

Proposition de sujet de P-SAT — Année 2025-2026

Analyse de la personnalisation et des dark patterns via la plateforme crowdsourcing STETOSCOPE

Louis KUSNO (louis.kusno@insa-lyon.fr)
Melisse COCHET (melisse.cochet@insa-lyon.fr)
Jixiang SUN (jixiang.sun@insa-lyon.fr)

Encadrant.e.s au département informatique

Antoine BOUTET (antoine.boutet@insa-lyon.fr)

Résumé

Contexte scientifique

La personnalisation des contenus Web et mobiles ainsi que les « dark patterns » sont désormais omniprésents dans l'écosystème numérique, influençant les choix des utilisateurs de manière opaque et parfois manipulatrice. Ces pratiques se manifestent sous différentes formes : prix variables selon le profil utilisateur, classements de recherche biaisés, ou encore compteurs de stock trompeurs exploitant les biais cognitifs humains.

L'analyse de ces phénomènes à grande échelle se heurte toutefois à des obstacles techniques majeurs. Les approches traditionnelles basées sur des robots d'exploration sont limitées par les systèmes anti-fraude des plateformes qui détectent et bloquent ces agents automatisés. STETOSCOPE propose une solution innovante à ce défi : une plateforme de crowdsourcing qui collecte des données authentiques (captures d'écran) auprès d'utilisateurs réels, permettant ainsi d'auditer la transparence des algorithmes sur le Web et le mobile sans déclencher les défenses anti-bots.

Objectifs

Ce projet P-SAT vise à auditer les pratiques de manipulation de plusieurs plateformes majeures du commerce en ligne (Amazon, Booking, AliExpress, Temu). Plus spécifiquement, nous cherchons à détecter et quantifier quatre types de manipulations : la discrimination par les prix (personnalisation tarifaire), le tri biaisé des résultats de recherche, la véracité des compteurs d'urgence (stocks limités, nombre de vues), et l'authenticité des rabais promotionnels.

Pour atteindre ces objectifs, le projet s'articule autour de deux axes techniques : d'une part, faire évoluer STETOSCOPE vers un outil véritablement multi-plateforme (Android, iOS, Web) pour élargir la base de participants ; d'autre part, automatiser l'extraction de données complexes à partir des captures d'écran grâce à des techniques avancées (LLM, OCR, vision par ordinateur).

Méthodologie

La démarche adoptée repose sur cinq piliers complémentaires :

- **Conception de protocoles d'audit** : Définition rigoureuse de scénarios d'étude reproductibles via l'interface d'administration de STETOSCOPE, garantissant la cohérence des données collectées.

- **Extraction intelligente par LLM** : Développement de scripts Python exploitant des modèles de vision (type GPT-4 Vision) pour extraire automatiquement des données structurées (JSON/CSV) à partir des captures d'écran.
- **Visualisation et analyse avancées** : Utilisation de Power BI pour transformer les données brutes en visualisations graphiques, pour identifier de potentiels discriminations et représenter les résultats.
- **Extension multi-plateforme** : Développement d'une interface Web complémentaire permettant l'intégration d'utilisateurs iOS et Desktop, jusqu'ici exclus de l'application Android native.
- **Campagne participative** : Organisation et coordination d'une campagne de collecte auprès d'étudiants insaliens volontaires.

Mots-clés

Dark Patterns, Crowdsourcing, Personnalisation des prix, Discrimination de recherche, Manipulation, LLM, Power BI.

Contexte de travail

Ce projet s'inscrit dans le cadre du cours P-SAT (Projet Scientifique, Artistique et Technique) sous la direction d'Antoine Boutet, enseignant-chercheur au département Informatique de l'INSA Lyon et dans le groupe Inria Privatics. L'équipe projet est composée de trois étudiants de 5^e année du département informatique.

Collaboration interdisciplinaire : Jusqu'à début décembre, le projet a été réalisé en collaboration avec 2 étudiants étrangers du département Télécommunications : Isabelle Ott Kiraly et Andy Vu Ngoc. Ils ont contribué à la rédaction des protocoles d'étude sur STETOSCOPE, à la communication de la campagne de collecte de données, ainsi qu'à l'analyse bibliographique.

