



INTELLIGENCE ARTIFICIELLE

COMMENT L'ACPR SE PRÉPARE-T-ELLE ?

JULIEN URI
PÔLE FINTECH-INNOVATION
5 JUIN 2024



L'IA, UNE OPPORTUNITÉ POUR « AUGMENTER » LA SUPERVISION

- Depuis 2018, l'ACPR a mis en place un programme « SupTech »
 - SupTech : usage des « nouvelles **technologies pour la supervision** »
 - Créer de nouveaux outils : augmenter nos capacités d'analyse
 - Comprendre les nouvelles technologies pour contrôler
 - Une tendance perceptible à l'échelle internationale : initiatives de la BRI (BIS Innovation Hub), du MSU, de nombreuses autorités nationales.
- De nombreux cas d'usage expérimentés et développés
 - Traduction, transcription, pré-analyse de rapports, etc.
 - Une démarche expérimentale avec plusieurs enjeux :
 - Assurer l'adhésion des utilisateurs par la co-construction
 - Assurer la compatibilité des outils avec les systèmes informatiques existants
 - Suivre les progrès et l'état de l'art des nouvelles technologies.



EXEMPLE : LE CONTRÔLE DE LA LCB-FT

➤ LUCIA, un outil pour le contrôle sur place :

- Algorithmes de *data mining* et de *machine learning* pour extraire des **signaux faibles** à partir des données d'opérations et de connaissance client collectées dans le cadre des investigations
- Restitue visuellement l'information sous la forme d'une **cartographie intelligente** des risques en groupant automatiquement les clients par profils types de risque
- Permet d'**investiguer les dossiers individuels** en mettant en exergue les faits saillants de risque des clients, et en donnant accès à un relevé d'opérations enrichi
- Facilite l'analyse de l'**environnement économique du client** en construisant des modèles de graphes relationnels.
- **Parti pris méthodologique** : exploiter uniquement les données remises par l'établissement ou des sources d'information accessibles (rapports TRACFIN, appels à vigilance, données ouvertes, presse...)



The screenshot shows the LUCIA application interface. At the top, there is a 'LUCIA' logo and a 'demo' label. Below this, there are several sections for configuration:

- Selectionnez les scénarios de filtrage à inclure :** A dropdown menu is set to 'Client: atypique ... x'.
- Selectionnez les scénarios de filtrage à exclure :** A dropdown menu is set to 'DS x'.
- Gestion des scénarios de filtrage
- Selectionnez l'axe de segmentation :** A dropdown menu is set to 'Niveau de risque LCB-FT'.
- Qualité des données
- Préférences

At the bottom, there is a logo for 'BANQUE DE FRANCE EUROPOSITÉMI v1.2'.

LUC!A

Opération(s) à risque identifiée(s)

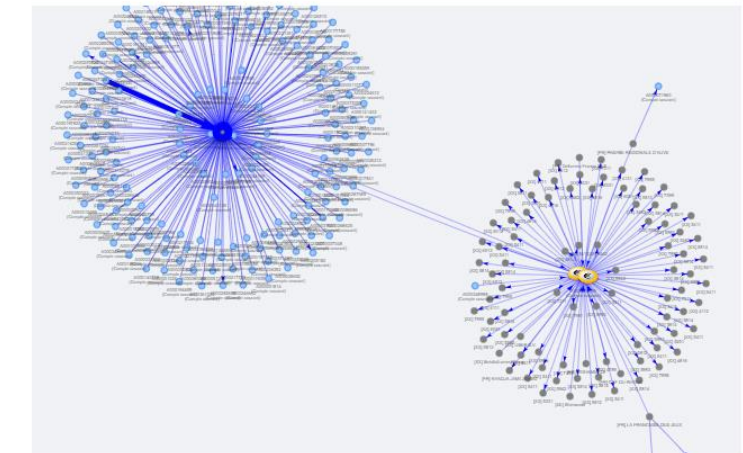
Date	Sens	Montant	Devise	Type d'opération	Contrepartie	Pays	IBAN
2021-09-21	D	400.00	EUR	SCT	KHADJA JAMI JAWED	France (FR)	FR8320000000000000000006210
2021-09-27	C	175.84	EUR	SCT	UBER B.V.	Pays-Bas (NL)	XX9999000000000000000002983
2021-09-27	C	176.41	EUR	SCT	Deliveroo France SAS	France (FR)	FR7611000000000000000002056
2021-09-30	C	78.71	EUR	SCT	UBER B.V.	Pays-Bas (NL)	XX9999000000000000000002983
2021-09-30	C	95.01	EUR	SCT	Deliveroo France SAS	France (FR)	FR7611000000000000000002056
2021-10-04	C	262.48	EUR	SCT	Deliveroo France SAS	France (FR)	FR7611000000000000000002056
2021-10-04	C	65.56	EUR	SCT	UBER B.V.	Pays-Bas (NL)	XX9999000000000000000002983
2021-10-06	C	159.83	EUR	SCT	UBER B.V.	Pays-Bas (NL)	XX9999000000000000000002983
2021-10-06	C	34.00	EUR	SCT	Deliveroo France SAS	France (FR)	FR7611000000000000000002056
2021-10-11	C	144.12	EUR	SCT	UBER B.V.	Pays-Bas (NL)	XX9999000000000000000002983

