

Année 2025-2026

## Analyse de la personnalisation et des dark patterns via la plateforme de crowdsourcing Stetoscope

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

### Encadrant 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 peut être faite à l'aide de robots, mais ils sont souvent limités par les systèmes anti-fraude des plateformes. Stetoscope propose donc une solution innovante à ce défi : une plateforme de crowdsourcing qui collecte des données authentiques sous forme de captures d'écran auprès d'utilisateurs réels. Il permet ainsi d'auditer la transparence des algorithmes sans déclencher les défenses anti-bots. De plus, les utilisateurs lambda possèdent un historique de navigation ou des préférences de cookies susceptibles d'influencer la personnalisation de contenu.

#### Objectifs

Ce projet P-SAT vise à auditer les pratiques de manipulation de plusieurs plateformes majeures du commerce en ligne (Amazon, Booking, AliExpress, Temu entre autres). 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 promotions. Pour cela, nous mobiliserons de véritables utilisateurs pour collecter leur expérience sur les plateformes via Stetoscope. Techniquement, nos objectifs seront d'élargir l'utilisation de l'application et d'automatiser l'extraction de données pertinentes à partir des captures d'écrans récoltées.

#### Méthodologie

Pour mener à bien ce projet, nous avons prévu un plan d'action bien concret. Nous commençons par préparer nos campagnes de test directement sur l'interface d'administration de Stetoscope en choisissant les scénarios d'achat et les sites à surveiller. Le gros du travail technique se jouera ensuite sur le traitement des données : nous développons des scripts Python qui s'appuient sur des modèles LLM de vision pour transformer automatiquement les captures d'écran en données structurées. En parallèle, nous nous occupons de l'accès à l'outil aux utilisateurs d'iPhone et d'ordinateurs en développant une interface Web, car l'application n'était disponible que sur Android jusqu'ici. Une fois toutes ces informations en main, nous utiliserons Power BI pour créer des visuels parlants et essayer de dénicher des preuves de discrimination tarifaire ou de tri biaisé.

Nous finirons par une campagne de collecte grande nature auprès des étudiants de l'INSA pour confronter nos outils à la réalité du terrain.

## 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<sup>e</sup> 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 Beaufils ([tao.beaufils@inria.fr](mailto:tao.beaufils@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

L'économie numérique 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. Cependant, dans la pratique, ces systèmes de recommandation fonctionnent comme des « boîtes noires » : 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 de prix (un même produit affiché à des tarifs différents selon le profil utilisateur), ou de recherche (des résultats ordonnés différemment pour orienter les choix d'achat). S'ajoutent à cela les « dark patterns », des mécanismes d'interface conçus pour exploiter les biais cognitifs humains : faux compteurs de stock créant un sentiment d'urgence, promotions mensongères, ou encore boutons trompeurs compliquant la désinscription.

## 2 État de l'art

### 2.1 Personnalisation, Profilage et Discrimination

Le profilage des utilisateurs est devenu un pilier de l'économie numérique. Cette collecte massive de données personnelles permet une personnalisation poussée des contenus, mais soulève des défis majeurs en matière de transparence [2]. Des études fondatrices ont démontré que cette personnalisation ne se limite pas à l'amélioration de l'expérience utilisateur, mais peut dériver vers des pratiques discriminatoires.

Hannak et al. [3] ont notamment documenté des phénomènes de *discrimination de prix* (variation des tarifs pour un même produit) et de *steering* (réordonnancement des résultats de recherche pour orienter l'utilisateur vers des options plus coûteuses). Malgré ces preuves empiriques, les causes racines — qu'elles soient liées à la localisation, à l'historique de navigation ou au terminal utilisé — restent complexes à isoler sans une collecte de données rigoureuse et à grande échelle.

## 2.2 Les Dark Patterns

Les manipulations d’interface, ou « dark patterns », sont désormais omniprésentes. Mathur et al. [4] les définissent comme des choix de conception qui exploitent les biais cognitifs pour inciter les utilisateurs à prendre des décisions contraires à leurs intérêts (ex : urgence artificielle, difficulté de désinscription).

La recherche actuelle tente de standardiser ces concepts [5], afin de les classifier selon leur intention et leur impact sur l’autonomie de l’utilisateur. Loin d’être accidentelles, ces techniques sont utilisées volontairement pour pousser à l’action, entrant souvent en conflit avec le RGPD [6].

## 2.3 Outils de Détection et Limites de l’Automatisation

Pour débusquer ces pratiques, plusieurs méthodes ont été testées par la communauté scientifique. Au début, on s’appuyait surtout sur des règles fixes pour scanner le code source, comme avec l’outil *DarkDialogs* [8], ce qui fonctionne assez bien pour des éléments statiques comme les bannières de cookies. Mais face à des interfaces de plus en plus dynamiques, les chercheurs se sont tournés vers le Machine Learning [9] pour essayer de « comprendre » visuellement la manipulation plutôt que de simplement lire le code. Sur mobile, des approches hybrides comme *UI Guard* [10] vont encore plus loin en combinant vision par ordinateur et analyse de texte (NLP) pour ne rien laisser passer.

