NDI

Analyse de données pour la surveillance des réseaux sociaux

Conseils sur les techniques, les outils et les méthodes de surveillance et d'analyse des réseaux sociaux

Mai 2020

Table des matières

Auteurs

Nick Monaco est le directeur du Digital Intelligence Lab (DigIntel) à l'Institute for the Future. Expert de la désinformation en ligne et de l'utilisation de robots politiques, il est spécialisé dans la désinformation chinoise. Au cours de son travail, il a consulté des décideurs et des technologues du gouvernement et de l'industrie sur la meilleure façon de lutter contre la désinformation et de maintenir l'intégrité des élections dans les pays du monde entier. Il a précédemment travaillé sur ces questions chez Graphika, une société d'analyse des réseaux sociaux et de renseignement sur les menaces, ainsi qu'au sein de Jigsaw, groupe de réflexion de Google sur les droits numériques. Il est également chercheur affilié au Computational Propaganda Project (ComProp) de l'Oxford Internet Institute.

Daniel Arnaudo est conseiller au NDI pour les stratégies d'information, couvrant l'intersection de la démocratie et de la technologie avec une responsabilité particulière pour élaborer des programmes de suivi de la désinformation et de promotion de l'intégrité de l'information dans le monde entier. Parallèlement, il est chercheur en cybersécurité à la Jackson School of International Studies de l'Université de Washington, où il a travaillé sur des projets au Brésil, en Birmanie et aux États-Unis. Récemment, il a également collaboré au Computational Propaganda Project de l'Oxford Internet Institute. Ses recherches portent sur les campagnes politiques en ligne, les droits numériques, la cybersécurité et les technologies de l'information et de la communication pour le développement.

À propos du NDI

Le National Democratic Institute (NDI) est une organisation non gouvernementale, non partisane et à but non lucratif qui répond aux aspirations des peuples du monde entier à vivre dans des sociétés démocratiques qui reconnaissent et promeuvent les droits humains fondamentaux.

Depuis l'ouverture de ses portes en 1983, le NDI et ses partenaires locaux ont travaillé pour soutenir et renforcer les institutions et les pratiques démocratiques en renforçant les partis politiques, les organisations civiques et les parlements, en protégeant les élections et en promouvant la participation des citoyens, l'ouverture et la responsabilité du gouvernement.

Avec des membres du personnel et des professionnels politiques bénévoles de plus de 100 pays, le NDI rassemble des individus et des groupes pour partager des idées, des connaissances, des expériences et des compétences. Les partenaires bénéficient d'une large exposition aux meilleures pratiques en matière de développement démocratique international, qui peuvent être adaptées aux besoins de leurs propres pays. L'approche multinationale du NDI renforce le fait que, s'il n'y a pas de modèle démocratique unique, certains principes fondamentaux sont partagés par toutes les démocraties.

Le travail de l'Institut respecte les principes inscrits dans la Déclaration universelle des droits de l'homme. Il favorise également le développement de canaux de communication institutionnalisés entre les citoyens, les institutions politiques et les élus, et renforce leur capacité à améliorer la qualité de vie de tous les citoyens. Pour plus d'informations sur le NDI, veuillez visiter www.ndi.org.

Remerciements

L'Institut remercie le National Endowment for Democracy (NED) pour son soutien à la création de ce guide. Le NED est une fondation privée à but non lucratif dédiée à la croissance et au renforcement des institutions démocratiques dans le monde. Chaque année, le NED accorde plus de 1 600 subventions pour soutenir les projets de groupes non gouvernementaux à l'étranger qui œuvrent pour des objectifs démocratiques dans plus de 90 pays. Depuis sa fondation en 1983, la fondation est restée à la tête des luttes démocratiques partout dans le monde, tout en évoluant en une institution aux multiples facettes qui est une plaque tournante d'activités, de ressources et d'échanges intellectuels pour les militants, les professionnels et les universitaires de la démocratie du monde entier.

Il remercie également l'Institute for the future (IFTF) pour sa coopération et son partenariat dans la diffusion de ce guide. L'IFTF est une organisation à but non lucratif consacrée à la conclusion de contrats à terme civiques. L'IFTF est la première organisation mondiale de contrats à terme. Depuis plus de 50 ans, les entreprises, les gouvernements et les organisations à impact social s'appuient sur les prévisions mondiales, de la recherche personnalisée et de la formation prospective de l'IFTF pour faire face aux changements complexes et élaborer des stratégies d'envergure mondiale. Les méthodes et les ensembles d'outils de l'IFTF produisent des vues cohérentes des possibilités de transformation dans tous les secteurs qui, ensemble, soutiennent un avenir plus durable.

