



03

DATA ONBOARDING

- INTRODUCTION
- BÉNÉFICES ET CAS D'USAGE
- MISE EN PLACE D'UNE SOLUTION DE DATA ONBOARDING
- CARTOGRAPHIE DES ACTEURS
- CHALLENGES DU MARCHÉ

Chapitre réalisé en collaboration avec

DARWIN
smart digital agency

#1 INTRODUCTION

A. CHIFFRES CLÉS DU MARCHÉ

- **Un Français sur trois** utilise quotidiennement 2 écrans ou plus pour surfer sur Internet (Médiamétrie – janvier 2017).
- **36% des internautes français** utilisaient un adblocker en novembre dernier (IPSOS).
- Selon LiveRamp, pour ces actions marketing digitales, un annonceur exploite aux Etats-Unis **entre 20 à 25 plateformes technologiques différentes** pour ses actions marketing. En France, ce chiffre serait compris entre 7 et 10.
- Parce que de plus en plus de marketers adoptent une approche omnicanal, le marché de l'Onboarding aux Etats-Unis est en très forte croissance, avec **une hausse de 30M\$ en 2012 à 250M \$ en 2016, et prévoit d'atteindre le milliard de \$ en 2020**, selon le Winterburry Group Report "The State of Consumer Data Onboarding".

B. LE RÔLE STRATÉGIQUE DE L'ONBOARDING

Le Data Onboarding est une solution permettant de digitaliser la donnée. Elle consiste à unifier et réconcilier les données offline d'un individu avec ses identifiants digitaux cross-device (cookies ou identifiants mobiles).

À l'ère du People-based Marketing, le Data Onboarding permet aux marketers de délivrer des expériences réellement ciblées, personnalisées et mesurables et ce, dans un univers cross-device et omnicanal.

Les solutions de Data Onboarding sont une des pierres angulaires de l'audience planning. La marque s'adresse directement à l'individu plutôt qu'un ciblage mass-market et place le client au centre de sa stratégie marketing.

Le consommateur est quant à lui désormais exposé à une réponse à ses besoins, plutôt qu'à une publicité.

#2 BÉNÉFICES ET CAS D'USAGE

Des campagnes marketing plus pertinentes avec des résultats significatifs, pour les consommateurs et les marques, tel est l'objectif de cette solution technologique innovante.

Il existe trois principaux cas d'usage : le targeting, la personnalisation du site web et la mesure de la performance des campagnes digitales sur les transactions en magasin physique.

A. TARGETING

Il s'agit certainement du plus répandu aujourd'hui. Grâce au Data Onboarding, les marques ont la possibilité de mettre en place du People-based Marketing.

Ainsi, elles peuvent personnaliser l'expérience client et leur communication auprès de chaque individu, et ce, sur la globalité des parcours en cross-device et en omnicanal.

La technologie du Data Onboarding est liée aux DSP / Trading Desk / Partenaires Médias utilisés par les annonceurs afin d'activer l'audience digitalisée selon des problématiques de ciblage adaptées, dont les principales sont :

- **CRM Display** : l'activation des segments issus du CRM qui permet notamment de capitaliser sur la donnée magasin afin d'optimiser l'achat média digital et diffuser un message publicitaire adapté dans des logiques de fidélisation et/ou de réachat.
- **Réactivation des inactifs** : les bases PRM/CRM sont souvent constituées d'emails de prospects et de clients

qui n'interagissent plus avec la marque et ne réagissent plus aux emails commerciaux. Le Data Onboarding est une solution qui permettra de leur adresser un message dédié et plus engageant afin de les réactiver dans un cycle de vie.

- **Extension d'audiences / look-a-like** : dans le cadre d'une stratégie d'acquisition pure, le Data Onboarding permet d'élargir les audiences au look-a-like et d'activer des messages de targeting pertinents grâce à la modélisation de données comportementales.

- **Mise en repoussoir** : l'exclusion d'une audience peut être liée à :

> La volonté de gérer la pression marketing exercée sur une audience ;

> La rationalisation de l'achat media sur une cible ultra réactive et fidèle sur laquelle l'investissement peut être modéré.

B. PERSONNALISATION ON SITE

Le but est ici de recentrer la donnée sur l'individu pour mieux le connaître et personnaliser son expérience client ou prospect.

La personnalisation se matérialise par la mise en place de contenus dynamiques via des technologies d'AB testing.

