



# **LA GOUVERNANCE DES ALGORITHMES D'IA ET LES RISQUES ASSOCIÉS**

**Laurent Dupont  
13 avril 2021**

--

**Webinaire EIFR « IA et compliance : quels modèles  
opérationnels ? »**

PÔLE FINTECH-INNOVATION



# AGENDA

1. Travaux de l'ACPR en IA
2. Principes de développement de l'IA en finance
3. Lignes directrices de gouvernance
4. Usages observés et réponse à la consultation
5. Méthodologie proposée pour l'audit
6. Prochaines étapes



# I. TRAVAUX DE L'ACPR EN IA

PÔLE FINTECH-INNOVATION

8/07/2020



# POURQUOI LE PÔLE FINTECH-INNOVATION S'INTÉRESSE-T-IL À L'IA ?

**Être un point  
d'entrée pour les  
fintechs**

**Rendre compte des  
tendances de  
l'innovation**

- 50% des nouveaux projets bancaires intègrent de l'IA (étude 2018)

**Contribuer aux  
travaux  
réglementaires**

- Paquet finance digitale
- Cadre général pour l'IA (2021)
- « Guidance » sectorielle

**Promouvoir  
l'innovation pour le  
contrôle (SupTech)**

- IA et Data Analytics comme facteurs-clés de l'amélioration et de l'innovation des méthodes de contrôle



# TROIS QUESTIONS SOULEVÉES PAR L'UTILISATION D'IA EN FINANCE

1. Comment évaluer un algorithme de ML (de façon démontrable) ?
2. Comment adapter la gouvernance à l'IA ?
3. Comment adapter les méthodes d'audit à l'IA ?

3 séries d'ateliers :

- **LCB-FT**: l'IA peut-elle améliorer les dispositifs ?
- **Modèles de risques**: comment et à quelle condition peuvent-ils intégrer de l'IA ?
- **Protection de la clientèle**: comment garantir que les algorithmes prennent correctement en compte l'intérêt du client ?



# ÉLÉMENTS DE RÉPONSE DU DOCUMENT DE RÉFLEXION DE JUIN 2020

## 1. Principes techniques de développement du ML

- Évaluation au cours du cycle de vie
- Nombreux arbitrages nécessaires
- L'explicabilité comme principe directeur

## 2. Principes de gouvernance

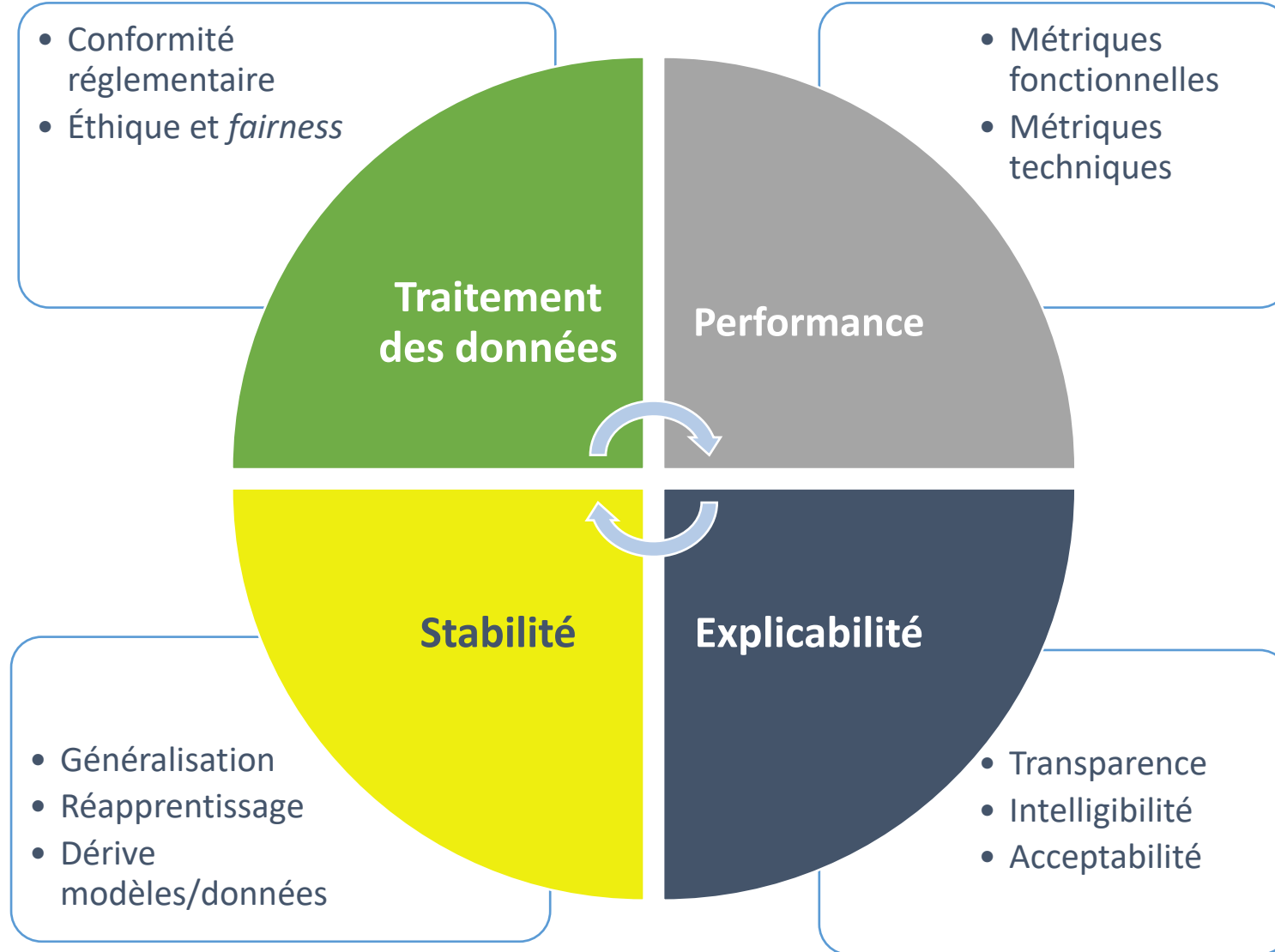
- Intégration de l'IA dans les processus
- Contrôle interne des processus incluant de l'IA
- Sécurité et risque de tiers