Conception et impression : Ironmark, 2020

Introduction

Les réseaux sociaux sont devenus une partie de plus en plus importante des conversations que les citoyens, les candidats, les partis et les organisations connexes engagent pour des événements politiques, des élections, des référendums ainsi que des votes sur des projets de loi, des grèves et d'autres formes d'activité politique. Il est essentiel que les chercheurs, les observateurs électoraux, les organisations de la société civile et les citoyens ordinaires se dotent d'outils, de méthodes et de pratiques pour aider à la collecte et à l'analyse des données de l'espace en ligne. Des membres d'organisations internationales de la société civile, notamment des chercheurs, des directeurs de programmes, des militants et d'autres intervenants, participent à divers programmes à l'échelle internationale, notamment dans le cadre de missions d'observation électorale, pour aider les groupes locaux à développer leur propre capacité de surveillance, ou observer les discours de haine, les tendances politiques et de nombreux autres sujets.

Ce guide a pour but d'aider les chercheurs, les observateurs électoraux, les technologues et autres parties prenantes à comprendre les meilleures pratiques ainsi que les outils et les méthodes pour développer l'observation et la surveillance en ligne des réseaux sociaux. Il présente une introduction aux concepts pertinents à comprendre lors de l'étude de ces questions, ainsi qu'un examen de la façon de construire une image complète du contexte sociotechnique dans un pays ou une région, y compris la présence en ligne des partis locaux, les réseaux sociaux et les taux de pénétration d'Internet, les médias locaux, les divisions ethniques et religieuses et de nombreux autres facteurs qui se manifestent dans l'espace en ligne.

Cette ressource contient des informations sur les sujets clés suivants :

- Collaboration : Les chercheurs doivent envisager les partenaires potentiels et les moyens de les choisir, qu'il s'agisse de divers types d'organisations locales, d'ONG internationales ayant une expertise dans le domaine ou d'entreprises privées allant des plus petites ayant une expertise dans le domaine aux grandes sociétés multinationales qui contrôlent les plateformes de réseaux sociaux et d'autres technologies qui atteignent les réseaux mondiaux. Une section de ce guide passera en revue les options et les considérations potentielles pour ces collaborations, en citant des exemples et en indiquant les avantages et les risques de travailler avec différents groupes.
- Méthodologie : En ce qui concerne la méthodologie, le guide examine différentes méthodes de collecte de données, y compris les considérations relatives aux différentes plateformes, les méthodes de travail avec les interfaces de programmation d'applications (API) de chacune et les différentes méthodes d'extraction de contenu.
- Cartographie et visualisation : Cette section contient des conseils sur l'élaboration et la lecture de cartes de réseaux. Le guide présente les termes techniques clés ainsi que des méthodes pour construire des cartes de l'espace en ligne, des exemples de recherche sur le terrain et les limites potentielles des cartes elles-mêmes.
- Analyse : Cette section couvre l'analyse des différents types d'entités et d'individus qui participent à la conversation. Les conseils sur ce sujet incluent un aperçu des comptes individuels, des influenceurs et des groupes, ainsi que de leurs rôles dans l'écosystème en ligne, et abordent également les façons dont les organes de presse et d'autres ressources externes deviennent d'importantes sources de contenu.
- Contenu : Quant au contenu en ligne lui-même, le guide examine différents aspects des publications, des tweets et d'autres formes de réseaux sociaux. Cette section décrit comment détecter différents types de propagande informatique en réseau, allant des botnets et des fermes à trolls à d'autres formes potentielles de manipulation. Elle examine diverses formes malveillantes de contenu manipulateur allant de la désinformation au discours de haine, ainsi que leurs cibles potentielles.
- Outils : Pour soutenir ces techniques de recherche, le guide répertorie et examine les différents types d'outils utiles à l'élaboration d'analyses pour divers aspects de la surveillance des réseaux sociaux. Cet inventaire

comprend des outils de collecte sur diverses plateformes, ainsi que des ressources pour l'analyse de réseaux, la visualisation de données et la recherche de renseignements open source (d'origine sources ouvertes).

- Réponses : Enfin, le guide présente un examen des réponses potentielles qui peuvent être éclairées et enrichies par ces analyses. Cela implique de collecter des données pour la recherche, d'élaboration de la documentation et de créer des mécanismes de signalement en collaboration avec les plateformes, les régulateurs gouvernementaux et les organisations de surveillance des élections. Le guide se termine par des recommandations pour développer la recherche dans des domaines futurs et élaborer des évaluations critiques des domaines du secteur en évolution.

Contexte

Lorsqu'ils pénètrent dans un nouvel environnement, les analystes doivent considérer de nombreux acteurs, réseaux et groupes sociopolitiques, ainsi que les systèmes techniques. Ils doivent examiner divers aspects des environnements informationnels, sociaux et politiques dans lesquels ils travaillent, en tenant compte des réseaux et de l'organisation du pays lui-même, de la région au sens large et de sa place dans le système mondial. Les informations ne passent pas toujours par les réseaux de médias en ligne ou traditionnels ; elles circulent en grande partie par le biais du bouche à oreille, de rumeurs, des médias traditionnels et d'autres méthodes. Cependant, ces discussions se retrouvent souvent en ligne, où elles peuvent être mieux suivies et comprises.

