Artificial Intelligence, robo-advisors and supervision in finance

Astrid Bertrand

SCOR Fondation
20 juin 2024

Pôle Fintech-Innovation ACPR + Télécom Paris
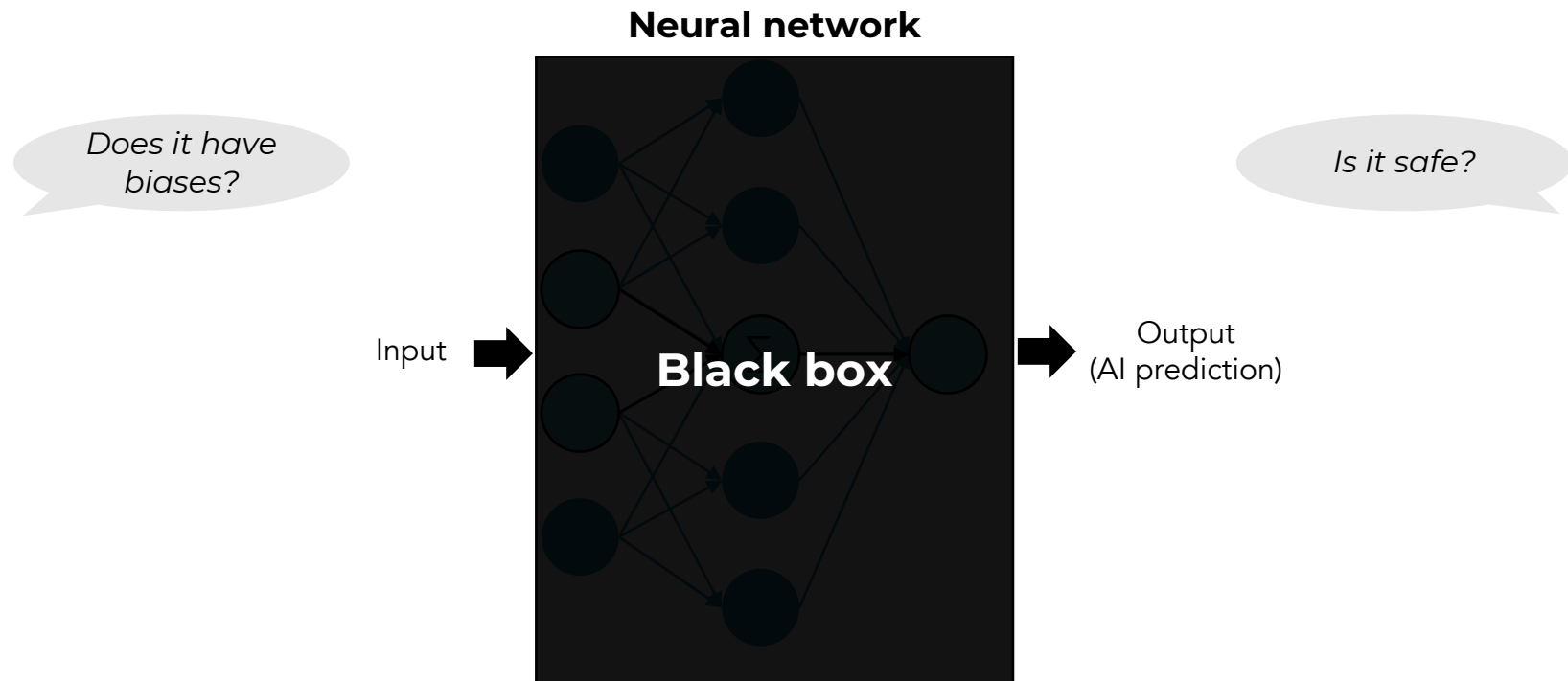
Agenda

Introduction to explainability

The role of explanations for life insurance robo-advisor customers

The role of explanations for supervisors in AML-CFT

AI explainability

Neural network

Does it have biases?