## 3. Approche méthodologique à l'audit

- Étudier les technologies et méthodes du marché
- Proposer une « boîte à outils du superviseur »

## II. PRINCIPES DE DÉVELOPPEMENT DE L'IA EN FINANCE

# PRINCIPES TECHNIQUES DE CONCEPTION DU ML







# EXPLICABILITÉ : *DÉFINITIONS ET ENJEUX*

## 1. Le « comment » i.e. la transparence

- ✓ Validation et surveillance internes
- ✓ Auditabilité par le superviseur

## 2. Le « pourquoi » i.e. l'interprétabilité

- ✓ Compréhension du comportement du système par les opérateurs humains qui interagissent avec lui
- ✓ Compréhension par le client ou utilisateur final
- ✓ Acceptabilité sociale ou éthique de la solution



# EXPLICABILITÉ : NIVEAUX D'EXPLICATION

## Explication de niveau 1 : observation

Elle répond sous un angle technique à la question : « *Que fait l'algorithme ?* », ou sous un angle plus fonctionnel : « *À quoi sert l'algorithme ?* ». Ce niveau d'explication peut être obtenu :

- de
- (it
- de
- de
- ur

## Explication de niveau 2 : justification

Elle répond à la question : « *Pourquoi l'algorithme donne-t-il tel résultat (en général ou dans une situation précise) ?* ». Ce niveau d'explication peut être obtenu :

- soit
- 4), é
- cont
- soit
- appr

## Explication de niveau 3 : approximation

Elle fournit une réponse, souvent inductive, à la question : « *Comment fonctionne l'algorithme ?* ». Ce niveau d'explication peut être obtenu, en sus des méthodes des niveaux 1 et 2 :

- pa
- M
- pa
- se
- de
- hy

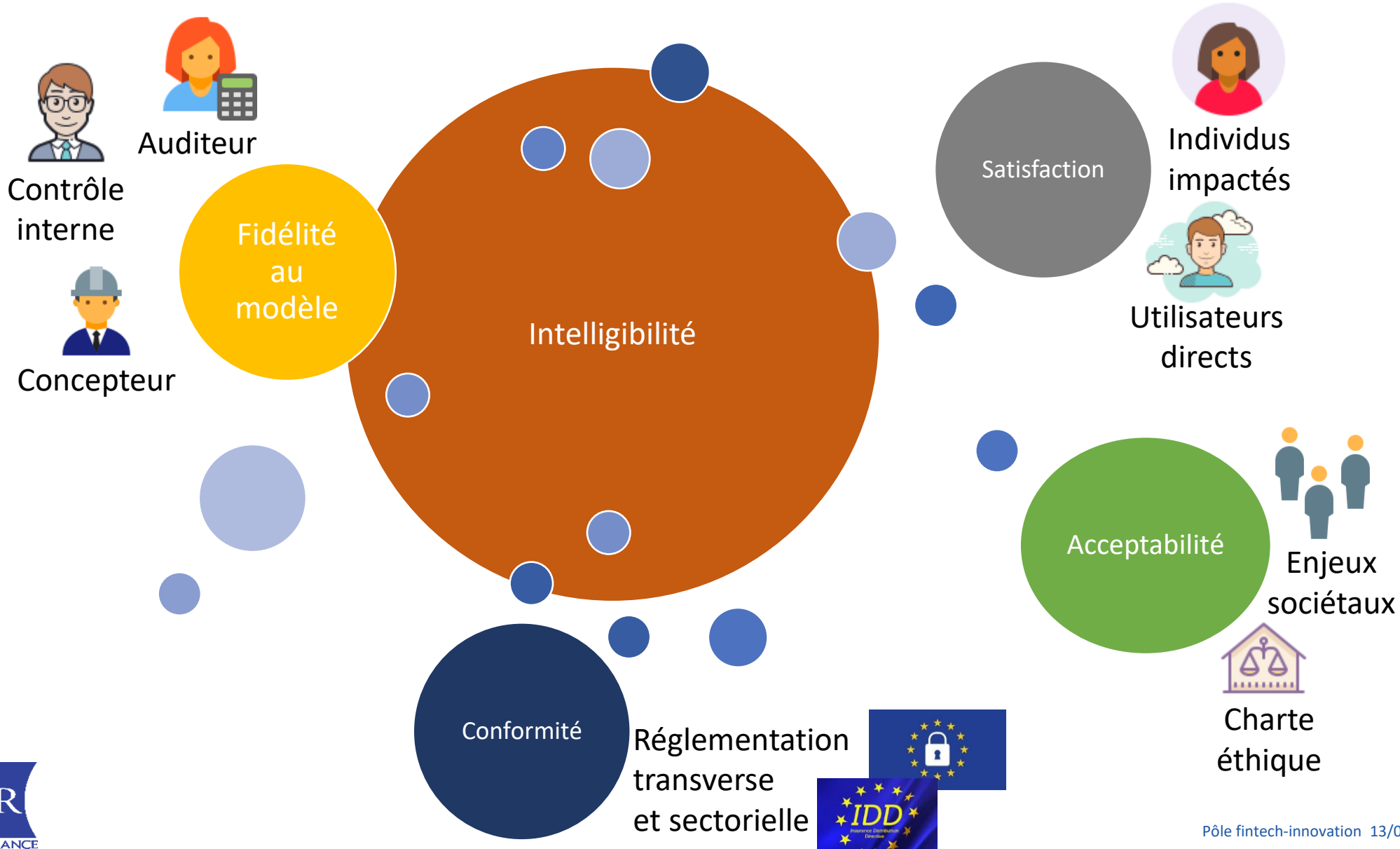
## Explication de niveau 4 : réplique

Elle fournit une réponse démontrable à la question : « *Comment prouver que l'algorithme fonctionne correctement ?* ».

Ce niveau d'explication peut être obtenu, en sus des méthodes des niveaux 1 à 3, par une analyse détaillée de l'algorithme, des modèles et des données. Dans la pratique, cela n'est possible que par une revue ligne à ligne du code source, une étude exhaustive des jeux de données utilisés, et un examen de l'ensemble des paramètres du modèle.



