WEBINAIRE - 21 janvier 2021

Vers un nouveau cadre de référence de l'Intelligence Artificielle

Intelligence et robotique : principe éthique, juridique, conformité et audit

Accueil/Introduction

William Nahum,

président de l'Académie des Sciences
et Techniques Comptables et Financières

Webinaire >> 21 janvier 2021

Présentation

Serge Yablonsky,
expert-comptable, commissaire aux comptes,
CGEIT, CISA, CRISC, SYC Consultants

Webinaire >> 21 janvier 2021

Présentation

Les intervenants ont préféré le distanciel au bal masqué

Composition du groupe de recherche

- Avocats, enseignants, auditeurs, juristes, tous praticiens de l'audit informatique de la gouvernance des SI
- Un noyau de membres très actifs, intervenant aujourd'hui
- Un panel élargi comprenant des personnalités impliquées dans l'IA

Objectifs du groupe de recherche

- A l'origine l'audit de l'IA « commissariat aux algorithmes et aux données »
- AUDIT = opinion de conformité à un référentiel de **bonnes pratiques** après avoir effectué des travaux conformes à un référentiel d'audit (étapes de travail, outil, tests, …)
- En matière d'IA :
    - Référentiel de bonnes pratiques : plusieurs mais incomplets
    - Référentiel de Gouvernance : COBIT notamment mais incomplet pour l'IA
    - Référentiel d'audit : guide d'audit de COBIT notamment mais incomplet pour l'IA
- Notre approche à ce jour :
    - Emettre des recommandations pour combler les manques des référentiels

Déroulé du webinaire

- Un aperçu des législations internationales actuelles avec un focus sur les projets de règlement du parlement Européen par Alain Bensoussan, avocat,
- Les 5 recommandations du groupe de travail
- Le commentaire de la CNIL par Sophie Nerbonne, directrice chargée de co-régulation économique.
- Débat avec la salle

Le webinaire sera disponible en replay sur le site www.lacademie.info

Une synthèse des travaux sera téléchargeable sur le site www.lacademie.info

Merci de nous aider à illustrer des cas d'audit de l'IA, partagez vos expériences ! Contact : s.yablonsky@syc-france.com

Plan de la conf

Les exigences

Transparence | Equité | Humanité ...

Patrick STACHTCHENKO

La gestion des risques

Non reproductibilité des résultats | Non traçabilité | Biais et erreurs | Responsable ? ...

Camille ROSENTHAL SABROUX

Les responsbilités

Concepteur / développeur | Fournisseur de data | Propriétaire ...

Jean Laurent LIENHARDT

Les bonnes pratiques

Bouton rouge | Tests adaptés et outillés | Traçabilité ...

Claude SALZMAN

L'audit et la certification

Exigences attendues | Niveau de confiance | Opinion ?

Patrick STACHTCHENKO

IA : historique, principes éthiques et juridiques, responsabilité civile

Alain Bensoussan,
Avocat à la Cour,
chargé d'enseignement à Sciences-Po

Webinaire >> 21 janvier 2021

cademie
ES TECHNIQUES COMPTABLES FINANCIERES

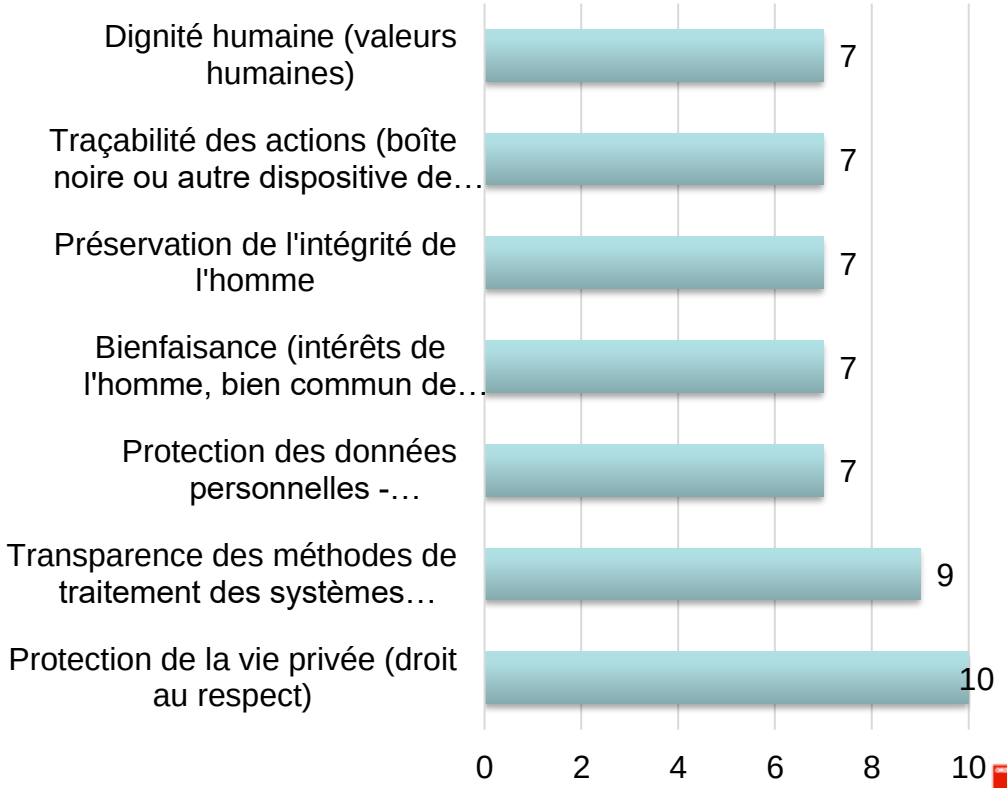
LEXING
ALAIN BENSOUSSAN AVOCATS

[21/01/2021 – WEBINAIRE GRATUIT] VERS UN NOUVEAU CADRE DE RÉFÉRENCE DE L'INTELLIGENCE ARTIFICIELLE

Alain Bensoussan
Avocat à la Cour
Chargé d'enseignement à Sciences Po

Introduction

- Marché mondial
  - IA et Robots
- Initiative
  - Tous les continents
- Problématique
  - Horizontale Verticale Mixte
- Situation générale
  - Réglementation sectorielle



1. Architecture

1. **Un cadre juridique (3 axes)**
   1. Aspects éthiques
   2. Responsabilité civile
   3. Propriété intellectuelle

