



ÉCOLE CENTRALE LYON

INFO
RAPPORT DU BE

MSO 4.4 - Rapport méthodologique de veille

Élèves :
Loïc DENIS

Enseignant :
Daniel MULLER
Mohsen ARDABILIAN
Nicolas JARDIN

12 mars 2021

Table des matières

1	Introduction	2
2	Démarche	2
2.1	Sources surveillées	2
2.2	Outils utilisés	3
2.3	Pipeline	4
3	Conclusion	5

Liste des tableaux

Table des figures

2	Sources ciblées par FeedBro	3
3	Résultat de l'acquisition via TweetDeck	4
4	Pipeline du dispositif de veille	5

1 Introduction

Pour mon sujet de veille, j'ai décidé de m'intéresser à l'utilisation des données personnelles par les réseaux sociaux. Depuis que des données personnelles d'utilisateurs ont été utilisées à leur insu à des fins politiques il y a quelques années, cette question est devenue très importante et l'exigence de chacun est devenue beaucoup plus ferme quant à la sécurité de ses informations privées. Dans cette veille, je m'intéresse à l'évolution des solutions qui sont apportées à cette problématique. Nous nous intéresseront en particulier aux solutions apportées par les réseaux sociaux eux-mêmes, mais aussi à l'évolution de la législation en générale.

2 Démarche

2.1 Sources surveillées

Journaux

Pour cette veille, j'ai pensé qu'il serait judicieux de surveiller en priorité des quotidiens spécialisés dans les nouvelles technologies, comme par exemple :

- The Verge
- The Guardian
- Le Monde
- TechCrunch
- The Next Web
- Mashable
- MIT Technologies Review
- Venture Beat Feed
- Le Journal du Net
- Numérama



(a) The Verge



(b) The Guardian

Les journaux sont volontairement aussi bien francophones qu'anglophones pour aussi bien avoir des informations sur les législations françaises que sur les avancées propres aux réseaux sociaux, qui eux sont internationaux.

Comptes Twitter

La veille va aussi s'appuyer sur Twitter en général, grâce à des alertes, mais aussi sur des comptes twitter en particulier comme :

- @analyticbridge
- @DataFeel
- @creapills
- @VeilleInnov

Ces comptes ont été choisis dans la mesure où ils mettent en avant des sujets d'actualités concernant le traitement des données et/ou les innovations technologiques.

2.2 Outils utilisés

Feedbro

L'outil qui a majoritairement été utilisé pour cette veille est Feedbro. Grâce à cet outil, nous avons pu suivre les actualités des journaux mentionnés ci-dessus, en mettant également en place des règles de tri de feed. Par exemple, les articles des journaux sélectionnés qui contiennent au moins un des mots "données" et "data" sont marqués comme importants, et ceux qui contiennent "leak" ou "fuite" sont marqués comme très importants. Ce sont ces articles qui vont être consultés en priorité.

Alertes Google

En complément des journaux, une mise en place d'une série d'Alertes Google avec les mots clefs suivants a été faite :

- "data" privacy
- "données personnelles"
- data loss
- data security
- gestion "données"
- protection "données"

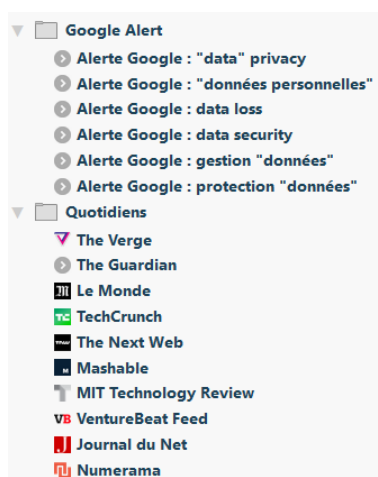


FIGURE 2 – Sources ciblées par FeedBro

TweetDeck

En complément de notre feed RSS, nous allons utiliser l'outil TweetDeck pour gérer l'acquisition des données à partir de Twitter. En plus des comptes mentionnés ci-dessus,

nous allons également faire une acquisition permanente par mots-clefs que nous pensons pertinents :

- "données personnelles"
- "data analytics"
- "privacy"



FIGURE 3 – Résultat de l'acquisition via TweetDeck

2.3 Pipeline

Pour résumer notre dispositif de veille, nous pouvons modéliser notre pipeline avec le schéma 4.

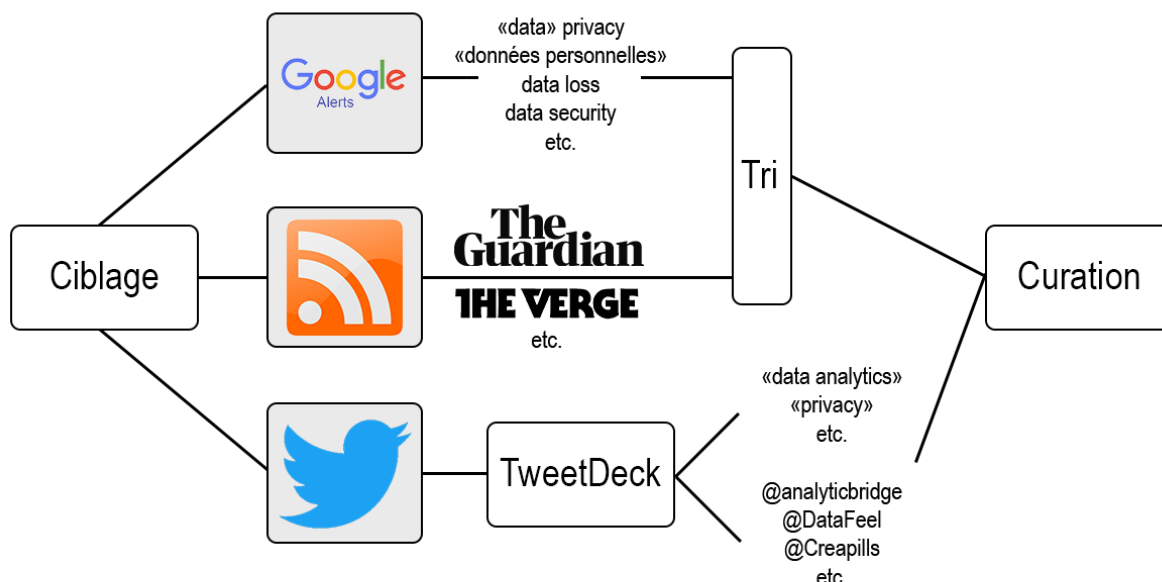


FIGURE 4 – Pipeline du dispositif de veille

3 Conclusion

Globalement, les résultats obtenus grâce à ce dispositif sont assez satisfaisants, néanmoins il peut être largement amélioré. En effet, les articles relevés par FeedBro étaient très nombreux et pas tous pertinents ; l'avantage est bien-sûr l'exhaustivité de mes résultats, mais il restait donc une assez grosse partie de tri manuel assez fastidieuse. Il aurait été très utile et probablement rentable de faire plus d'efforts sur l'amélioration de l'acquisition automatisé en affinant les mots-clefs, en testant d'autres choses, en prenant un peu plus de risque. Au contraire, ce dispositif de veille sous exploite le réseau Twitter. Ici aussi, il aurait été intéressant de passer plus de temps sur la recherche de mots clefs, de hashtags spécifiques et de comptes plus spécialisés.