



AfIA – FIIA

**6^{ème} Forum
Industriel de l'IA**

**IA de confiance :
responsabilité,
robustesse,
transparence**

POUR UNE IA DIGNE DE CONFIANCE

De quoi parlons-nous vraiment ?

1. POUR UNE PÉDAGOGIE DE L'INTELLIGENCE ARTIFICIELLE

POUR UNE PÉDAGOGIE DE L'IA

Charte ou
Livre Blanc,
pour un usage
responsable de l'IA

IA = opportunités
+ questionnements

- Le développement de l'IA est **rapide** dans un nombre croissant de secteurs et de cas d'usages, permettant d'**enrichir**, en qualité et en intensité, les **interactions** de l'entreprise avec ses parties prenantes
- Les technologies d'IA sont autant d'**opportunités** d'amélioration de l'expérience client, d'augmentation de la productivité, mais aussi de renforcement des fonctions de compliance et de gestion des risques

*Pour autant, l'usage de l'Intelligence Artificielle soulève de nombreuses **questions** d'ordre épistémologique.*

L'IA nécessite
une attention
particulière

POUR UNE PÉDAGOGIE DE L'IA

Démythifier l'Intelligence Artificielle

IA = objets auxiliaires de
l'intelligence humaine

- Les Intelligences Artificielles conçues et développées aujourd'hui, partout dans le monde, ne sont que des **objets techniques**,
- donc potentiellement **rien d'autre** que des **objets auxiliaires** de l'intelligence humaine

Quels sont les usages de l'IA que les hommes ou les organisations humaines, agissant en responsabilité, s'autoriseront ou s'interdiront ?

POUR UNE PÉDAGOGIE DE L'IA

Acceptabilité de
la technologie et
besoin de
pédagogie

Acceptabilité sociale
= maîtrise + culture

L'**acceptabilité sociale** de toute **technologie** dépend de deux facteurs :

- de la capacité effective des entreprises d'en **maîtriser** la mise en œuvre et les effets
- du degré de **culture** théorique et pratique, scientifique, technique et industrielle, irriguant la société, entourant les innovations les plus récentes

*Expliquer et rendre concret dans la réalité opérationnelle d'aujourd'hui
ce qu'est l'IA tout autant que ce qu'elle n'est pas*

2. QU'EST-CE QUE L'INTELLIGENCE ARTIFICIELLE ?

QU'EST-CE QUE L'INTELLIGENCE ARTIFICIELLE ?

Définition fonctionnelle

« On peut définir un système d'intelligence artificielle, comme un ensemble de technologies, visant à répliquer des capacités cognitives humaines, **pour atteindre les objectifs qui lui sont assignés (*)**, en tenant compte des contraintes et des obstacles de l'environnement ».



Sur le court et moyen termes, la définition est volontairement restreinte à l'IA faible

(*) cette définition exclut volontairement tout système doté d'intentionnalité (i.e. tout système susceptible de s'assigner ses propres objectifs)

QU'EST-CE QUE L'INTELLIGENCE ARTIFICIELLE ?

L'Intelligence
Artificielle est
un artefact



- Comme tout objet façonné par l'homme, un système d'Intelligence Artificielle est un **artefact** aux propriétés uniques
- Un système d'IA constitue ainsi un **phénomène**, c'est-à-dire un **objet technique et social à la fois**, créé de toutes pièces par les conditions expérimentales qui président à son existence et à son fonctionnement

« L'objet technique, en tant que produit des sociétés humaines, incorpore des représentations, des valeurs et des rapports de force » Gilbert Simondon

3. QU'EST-CE QUE L'IA DIGNE DE CONFIANCE ?

QU'EST-CE QUE L'IA DIGNE DE CONFIANCE ?

Une IA licite,
éthique et
robuste



Pour être réputée digne de confiance :

1. une IA doit être **licite**, en assurant le respect des législations et réglementations applicables ;
2. elle doit être **éthique**, en assurant l'adhésion à des principes et valeurs éthiques ;
3. et elle doit être **robuste**, tant sur le plan tant technique que social

La notion d'*Intelligence Artificielle digne de confiance* a été introduite par le *High Level Expert Group* de la Commission européenne.