À l'échelle internationale, divers groupes emploient des tactiques pour manipuler la perception du public sur les candidats et les problèmes, affaiblir la confiance dans les processus démocratiques et créer une confusion chez les électeurs au sujet des lieux de scrutin, de leur inscription ou du système électoral lui-même. L'un des principaux objectifs de ce type de recherche est d'aider à comprendre comment les réseaux de propagande informatique[1] fonctionnent en ligne, pour contribuer à exposer leur fonctionnement, documenter des cas pour la recherche et alerter potentiellement les autorités ou les sociétés de réseaux sociaux contre la manipulation et les abus en ligne. Détecter l'automatisation, les faux comptes, les faux contenus, les mauvaises sources d'information et autres manipulations pourraient tous être des objectifs de ce genre de projet pour décrire la propagande informatique dans différents cas.

Les analystes des réseaux sociaux qui étudient les questions relatives à l'information dans des contextes du monde entier doivent prendre en compte de nombreux facteurs lorsqu'ils élaborent des rapports. Ils doivent déployer des outils, rechercher les lois et les institutions régissant l'espace informationnel et utiliser des méthodes hors ligne, telles que des entretiens avec des représentants du gouvernement, des partis, des médias et des candidats, pour dresser le tableau le plus précis possible de la désinformation dans l'environnement électoral.

La désinformation est un sujet difficile à appréhender car ce n'est généralement pas un terme très bien défini. Un élément clé de la désinformation est le concept d'intention, en ce sens que la désinformation est transmise avec l'intention de tromper, tandis que la fausse information est un contenu incorrect sans l'intention nécessaire de falsifier. Pour en savoir plus sur les définitions et sur le contexte de ces sujets, voir le rapport Data and Society Lexicon of Lies et Information Disorder de First Draft, ainsi que le guide du NDI intitulé Supporting Information Integrity and Civil Political Discourse[2], qui a été traduit en albanais, arabe, anglais, français, russe, serbe et espagnol. D'autres sources sont indiquées dans la section des références. Outre la désinformation, un chercheur doit prendre en compte de nombreux types de contenus différents, certains inoffensifs, d'autres malveillants ou bien encore positifs.

1 Selon l'Oxford Internet Institute, dont le groupe de recherche a aidé à définir le terme et a lancé une grande partie de la recherche autour de ces questions : « La propagande informatique est l'utilisation d'algorithmes, d'automatisation et de curation humaine pour diffuser délibérément des informations trompeuses sur les réseaux de réseaux » (Woolley et Howard, 2017, 4).

2 Voir https://www.ndi.org/publications/supporting-information-integrity-and-civil-physical-discourse

Il est également important de tenir compte du rôle des médias et de la manière dont ils opèrent et influencent les systèmes de médias en ligne et de réseaux sociaux. Examinez les principales sources de journaux imprimés, radiophoniques et télévisés et évaluez leurs parts de marché, leurs liens avec les partis politiques, la participation des secteurs public et privé et leur influence sur l'espace en ligne. Beaucoup ont une présence en ligne considérable, et en particulier lorsqu'elles sont liées à un grand parti politique, peuvent avoir une énorme influence sur la ligne éditoriale et les préférences politiques d'une organisation, ainsi que sur la propension à obtenir un soutien viral, organique ou généré artificiellement. La polarisation dans ce contexte peut être considérée comme critique par rapport aux réseaux, elle peut devenir facilement apparente lorsque des groupes distincts plus importants se forment dans l'opposition et en coalition, ce qui peut également signaler les niveaux de propagande informatique et de désinformation présents.

Le discours de haine comprend souvent de la désinformation pour dénigrer les cibles des campagnes avec de fausses attaques. Cela cible particulièrement les femmes, les minorités et les autres groupes vulnérables. Le travail de recherche de l'équipe Gender, Women and Democracy (Genre, Femmes et Démocratie) du NDI sur la violence à l'égard des femmes en politique (VCF-P) note que « lorsque les attaques contre les femmes actives en politique sont canalisées en ligne, la portée étendue des plateformes de réseaux sociaux amplifie les effets de la violence psychologique en donnant l'impression que ces effets sont anonymes, sans frontières et durables, sapant ainsi le sentiment de sécurité personnelle des femmes d'une manière inconnue des hommes. La plupart des acteurs étatiques et non étatiques qui commettent des actes de VCF-P en ligne se mobilisent à travers des réseaux transnationaux. L'utilisation abusive, par les États, les organisations et les individus, des libertés mêmes que l'espace de l'information est censé permettre de s'exprimer, est devenue l'une des plus grandes menaces à son intégrité » (Zeiter et al., 2019, 4). En conséquence, les chercheurs doivent savoir à quel point les femmes sont particulièrement vulnérables à ces attaques en ligne et envisager des moyens de documenter et de signaler ce type d'abus aux plateformes. Le NDI a piloté l'utilisation de lexiques de termes de discours de haine pour étudier les réseaux sociaux avec des études de cas en Colombie, en Indonésie et au Kenya. Les méthodes d'élaboration de ces lexiques et les études de cas qui sont détaillées dans le rapport de cette recherche « Tweets that Chill : Analyzing Violence Against Women in Politics » (Des tweets qui font froid dans le dos : analyser la violence contre les femmes en politique).

