

Réponses aux questions de la commission « Les Lumières à l'ère du numérique »

Table des matières

Propos liminaire.....	2
Pouvez-vous s'il vous plait nous expliquer les initiatives prises par les agences media françaises pour assécher le financement de sites conspirationnistes et alimentant la haine en ligne ? Quelle en est la portée et quelles en sont les limites ?.....	2
Limiter le risque au maximum avec une démarche de vérification a priori et de correction a posteriori.....	3
Adverification.....	3
Périmètres pertinents.....	4
Search.....	4
Social.....	4
Display.....	5
Des mesures sont prises au quotidien par les agences pour limiter les risques.....	6
Y a-t-il des échanges et discussion à ce sujet au sein de l'UDECAM, et un consensus se dégage-t-il ? Si oui, lequel ? Dans quelle mesure les pratiques différent-elles ?.....	7
Vous avez récemment affirmé dans la presse que l'assèchement des revenus publicitaires pour les sites de fake news était votre combat de la décennie. Qu'allez-vous faire ? Sur quoi faut-il agir selon vous, à la fois au sein de Publicis et de l'UDECAM ?.....	8
La Commission a pour but de proposer des solutions concrètes pour réduire le financement de la désinformation via la publicité programmatique. Quelles sont les propositions qui vous semblent devoir être mises prioritairement en œuvre au sein des agences médias ? Y a-t-il d'autres actions qui vous paraissent essentielles, de la part d'autres acteurs ?.....	8
Les entreprises offrant des listes d'exclusion de sites web conspirationnistes et incitant à la haine existent. Lesquelles connaissez-vous ? Ne pensez-vous pas que les agences de médias françaises pourraient travailler directement avec ces acteurs (l'un ou l'autre), en intégrant leurs listes par défaut ?.....	9
Quels sont d'après vous les acteurs (personne ou institution) que la commission devrait consulter sur le sujet qui nous préoccupe ?.....	10
Annexe 1 - Actions mises en place chez PMF pour la Brand Safety.....	11
Définition Brand Safety.....	11
Typologie des contenus/contextes et part dans le volume de pages publiées.....	11
Les mesures et actions mises en place au sein de Publicis Media.....	11
Intégration de filtre de contrôle en natif dans les outils programmatiques.....	12
Intégration d'outils tiers permettant une vérification des contenus en amont de l'impression.....	12
Paramétrage des campagnes dans les process d'achat.....	12
Pour aller plus loin, signature d'un partenariat avec Newsguard.....	12
Publicis est membre actif du Global Alliance for Responsible Media (GARM).....	13

Au quotidien, les actions mises en place et à suivre dans le temps	13
Audit des comptes	13
De l'exclusion vers l'inclusion	14
Annexe 2 - Integral Ad Science (IAS)	14

Propos liminaire

Cette note répond aux questions posées par la commission « Les Lumières à l'ère du numérique ». Certains exemples sont relatifs à des pratiques existantes chez Publicis. Le cas échéant cette situation est explicitement précisée.

D'autres points concernent l'ensemble des agences membre de l'Udecam qui est le syndicat patronal des agences media en France.

Pouvez-vous s'il vous plait nous expliquer les initiatives prises par les agences media françaises pour assécher le financement de sites conspirationnistes et alimentant la haine en ligne ? Quelle en est la portée et quelles en sont les limites ?

La notion de « fake news » ou « d'infox » est complexe et fortement polysémique. La définition même de ce qu'est une infox n'est pas évidente. Bien souvent les infox mélangent informations « vraies » et « fausses » rendant leur détection très ardue par des moyens automatisés (et parfois même par des lecteurs humains). De plus, certains titres de presse qui sont régulièrement cités comme diffusant des infox ont également l'agrément IPG.

Les agences media sont fortement engagées pour permettre l'expression de la diversité des opinions dès lors, il ne leur appartient pas de qualifier ce qui est une infox ou ce qui n'en est pas. Par conséquent, les agences font appel à des sociétés tierces qui fournissent des listes noires ou des listes blanches de sites sur lesquels il est possible de faire (ou pas) de la publicité. Cela permet d'éviter tout site qui pourrait ne pas être « brand safe », tels que les sites de fake news ou incitant à la haine. En plus de ces approches par liste, des sociétés dites « d'adverification » fournissent des outils de contrôle a priori (fondée sur l'historique) et de mesure a posteriori qui permettent une amélioration continue de ces listes noires.

Pendant, certains sites d'infox ont une démarche proactive pour échapper aux mesureurs : changements de nom de domaine, modification des IP, etc. ce qui impose aux mesureurs un contrôle systématique de tous les sites et applications monétisant leurs inventaires. Par rapport à d'autres types de sites illégaux comme ceux de streaming illégal il est très difficile d'identifier un site de fake news : rien ne ressemble plus à un site de fake news que le site d'un titre de presse légitime. Du côté des agences et des annonceurs cela signifie une mise à jour permanente de nos listes d'exclusion avec cette limite qu'une tolérance un peu trop grande peut laisser passer quelques sites complotistes ou incitant à la haine au travers du filet et qu'à l'inverse une sévérité trop grande risque de remettre en question la survie de sites d'information de plus petite taille.

