

LE PENSE PAS BÊTE DU



LE GUIDE DE LA BRAND SAFETY ET DE LA BRAND SUITABILITY



A l'origine de la publicité digitale, **le mode de commercialisation reposait uniquement sur le gré à gré** : l'annonceur et son agence choisissaient les médias et les rubriques sur lesquels diffuser les campagnes en fonction, notamment, de l'adéquation des contenus des sites web avec les valeurs de la marque, les objectifs de la campagne et le profil du public-cible. Le choix d'un univers de diffusion d'une campagne et son contrôle était au cœur du media planning.

Avec la croissance exponentielle du nombre d'éditeurs, de contenus et donc de volume d'inventaire, de nouveaux modes de transaction sont apparus, en parallèle de la commercialisation gré à gré, pour faciliter l'achat et la vente de ces emplacements et optimiser le ROI des investissements publicitaires. Les ad-networks, dans un premier temps, ont proposé des inventaires

agrégés entre plusieurs sites aux univers proches où media planning et ciblage contextuel restaient de mise. Puis, le programmatique, à partir des années 2010, a permis d'automatiser l'achat et la vente d'une infinité d'espaces publicitaires sur la base de nouveaux critères, issus de la connaissance des audiences. **Le media planning se complète alors de l'audience planning**. En ciblant non plus des contextes de diffusion mais des catégories d'individus en affinité avec les objectifs de marques, la diffusion des campagnes ne se fait plus sur quelques dizaines de sites mais sur des milliers de sites.

Face à la dispersion de leurs campagnes digitales, les annonceurs et leurs agences ont en effet besoin de s'assurer que les contextes de diffusion sont appropriés pour leurs marques. On parle ainsi de **brand safety** et les technologies qui y sont associées permettent

d'analyser le contenu et de cibler ou d'exclure des URL en fonction de mots-clés. Ce sont l'IAB en 2012, puis le GARM, qui ont élaboré des standards pour catégoriser et filtrer les contenus jugés universellement indésirables.

Avec la forte croissance du programmatique, et de l'open auction en particulier, l'adoption de ces filtres a pris de l'ampleur, en particulier à la suite de plusieurs scandales impliquant de grandes plateformes sur lesquelles des marques avaient été placées en face de contenus illicites de type terroristes ou pédopornographiques. Plus récemment, les questions liées aux fake news et à la haine en ligne ont renforcé la vigilance des annonceurs et des agences quant à la possibilité de voir les marques adossées à un contenu inapproprié.

Il existe différentes manières d'opérer ces filtres qu'il s'agisse de noms de domaines, de catégories thématiques ou de listes de mots-clés. Ces fonctionnalités sont proposées par la plupart des outils d'adverification et / ou directement intégrées dans les outils

d'achat et vente programmatiques (DSP et SSP). A cela, se sont ajoutées de nouvelles fonctionnalités liées à la notion de **brand suitability**. En effet, au-delà des catégories standards de **brand safety**, les annonceurs peuvent considérer en plus des critères qui leur sont propres, notamment au travers de block lists personnalisées de mots-clés, c'est le sujet de la **brand suitability**.

L'utilisation des techniques de **brand safety** et **brand suitability** peuvent aussi avoir des effets contre-productifs, réduisant considérablement la disponibilité des inventaires publicitaires et entraînant des problèmes de diffusion et de couverture des campagnes ainsi qu'un manque à gagner important pour les éditeurs.

Le fonctionnement des outils et méthodologies liés à la **brand safety** et à la **brand suitability** est complexe et implique de nombreuses parties prenantes. Ce pense-pas-bête pratique a pour ambition d'en expliquer les grands principes.

DÉFINITIONS

■ AD VERIFICATION :

Outils et services permettant de contrôler, avant ou après diffusion, le contexte d'affichage publicitaire. Il peut s'agir de mesure de la visibilité, de la fraude ou encore de la **brand safety** et de la **brand suitability**.

■ BRAND SAFETY :

Pratiques et outils mis en place pour garantir la diffusion de la campagne de publicité en ligne dans un environnement « propre », et donc qu'elle ne soit pas diffusée aux côtés de contenus jugés inappropriés à l'aune de catégories communément définies par la profession. Ces règles d'exclusion sont appliquées à chaque campagne afin de permettre aux annonceurs de protéger la réputation de leur marque et d'assécher les revenus de créateurs de contenus dangereux ou illégaux. En fonction de chacune de ces catégories, les marques peuvent définir leur seuil de tolérance face au risque : faible, moyen ou élevé.