Lors de l'élaboration de stratégies de collecte de données, réfléchissez à la manière dont ces opérations d'information utilisant le discours de haine et la propagande informatique fonctionnent dans le contexte dans lequel vous travaillez. Le contenu ultérieur de ce guide aidera les analystes à apprendre à identifier ces campagnes, ces réseaux et utilisateurs en ligne, mais les chercheurs doivent comprendre globalement quel type de propagande informatique, de discours préjudiciable ou d'autres schémas ils recherchent avant de travailler.

Il est également important de bien comprendre à la fois l'environnement réglementaire local (y compris la réglementation du financement des élections et des campagnes, qui peut aider à éclairer les rapports), ainsi que les règles de modération du contenu et les politiques qui régissent les plateformes étudiées. C'est ainsi que les chercheurs peuvent concevoir leurs études de manière légale et éthique, et alerter les entreprises, et potentiellement les gouvernements, des abus et autres actions illégales et négatives en ligne. Cela aide également les chercheurs à mieux comprendre les pays qu'ils étudient.

La compréhension des conditions d'utilisation des différentes sociétés de réseaux sociaux représente un aspect essentiel de ce rapport. Les normes clés à comprendre pour Facebook, Twitter et YouTube avec des liens vers les politiques sont répertoriées ci-dessous.

Normes communautaires de Facebook	• Informations substantielles sur le signalement de différents scénarios, ici. • Signaler la page d'un imposteur d'un personnage public , ici. • Comment signaler des choses » détaillé selon les types de post, ici. • Comment marquer un post comme étant une fausse nouvelle ? ici.
Règles de Twitter	• Aperçu général des signalements de violations ici • Comment signaler un tweet, un compte abusif ou un message individuel ici • Signaler un compte d'usurpation d'identité, ici • Signaler des instructions de spam, ici
Directives de la communauté YouTube	• Comment signaler un contenu inapproprié, détaillé selon le type de post, ici • « Signaler une prédiction de recherche YouTube , ici • « Autres options de signalement , ici • « Comment signaler des spams ou des contenus trompeurs , ici (fin de page) • Outil de signalement YouTube, ici

Certaines entreprises, telles que Facebook, exigent que les utilisateurs s'identifient avec des informations réelles, de telle sorte que le simple fait d'identifier un compte qui n'est pas une personne réelle, ou un groupe connecté à ce faux compte peut entraîner un retrait. D'autres, comme Twitter, n'interdisent pas l'anonymat mais ont des interdictions contre les discours de haine ou l'amplification artificielle qui peuvent être identifiés et signalés par la recherche. Il est important de vous familiariser avec les différents codes et mécanismes de signalement de contenu, qui sont décrits dans le tableau ci-dessus.

Le signalement est important, tout comme la documentation des campagnes, des comptes et des contenus pertinents afin que les signalements puissent être vérifiés. Pensez à utiliser des systèmes qui peuvent être facilement sauvegardés et consultés, en termes d'annotation et de flux de travail. Comme indiqué, les normes communautaires de Facebook offrent des mécanismes pour signaler les faux comptes et les formes de contenu négatives, et pour demander aux vérificateurs de faits de vérifier et potentiellement de supprimer les contenus préjudiciables. Les règles de Twitter prévoient différentes formes de signalement pour l'usurpation d'identité, le spam et certaines formes de discours haineux ou autrement interdits. Les politiques de YouTube sont axées sur les médias et les droits d'auteur, ainsi que sur les formes de discours explicites, haineuses ou les autres formes de discours préjudiciables. En tant que propriétaires de YouTube, les termes de référence de Google vous donnent une idée de la manière dont vous pouvez non seulement signaler des comptes, mais également enquêter sur les comptes liés au service de streaming (diffusion) vidéo.

En termes de réglementations gouvernementales, les analystes doivent tenir compte des lois sur la protection des données telles que le Règlement général sur la protection des données de l'Union européenne, qui contient des aspects couvrant la collecte de renseignements permettant d'identifier une personne en Europe ou sur les Européens se trouvant partout dans le monde, ainsi que sur les entreprises qui y opèrent.[3] La collecte de données auprès de groupes, d'utilisateurs et de réseaux privés pourrait faire exception à ces lois ainsi que les conditions d'utilisation des plateformes.