Exporter les 140 opération(s)

Analyse des transferts de fonds

Profondeur d'analyse maximale:

2



Exporter le graphe

L'ARRIVÉE DE L'IA GÉNÉRATIVE (1)

➤ Les enjeux

- Transformation de nos métiers : risques et opportunités ?
- Acculturation des équipes.
- Questions techniques : contraintes techniques ? Fiabilité ?

➤ Proposer une approche expérimentale pour éclairer ces enjeux

- Un « hackathon » associant des *data scientists* externes à des agents de l'ACPR (8 équipes constituées) pour proposer des cas d'usage ancrés dans la supervision
- Les résultats :
 - Démonstration concrète des différents usages possibles des *large language models* pour la supervision : recherche avancée dans un corpus, résumé, pré-contrôle de conformité, transformation de texte en code, constitution de bases de données, utilisation de données externes pour enrichir de l'information, vérification d'un LLM par un autre LLM
 - Possibilité de combiner avec d'autres briques logicielles
 - Fort engagement des collègues de l'ACPR dans l'expérimentation



L'ARRIVÉE DE L'IA GÉNÉRATIVE (2)



➤ Les prochaines étapes

- Expérimentations plus longues pour approfondir 4 cas d'usage
- Réflexion, dans le cadre du plan stratégique, sur les infrastructures et le déploiement technique de ces nouveaux outils

➤ Mais aussi des questions pour l'avenir

- Quelle est la part exacte de l'analyse, de l'effort, que nous voulons voir subsister chez les humains ?
- Comment les LLM vont-ils transformer notre relation à l'information et au reporting?
 - Aura-t-on recours de façon croissante à de l'information non structurée ?
 - En retour, et en tenant compte également des possibilités de « se passer » de code dans certaines situations, quelle transformation envisager pour les systèmes d'information utilisés ?



SE PRÉPARER À CONTRÔLER L'USAGE DE L'IA

- Pourquoi ?

- Contrôler la conformité au règlement européen IA mais aussi l'usage de l'IA au regard de nos propres objectifs sectoriels

- Les enjeux

- Se doter de méthodologies appropriées, des compétences adaptées
- Faire connaître nos attentes au secteur supervisé

- Effectuer les apprentissages nécessaires

- 2018 : 1^{er} document de réflexion
 - ✓ Cas d'usage dans le secteur, risques et opportunités, enjeux pour les autorités de contrôle
- 2020 : 2^{ème} document de réflexion
 - ✓ Leçons tirées d'ateliers pratiques avec des établissements volontaires sur des cas d'usage précis
 - ✓ Principes techniques, principes de gouvernance, perspective pour l'audit des algorithmes
- 2021 : Tech sprint sur l'explicabilité
 - ✓ État de l'art des techniques d'explicabilité des « boîtes noires »
- 2021-2024 : Travaux de recherche sur l'explicabilité
 - ✓ Pour le client, pour le contrôleur
- 2024 : Travaux en cours sur l'audit des algorithmes d'IA



