

Capacitation des utilisateurs sur les réseaux sociaux

11/06/2025

Nicolas Soulié
Maître de conférences
Economie numérique

nicolas.soulie@imt-bs.eu

Adrian Popescu
Ingénieur-chercheur
Intelligence artificielle

adrian.popescu@cea.fr

Jérôme Deshayes
Ingénieur-chercheur
Intelligence artificielle

jerome.deshayes-chossart@cea.fr



Rappel des objectifs

- Travailler sur les risques et opportunités du partage des données personnelles sur les réseaux sociaux
 - Focus sur les photos car partagées fréquemment et porteuses d'informations implicites
- Modéliser des situations impactantes : recherche d'un crédit bancaire, d'une assurance, d'un logement, d'un travail
- Volet étude utilisateurs : testing en ligne de l'influence des contenus partagés
- Volet algorithmique : intégration dans un prototype d'application mobile d'une analyse automatique des photos et de modèles algorithmiques des situations

Avancement application YDSYO

Création d'une association pour publier l'app

- Structure pérenne dédiée à la mise en place d'applications type YDSYO
- Date de création : 11/02/2025
- Structure
 - Président: Bertrand DELEZOIDE
 - Trésorier : Jérôme DESHAYES
 - Secrétaire : Adrian Popescu
- Dernier round de discussion avec le CEA pour publier le prototype d'application
 - Ok donné en mars 2025

VIOLETAI
Activité

Mission

Capacitation des utilisateurs de systèmes numériques par l'explication du fonctionnement des algorithmes d'intelligence artificielle exploitant des données personnelles partagées en ligne, sensibilisation à la protection de la vie privée et à la personnalisation des services, assurance de la transparence et de la sécurité des données en fournissant des moyens d'information et de protection aux utilisateurs dans leurs interactions numériques, exploitation des résultats de projets de recherche transférés pour une publication à but non lucratif, et sensibilisation du public aux droits associés aux données personnelles

Commune

Sceaux (92330)

Domaine

- Recherche médicale

Contact

Informations déclarées à l'administration :

- La mairie peut avoir des informations : [Contacter la mairie](#)

Autour de l'association

Secteur

Les associations du département

- Recherche médicale - [Voir les associations proches](#)
- Toutes les associations - [Voir les autres associations du département](#)

Services pour cette association

Partenaires

- Vous souhaitez diffuser vos offres auprès des associations, ce sera bientôt possible ici grâce à [votre compte assoce](#)

Assoce

- Vous êtes une administration publique ? Débloquez l'accès à toutes les données du site avec [votre compte Assoce](#)
- Vous souhaitez diffuser vos actions, vos informations de contact, ce sera bientôt possible sur [votre compte assoce](#)

Mise à jour d'YDSYO

- Coeur algorithmique inchangé
- Modélisation de quatre situations impactantes
 - Recherche crédit bancaire, appartement, travail IT ou restauration
- EfficientDet-lite4 pour la détection visuelle
 - Modèle fonctionnant sur la plupart des téléphones Android
 - Quantification - performances techniquement sous-optimales
 - Surtout par rapport aux gros VLMs disponibles actuellement
- Base d'utilisateurs
 - 1000 profils avec 100 images par profil
 - Extraite à partir de Flickr et anonymisée
 - Communauté suffisamment large pour situer les utilisateur