Certains sites proposant contenus et services parmi les plus reconnus peuvent être considérés comme **brand safe by design**. Les médias et services répondant aux règles déontologiques communes à la profession, se portant garant des contenus qu'ils diffusent, sont par définition brand safe. A titre d'exemple, un article traitant du terrorisme, publié sur un site d'actualité renommé, ne fait en aucun cas l'apologie d'actes illégaux. Pour s'assurer de ne pas écarter leur inventaire après utilisation de listes de mots-clés, il est possible de les whitelister automatiquement via les outils des mesureurs ou les plateformes d'achat programmatique.

GARM : BRAND SAFETY FLOOR + SUITABILITY FRAMEWORK

CATEGORY (Video / Audio / Text)	BRAND SAFETY « FLOOR »
Adult & Explicit Sexual Content	<ul style="list-style-type: none"> • Illegal sale, distribution, and consumption of chil pornography • Explicit or gratuitous depiction of sexual acts, and/or display of genitals, real or animated
Arms & Ammunition	<ul style="list-style-type: none"> • Promotion and advocacy of Sales of illegal arms, rifles, and handguns • Instructive content on how to obtain, make, distribute, or use illegal arms • Glamorization of illegal arms for the purpose of harm to others • Use of illegal arms in unregulated environments
Crime & Harmful acts to individuals and Society and Human Right Violations	<ul style="list-style-type: none"> • Graphic promotion, advocacy, and depiction of willful harm and actual unlawful criminal activity - murder, manslaughter & harm to others. Explicit violations/demeaning offenses of Human Rights (eg, trafficking, slavery, etc.)
Death of Injury	<ul style="list-style-type: none"> • Promotion or advocacy of Death or Injury • Murder or Willful bodily harm to others • Graphic depictions of willful harm to others
Online piracy	<ul style="list-style-type: none"> • Pirating, Copyright infringement, & Counterfeiting
Hate speech & acts of aggression	<ul style="list-style-type: none"> • Unlawful acts of aggression based on race, nationality, ethnicity, religious affiliation, gender, or sexual image or preference • Behavior or commentary that incites such hateful acts, including bullying
Military Conflict	<ul style="list-style-type: none"> • Incendiary content provoking, enticing, or evoking military aggression • Live action footage/photo of military actions & genocide or other war crimes
Obscenity and Profanity	<ul style="list-style-type: none"> • Excessive use of profane language or gestures and other repulsive actions with the intent to shock, offend, or insult
Illegal Drugs	<ul style="list-style-type: none"> • Promotion or sale of illegal drug use - including abuse of prescription drugs. Federal jurisdiction applies, but allowable where legal local jurisdiction can be effectively managed
Spam or Harmful Content	<ul style="list-style-type: none"> • Malware/Phishing
Terrorism	<ul style="list-style-type: none"> • Promotion and advocacy of graphic terrorist activity involving defamation, physical and/or emotional harm of individuals, communities, and society
Tobacco / e-cigarettes / Vaping	<ul style="list-style-type: none"> • Promotion and advocacy of Tobacco and e-cigarette (Vaping) & Alcohol use to minors
Sensitive Social Issue / Violations of Human Rights	<ul style="list-style-type: none"> • Disrespectful and harmful treatment of sensitive social topics (eg, abortion, extreme political positions, etc) • Acts, language, and gestures deemed illegal, not otherwise outlined in this framework (eg, harm to self or other and animal cruelty) • Targeted harassment of individuals and groups

■ BRAND SUITABILITY :

En utilisant les outils et méthodologies de **brand safety**, certains annonceurs excluent ou privilégient des environnements de diffusion en fonction de critères propres à leur entreprise, à une marque ou à une campagne. La **brand suitability** est donc une version brand centric du concept de **brand safety**, une customisation de l'univers de diffusion d'une campagne au cas par cas.

Les catégories et les niveaux de risques sont définis dans le **Brand Safety Floor Framework**, initialement rédigé par les 4A et l'IAB et soutenu par la Global Alliance for Responsible Media (GARM).