2. **Un cadre technique (3 axes)**
   1. IA « un système (..) qui fait preuve d'un comportement intelligent »
   2. Technologie « robotique »
   3. Technologies connexes (Capteurs et Commandes)

2. Structure technique

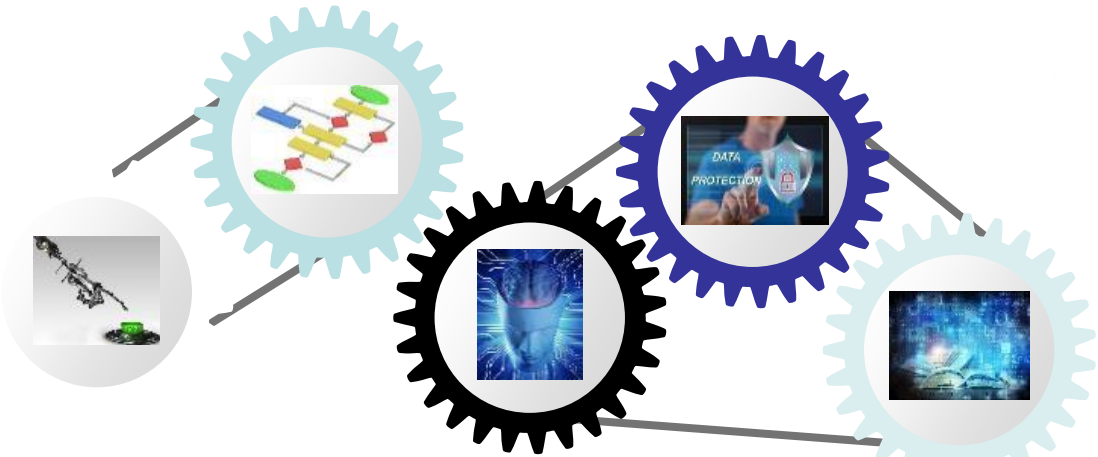
CONCEPT

Algorithme
Logiciel
Autonomie
Robotique
Capteur
Commande

PIVOT
Système
Haut risque
Sectoriel
Finalité

DATA

Données personnelles
Données non personnelles
Lac des données



Faible | Moyenne | Forte

3. Textes

P9_TA(2020)0275

Cadre pour les aspects éthiques de l'intelligence artificielle, de la robotique et des technologies connexes

Résolution du Parlement européen du 20 octobre 2020 contenant des recommandations à la Commission concernant un cadre pour les aspects éthiques de l'intelligence artificielle, de la robotique et des technologies connexes (2020/2012(INL))

P9_TA(2020)0276

Un régime de responsabilité civile pour l'intelligence artificielle

Résolution du Parlement européen du 20 octobre 2020 contenant des recommandations à la Commission sur un régime de responsabilité civile pour l'intelligence artificielle (2020/2014(INL))

P9_TA(2020)0277

Les droits de propriété intellectuelle pour le développement des technologies liées à l'intelligence artificielle

Résolution du Parlement européen du 20 octobre 2020 sur les droits de propriété intellectuelle pour le développement des technologies liées à l'intelligence artificielle (2020/2015(INI))

3. Structure juridique (1)

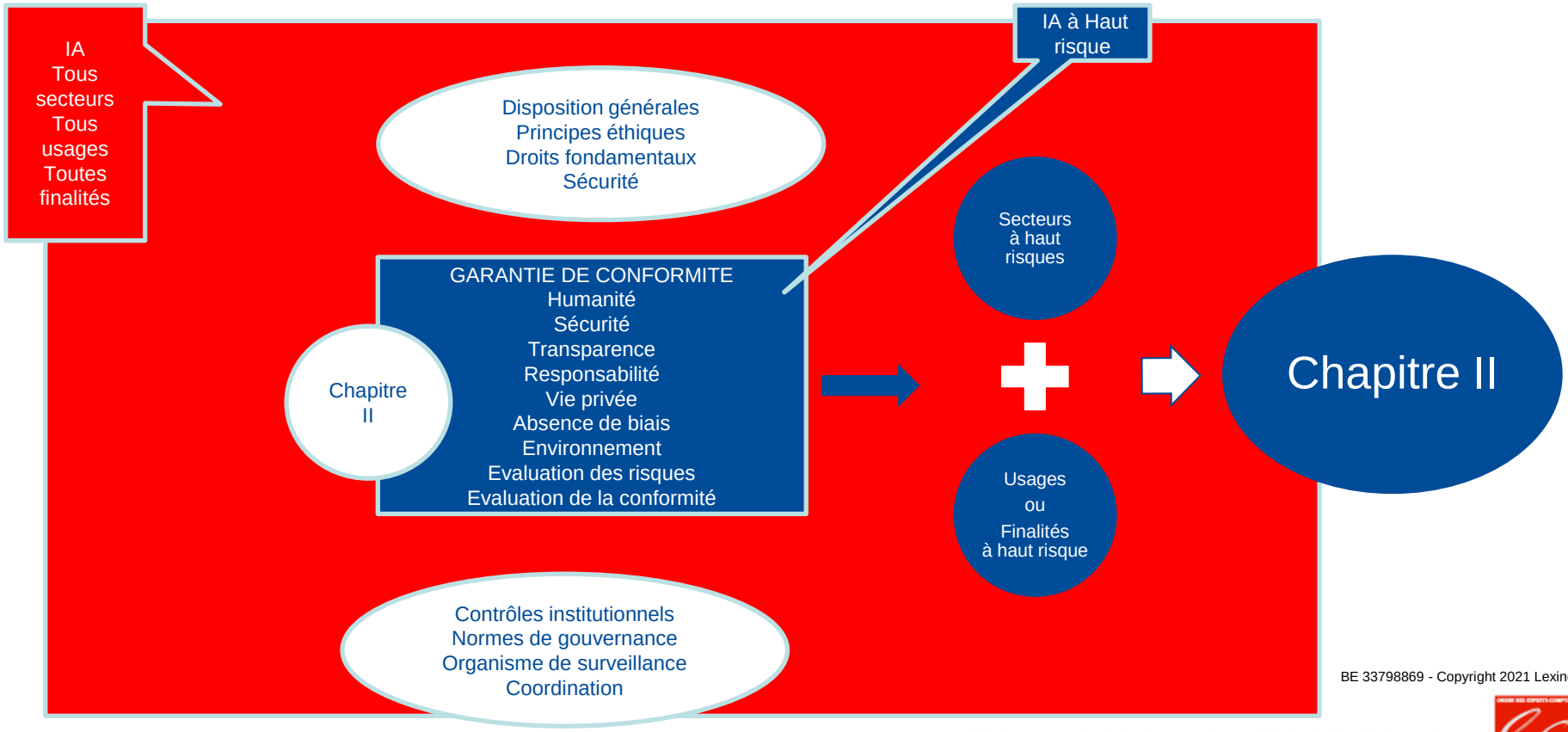
1. Deux régimes juridiques
    1. Général : principes éthiques et les droits fondamentaux
    2. Spécial : haut risque, garantie de conformité et droits particuliers
2. Régulation
    1. Normes de gouvernance
    2. Organisme de surveillance
    3. Coordination UE

ANNEXE

Liste exhaustive et cumulative des secteurs à haut risque et des usages ou finalités à haut risque comportant un risque de porter atteinte aux droits fondamentaux et aux règles de sécurité.

