

Un radar hi-Tech pour surveiller le web social

Aurélien Vathonne

Radarly est un outil de surveillance du web qui a vu le jour en 2012. Il a été conçu par la société Linkfluence qui depuis 2006, s'est fait connaître en tant qu'institut d'étude « 2.0 », pionnier de l'analyse sociologique du web, à une époque où le concept-même de web social était encore balbutiant.

Radarly arrive régulièrement dans le top des meilleurs outils de social monitoring. Certains blogueurs l'ont même qualifié de « Rolls-Royce de la veille sur le web social »¹ Et de fait, l'un des principes qui ont présidé à sa conception est d'extraire « le meilleur de la veille et de l'IE » pour la collecte et l'analyse des données et « le meilleur du web » pour l'interface et le design. Après le rachat de son concurrent Trendy Buzz

¹ <https://ateliercartographie.wordpress.com/2014/05/15/linkfluence/>

en 2014, Linkfluence a dévoilé en 2015 la V2 de Radarly, avec notamment des fonctionnalités évoluées d'analyse.

Pour bien comprendre comment fonctionne Radarly, il faut prendre conscience du fait qu'il ne s'agit pas d'un « Google des réseaux sociaux ». Au démarrage d'un projet, la première étape consiste d'abord à délimiter en fonction de critères assez larges, un ou plusieurs périmètres de données qui vont se mettre à jour en temps réel avec les nouvelles publications, et qui vont constituer le matériau de base du projet. C'est d'ailleurs la volumétrie de ce périmètre qui détermine une partie du prix de facturation.

Une deuxième étape va consister à paramétrer l'outil en construisant notamment des requêtes au sein de ce corpus de départ, dont les résultats

seront consultables et analysables au sein de tableaux de bord dynamiques (dashboards) (voir fig.1)

COUVERTURE

Typologie des sources du corpus de Radarly :

- Réseaux sociaux (voir détails plus bas)
- Medias en ligne
- Autres sites web
- Blogs
- Vidéos
- Forums
- Commentaires : considérer les commentaires (publiés à la suite de billets de blogs ou d'articles) comme un type de source à part entière est une originalité de Radarly.