Pour résumer, il existe trois stratégies différentes pour répondre à trois objectifs différents :

- **LA BRAND SAFETY** : Garantir aux campagnes de publicité un environnement de diffusion sûr, selon 13 catégories définies par l'interprofession (cf tableau page 3).
- **LA BRAND SUITABILITY** : Garantir aux campagnes de publicité un environnement de diffusion sûr, en fonction de critères propres à une marque.
- **LE CIBLAGE** : Sélection des médias, des catégories thématiques, des contextes, données socio-démographiques, comportementales, géographiques, etc.

CHAPITRE 1 : LES PARTIES PRENANTES DE LA BRAND SAFETY

La **brand safety** et la **brand suitability** concernent l'ensemble de la chaîne de valeur publicitaire. Chacun, à son échelle, peut contribuer à lutter contre les risques

liés aux contextes inappropriés qu'il s'agisse de garantir un environnement web fiable ou d'éviter les effets délétères d'un filtrage poussé à outrance.



**ANNONCEURS
&
AGENCES MEDIAS**



**ADTECH
Ad verification
DSP & SSP
Fournisseurs data
Adservers**



**EDITEURS
&
RÉGIES**



INTERNAUTES

Points de vigilance

- Longueur & pertinence des listes de mots clés
 - Impact sur le reach du fait de l'éviction de catégories jugées négatives
 - Potentielles discriminations d'audiences en lien avec les contenus lus
- Superposition de segments de ciblage
- Pour les éditeurs / plateformes non identifiés « brand safe by design », ne pas monétiser des contenus inadéquats (infox ou UGC non approprié)
- Potentielle exclusion de cibles publicitaires du fait d'un filtrage discriminant

- **LES ANNONCEURS**, via leurs agences, cherchent à assurer au message publicitaire et à la marque un environnement de qualité pour engager l'audience des éditeurs qu'ils ciblent. Pour y parvenir, ils utilisent des technologies publicitaires et/ou communiquent directement avec les régies des éditeurs.
- **LES ADTECH / TECHNOLOGIES PUBLICITAIRES** proposent des outils de **brand safety** et

des segments de données de ciblage aux acheteurs.

- **LES ÉDITEURS ET LEURS RÉGIES** produisent des contenus fiables dans des contextes maîtrisés à destination de leur audience.
- **LES INTERNAUTES** souhaitent accéder à des contenus informatifs ou divertissants et des services de confiance dans des environnements sûrs.

CHAPITRE 2 : PRATIQUES ET OUTILS

COMMENT FONCTIONNENT LES OUTILS DE MESURES TIERCES ?

De nombreux acteurs spécialisés dans l'*ad verification* proposent leurs services en France. Ces outils reposent sur les mêmes principes méthodologiques :

- Les acheteurs paramètrent l'outil en définissant leur niveau de tolérance (faible, modéré, élevé) en matière de **brand safety** ainsi que des catégories custom et/ou des listes d'inclusion ou d'exclusion le cas échéant. Ces listes peuvent concerner des noms de domaines, des zones géographiques, des marques ou des mots-clés.

*Tous les mesureurs ne proposent pas de catégories customisées (**brand suitability**).*

- Un *crawler* (robot) scanne les éléments d'une page web pour détecter, en temps réel, des environnements *unsafe* (inappropriés) ou *unsuitable* (inadaptés) en fonction des catégories standardisées par l'interprofession, de catégories *custom*, de listes d'inclusion ou d'exclusion de mots-clés. En parallèle, un tag ou script est intégré à la création publicitaire pour *crawler* le contexte de diffusion.

En fonction des mesureurs et des cas, l'analyse peut être réalisée seulement sur le domaine et les sous-domaines, sur l'URL, voire aller jusqu'à l'ensemble des contenus d'une page (textes, images, commentaires, liens sortants et entrants, métadonnées).

- Les données collectées (mots, phrases et expressions régulières) sont confrontées aux listes d'inclusion ou d'exclusion (domaines ou mots-clés) utilisées par

la marque ou analysées via la *machine learning* pour les faire correspondre à des modèles (catégories). Ces derniers permettent, *in fine*, de classer les pages en fonction de leur potentiel de risques (faible, modéré, élevé), en temps réel.

Depuis quelques années, l'analyse sémantique vient enrichir l'offre de certains mesureurs afin d'ajouter une interprétation de la tonalité du contenu (positif, neutre, négatif) et éviter les erreurs d'interprétation.

