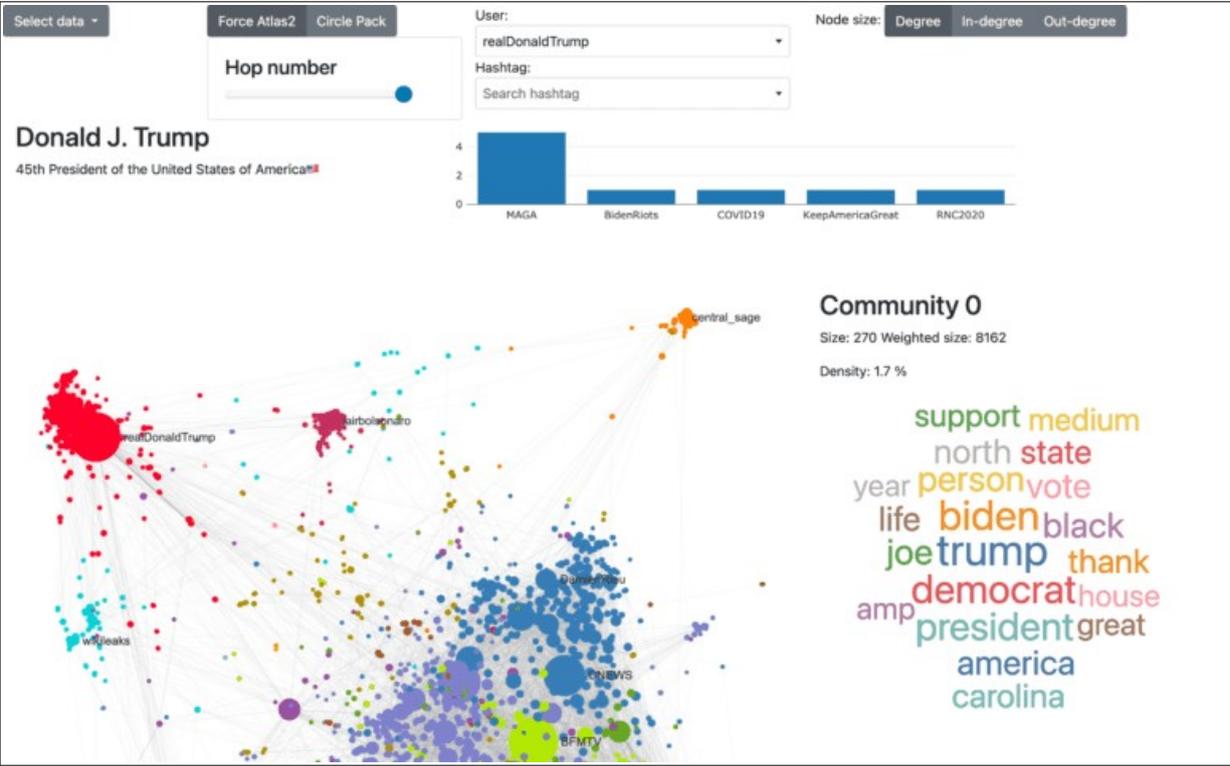


Pas-à-pas pour exploiter l'outil de visualisation SAD2

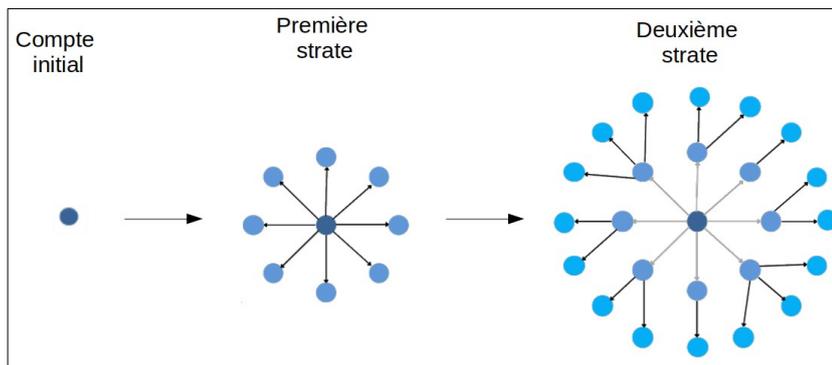


Avant propos... c'est quoi SAD et pourquoi l'utiliser ?