Des réseaux tels que la Design 4 Democracy Coalition (D4D Coalition) peuvent aider à défendre les principes démocratiques dans les entreprises technologiques, par exemple, pour attirer l'attention des plateformes sur des campagnes d'influence à grande échelle à des fins de propagande électorale, de discours de haine ou autres. La D4D Coalition est composée d'organisations internationales et nationales de la société civile s'engageant avec des

3 https://gdpr.eu/

entreprises technologiques pour intégrer et soutenir les principes démocratiques, y compris dans le contexte de la modération de contenu, de l'élaboration de politiques et des considérations relatives aux produits. La Coalition (sous la direction du NDI, de l'International Republican Institute, de la Fondation internationale pour les systèmes électoraux et d'International IDEA) relie les acteurs de la société civile et de la démocratie dans divers contextes avec bon nombre des entreprises technologiques les plus influentes (y compris Facebook, Microsoft et Twitter) pour encourager le partage de l'information et faire progresser les stratégies de promotion de l'intégrité de l'information et de protection des processus démocratiques. De tels efforts peuvent être soutenus et améliorés par une documentation, une analyse et des rapports solides à l'aide des outils, méthodes et tactiques décrits ici.

Pour plus de détails sur l'élaboration de stratégies d'analyse des données sur les réseaux sociaux lors des élections, consultez le document d'orientation du NDI sur la désinformation et l'intégrité électorale[4], ainsi que le Guide de Supporting Democracy pour la société civile sur la surveillance des réseaux sociaux pour les élections.[5] Ces guides abordent les méthodologies et les considérations réglementaires pour les observateurs en termes de surveillance des réseaux sociaux, ainsi que la possibilité d'intégrer les données collectées en ligne dans des missions d'observation électorale traditionnelles.

Travailler avec la collecte de données

La collecte de données est la première étape d'une analyse rigoureuse de l'activité en ligne autour d'une élection. En tant qu'analyste sur le terrain, la première étape de la collecte de données est une enquête sur le paysage des réseaux sociaux/en ligne dans la région que vous observez. Parmi les questions importantes à envisager, on peut citer :

- Quelles sont les plateformes les plus populaires dans la région ? Quels sont les différents taux de pénétration des plateformes ? Internet World Stats, l'Union internationale des télécommunications, le rapport Freedom on the Net de Freedom House ou Facebook lui-même peuvent être de bonnes ressources sur ce point. Un pays avec des taux d'engagement élevés sur Facebook mais des taux de pénétration faibles sur Twitter (comme Taïwan) produirait probablement les informations les plus précieuses sur Facebook.
- Quels sont les sites Web populaires pour les actualités ? Lesquels d'entre eux sont des médias traditionnels ? Lesquels ont été créés plus récemment ?
- Quels hashtags sont les plus pertinents pour l'élection en question ? De même, quels comptes officiels représentent les partis, les candidats et leurs campagnes dans ces élections ? Souvent, un candidat aura plus d'un compte ou d'une page concernant une élection, comme un compte Twitter personnel et un compte Twitter de campagne officiel. En dresser la liste est une bonne première étape pour collecter des données pertinentes.

Une fois que vous avez identifié les plateformes et les sites Web de réseaux sociaux les plus importants du pays, vous êtes prêt.e à commencer à collecter des données sur les plateformes et les sites Web pertinents. Dans cette section, nous examinerons les options dont disposent les chercheurs pour la collecte de données.

Méthodes de collecte

Lors de la collecte de données sur les réseaux sociaux en ligne, les chercheurs disposent de plusieurs options. Les trois principales options sont l'utilisation d'outils de collecte tiers, l'interaction directe avec l'API d'une plateforme ou le recours au « web scraping » (extraction de contenu de sites Web, ou grattage web). Dans cette section, nous allons explorer les détails de ces trois méthodes et les différences qui existent entre elles.

4 https://www.ndi.org/publications/disinformation-and-electoral-integrity-guidance-document-ndi-elections-programs

5 Supporting Democracy est mis en œuvre par un consortium composé de SOFRECO, Democracy Reporting International (DRI) et le NDI. https://www.ndi.org/sites/default/files/social-media-DEF.pdf

Outils tiers (accès indirect)

Pour les chercheurs qui ont des délais serrés ou qui n'ont pas la capacité d'interagir avec des sites Web via du code informatique, les outils tiers peuvent être une option utile. Ces outils interagissent généralement avec les API d'une ou plusieurs plateformes cibles de manière invisible en arrière-plan et présentent les données dans une présentation graphique facile à utiliser telle qu'un tableau de bord. Pour Facebook, qui n'autorise actuellement pas les chercheurs externes ou les entreprises à utiliser son API, CrowdTangle est l'une des meilleures options pour surveiller la portée des pages, des groupes et des URL sur Facebook, et montre également aux utilisateurs la portée des URL sur Twitter, Instagram et Reddit. L'extension CrowdTangle[6] permet à un utilisateur de voir une estimation en temps réel du nombre de réactions qu'une publication, une page ou une URL a suscitées sur ces plateformes, ce qui peut être un moyen utile de surveiller les tendances du contenu sur plusieurs plateformes au quotidien.