Limiter le risque au maximum avec une démarche de vérification a priori et de correction a posteriori

Les agences media ont pris ce sujet très au sérieux dès l'apparition de l'achat programmatique (2011 sur le marché français). Pour cette raison, 99% des campagnes sont mesurées avec un outil d'adverification.

Le fonctionnement des outils d'adverification sur la brand safety consiste à « scanner » les contenus de chaque site qui met en vente son inventaire publicitaire sur les places de marché. Cependant, cette opération ne se fait pas en temps réel. En effet dans l'achat programmatique, quand une enchère est proposée (bid request) il ne se passe en moyenne que 120-150 millisecondes avant que l'acheteur ne soumette une enchère (bid response). Le gagnant de l'enchère diffuse ensuite sa publicité. Ainsi pour chaque nouveau site ou application qu'un mesureur voit passer pour la première fois, il va laisser passer la publicité mais un processus d'audit pour classer le site va être lancé. Ce processus peut prendre quelques heures ou jours en fonction de la quantité de contenu à analyser.

Une fois le site catégorisé, s'il ne respecte pas les critères de brand safety, il va intégrer la liste noire et lorsqu'à nouveau une enchère sera soumise, l'acheteur sera alerté et aucune enchère ne sera proposée. On estime que chaque jour en France ce sont entre 15 et 20 milliards d'enchères qui sont soumises sur les places de marché. Pour les sites ayant un gros volume de contenu, les mesureurs vont scanner l'ensemble des pages 2 à 3 fois par jour et pour les plus petits quelques fois par semaine. Ainsi, si entre deux contrôles un contenu préjudiciable est publié il ne sera pas détecté tout de suite ce qui peut mener à faire apparaître des publicités sur des sites complotistes ou de fake news mais pour une durée très limitée.

Les sites éditoriaux producteurs de contenus ont en général une ligne éditoriale clairement établie et si un site ou une application a été catégorisé comme non « brand safe », toute nouvelle publication sera considérée comme non « brand safe ». A titre d'exemple, lefigaro.fr est considéré comme brand safe par les mesureurs a priori. En revanche le site francesoir.fr, sur lequel des articles de désinformation ont déjà été publiés, est considéré a priori comme non conforme dans son ensemble, quand bien même un nouvel article publié porterait sur une information réelle et vérifiée.

En complément, il convient également de rappeler les définitions de l'adverification, qui inclut notamment les sites de désinformation ou complotistes, celle des périmètres pertinents concernés par ces contenus ainsi que le modèle d'affaires ou de recherche d'influence des sites d'infox et les flux financiers qui les rémunèrent. Cette double définition permettra de mieux délimiter le périmètre sur lequel œuvrent les agences media dans la publicité numérique et donc les responsabilités qui leur incombent.

Adverification

Il s'agit de faire mesurer par un tiers de confiance, la qualité globale de la diffusion d'une annonce publicitaire digitale. Cette mesure est limitée aux publicités de type « display » ou « vidéo ». En particulier cette vérification vise à éviter qu'une annonce ne se trouve à proximité de contenus qui soient incompatibles avec la marque tels que ceux qui propagent la haine ou des infox.

Il y a trois dimensions qui sont contrôlées :

- **La visibilité** : il s'agit de la durée d'exposition de la publicité en fonction du pourcentage de sa surface visible à l'écran.
- **La fraude** : il s'agit de s'assurer que l'affichage de la publicité et le clic sur la publicité sont bien réalisés par des humains et non des robots.
- **La « brand safety » et par extension la « brand suitability »** : il s'agit de s'assurer que le contexte de diffusion sur une page ou un newsfeed soit conforme aux valeurs de la marque et ne contienne pas de contenus illicites (drogue, alcool, violence, haine, désinformation, streaming illégal, contenus pour adultes, etc.).

Périmètres pertinents

Dans le domaine de la publicité en ligne il y a trois marchés pertinents tels que nous avons l'habitude de les définir notamment pour les autorités de concurrence. Il s'agit des liens sponsorisés sur les moteurs de recherche (le « search »), les réseaux sociaux (le « social ») et les annonces sous forme de bannière ou de vidéo (le « display »).

Search

Les moteurs de recherches lisent, indexent et « historisent » les millions de contenus publiés chaque jour sur internet. Lorsque l'utilisateur effectue une recherche, le moteur de recherche affiche les résultats les plus pertinents par rapport à la requête. Dans son combat pour lutter contre les « infox » notamment, un moteur de recherche peut désindexer les contenus concernés ou au moins faire en sorte qu'ils n'apparaissent pas dans les premiers résultats.

Google (91,2% de part de marché en France selon statcounter.com en septembre 2021) a modifié son algorithme et formé des Search Quality Raters pour contrôler l'acuité de l'algorithme afin de réduire le quality score des sites d'infox et empêcher leur affichage dans les premières pages de résultats du moteur. En termes d'impact sur le modèle d'affaires de ces sites cela signifie moins d'audience et donc moins de revenus mais c'est aussi une façon de réduire le pouvoir d'influence sociale et politique de ces sites. Dans le cas du search, la publicité ne finance pas directement les sites d'infox : 100% des recettes publicitaires du search sont au bénéfice du moteur de recherche.

