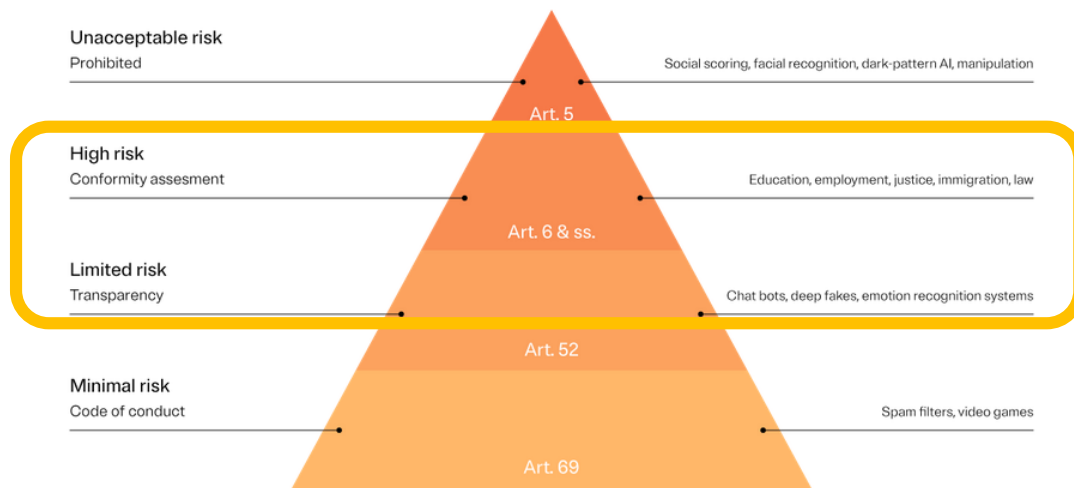
Utiliser des outils ou méthodes dédiés aidant à la validation

## EXPLICABILITE



Construire la collaboration Homme - IA

# ▶ AI ACT : FRAMEWORK GLOBAL EUROPEEN



- **Unacceptable Risk** — Social scoring systems, real-time remote biometric verification.

- **High Risk** — Credit scoring systems, automated insurance claims .

For processes that fall into this bucket, companies need to conduct a conformity assessment and register it with an EU database before the model is available to the public.

Apart from this, these high-risk processes require detailed logs and human oversight as well.

- **Limited Risk** — Chatbots, personalization.

For limited risk processes, companies need to ensure that they're being completely transparent with their customers about what AI is being used for and the data involved.

- **Minimal Risk** — For any processes that companies use that fall into the "minimal risk" bucket, the draft EU AI Act encourages providers to have a code of conduct in place that ensures AI is being used ethically.

# ▶ LE CONTEXTE NORMATIF AUTOUR L'AI ACT



JTC1 / SC 42  
SC65A/JWG21

S'appuie

Normatif  
« horizontal »