Le gros souci aujourd’hui, c’est que les plateformes sont devenues très fortes pour bloquer les robots. Comme l’expliquent Venugopalan et al. [11], les techniques de *fingerprinting* permettent désormais aux sites de griller immédiatement si c’est un humain ou un script qui navigue. Cette « course à l’armement » rend les audits automatisés souvent partiels ou carrément faussés, car le contenu affiché à un robot n’est pas forcément le même que celui d’un vrai utilisateur. C’est là que le projet *Stetoscope* [1] entre en jeu : au lieu de lutter contre les anti-bots, on passe par le crowdsourcing. En utilisant de vraies captures d’écran prises par des gens, on s’assure d’avoir des données authentiques que les algorithmes de détection ne peuvent pas bloquer.

## 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 4, permet de coordonner les différentes phases du processus de collecte.

Le fonctionnement opérationnel se décline en quatre étapes successives. Tout d’abord, l’administrateur configure une campagne via l’interface web (voir Figure 3), 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.

### 3.2 Limites initiales

Bien que fonctionnel, l’outil présentait plusieurs limitations que nous avons essayé d’améliorer. 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. De plus, l’interface administrateur ne communiquait pas d’informations sur l’avancement de l’analyse, ce qui a compliqué son utilisation. L’écosystème de collecte était exclusivement restreint aux utilisateurs d’Android, excluant de fait une large population d’utilisateurs. Enfin, les capacités de visualisation intégrées au tableau de bord d’origine étaient assez limitées, ne permettant pas d’effectuer des analyses croisées ou un filtrage dynamique.

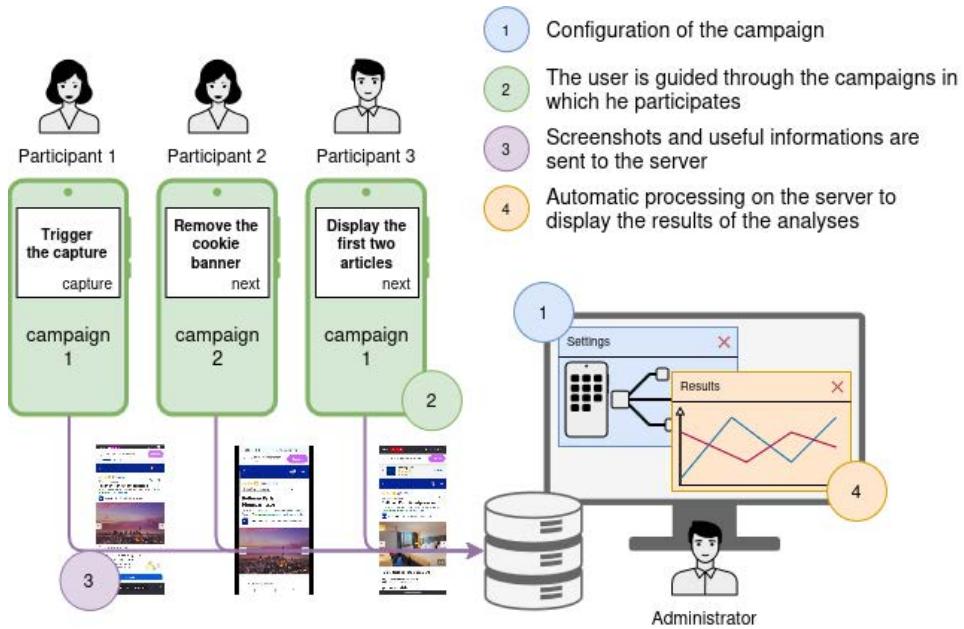


FIGURE 1 – Architecture de Stetoscope : 1) Campagne configurée, 2) Guidage de l’utilisateur, 3) Capture d’écran, 4) Traitement automatique.



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

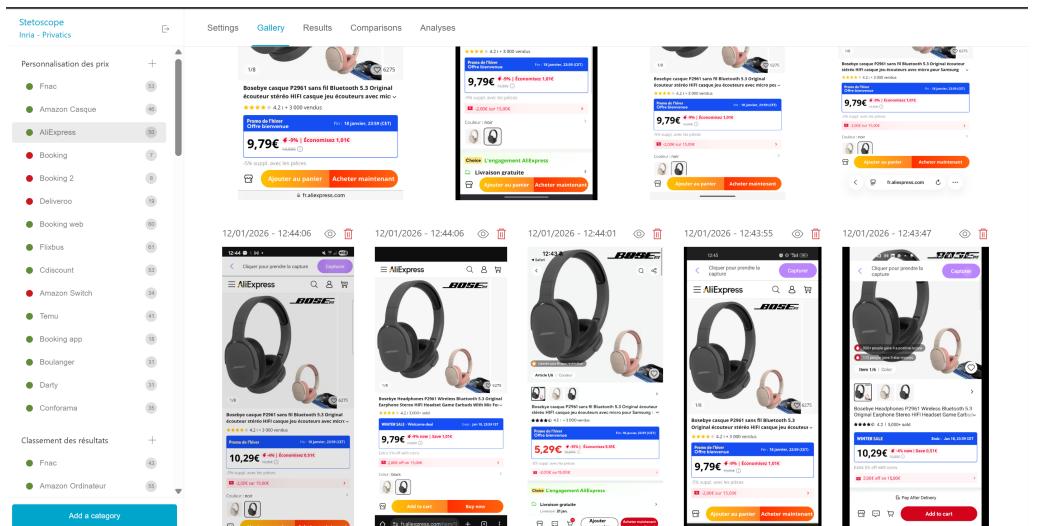


FIGURE 3 – Tableau de bord de l’administrateur pour la gestion et l’analyse.