Social

Les réseaux sociaux ont la particularité de ne pas produire directement de contenus vidéos ou éditoriaux. Ce sont des plateformes de partage de contenus où les utilisateurs vont publier des contenus propres (« user generated content » – UGC) ou des contenus issus de sites et d'applications tiers. Ces deux types de contenus peuvent être à risque pour les marques. La difficulté pour les mesureurs est de plusieurs ordres.

En premier lieu, à la différence d'un site éditorial construit qui affiche un article qui sera le même pour tout le monde, sur les réseaux sociaux, on ne parle pas de page mais d'un fil infini (feed) de contenus construit algorithmiquement. Ce fil est unique pour chaque utilisateur. Autre difficulté, les réseaux sociaux appelés aussi « Walled Garden », soucieux de préserver la donnée de leurs utilisateurs pour leur propre usage, bloquent l'utilisation de code java script et de traceurs nécessaires au fonctionnement des outils d'advertising. Quand un outil d'advertising dit mesurer ces plateformes, en réalité, le contenu contrôlé par le mesureur sera transmis, serveur à serveur, par la plateforme. En aucun cas le mesureur ne contrôle directement le feed unique des utilisateurs in situ. En termes de modèle d'affaires, les plateformes sociales ne font pas de partage de revenus avec les émetteurs des contenus qu'ils soient UGC ou émanant d'un éditeur. La seule exception est celle des contenus vidéo des partenaires officiels des plateformes qui sont modérés par la plateforme en amont et dans lesquels les marques peuvent acheter des formats « instream ».

Autre exception : le réseau Facebook Audience Network (FAN) qui permet à des sites tiers de monétiser leurs inventaires publicitaires par l'intermédiaire de Facebook. De la même façon que pour la vidéo, les sites intégrés dans FAN sont modérés a priori par les équipes éditeurs de la plateforme. Ainsi, il n'y a pas de bénéfice financier direct pour les sites d'infox mais un bénéfice indirect via la captation d'audience qui leur permettra de mieux monétiser leur inventaire sur les places de marché.

Cette source de trafic est d'autant plus conséquente que les révélations récentes de la lanceuse d'alerte Frances Haugen (ex employée de Facebook qui sera auditionnée à l'assemblée le 10/11/2021) a révélé que les algorithmes de Facebook favorisent les contenus « polémiques », plus à même, selon l'entreprise, de permettre des « interactions sociales significatives » (meaningful social interactions) et ainsi de créer toujours plus d'inventaires publicitaires. Cela constitue un risque réel pour les annonceurs qui utilisent ces réseaux sociaux, que leur marque puisse être associée à ce type de contenus.

Concernant, le marché des réseaux sociaux, les plateformes déclarent des chiffres de visibilité aux mesureurs agréés mais ils ne contrôlent pas la brand safety de leurs inventaires. De plus, la visibilité et une partie de la brand safety sur les contenus instream ou alors sur le Facebook Audience Network ne sont pas non plus contrôlés. IAS travaille en exclusivité avec TikTok. Les serveurs de TikTok et de IAS sont interconnectés pour permettre à IAS de récupérer toutes les vidéos, elles sont scannées (son, sous-titre, etc.) et si deux vidéos dans le feed sont brand safe alors une publicité peut être placée entre ces 2 vidéos. Twitter a le même type de discussion avec IAS. Facebook pour le moment ne se positionne pas sur cette possibilité.

Display

Le marché du display est le plus concerné par les problématiques de brand safety mais c'est aussi le mieux mesuré par les outils d'adverification sous réserves des limites techniques évoquées ci-dessus. Nous pouvons identifier trois modalités d'accès à ces inventaires :

- L'achat direct, dit aussi achat de gré à gré, où l'acheteur s'adresse directement à un éditeur via sa régie publicitaire et entre en négociation pour acheter de l'espace publicitaire. Dans ce cas précis la sélection des sites est extrêmement contrôlée et le risque d'afficher sa publicité dans un environnement non propice à la marque est inexistante.
- L'achat via des réseaux de sites (Adnetworks ou via des Trading Desk indépendant), dans ce cas l'acheteur sélectionne ce réseau en général car celui-ci apporte un service et une valeur ajoutée précise, soit parce qu'il a fait une curation de sites sur des univers thématiques précis, ou parce que le site apporte de la donnée de ciblage exclusive, ou bien encore qu'il propose l'intégration de formats publicitaires uniques. Dans la majorité des cas ces réseaux sélectionnent scrupuleusement les sites partenaires pour garantir aux acheteurs la qualité éditoriale. Cependant il y a toujours des risques, notamment dans le cas particulier de la diffusion de « fake news ».
- Enfin, il y a l'achat programmatique qui consiste à acheter de l'inventaire sur les places de marché sur des sites filtrés soit par de la data, soit pour leur contexte éditorial, soit par de la performance publicitaire, etc. ; les critères peuvent être nombreux. En général, une campagne programmatique va être publiée sur des milliers de sites qui ne sont en général pas sélectionnés un à un (même s'il s'agit d'une option, notamment en vidéo) mais par les algorithmes en fonction des critères de ciblage choisis.

Quelles que soient les modalités d'achat de ces inventaires, les annonceurs et leurs agences utilisent systématiquement des outils d'adverification qui permettent d'identifier rapidement les sites non conformes. Même si le risque zéro n'existe pas il est extrêmement réduit.