4. QU'EST-CE QUE L'ETHIQUE ?

QU'EST-CE QUE L'ETHIQUE ?

L'éthique

« La nature de l'équitable est d'être un correctif de la loi [...] De même que la règle () épouse les contours de la pierre et n'est pas rigide, ainsi le décret est adapté aux faits. » Aristote, Ethique à Nicomaque*



() : une règle, faite d'un plomb malléable, utilisée autrefois par les architectes dans les constructions sur l'île de Lesbos*

QU'EST-CE QUE L'ETHIQUE ?

Définition

« L'éthique est la recherche, au travers d'un dialogue raisonnable, d'une mise en pratique la plus **juste** possible de principes relevant de valeurs et du droit.

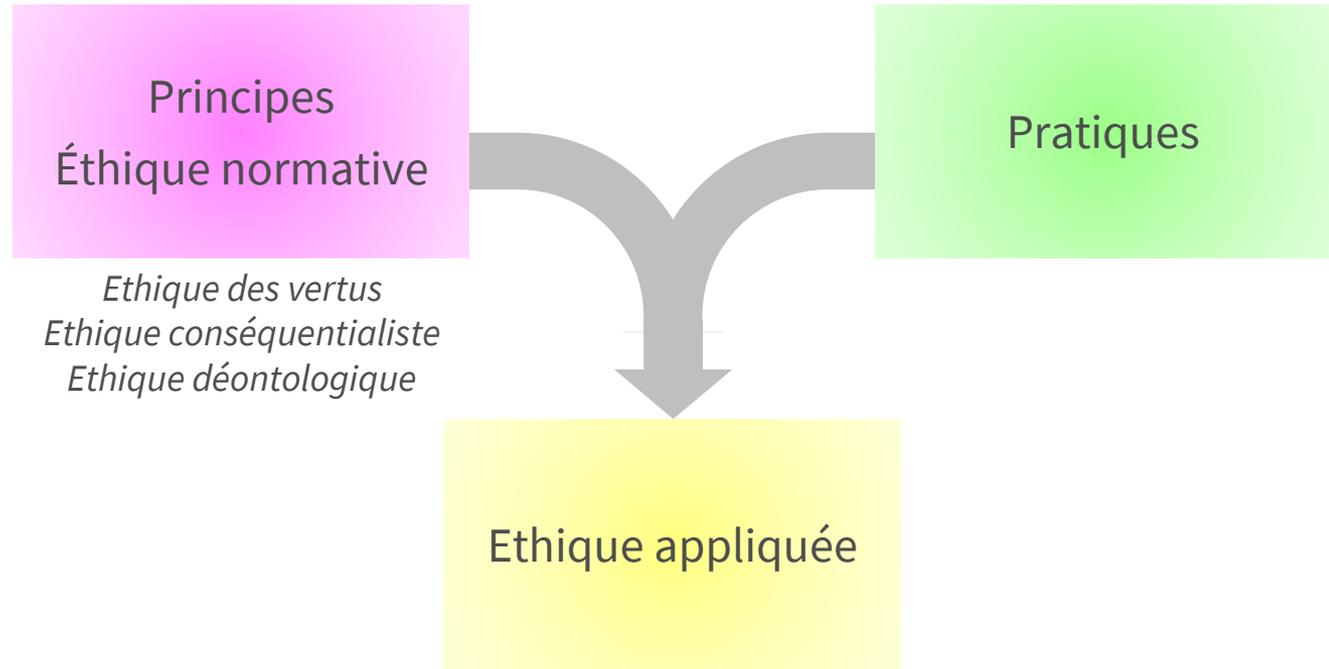


L'**éthique** vise ce qui est **juste**

L'éthique est un *impératif hypothétique* : l'action dépend d'hypothèses et de contraintes. L'éthique, en se soumettant à l'approbation du collectif, renvoie à l'*humanité*

QU'EST-CE QUE L'ETHIQUE ?

**Ethique
appliquée et
éthique
normative**

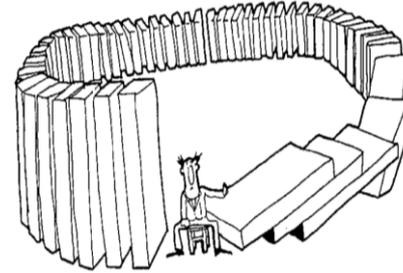


Trois courants

*La résolution de dilemmes éthiques résulte d'une **éthique appliquée***

QU'EST-CE QUE L'ETHIQUE ?