QUELQUES MOTS SUR L'EXPLICABILITÉ DE L'IA

- L'explicabilité, une notion encore floue :
 - Le règlement IA
 - Mentionne cette notion sans la définir : « *"Transparence" renvoie au fait que les systèmes d'IA sont développés et utilisés de manière à permettre une traçabilité et une explicabilité appropriées...* »
 - Parle plutôt d'« interprétation » des résultats (« sorties ») du système d'IA
 - Les différentes disciplines ont des compréhensions différentes de la notion :
 - Data science : une approche technique (importance des variables, modèles simplifiés, explications « contrefactuelles ») très riche mais complexe et encore jeune ;
 - Explication « globale » ≠ explication « locale »
 - Sciences sociales et « interactions homme-machine » : prise en compte des objectifs de l'explication (justifier, contrôler, améliorer, apprendre), du processus d'explication humain, du contexte et des biais cognitifs
 - Droit : notion « intuitive » plutôt centrée sur le cas particulier (« justification », donner la possibilité de contester), peu de critères aujourd'hui pour définir ce qu'est un modèle ou un système « explicable »
- Les établissements devront mettre en œuvre différents niveaux d'explication pour différentes audiences :
 - Description, justification (par rapport à des normes « externes »), approximation, explication des causes...
 - Clients, utilisateur dans l'entreprise, contrôleur interne, organes de gouvernance, auditeur – interne ou externe...



L'EXPLICABILITÉ DE L'IA : LES DIFFICULTÉS DE MISE EN ŒUVRE (1)

1/ L'explicabilité pour les contrôleurs – étude de cas 2023 (contrôle de la LCB-FT)

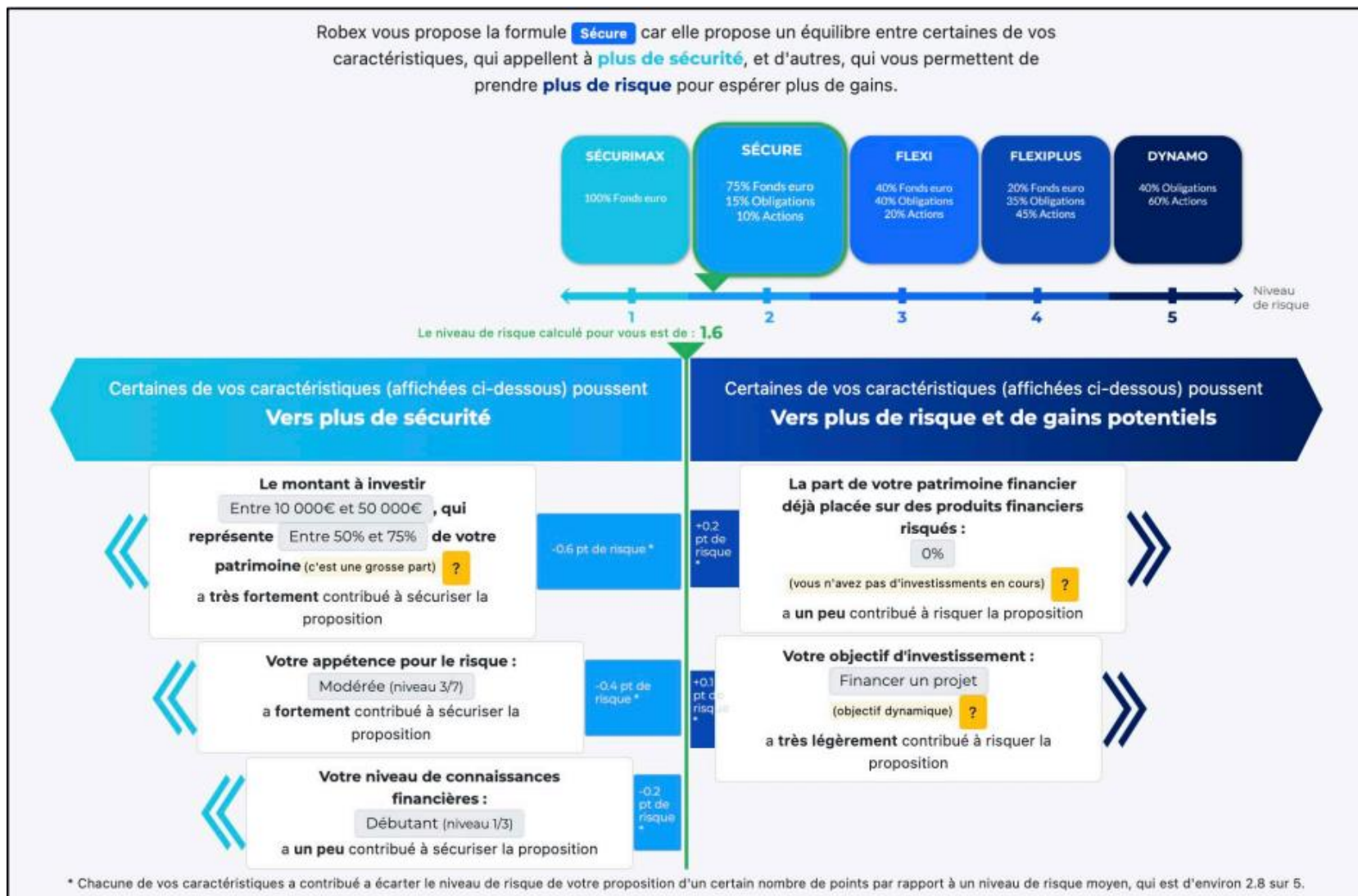
- Une dualité d'objectifs à concilier :
 - Comment le système fonctionne-t-il réellement ? (explication intrinsèque)
 - Le système est-il conforme aux exigences réglementaires ? (justification)
- Une simple explication/justification ne peut répondre à l'ensemble des objectifs d'un contrôleur
- Elle peut toutefois contribuer à ces objectifs (vérifier la prise en compte effective des profils de risque, détecter des erreurs méthodologiques, comprendre des alertes, s'assurer de la maîtrise du système d'IA, mesurer l'efficacité du dispositif)
- Pour cela :
 - Il faut concevoir des explications qui prennent en compte les objectifs et les biais cognitifs des utilisateurs (et des contrôleurs)
 - Il faut aussi « former » les utilisateurs pour 1) diminuer les biais cognitifs, 2) comprendre la complémentarité des objectifs (justification/explication)