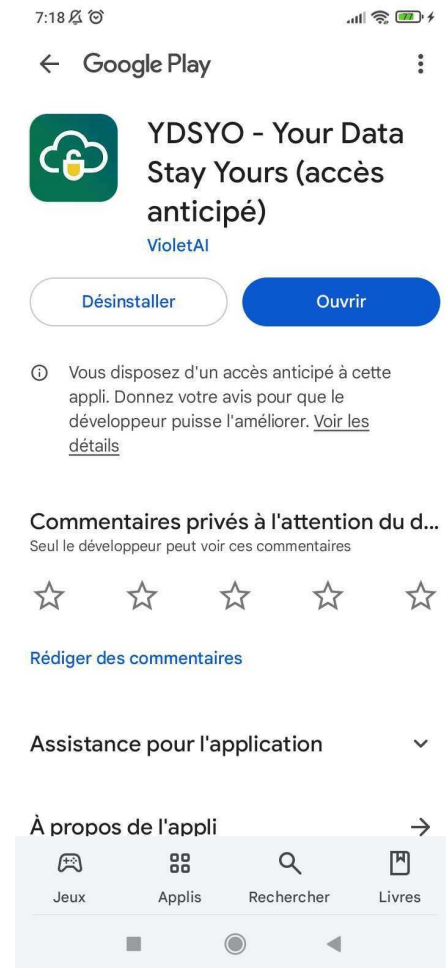


Mise à jour d'YDSYO

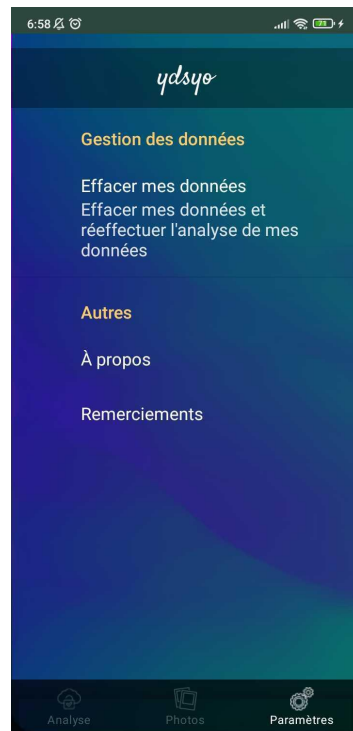
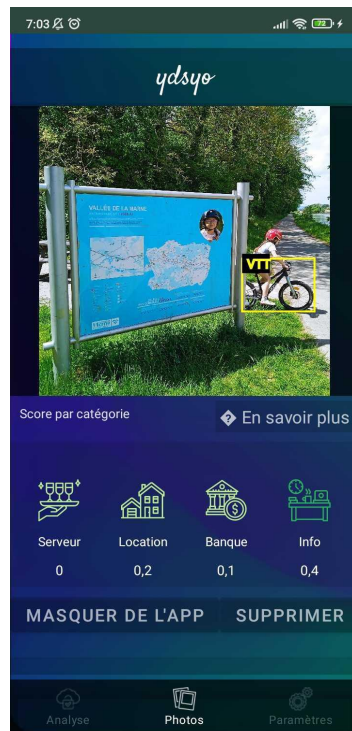
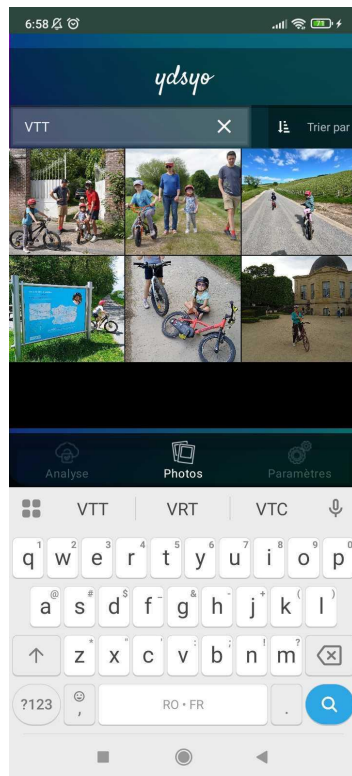
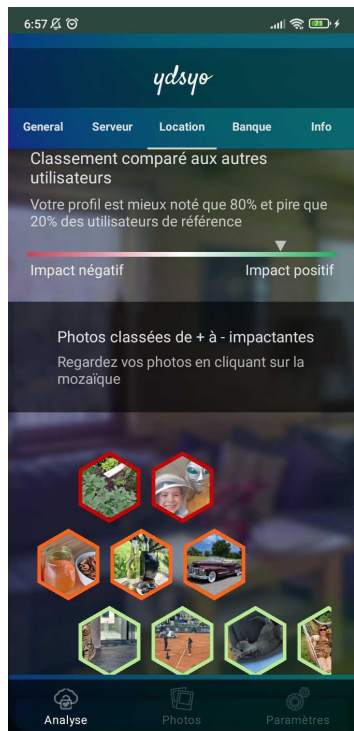
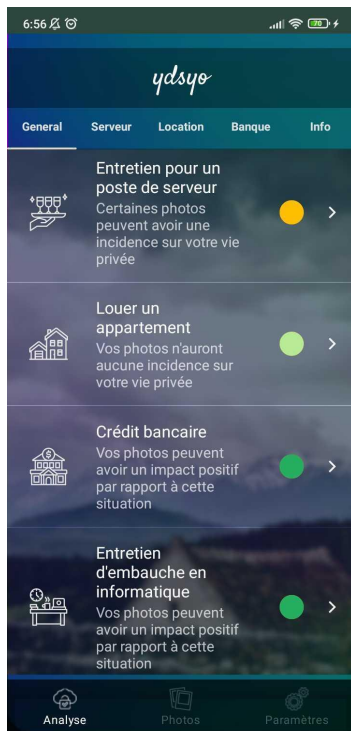
- Correction des erreurs
 - Mise à jour des librairies “deprecated”
 - Librairies Butterknife, TFLite...
 - Mise à jour des versions minimales et recommandées SDK pour s'accorder aux appareils actuels
- Lancement des tests fermés qui permettent de découvrir des erreurs “silencieuses”
 - Certains appareils n'arrivaient plus à importer leur photo -> mise à jour de la librairie PermissionsDispatcher