Le dernier rapport sur la brand safety publié par Integral Ad Science (IAS) montre une baisse des risques relatifs à la brand safety sur le premier semestre 2021 par rapport au premier semestre 2020. Selon IAS, les principaux risques de brand safety au 1^{er} semestre 2021 étaient liés aux catégories suivantes : contenu pour adultes, alcool, violence, langage offensant et contenu controversé (ce qui inclut les fausses informations).

En agence, nous observons des taux moyens de contenus non brand safe monétisés inférieurs aux moyennes observées par IAS car nous apportons une attention particulière à la sélection des contextes de diffusion en amont des campagnes. Pour cela les agences mènent plusieurs actions parmi lesquelles :

- Nous créons nous-même des places de marchés privées où chaque éditeur est sélectionné par l'humain,
- Nous n'achetons pas sur les places de marchés peu scrupuleuses qui n'offrent pas de garanties suffisantes en termes de filtrage de la qualité des éditeurs,
- Nous appliquons notre liste noire sur chaque campagne,
- Nous utilisons également régulièrement, pour les annonceurs qui acceptent de prendre en charge un surcoût, les fonctionnalités pre-bid (pré-enchère) qui permettent de ne pas enchérir sur les sites non-brand safe.

Le modèle d'affaires des sites complotistes ou de désinformation est la monétisation de leurs contenus directement sur ces places de marché dont ils tirent l'essentiel de leurs revenus.

Périmètre et portée des agences media de l'UDECAM et de leurs clients sur le total des investissements digitaux :

<i>Brand Safety</i>	Existence du	Contrôle du Risque	Part de Marché des agences de
SEARCH	∅	∅ (risque inévitable)	15% (85% liste longue)
SOCIAL	**** (Editeurs + UGC)	Faible mais en amélioration	20% (80% liste longue)
DISPLAY	*** (Editeurs)	Fort (<1,5%)	66% en moyenne (34% en direct avec les éditeurs, les Adnetworks et

**Poids des investissements publicitaires des Agences de l'UDECAM sur le total des dépenses publicitaires mesurées par l'Observatoire de l'èPub SRI-Oliver Wyman-UDECAM*

Des mesures sont prises au quotidien par les agences pour limiter les risques

Les agences de l'UDECAM utilisent depuis 2011 et l'arrivée conjointe de l'achat programmatique en France des outils d'adverification. De plus, pour s'assurer de limiter les risques de diffuser dans de tels environnements nous avons des clauses contractuelles avec nos clients qui exigent qu'un outil d'adverification soit utilisé pour toutes les campagnes, qu'il soit souscrit directement par l'annonceur ou que le prestataire mesureur de l'agence lui soit facturé.

L'ensemble des agences de l'Udecam ont des contrats directs avec les mesureurs (Integral Ad Science – IAS, MOAT, Double Verify, Adloox, etc.) et travaillent régulièrement sur les améliorations des produits.

A titre d'exemple, voici certaines des mesures prises chez **Publicis Media (chacune des agences ayant leur propre Roadmap)**

- Dans le cas de notre collaboration avec les adnetworks et les trading desk indépendant nous imposons l'utilisation d'un outil d'adverification et nous les alertons pour qu'ils bloquent tous les sites litigieux que nous aurions remarqué.
- Dans chaque agence, nous diffusons et mettons à jour régulièrement un guide des bonnes pratiques pour limiter au maximum les risques liés à la Brand Safety.
- Concernant les réseaux sociaux et les mesureurs, nous travaillons très régulièrement à l'amélioration des intégrations techniques permettant de limiter les risques de brand safety.

Vous trouverez en annexe l'exemple des mesures prises au sein de Publicis Media « Actions mises en place chez PMF pour la Brand Safety » et un comparatif Integral Ad Science des mesures possibles sur chaque plateforme sociale et le planning de mise en place sur les mois années à venir.

Y a-t-il des échanges et discussion à ce sujet au sein de l'UDECAM, et un consensus se dégage-t-il ? Si oui, lequel ? Dans quelle mesure les pratiques différent-elles ?

Il y a des échanges nourris au sein de l'Udecam. L'ensemble des agences est sensibilisé à la nécessité de ne pas financer les infox et (Source étude Globale Mediamétrie-Nielsen Septembre 2021) et en affichant une annonce sur un de ces contenus incitant à la haine. Il faut bien comprendre que les agences n'ont aucun intérêt (au contraire) à financer ces sites : les inventaires existant sur le web sont gigantesques (environ 15 à 20 milliards de pages vues par jour pour 45,9 millions d'utilisateurs par jour moyen sites l'agence risque d'entacher la réputation de la marque pour laquelle elle travaille. C'est donc une chose que nous souhaitons éviter à tout prix. Par extension, le consensus est que l'une des prérogatives de l'Udecam est avant tout de s'assurer de diffuser les annonces de ses clients annonceurs dans des contextes brand safe et brand suitable ce qui exclut de facto les sites d'infox ou les contenus incitant à la haine.