## 4 Amélioration de Stetoscope : Nos Contributions

Afin de pallier les limitations précédemment identifiées, nous avons développé trois axes d’amélioration.

### 4.1 Extension multi-plateforme (Web)

Afin d’étendre les collectes au-delà des appareils 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.

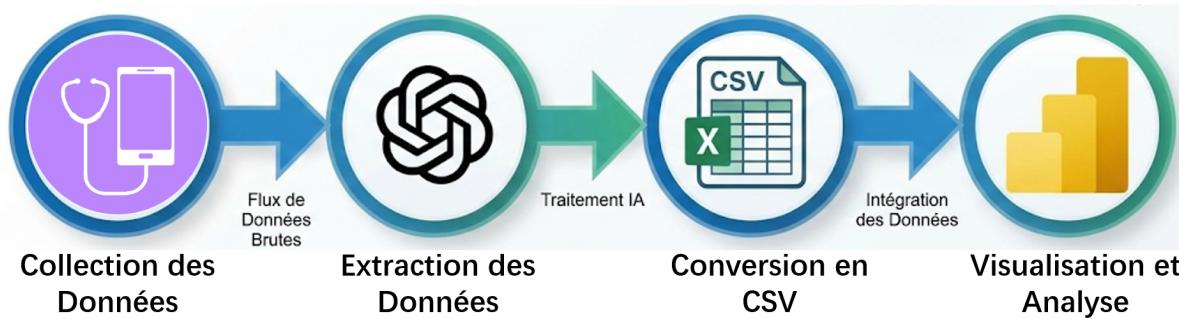


FIGURE 4 – Pipeline de traitement des données : de la collecte via Stetoscope à la visualisation Power BI.

moderne. L'impact de cette évolution est significatif, puisqu'il est maintenant possible de participer aux campagnes depuis un iOS ou un ordinateur. Par conséquent, nous sommes capables d'intégrer n'importe quel utilisateur et de diversifier les appareils étudiés(voir Figures 5 et 6).

**Available Tasks**

**Classement des résultats**

- AliExpress**  
Analyse de la personnalisation de résultat pour les mots-clés "ordinateur portable".
- Amazon Ordinateur**  
Analyse de la personnalisation de résultat pour les mots-clés "ordinateur portable".
- Amazon Switch**  
Analyse de la personnalisation de résultat pour les mots-clés "nintendo switch" - PSAT
- Booking New York**  
Analyse de la personnalisation de résultat pour New York en février - PSAT
- Booking Paris**  
Analyse de la personnalisation de résultat pour Paris en février - PSAT

Back to Home

**← Back to Tasks**

**AliExpress**

Analyse de la personnalisation de résultat pour les mots-clés "ordinateur portable".

**Instructions**

1. Accepter les cookies
2. Défiler pour pouvoir voir à minima les trois premiers résultats (description & prix)

**Open Target Website**

<https://fr.aliexpress.com/w/wholesale-ordinateur-portable.html>

**Upload Evidence**

浏览... 未选择文件。

No screenshot selected yet.

**Submit Task**

FIGURE 5 – Interface de la liste des tâches de la version Web de STETOSCOPE.

FIGURE 6 – Interface détaillée d'une tâche sur la version Web (exemple AliExpress).

## 4.2 Analyse avancée par LLM

L'extraction de données par Regex a très vite été insuffisante pour exploiter les résultats, donc nous avons implémenté un pipeline indépendant de Stetoscope qui exploite les grands modèles de langage (LLM). Cette implémentation repose sur un script Python qui transmet les captures d'écran à l'API d'OpenAI.

Nous avons fait attention au prompt utilisé pour que le modèle soit 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 ou l'analyse de l'ordre d'apparition des résultats. 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 les données brutes récoltées, nous avons intégré Power BI comme outil de visualisation. Nous avons choisi cette solution car elle offre de très bonnes capacités de filtrage et d'agrégation, et que nous avions jamais eu l'occasion de l'utiliser. Nous avons donc pu réaliser différents graphiques adaptés aux campagnes pour visualiser les variations de prix selon l'utilisateur ou la période temporelle par exemple. (voir Figure 7).