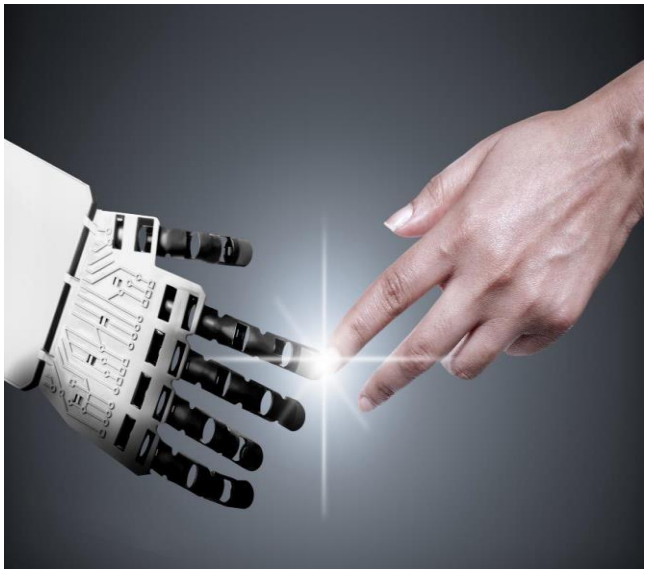
Secteurs à haut risque	• Emploi • Éducation • Soins de santé • Transports • Énergie • Secteur public (asile, migration, contrôles aux frontières, système judiciaire et services de sécurité sociale) • Défense et sécurité • Finance, banque et assurance
Usages ou finalités à haut risque	• Recrutement • Notation et évaluation des étudiants • Affectation de fonds publics • Octroi de prêts • Commerce, courtage, fiscalité, etc. • Traitements et procédures médicaux • Processus électoraux et campagnes politiques • Décisions du secteur public ayant une incidence importante et directe sur les droits et obligations des personnes physiques ou morales • Conduite automatisée • Gestion du trafic • Systèmes militaires autonomes • Production et distribution d'énergie • Gestion des déchets • Contrôle des émissions

3. Structure juridique (2)



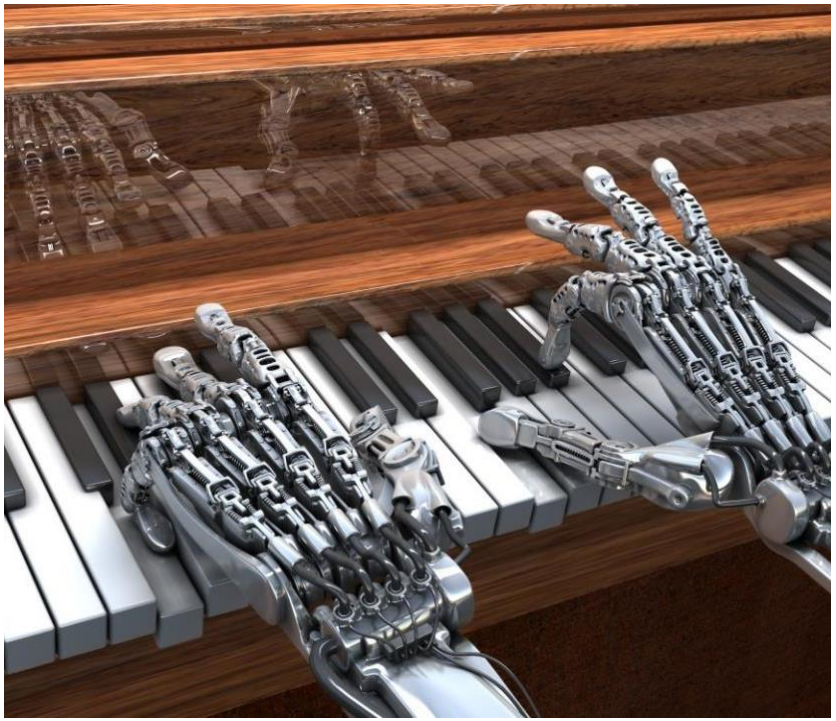
4. Déclinaison : Responsabilité

1. **Responsabilité objective des opérateurs**
   1. IA à haut risque
2. **Indemnisation**
   1. Deux millions (Décès et préjudices corporels)
   2. Un million (Préjudices immatériels)
3. **Prescription**
   1. 10 ans (Dommage)
   2. 30 ans (Exploitation)



5. Déclinaison : Propriété intellectuelle

1. **Orientation et non un projet de règlement**
    1. Suggestion d'un règlement
2. **Absence de réglementation spécifique**
    1. Analyse d'impact pour les droits de PI
3. **Open data**
    1. Exigence de développement
4. **Absence de reconnaissance des droits d'une IA**
    1. Qualité humaine de la création



5 recommandations

>> Les exigences
>> La gestion des risques
>> Les responsabilités
>> Les bonnes pratiques
>> L'audit et la certification

Webinaire >> 21 janvier 2021

Les exigences

Patrick Stachtchenko,
Chargé de cours en Gouvernance, Cyber Sécurité, Audit IA et Référentiels, CGEIT, CRISC et CISA

Confiance/Exigences : Socle commun de compréhension

Quelles exigences (a minima / pour qui / pour quel contexte)? Quels Qualificatifs / Caractéristiques ? Quelles règlementations / contractualisations ?