- Les résultats de ces analyses sont hébergés dans des bases de données directement branchées aux adservers des régies, pour les achats en gré à gré, et aux DSP, pour les achats programmatiques.

Outre ce socle méthodologique commun, quelques particularités :

- **LES ENVIRONNEMENTS** : si tous les acteurs opérant en France mesurent le web et le web mobile, seuls certains peuvent analyser les contenus in-app ou ceux diffusés par les *walled gardens*. L'univers in-app, en effet, est plus complexe à mesurer, puisque l'accord de l'éditeur est obligatoire pour intégrer le SDK de contrôle à l'application. Par ailleurs, les mesureurs ne peuvent pas proposer les mêmes fonctionnalités in-app que sur le web. A titre d'exemple, il est

possible de détecter la présence d'UGC mais pas de l'analyser. Les *walled gardens*, quant à eux, ne signent des accords qu'avec quelques acteurs de la **brand safety** en fonction de leurs références clients.

- **LES MÉDIAS OU TYPOLOGIES DE CONTENUS** : si le texte est universellement analysé, il est plus difficile de traiter l'image, et surtout la vidéo et l'audio. En cause ? Le volume de données à traiter. Il faut ajouter à cela que les éditeurs utilisent une grande diversité de *players*. Pour traiter le média vidéo, les mesureurs se concentrent sur l'analyse d'éléments annexes aux images tels que les méta descriptions, les commentaires, les liens entrants et sortants voire, pour certains, les sous-titres.

- **LES PRODUCTEURS DE CONTENUS OU FOURNISSEURS DE SERVICES** : la majorité des éditeurs de contenus et de services acceptent les tags des mesureurs tiers. En revanche, nous l'avons vu plus haut, il est plus complexe de travailler avec les *walled gardens*. Soit parce qu'ils n'acceptent pas de collaborer avec tous les acteurs, soit parce que leur fonctionnement ne permet pas d'analyser facilement les environnements publicitaires. C'est le cas, notamment, des réseaux sociaux pour lesquels un flux de contenu est propre à son utilisateur.

Pour en savoir plus sur les spécificités des offres des mesureurs présents sur le marché français, nous vous invitons à consulter [les grilles d'analyses réalisées par le CESP](#).

DES ACTIONS TOUT AU LONG DE LA CHAÎNE DE VALEUR

La mesure s'applique différemment selon les technologies, les méthodologies et les modes de transaction utilisés.

MODES DE COMMERCIALISATION : EN PROGRAMMATIQUE

- Les technologies de mesure sont intégrées aux principaux DSP et SSP du marché. Chez une partie des mesureurs, les catégories de ciblage liées à la **brand safety** sont disponibles en *pré-bid*. Pour d'autres la mesure a lieu après que l'enchère est passée. Dans ce dernier cas, la mesure étant plus rapide que le chargement de la bannière, les utilisateurs n'ont pas le temps de voir la publicité.
- Faute de pouvoir faire passer les *crawlers* sur toutes les pages web en continu, les acteurs *pré-bid* opèrent également un contrôle *post-bid*. Dans le cas où une page est signalée comme inappropriée en *post-bid*, la bannière est bloquée avant d'avoir atteint *adserver*. Si elle est tout de même diffusée, elle peut être masquée par un *layer*. Dans les deux cas, certaines agences demandent a posteriori des compensations aux éditeurs et leurs régies.

EN GRÉ À GRÉ

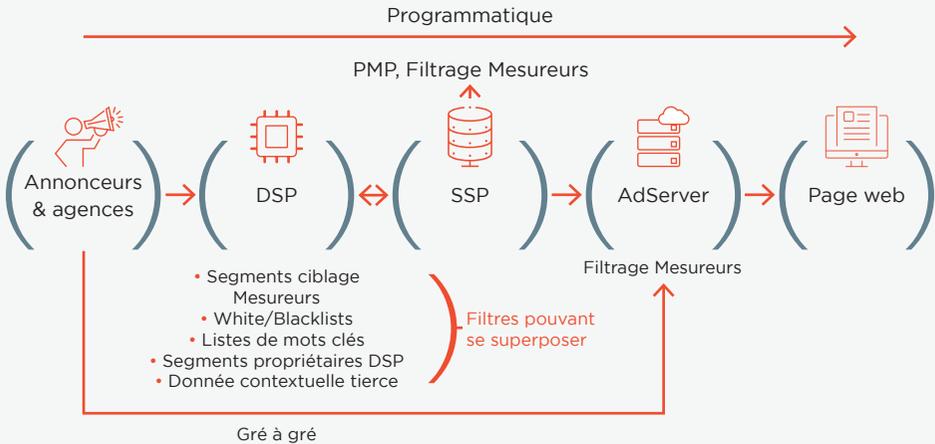
- Les impressions inappropriées sont filtrées avant qu'elles n'atteignent *adserver*. Elles ne sont donc pas diffusées.