La plateforme twitter est un lieu d'exposition et d'échange d'opinions, les interactions entre comptes qu'elle propose suivent des dynamiques sociales spécifiques. Dans cet ensemble, les mentions et retweets sont régulièrement étudiés en sciences sociales et il ressort de ces travaux que des comptes se retweetant et/ou se mentionnant régulièrement entre eux permettent d'identifier des formes de communautés idéologiquement proches. Dès lors, l'étude de ces groupes et des liens qu'ils entretiennent entre eux permet d'avoir accès à une partie d'une controverse présente sur le web.

L'outil SAD2 permet aux utilisateurs de collecter des ensembles de comptes à partir de quelques personnalités identifiées au préalable, puis d'explorer les communautés de retweets dans un graphe dynamique.

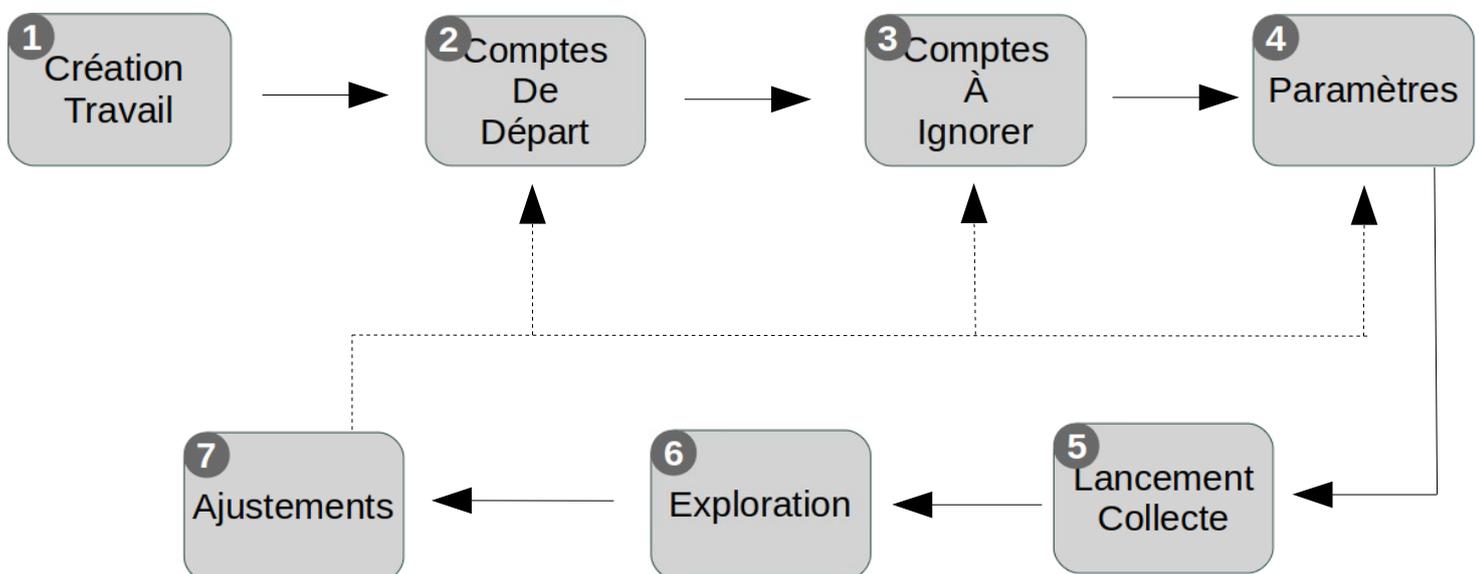
Son mode de collecte, qui est aussi sa spécificité, part de quelques comptes initiaux qui sont renseignés manuellement par l'utilisateur et sélectionnés en fonction des besoins de l'enquête. Cette première liste permettra de récolter les 200 derniers tweets de chaque personne et identifier les comptes qu'elle a retweetés et mentionnés. Cette nouvelle liste constituera une première strate qui servira de base à une nouvelle collecte des 200 derniers tweets de ces comptes. L'utilisateur peut choisir le nombre de strates à effectuer afin d'agrandir le réseau à étudier.



Une fois les comptes collectés, les liens qu'ils entretiennent tous entre eux permettent de les relier et de calculer les groupes les plus connectés entre eux. L'outil propose une personnalisation de nombreux paramètres décrits dans la documentation. Ce pas-à-pas en vous guidant sur la marche à suivre vous permettra de commencer à l'utiliser.

L'illustration ci-dessous reprend les différentes étapes à effectuer, en cliquant sur chaque bloc, vous pourrez accéder aux explications.