- Traditionnelles mais avec fortes spécificités IA :
  - **Performance** (finalités/usages): Efficacité, Efficience, Pertinence, Minimisation, Proportionnalité,
  - **Fiabilité** : Sincérité, Loyauté, Conformité, Régularité, Exactitude, …
  - **Sécurité** : Protection de la Vie Privée (RGPD), Sureté, Confidentialité, Intégrité, Disponibilité, Validité,..
  - **Résilience**: Pérennité, Extensibilité, Transférabilité, Robustesse, Durabilité, Agilité, Anti-Fragilité,
- Spécifiques IA
  - **Transparence** : Explicabilité, Interprétabilité, Traçabilité, Imputabilité, Auditabilité, ...
  - **Equité** : Non discrimination (Biais)? Neutralité, Impartialité, Diversité, …
  - **Humanité** : Ethique, Libre arbitre, Utilisation responsable, Dignité, «Bouton rouge», …

Quelles solutions ? Quelles certifications ? Quel niveau de confiance (tolérance à l'erreur) ? Quelle soutenabilité ? Quel équilibre/priorités entre exigences ? Quels indicateurs clés ? Quelles responsabilités ?

Exigences : Transparence

Problématiques

- **Objectifs** : Quoi ? Quelles assertions ?
- **Opacité** : Comment ? Boîte noire / Apprentissage profond, Reproductibilité, …
- **Déséquilibres à l'accès**, Décisions individuelles : Recours, Preuves, Alertes, …
- **Responsabilités** : Diffuses, Développeur / Déployeur, «Provider / Controller», Utilisateurs, …
- **Contraintes** : Droits et besoins de savoir vs de confidentialité (données personnelles vs économiques / industrielles, RGPD vs secret des affaires) ? Circonstances ? Concurrence loyale ? Exactitude ? …

Réponses

- **Transparence « by design »** : De bout en bout (Conception, Développement, Exploitation, …)
- **Traçabilité / Piste d'audit**: Décisions, Opérations, Actions, Incidents, Réponses, ..; Boîte noire
- **Outils / Techniques** : En cours de développement, Limites, Gouvernance, Analyse d'impact, ...

Exigences : Equité / Non Discrimination

Problématiques

- **Types d'équité** : de groupe / individuelle, opportunités / résultats, Quotas, …
- **Nature des biais :** cognitifs, techniques, volontaires, « Nudging », discriminants, de restitution, erreurs, …
- **Sources diverses** : données entraînement, données exploitation, modèles, algorithmes, variables absents, interprétation, …
- **Accentuation des biais** : généralisation, opacité, …

Réponses

- **Outils / Techniques** : Développement, analyse données,… (sous-groupes, anti-classification, calibration, ...), analyse d'impact
- **Transparence** : biais volontaires, biais techniques (probabilistes, aléas,..), corrélation/causalité, méthodes déductives vs inductives, signaux faibles, « cygnes noirs », degré de maturité, notions d'exactitude statistique et marges d'erreur, tolérance, impacts des faux positifs/faux négatifs, contreparties, évolutions dans le temps, …

La gestion des risques

Camille Rosenthal-Sabroux,
professeur émérite, LAMSADE
Université Paris Dauphine PSL

Webinaire >> 21 janvier 2021

Applications à « haut risque »

Selon la Commission Européenne, une application d'IA devra être considérée comme étant à haut risque en fonction de ce qui est en jeu, en examinant si des risques importants sont associés à la fois au secteur et à l'utilisation envisagée, notamment du point de vue de la sécurité des personnes et des biens, des droits des consommateurs et des droits fondamentaux.

Même si une application n'est pas classée à « haut risque », elle reste entièrement soumise aux règles de l'UE en vigueur.

Secteurs à « haut risque »

- Emploi
- Éducation
- Soins de santé
- Transports
- Énergie
- Secteur public (droit d'asile, migrations, contrôles aux frontières, système judiciaire et services de Sécurité sociale)
- Défense et sécurité
- Finance, banque et assurance

Usages ou finalités à « haut risque »

- Recrutement
- Notation et évaluation des étudiants
- Affectation de fonds publics
- Octroi de prêts
- Commerce, courtage, fiscalité, etc.
- Traitements et procédures médicaux
- Processus électoraux et campagnes politiques
- Décisions du secteur public ayant une incidence importante et directe sur les droits et obligations des personnes physiques ou morales
- Conduite automatisée
- Gestion du trafic
- Systèmes militaires autonomes
- Production et distribution d'énergie
- Gestion des déchets
- Contrôle des émissions

Elaborer des modèles spécifiques de classification des risques

Plusieurs types d'IA, de techniques d'apprentissage, de raisonnements, d'outils et de dispositifs.

Plusieurs types de services et de fonctionnalités.

Plusieurs types d'acteurs concernés.

➔ IA des opportunités mais

DES NOUVEAUX RISQUES

Nouveaux risques

Perte de contrôle et du libre arbitre

Enfermement algorithmique

Opacité et non reproductibilité des résultats

Renforcement de certains biais

Absence de traçabilité

Imputabilité

Plus grande dépendance

Fragilité

Non protection des données personnelles, économiques, industrielles

Non responsabilisation des différentes parties prenantes….

Nouveaux principes

Minimisation des risques

Précaution

Vigilance

Proportionnalité

Non-malfaisance

Alerte…

Démarche de gestion des risques

Simple

Souple

Flexible

Facile

Alignée aux principaux référentiels internationaux

Favorisant une large adhésion

➔ **Vigilance**
➔ **L'environnement évolue, les technologies évoluent très rapidement, les classifications doivent évoluer aussi !**

Les responsabilités

Jean-Laurent Lienhardt,

expert-comptable mémorialiste

Webinaire >> 21 janvier 2021

Elaborer une démarche de gouvernance adaptée avec l'affectation des responsabilités associées

- **Problématiques**
  - Qui porte la responsabilité relative aux décisions/actions de l'IA en cas de problème ou de contestation ?
  - De quoi les différents responsables sont-ils responsables et jusqu'où (nature de la responsabilité) ?
  - Qui porte la charge de la preuve en cas de litiges et dans quelles mesures ?
- **Acteurs de la chaîne de valeur**
  - Concepteurs ;
  - Fournisseurs de technologie ou de services ;
  - Fournisseurs de Data (données brutes) ;
  - Propriétaire de l'algorithme ;
  - L'IA (en elle-même) ;
  - Data Scientiste et Opérateur de données ;
  - Déployeur/Opérateur de l'algorithme
  - L'Etat ;
  - L'utilisateur ;
  - Tout autre partie prenante en lien direct avec l'IA concernée.