Certains outils tiers, tels que Sysomos et Brandwatch, proposent des abonnements payants coûteux, mais offrent un accès à une grande quantité de données qui peuvent être utiles pour surveiller les campagnes de hashtag et d'autres contenus tendances sur les réseaux sociaux.

En plus d'être visualisées sur un navigateur Web, les données de ces outils peuvent souvent être exportées vers un fichier lisible par une machine, tel qu'un fichier de valeurs séparées par des virgules (CSV), qui peut à son tour être manipulé par un data scientist (expert en mégadonnées) pour interroger les données de manière nouvelle et utile.

Facebook : Sysomos, Brandwatch

Twitter : Twitonomy

Accès direct - API et web scraping : Quelle est la différence ?

Pour une interaction plus directe avec les plateformes et les sites Web, les chercheurs ont deux possibilités alternatives : utiliser une interface de programme d'application (API) ou collecter les informations directement à partir du code source de la page Web, une pratique connue sous le nom de « web scraping ». Il est important de comprendre la différence entre les API et le Web scraping : l'extraction de données à partir d'API est dans la plupart des cas licite et éthique, car les données des API sont délibérément régulées par les plateformes et structurées de manière à ne pas violer les droits des utilisateurs. Le web scraping est, dans de nombreux cas, une violation des conditions d'utilisation et est plus difficile à réguler. Dans de nombreux cas, le web scraping peut être illicite.

Il est important de comprendre la différence entre ces deux méthodes de collecte de données. Il est également important de savoir que vous entendrez parfois des chercheurs eux-mêmes se référer à tort à des données extraites d'une API comme étant « grattées ». La distinction est importante, non seulement pour des raisons pratiques telles que le gain de temps et d'efforts, mais également au regard des violations éthiques et légales qui peuvent résulter du web scraping.

En gardant à l'esprit cette analogie de base de la différence entre les deux approches, plongeons-nous dans les détails des API et du web scraping.

API

Pour les chercheurs intéressés par une approche plus pratique de la collecte de données, de nombreuses plateformes ont une forme d'accès plus direct aux données sous la forme d'interfaces de programmation d'applications (API). De manière générale, les API sont un moyen pour les utilisateurs d'interagir facilement avec

6 CrowdTangle est actuellement disponible pour les universitaires et les chercheurs sur une base sélective. Vous et votre équipe pouvez postuler sur CrowdTangle.com. La version complète comprend des données actuelles et historiques sur Facebook et Instagram. Une extension gratuite CrowdTangle est également disponible sur le site Web. Cette extension prend une URL comme entrée et propose les 500 publications publiques les plus récentes qui ont suscité une traction significative en citant l'URL sur Facebook, Instagram, Reddit et Twitter. La plateforme complète et l'extension CrowdTangle peuvent être utiles pour les enquêtes.

un site Web ou une plateforme de réseaux sociaux via un code informatique. Cela permet à un utilisateur de traiter rapidement beaucoup plus de données, et de générer à son tour des informations plus approfondies sur l'activité en ligne qu'il ne pourrait autrement le faire manuellement.

Types d'API

Il est important de connaître deux caractéristiques de l'API avant de commencer la collecte des données : l'ouverture et la durée.

- Ouverture : Les API se présentent sous différentes formes : les API ouvertes permettent à toute personne de collecter des données, tandis que les API authentifiées exigent qu'un utilisateur se soumette à une certaine forme de vérification avant d'autoriser la collecte de données.
  - API ouvertes : Venmo, l'application utilisée pour les paiements électroniques aux États-Unis, dispose d'une API publique qui permet à quiconque d'afficher un certain nombre des transactions publiques les plus récentes sur l'application. Vous pouvez trouver des listes d'API ouvertes partout sur Internet, comme cette liste sur Github. Any-api.com dispose également d'une liste de plusieurs API accessibles au public pour les utilisateurs intéressés, dont beaucoup sont ouvertes.
  - API authentifiées : La plupart des API avec lesquelles vous traiterez pour la collecte de données sur les réseaux sociaux (Twitter, Reddit, etc.) nécessitent qu'un utilisateur s'authentifie avant de pouvoir commencer à collecter des données.
- Durée : Les sites web et les plateformes structurent aussi généralement leurs API de manière différente selon le moment.
  - Collecte de données historiques : La plupart des API vous permettent de collecter une certaine forme de données historiques sur leurs sites - Twitter et Reddit le permettent notamment. Cette forme d'API, que nous appellerons API historique, vous permet d'extraire rétroactivement les données qui ont été générées avant que vous ne fassiez la requête. Il est important de noter que les données publiées après que vous ayez effectué la requête ne sont pas disponibles pour la collecte.
  - Streaming de données en temps réel : L'ingestion et le téléchargement de données telles qu'elles se produisent en temps réel est un processus appelé streaming. Si Twitter est une plateforme pertinente pour l'élection ou la période que vous prévoyez de surveiller, le streaming est probablement votre meilleure option pour la collecte de données. Lors du streaming de données, vous collectez des Tweets en temps réel en fonction d'une requête spécifiée (par exemple, tous les tweets utilisant le hashtag #election, ou tous les tweets mentionnant des comptes d'intérêt, citant des URL d'intérêt, etc.)