Is it safe?

Input

Black box

Output
(AI prediction)

Explainability

AI

Bank loan denied ❌

Explainable AI

Why was my loan denied?

"Your loan was denied because you have too many credits in progress"

The outcome would be different if you changed X….

Is this AI system trustworthy ? compliant ?

Consumer protection

..?

Protection from financial crime

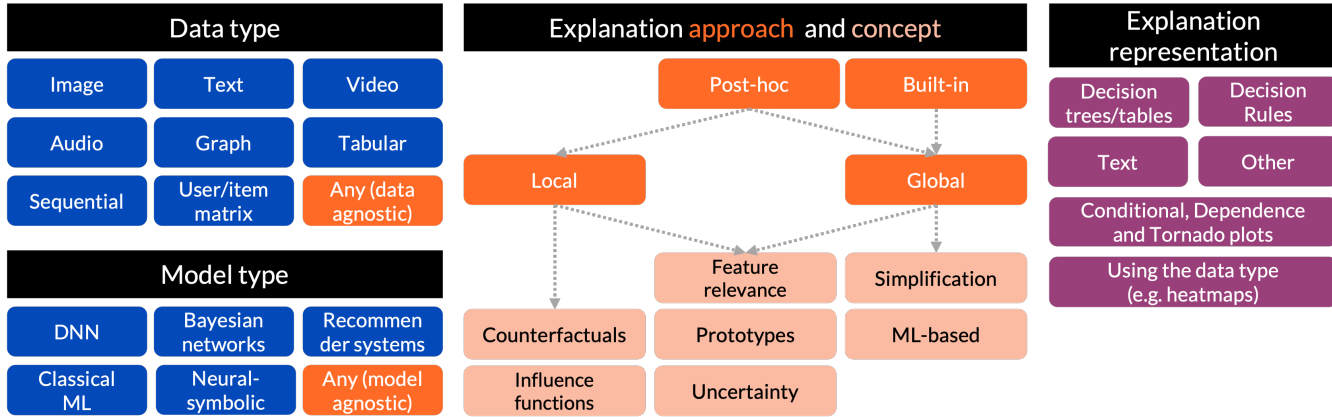Human biases and needs

What does XAI look like?

Data type

Image	Text	Video
Audio	Graph	Tabular
Sequential	User/item matrix	Any (data agnostic)

Model type

DNN	Bayesian networks	Recommender systems
Classical ML	Neural-symbolic	Any (model agnostic)

Explanation approach and concept

- Post-hoc
- Built-in
- Local
- Global
- Feature relevance
- Simplification
- Counterfactuals
- Prototypes
- ML-based
- Influence functions
- Uncertainty

Explanation representation

Decision trees/tables	Decision Rules
Text	Other
Conditional, Dependence and Tornado plots	
Using the data type (e.g. heatmaps)	

Figured inspired from « Anecdotal Evidence to Quantitative Evaluation Methods: A Systematic Review on Evaluating Explainable AI», Nauta et al. 2023
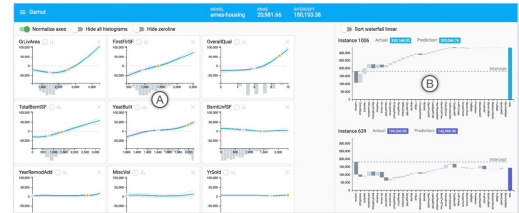
The role of explanations for customers

Context

The duty of **information and advice** in life insurance (L.522-5 CdA)

- **Formalize the reasons** for the appropriateness of the proposed contract **in relation to the** requirements and **needs expressed.**.

The underlying goals of the regulation:

- Allow the customer to understand

- Allow the customer to calibrate his confidence

- Strengthen the responsibility of insurance distributors

! **Currently nebulous / generic explanations**

! **Increasingly AI-based and real-time explanations**

Empowering customers of robo-advisors with explainability

- **Questions**
  What are the roles of explanations in financial investment services to protect customers?
  How effective are different representations of hybrid textual and graphical explanations to meet the aforementioned expectations?