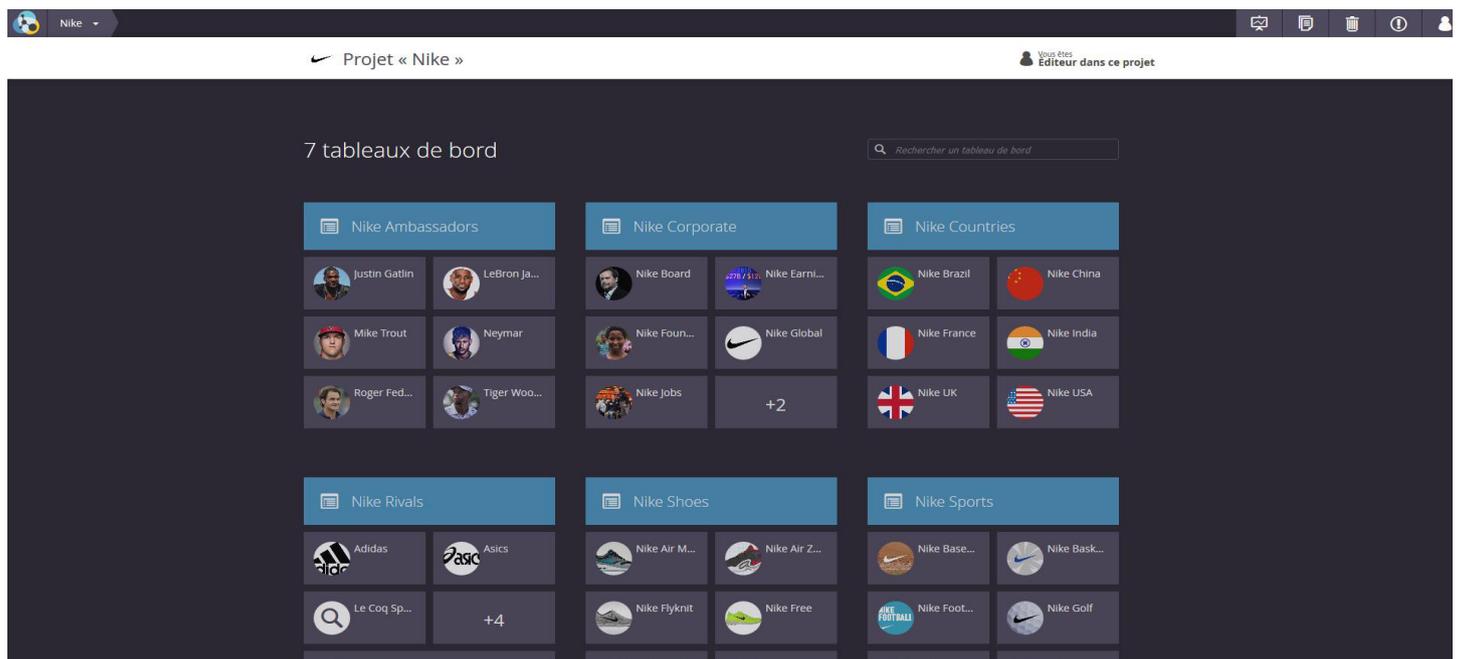


Figure 1 : Vue de la page d'accueil de Radarly, avec ses tableaux de bord (ici projet « Nike »)

Nombre de sources : En tout, 300 millions de sources dont 80 à 100 000 sources medias

Sources gratuites / payantes : Gratuites uniquement

Quels médias sociaux sont surveillés ?
Facebook (pages publiques) ², Twitter, Instagram, Google+, Youtube, Dailymotion, blogs.

Nombre de langues couvertes : 60 langues

Nombre de pays couverts : Tous pays

Réseaux sociaux locaux : peuvent être ajoutés à la demande et sur devis

PARAMETRAGE DE LA VEILLE

Types d'accès : Il existe quatre types d'accès qui vont du simple lecteur au propriétaire du radar (super administrateur), en passant par l'administrateur et l'éditeur (avec des droits de lecture/écriture). Un même utilisateur peut gérer ou avoir accès à plusieurs projets étanches, et au sein de chacun des projets, l'administrateur peut donner accès à certains tableaux de bord et pas à d'autres.

Qui peut créer les requêtes ?
L'utilisateur, pour peu qu'il ait au moins les droits d'éditeur, est autonome dans la création de requêtes au sein du corpus de départ. L'administrateur a en outre la main sur la « requête-mère » c'est-à-dire celle qui détermine le corpus de données.

² Pour pallier la fermeture récente de l'API search de Facebook, les concepteurs d'outils de social media monitoring ont cherché des solutions : alors que Synthesio a développé un partenariat avec Datasift qui récupère les données des profils/pages privées, les anonymise, les agrège et fournit des statistiques, Linkfluence a de son côté développé une technologie maison : un large échantillon pertinent et qualifié des groupes et pages Facebook en accès public (représentatif par secteurs et pays).

Fonctionnalités de recherche : Par type de sources, pays, langue. On peut préciser que la recherche textuelle doit se faire sur le titre uniquement ou sur le texte (quand cette précision a du sens, ce qui n'est pas le cas par exemple pour une recherche sur les tweets), ou encore que la recherche d'un mot doit se faire au sein des hashtags uniquement. On peut inclure dans une requête le flux provenant d'un compte Twitter ou d'un compte Instagram ou encore d'une page Facebook spécifiques. Enfin on peut rechercher l'information provenant d'un site web particulier.