# AUDIENCE D'UNE EXPLICATION



# NIVEAUX D'EXPLICATION – EXEMPLE 1

Audience de l'explication	Contexte	Risque associé	Niveau d'explication requis
Client	Processus d'indemnisation	Risque opérationnel (insatisfaction du client)	1
Contrôleur interne	Vérification au quotidien du bon fonctionnement du processus	- Risque opérationnel - Risque de conformité (respect du contrat) - Risque financier	2
Auditeur	Évaluation de l'algorithme	- Risque opérationnel - Risque de conformité (respect du contrat) - Risque financier	3

## Domaine

Contrats d'assurance

## Processus métier

Gestion de contrat

## Fonctionnalité

Propositions d'indemnisation



## NIVEAUX D'EXPLICATION – EXEMPLE 2

Audience de l'explication	Contexte	Risque associé	Niveau d'explication requis
Équipe de validation	Validation des modèles, et de la politique de changement de modèle	- Risque de modèle (de solvabilité) - Risque de conformité	4
Organes d'administration, de gestion ou de surveillance	Approbation	- Risque de modèle (de solvabilité) - Risque de conformité	2

### Domaine

Modèles internes

### Processus métier

Conception du modèle

### Fonctionnalité

Calcul des ratios de solvabilité

## NIVEAUX D'EXPLICATION – EXEMPLE 3

Audience de l'explication	Contexte	Risque associé	Niveau d'explication requis
Agent en niveau 2	Analyse des alertes	Néant (si le comportement de l'analyste n'est pas modifié par l'existence de l'algorithme)	1
Contrôleur de l'algorithme	Contrôle permanent	<ul style="list-style-type: none"> <li>- Risque opérationnel (faux positifs et faux négatifs)</li> <li>- Risque de conformité (obligation de résultat)</li> </ul>	2
Auditeur	Contrôle périodique	<ul style="list-style-type: none"> <li>- Risque opérationnel (faux positifs et faux négatifs)</li> <li>- Risque de conformité (obligation de résultat)</li> </ul>	3