Support technique et ressources : Tao Beaufiles (tao.beaufiles@inria.fr) s'est occupé du développement de Stetoscope pendant son Projet de Fin d'Études (PFE). Des échanges avec lui nous ont permis de mieux comprendre l'architecture logicielle existante et de planifier les évolutions futures. Pour les phases de développement et de test, deux smartphones Android nous ont été mis à disposition. Le suivi du projet est assuré par des réunions hebdomadaires avec Antoine, garantissant un accompagnement continu.

1 Introduction

1.1 Contexte

L'économie numérique contemporaine s'appuie massivement sur le profilage des utilisateurs et la personnalisation algorithmique des contenus. En théorie, cette personnalisation vise à améliorer l'expérience utilisateur en proposant des produits, services ou informations adaptés aux préférences individuelles. Dans la pratique, cependant, ces systèmes de recommandation fonctionnent comme des « boîtes noires » algorithmiques : ni les utilisateurs ni les régulateurs ne savent précisément quelles données personnelles alimentent ces profils, ni selon quels critères les décisions d'affichage sont prises.

Ce manque de transparence crée un terrain propice aux pratiques abusives. On observe ainsi des phénomènes de discrimination par les prix (un même produit affiché à des tarifs différents selon le profil utilisateur) ou de discrimination par la recherche (des résultats ordonnés différemment pour orienter les choix d'achat). S'ajoutent à cela les « dark patterns », ces mécanismes d'interface délibérément conçus pour exploiter les biais cognitifs humains : faux compteurs de stock créant un sentiment d'urgence, promotions fallacieuses, ou encore boutons trompeurs compliquant la désinscription.

1.2 Problématique

Auditer ces pratiques à grande échelle représente un défi technique et méthodologique considérable. Les approches traditionnelles d'audit algorithmique reposent sur des robots d'exploration qui simulent le

comportement d'utilisateurs en parcourant systématiquement les plateformes Web. Cette méthode présente toutefois trois limitations fondamentales :

1. **Absence d'historique authentique** : Les bots ne possèdent pas l'historique de navigation, les cookies ou le profil comportemental d'un utilisateur réel. Or, c'est précisément sur ces données que repose la personnalisation. Un bot voit donc une version « neutre » ou « par défaut » de la plateforme, pas la version personnalisée que verrait un utilisateur lambda.
2. **Détection et blocage** : Les plateformes commerciales ont développé des systèmes anti-fraude sophistiqués (CAPTCHA, fingerprinting, analyse comportementale) qui détectent et bloquent rapidement le trafic automatisé, rendant les campagnes d'audit à grande échelle techniquement impossibles sans contournement constant.
3. **Difficultés sur mobile** : L'automatisation sur smartphones (via des frameworks d'instrumentation comme Frida ou Appium) nécessite souvent un accès root ou des modifications système détectables par les applications. De plus, l'écosystème mobile (iOS notamment) est particulièrement fermé, rendant l'automatisation encore plus complexe que sur le Web.

Ces obstacles expliquent pourquoi, malgré l'importance sociétale des dark patterns, les études empiriques à grande échelle restent rares, particulièrement sur les plateformes mobiles qui représentent pourtant la majorité du trafic e-commerce.

1.3 Objectifs du projet

L'objectif principal de ce P-SAT est de contourner ces limitations méthodologiques en adoptant une approche radicalement différente : le crowdsourcing, ou participation citoyenne volontaire. Plutôt que de chercher à simuler artificiellement des utilisateurs réels, nous mobilisons de véritables utilisateurs humains pour collecter des données authentiques sur leur expérience personnalisée des plateformes.

Concrètement, le projet s'appuie sur STETOSCOPE, une plateforme de collecte collaborative, pour atteindre trois objectifs opérationnels :

- **Auditer les pratiques de personnalisation** sur des plateformes commerciales majeures (Amazon, Booking.com, AliExpress, Temu) en collectant les données d'affichage réelles vues par différents utilisateurs pour un même produit ou une même recherche.
- **Détecter et documenter des manipulations concrètes** : variations de prix injustifiées (discrimination tarifaire), différences dans l'ordre des résultats de recherche (discrimination algorithmique), faux compteurs de stock ou de vues (dark patterns d'urgence), promotions trompeuses (réductions fictives).
- **Automatiser l'extraction et l'analyse de données** visuelles complexes à partir des captures d'écran fournies par les participants, en exploitant des technologies avancées : reconnaissance optique de caractères (OCR), expressions régulières (Regex), et modèles de vision par apprentissage profond (LLM multimodaux).