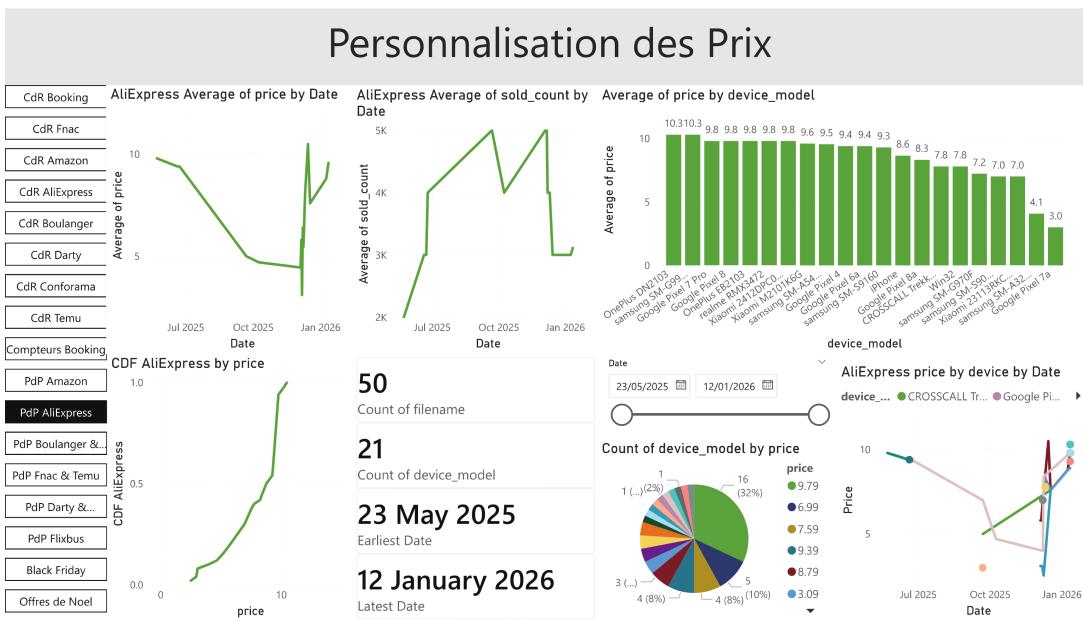


FIGURE 7 – Tableau de bord Power BI présentant l'analyse de la personnalisation des prix sur AliExpress.

## 5 Campagnes de Collecte de Données

### 5.1 Types de campagnes

Depuis la plateforme d'administration de Stetoscope, nous avons défini différentes catégories de campagnes. Celles ci sont regroupées dans le tableau ci-dessous :

Nom	Description
Personnalisation des prix	Observation du prix pour un même article
Classement des résultats	Observation des résultats proposés suite à une même recherche d'articles
Compteurs	Observation des compteurs incitant à des actions rapides (nombre d'expériences, stock restant...)
Personnalisation genrée	Observation de l'influence du genre sur les résultats d'une recherche
Black Friday	Analyse des prix pendant et après le Black Friday
Offres de Noël	Analyse des prix avant, pendant et après les offres de Noël

Nous avons conduits ces campagnes sur différents types de plateformes que nous avons identifiées avec Antoine. On peut distinguer les plateformes selon 4 catégories :

- E-commerce : AliExpress, Amazon, Boulanger, Cdiscount, Conforama, Darty, Fnac, Temu
- Réservation d'hôtels : Booking
- Réservation de transports : Flixbus
- Recrutement : Indeed

Celles-ci seront potentiellement explorées par le PReN à l'aide de robots, donc l'objectif sera de comparer les résultats obtenus ultérieurement.

### 5.2 Organisation des campagnes

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

La première approche repose sur une collecte régulière effectuée avec nos appareils personnels. Les captures pourront ainsi être faites durant différentes périodes et de manière répétée. Nous avons été plus vigilants pendant les périodes de fortes promotions, tels que le Black Friday ou Noël. Le but était donc de comparer l'évolution des prix pour un même article au cours du temps, et de vérifier la véracité des réductions affichées.

La deuxième approche vise à toucher un plus grand nombre d'utilisateurs simultanément. Une campagne a été organisée à l'INSA de Lyon le 12 janvier 2026, avec l'aide de notre tuteur Antoine. Cet événement a réuni plus de vingt participants autour d'une présentation pédagogique traitant des enjeux de la personnalisation, puis d'une collecte de données avec l'application Stetoscope (voir Figure 8). Afin d'attirer les étudiants et de rendre le moment plus convivial, le laboratoire de l'INRIA a financé l'achat de pizzas qui ont été distribuées gratuitement pour inciter des participants à venir. 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 8 – 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. Pour le classement des résultats on a constaté que si les résultats diffèrent par appareils on mettait un ✓ même si le prix moyen ne changeait guère. Ceci est juste une observation, nous allons plonger plus en détail dans les sections suivantes sur les pratiques que nous jugeons les plus discriminatoires.

TABLE 1 – Tableau récapitulatif des mécanismes constatés pour chaque plateforme de e-commerce

Plateformes	Personnalisation de Prix	Classement des Résultats
AliExpress	✓	✓
Amazon	✗	✓
Booking	✓	✓
Boulanger	✗	✗
Cdiscount	✓	✗
Conforama	✗	✗
Darty	✗	✓
Fnac	✗	✓
Temu	✗	✓

Autre que ces plateformes, nous avons également organisé une campagne pour différencier les résultats d'une recherche d'emploi sur Indeed selon le genre de l'utilisateur, mais les mêmes offres sont systématiquement affichés.