Corpus « horizontal » - hEN

Conformité légale

Mandate



Régulateur européen  
Institutionalise



IA ACT

Conformité légale

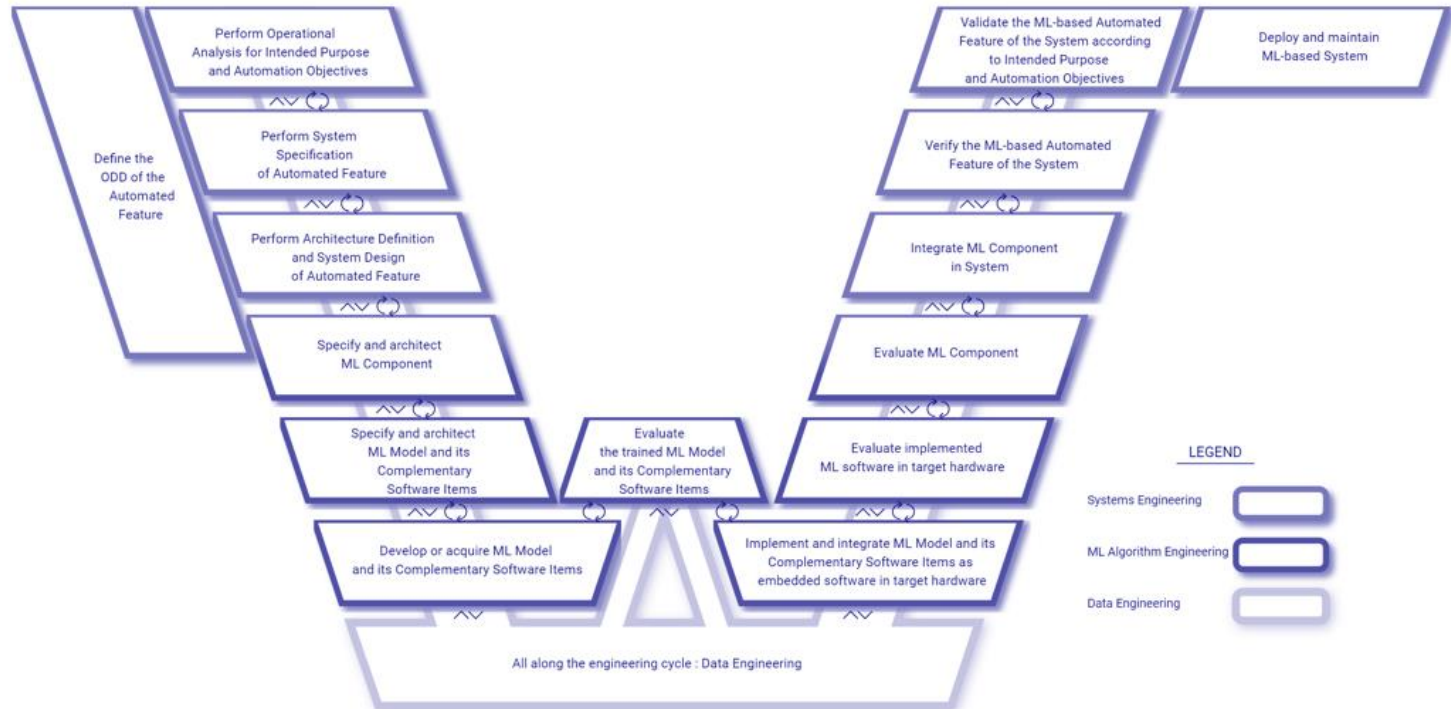
Normatif  
« Vertical »