La personnalisation on site peut aussi bien servir des problématiques d'UX (expérience client) que d'amélioration de la conversion (Conversion Rate Optimization).

C. MESURE DE LA PERFORMANCE DES CAMPAGNES DIGITALES SUR LES TRANSACTIONS EN MAGASIN PHYSIQUE

Il s'agit de l'étape la plus ultime qui consiste à enrichir l'analyse des campagnes et la connaissance client.

La réconciliation des données omnicanales et cross-device liées à un individu, permet d'obtenir une vision globale de son parcours d'achat, jusqu'à la conversion en magasin.

En clair, le Data Onboarding fournit un point de contact qui crée une passerelle entre deux univers jusque-là cloisonnés.

#3 MISE EN PLACE D'UNE SOLUTION DE DATA ONBOARDING

La base PRM/CRM d'une entreprise est uploadée sur une plateforme d'Onboarding et chacune des lignes est comparée (matchée) au référentiel de l'onboarder. Il s'agit d'un ID graph composé d'identifiants digitaux.

Cela se matérialise par un taux de match qui est le KPI principal. Il s'agit du pourcentage de profils communs entre la base PRM/CRM de la marque et l'ID graph de l'onboarder.

Il existe différentes méthodes de réconciliation des données, dont les principales sont la méthode déterministe et la méthode probabiliste.

Si l'on devait résumer ce processus technique, les 5 étapes suivantes seraient les plus représentatives :

1. Chargement : upload de la base de données contenant un maximum de champs d'identification (email, nom/prénom,

téléphone, adresse postale, code postal...). Il est impératif de segmenter la base de données en fonction des cas d'usage car elle sera le socle de la stratégie d'activation.

2. Anonymisation de la donnée via cryptage des informations et suppression des données personnelles.

3. Matching : réconciliation des données fournies par la marque avec l'ID graph de l'onboarder. Cette réconciliation interviendra idéalement sur deux clés d'identification. Exemple : email + nom/prénom ou nom/prénom + adresse postale. D'une manière générale, les clés d'identification comprenant l'email affichent les meilleurs taux de match.

4. Distribution : restitution des données digitalisées auprès des différentes plateformes utilisées par l'annonceur

5. Exécution de la campagne : activation des segments d'audience et personnalisation et/ou mesure de la performance des campagnes digitales sur la partie retail.

#4 CARTOGRAPHIE DES ACTEURS

A. MÉTHODOLOGIE ET CRITÈRES DE COMPARAISON

Il existe 2 types de solutions de Data Onboarding :

- **Direct Onboarding** : les marques intègrent directement leur base de données PRM/CRM sur Google, Facebook, ou tout autre environnement permettant de matcher les audiences.
- **Third Party Onboarding** : les marques peuvent également onboarder leur data via des prestataires (onboarders) qui collectent des informations à partir d'un réseau de fournisseurs de data (ex. : publishers).

Notre enquête concerne uniquement l'Onboarding Third Party. Liveramp, Temelio, Graphinium et Squadata y seront considérés comme les quatre prestataires principaux du Data Onboarding en France.

La méthodologie de ce benchmark permet d'identifier, comprendre et analyser leurs similitudes et différences sur des critères déterminants, notamment :

- **Les informations légales et juridiques liées aux sociétés** : date de création, siège social, nombre d'employés, types de clients... Quelle protection juridique de la donnée est fournie par l'onboarder ?
- **Technologie et hébergement de la data** : quelle est la taille minimum requise de la base PRM/CRM ? Quelles connexions ou synchronisations sont possibles avec les outils d'activation ? Quel est le niveau de précision et de fiabilité ?

- **Data Management** : quel est le taux de matching moyen selon les thématiques ? S'agit-il d'une approche déterministe, probabiliste ou couplée ? Quelle est la fréquence de mise à jour de la donnée ?

- **Gestion de la segmentation** : quels sont les types de segments proposés et à quelles fins ?

Autant de questions qu'il est donc essentiel de se poser pour une marque qui se lance dans une opération de Data Onboarding. En effet, sélectionner le partenaire le mieux adapté à son profil, à ses ressources, à ses contraintes et problématiques sera la clé du succès sur le long terme.