Mise à jour d'YDSYO

- Publication de l'application en accès anticipé sur Google PlayStore réalisée
- Phase de test commencée le 10/06/2025 pour deux semaines
- 16 testeurs actuellement (12 minimum)
- Publication effective à l'issue de la phase de test



Auto-analyse



Améliorations possibles

- Enrichir la base d'apprentissage des modèles edge avec des annotations semi-automatiques par LMM
- Optimiser les performances avec des grands modèles multimodaux (LMM) au lieu de modèles edge
- Compromis sur la vie privée - envoi des photos sur des serveurs GPU
 - Utiliser un serveur dédié avec un LMM hébergé



Describe this photo. Be precise.

Thought for 14 seconds >

A light-grey, short-haired cat is lying on its right side on a grey, textured fabric surface (likely a bedspread). The cat's body runs almost straight from the top to the bottom of the frame, with its head near the upper left and its hind legs extended together toward the lower right. Its front paws are tucked close to its chest, while the rear legs are stretched out and parallel. The cat is wearing a dark collar fitted with a circular tracking tag—black rim with a pale centre—positioned just behind its left ear. Soft light from the left creates gentle shadows, and a few small specks of lint are scattered on the fabric around the cat.

Bilan scientifique

© 2021 IEEE. This is the author's version of the article that has been published in the proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. The final version of this record is available at <https://ieeexplore.ieee.org/document/9706741> (Van-Khoa Nguyen, Adrian Popescu, Jérôme Deshayes-Chossart, "Unveiling Real-Life Effects of Online Photo Sharing," IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 2022, pp. 3859-3869, doi: 10.1109/WACV51458.2022.00391).

Unveiling Real-Life Effects of Online Photo Sharing

Van-Khoa Nguyen, Adrian Popescu, Jérôme Deshayes-Chossart
Université Paris-Saclay, CEA, List, F-91120, Palaiseau, France
khoa.v18nguyen@gmail.com, [\[adrian.popescu, jerome.deshayes-chossart\]@cea.fr](mailto:[adrian.popescu, jerome.deshayes-chossart]@cea.fr)

Abstract

Social networks give free access to their services in exchange for the right to exploit their users' data. Data sharing is done in an initial context which is chosen by the users. However, data are used by social networks and third parties in different contexts which are often not transparent. In order to unveil such usages, we propose an approach which focuses on the effects of data sharing in impactful real-life situations. Focus is put on visual content because of its strong influence in shaping online user profiles. The approach relies on three components: (1) a set of visual objects with associated situation impact ratings obtained by crowdsourcing, (2) a corresponding set of object detectors for mining users' photos and (3) a ground truth dataset made of 500 visual user profiles which are manually rated per situation. These components are combined in LERVUP, a method which learns to rate visual user profiles in each situation. LERVUP exploits a new image descriptor which aggregates object ratings and object detections at user level and an attention mechanism which boosts highly-rated objects to prevent them from being overwhelmed by low-rated ones. Performance is evaluated per situation by measuring the correlation between the automatic ranking of profile ratings and a manual ground truth. Results indicate that LERVUP is effective since a strong correlation of the implementation of the user awareness ab

awareness about data processing through feedback contextualization. To do this, we introduce a plausible decision-making system which combines machine learning and domain knowledge. User awareness is increased by linking the sharing process to impactful situations such as searching for a job, an accommodation, or a bank credit. Photos are in focus because they constitute a large part of shared data and contribute strongly to shaping user profiles [2]. The main technical contribution is a method that rates visual user profiles and individual photos in a given situation by exploiting situation models, visual detectors and a dedicated photographic profiles dataset. The proposed method, named LERVUP from LERning to Rate Visual User Profiles, learns a ranking of user profiles which attempts to reproduce human profiles ranking. LERVUP exploits a new descriptor which combines object impact ratings and object detections in a compact form. The contributions of objects with high ratings are boosted in order to mimic the way humans assess photographic content. We compare manual and automatic rankings of user profile ratings and obtain a positive correlation between them. This result holds promise to help users better understand the effects of online data sharing and, ultimately, to better control their data.