De même pour Flixbus, nous avons observé des prix qui fluctuent et une cohérence entre les utilisateurs pour un même trajet.

### 6.1 Personnalisation des prix

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 dans le temps.

Les observations réalisées sur la plateforme AliExpress pour un casque audio ont été particulièrement révélatrices à cet égard. Du 22 mai 2025 au 12 janvier 2026, son prix a fluctué entre 2,53€ et 10,49€ sans aucune justification apparente (voir figure 9). De plus, nous avons constaté des différences entre utilisateurs au même moment. Lors de notre campagne du 12 janvier, nous avons récolté 18 captures d'écrans sur 12 modèles d'appareil différents (dont 5 iphones), et on observe une gamme de prix allant de 5,29 € à 10,29 € sur le graphique circulaire ???. L'analyse de ces données suggère fortement une stratégie de personnalisation tarifaire basée sur l'historique d'achat ou sur le profil supposé de l'utilisateur.

De la même manière, nous avons observé des réductions différentes selon l'utilisateur pour des écouteurs sur Cdiscount, et des prix de chambres d'hôtels contradictoires sur Booking.

### 6.2 Classement des résultats

Un second axe d'analyse portait sur la distribution de prix des produits affichés lors d'une recherche standardisée

Comme celle du terme « ordinateur portable » sur Amazon, nous avons observé qu'à requête identique, les prix moyens des premiers articles mis en avant présentait des disparités majeures, allant de 220€ à plus de 1100€. Voir figure 11. 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.

Par ailleurs il est aussi intéressant de noter que sur Booking.com

Cependant, comme nos collectes sont limitées cet écart n'est pas toujours probabilistiquement significatif. Sur la même plateforme (Amazon) pour une autre recherche de "nintendo switch", nous avons aussi un écart