**Applicabilité 2026**

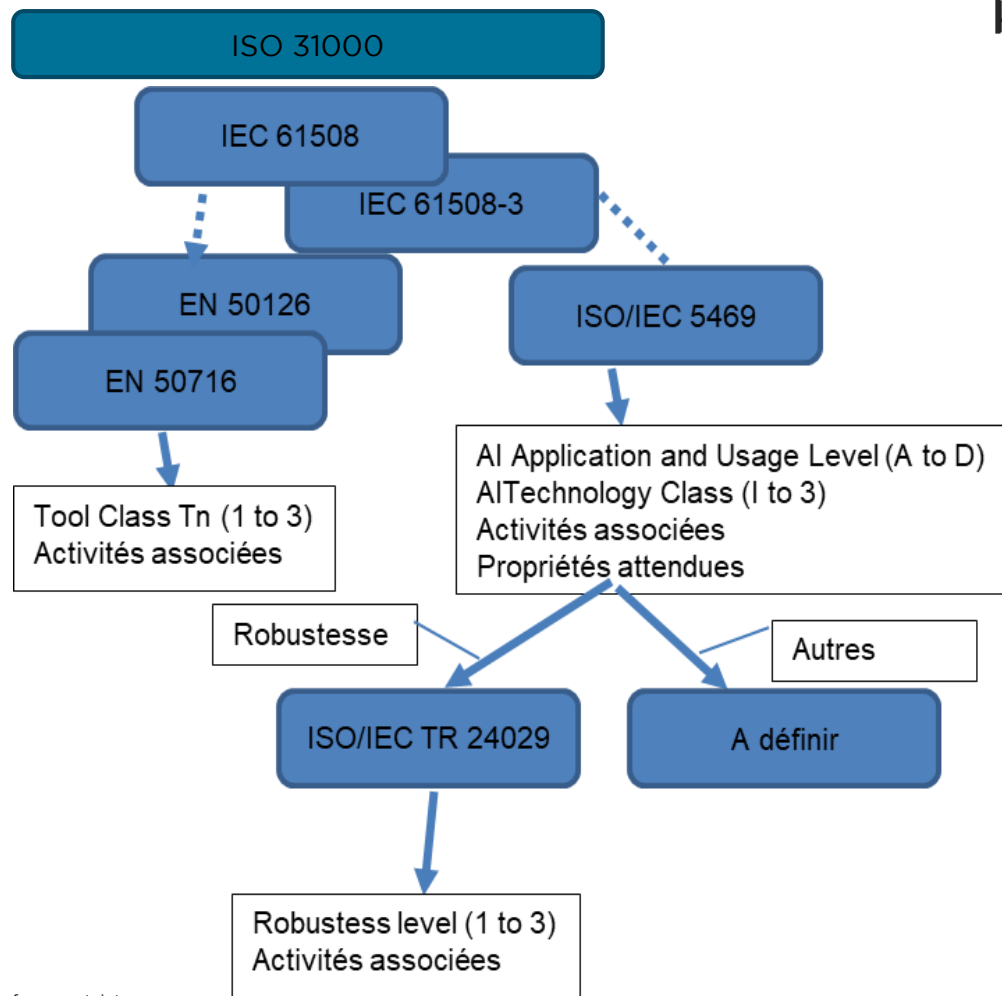
Secteurs – corpus « verticaux »