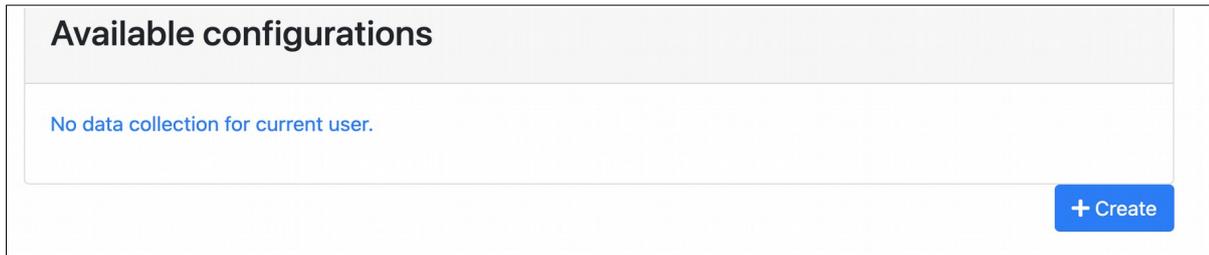
L'adresse pour la connexion est <https://lts2-imi-demo.epfl.ch>



1. Création du travail

[\[retour au menu\]](#)

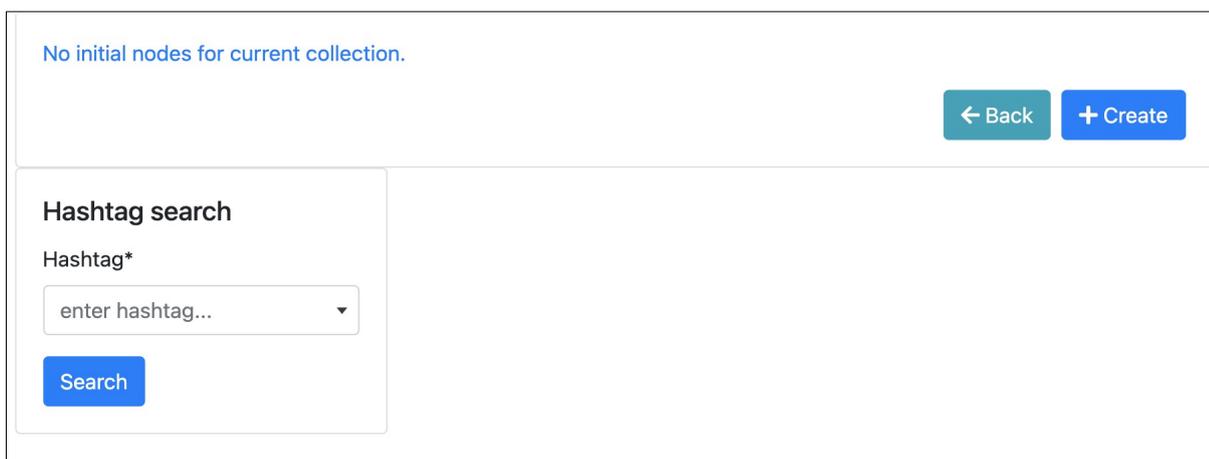
Lors de la première connexion, l'écran d'accueil est presque vide (c'est normal). Il faut d'abord créer une configuration pour la collecte avec le bouton "Create".



2. Comptes de départ

[\[retour au menu\]](#)

Sans compte(s) de départ défini(s), il n'est pas possible de démarrer la collecte, il faut en ajouter (même un petit nombre). En cliquant sur , la vue qui permet de définir ces comptes s'affiche alors:



Il est possible d'ajouter les comptes de manière manuelle (bouton "Create"), on peut alors ajouter les comptes Twitter un par un (NB: ne pas mettre le '@' initial):



Au lieu d'ajouter manuellement des comptes de départ, il est également possible de rechercher les "meilleurs" tweets avec un hashtag (comme dans la recherche avec Twitter, les résultats sont un mix entre les tweets les plus populaires et les plus récents).