L'EXPLICABILITÉ DE L'IA : LES DIFFICULTÉS DE MISE EN ŒUVRE (2)

2/ L'explicabilité pour les clients – test d'explications pour un « robo-advisor » en assurance-vie

- Une expérience réaliste : un *robo-advisor* (simplifié) effectuant toutes les étapes (réglementaires) du processus de souscription (devoir de conseil)
 - Tests de plusieurs types d'explication (graphes, graphes interactifs, explication « littéraire », pas d'explication)
 - 2 situations : une préconisation cohérente, une préconisation incohérente avec le profil du client (incorrecte)
 - Mesure de l'effet sur le client : accord, compréhension, confiance...
- Des conclusions inattendues
 - Les explications n'aident pas significativement à la compréhension du produit
 - Les propositions incorrectes sont moins bien comprises mais pas forcément rejetées, notamment par les clients ayant un niveau d'éducation plus faible
 - Les explications sous forme de conversation augmentent à tort la confiance des utilisateurs dans les propositions incorrectes
- Faudra-t-il modifier les exigences réglementaires pour assurer les objectifs réglementaires ?
 - S'inspirer de la recherche en interaction humain/machine : quid du profil du client (biais cognitif, niveau d'étude) ? Imposer des étapes qui améliorent la compréhension/la décision du client ? (comparaison entre produits ? test de compréhension du produit proposé...)

L'EXPLICABILITÉ DE L'IA – LES DIFFICULTÉS DE MISE EN ŒUVRE (3)



Astrid Bertrand, James R. Eagan, Winston Maxwell, *Questioning the ability of feature-based explanations to empower non-experts in robo-advised financial decision-making*, FAcT '23: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (June 2023).



CONCLUSION

- Les réflexions de l'ACPR sur la mise en œuvre pratique de la réglementation et du contrôle de l'IA dans le secteur financier font apparaître :
 - La nécessité de concilier les approches scientifiques (naissantes) et juridiques
 - Qu'est-ce qu'une bonne explication ? Qu'est-ce que l' « équité » ? etc.
 - Le rôle central de l'interaction entre les humains et les machines
 - Les algorithmes devront être conçus pour permettre le respect des objectifs réglementaires, en fonction des biais cognitifs humains
 - Mais les humains (les contrôleurs notamment) devront aussi s'adapter à ces nouvelles interactions
- À terme, ces « travaux pratiques » changeront les méthodes de contrôle et peut-être certains aspects de la réglementation