Au sein de l'Udecam nous sommes d'accord pour dire que le marché du display est plutôt bien maîtrisé notamment grâce aux outils d'adverification. Certes certaines annonces vont apparaître sur des sites d'infox ou incitant à la haine mais ce phénomène est principalement dû à l'ingéniosité des éditeurs de ces sites qui vont créer de nouveaux noms de domaine ou changer leur adresse IP pour passer au travers des mailles du filet. Il va sans dire que dès qu'une telle opération est détectée (grâce au reporting a posteriori de nos outils d'adverification) nous faisons évoluer nos règles de filtrage pour s'assurer que cela ne se reproduise plus et que le site en question intègre notre liste noire commune à tous les annonceurs.

Par conséquent, nos discussions portent plutôt sur les réseaux sociaux. Ces derniers vont avoir tendance à relayer des infox mais les sommes investies ne viennent pas enrichir directement les sites de fake news relayés sur les réseaux sociaux. Sur ce point, les plateformes de réseaux sociaux sont les seules à bénéficier financièrement des investissements des annonceurs et ce sont leurs algorithmes qui devraient effectuer un filtrage. En revanche le bénéfice indirect pour ces sites est la génération d'audience que permet le partage de leurs contenus sur les réseaux sociaux : dès lors qu'un utilisateur clique sur le lien éditorial, le site concerné pourra monétiser ses inventaires à des acheteurs peu regardant (long tail d'annonceurs qui ne passent pas par des agences, adnetworks, trading desks indépendants, etc.).

Au sein des agences nous sommes tous alignés sur la nécessité et la systématisation de l'utilisation d'outils

Vous avez récemment affirmé dans la presse que l'assèchement des revenus publicitaires pour les sites de fake news était votre combat de la décennie. Qu'allez-vous faire ? Sur quoi faut-il agir selon vous, à la fois au sein de Publicis et de l'UDECAM ?

Les media ont toujours été la base de nos démocraties. La publicité a permis de financer les media d'information et partant permet de supporter la diversité des courants de pensée. Nous constatons une défiance en augmentation constante vis-à-vis des media d'information. Parmi les causes de cette défiance se trouve la diffusion de fake news qui sont souvent plus attrayantes que l'information de qualité : elles offrent des solutions simples à des problèmes complexes, satisfont le goût du lecteur pour le sensationnel, etc. Cependant, nous sommes conscients du caractère nocif de ces contenus pour la société et nous considérons qu'il est de notre devoir moral en tant qu'acteur responsable de lutter contre.

Par ailleurs, ces contenus sont nocifs pour les marques que nous accompagnons et il est donc également de notre responsabilité en tant qu'acteur économique et vis-à-vis de nos clients de veiller à ne pas financer ces contenus. Nous allons donc continuer, au sein de l'UDECAM et individuellement, à investir dans les moyens de lutter contre ces sites aux côtés d'autres associations professionnelles concernées par ce sujet comme l'UDM, le SRI, l'IAB, le GESTE, la MMA, etc.

Aujourd'hui les agences outrepassent largement le rôle de simple centrale d'achat que l'on pourrait leur assigner en adoptant une attitude agressive vis-à-vis des sites d'infox, en investissant dans les outils d'adverification, en formant leurs collaborateurs.

Il est probable qu'il faut agir sur l'offre en créant des obligations aux vendeurs d'espaces publicitaires de garantir la brand safety des emplacements qu'ils commercialisent. Il pourrait également être pertinent d'investir dans la pédagogie pour permettre aux citoyens d'identifier ces sites.

La Commission a pour but de proposer des solutions concrètes pour réduire le financement de la désinformation via la publicité programmatique. Quelles sont les propositions qui vous semblent devoir être mises prioritairement en œuvre au sein des agences médias ? Y a-t-il d'autres actions qui vous paraissent essentielles, de la part d'autres acteurs ?

Au sein des agences media le problème ne réside pas tant dans ce qu'il faudrait faire en plus comme action que sur l'efficacité des outils existants. Par l'indépendance et l'agnosticité que nous garantissons à nos clients nous sommes obligés de recourir à des outils de vérification tiers et les garanties de brand safety sont au cœur de nos

Il y a donc pour nous 3 chantiers prioritaires :

- Un chantier sur l'amélioration des outils tiers existants, notamment sur la performance et l'exhaustivité des listes noires de sites à ne pas utiliser. Une des mesures que nous pourrions proposer au sein de l'Udecam serait la mise en œuvre d'un guichet ouvert pour déclarer les sites de fake news qui nous permettrait d'améliorer encore ces listes noires.
- Un chantier à mener avec les plateformes de réseaux sociaux, adnetwork, réseaux Native Ads, intermédiaires à la performance type Trading Desk Indépendant, Google Display Network, Facebook Audience Network et les SSP programmatiques (Supply Side Platform) tiers 2 et 3 pour veiller à ce que les inventaires qu'ils commercialisent soient compatibles avec l'exigence de lutte contre les fake news et les contenus haineux.
- Un dernier chantier sur l'intégration de nouveaux leviers qui sont sur des environnements technologiques différents comme les TV connectées dont le volume croit considérablement d'une année sur l'autre, le digital OOH ou l'audio avec le streaming et les podcasts.

