



HAL
open science

Towards Informed Decision-making: Triggering Curiosity in Explanations to Non-expert Users

Astrid Bertrand

► **To cite this version:**

Astrid Bertrand. Towards Informed Decision-making: Triggering Curiosity in Explanations to Non-expert Users. IHM '22 - Workshop on XAI, Apr 2022, Namur, Belgium. hal-03651368

HAL Id: hal-03651368

<https://hal.science/hal-03651368>

Submitted on 27 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Towards Informed Decision-making: Triggering Curiosity in Explanations to Non-expert Users

Vers une Prise de Décision Éclairée : Susciter la Curiosité dans les Explications aux Utilisateurs Non Experts

Astrid Bertrand

astrid.bertrand@telecom-paris.fr

Télécom Paris, Institut Polytechnique de Paris
Paris, France

ABSTRACT

As AI systems gain performance, their adoption expands to areas considered critical. In finance, increasingly sophisticated recommender systems known as "robo-advisors" are democratizing online access to life insurance. Their users are confronted with a choice that will affect their quality of life, even though they may not have any financial knowledge. To address this asymmetry of information between advisor and client, the French and European legislation [4] require insurance providers to produce "clear, precise and non-misleading" explanations to guide potential customers towards an "informed" decision. The research area of explainability thus has an important role to play in producing automatic explanations of robo-advisors that are relevant to the client. In this study conducted with the ACPR¹, we develop a life insurance recommendation interface. We build two explanatory interfaces: a "static" one with local and global explanations, and a "stimulating" one with explanations adapted to the user's knowledge and conveyed through questions: "do you know why X" [2]? We draw on the expertise of experts in customer protection in insurance as well as on the explainability literature on over-reliance issues [1, 3]. In the coming months, we plan to conduct a large-scale study on potential clients to investigate the impact of these different forms of explanations on "informed decision-making".

KEYWORDS

Explainable AI, Robo-advisor, User study, Cognitive bias

RÉSUMÉ

À mesure que l'IA gagne en performance, son adoption s'étend à des domaines considérés comme critiques. En finance, des systèmes de recommandation de plus en plus sophistiqués démocratisent l'accès en ligne à l'assurance-vie. Leurs utilisateurs de ces "robo-advisors" sont confrontés à un choix impactant leur qualité de vie,

¹The French Prudential Supervision and Resolution Authority

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

IHM '22, April 05–08, 2022, Namur, Belgique

© 2022 Association for Computing Machinery.

ACM ISBN XXX-X-XXXX-XXXX-X/XX/XX...\$15.00

<https://doi.org/10.1145/XXXXXXXXXX.XXXXXXXXXX>

alors qu'ils n'ont peut-être pas de connaissances financières. Pour remédier à cette asymétrie d'information entre client et conseiller, le droit français et européen [4] impose aux prestataires d'assurance de produire des explications "claires, précises et non trompeuses" pour guider les clients potentiels vers une décision "éclairée". Le champ de recherche de l'explicabilité a donc un rôle important à jouer pour produire des explications automatiques et pertinentes des robo-advisors. Dans cette étude menée avec l'ACPR², nous développons une application simplifiée de recommandation de produit d'assurance vie. Nous construisons deux interfaces explicatives : une « statique » pourvue d'explications locales et globales, une autre "stimulante" avec des explications adaptées aux connaissances de l'utilisateur et transmises par le biais de questions : « savez-vous pourquoi X ? » [2] ? Nous nous basons sur l'expertise d'experts en protection des consommateurs en assurance ainsi que sur la littérature en explicabilité sur les problèmes de surconfiance [1, 3]. Dans les prochains mois, nous prévoyons de mener une étude auprès d'une large base de clients potentiels, pour étudier l'impact de ces différentes formes d'explications sur la "prise de décision éclairée".

MOTS-CLÉS

IA explicable, Robo-advisor, Étude utilisateur, Bias cognitifs

ACM Reference Format:

Astrid Bertrand. 2022. Towards Informed Decision-making: Triggering Curiosity in Explanations to Non-expert Users. In *IHM '22 : 33^e conférence Francophone sur l'Interaction Homme-Machine, April 05–08, 2022, Namur, Belgique*. ACM, New York, NY, USA, 1 page. <https://doi.org/10.1145/XXXXXXXXXX.XXXXXXXXXX>

REFERENCES

- [1] Zana Buçinca, Maja Barbara Malaya, and Krzysztof Z. Gajos. 2021. To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making. *Proceedings of the ACM on Human-Computer Interaction 5*, CSCW1 (2021), 188:1–188:21. <https://doi.org/10.1145/3449287>
- [2] Tim Miller. 2019. Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence 267* (Feb. 2019), 1–38. <https://doi.org/10.1016/j.artint.2018.07.007>
- [3] Mahsan Nourani, Chiradeep Roy, Jeremy E Block, Donald R Honeycutt, Tahrira Rahman, Eric Ragan, and Vibhav Gogate. 2021. Anchoring Bias Affects Mental Model Formation and User Reliance in Explainable AI Systems. In *26th International Conference on Intelligent User Interfaces*. Association for Computing Machinery, New York, NY, USA, 340–350. <https://doi.org/10.1145/3397481.3450639>
- [4] The European Parliament and the European Council. 2016. Directive (EU) 2016/97 on insurance distribution.

²L'Autorité de Contrôle Prudenciel et de Résolution