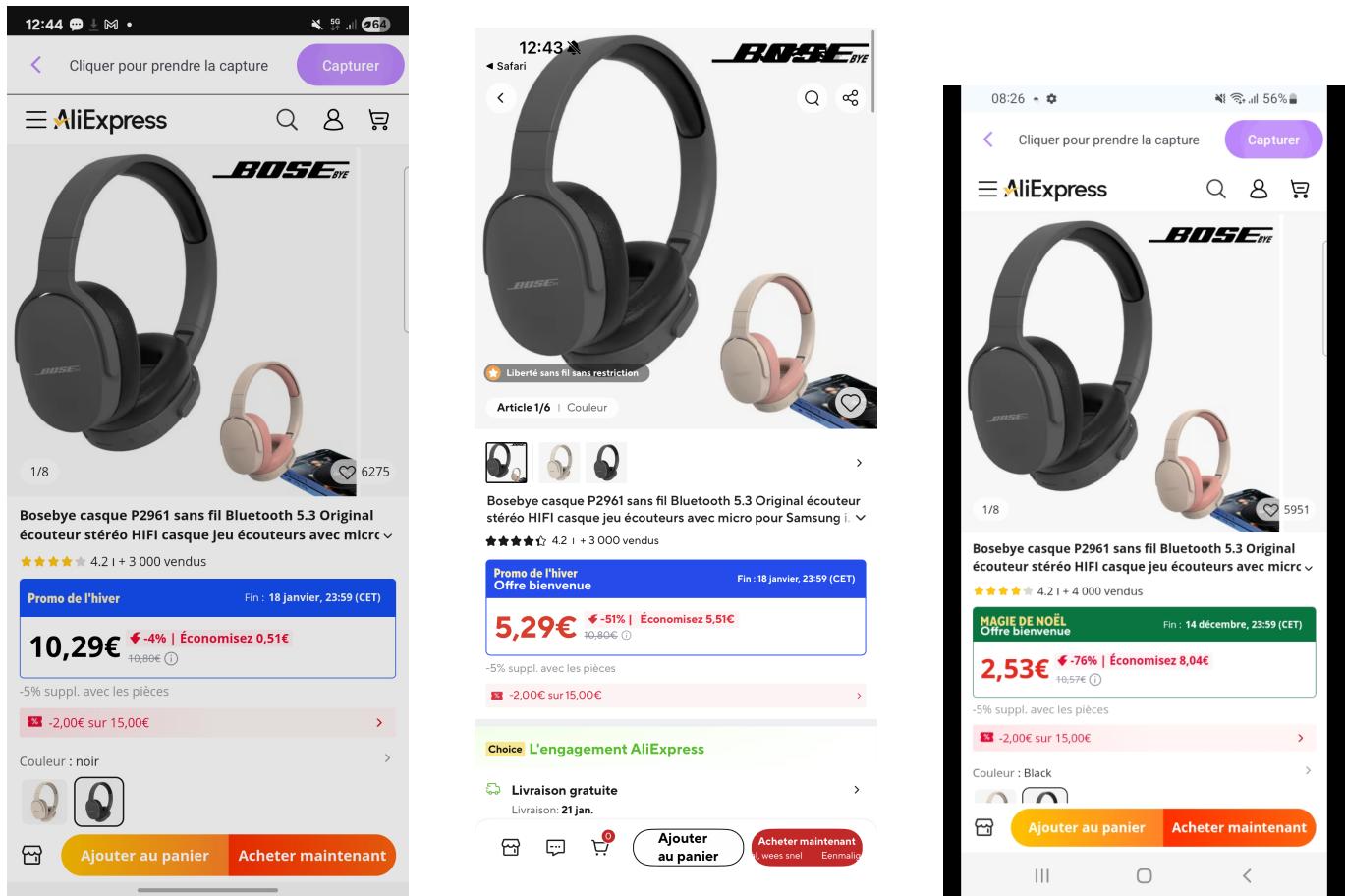


FIGURE 9 – Preuve de prix différenciés : même produit et même instant, mais tarifs divergents sur AliExpress

des moyennes des prix de 152 € à 360 € voir figure 12. Par contre cet différence peut s'expliquer car pour la moyenne la plus basse nous avons eu un article à 19,91 € qui n'est pas apparue ailleurs. Cela met en cause la limite de notre collecte, pour pouvoir avoir de meilleurs résultats il nous faudrait de plus de 50 capture d'écrans au même moment au lieu des 21 que nous avons réussi à avoir.

### 6.3 Incohérence des compteurs

Parallèlement aux questions de prix, nous nous sommes penchés sur la véracité des mécanismes de pression sociale, tels que les compteurs d'urgence par exemple : "Plus que 2 articles disponibles". Cet outils de conversion (un nudge) déclenche deux bias cognitifs. La peur de manquer (FOMO - Fear Of Missing Out) "Si je ne réserve pas maintenant, quelqu'un d'autre va prendre la chambre" ainsi que la validation sociale "Si 20 personnes ont réservé aujourd'hui, c'est que l'hôtel est bien"

Sur Booking.com, l'étude des compteurs d'expériences vécu (qui semble être décrire le nombre d'avis et de commentaire) a mis en lumière différentes tendances. Nous avons détecté un cas où la valeur affichée augmentait ainsi qu'un autre cas où elle diminuait, dans tous les deux cas elle semble suivre une tendance linéaire. Un observation étrange pour une donnée cumulative qui peut nous faire penser à une défaillance technique du système, soit, plus probablement, l'utilisation de compteurs fictifs destinés à simuler une popularité ou une urgence artificielle.

Autre part, sur Amazon nous avons constaté que sur certains appareils un nudge est présent avec le nombre d'article disponible alors que sur la majorité des autres mobile la seul le fait que l'article est en stock est affiché.

Count of device\_model by price

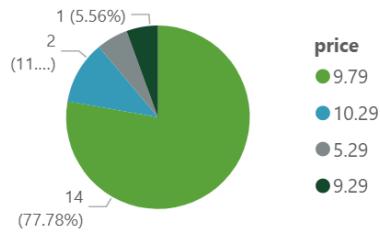


FIGURE 10 – Données du 12 janvier 2026 pour 18 capture d'écrans et 12 types d'appareilles

Average of price by device\_model

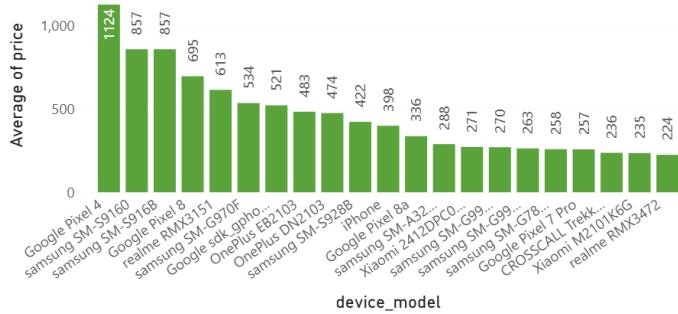


FIGURE 11 – Données du 23 may 2025 jusqu'au 12 janvier 2026 pour 54 capture d'écrans et 22 types d'appareilles

## 6.4 Fausses promotions

Enfin, nous avons pu observer certaines pratiques déloyales pendant des périodes de fortes affluences, tels que le Black Friday ou Noël. Sur la plateforme Temu, un article affiché avec une remise spectaculaire de 72% pour un prix final de 136€ durant le Black Friday a été retrouvé à 150€ peu de temps après la fin des promotions. De plus, la promotion est clairement mise en avant avec une bannière rouge "Promos Black Friday" pour inciter à l'achat, alors qu'en réalité il n'y a pas de bénéfice réel pour le consommateur (voir Figure 14). Sur Amazon/Fnac/Cdiscount, les offres promotionnelles du Black Friday étaient toujours actives plus d'une semaine après l'événement. Encore aujourd'hui, certaines réductions sont toujours en cours même si elles ont été légèrement abaissées. Par exemple, un sèche cheveux Dyson sur Amazon bénéficie mi-janvier de 14% de réduction, contre 20% lors du Black Friday.