Une chaîne de responsabilité complexe

- En règle générale, le processus de décision dépend grandement de la responsabilité qu'oblige cette décision. Pourtant plusieurs facteurs spécifiques à l'IA viennent complexifier ce modèle :
  - Aspect diffus
  - Non-traçabilité des actions
  - Non traçabilité des résultats aux sources
  - Manque d'éléments probants (jeu de données, biais, etc.)
  - Insécurités juridiques (typologie de responsabilité pénale, contractuelle, civile, etc.)

- En complément des éléments précités se pose la question de la détermination de l'obligation ou des obligations assortie(s) à l'IA : obligation de **moyen** ou de **résultat.**

- En parallèle, comment gérer la responsabilité relative aux comportements humains outrepassant le contrôle de l'IA, pour imposer leurs décisions sur les résultats prédictifs issus de l'IA, réputés plus « fiables » ? L'utilisateur pourrait-il voir sa responsabilité engagée au titre d'une perte de chance ?

L’assurance un enjeu majeur de l’IA

- L’encadrement de la responsabilité des acteurs est actuellement un des défis majeurs de l’IA, au même titre que l’assurance allouée à cette responsabilité.

- Ainsi pour l’heure une ébauche réglementaire prévoit l’encadrement reposant sur la typologie des dommages concernés (immatériel / matériel). Pour autant, il apparaît difficile d’encadrer les préjudices moraux en lien avec l’IA (Quid de l’accident d’une voiture autonome, par exemple).
  - Ainsi qu’en est-il du droit de recourt de l’utilisateur ? Comment l’appliquer, vers quel acteur ?
  - En parallèle, comment s’assurer de la protection des droits des utilisateurs face au caractère diffus des risques et de la chaîne de responsabilité ?

- Quid des lanceurs d’alerte ?
  - Comment les protéger face aux implications de l’IA dénoncés ?
  - Quels sont les moyens à leur disposition pour lancer l’alerte ? Qui est responsable ?

La réponse aux problématiques de

- Référentiel Global
  - Responsabilisation des concepteurs et des constructeurs (premiers concernés par les conséquences de l'utilisation de l'IA) – Charge de l'influence.
  - L'encadrement stricte de certaines pratiques telles que le profilage et des prises de décisions automatisées en lien avec le traitement des données personnelles (RGPD).
  - L'encadrement dans la constitution des bases de données nécessaires à l'apprentissage (Traitement des biais positif et/ou négatif) – Machine learning.

- Responsabilité Juridique
  - Adaptation sectorielle des régimes de responsabilité pour envisager les responsabilités sectorielles de la chaîne d'acteur.
  - Création d'un régime autonome de responsabilité notamment relatif :
    - Au concepteur afin de permettre l'engagement de sa responsabilité
    - À l'utilisateur de par l'usage qu'il prévoit de l'IA
  - Matrice d'imputation des responsabilités

Les bonnes pratiques

Claude Salzman,

consultant en Système d'Information, Président du Club Européen de la Gouvernance des Systèmes d'Information

Webinaire >> 21 janvier 2021

Pas d’audit sans bonnes pratiques

- Les bonnes pratiques sont ce que tout professionnel expérimenté connait et sait qu’il faut appliquer.
- Malheureusement, parfois, elles sont négligées.
- Ces oublis peuvent se traduire par des dérives significatives.
- Des référentiels largement admis recensent les bonnes pratiques : COBIT, PMBOOK, CMMI, Itil, ISO, …
- A ce jour il n’y a pas de référentiel de bonnes pratiques en matière d’Intelligence Artificielle.

Les 9 principes éthiques proposés par le règlement européen du 20/10/2020 sur l'IA

Le projet de règlement dans son annexe identifie 9 principes éthiques, bonnes pratiques à mettre en œuvre :

1. IA axée sur l'humain, développée et contrôlée par l'homme
2. Evaluation de la conformité des applications à haut risque
3. Sécurité, transparence et responsabilité
4. Garanties et solutions contre les biais et la discrimination
5. Droit de recours
6. Responsabilité sociale et égalité entre les genres
7. IA et robotique durables sur le plan environnemental
8. Respect de la vie privée et limitation de la reconnaissance faciale
9. Bonne gouvernance y compris des données utilisées et produites

Quelques bonnes pratiques identifiées dans le domaine de l'Intelligence Artificielle

1. Contrôle humain du système d'Intelligence Artificielle
2. Respect de la vie privée
3. Transparence
4. Egalité de traitement, non-discrimination
5. S'assurer de la robustesse technique et du niveau de sécurité suffisant du système d'Intelligence Artificielle
6. La gouvernance des données
7. Respecter le cycle de développement
8. Nécessité d'une conception
9. Rédaction d'un document de spécification
10. Existence d'une documentation suffisante
11. S'assurer que des tests suffisants ont été effectués
12. Disposer d'un environnement de tests avec des données de tests
13. Conformité de l'algorithme aux spécifications des professionnels
14. Evaluation de la satisfaction des utilisateurs

Quelques bonnes pratiques identifiées dans le domaine de l'Intelligence Artificielle

1. Contrôle humain du système d'Intelligence Artificielle
2. Respect de la vie privée
3. Transparence
4. Egalité de traitement, non-discrimination
5. S'assurer de la robustesse technique et du niveau de sécurité suffisant du système d'Intelligence Artificielle
6. La gouvernance des données
7. Respecter le cycle de développement
8. Nécessité d'une conception.
9. Rédaction d'un document de spécification
10. Existence d'une documentation suffisante.
11. S'assurer que des tests suffisants ont été effectués.
12. Disposer d'un environnement de tests avec des données de tests.
13. Conformité de l'algorithme aux spécifications des professionnels.
14. Evaluation de la satisfaction des utilisateurs

Quelques exemples de bonnes pratiques dans le domaine de l'Intelligence Artificielle

1. Y-a-t il eu une analyse des risques ?
2. Existe-t-il une note de cadrage de la future application ?
3. Est-ce que le système dispose d'outil de traçabilité permettant de comprendre les décisions prises ?
4. Existe-t-il un jeu d'essais de référence et peut-il être facilement rejoué ?
5. Est-ce que des biais ont été constatés et quelle est leur cause ?
6. Est-ce qu'en matière de données respecte-t-on le RGPD ?