### Domaine

Sécurité financière

### Processus métier

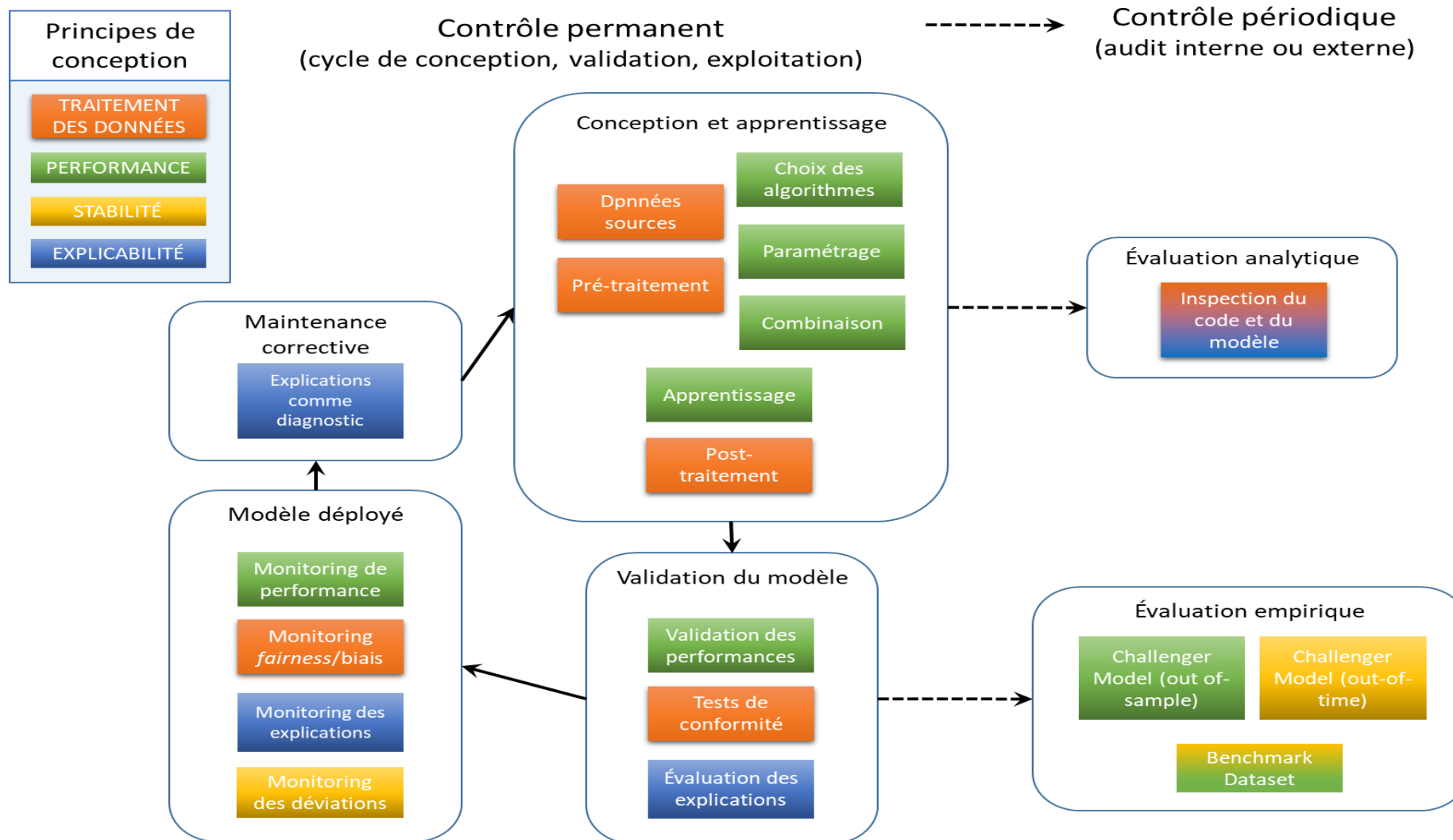
Gel des avoirs

### Fonctionnalité

Remontée d'alertes



# ÉVALUATION DU ML AU COURS DE SON CYCLE DE VIE





# III. LIGNES DIRECTRICES DE GOUVERNANCE DE L'IA EN FINANCE

PÔLE FINTECH-INNOVATION

8/07/2020

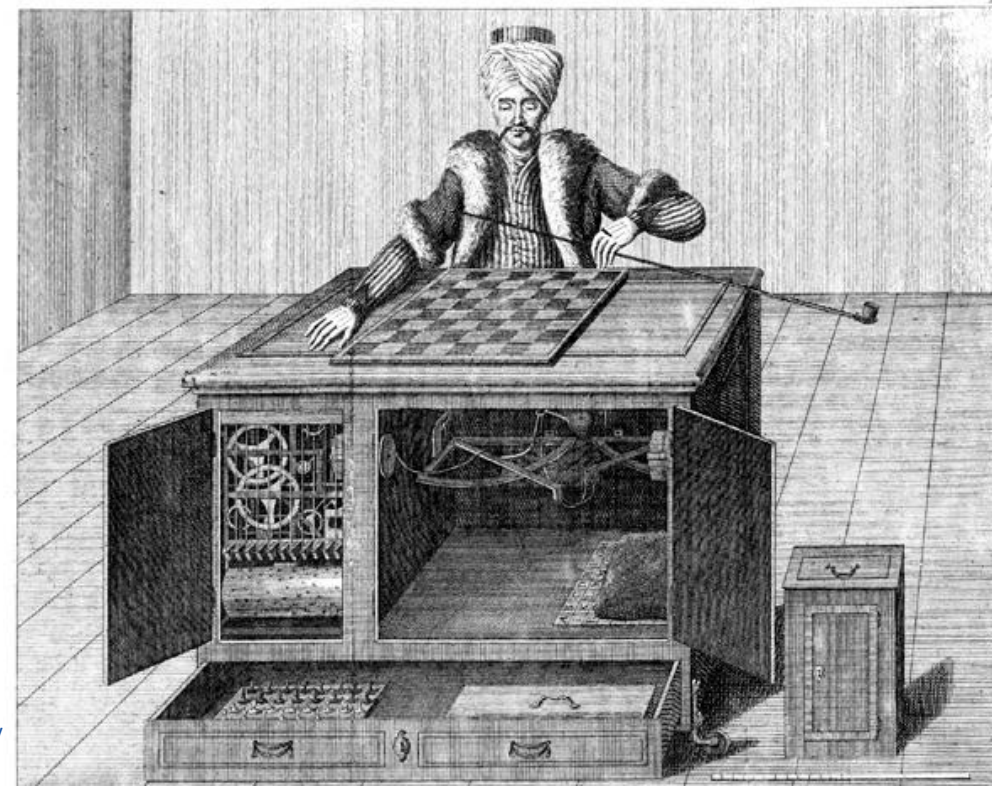




# GOVERNANCE ET CONTRÔLE DES ALGORITHMES D'IA

Quelques points-clés :

- Intégration de l'algorithme dans les processus
  - Quel est son objectif ?
  - Comment s'insère-t-il dans les processus existants ?
  - Interactions humain/algorithme ?
  - Quel choix d'industrialisation ?
- Processus de validation initiale et d'évaluation continue
- Sécurité et externalisation
- Audit





## IMPACT DE L'IA SUR LE CONTRÔLE INTERNE

- Impact organisationnel
- Processus de validation initiale, continue, technique
- Gestion des risques réglementaire, juridique, opérationnel, financier

**Exemple LCB-FT** : maintien de la capacité d'évaluation future

**Exemple Credit scoring**: validation aux niveaux

- Data Scientists (en local+ groupe)
- Équipe Validation Groupe
- DSI
- Comité Risques Groupe
- BCE