- **Methodology**

Robex, the explainable and simplified robo-advisor

Profiling Questionnaire
- Objective
- Amount to be invested in % of financial assets
- Share of assets already invested in risky products
- Risk aversion
- Knowledge & Experience

Fintech-Innovation ACPR + Télécom Paris

Robex, the explainable and simplified robo-advisor
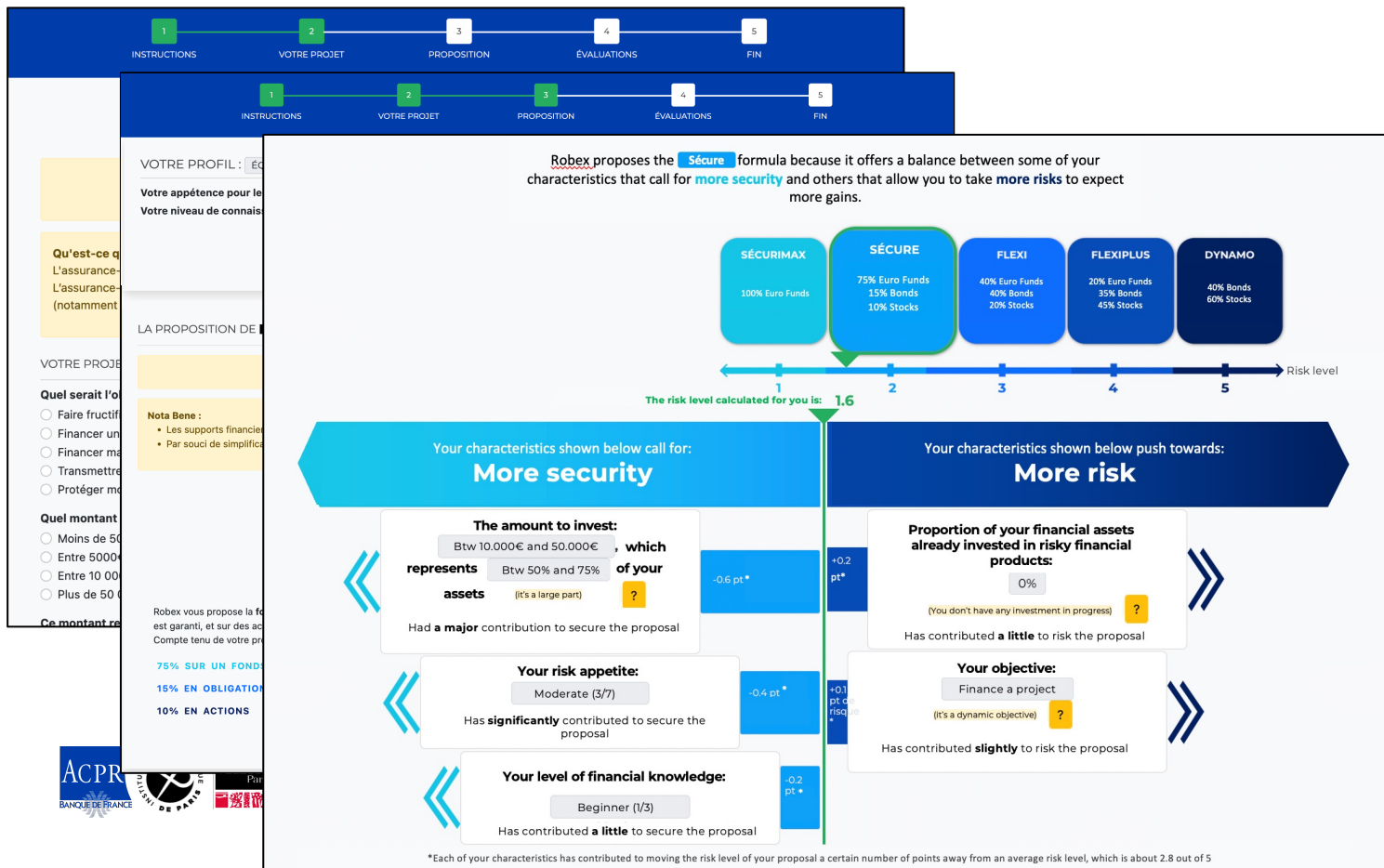