Des réductions trompeuses ont aussi pu être capturées pendant la période de Noël sur Conforama et Fnac. En effet, les promotions de certains articles ont disparu en janvier, mais le prix affiché reste le même. Cela suggère que le prix de référence était gonflé artificiellement pour simuler une bonne affaire, et que le prix "soldé" correspondait en réalité au prix habituel de l'article.

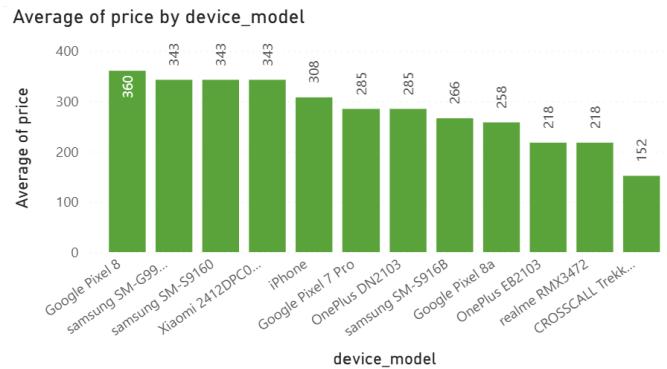


FIGURE 12 – Données du 12 janvier 2026 pour 21 capture d'écrans et 12 types d'appareils

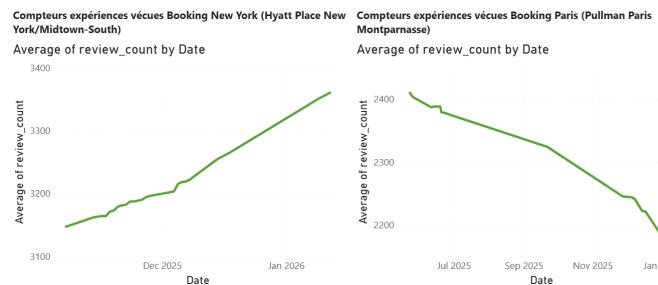


FIGURE 13 – Données du ...

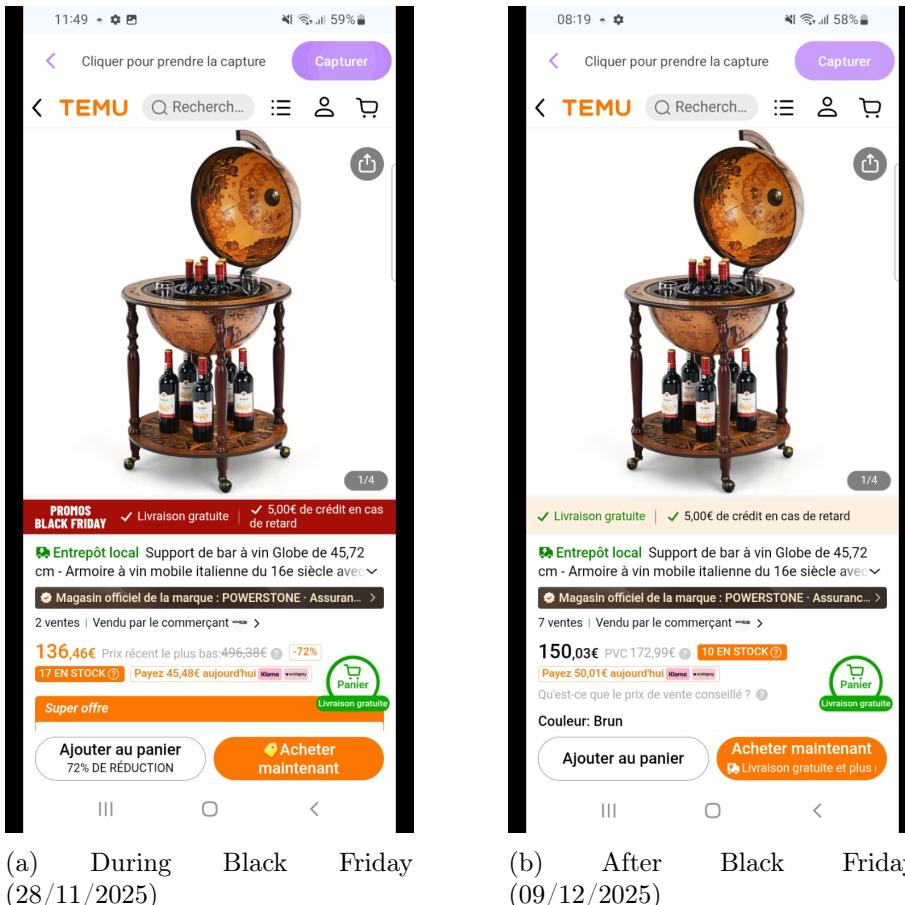


FIGURE 14 – Réduction du Black Friday incohérente par rapport au prix affiché après la période de promotion