**Exemple en assurance** : traitement des déclarations de sinistres



# INTERACTION HOMME/MACHINE (1/3)

## 1. Processus traditionnel

Bénéfices

Risques

Explicabilité du processus

Latence prédictive / décisionnelle

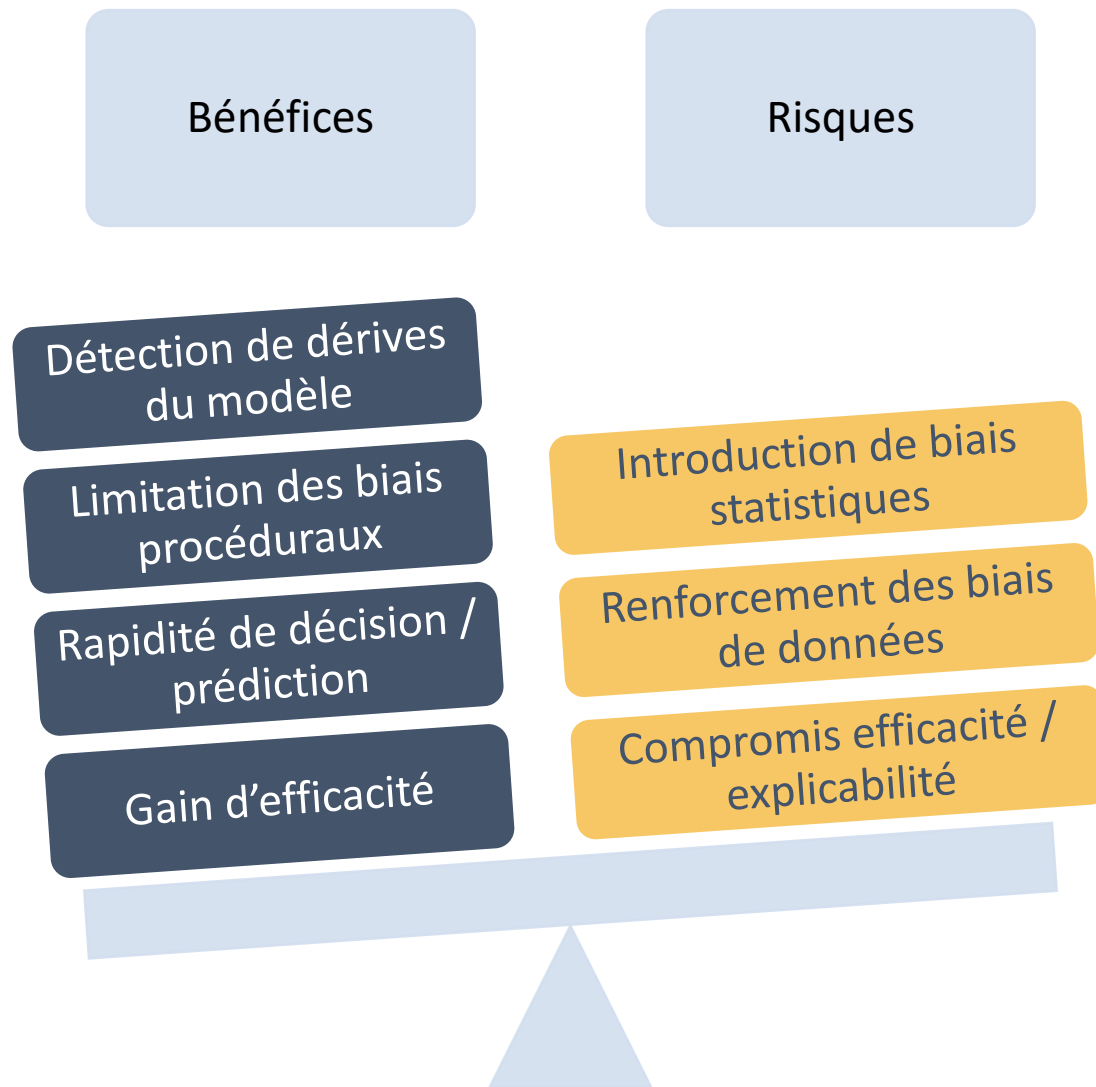
Introduction de biais de jugement





# INTERACTION HOMME/MACHINE (2/3)

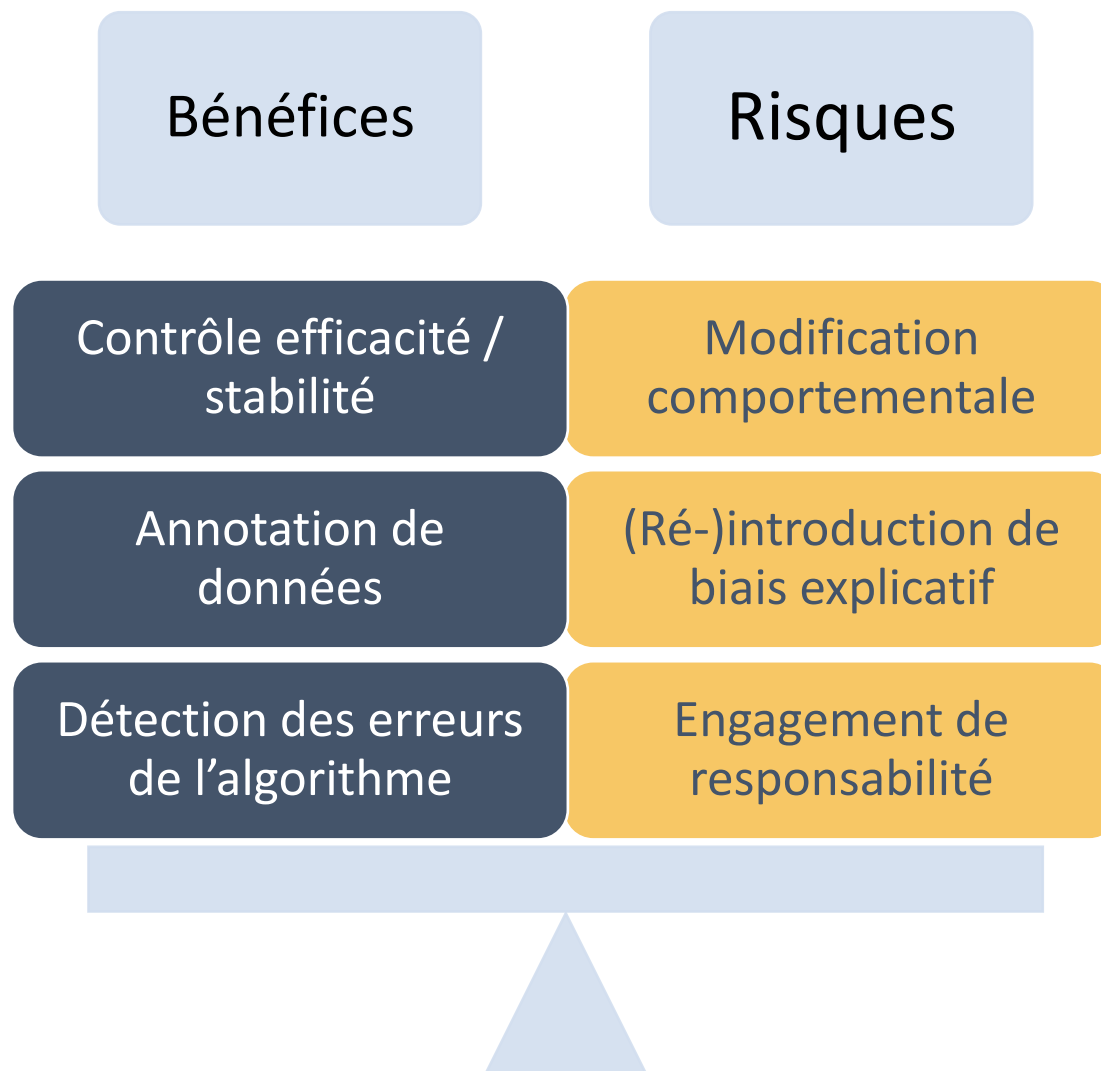
1. Processus traditionnel
2. Introduction de l'IA





## INTERACTION HOMME/MACHINE (3/3)

1. Processus traditionnel
2. Introduction de l'IA
3. Intervention humaine





## **IV. USAGES OBSERVÉS ET RÉPONSE À LA CONSULTATION**

PÔLE FINTECH-INNOVATION

8/07/2020



# USAGES OBSERVÉS DE L'IA : *CHAÎNE DE VALEUR EN ASSURANCE*

## Marketing & distribution

- Pré-remplissage de devis d'assurance habitation
- *Customer churn & retention*
- Souscription & tarification
- *Claims management*
  - Automatisation partielle du processus de traitement
  - Détection de fraude
  - *Subrogation referral*



# USAGES OBSERVÉS DE L'IA : *LES MODÈLES DE RISQUE DE CRÉDIT*

## En production

- Modèle avancé de risque de crédit des entreprises
- Modèle hybride de probabilité de défaut

## En expérimentation

- Calcul séquentiel de score d'octroi : règles métier → ML avancé → forçage humain