Recommendation
Rules inspired by discussions and documents provided by ACPR supervisors.
Each input variable is associated with a risk score (via coefficients associated with each question).

Robex, the explainable and simplified robo-advisor

Explanation
Explanation based on SHAP.
The weight that each of the 5 input variables had on the risk of the final proposal is evaluated.

An experiment to study user's "empowerment" through explanations

1 - Profiling questionnaire

2 - Participants either receive a...

Incorrect Recommendation

No explanation (Control)

Graphic Static explanation

Graphic Mutable explanation

Dialogic explanation

Correct Recommendation

No explanation (Control)

Graphic Static explanation

Graphic Mutable explanation

Dialogic explanation

3 - User makes a choice (accept or reject the recommendation)

4 – Post-study questionnaire

Explanation versions

Content
- Links between user characteristics and the proposed contract
- Important definitions
- Descriptions of the effects of complex user input parameters

Formats
- Graph-static (A)
- Graph-mutable (B)
- Conversational (C)

Explanations do not help much in the calibration of trust / reliance

Trust
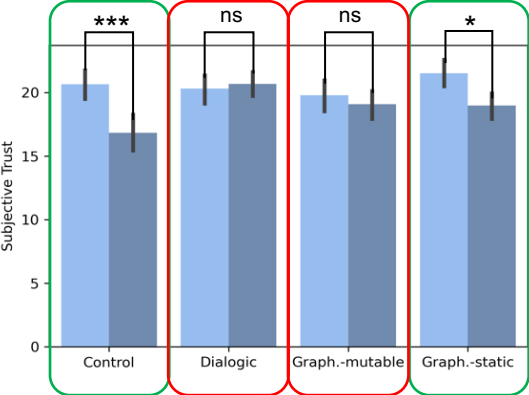Benevolence / Competence of McKnight's framework
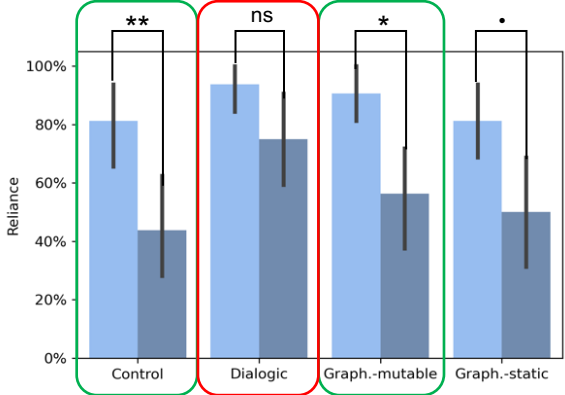
Reliance
Acceptance of the recommendation

Comprehension
of the recommended product as measured through domain-specific questions