LES BLOCK LISTS, CATÉGORIES THÉMATIQUES ET DONNÉES TIERCES :

- Les listes d'exclusion ou d'inclusion (*whitelists*, *blacklists* et listes de mots-clés) peuvent être confiées aux mesureurs ou directement opérées dans la DSP, le SSP ou la régie.
- Des catégories thématiques spécifiques à un DSP ou des données contextuelles tierces peuvent également être activées par l'acheteur ou par la régie .

De plus, les transformations réglementaires et techniques en cours (RGPD et disparition progressive des cookies tiers) incitent éditeurs et régies à travailler leurs *assets*. La tendance est donc d'approfondir la segmentation des pages web pour proposer des offres de ciblage contextuel « nouvelle génération ». L'évolution des technologies d'IA basées sur l'analyse sémantique et le langage naturel permet, en effet, une compréhension plus poussée des contenus par les machines, et donc une segmentation contextuelle plus fine. Ces segments contribuent également à faciliter la visibilité des annonceurs sur la *safety/suitability* de leurs inventaires, en atténuant les risques de « faux positifs ».

BRAND SAFETY ET BRAND SUITABILITY DANS LA CHAÎNE DE MONÉTISATION



CHAPITRE 3 : BONNES PRATIQUES ET RECOMMANDATIONS

L'image de marque est importante, beaucoup d'énergie, de temps, d'argent y sont consacrés. Diffuser les messages dans des environnements adéquats est donc capital,

d'autant que cela permet aussi d'atteindre les consommateurs les plus en affinité avec la marque.

Nous avons vu plus haut que l'utilisation d'outils de brand safety proposés par les mesureurs indépendants est désormais répandue. Mais il ne faut pas oublier que la technologie peut comporter des biais aux effets contre-productifs. S'il est légitime pour un annonceur de chercher à protéger l'intégrité de sa marque, d'éviter de cautionner ou de financer indirectement des contenus qui vont à l'encontre de ses valeurs voire des organisations illicites, il faut toutefois, prendre des précautions pour éviter les « faux positifs » liés pourtant à des contenus professionnels et neutres (ex : termes polysémiques comme « coup de poing » et « tout à coup »).

Quant à la brand suitability, il est compréhensible que certains contextes ne conviennent pas à certaines marques (par exemple : une bannière pour un produit très sucré placé en face d'un article sur l'obésité).

Il est également crucial de distinguer stratégie de brand safety/suitability et stratégie de ciblage. Il est courant que les block lists contiennent un trop grand volume de mots-clés pour écarter des contextes. Or, ce filtrage devrait plutôt être opéré à travers une stratégie de ciblage (sélection des médias, des catégories thématiques, data, etc.) plutôt que d'exclure a priori des contenus à potentiel.

Pour palier les effets nuisibles d'une **brand safety/suitability** inadaptée ou poussée à l'extrême, et bien que chaque marque ait des particularités liées à ses campagnes,

ses valeurs et objectifs, **voici 3 conseils universels pour bâtir une stratégie pertinente et efficace et quelques bonnes pratiques à respecter :**

1. ASSURER LE SOCLE DE LA PROTECTION DES MARQUES

- Veiller à éviter les 13 catégories décrites comme inappropriées à la monétisation publicitaire.
- Choisir le mode d'achat qui convient le mieux à la prise de risque acceptée par l'annonceur.
- Choisir les *devices* et environnements sur lesquels communiquer. Les outils de mesure sont complètement opérationnels sur le display desktop et web mobile. En revanche, tous ne mesurent pas la **brand safety** sur Appli et, lorsqu'ils le font, n'utilisent pas la même méthodologie que pour le web.

La **brand safety** en vidéo et sur les médias sociaux restent encore des défis.