Opérateurs disponibles

Les opérateurs sont très nombreux et sophistiqués : on trouve bien sûr les classiques opérateurs booléens *AND OR NOT*, la recherche d'expression exacte « entre guillemets », les troncatures * et ? mais aussi, ce qui est déjà plus rare, des opérateurs de proximité entre deux mots (en utilisant le *tilde*) ou entre deux groupes de mots (en utilisant *NEAR*).

exemple : « danone yaourt »~3

-> recherche les mots danone et yaourt éloignés de 3 mots au maximum

(mot1 OR mot2 OR mot3 OR mot4)

NEAR10 (mot5 OR mot6 OR mot7

recherche l'un des mots du groupe 1

éloigné de 10 mots au maximum

de l'un des mots du groupe 2

On peut même spécifier l'ordre des mots, avec les chevrons (opérateur <<), assorti d'une possibilité de préciser le nombre de mots minimum entre chaque terme.

exemple : mot1 <<5 mot2 <<6 mot3

Mais ce n'est pas tout, puisque Radarly offre également un opérateur *quorum* (permettant d'exiger la présence de *n* mots au sein d'une liste de mots définis),

ainsi qu'un opérateur permettant de spécifier la casse (majuscule ou minuscule) sur la totalité du mot ou seulement la première lettre. C'est à notre connaissance, le seul outil de veille sur les médias sociaux à proposer ce type d'opérateurs.

Dossiers / classement par thématique :

Oui, la catégorisation des publications s'effectue automatiquement grâce à la fonction **clusters** qui repose sur la technologie du « fuzzy matching » et qui permet de regrouper les publications suffisamment similaires, mais aussi de les quantifier, et de les qualifier (tonalité). Cette fonctionnalité est particulièrement efficace sur de grands sujets médiatiques. En revanche, elle n'est pas encore opérationnelle sur les « back data » (historique de données qui peut être chargé dans l'outil au démarrage d'un projet).

Fonctionnalités d'alertes : Oui, sur le volume ou l'auteur. Il s'agit plutôt d'une notification qui va permettre d'alerter l'utilisateur lorsqu'une requête génère des résultats dépassant un certain volume ou encore lorsqu'un auteur sur Twitter (ou n'importe quelle autre plateforme), préalablement identifié comme « à surveiller », publie quelque chose. En revanche il n'existe pas d'envoi d'alerte avec des contenus.

Peut-on désambiguïser des termes de recherche ? (exemple orange/Orange) ?

Oui, avec l'opérateur qui permet de préciser la casse.

Peut-on sélectionner un corpus précis de sources ?

On peut sélectionner des types de sources (Twitter, médias en lignes, blogs ...), ou encore sélectionner précisément dans la requête un ou plusieurs comptes Twitter ou Instagram ou pages Facebook, ou encore certains sites web spécifiques.

Ajout manuel de sources :

C'était possible dans l'ancienne version de Radarly. Cette fonctionnalité a été supprimée dans la V2 mais devrait être rétablie courant 2016, au moyen d'un bookmarklet.

Nombre d'accès utilisateurs : illimité.

Radarly est facturé au projet, quel que soit le nombre d'utilisateurs.

Nombre de requêtes : au sein d'un

tableau de bord, l'utilisateur peut créer autant de requêtes qu'il le souhaite. Néanmoins, pour des impératifs de lisibilité, il est recommandé de ne pas dépasser une dizaine de requêtes par dashboard, et au-delà de créer de nouveaux dashboards.

Fréquence de mise à jour : Temps réel à la seconde.

ANALYSE DES RESULTATS

Filtres disponibles pour visualiser les résultats : Très nombreux filtres : des

filtres standard tels que le type de source (plateforme), le pays, la langue, la restriction aux seules images/vidéos, mais aussi des filtres plus originaux tels que la tonalité (négative, positive, neutre, mitigée), les mots-clés (tags), les influenceurs, le genre de l'auteur (masculin/féminin) ... On peut également ajouter des filtres manuels (c'est-à-dire construits par l'utilisateur), mais cela suppose que chaque donnée soit qualifiée manuellement par l'utilisateur au préalable.

Fonctionnalités d'analyse (graphiques, cartographie, nuages de mots-clés) et statistiques

Toutes ces fonctionnalités se retrouvent dans l'onglet *Analytics* d'un tableau de bord. Cette partie *Analytics* peut être consultée sous forme de synthèse ou dans sa forme détaillée. Dans les deux cas, une réglette verticale sur la gauche de l'écran permet de modifier de façon dynamique la période de temps analysée.