Correct proposal
Incorrect proposal

• p ≤ 0,07   * p ≤ 0,05   ** p ≤ 0,01
*** p ≤ 0,001   ns non significant

Explanations do not help much in the understanding of the proposal

Trust

Benevolence / Competence of McKnight's framework

Reliance

Acceptance of the recommendation

Comprehension

of the recommended product as measured through domain-specific questions

Correct proposal
Incorrect proposal

- • p ≤ 0,07   * p ≤ 0,05   ** p ≤ 0,01
- *** p ≤ 0,001   ns non significant

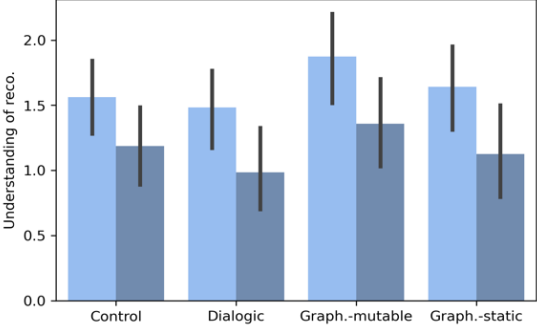# Dialogic explanations increase subjective trust in the proposal
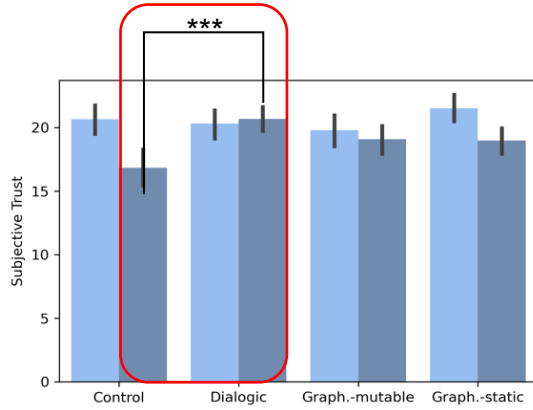
Trust
Benevolence / Competence of McKnight's framework
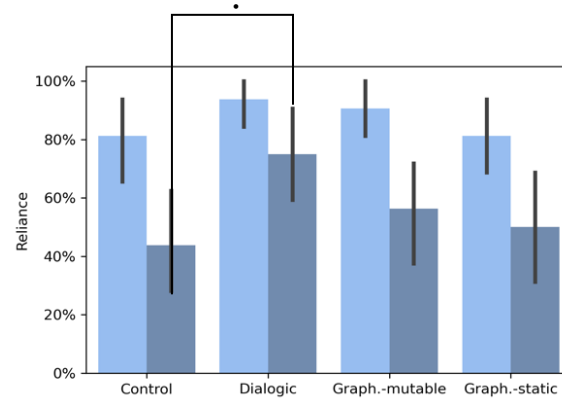
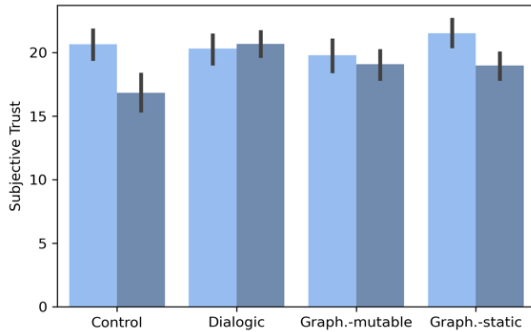**Reliance**
Acceptance of the recommendation

Comprehension
of the recommended product as measured through domain-specific questions