Les bonnes pratiques sont très évolutives

- Attention : les bonnes pratiques de demain risquent de ne pas être forcement celles d'aujourd'hui.
- Car les technologies de l'Intelligence Artificielle évoluent très vite.
- Des solutions technologies sont en phase d'émergence notamment en ce qui concerne les tests.
- Définir les tests et les outils adaptés et suffisants associés par catégorie d'Intelligence Artificielle pour les phases d'apprentissage et opérationnelle :
  - Test de la performance
  - Fiabilité des algorithmes
  - Contrôle de l'apprentissage profond

L’audit et la certification

Patrick Stachtchenko,
chargé de cours en Gouvernance, Cyber Sécurité, Audit IA et Référentiels, CGEIT, CRISC et CISA

Webinaire >> 21 janvier 2021

Types de certification

Certifications légales : missions **d'intérêt général** (idem commissariat aux comptes)

- Missions pour les IA à **Haut Risque**
  - **Critères** : types d'activités, situations critiques, …
  - **Règles / Normes** communes (UE, International, …) : instances professionnelles
  - **Expertises** nécessaires : qualifications/certifications (techniques IA, gouvernance/management, audit, indépendance, …)
    - Commissaires aux IA

Certifications/Missions contractuelles : missions d'audit pour les parties prenantes concernées

- Missions au cas par cas
- **Labels** à définir
  - **Critères** : types d'activités, situations sensibles, niveau de confiance, …
  - **Règles/normes** : instances professionnelles

Certifications : Conditions nécessaires

- **Prescripteurs et destinataires** : qui demande ? (contexte, pourquoi, …), pour qui ? (attentes)
- **Périmètre**
  - Quels sont les objets et affirmations à certifier ?
  - Quelles sont les qualités/exigences à certifier ?
  - Quelles sont les acteurs devant faire l'objet d'études ? Quelle période ? …
- **Opinion** : certification, attestation, opinion circonstanciée, ….
  - Sur les résultats et/ou sur les moyens mis en œuvre (bonnes pratiques) ?
  - Sans réserve, avec réserve, refus, …?
- **Niveau de confiance** attendu pour les assertions et la certification (risques acceptables) ?
- **Démarches et Référentiels** : simples, flexibles, alignés, … (à base de modèles, …)
  - Bonnes pratiques en matière de résultats attendus
  - Bonnes pratiques de gouvernance et de management mises en œuvre
  - Bonnes pratiques en termes de diligences pour l'auditeur
- **Niveau de soutenabilité** technique et financière acceptable : propositions de valeur ? …
- **Responsabilités** : obligation de moyens / résultats, attentes, assurances, …

Certifications : les réponses

- **Illustrations** des types d'affirmations et d'opinions pour différents contextes (simple, boîte noire, …)
  - Quelles **propositions de valeur** possibles / acceptables ?
  - Responsabilités
- **Adaptation des démarches et des guides d'audit** : COBIT 2019, ITIL V4, ISO, …
  - Les différents **moyens/leviers** :
    - Principes, directives, référentiels, standards, ..
    - Structures organisationnelles (« IA Officer » ?)
    - Culture, éthique, comportements, incitations, aspects dissuasifs, ..
    - Informations
    - Aptitudes, compétences, savoir-faire, …
    - Processus, procédures, …
    - Outils et services : infrastructures, logiciels, …
  - **Cycle de vie** : planification, conception, développement, exploitation, …
  - **Indicateurs de résultats et de moyens** (on ne contrôle que ce qui se mesure !)

Certifications : la suite

- **Conditions pas** encore **réunies** pour fournir des **certifications « holistiques »**
  - Propositions de valeur non validées : assertions et faisabilité technique et financière
  - **Pas encore de référentiels** techniques, de gouvernance, de management et d'audit largement diffusés et acceptés
  - **Pas de structures institutionnelles/professionnelles** (normes, qualifications, …)
  - Nécessité d'effectuer des **missions pilotes**

- Néanmoins, possibilité de fournir des **opinions circonstanciées**
  - Certaines **assertions** (Vie privée, Sécurité, Transparence, …)
  - Certains **objets** (Algorithmes, Données, …)
  - Certains **leviers** (Structures organisationnelles, Outils, Technologies, Directives, Processus, …)

L’IA vue par la CNIL

Sophie Nerbonne,
directrice chargée de co-régulation économique
à la CNIL

Webinaire >> 21 janvier 2021

l'Academie
SCIENCES TECHNIQUES COMPTABLES FINANCIERES

LES CHIFFRES CLÉS 2019

CONSEILLER & RÉGLEMENTER

33 AUDITIONS PARLEMENTAIRES

362 AUTORISATIONS DE RECHERCHE SANTÉ

160 DÉLIBÉRATIONS DONT : 117 AVIS SUR DES PROJETS DE TEXTE

INFORMER

145 913 APPELS REÇUS

17 302 REQUÊTES REÇUES PAR VOIE ÉLECTRONIQUE

8 millions DE VISITES SUR CNIL.FR

115 700 *FOLLOWERS* SUR TWITTER

35 000 FANS SUR FACEBOOK

115 000 ABONNÉS SUR LINKEDIN

ACCOMPAGNER LA CONFORMITÉ

64 900 ORGANISMES ONT DÉSIGNÉ UN DÉLÉGUÉ À LA PROTECTION DES DONNÉES (DPO)

21 000 DPO DÉSIGNÉS

+ 31 % PAR RAPPORT À 2018

62 000 COMPTES CRÉÉS SUR LE MOOC* ATELIER RGPD**

2 287 NOTIFICATIONS DE VIOLATIONS DE DONNÉES

PROTÉGER

14 137 PLAINTES

+ 27 % PAR RAPPORT À 2018

4 517 DEMANDES DE DROIT D'ACCÈS INDIRECT

+ 6 % PAR RAPPORT À 2018

3 573 VÉRIFICATIONS EFFECTUÉES

CONTRÔLER & SANCTIONNER

300 CONTRÔLES ONT ÉTÉ EFFECTUÉS DONT :

53 CONTRÔLES EN LIGNE

45 CONTRÔLES SUR PIÈCES

42 MISES EN DEMEURE DONT :

2 PUBLIQUES

2 RAPPELS À L'ORDRE

2 AVERTISSEMENTS

8 SANCTIONS DONT :

7 AMENDES D'UN MONTANT TOTAL DE 51 370 000 EUROS

5 INJONCTIONS SOUS ASTREINTE

2 NON-LIEUX

RESSOURCES HUMAINES

BUDGET : 18,5 MILLIONS D'EUROS

215 emplois

39 ans Âge moyen

48 % DES POSTES OCCUPÉS PAR DES JURISTES

22 % PAR DES ASSISTANTS

19 % PAR DES INGÉNIEURS / AUDITEURS DES SYSTÈMES D'INFORMATION

80 % DES AGENTS OCCUPENT UN POSTE DE CATÉGORIE A

58 % DES AGENTS TRAVAILLANT À LA CNIL SONT ARRIVÉS ENTRE 2014 ET 2019

8 ANS ANCIENNETÉ MOYENNE DES AGENTS DE LA CNIL

Une double responsabilité de la CNIL au regard de l'IA

- **Au regard du RGPD et de la loi LIL2** : ceux qui conçoivent et mettent en œuvre des traitements d'IA doivent « garantir un traitement équitable et transparent », fournir des « informations utiles concernant la logique sous-jacente »,« d'une façon concise, transparente, compréhensible et aisément accessible ». Nombreuses prises de position de la Cnil au travers de sa doctrine notamment sur les sujets de :