Hashtag search

Hashtag*

#Topchef

Search

Add users

Users list

Save

Show 10 entries

#	User	Text	❤️	🗨️
1	ubereats_fr 55k 👤 768 ❤️ Mar 2019	On ne sait pas qui va remporter la grande finale de #TopChef ce soir mais une chose est sûre : on offre 10€ de créd... https://t.co/vFXjREsRTF Wed 9 Jun 2021	17	228
2	flogazan 183k 👤 8k ❤️ Apr 2009	Bravo à Mohamed pour sa victoire dans #TopChef ! Je fonce sur CNews pour le débrief de cette victoire 🙌 Wed 9 Jun 2021	782	49
3	M6	Ce soir, ne regardez pas #TopChef ! Allez au restaurant, et en rentrant, consommez @6play sans modération ! https://t.co/4vNXqEAYx7	1648	412

On peut alors depuis cette vue sélectionner les utilisateurs à ajouter à la liste des comptes de départ (). À la différence de l'ajout manuel, on peut sélectionner plusieurs utilisateurs à ajouter. Cliquer "save" pour ajouter les utilisateurs sélectionnés à la liste des comptes de départ. Il est également possible de suivre d'autres hashtags présents dans les tweets affichés.

3. Comptes à ignorer

[\[retour au menu\]](#)

Il est possible de définir une liste de comptes à toujours ignorer lors de la collecte. Certains comptes n'apportent que peu d'informations, tout en étant très souvent mentionnés. Par exemple les bots qui compilent les "threads" en une seule page ("threadreader"), ou encore les gros comptes de type "Youtube" vont se retrouver mentionnés très souvent, par des personnes très différentes n'ayant pas de rapport avec le sujet de la collecte.

L'ajout et la suppression de membres de cette liste noire fonctionnent sur le même principe que les comptes initiaux (excepté la partie relative aux hashtags).

4. Paramètres

[\[retour au menu\]](#)

La configuration a besoin de quelques paramètres de base

Name*

Name

Source*

Collection source

Depth*

Collection depth

Sampling probability*

Sampling probability (percent)

Collect by*

Collect by

[← Cancel](#) [Create](#)

- Un nom (pour facilement s’y référer plus tard)
- la source (pour l’instant seulement Twitter)
- la profondeur de collecte, i.e. le nombre de “strates” (ou “hops”) maximum depuis les points de départ. Attention à ce paramètre, la collecte peut prendre beaucoup de temps lorsqu’on l’augmente. Il est conseillé de partir avec une valeur pas trop élevée (2 ou 3) et d’augmenter progressivement (sachant que le nombre de noeuds visités va croître très rapidement)
- la probabilité d’échantillonnage: à chaque “saut”, on va regarder toutes les nouvelles connexions possibles (retweets et mentions) et suivre seulement une fraction (définies par ce paramètre). Cela permet de ne pas se retrouver avec un réseau énorme après un petit nombre de “sauts”. Une valeur entre 10 et 20% est généralement un bon choix. Elle permet de collecter les nœuds importants du réseau tout en explorant quand même d’autres parties de celui-ci.
- Le dernier paramètre définit la méthode de collecte, on peut le laisser à sa valeur par défaut pour faire simple.

5. Lancer la collecte

[\[retour au menu\]](#)

Une fois qu’il existe au moins une configuration de collecte, l’écran d’accueil ressemble à ce qui suit:

# ↑	Name ↕	Source	Depth	Sampling Proba. (%)	Collect by	Nb. initial nodes	
1	test		2	10	coreball	0	

Showing 1 to 1 of 1 entries

Previous **1** Next

Les icônes de la dernière colonne permettent de, de gauche à droite:

- 1 démarrer une collecte
- 2 voir les collectes en cours ou finies pour cette configuration
- 3 éditer les paramètres de la collecte