Dans la synthèse (voir fig. 2), l'utilisateur peut accéder à une courbe du volume

de données générées au cours de la période et pour chaque requête au sein du tableau de bord, on retrouve notamment le nombre de publications, leur répartition sur les différents types de plateforme, et leur tonalité (dont l'écrasante majorité est souvent neutre).

Dans la vue détaillée, on trouve en outre un nuage de mots-clés, la répartition par type de plateforme au jour le jour, la répartition géographique (sous forme de carte) et par langue, ainsi qu'une vue globale de la tonalité (sur l'ensemble des requêtes).

Analyse automatique du sentiment : Oui (tonalité positive, négative, neutre, mitigée)

Possibilité de corriger / supprimer des éléments manuellement : On peut corriger l'analyse du sentiment, la langue, ajouter différentes catégories de tags ou supprimer complètement des publications.

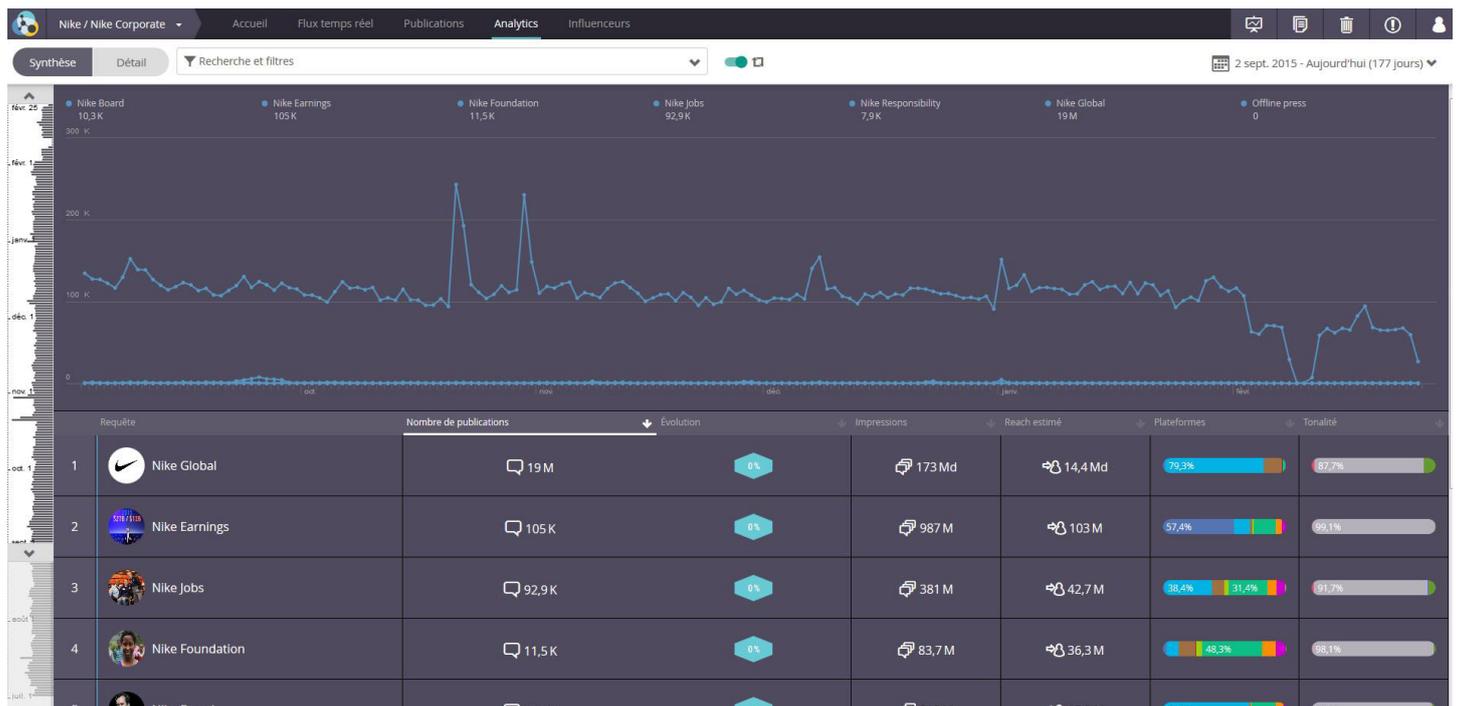


Figure 2 : vue de la page synthèse des Analytics

Comment l'outil traite les données dupliquées (retweets par exemple) ?