  - biométrie (e.g. reconnaissance faciale, reconnaissance vocale),
  - traitement automatique de la parole (e.g. utilisation d'assistants vocaux),
  - justice prédictive (e.g. projet DataJust pour l'évaluation des indemnisations des préjudices corporels),
  - vision par ordinateur (e.g. détection de comportements suspects, détection du port du masque),
  - profilage (e.g. Parcoursup, évaluation des « compétences sociales » des candidats lors de recrutements)

Une double responsabilité de la CNIL au regard de l'IA

- **Au regard de la mission confiée par la loi pour une République Numérique de 2016 :** conduire une réflexion sur les enjeux éthiques et les questions de société soulevées par l'évolution des technologies numériques. Rapport (décembre 2017) « Comment permettre à l'homme de garder la main ?

Les enjeux éthiques des algorithmes et de l'intelligence artificielle » (deux principes fondamentaux : le principe de loyauté et celui de vigilance).

Articulation entre RGPD et les 3 piliers de l'approche européenne de l'IA

- Etre en avance sur les développements technologiques et encourager l'adoption de l'IA par les secteurs public et privé.
- Se préparer aux changements socio-économiques induits par l'IA.
- Assurer un cadre éthique et juridique approprié : **Bâtir la confiance dans l'IA centrée sur l'humain.**

➢ Suivi des travaux de la Commission Européenne (CE) : en 2018 groupe d'experts de haut-niveau sur l'IA.
(AI HLEG), 02.20 livre blanc sur l'Intelligence Artificielle,

➢ Trois résolutions du PE le 20.10.20 présageant les propositions CE début 2021 :
- un cadre éthique pour l'IA: 24 articles entre principes éthiques, obligations sur technologies à haut risque et contrôle par les instances européennes et nationales ;
- protection des droits de PI : 22 recommandations pour préserver la souveraineté numérique et industrielle de l'UE, assurer sa compétitivité et protéger l'innovation et les droits de PI ;
- régime spécial de responsabilité en cas de dommage causés par l'IA: 14 articles distinguant les dommages causés par les systèmes d'IA à haut risque et ceux causés par les autres systèmes d'IA.

Portée du principe d'analyse par les risques du responsable du traitement

- **Dans le RGPD** : Principe de responsabilité *(accountability)* et notion de *traitements à haut risque et traitement de données sensibles* : Le RT détermine les mesures techniques et organisationnelles appropriées à mettre en œuvre pour s'assurer et pouvoir démontrer que le traitement est effectué conformément au RGPD, avec la prise en compte de :
  - ✓ la nature, de la portée, du contexte et des finalités du traitement,
  - ✓ des risques, dont le degré de probabilité et de gravité varie, pour les droits et libertés des personnes physiques.

  Ces mesures sont réexaminées et actualisées si nécessaire.

- **Dans la résolution du PE** : notion d'« IA à haut risque » (secteur porteur de risques importants et utilisation comportant elle-même des risques)
  - ✓ Pour les IA à haut risque, un examen de conformité préalable sera à mener : inter-régulation ?
  - ✓ Quid d'une intégration de la méthodologie et de l'outil PIA dans l'évaluation de systèmes d'IA à hauts risques?

Accompagnement des acteurs au travers des outils de conformité prévus par le RGPD

- **Analyse d'impact sur les risques pour les personnes** (une même analyse peut porter sur un ensemble d'opérations de traitement similaires qui présentent des risques élevés similaires :
  - Socle commun : mutualisant certains éléments de l'AIPD (description des opérations de traitement envisagées, des finalités poursuivies, l'évaluation de la nécessité et de la proportionnalité des opérations de traitement mises en œuvre au regard des finalités poursuivies de l'AIPD)
  - Spécificités tenant à la situation propre : risques liés à la nature des données traitées ou des mesures de sécurité

  Avec l'appui méthodologique de la CNIL prévu dans sa réalisation

- **Code de conduite** (accompagnement méthodologique dans la rédaction)
- **Mécanisme de certification** (idem)
- **Application du « privacy by design »** pour le développement de systèmes d'IA avec divers angles :
  - Le principe de minimisation des données pour l'IA ;
  - La légitimité des finalités dans le cadre d'une démarche exploratoire ;
  - La question de la réutilisation ultérieure des données pour une finalité compatible et notamment le cas de la recherche ;
  - L'anonymisation des données : expertise reconnue de la CNIL en matière d'anonymisation de données (critères d'individualisation / de corrélation / d'inférence) résultant de l'avis du G29 et examen ponctuel de dispositifs en attendant l'homologation et la certification de méthodes d'anonymisation prévues par la loi pour une République Numérique ;
  - La génération de données synthétiques (dériver d'un jeu de données réel de caractéristiques sur les individus le composant afin de générer des individus fictifs présentant des caractéristiques réalistes par rapport au jeu de données initial).

La constitution de bases de données d'apprentissage pour l'IA en conformité avec le RGPD

- **Enjeu principal des entreprises qui font de l'IA** car le recours aux « bases de données d'apprentissage » est essentiel pour les systèmes d'IA reposant sur l'apprentissage automatique (machine learning), qu'il s'agisse d'apprentissage supervisé, non-supervisé ou par renforcement, afin d'entrainer, paramétrer, tester et valider les systèmes d'IA.

- **Nombreuses consultations de la CNIL sur la constitution de telles bases de données, exemples des services de caméras intelligentes ou de reconnaissance faciale :**

  - Sur les possibilités de collecte directe de données vidéo à des fins de constitutions de telles bases.
  - Possibilités de réutilisation de données vidéo collectées par des dispositifs de vidéosurveillance/vidéoprotection.
  - Mesures techniques et organisationnelles à mettre en œuvre pour protéger les personnes concernées (floutage, etc.)