Ce projet s'inscrit ainsi dans une démarche de transparence algorithmique et de protection des consommateurs, en fournissant des preuves empiriques des pratiques de manipulation numérique.

2 État de l'art

2.1 La collecte de données et la personnalisation

Le profilage est la norme sur Internet. Des études récentes montrent que même les assistants vocaux profilent les utilisateurs pour cibler les publicités. Pourtant, déterminer précisément quelles informations sont collectées reste un défi. Des discriminations de prix (proposer un prix différent pour le même produit selon l'utilisateur) et de recherche (ordonner les résultats différemment pour influencer l'achat) ont été documentées, mais leurs causes racines restent mal comprises.

2.2 Les outils de mesure existants

Les outils actuels pour détecter les manipulations se classent en plusieurs catégories :

- **Systèmes basés sur des règles** : Ils scannent le code des pages web à la recherche de signaux prédéfinis (ex : bannières de cookies trompeuses). Cette méthode échoue face aux tactiques subtiles ou changeantes.
- **Systèmes basés sur les données (Machine Learning)** : Ils utilisent des modèles pour identifier visuellement ou textuellement des dark patterns.
- **Approches hybrides** : Outils comme UI Guard qui analysent à la fois la mise en page et le langage.

Cependant, avec la migration massive des usages vers le mobile, l'automatisation de ces audits est devenue ardue, les plateformes utilisant des techniques de « fingerprinting » avancées pour exclure tout trafic non-humain. C'est ici que l'approche participative de STETOSCOPE se distingue, en réintégrant l'humain dans la boucle de collecte.

3 La plateforme STETOSCOPE

3.1 Vue d'ensemble de l'architecture

STETOSCOPE est une plateforme de crowdsourcing initialement développée pour auditer la transparence des plateformes numériques sur Android. Elle se compose d'une application mobile pour les participants et d'un serveur backend pour l'administration. L'architecture globale, illustrée dans la Figure 1, permet de coordonner les différentes phases du processus de collecte.

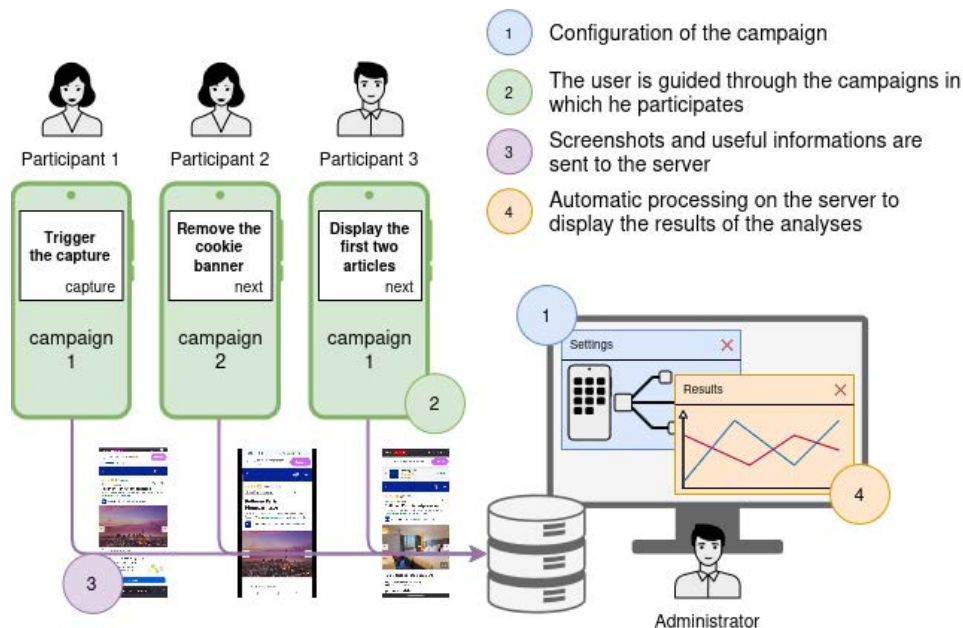


FIGURE 1 – Architecture de STETOSCOPE : 1) Campagne configurée, 2) Guidage de l'utilisateur, 3) Capture d'écran, 4) Traitement automatique.