2. Related work

1. Introduction

The ubiquitous use of online social networks shows that their services are appealing to users to implement a business model in which access to user data is exchanged for user data monetization [11]. Intrusiveness is likely to grow with the wide usage of AI techniques to infer actionable information from users' data. Automatic in

ferent people and algorithms, an affinity-based network of positive or negative information is present in contexts [4]. This interpretation provided by human and/or technical biases. Human biases are often studied in social factors such as gender or ethnicity [50]. Implicit biases [18] might influence one's decisions without that person being conscious of them. Explicit biases are assumed and used intentionally. Technical

© Owner/Author(s) 2023. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in Hugo Schindler, Adrian Popescu, Khoa Nguyen, and Jérôme Deshayes-Chossart. 2023. Raising User Awareness about the Consequences of Online Photo Sharing. In Proceedings of the 2023 ACM International Conference on Multimedia Retrieval (ICMR '23). Association for Computing Machinery, New York, NY, USA, 10–19. <https://doi.org/10.1145/3591106.3592290>

Raising User Awareness about the Consequences of Online Photo Sharing

Hugo Schindler
hugo-schindler@orange.fr
Université Paris-Saclay, CEA, List
F-91120, Palaiseau, France

Adrian Popescu
adrian.popescu@cea.fr
Université Paris-Saclay, CEA, List
F-91120, Palaiseau, France

Van-Khoa Nguyen
van-khoa.nguyen@etu.unige.ch
University of Geneva
Geneva, Switzerland

Jérôme Deshayes-Chossart
jerome.deshayes-chossart@cea.fr
Université Paris-Saclay, CEA, List
F-91120, Palaiseau, France

ABSTRACT

Online social networks use AI techniques to automatically infer profiles from users' shared data. However, these inferences and their effects remain, to a large extent, opaque to the users themselves. We propose a method which raises user awareness about the potential use of their profiles in impactful situations, such as searching for a job or an accommodation. These situations illustrate usage contexts that users might not have anticipated when deciding to share their data. User photographic profiles are described by automatic object detections in profile photos, and associated object ratings in situations. Human ratings of the profiles per situation are also available for training. These data are represented as graph structures which are fed into graph neural networks in order to learn how to automatically rate a profile. An adaptation of the learning procedure per situation is proposed since the same profile is likely to be interpreted differently, depending on the context. Automatic profile ratings are compared to one another in order to inform individual users of their standing with respect to others. Our method is evaluated on a public dataset, and consistently outperforms competitive baselines. An ablation study gives insights about the role of its main components.

CCS CONCEPTS

• Computing methodologies → Neural networks; • Security and privacy → Privacy protections.

KEYWORDS

images, graph neural networks

ACM Reference Format:
Hugo Schindler, Adrian Popescu, Khoa Nguyen, and Jérôme Deshayes-Chossart. 2023. Raising User Awareness about the Consequences of Online Photo Sharing. In *International Conference on Multimedia Retrieval (ICMR '23)*, June 12–15, 2023, Thessaloniki, Greece. ACM, 10 pages. <https://doi.org/10.1145/3591106.3592290>

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 the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or

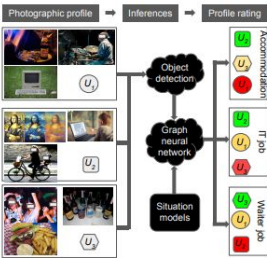


Figure 1: Potential consequences of personal data sharing for 3 users in 3 real-life situations. Users' photos include information about their interests and lifestyle which can be inferred and aggregated to compute the appeal of a profile in context. The aggregation is done with a graph neural net trained to rate profiles. The network is trained with object detections and ratings per situation and uses human ratings to model several situations. Modeling information can be done with the two top row photos when searching for accommodation, but positively for a waiter job. Top row photos are positive for IT job, but their photos with high ratings are problematic if searching for a waiter job. Profile ratings are used to raise awareness about the consequences of data sharing.