La présence de données personnelles dans les modèles d'IA

- **Enjeu de sécurité juridique pour les RT :** premiers éléments publiés par l'ICO.

- **Les méthodes d'apprentissage automatique** (en particulier « supervisé ») se basent sur des données annotées pour dériver un modèle (description mathématique approximative du mécanisme qui a généré les données)

  - Caractère anonyme ou non des modèles statistiques appris ? (mémorisation de données à caractère personnel dans les modèles d'apprentissage automatique ?)

  - Implications juridiques : quelles conditions de présence de DCP dans les modèles et quelles modalités d'application RGPD.

  - Mesures techniques et organisationnelles à mettre en œuvre.

La transparence des algorithmes d'IA

- **Enjeu fondamental pour les droits des personnes concernées.**

- **Obligations de transparence résultant :**
  - du RGPD : profilage (Art.13-2-f), obligation d'information (Art. 15-1-h)
  - du CRPA (code des relations entre le public et l'administration) : pour les organismes publics (administration d'Etat, collectivité, organisme de droit public ou de droit privé intervenant dans le cadre d'une mission de service public) en cas d'usage d'algorithmes (d'IA ou non) pour prendre, ou aider à prendre, une décision administrative sur un administré (Cf. exemple Parcours sup infra)

- **L'explicabilité de l'algorithme doit permettre aux individus :**
  - un usage informé (quelle pédagogie des systèmes complexes face à la diversité des publics en contact avec ces algorithmes ?)
  - d'exercer leur droit à ne pas être sujet à une décision entièrement automatisée (art.22 RGPD)
  - modalités d'exercice des droits sur les traitements d'IA avec éventuels aménagements et limites liés à cette technologie illustrés par Parcours sup.

Focus sur Parcoursup (enregistrement des candidatures des bacheliers pour proposer une affectation après le classement opéré par les établissements)

❖ **Rappel sur la qualité de RT ?** qui définit les finalités et les moyens du traitement :

- Les établissements d'enseignement supérieur qui décident des modalités et des critères d'examen des candidatures, ou choisissent de recourir à l'outil d'aide à la décision de Parcoursup (pour le paramétrer en fonction de leurs besoins ou de l'utiliser tel que proposé par Parcoursup).
- Le ministère est RT de la plateforme nationale Parcoursup et ST pour le traitement des DP nécessaires au classement des candidatures pour le compte de ces établissements (nécessité d'un contrat ou convention cadre reprenant l'article 28 RGPD).

❖ **Transparence** : L'article L. 612-3 du Code de l'éducation aménage l'obligation pour les administrations de publier en ligne les règles définissant les principaux traitements algorithmiques utilisés dans l'accomplissement de leurs missions lorsqu'ils fondent des décisions individuelles (art L. 312-1-3 du Code des relations entre le public et l'administration - CRPA) : Une telle obligation de publication est réputée satisfaite si les candidats sont informés par les établissements qu'ils peuvent demander la communication des critères et modalités d'examen de leurs candidatures ainsi que des motifs pédagogiques qui justifient la décision prise.

- Décision du Conseil d'État (4e et 1ère chambres réunies, 12/06/2019, 427916) : ces dispositions spéciales réservent aux seuls candidats le droit d'accès aux documents relatifs aux traitements algorithmiques utilisés et pour les seules informations relatives aux critères et modalités d'examen de leur candidature.

- Décision du CC (n° 2020-834 QPC du 03/04/2020) : chaque établissement publie, à l'issue de la procédure nationale de préinscription et dans le respect de la vie privée des candidats, les critères en fonction desquels les candidatures ont été examinées (rapport) et peut donner des précisions sur les traitements algorithmiques utilisés pour procéder à cet examen.

Focus sur Parcoursup (suite)

❖ **Information :**

- L'information au titre du code de l'éducation sur la possibilité d'obtenir des informations relatives aux critères et modalités d'examen de leurs candidatures ainsi que des motifs, s'applique.

- L'information RGPD (art 15.1.h) sur les caractéristiques principales du traitement algorithmique (logique sous-jacente et principaux paramètres de l'algorithme, importance et conséquences prévues de cet algorithme) n'est obligatoire qu' en cas de décision entièrement automatisée. Pas le cas de Parcoursup (commissions d'examen des candidatures) mais même si cette information n'est pas exigée, elle constitue une bonne pratique recommandée et encouragée par la CNIL.

Le risque de discrimination et la question des biais

Enjeu majeur en termes de droits fondamentaux :

- La notion d'IA est désormais inséparable de celle de biais : les algorithmes d'IA embarquent le système de valeurs et de représentation du monde des concepteurs, qu'ils en soient conscients ou pas ;

- Détecter la présence de biais dans les décisions fournies par un système d'IA, voire travailler à introduire de la discrimination positive sont des enjeux essentiels pour s'assurer que les systèmes d'IA demeurent au service des utilisateurs et de la société ;

- Articuler RGPD avec le principe de loyauté (article 5) avec la question de la discrimination qui est de la responsabilité du Défenseur des droits (cf. séminaire « Algorithmes et discriminations » organisé par le DDD/CNIL en mai dernier).

La construction d'un cadre d'audit de solutions d'IA

- Quel pouvoir des régulateurs à auditer, constater et le cas échéant sanctionner des acteurs les mettant en œuvre ? Cf. travaux de l'ACPR (Gouvernance des algorithmes d'IA dans le secteur financier) ; l'ICO a publié le document Guidance on the AI auditing framework.

- Importance de la construction d'un cadre d'audit visant à s'assurer de la licéité et de la loyauté des divers systèmes mis en œuvre :
  - black box (pas d'hypothèse sur le code ou le modèle sous-jacent du système)
  - white box (code et modèle accessibles pour analyse)
  - modèle hybride (données d'apprentissage accessibles pour analyse)

- Le grand défi « Sécuriser, certifier et fiabiliser les systèmes fondés sur l'intelligence artificielle » pour la mise en œuvre concrète de telles méthodes de supervision des systèmes d'IA (rapports « France IA » et « Donner un sens à l'Intelligence Artificielle » indiquant de possibles responsabilités du Laboratoire National de Métrologie et d'évaluation (LNE).

WEBINAIRE - 21 janvier 2021

Vers un nouveau cadre de référence de l'Intelligence Artificielle

Intelligence et robotique : principe éthique, juridique, conformité et audit