## Objectifs de l'IA

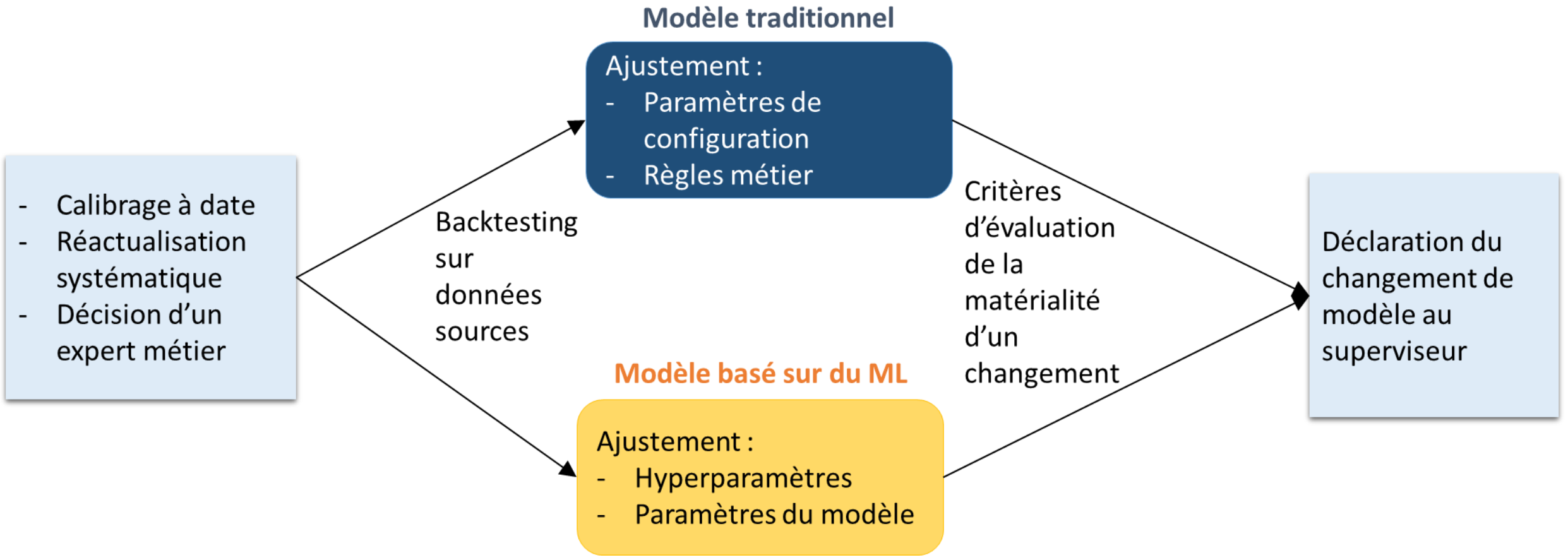
- Réduire la dépendance aux fournisseurs de données externes
- Mettre en qualité les données
- Améliorer le pouvoir prédictif des modèles (donc l'estimation du capital réglementaire)
- Accroître la transparence et l'auditabilité des modèles





# LE CAS DES MODÈLES INTERNES

## COMPARAISON ENTRE MODÈLES TRADITIONNELS ET ML





## USAGES OBSERVÉS DE L'IA : *PROCESSUS LCB-FT*

- **Processus de gel des avoirs**
  - Remontée directe en N2 (au lieu de N1) de certaines alertes
  - Annotation humaine en parallèle par N1
- **Détection de transactions suspectes**
  - Création de nouvelles alertes envoyées en N2
  - Routage N2 → N1 d'alertes issues des règles métier
  - Adaptation de modèles particuliers aux professionnels/associations
- **LCB + FT**
  - Profilage client
- **FT**
  - Identification de signaux faibles

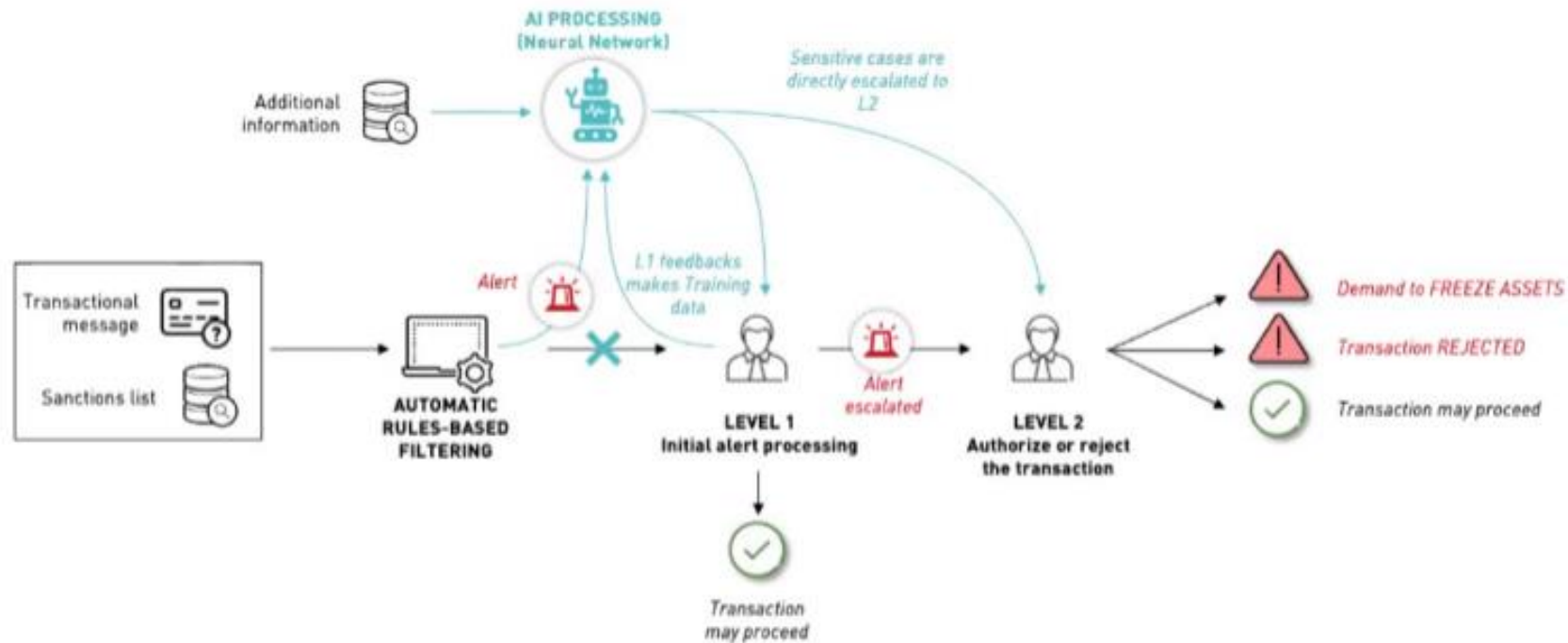
en production

en expérimentation

Pôle fintech-innovation 13/04/2021

# USAGES OBSERVÉS DE L'IA : *PROCESSUS LCB-FT*

- Processus de gel des avoirs
  - Remontée directe en N2 (au lieu de N1) de certaines alertes
  - Annotation humaine en parallèle par N1