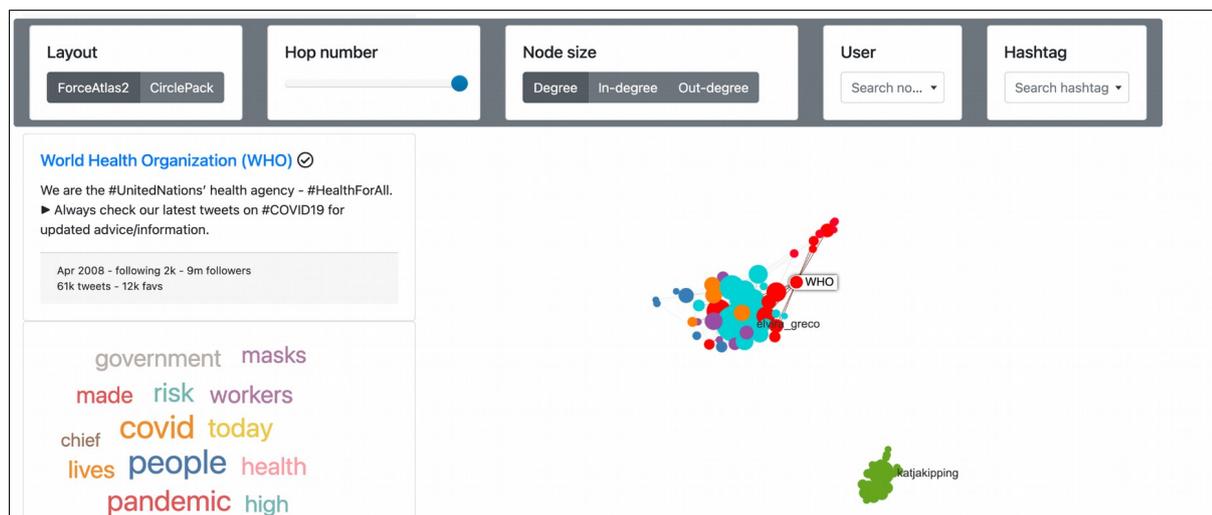
- 4 éditer les noeuds de départ
- 5 éditer la “liste noire”
- 6 supprimer la configuration (et toutes les collectes effectuées avec celle-ci)

6. Explorer

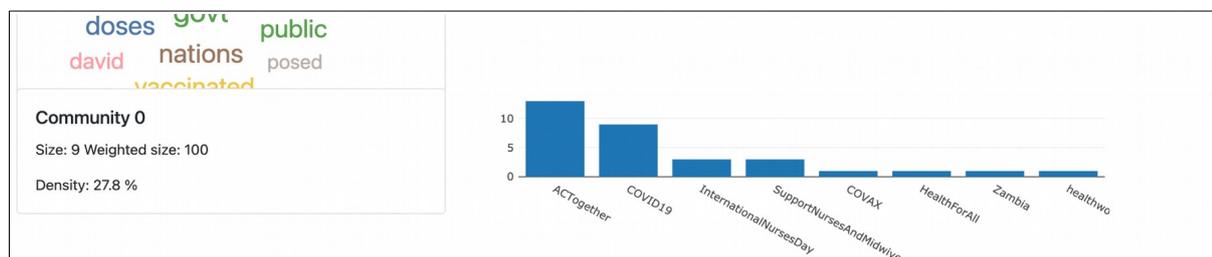
[\[retour au menu\]](#)

Une fois la collecte terminée avec succès, on peut visualiser et explorer le graphe . Les nœuds sont colorés en fonction de la communauté à laquelle ils appartiennent. Une communauté est un ensemble de nœuds fortement connectés entre eux et au moins connectés au reste du graphe.

- Choix du layout
 - “ForceAtlas2” les nœuds connectés s’attirent, ceux qui ne sont pas connectés se repoussent. Cela permet d’estimer la proximité des différentes communautés..
 - “CirclePack” organise les nœuds par communautés (et par degré). L’affichage de chaque communauté est plus clair, mais les connexions entre communautés sont moins claires.
- Filtrage par “hop”: permet de visualiser la progression de la construction du graphe
- Choix de la taille des noeuds:
 - “Degree”: taille proportionnelle au degré, i.e. au nombre total de connexions du noeud (entrantes + sortantes = mentions/RT + tweets)
 - “In degree”: taille proportionnelle au nombre de connexions entrantes (i.e. mentions et RT)
 - “Out degree”: taille proportionnelle au nombre de connexions sortantes (i.e. nombre de tweets)



Lorsqu’on clique sur un nœud, les informations du profil (vérifié, nombre de followers, date de création, etc.) s’affichent, ainsi que le lexique de la communauté à laquelle il appartient. Dans la partie inférieure (il faut parfois scroller suivant la taille de l’écran à disposition) on peut voir des statistiques sur la communauté et les hashtags employés par l’utilisateur.



7. Ajuster

[\[retour au menu\]](#)

Après avoir exploré le graphique, il se peut qu'il y ait certains éléments à affiner. Par exemple ignorer un compte qui prend trop de place dans ce graphe, ou encore ajouter une strate de plus.

Pour cela vous devrez cliquer sur l'onglet « Data collection » (en haut de la page) pour revenir au tableau de bord et pouvoir accéder à nouveau aux différents paramètres.

Une fois les éléments nécessaires modifiés, vous devrez relancer une collecte par l'onglet .