Le fonctionnement opérationnel se décline en quatre étapes successives. Tout d'abord, l'administrateur configure une campagne via l'interface web, définissant les scénarios et les instructions de guidage. Ensuite, le participant sélectionne cette campagne dans l'application mobile (voir Figure 2), où une bannière en surimpression le guide jusqu'à l'information cible. Une fois l'objectif atteint, l'utilisateur déclenche manuellement une capture d'écran qui est instantanément transmise au serveur backend. Enfin, le serveur traite automatiquement ces captures pour en extraire les données pertinentes et les structurer.

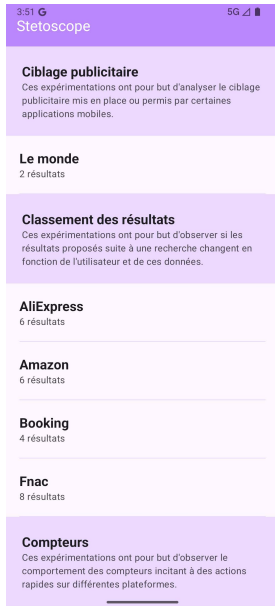


FIGURE 2 – Interface de l'application mobile listant les campagnes.

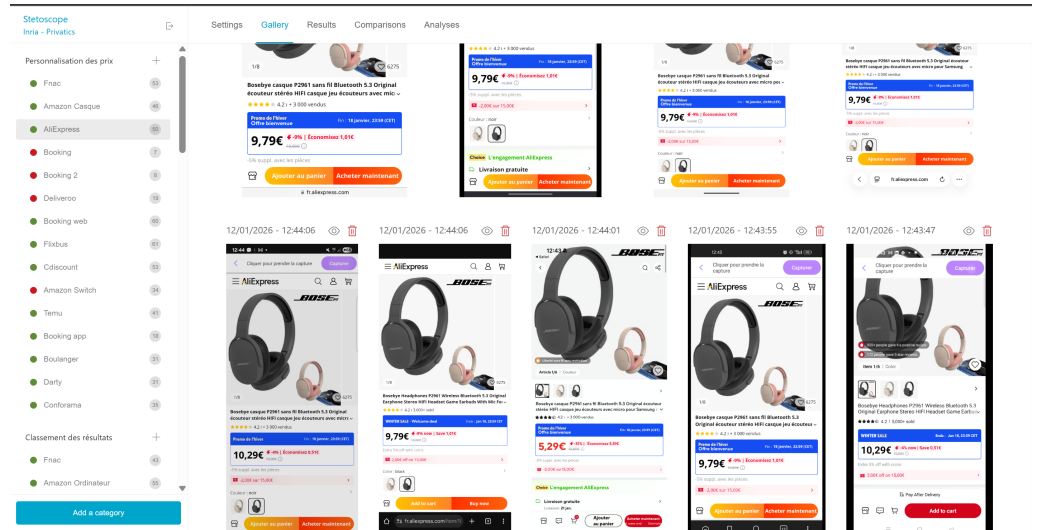


FIGURE 3 – Tableau de bord de l'administrateur pour la gestion et l'analyse.

3.2 Limites initiales

Bien que fonctionnel, l'outil présentait plusieurs limitations majeures avant notre intervention. Sur le plan technique, l'extraction des données reposait exclusivement sur la reconnaissance optique de caractères (OCR) couplée à des expressions régulières (Regex). Cette méthode, bien que performante sur des structures fixes, manquait cruellement de flexibilité face aux mises à jour fréquentes des interfaces web. De plus, l'écosystème de collecte était exclusivement restreint aux utilisateurs d'Android, excluant de fait une large population d'utilisateurs d'iOS ou de navigateurs de bureau. Enfin, les capacités de visualisation intégrées au tableau de bord d'origine étaient limitées, ne permettant pas d'effectuer des analyses croisées complexes ou un filtrage dynamique des données récoltées.