IEEE WACV 2022

(rang A)

ACM ICMR 2023

(rang B)

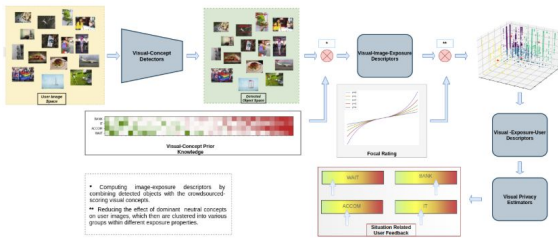
Bilan technique

README

LERVUP: LEarning to Rate Visual User Profiles

Official repository.

Paper



- Computing image exposure descriptors by certifying selected objects with the unsupervised learning visual concepts.
- Reducing the effect of dominant visual concepts on user ratings, which have been clustered into various groups within different exposure properties.

Unveiling Real-Life Effects of Online Photo Sharing [arXiv]

If you find our work useful, please cite:

Code

AIcrowd

Search...

Official Round: Completed

ImageCLEF 2022 Aware

By ImageCLEF

2639 25 0 10 1 Share

Overview Leaderboard Discussion Insights Resources Submissions Rules

Dataset

7:18

Google Play

YDSYO - Your Data Stay Yours (accès anticipé)
VioletAI

Désinstaller Ouvrir

Vous disposez d'un accès anticipé à cette appli. Donnez votre avis pour que le développeur puisse l'améliorer. [Voir les détails](#)

Commentaires privés à l'attention du d...
Seul le développeur peut voir ces commentaires

☆ ☆ ☆ ☆ ☆

Rédiger des commentaires

Assister **Prototype**

À propos de l'appli

Jeux Appis Rechercher Livres

Testing utilisateurs

Testing: principe général

Nouveaux clients potentiels: quel impact de la réputation en ligne ?

- Informations personnelles disponibles sur les médias sociaux
- Pratique indiquée dans les CGU
- Cas des comportements à risque: consommation d'alcool

Testing:

- Création de faux profils “identiques” sur Facebook + GMail
- Alimentation avec des photos signalant ou non des comportements à risque
- Envoi de demande de devis Assurance Auto (+ immobilier et banque)

Testing 1ère vague

- Testing avec 5 profils masculins
- Assurance Auto
- Pour chaque profil, 8 photos contenant ou non le signal “alcool”:

Force du signal “alcool”				
0% Lucas Fasseaux	25% Hugo Dasseaux	50% Thomas Basseaux	75% Maxime Casseaux	100% Alexandre Gasseaux

	Photo avec signal
	Photo sans signal

Testing: photos



Testing: photo



Testing 1ère vague: résultat

	Force signal “Alcool”				
	0% Lucas Fasseaux	25% Hugo Dasseaux	50% Thomas Basseaux	75% Maxime Casseaux	100% Alexandre Gasseaux
Nombre d’envois	1080	1076	1081	1079	1078
% retour répondeur	28.80%	32.62%	29.97%	31.79%	31.26%
% retour mail	51.48%	50.37%	53.10%	51.07%	53.06%

- Taux très similaires
- Photos pas assez différenciantes

Testing 2ième vague: photos

Photos:

- Signal “alcool” plus fort:
 - Alcool fort (whisky + vodka) + attitude “trash”
 - 16 photos avec 2 X 2 photos par semaine
- Personnage principal + 1 ou 2 amis (évalués en amont)

Testing:

- Création faux profils “identiques” sur Facebook + GMail: Quentin Jasseaux, Théo Lasseaux, Alexis Masseaux, Nicolas Nasseaux et Antoine Passeaux
- 5 signaux: i) Contrôle, ii) Soft drinks, iii) Alcool fort, iv) Alcool fort + Trash, et v) Trash

Profil: Contrôle (Quentin Jasseaux)



Profil: Soft drinks (Alexis Masseaux)



Profil: Alcool fort (Théo Lasseaux)



Profil: Alcool fort + Trash (Nicolas Nasseaux)



Profil: Trash (Antoine Passeaux)



Testing 2ième vague: profils et mails

Recherche profil FB:

- Combinaison Nom et Prénom unique sur FB
- Photo et information dans signature du mail de demande de devis



Téléphone : 07.68.62.39.24

Comptable (Promotion 2023-2024, IAE Gustave Eiffel)

Moteurs de recherche:

- Google: pas présent dans les résultats (trois 1ères pages)
- Bing: 1er résultat

Testing 2ième vague: résultat

3 types de réponse non “exploitable”:

- Mail bloqué
- Adresse incorrecte
- Réponse automatique

5 types de réponse “exploitable”:

- Demande d'information
- Demande d'information avec formule politesse
- Demande d'information et contact
- Demande contact
- Téléphone: message et SMS

Testing 2ième vague: résultat mail

	Contrôle Quentin Jasseaux	Soft drinks Alexis Masseaux	Alcool Théo Lasseaux	Alcool+Trash Nicolas Nasseaux	Trash Antoine Passeaux
Nombre d'envois	520	518	519	520	519
Envois non exploitables	61	63	56	71	62
Envois exploitables	459	455	463	449	457
Complément d'info "sec"	42,2	58,2	55,8	46,5	45,3
Compl. d'info + politesse	17,8	7,8	4,7	17,6	22,1
Compl. d'info + contact	24,4	19,6	19,2	28,3	22,0
Prise contact	15,0	13,1	20,4	6,3	9,4
% moyen de retour	62,3	60,5	62,5	62,0	58,3

Testing 2ième vague: résultat téléphone

	Contrôle Quentin Jasseaux	Soft drinks Alexis Masseaux	Alcool Théo Lasseaux	Alcool+Trash Nicolas Nasseaux	Trash Antoine Passeaux
Envois exploitables	459	455	463	449	457
Nombre de message	41,5	36,4	40,9	40,4	44,0
SMS	4	0	0	1	0
Appels infructueux/ mails reçus	16,4	10,3	11,4	12,6	12,6

- Un peu plus de rappels pour contrôle et soft drink
- Résultat “anecdotique”: 3 ou 4 rappels + protocole (ne pas répondre)

Testing 3ième vague: agences immobilières

Contexte:

- Recherche d'un logement: appartement ~50m²
- Mail similaire (sauf Paris/Province)

Protocole identique 2ième vague:

- Faux profils “identiques” sur Facebook + GMail: Quentin Jasseaux, Théo Lasseaux, Alexis Masseaux, Nicolas Nasseaux et Antoine Passeaux
- 5 signaux: i) Contrôle, ii) Soft drinks, iii) Alcool fort, iv) Alcool fort + Trash, et v) Trash

Testing 3ième vague: résultats

Testing:

- Randomisation: département, taille ville, franchisée/non franchisée
- Envoi de 500 demandes/profil

3 types de réponse non “exploitable”: i) mail bloqué, ii) adresse incorrecte et iii) réponse automatique

3 types de réponse “exploitable”:

- Demande d’information (mail ou lien dossier)
- Demande contact
- Demande d’information et contact
- Téléphone: message et SMS

Testing: 3ième vague, résultat mail

	Contrôle Quentin Jasseaux	Soft drinks Alexis Masseaux	Alcool Théo Lasseaux	Alcool+Trash Nicolas Nasseaux	Trash Antoine Passeaux
Nombre d'envois	512	509	515	508	511
Envois non exploitables	72	68	77	65	66
Envois exploitables	440	441	438	443	445
Demande information	23,7	26,8	21,2	25,3	23,5
Demande contact	10,2	9,5	11,7	12,8	11,5
Demande information + contact	7,1	5,9	8,2	6,0	6,7
% moyen de retour	40,9	42,2	41,1	44,1	41,7

Testing 3ième vague: résultat téléphone

	Contrôle Quentin Jasseaux	Soft drinks Alexis Masseaux	Alcool Théo Lasseaux	Alcool+Trash Nicolas Nasseaux	Trash Antoine Passeaux
Envois exploitables	440	441	438	443	445
Nombre de message (%)	12,4	16,2	15,3	11,8	14,9
SMS (nombre)	1	2	1	3	4

- Plus de relance pour profils Contrôle et Soft drinks ?

Conclusion

3 vagues de testing:

- 2 x Assurance + 1 x agence immobilière
- 1000 demandes + 2 x 500 demandes
- Taux de retour mail/téléphone très similaires

Données personnelles et accès aux services:

- Pas d'évidence d'usage de données personnelles issues des médias sociaux (Facebook) lors des 1ers échanges avec clients potentiels
- Rappels plus insistants pour “bons” profils: nombre très faible (p/r 500 demandes de devis) + protocole (pas de rappel)

Merci pour votre attention !