Si les agences travaillent sur ces sujets et sont souvent pointées du doigt nous aimerions être soutenus par d'autres acteurs dans cette démarche. En effet, certains acteurs comme les réseaux sociaux de plus petite taille que Facebook par ex., les trading desk indépendants, etc. se cachent derrière leur rôle d'intermédiaire technique alors que c'est en réalité de leur responsabilité de vérifier que leurs inventaires ne contiennent pas de contenus contraires aux réglementations en vigueur. A minima, par les limites qu'ils mettent ce sont les seuls à pouvoir techniquement assurer ce rôle. Aujourd'hui ces intermédiaires dans la chaîne de valeur échappent à leurs responsabilités alors même qu'ils sont la clé de cette lutte contre les sites complotistes et incitant à la haine.

Les entreprises offrant des listes d'exclusion de sites web conspirationnistes et incitant à la haine existent. Lesquelles connaissez-vous ? Ne pensez-vous pas que les agences de médias françaises pourraient travailler directement avec ces acteurs (l'un ou l'autre), en intégrant leurs listes par défaut ?

Oui nous les connaissons et nous contribuons à les enrichir via les URL black listées qui nous remontent tous les jours. Publicis a par exemple un partenariat mondial avec Newsguard et nous complétons notre black list des URLs et application fournies par Newsguard. De plus, toutes les impressions sont mesurées par des outils d'ad vérification ce qui permet a posteriori d'améliorer les listes d'exclusion avec les sites repérés par les agences media.

Cependant, ainsi que nous le mentionnions, ces listes d'exclusion sont un dispositif malheureusement imparfait en raison de l'inventivité des éditeurs de sites web conspirationnistes et incitant à la haine qui font beaucoup pour les contourner.

Nous précisons que ces outils d'adverification ne se limitent pas à signaler et bloquer la diffusion sur des sites conspirationnistes ou alimentant la haine en ligne. Ils limitent, signalent et bloquent également les environnements liés à l'alcool, la drogues, les sites adultes, les sites de téléchargement ou de streaming illégaux et les sites violents.

De façon générale, pour s'assurer de limiter les risques de diffuser dans des environnements comme les sites conspirationnistes ou incitant à la haine, nous avons des clauses contractuelles qui exigent qu'un outil d'adverification soit utilisé et souscrit sur toutes les campagnes des marques. Si l'annonceur n'a pas de contrat direct avec un prestataire nous lui proposons d'utiliser les plateformes sous licence en contrat direct Publicis, que nous leur refacturons en toute transparence.

Quels sont d'après vous les acteurs (personne ou institution) que la commission devrait consulter sur le sujet qui nous préoccupe ?

Il y a plusieurs personnes que la commission pourrait consulter sur ce sujet :

- Les éditeurs de solutions d'adverification qui proposent des produits utilisés par les agences pour identifier au mieux les sites complotistes ou propageant la haine : Buster.ai, Newsguard, Integral Ad Science, Double Verify, MOAT, Adloox, Human sur la Fraude.
- Les plateformes qui constituent autant de « rampes de lancement » pour les contenus qui préoccupent la commission ou certains qui, au contraire, ont mis en place des dispositifs de filtrage efficaces : Facebook, TikTok (qui a mis en place un dispositif de contrôle avec IAS que nous avons testé récemment en exclusivité et qui affiche un score de 98,7% d'impressions Brand safe), Twitter, Leboncoin et Amazon (qui ont mis en place des solutions de lutte contre la contrefaçon automatisés et dont les mécaniques pourraient être utilisées pour lutter contre les contenus qui intéressent la commission).
- Les intermédiaires à la vente qui font de la curation de chaînes et vidéo Youtube (Youtube représente en France, à lui seul, entre 40 et 45% des investissements publicitaires vidéo) garantissant la brand safety aux marques tels que Zefr et Channel Factory.
- Les équipes de Monsieur l'ambassadeur pour le numérique Henri Verdier.

Annexe 1 - Actions mises en place chez PMF pour la Brand Safety

Définition Brand Safety

Le principe de brand safety consiste à prévenir la diffusion de publicités sur des contextes pouvant être néfastes pour l'image de la marque, soit sur des sites et contenus néfastes (sexe, drogue, violence, raciste, etc.), soit sur des contextes auxquels la marque ne souhaite pas être associée (actualités spécifiques, événements spécifiques, etc.).

Typologie des contenus/contextes et part dans le volume de pages publiées

Certains contextes sont clairement identifiés : du contenu à caractère sexuel, terroriste, violent... classés en 11 catégories par l'IAB. Ces contenus peuvent être identifiés grâce à des mots-clés, des URLs et ils laissent peu de place à l'interprétation. Les technologies de brand safety sont capables de les identifier clairement, de manière automatique et de bloquer la diffusion sur ces environnements.

D'autres sites ayant une ligne éditoriale claire avec la publication de contenus haineux ou discriminatoires sont également identifiables – pour les plus importants d'entre eux.

Depuis quelques années, d'autres types de contenus se développent, notamment autour des théories complotistes et d'informations erronées. Ces contenus sont plus difficilement identifiables.

En effet, les outils de détections traditionnels ne peuvent déterminer automatiquement quels contenus partagent une propagande complotiste ou de fausses informations, car cela nécessite une intervention et une interprétation humaine.