4 Amélioration de STETOSCOPE : Nos Contributions

Afin de pallier les limitations précédemment identifiées, nous avons développé trois axes d'amélioration majeurs qui constituent le cœur de notre apport au projet.

4.1 Extension multi-plateforme (Web)

Afin d'étendre la portée de la collecte au-delà de l'écosystème Android, nous avons développé une version Web de STETOSCOPE. Sur le plan technologique, cette extension interagit directement avec les API backend existantes, permettant ainsi de soumettre des captures d'écran et des métadonnées depuis n'importe quel navigateur moderne. L'impact de cette évolution est significatif, puisqu'elle rend la collecte possible sur des plateformes jusqu'alors inaccessibles comme iOS ou les ordinateurs de bureau. Cette versatilité permet de diversifier considérablement les profils utilisateurs et d'enrichir les environnements d'audit (voir Figures 4 et 5).

4.2 Analyse avancée par LLM

L'extraction de données par Regex montrant ses limites face à la versatilité des interfaces de e-commerce, nous avons implémenté un pipeline innovant exploitant les modèles de langage de grande taille (LLM). Le

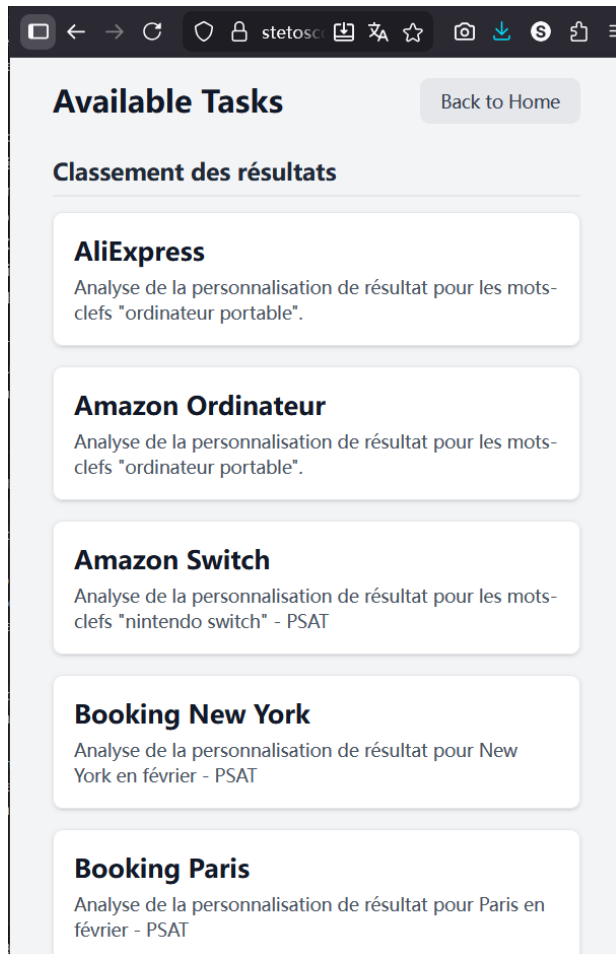


FIGURE 4 – Interface de la liste des tâches de la version Web de STETOSCOPE.