Correct proposal
Incorrect proposal

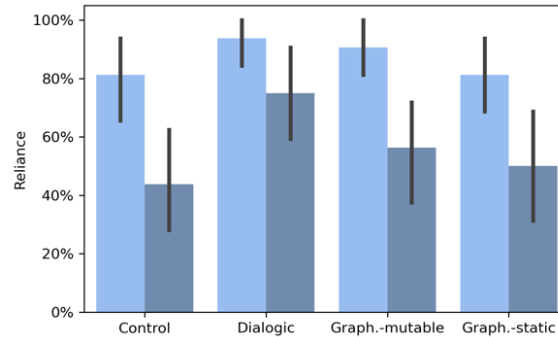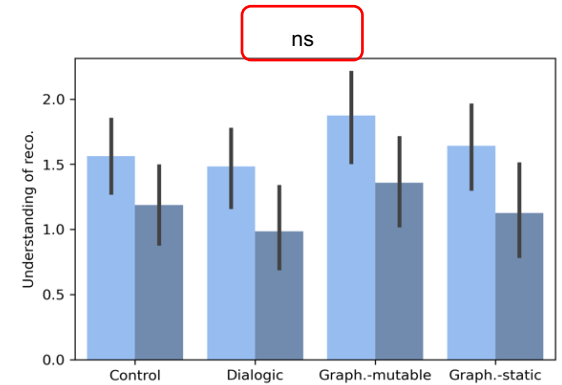• p ≤ 0,07   * p ≤ 0,05   ** p ≤ 0,01
*** p ≤ 0,001   ns non significant
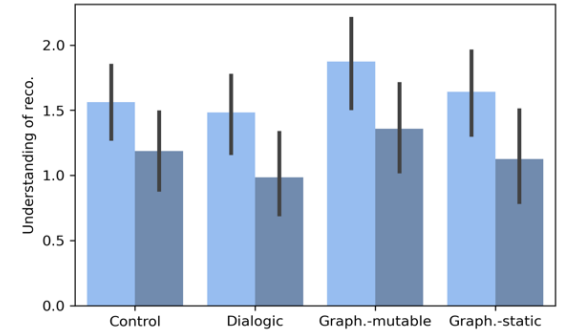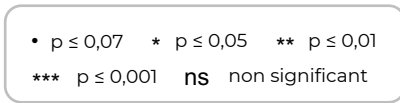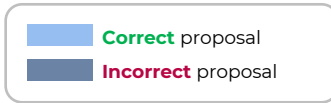
Takeways

1. **Conversational explanations (*i.e.* ChatGPT) can persuade and mislead users more than graphical explanations**

2. **"Legal" explanations are not always useful, nor even neutral: they can be misleading. Our findings debunk the idea from self-governance theories that explanations help the responsible consumer.**

3. **Explanations are still useful for the accountability of intermediaries.**

The role of explanations for supervisors

In Europe, €200 billion in criminal funds circulate every year
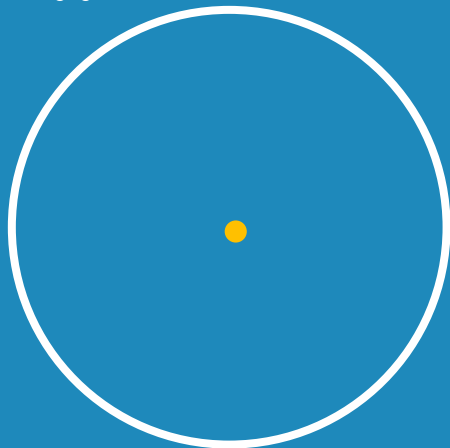
...Yet only **1%** of criminal funds are confiscated

[Europol, 2016]

Current AML-CFT systems are rule-based, extremely costly, and largely ineffective.

AI can help improve AML systems but creates **regulatory uncertainty.**

Explainability for AML-CFT

Financial data

"Black box" AI

AML-CFT alert

Why was the alert generated?

How can I debug?

Is the system compliant?