B. BENCHMARK DES SOLUTIONS

| | LIVERAMP | SQUADATA | TEMELIO | GRAPHINIUM |
|---|--|------------------------------------|------------------------------------|---|
| Année de création | 2009 | 2009 | 2011 | 2012 |
| Nombre d'employés | 501 à 1000 | 1 à 10 | 1 à 10 | 1 à 10 |
| CA | Non renseigné | 1 à 2 M de € | Non renseigné | Non renseigné |
| Type de clients | TPE, PME, ETI, Grandes entreprises, Administration | PME, ETI, Grandes entreprises | ETI, Grandes entreprises | PME, ETI, Grandes entreprises, Administration |
| Modèle économique | CPM en entrée / en sortie | CPM en sortie | CPM en entrée | CPM en sortie |
| Volume de la base de données clients (minimum) | 250 000 lignes | 100 000 lignes | 150 000 lignes | 100 000 lignes |
| Nombre de solutions d'activation synchronisées | 550 | 4 | Plus de 30 | Plus de 20 |
| Hébergement de la data | Europe | France | France | France |
| Taille de la base cookie pool France | 80 M | 70 M | 90 M | 60 M |
| Taille de la base cookie pool Monde | 350 M | 30 M | - | - |
| Types de données gérées | Offline, 1st, 2nd & 3rd party data | Offline, 1st, 2nd & 3rd party data | Offline, 1st, 2nd & 3rd party data | Offline, 1st, 2nd & 3rd party data |
| Niveaux de matching gérés | Déterministe / Household | Déterministe | Déterministe / Household | Déterministe |
| Taux de matching moyen (déterministe) | Entre 35% et 55% | 15% | Entre 25% et 50% | Entre 30% et 50% |
| Nombre moyen de cookie par individu | 3,8 | 2,5 | 2,5 | Entre 1,4 et 2 |

C. CRITÈRES DE SÉLECTION D'UNE SOLUTION DE DATA ONBOARDING

La qualification de ces paramètres permet de tirer des premiers enseignements sur l'offre et d'identifier les aspects essentiels et stratégiques que chaque marque devra évaluer pour mieux appréhender les valeurs ajoutées des acteurs consultés.

Avant toute chose, et étant donné le contexte actuel auquel les marques font face avec la mise en application du nouveau règlement sur la protection des données (RGPD), **le fait que la data soit hébergée en France ou en Europe** constitue un élément de différenciation important.

Les sources de collecte varient selon l'onboarder. Elles sont généralement directement en lien avec leur propre histoire, leur filiation ...

Le fait de comprendre **les méthodes de réconciliation des données** au sein de l'ID graph est la clef de voûte du Data Onboarding. Les onboarders appliquent principalement deux méthodes de réconciliation. Il s'agit des méthodes **déterministes** et **probabilistes** :

- **La méthode déterministe** semble clairement la plus fiable et celle à favoriser car la data est croisée avec des données individuelles anonymes sûres issues d'un espace loggué, d'une interaction avec une publicité ou même de l'utilisation d'une carte de fidélité.
- **L'approche probabiliste** utilise des données comportementales et anonymes liées à la navigation des utilisateurs. Elle peut, entre autres, venir en renfort d'une réconciliation déterministe trop faible. La marque aura alors intérêt à opter pour un onboarder capable de coupler une approche déterministe à une approche probabiliste pour assurer la pertinence de sa stratégie d'activation.

Le positionnement stratégique de l'onboarder à l'étranger constitue également un élément différenciant,

particulièrement pour les marques ayant une activité internationale. Deux types d'approches vont coexister :

- Soit l'onboarder possède, par sa filiation avec un groupe international, de la data propriétaire dans certains pays et bénéficie donc d'une expertise locale.
- Soit l'onboarder n'a pas cette spécificité et sous-traitera donc les opérations à des pairs reconnus dans leurs pays respectifs et propriétaires de leur Data.

La taille de l'onboarder aura un impact sur la façon dont la marque pourra opérer l'activation, et sur la réactivité attendue.

En effet, d'une certaine manière, un partenaire d'une moindre taille démontrera une agilité ou une capacité à initier des preuves de concept (POC) ou des développements ad hoc dans le but d'apporter à la marque un accompagnement sur-mesure.

L'aptitude à se connecter avec l'ensemble des technologies du marché est essentielle afin de pouvoir immédiatement activer les données ou segments après digitalisation. A ce jour, il est clair que la totalité des onboarders possèdent cette capacité à mettre en place des connecteurs sur demande.

La marque doit néanmoins s'assurer que le partenaire ne favorise pas un écosystème technologique qui ne sera pas adapté à sa stratégie d'activation.