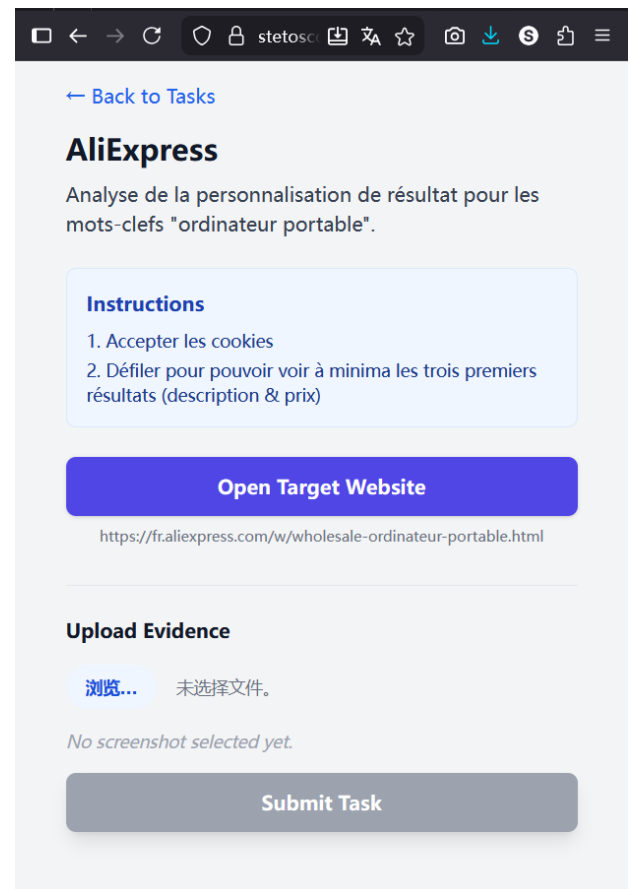


FIGURE 5 – Interface détaillée d'une tâche sur la version Web (exemple AliExpress).

coeur du dispositif repose sur un script Python qui transmet les captures d'écran à l'API d'OpenAI, en utilisant des modèles dotés de capacités de vision. Grâce à une ingénierie de prompt rigoureuse, le modèle est capable d'extraire des informations de manière sémantique. Contrairement à l'approche par Regex, le LLM peut traiter des cas complexes tels que l'identification simultanée de plusieurs produits sur une même capture, l'analyse de l'ordre d'apparition des résultats, ou encore la distinction critique entre les résultats de recherche organiques et les publicités sponsorisées. En sortie, les données sont structurées sous format JSON avant d'être converties en CSV pour faciliter les phases ultérieures d'analyse statistique.

4.3 Visualisation et Analyse via Power BI

Pour transformer la masse de données brutes récoltée en informations stratégiques, nous avons intégré Power BI comme outil pivot de visualisation. Cette solution offre des capacités de filtrage et d'agrégation bien supérieures au tableau de bord d'origine. Désormais, il est possible d'agréger les variations de prix en fonction de critères précis tels que le modèle de smartphone, la localisation géographique ou la période temporelle. L'interactivité des graphiques Power BI permet de mettre rapidement en évidence des corrélations complexes et de documenter de manière visuelle et indiscutable les pratiques de personnalisation algorithmique identifiées au cours du projet (voir Figure 6).

5 Campagnes de Collecte de Données

La collecte des données a été structurée autour de trois approches complémentaires, permettant d'obtenir un jeu de données à la fois régulier et ponctuellement riche.