Ethique
appliquée et
éthique
normative



Ethique des vertus
ou **éthique de conviction**
(« en amont »,
« a priori »)

Ethique conséquentialiste
ou **éthique de responsabilité**
(« en aval »,
« a posteriori »)

Ethique déontologique
(« entre amont
et aval »)

Trois courants

QU'EST-CE QUE L'ETHIQUE ?

L'éthique est adossée à la notion de responsabilité



Être responsable, c'est **répondre de soi** et de ses actes devant autrui (approche juridique)



Être responsable, c'est **répondre d'autrui** devant autrui (attention bienveillante, souci pour l'autre, vulnérable et fragile (Emmanuel Levinas, « *Le Visage de l'Autre* »))



Être responsable, c'est **répondre de ce qui nous est confié**, fragile et périssable, devant les générations à venir (Hans Jonas, « *Le principe responsabilité* »)

Trois conceptions

5. APPROCHE PRUDENTIELLE

APPROCHE PRUDENTIELLE

Le socle réglementaire et les dispositifs de gouvernance



Beaucoup d'entreprises disposent déjà de **dispositifs de gouvernance** stricts permettant :

- d'encadrer leurs **processus de décision**,
- de vérifier que ceux-ci reposent sur des **modes opératoires** explicites
- de soumettre les **résultats de calculs** produits par leurs systèmes techniques à des processus d'**homologation**, de **validation**, et de **surveillance continue**, très stricts

La plupart des entreprises ne partent pas de rien !

APPROCHE PRUDENTIELLE

Qu'est-ce qu'une
décision ?



Selon le sens commun et le dictionnaire, une **décision** est à la fois :

1. un **jugement** qui apporte une solution,
2. mais aussi la fin d'une **délibération** (ce qui suppose un **examen conscient** et réfléchi) dans un acte volontaire de faire ou de ne pas faire quelque chose.

En l'absence d'IA forte,
la notion de
« personne robot »
n'a pas de sens !

*A ce jour, **aucune IA ne prend de décisions** : seule une entité réputée responsable (humain ou organisation humaine) est susceptible de le faire*

APPROCHE PRUDENTIELLE

Principe de neutralité technologique

IA = objet technique

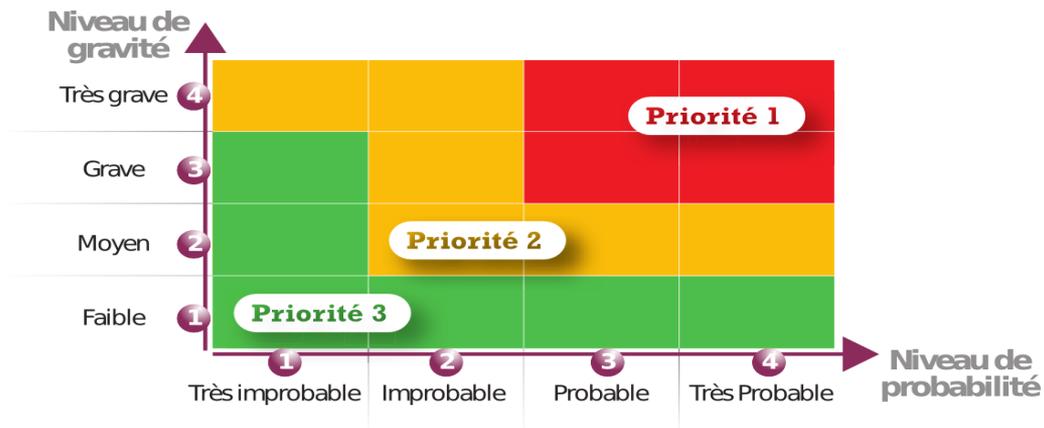
Tout projet de réglementation de l'IA devrait **se limiter** aux exigences pour faire face aux risques et aux problèmes spécifiquement **liés aux usages** de l'IA, pour autant :

- qu'ils ne soient pas déjà couverts par les réglementations existantes
- ou qu'il ne soit pas possible d'ajuster ou de compléter le cadre réglementaire existant