Selon Integral Ad Science¹, sur le marché français, le volume de pages représentant un risque est de :

- Display Desktop : 4,5% ayant un risque moyen et 0,4% un risque élevé
- Display Mobile : 6,6% ayant un risque moyen et 1,1% un risque élevé
- Vidéo Desktop : 7,8% ayant un risque moyen et 1,8% un risque élevé
- Vidéo Mobile : 6,6% ayant un risque moyen et 1,1% un risque élevé

A titre indicatif, sur les dépenses du premier semestre 2021 via le DSP DV360 pour le périmètre Publicis Media France, et grâce aux mesures de contrôles supplémentaires mises en place, la part d'impressions à risque est de 0,07% des impressions et 0,018% du media cost a été diffusé sur des sites de la blocklist Newsguard (soit moins de 24€ en moyenne par annonceur sur la période).

Les mesures et actions mises en place au sein de Publicis Media

Différentes actions sont effectuées sur les activations digitales de manière à réduire au maximum le risque de diffusion des campagnes sur des sites ou contextes problématiques. Elles reposent sur des process et cadrages en amont de la diffusion des campagnes et des outils embarqués nativement par les plateformes d'achats utilisées ou ajoutées lors de la diffusion pour renforcer le contrôle des contenus/contextes qui vont être utilisés.

Malgré l'ensemble de ces actions et contrôles, le risque « zéro » n'existe pas puisque d'une part la qualification des contenus reste pour certains, liée à l'interprétation humaine, et d'autre part certains sites identifiés développent des stratégies de changement de nom de domaine qui rendent leur traçabilité très complexe.

¹ IAS - Baromètre de la qualité média H2 2020

Intégration de filtre de contrôle en natif dans les outils programmatiques

Les plateformes d'achats programmatiques intègrent par défaut des filtres pour exclure les plus gros sites problématiques. Ces filtres sont très larges et ne sauraient pas être la seule mesure pour assurer la protection.

Les plateformes de vente d'inventaire programmatique appliquent également des filtres pour ne pas intégrer dans leur réseau des sites potentiellement néfastes. Ces filtres ne sont pas infaillibles : les sites peuvent passer par des intermédiaires pour être intégrés dans les plateformes, trompant les outils et audits mis en place.

Intégration d'outils tiers permettant une vérification des contenus en amont de l'impression

Plusieurs outils tiers sont disponibles afin de prévenir la diffusion sur des sites néfastes : Integral Ad Science, DoubleVerify, MOAT, Adloox, Grapeshot, etc.

Ils permettent d'empêcher la diffusion sur des pages pouvant être néfastes. Ces outils fonctionnent en scannant le web afin d'enregistrer et catégoriser le contenu des pages. Quand une opportunité d'acheter une impression publicitaire se présente, ils recherchent dans leur base de données si la page a déjà été scannée et si elle présente un risque.

- Si elle présente un risque, ils empêchent l'achat de l'espace
- Si elle n'est pas connue, en fonction des réglages de l'outil, ils peuvent laisser faire l'achat de l'impression. Après l'achat, mais avant l'affichage, l'outil va scanner la page à la recherche de mots-clés potentiellement dangereux. Si des mots-clés interdits sont détectés, la publicité est remplacée par une image neutre et la marque annonceur n'est pas associée au contenu sensible.

Les limites des outils tiers :

- Le scan n'est pas en temps réel : il peut y avoir un écart de quelques heures ou de plusieurs jours en fonction de l'audience du site entre la publication d'un article et son indexation dans les outils.
- Ils sont globalement efficaces pour détecter certains types de contenus néfastes : violence, terrorisme, drogue. En revanche, sur la diffusion de fausse information, ils sont limités par la nécessité de comprendre le contenu de manière précise avec les nuances nécessaires - que seul un humain peut détecter à date.

Un panorama et un audit complet des outils tiers sont disponibles auprès des équipes Data&Tech (Antoine Barbet)

Paramétrage des campagnes dans les process d'achat

Une liste d'exclusion de plus de 66 000 sites et applications est régulièrement mise à jour, est intégrée par défaut sur l'ensemble de nos plateformes d'achats.

Cette liste est également communiquée aux partenaires externes (Adnetwork et Trading Desk indépendant) opérant les campagnes afin qu'elle soit appliquée sur leurs outils.

Pour aller plus loin, signature d'un partenariat avec Newsguard

Pour pallier les limites des outils traditionnels, notamment liées à l'exactitude des informations et au positionnement des contenus (complotiste par exemple), Publicis a signé début juin un partenariat avec Newsguard. Ce partenariat permet de qualifier les contenus de plus de 6 000 sites et est mis gratuitement au service des campagnes de nos annonceurs.

NewsGuard a été fondée en 2018 pour fournir des outils publicitaires permettant d'assurer la transparence des notes de confiance sur plus de 6 000 sites/applications d'actualité et d'information.

Ces 6 000 propriétés représentent 95 % des engagements en ligne dans les pays où Newsguard opère. L'équipe de NewsGuard comprend un large éventail de rédacteurs et de journalistes d'horizons divers, opérant aux États-Unis, au

Royaume-Uni, en Allemagne, en France et en Italie. Leur équipe évalue la crédibilité et la transparence des sites d'informations et d'actualités en se basant sur neuf critères journalistiques, notamment :

- La publication répétée de faux contenus par le site
- La transparence sur l'actionnariat du site et son financement
- Le processus de correction des erreurs factuelles lorsqu'elles se produisent

Sur la base de ces neuf critères, chaque site reçoit une note globale avec un code couleur, vert (généralement digne de confiance) ou rouge (généralement non-digne de confiance), une note de confiance de 0 à 100 points, ainsi qu'une "étiquette nutritionnelle" détaillée expliquant qui est à l'origine du site, pourquoi il a reçu cette note et ce score, et des exemples spécifiques de tout problème identifié par Newsguard.