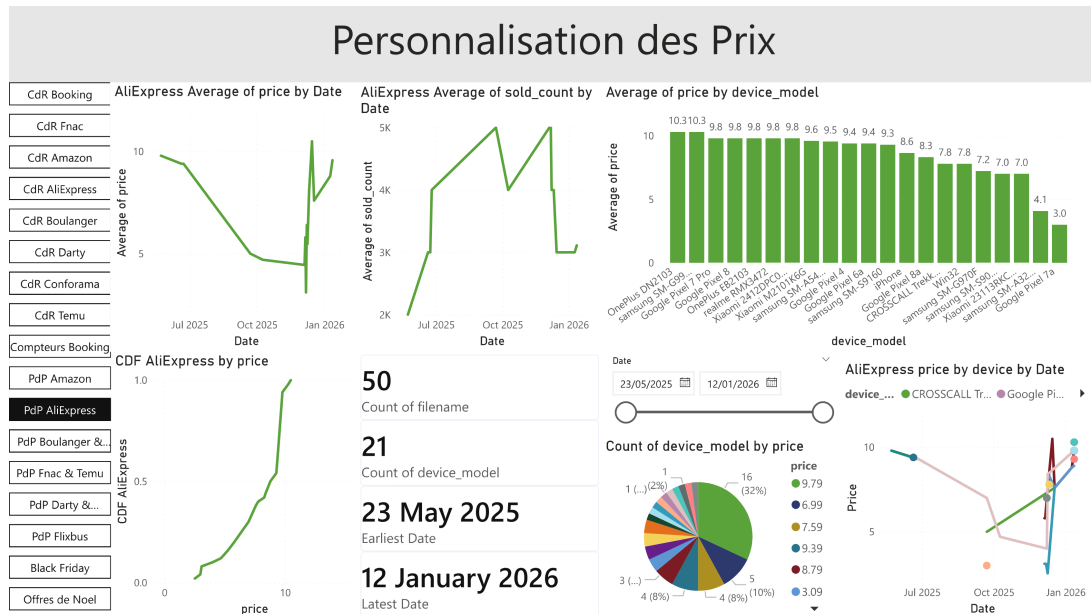


FIGURE 6 – Tableau de bord Power BI présentant l'analyse de la personnalisation des prix sur AliExpress.

5.1 Méthodologies de collecte

La première approche repose sur une collecte quotidienne assurée par les trois membres de l'équipe projet. En utilisant l'application STETOSCOPE au fil de leur navigation personnelle, ils ont pu documenter de nombreux cas de personnalisation des prix rencontrés "en conditions réelles". Cette source de données continue a été complétée par un suivi saisonnier et événementiel rigoureux, ciblant particulièrement les périodes de fortes promotions telles que le Black Friday et les fêtes de Noël. Durant ces phases critiques, l'effort de collecte s'est focalisé sur les articles arborant des labels promotionnels afin d'auditer la véracité des rabais affichés.

Enfin, une campagne d'envergure sous forme d'atelier participatif a été organisée à l'INSA Lyon le 12 janvier 2026, en collaboration étroite avec Antoine BOUTET. Cet événement a réuni plus de vingt participants autour d'une présentation pédagogique traitant des dark patterns et des enjeux de la personnalisation algorithmique, ponctuée de sessions interactives (voir Figure 7). Pour encourager la participation et garantir un climat convivial, une distribution de pizzas gratuites a été organisée. Sur le plan scientifique, cet atelier a permis de générer un pic massif de données synchronisées, facilitant la comparaison directe des profils et l'identification des biais algorithmiques.



FIGURE 7 – Affiche de la campagne de collecte organisée à l'INSA Lyon.

6 Analyse des Résultats

Dans cette section, nous détaillons les observations issues des différentes campagnes de collecte, en mettant en lumière les pratiques de personnalisation et les dark patterns identifiés.

6.1 Personnalisation des prix sur AliExpress

L'une des premières campagnes menées visait à vérifier si le prix d'un même article fluctuait en fonction du profil de l'utilisateur ou du terminal utilisé. Les observations réalisées sur la plateforme AliExpress ont été particulièrement révélatrices à cet égard. Des variations de prix significatives ont été constatées pour

des produits identiques, les tarifs oscillant entre 3€ et 8,50€ sans qu'aucune justification apparente, comme d'éventuels frais de livraison différenciés, ne puisse expliquer de tels écarts. L'analyse de ces données suggère fortement la mise en oeuvre d'une stratégie de personnalisation tarifaire basée sur l'historique d'achat ou sur le profil socio-démographique supposé de l'utilisateur.

6.2 Classement des résultats sur Amazon

Un second axe d'analyse portait sur l'ordre d'affichage des produits lors d'une recherche standardisée, comme celle du terme « ordinateur portable ». Sur Amazon, nous avons observé qu'à requête identique, l'ordre des résultats proposés variait considérablement d'un participant à l'autre. Plus frappant encore, le prix moyen des premiers articles mis en avant présentait des disparités majeures, allant de 235€ à plus de 1100€. Ces résultats tendent à démontrer que la plateforme oriente activement certains profils vers des gammes de produits nettement plus onéreuses, influençant ainsi indirectement les décisions d'achat en limitant la visibilité des alternatives plus économiques.

6.3 Incohérence des compteurs sur Booking.com

Parallèlement aux questions de prix, nous nous sommes penchés sur la véracité des mécanismes de pression sociale, tels que les compteurs d'urgence (ex : « Plus que 2 chambres disponibles »). Sur Booking.com, l'étude des compteurs d'avis a mis en lumière des anomalies flagrantes : nous avons détecté des cas où la valeur affichée décroissait au fil du temps. Cette observation est mathématiquement aberrante pour une donnée cumulative et laisse supposer soit une défaillance technique du système, soit, plus probablement, l'utilisation de compteurs fictifs destinés à simuler une popularité ou une urgence artificielle.