# Le “Body of Knowledge”

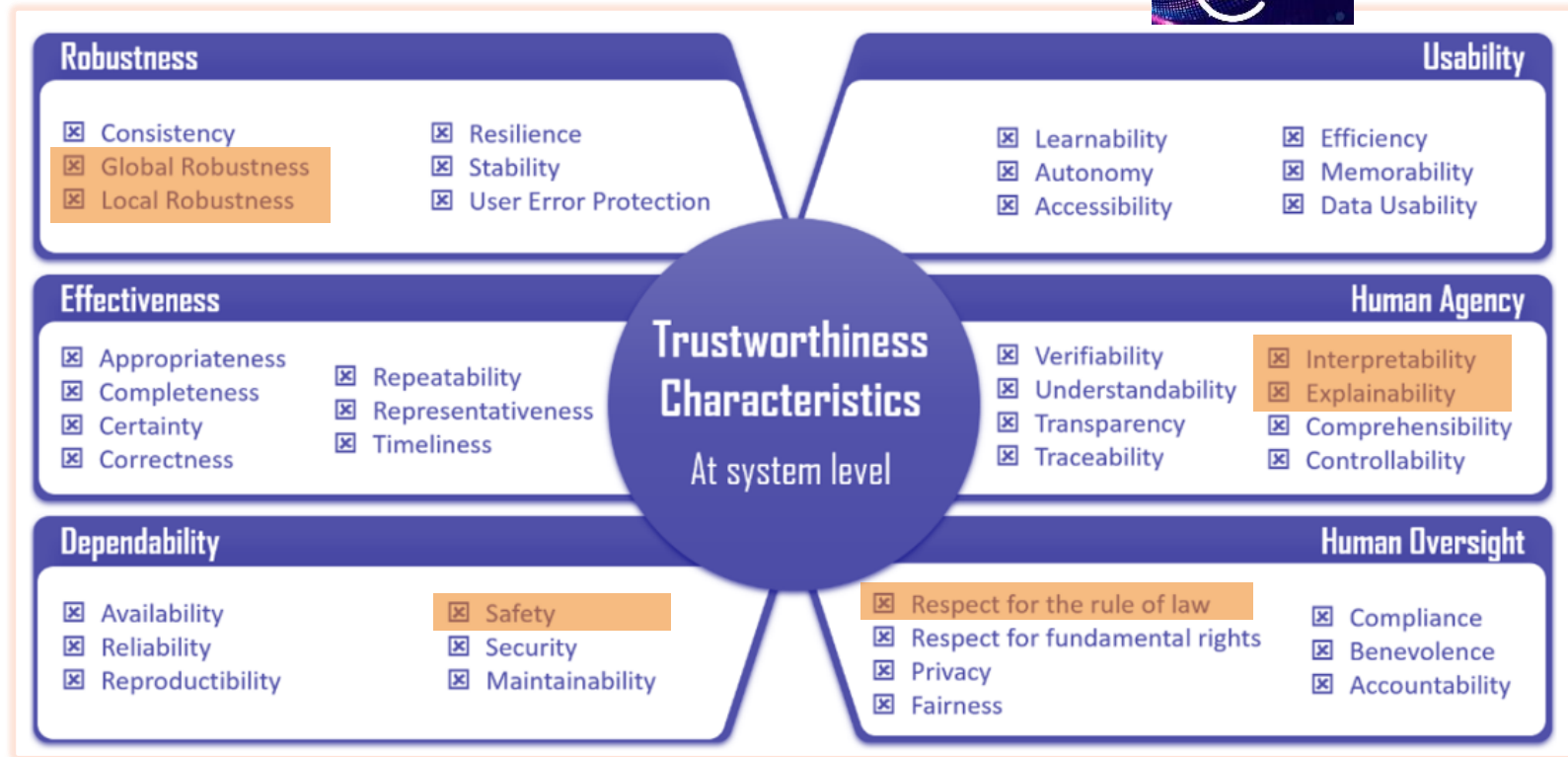


<https://www.trustworthy-ai-association.eu/our-methodology-and-tools/>





# Les principaux axes de travail

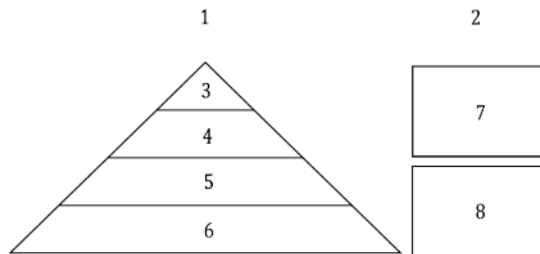




/04

DEMARCHE

AI Technology Class = > AI application and usage level	AI technology Class I	AI technology Class II	AI technology Class III
Usage Level A1 <sup>a</sup>	Application of risk reduction concepts of existing functional safety International Standards possible	Appropriate set of requirements <sup>c</sup>	At the time of publication of this document no appropriate set of properties with related methods and techniques is known to achieve sufficiently reduction of risk
Usage Level A2 <sup>a</sup>		Appropriate set of requirements <sup>c</sup>	
Usage Level B1 <sup>a</sup>		Appropriate set of requirements <sup>c</sup>	
Usage Level B2 <sup>a</sup>		Appropriate set of requirements <sup>c</sup>	
Usage Level C <sup>a</sup>		Appropriate set of requirements <sup>c</sup>	
Usage Level D <sup>b</sup>	Application of risk reduction concepts of existing functional safety International Standards		
<sup>a</sup> Static (offline) (during development) teaching or learning only.			
<sup>b</sup> Dynamic (online) teaching or learning possible.			
<sup>c</sup> The appropriate set of requirements for each usage level is established by application of risk reduction concepts of existing functional safety International Standards and additional considerations based on the literature review performed in <a href="#">Clauses 8, 9, 10 and 11</a> . Examples are provided in <a href="#">Annex B</a> .			