# RÉPONSE À LA CONSULTATION : ALGORITHMES ET MÉTHODES





# RÉPONSE À LA CONSULTATION : CAS D'USAGE





# RÉPONSE À LA CONSULTATION : RISQUES INDUITS PAR L'IA

■ Opérationnel ■ Juridique/réputationnel ■ Financier ■ Éthique/sociétal

Opérationnel			Éthique/sociétal		Juridique/réputationnel
Défaut de généralisation	Manque de reproductibilité				Sécurité de l'IA
Risque IT (maintenance, intégration)	Réversibilité		Manque d'intelligibilité/explicabilité	Biais discriminatoire	Conformité réglementaire
			Traitement des données "unfair"		Responsabilité légale
Dérive de modèle	Rétroaction positive	Rétroacti... négative	Financier	Fraude non détectée	Détérioration de la réponse client

## V. MÉTHODOLOGIE PROPOSÉE POUR L'AUDIT D'IA EN FINANCE



# SUPERVISION DE L'IA EN FINANCE

## PROPOSITION DE « BOÎTE À OUTILS DU SUPERVISEUR »

Objet audité

Approches analytiques



Approches empiriques

Outil d'audit...	...des données	...de l'algorithme ou service d'IA	...du modèle
<i>Standards descriptifs</i>	<i>Datasheets, Data Statements, etc.</i>	<i>Factsheets</i>	<i>Model Cards</i>
<i>Analyse statique</i>	Analyse exploratoire	Code source	Quelques standards en <i>Deep Learning</i> uniquement
<i>Méthodes explicatives</i>	Pré-modélisation	Méthodes conjointes (TED, etc.)	Post-modélisation (locales/globales, spécifiques/agnostiques au modèle)
<i>Méthodes empiriques d'audit</i>	Jeux de données de <i>benchmarking</i>	Algorithmes <i>challengers</i>	Modèles <i>challengers</i>





## VI. PROCHAINES ÉTAPES

PÔLE FINTECH-INNOVATION

8/07/2020



# IA : LES PROCHAINES ÉTAPES

## Poursuivre la courbe d'apprentissage

- Travaux expérimentaux (ex : *fairness*)
- Ateliers pratiques (ex : *data sharing*)