6.4 Fausses promotions sur Temu lors du Black Friday

Enfin, nos analyses saisonnières ont révélé des pratiques de prix barrés potentiellement trompeuses lors d'événements comme le Black Friday. Sur la plateforme Temu, un article affiché avec une remise spectaculaire de « -72% » pour un prix final de 136€ durant l'événement a été retrouvé à 150€ peu de temps après la fin des promotions. L'absence de corrélation entre la réduction affichée et la réalité de la variation tarifaire suggère que ces badges promotionnels sont souvent utilisés comme de simples artifices visuels pour inciter à l'achat immédiat, sans bénéfice réel pour le consommateur.

7 Analyse et Discussion

7.1 Apports du projet

L'approche participative adoptée par STETOSCOPE prouve sa pertinence pour auditer les algorithmes de personnalisation, particulièrement dans l'écosystème mobile souvent opaque. En impliquant directement l'utilisateur final, nous parvenons à contourner les systèmes de protection anti-bots et à collecter des données authentiques reflétant l'expérience réelle des consommateurs.

Nos contributions techniques ont apporté une valeur ajoutée significative à cette démarche. D'une part, l'introduction des modèles de langage de grande taille (LLM) pour l'extraction automatique des données a considérablement accru la flexibilité du système, permettant d'analyser des captures d'écran visuellement complexes là où les méthodes traditionnelles par Regex échouaient. D'autre part, le passage à Power BI a transfiguré l'exploitation de ces données. Cet outil permet de déceler avec une précision accrue des tendances de personnalisation jusque-là invisibles, en croisant par exemple les variations tarifaires avec la marque du terminal ou la localisation géographique précise de l'utilisateur.

7.2 Limitations et Éthique

Bien que nous puissions observer avec précision les résultats de la personnalisation (l'output), le système algorithmique des plateformes reste une boîte noire quant aux mécanismes décisionnels internes. De plus, la collecte de données privées via des captures d'écran a nécessité une attention particulière pour garantir

une conformité totale avec le RGPD. Des protocoles de floutage automatique ont été mis en place pour protéger les informations sensibles des participants, sous la supervision du DPO (Délégué à la Protection des Données).

8 Conclusion

Ce projet P-SAT a démontré la puissance du crowdsourcing pour lever le voile sur les pratiques opaques des plateformes numériques. En s'appuyant sur STETOSCOPE, nous avons pu collecter et analyser des preuves empiriques de manipulations algorithmiques et de dark patterns.

Nos travaux ont permis de transformer une plateforme de collecte Android initiale en un écosystème d'audit multi-plateforme performant. L'intégration de modèles de vision LLM et de tableaux de bord Power BI a radicalement amélioré la qualité et la profondeur de l'extraction et de l'analyse des données. Les campagnes menées ont confirmé des pratiques préoccupantes : discriminations tarifaires sur AliExpress, biais de classement sur Amazon, et promotions artificielles sur Temu.

À terme, ces outils pourraient être mis à disposition d'observatoires de la transparence numérique ou de régulateurs pour assurer une protection accrue des consommateurs face aux dérives de la personnalisation algorithmique.

Références

- [1] Beaufils, T., Boutet, A., et al. *STETOSCOPE : underStand TargEting and manipulaTiOnS via Collabo-rative Private data collection*. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies.
- [2] Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2018). *The dark (patterns) side of UX design*. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (pp. 1-14).
- [3] Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., & Narayanan, A. (2019). *Dark patterns at scale : Findings from a crawl of 11K shopping websites*. Proceedings of the ACM on Human-Computer Interaction, 3(CSCW), 1-32.
- [4] Hannak, A., Soeller, G., Lazer, D., Mislove, A., & Wilson, C. (2014). *Measuring price discrimination and steering on e-commerce web sites*. In Proceedings of the 2014 conference on internet measurement conference (pp. 305-318).
- [5] Regulation (EU) 2016/679 of the European Parliament and of the Council. *General Data Protection Regulation (GDPR)*. Official Journal of the European Union, 2016.