➤ 40 items

➤ Hardware

8 - Librairies

cuDNN

OpenCV

Tesseract

Leptonica

Pango

Cairo

ICU

Zlib

7 – Framework de Machine Learning

TensorFlow

Pytorch

6 – Code ou Compilateur

TensorRT

Python

G++

Autoconf

Automake

Libtool

Pkg-  
config

5 – Hardware

Hors Scope

4 – Executable Machine Code

CUDA

Ubuntu

Jetpack SDK

3 – ML Model

YOLO

Tesseract OCR

- Outil de Développement

Hors Scope

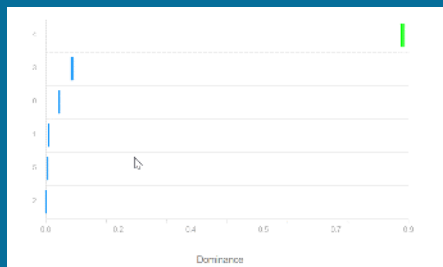
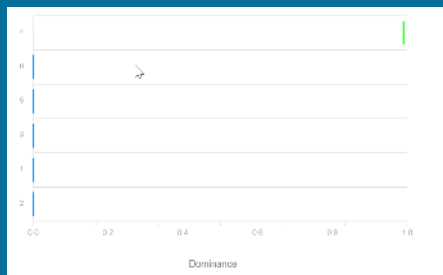


- Contexte : Ecran (Icônes, texte, compteur), LEDs
- Environnement : Illumination, communication
- Eléments dynamiques : densité d'information, chevauchement, Affichage furtif

- **Réponse fausse**, contenant une contre-vérité, une erreur factuelle

- Identification d'apparition d'un icône à tort :
  - Intempestif
  - Classement erroné
  - Superposition d'icônes
- Erreur
  - Identification à tort d'un icône prévu (G4)
  - Identification à tort d'un icône non prévu (G1)
  - Libellé message erroné (G4)
  - Occurrence message non sollicité (G4)

# ▶ ROBUSTESSE / EXPLICABILITE



Initial Picture

Non focus relevance

Focus relevance

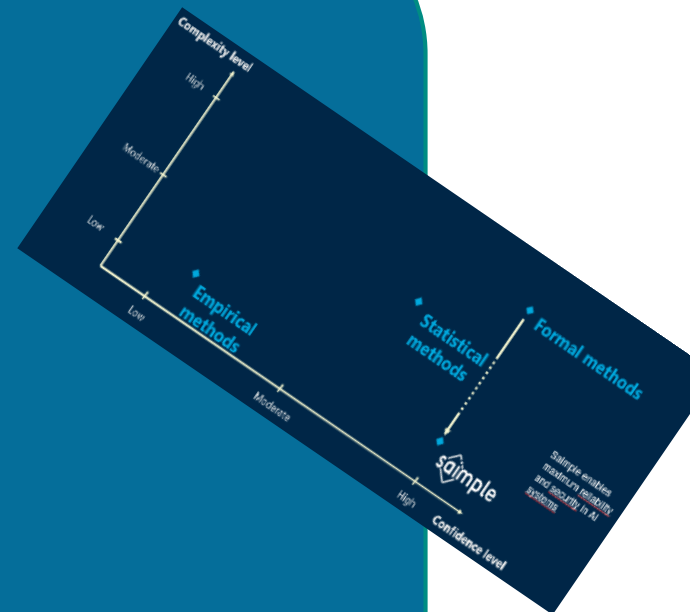
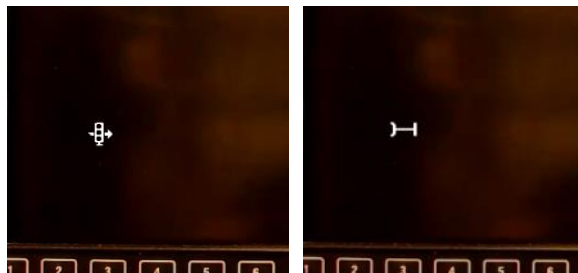