## Accompagner l'écosystème

- Séminaires et GT (FR/UE/RoW)
- Tech Sprint

## Édicter des bonnes pratiques

- Recommandations européennes transsectorielles
- Recommandations sectorielles

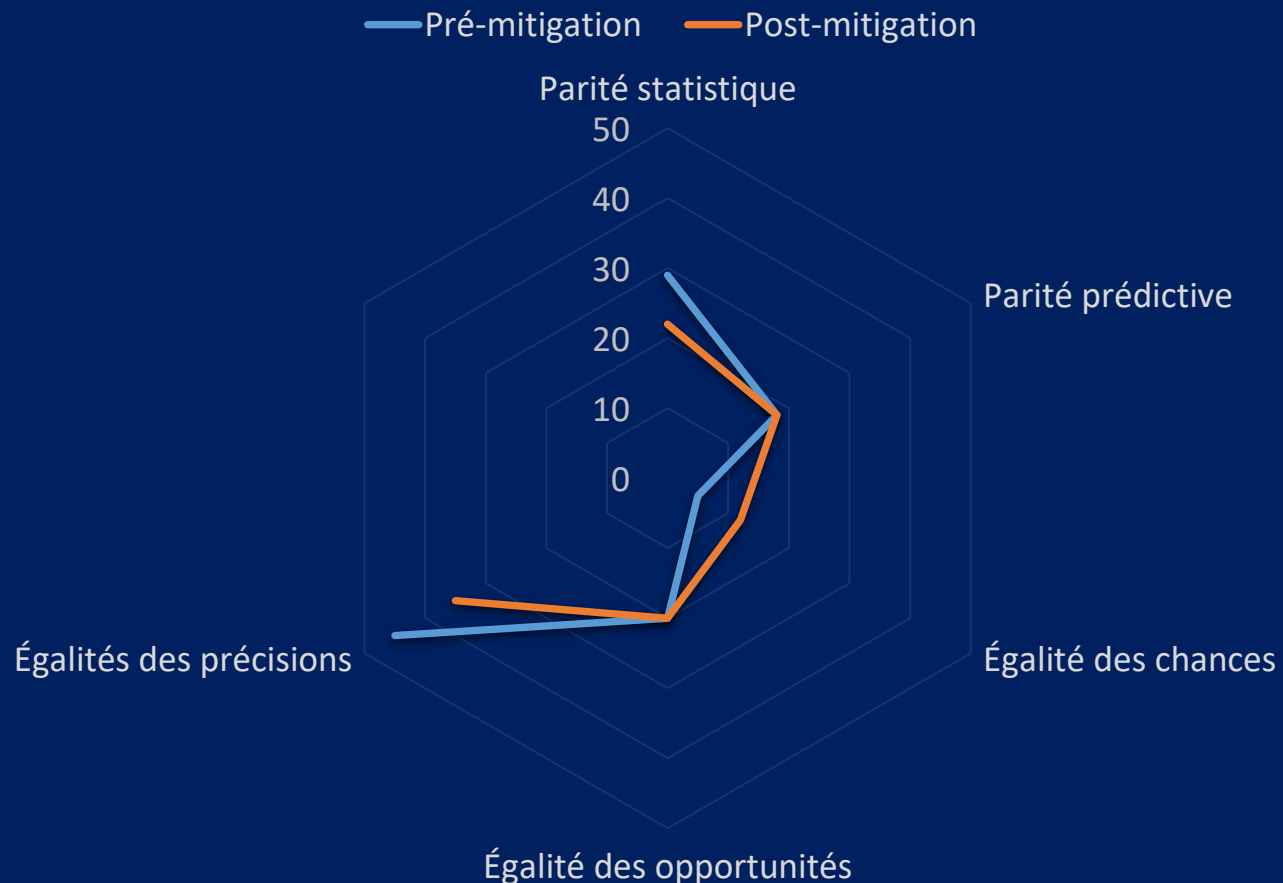
## Préparer le contrôle de demain

- Poursuivre le programme SupTech
- Construire la boîte à outils d'audit de l'IA



## TRAVAUX DE R&D EN *FAIRNESS*

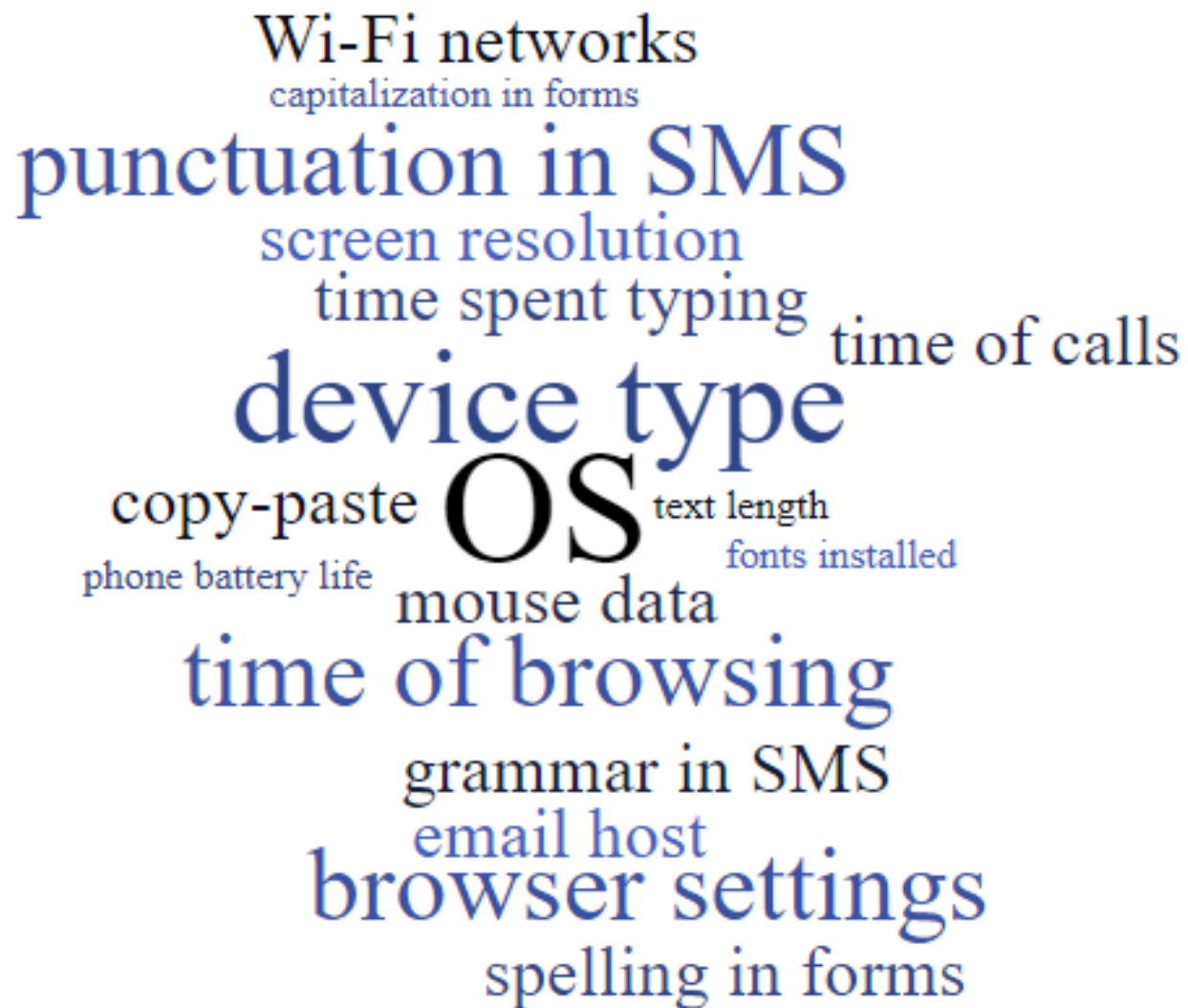
- Étude des métriques pertinentes en finance
  - *credit scoring*, détection de fraude, etc.
- Framework de diagnostic multidimensionnel, normalisé
- Méthodes de réduction des biais
  - en apprentissage
  - post-apprentissage





## TRAVAUX DE R&D EN *FAIRNESS*

- Impact du recours aux données alternatives
  - *Digital footprints* dans le domaine du *credit scoring*





# ATELIERS PRATIQUES EN *DATA SHARING & DATA POOLING*

## Enjeux réglementaires et techniques

- Protection des données personnelles ou sensibles
- Qualité et standardisation des données
- Clarification réglementaire
- Explicabilité et transparence
- Inclusion financière (*vs. de-risking*)
- Compétitivité entre acteurs
- Coûts et contraintes technologiques

## Technologies applicables

- Homomorphic encryption / functional encryption
- Zero knowledge proofs
- Secure multiparty computation (SMPC)
- Differential Privacy (centralized, local, hybrid)
- Machine Learning (supervised/unsupervised/reinforced)
- Federated Learning (horizontal, vertical)
- Deep Learning
- NLP
- Robotic process automation
- Network Analytics



# GOUVERNANCE DES ALGORITHMES D'IA DANS LE SECTEUR FINANCIER

*Merci pour votre attention.*

Pour toute question :

**[laurent.dupont@acpr.banque-france.fr](mailto:laurent.dupont@acpr.banque-france.fr)**

**[fintech-innovation@acpr.banque-france.fr](mailto:fintech-innovation@acpr.banque-france.fr)**