Case Handler

Data scientist
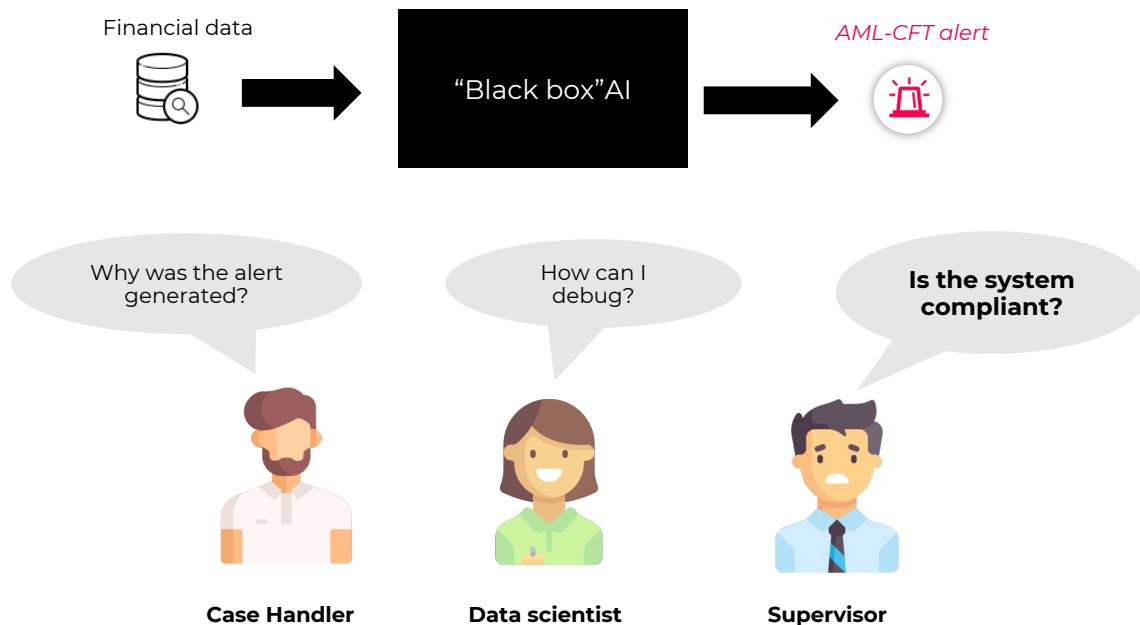
Supervisor

We explored the *user needs* and *legal requirements* of *supervisors* for AI justifiability* in AML-CFT.

* supervisors expect a ***justification,*** rather than an explanation, by regulatees that an AI system or decision complies with a legal standard, a rule, or an objective.

Understanding the supervisors' needs for explainable AI in financial crime detection

Improved ML-FT detection
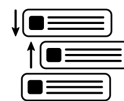
Alert management: priorisation and automatic closure

- **Questions**
  What are the **needs of supervisors** for justifiability* of AI systems?
  What are regulatory supervisors' current auditing practices and socio-technical context?
  How does AI opacity conflict with AML-CFT compliance requirements?

- **Methodology**
  We used a hybrid HCI and legal approach by
  - conducting **6 workshops** with 20 participants from the ACPR and from Crédit Agricole.
  - performing **a "compliance assessment"** of the AI case studies with AML-CFT laws

"We're going to end up with this like chickens with a knife...we won't be able to assess the adaptation to the risk". (P4)

Black-box AI

AML-CFT Laws
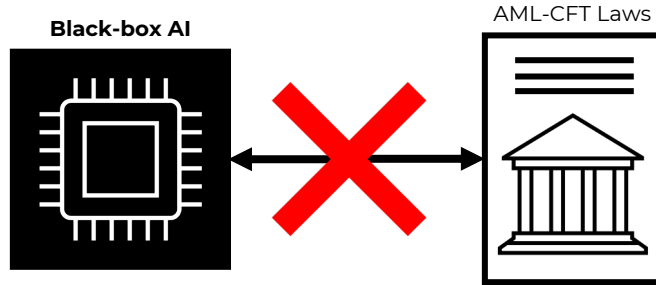
1. **Adapt the system to the bank's risk classification** *(are human and machine criteria aligned?)*

2. **Carefully examine transactions** *(is the system able to detect anomalies?)*

3. **Perform enhanced vigilance and SARs** *(are analysts able to understand alerts?)*

4. **Internal control requirements** *(is the bank able to detect incidents and have control over the purpose and operation of any device used?)*

5. **Allocate material and human resources** *(are human expertise and AI systems balanced? is collaboration between humans and AI smooth?)*