Sur la base de ces scores, tous les sites ayant une note inférieure à 50 (selon la recommandation de Newsguard) sont désormais exclus de la diffusion de nos outils. Elle est également communiquée à nos prestataires pour qu'ils puissent les exclure de toute diffusion.

Publicis est membre actif du Global Alliance for Responsible Media (GARM)

La Global Alliance for Responsible Media a défini en janvier 2020 un plan d'action pour éradiquer les contenus nuisibles en ligne et s'efforce de créer un environnement numérique plus durable et plus responsable qui protège les consommateurs, le secteur des médias, les marques et, par conséquent, la société.

Les membres de GARM ont fixé 3 priorités :

- Établir des normes de sécurité communes et universelles pour la publicité et les médias
- Améliorer et créer des outils communs de brand safety pour l'ensemble de l'industrie
- Favoriser la responsabilité mutuelle et des instances de contrôle et de surveillance indépendantes.

Dans le cadre de son positionnement de Performance Responsable, Publicis Media est engagée au quotidien avec l'ensemble des acteurs de la chaîne de valeur et le régulateur pour apporter toujours plus de sécurité et de performance à ses clients.

Au quotidien, les actions mises en place et à suivre dans le temps

Audit des comptes

Auditer nos comptes clients sur l'utilisation d'un outil de brand safety.

- Quel outil est utilisé, quelle est la nature du contrat (direct ou via Publicis).
- Quel périmètre d'utilisation de l'outil ?
 - o L'outil est-il utilisé sur 100% des stratégies media (du branding à la performance), sur tous les modes d'achat (direct y compris les réseaux et en programmation self-service) ?
 - o L'outil est-il activé sur toutes les marques ?

En fonction de cet audit, il est important de partager avec le client et de le sensibiliser aux risques pris si l'utilisation de ces outils n'est pas souhaitée. Il est important de recueillir leur consentement (optin/optout) sur l'utilisation ou non d'un outil de brand safety.

Ce processus est à renouveler à minima tous les ans.

De l'exclusion vers l'inclusion

Si les campagnes de Branding sont majoritairement diffusées sur des sites référents et ont par ailleurs souvent un mesureur brand safety intégré, les campagnes ayant des objectifs de performance ou les campagnes de retargeting diffusent sur des environnements plus large (réseaux) dont les contenus sont moins bien qualifiés. L'objectif de ces campagnes étant d'avoir une couverture maximale afin de laisser l'algorithme des plateformes d'achats avoir suffisamment de données pour obtenir les meilleures performances. L'objectif de ces campagnes étant lié au coût d'acquisition, elles font moins l'objet de mesureurs tiers, car ils viendraient alourdir le coût de la campagne avec des frais techniques additionnels.

Aussi, d'une manière générale et en particulier sur ces campagnes performances, et afin de limiter au maximum le risque (comme l'apparition d'un nouveau site pas encore identifié comme potentiellement problématique), il convient de changer la façon d'opérer pour utiliser uniquement une liste de sites définis en amont. Et ainsi passer d'une logique d'exclusion (listes de sites ou de contenus non voulus) à une logique d'inclusion (liste de sites souhaités car vérifiés en amont).

Cette liste devra être le plus large possible afin de garantir de bonnes performances, mais elle restera fermée pour éviter l'entrée de nouveaux sites ou les pratiques de changements de nom de domaine.

Annexe 2 - Integral Ad Science (IAS)

Tableau comparatif des mesures de brand safety, de visibilité et de fraude compatibles avec les principales plateformes sociales

Une couverture unique des plateformes Sociales

	Facebook	Instagram	IMM	Twitter	YouTube	Verizon Media Native	Spotify	Snapchat	Pinterest	LinkedIn	Twitch	Walmart Connect	TikTok	Amazon	Microsoft Audience Network
METRIC															
Viewability	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓ alpha	2021	2021	2021/2 022	2022	2021
Ad fraud	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓ alpha	2021	2021	2021/2 022	2022	2021
Brand safety	✓*		✓	2021	✓	Channel Science & Pre-screened YouTube Select Lineups				✓*	2021		2021	2022	2021
FORMAT															
Display	✓	✓	✓**			✓	✓	2021	✓	H2 2021	N/A	2021	N/A	2022	2021
Video	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓ alpha	2021	2021	2021	2022	2021
ENVIRONMENT															
Web	✓	N/A	✓	✓	✓	✓	✓***	N/A		✓ alpha	2021	2021	N/A	2022	2021
In-App	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓ alpha	2021	2021	2021	2022	N/A

* In stream video and instant articles only
 ** Display is only supported for Brand Safety
 *** Web leaderboard banners measurable with IAS JS tag

Content Allow Lists for In-Stream Video

* LinkedIn Audience Network Platform-wide Brand Protection: bid on brand safe and fraud free ads