*Telle que nous l'avons définie, l'IA (faible) à laquelle nous faisons référence n'est par essence **rien d'autre qu'un objet technique**, plus précisément, une technologie parmi d'autres avec des spécificités qu'il s'agit de préciser.*

APPROCHE PRUDENTIELLE

Approche par les risques



L'approche par les risques est « robuste » et communément partagée

- Assurer des **conditions de concurrence équitables** entre acteurs économiques, d'autant que la frontière entre secteurs d'activité est quelquefois difficile à établir, ... une même activité pouvant être opérée par différents secteurs d'activité
- Appliquer un **principe de proportionnalité** des moyens à mettre en œuvre (transparence, d'explicabilité, d'auditabilité, etc.) pour opérer le suivi et le contrôle de l'usage des systèmes d'IA.

6. PRINCIPES POUR UNE IA DIGNE DE CONFIANCE

PRINCIPES POUR UNE IA DIGNE DE CONFIANCE

Garantie humaine et contrôle humain



1. S'assurer, en phase d'avant-projet, du **bien-fondé éthique** des projets d'Intelligence Artificielle, tant du point de vue de leur **existence** que de leurs **conditions de mise en œuvre**
2. Exercer une **supervision humaine** tout au long du cycle de vie des systèmes d'Intelligence Artificielle
3. Veiller à ce que la mise en œuvre des systèmes d'Intelligence Artificielle soit respectueuse du **libre arbitre** des utilisateurs

PRINCIPES POUR UNE IA DIGNE DE CONFIANCE

Robustesse
technique et
sécurité



4. Tout au long de leur cycle de vie, s'assurer que les systèmes d'Intelligence Artificielle respectent des critères de **fiabilité**, de **sécurité** et d'**intégrité**

PRINCIPES POUR UNE IA DIGNE DE CONFIANCE

Respect de la
vie privée et
Gouvernance
des données



5. De manière générale et en particulier lors de l'utilisation des produits et services reposant sur des systèmes d'Intelligence Artificielle, **protéger** la **vie privée** et les **données à caractère personnel**

PRINCIPES POUR UNE IA DIGNE DE CONFIANCE

Transparence



6. Mettre à disposition des utilisateurs des **kits d'information**, adaptés à leurs profils, **permettant de comprendre** tant les **caractéristiques générales**, la **finalité** et le **périmètre d'utilisation** des systèmes d'Intelligence Artificielle, que les **dispositifs de gouvernance** qui les encadrent
7. Informer les utilisateurs lorsqu'ils interagissent avec des **chatbots**, plus généralement des agents conversationnels

PRINCIPES POUR UNE IA DIGNE DE CONFIANCE

Diversité, non-discrimination et équité



8. Lorsqu'on développe ou met en œuvre des systèmes techniques, y compris d'Intelligence Artificielle, le faire en respectant des principes d'**inclusion**, de **diversité** et de **non-discrimination**

PRINCIPES POUR UNE IA DIGNE DE CONFIANCE

Bien-être
social et
environnemental



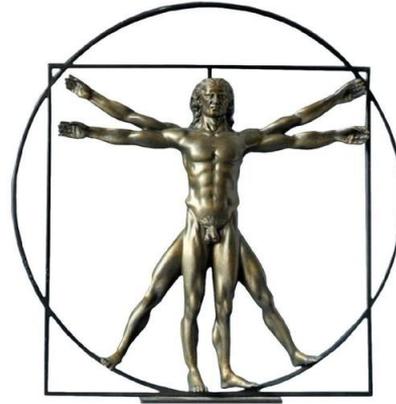
9. Comme pour toutes les technologies, **viser l'efficacité énergétique** des systèmes d'Intelligence Artificielle
10. **Préparer et accompagner les évolutions RH** (compétences, métiers) liées au développement et aux usages des systèmes d'Intelligence Artificielle

7. CONCLUSION

CONCLUSION

Pour une
approche
trans-
disciplinaire

Aborder l'*intelligence artificielle* dans toutes ses dimensions :
technique & **anthropologique**



Introduire / réintroduire, au travers d'une **approche inter- et transdisciplinaire**, l'ensemble des **Humanités** au sein des organisations humaines