- Au sein de ces catégories, l'annonceur/ agence doivent délimiter un niveau de risque accepté. Une certaine tolérance peut être acceptée en fonction des marques (ex : sexe // marques de préservatif) et en fonction des campagnes (produit d'une même marque destiné à des enfants // produit destiné à des adultes)
- Travailler avec des éditeurs **brand safe by design** qui se portent garants de leurs contenus et services.

2. LA BRAND SUITABILITY, AVEC SUBTILITÉ

- Se pencher sur les valeurs de marque pour établir une stratégie de **brand suitability** pertinente et long-terme.
- Certains contextes peuvent ne pas convenir à une marque en particulier. Pour ne pas manquer un inventaire de qualité, même s'il est lié à un sujet a priori *unsuitable*, certaines solutions basées sur la sémantique et les segments contextuels peuvent être plus pertinentes que les listes de mots-clés. Elles peuvent apporter plus de granularité dans l'exclusion de certains sujets et éviter les faux positifs créés par

l'utilisation de liste de mots-clés. (ex : voile religieux, la voile sport ou encore la Covid avec le nombre de morts ou la Covid et les applaudissements à 20h00).

- Se pencher sur le niveau de risque accepté.
- Dans le cas de listes *custom*, les réviser très régulièrement et limiter le nombre de mots-clés présents.
- Ne pas multiplier les couches de filtres au risque de voir baisser drastiquement les inventaires disponibles.

3. ENTREtenir UNE COLLABORATION ENTRE PARTENAIRES TECHNOLOGIQUES, RÉGIES ET AGENCES

- Conseils sur la **brand suitability** : demander conseils aux partenaires technologiques pour optimiser les listes de mots-clés, si elles sont utilisées.
- Se former régulièrement aux outils qui évoluent rapidement.
- Possibilité pour les agences de communiquer

aux régies les paramètres de *brand safety* de l'annonceur.

- Considérer que les sites premium sont **brand safe by design**, sur la majorité de leurs catégories thématiques.
- *Whitelister* les sites en fonction de l'affinité marques / annonceurs. *Blacklister* les sites

à proscrire des plans toujours en fonction de l'adéquation aux valeurs de marque annonceurs. Dans les 2 cas, les listes doivent être régulièrement révisées.

- Développer une culture média, rencontrer les éditeurs pour connaître leur ligne éditoriale et leurs offres.
- Développer une relation de proximité avec les éditeurs à privilégier. Ils sont à

même de proposer des offres alignées sur les objectifs et exigences agences / annonceurs.

- Assurer une meilleure communication entre les annonceurs, agences et régies autour des paramètres liés à la **brand safety** en amont des campagnes, pour éviter les diffusions en contexte inapproprié et potentiellement les demandes de compensation.

POUR UTILISER LES LISTES DE MOTS-CLÉS, QUELQUES BONNES PRATIQUES

- Utiliser une liste concise de mots clés, adapter les listes par marques/produits. Les restreindre à des termes très spécifiques, sans ambiguïté. D'autant plus qu'un annonceur peut avoir plusieurs listes (génériques et/ou par produits) et plusieurs DSP qui ont chacun leurs propres paramétrages.
- Si certains mots peuvent avoir un double sens, leur adjoindre un ou d'autre(s) mot(s) pour préciser. Ex : « coup » et « poing » pour ne pas être confondu avec « Tout à coup » / « Tout d'un coup ».
- Dans la même idée, ne pas scinder les expressions ou locutions (ex : « gilet jaune » et non « gilet » et « jaune »). Il vaut mieux choisir des expressions qui ont une forte probabilité d'apparaître dans les médias sélectionnés plutôt que de vouloir parer à toutes les éventualités.
- Les *block lists* sont majoritairement chargées dans les DSP et sont prioritaires puisqu'elles filtrent les *bid requests*. Si un outil d'adverification est ajouté : bien s'assurer auprès du prestataire qu'il n'y a pas de doublons / interférences entre les 13 catégories de **brand safety** et les filtres du DSP.
- Réviser régulièrement, a minima 2 fois par an, les listes de mots clés et les *black lists* de domaines. Les ajuster en conséquence.
- Analyser l'impact des listes sur la diffusion de vos campagnes et leurs coûts à l'aune de la qualité media recherchée. Une protection excessive peut entraîner une exclusion d'inventaire sûr et efficace, une portée quasi nulle des campagnes et une augmentation des CPM.