Enfin, le business model pratiqué peut avoir une incidence importante sur le choix de l'onboarder.

Certains acteurs proposent une facturation au CPM « en entrée », c'est-à-dire selon le nombre de lignes uploadées, tandis que d'autres appliquent un coût « en sortie », soit fonction du nombre de profils matchés.

Certains peuvent même modéliser leur business selon le nombre de profils activés en média.

Les modèles économiques des prestataires de Data Onboarding semblent tendre vers des coûts liés au nombre de profils matchés voire activés afin de permettre aux marques une meilleure maîtrise de leurs investissements et optimiser leur activation selon leurs objectifs.

#5 CHALLENGES DU MARCHÉ

A. AMÉLIORATION DU CROSS-DEVICE

Reconnaître un même utilisateur, qu'il soit sur desktop ou sur mobile, permettra à la marque une activation des campagnes en cross-device bien plus pertinente.

Pour répondre à cet enjeu, les onboarders mettent en place des partenariats avec de grands éditeurs d'applications qui sont en mesure de réconcilier des ID mobile et des cookies web.

Depuis peu, des connexions avec des grandes plateformes telles que Facebook, Twitter ou Spotify sont possibles chez certains et permettent d'obtenir de la donnée cross-device.

B. AMÉLIORATION DU TAUX DE MATCHING

Un taux de matching correct est généralement compris entre **40% et 60%** quand la clé est l'email, et est **supérieur à 80%** avec l'adresse postale, selon le Forrester report "The Forrester Wave™ : Enterprise Marketing Software Suites", Q2 2016. Dans un secteur où le reach règne, quels sont les moyens déployés par les onboarders pour accroître leur couverture et donc aller chercher des pourcentages de match supplémentaires ?

Entre le nombre de personnes qui ne se connectent jamais sur Internet et les adblocker qui compliquent l'exploitation des cookies, l'amélioration du taux de match semble être le premier challenge.

Les taux de match sont couramment plus élevés pour les marques dont l'activité est majoritairement digitale (entre 50 et 70%) et moins important pour les marques ayant un chiffre d'affaires essentiellement physique (entre 25 et 40%).

Ces taux de match peuvent être considérés comme satisfaisants s'il s'agit de mesurer le pouvoir contributeur de l'activité online sur l'offline mais cela devient plus problématique pour des usages tels que la personnalisation on site ou de l'achat média étant donné qu'une grande part de la donnée ne peut être réconciliée.

Par ailleurs, si 85% de la population française utilise Internet, certaines marques rencontreront naturellement plus de difficultés à digitaliser leurs données en fonction des profils qui constituent leur base.

A titre d'exemple, les seniors représentent une typologie de profils qui affiche les taux de match les plus faibles du fait de leur présence moins soutenue sur le web.

C. SYNERGIES ENTRE LES PARTENAIRES

Les solutions d'Onboarding sont certes concurrentes, mais aussi très souvent complémentaires selon les cas d'usage. Dépendantes de cette logique où le reach est maître, les synergies entre onboarders complémentaires permettent d'avoir une couverture supérieure visant à améliorer le taux de match.

La marque devra de son côté s'assurer que le taux de recoupement (ou chevauchement) entre les référentiels de chacun des onboarders soit le plus faible possible et que

l'uplift soit suffisamment significatif. Le taux de recouplement serait dans la plupart des cas assez faible étant donné la diversité des sources de données de chaque onboarder.

La marque pourra tirer parti de cette complémentarité notamment dans le but de retrouver un maximum de points de contact pour un individu et ainsi optimiser la complétude du profil.

Par exemple, une marque faisant appel à deux onboarders, pourra retrouver un client sur le navigateur Chrome d'un desktop via le premier, et ce même client sur le navigateur Safari d'un iPhone via le second.

D. UNE VÉRITABLE ALTERNATIVE OU COMPLÉMENTARITÉ AUX GAFA

Les marques semblent rencontrer certaines difficultés à confier leurs données les plus précieuses et sensibles à Google ou Facebook. Malgré l'existence des conditions générales d'utilisation des données onboardées, elles sont considérées comme des plateformes « floues » en matière de traitement, de stockage et d'utilisation des données.

Les onboarders third party offrent un cadre contractuel clair et transparent ainsi qu'un partenariat sur-mesure constituant un élément de réassurance indéniable pour les marques.

Par ailleurs, la synergie entre les onboarders third party met à disposition des marques une puissante offre sur le marché complémentaire aux GAFA.