Limitations à la collecte imposées par les API

Dans l'intérêt de préserver la confidentialité et la sécurité des utilisateurs, la plupart des plateformes de réseaux sociaux limitent la quantité de données qu'un utilisateur est autorisé à collecter. Nous explorons ici quelques-unes de ces limites pour vous familiariser avec les problèmes que vous pouvez rencontrer lors de la collecte de données.

- Limites de volume : La plupart des API limitent le volume de données qu'un utilisateur donné est autorisé à collecter. Par exemple, lors de la diffusion de données en temps réel sur Twitter, la plateforme limite la quantité de données qu'un utilisateur peut collecter à 1 % des données mondiales de diffusion en continu.
  - Limites de débit : Les limites de débit sont le type de limite de volume le plus courant que vous rencontrerez lors de l'extraction de données d'API. La plupart des API ont des limites de débit pour garantir qu'un seul utilisateur ou une seule application ne peut pas télécharger une quantité excessive de données (comme défini par la plateforme ou le site Web en question). Par exemple, Twitter limite le nombre de Tweets qu'un seul utilisateur peut télécharger dans un laps de temps de 15 minutes.
- Limite constituée par les données supprimées : Comme évoqué brièvement ci-dessus, les plateformes ont tendance à supprimer le contenu qui enfreint les règles et réglementations d'utilisation spécifiées - ces

réglementations ont plusieurs noms différents (normes communautaires, conditions d'utilisation, etc.). Cette remarque est particulièrement importante pour les contextes électoraux dans lesquels un comportement néfaste est susceptible de se produire - en particulier, le streaming de données sur Twitter vous permet de collecter et de conserver des données sur des acteurs néfastes en temps réel qui pourraient être supprimées de la plateforme ultérieurement. Une fois supprimées, les données sur ces acteurs ne sont pas disponibles. Si votre équipe souhaite capturer les mauvais acteurs, la désinformation et d'autres contenus pour une analyse ultérieure, le streaming en temps réel maximise vos chances d'y parvenir.

- Limitations du type de données : De même, la plupart des plateformes limitent le type de données qu'un utilisateur peut collecter depuis la plateforme. L'API de Twitter vous permettra de collecter certaines informations sur un utilisateur cible (telles que la quantité de ses posts, son quantité d'abonnés/followers et la date de création de son compte), mais elle ne vous permettra pas d'accéder à d'autres types de données auxquelles l'accès est restreint (telles l'adresse IP la plus fréquemment utilisée par l'utilisateur). Facebook ne permet actuellement pas du tout aux chercheurs de recueillir des informations sur d'autres utilisateurs ou d'autres pages via son API. Il est utile de se familiariser avec le type de données pouvant être collectées sur les plateformes cibles lors de la conception d'un projet de recherche.
- Limites de temps des données : Twitter limite la collecte de données historiques par les utilisateurs, via l'API de recherche Twitter, aux données générées pendant les 7 à 9 derniers jours. Les données datant de plus de 7 à 9 jours sur Twitter peuvent être achetées auprès de fournisseurs de données tels que GNIP, mais ne peuvent pas être collectées via un accès API standard.