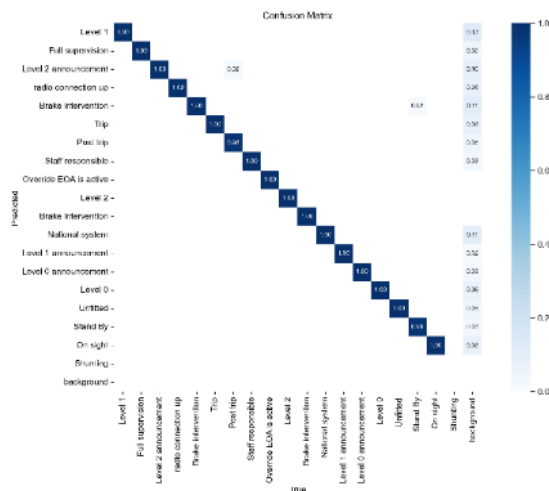
Image avec bruit Gaussien

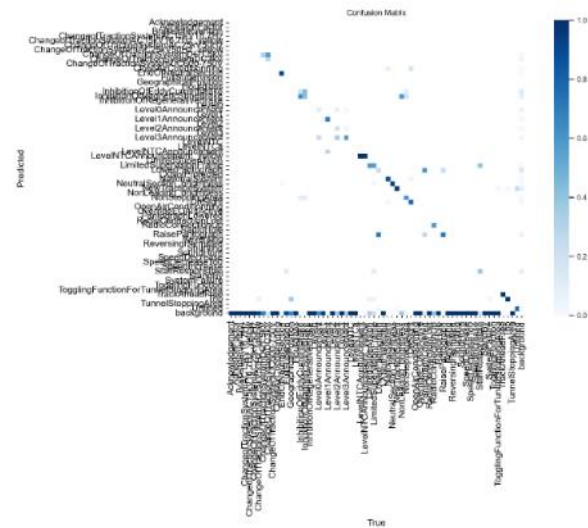
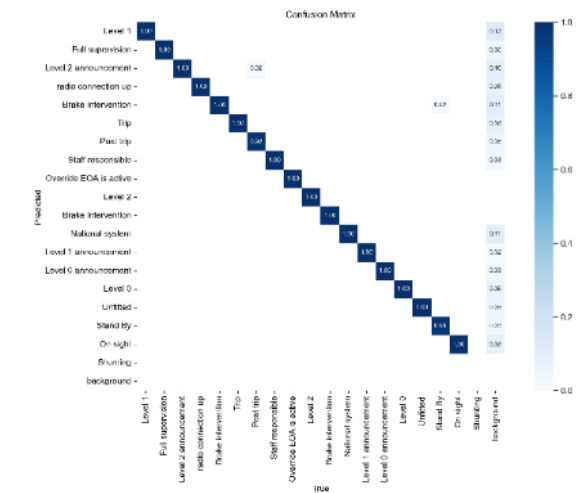
Image avec bruit de reflet

Types de données	Tentatives	Détections Correctes	Taux de réussite
Sans Bruit	78	45	57,7 %
Bruit Gaussien	78	33	42,3 %
Bruits de reflets	78	30	38,5 %
Bruits multiplicatifs	78	29	37,2 %
Bruit Blanc	78	32	41,0 %




## Le projet avant analyse





- Un cycle de développement spécifique
- Au-delà de l'effet Whaou
- Importance de l'ODD et de l'analyse de risque
- Choix du contexte normatif
- Impact majeur : définition du set d'apprentissage





AI ethics and safety,  
AI for social good, and  
sustainable AI have become  
central themes in all major AI  
conferences.

Francesca Rossi, AAAI President 2022-2025