## 7 Analyse et Discussion

### 7.1 Perception des utilisateurs

Pendant la récolte de données organisée mi-janvier, nous avons d'abord interrogé les étudiants sur leur perception de ces pratiques. Parmi eux, 88% avaient déjà été témoins d'une discrimination de prix, notamment l'augmentation de leurs billets de transport après une première consultation. Il est cependant difficile de savoir si cette augmentation était strictement personnel ou si elle avait eu lieu pour tous les utilisateurs. La majorité d'entre eux étaient conscients d'obtenir des résultats différents ou similaires pour une même recherche en fonction de la plateforme utilisée. Ils ont qualifiés les compteurs comme utiles et incitants à l'achat, mais parfois aussi malhonnêtes. Enfin, leurs réactions face aux réductions du Black Friday étaient plutôt partagées (voir ci dessous Figure 15) :

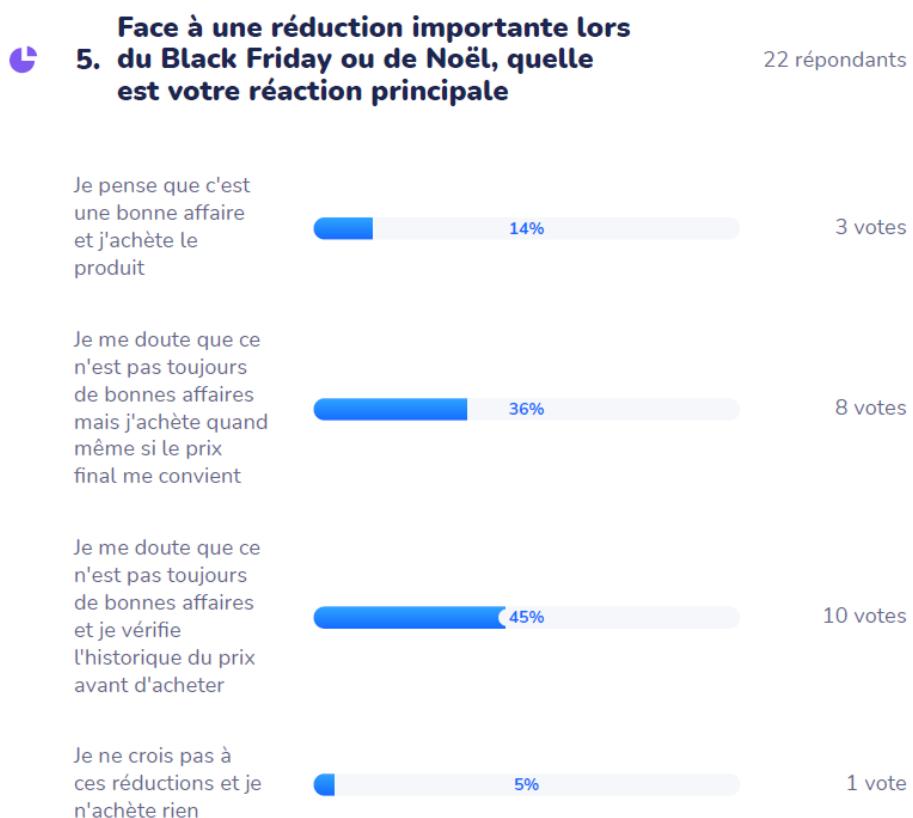


FIGURE 15 – Sondage réalisé auprès d'étudiants de l'INSA le 12/01/2026

### 7.2 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.

### 7.3 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

### Références

- [1] Beaufils, T., Boutet, A., et al. (2025). *STETOSCOPE : underStand TargEting and manipulaTiOnS via Collaborative Private data collection*. Proceedings of the ACM on IMWUT.
- [2] Melicher, W., et al. (2016). *Online Tracking : A 1-Million-Site Measurement and Analysis*. Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security.
- [3] Hannak, A., et al. (2014). *Measuring price discrimination and steering on e-commerce web sites*. In Proceedings of the 2014 conference on internet measurement conference (pp. 305-318).
- [4] Mathur, A., et al. (2021). *What Makes a Dark Pattern... Dark ? Design Attributes, Normative Considerations, and Measurement Methods*. arXiv :2101.04843.
- [5] Gray, C. M., et al. (2023). *An Ontology of Dark Patterns Knowledge : Foundations, Definitions, and a Pathway for Shared Knowledge-Building*. arXiv :2309.09640.
- [6] Bielova, N. (2023). *Présentation des enjeux informatiques des dark patterns*. Vidéo Canal-U, VDN.
- [7] Regulation (EU) 2016/679 (GDPR). (2016). *General Data Protection Regulation*. Official Journal of the European Union.
- [8] Kirkman, et al. (2023). *DarkDialogs : Automatic Detection of Dark Patterns in Cookie Banners*.
- [9] Yada, S., et al. (2022). *Dark patterns in e-commerce : A dataset and its baseline evaluations*. IEEE BigData.
- [10] Chen, J., et al. (2023). *UI Guard : Detecting Dark Patterns in Mobile Applications*. Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology.
- [11] Venugopalan, R., et al. (2025). *FP-Inconsistent : Measurement and Analysis of Fingerprint Inconsistencies in Evasive Bot Traffic*. IMC '25.