Les retweets sont inclus par défaut dans les calculs mais un bouton on/off accessible depuis tous les écrans, permet de les exclure de façon dynamique.

Rank/algorithm propriétaire pour évaluer l'influence et la fiabilité d'une source : L'influence est évaluée au moyen de plusieurs indicateurs : le score d'engagement d'un compte, c'est-à-dire sa capacité à générer une action (like, commentaire...), un score de viralité (capacité à se diffuser), et le reach (nombre estimé d'internautes ayant vu la publication, calculé à l'aide du nombre de followers + followers des gens qui suivent ce compte).

Identification d'influenceurs : Oui, dans l'onglet *Influenceurs*. Ces derniers peuvent être qualifiés en outre de façon personnalisée (tags, annotations).

Diffusion de la veille / export de données

Il existe de nombreuses fonctions d'export : les statistiques peuvent être exportées dans un tableur excel, mais il existe aussi une fonction de génération de rapport automatique mais néanmoins personnalisable (powerpoint) permettant de produire un livrable clé en main. Enfin, il est possible de créer une revue du web digitale au sein de laquelle on peut classer les publications par rubrique puis exporter le tout au format PDF, ou la consulter en ligne ou encore la partager par e-mail.

Tarifs

Il n'existe pas de formule standard car chaque projet fait l'objet d'un devis précis, en fonction du volume de données, du nombre de dashboards souhaités ou nécessaires, du degré d'autonomie des utilisateurs qui auront plus ou moins besoin d'encadrement et

de services rendus. Le prix de démarrage se situe aux alentours de 1 000 € HT / mois.

Recherche ponctuelle / antériorité

Les données sont collectées à partir du moment où le paramétrage de l'outil sur un projet/thème précis s'effectue. Néanmoins, il est possible d'acquérir des « back data » c'est-à-dire des données d'archives (toujours sur ce même projet ou thème), avec une antériorité qui peut remonter à 4/5 ans. Au sein du corpus de données défini, on peut réaliser autant de recherches ponctuelles que souhaité.

Autres

Personnalisation de la plateforme : la plateforme n'est pas disponible en marque blanche.

Partenariats/Liens/intégration avec d'autres outils : plusieurs partenariats sont en cours de mise en place mais Linkfluence ne communique pas encore dessus.

Formation, maintenance : Oui, inclus dans les services de *l'account manager* qui est la personne de contact principale.

Notre sentiment sur l'outil

Les possibilités de recherche et de filtres offertes par Radarly sont parmi les plus puissantes du marché sur ce type d'outils. Cette avance technologique se retrouve également dans la capacité à traiter en temps réel de gigantesques corpus de données, ou des fonctions de clustering automatique. Le design est en outre très réussi, nous avons beaucoup apprécié la timeline verticale qui permet de modifier la période de données analysées et de générer de nouveaux graphiques / statistiques de façon dynamique.

Les possibilités d'export sont particulièrement abouties, notamment avec l'ajout récent de la fonction « génération de rapport automatique

» et d'autre part avec la possibilité de réaliser une revue du web digitale.

Autre atout : la possibilité de charger des « back-data » (historique de données) avec une antériorité assez grande, pouvant remonter à 4 ou 5 ans.

Radarly peut s'adresser aussi bien à des utilisateurs « finaux » (responsables communication, marketing etc...) qu'à des utilisateurs plus aguerris aux systèmes de veille.

Tous ces points forts ont un coût, puisque Radarly est également l'un des outils les plus chers que nous ayons testés (à partir de 1 000 € / mois d'utilisation).

Pour plus d'informations

<http://linkfluence.com/produits/radarly/>