Supervisors' needs for model explainability and justifiability in AML-CFT

Supervisors' Questions on Banks' AI Models
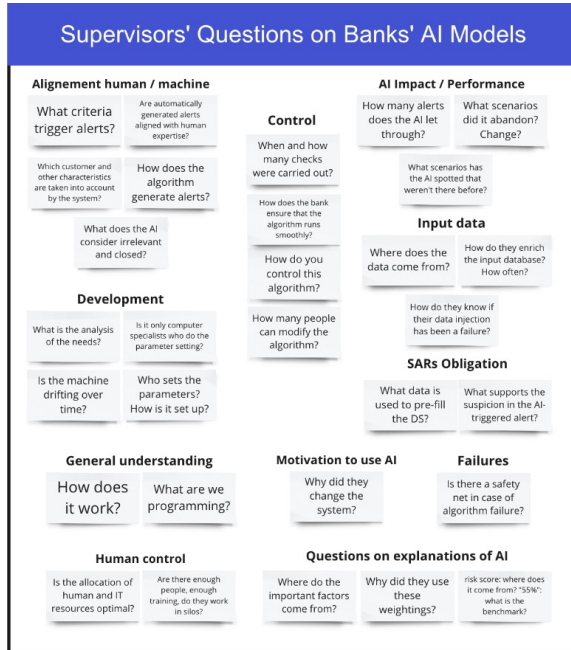
1. General comprehension

2. Demonstrate the legitimacy of SupTech AI tools

3. Measure global efficiency of the system

4. Establish the reprehensibility of sampled failure cases

5. Verify and challenge banks' understanding of AI:
   - Analysts are able to understand alerts
   - Decision criteria are aligned with human expertise
   - Model control is demonstrated

1. **AML-CFT illustrates the tension of using AI in a highly regulated environment**

2. **Current explainability techniques are incomplete and uncertain but can alleviate this tension**

3. **Explanations have a role of "trial evidence" for justifications of AI behavior**

Thank you for your attention !

Questions ?

Mesures

Variable	Mesure
Accord avec la recommandation	Choix: • La proposition de Robex paraît adaptée à ma situation • La proposition de Robex ne paraît pas adaptée à ma situation
Compréhension recommandation	1. À combien estimez-vous la part de fonds euro de la proposition qui vous a été faite ? 2. Sur une échelle de 1 à 5 (5 étant le plus risqué), à combien estimez-vous le risque de la proposition de Robex ? 3. Quelle est la particularité d'un fonds euro ?
Compréhension explications	1. Parmi vos caractéristiques et vos objectifs, quel facteur a eu le plus de poids dans la proposition qui vous a été faite ? 2. Comment la part de votre patrimoine financier déjà placée sur des produits financiers risqués, qui est pour vous : 0% de votre patrimoine financier, a-t-elle impacté la proposition faite par Robex ? 3. Comment votre objectif d'investissement, qui est Faire fructifier mon épargne a-t-il impacté la proposition faite par Robex ?
Confiance	1. Je pense que Robex agit dans mon meilleur intérêt 2. Robex veut comprendre mes besoins et mes préférences 3. Robex est compétent et efficace pour fournir des recommandations d'assurance-vie 4. Robex a l'expertise pour comprendre mes besoins et mes préférences 5. Robex remplit très bien son rôle de conseil en assurance-vie 6. 'aurais besoin d'un conseiller humain pour m'aider à choisir une formule d'assurance-vie
Engagement utilisateur	1. Je me suis senti(e) impliqué(e) dans ma tâche de choisir une formule d'assurance-vie 2. Le contenu du site de recommandation d'asurance-vie a attisé ma curiosité 3. J'ai été intéressé(e) par l'expérience
Charge cognitive	1. J'ai trouvé que c'était mentalement exigeant de lire et de comprendre la formule d'assurance-vie proposée et ses explications 2. J'ai dû faire des efforts pour lire et comprendre la formule d'assurance-vie proposée et ses explications