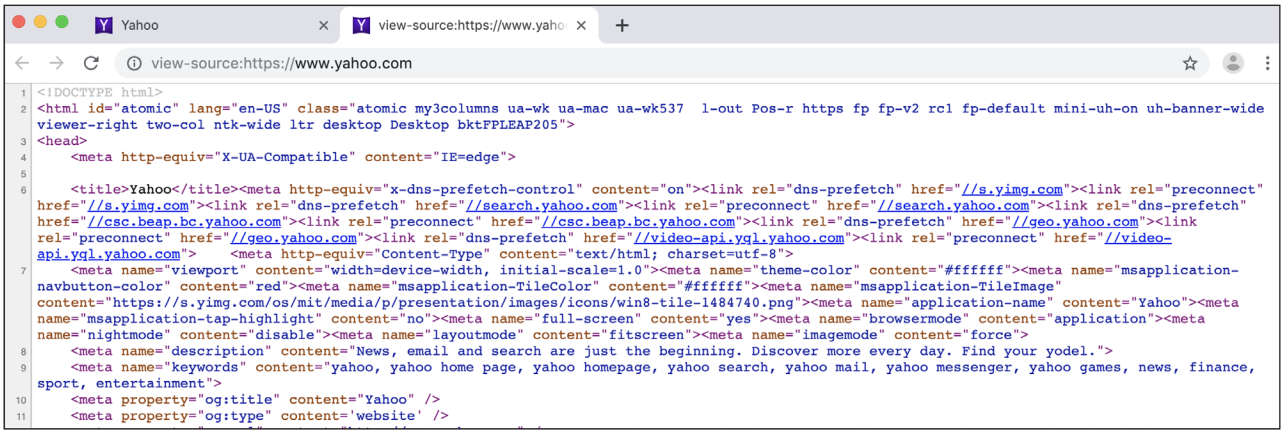
Un aspect intéressant des API est qu'elles ne sont pas spécifiques à un langage. Les données d'API peuvent être extraites à l'aide de Java, Python, R, Ruby, Perl ou de tout autre langage de programmation que vous ou votre équipe technique préférez. Bien que ce soit le cas, il existe des modules spécifiques dans des langages de programmation courants qui simplifient le processus d'interrogation des API en prenant en charge à votre place certaines des difficultés.

Un exemple de code avec des explications montrant comment utiliser le module rtweet de R pour collecter des données à partir de l'API de recherche et de streaming de Twitter est disponible à la fin de ce guide de terrain dans l'annexe : Exemple de code API - Collecte de données à partir d'API de recherche et de diffusion de Twitter avec le module Rtweet.

Web scraping

Bien que les API offrent un moyen simplifié de collecter des données d'une plateforme ou d'un service, elles ne sont pas la seule option pour collecter des données. Le Web scraping (grattage Web) est le processus d'extraction du code source d'un site Web cible et d'extraction des données pertinentes. Chaque page sur Internet est le résultat du code source qui la compose - HTML et autres langages Web dynamiques, scripts, liens hypertexte et sources multimédias. Le processus par lequel un navigateur prend le texte du code et le transforme en une page Web visuelle et interactive est appelé rendu. Vous pouvez afficher le code source sous-jacent de n'importe quelle page Web dans la plupart des navigateurs modernes - Google Chrome, Mozilla Firefox, Safari, Brave et Opera ont tous cette fonctionnalité.

Par exemple, lorsque vous utilisez Google Chrome, vous pouvez cliquer avec le bouton droit sur n'importe quelle page Web chargée (ou « rendue ») dans votre navigateur et cliquer sur l'option « Afficher la source de la page ». Chrome ouvrira un nouvel onglet qui affichera le code HTML et CSS utilisé pour charger la page Web que vous consultez. Vous trouverez ci-dessous un exemple tiré de yahoo.com.



Capture d'écran du code source HTML et CSS sous-jacent à yahoo.com (instantané pris début août 2019). Google Chrome et d'autres navigateurs modernes permettent aux utilisateurs d'afficher le code source de tout site Web qu'ils visitent - ce code source est ce qui est récupéré lorsqu'un site Web est « gratté ».

Préoccupations éthiques lors du Web scraping

Le grattage Web peut être un outil utile pour analyser les pages Web. Il convient cependant de noter que le grattage Web est une violation des conditions d'utilisation de la plupart des plateformes de réseaux sociaux et qu'il peut être facile d'enfreindre la loi lors du grattage de pages Web. Pour cette raison, il est important de toujours tenir compte de la vie privée des utilisateurs, des conditions générales et des lois pertinentes lors du grattage de sites Web. Il est toujours important de vous assurer que votre processus de collecte de données est à la fois éthique et licite. De nombreux chercheurs ne collectent des données que via des API et choisissent de ne pas gratter les sites Web pour ces raisons.

Lorsque vous utilisez des données collectées par d'autres équipes ou d'autres outils (scénario fréquent pour de nombreuses équipes), il est également primordial de s'assurer que les données avec lesquelles vous travaillez ont été obtenues de manière éthique. Par exemple, si les données ont été obtenues illégalement par grattage ou piratage, il est déconseillé de les utiliser pour un projet de recherche pour des raisons juridiques, éthiques et politiques, et ce type de projet peut entraîner des répercussions tant de la part des autorités gouvernementales que des entreprises. Ces considérations sont importantes à prendre en compte avant de commencer l'analyse des données.

Différences des plateformes : Sommaire

	API historique disponible	API de streaming disponible	Outils de collecte de données tiers
Twitter	Oui	Oui	CrowdTangle (extension)
Facebook	Non	Non	CrowdTangle
Gab	Oui, via Pushshift.io	Non	Non
Instagram	Non	Non	CrowdTangle
Reddit	Oui	Oui (certains modules ont cette fonctionnalité - par exemple Reddit SSE)	CrowdTangle (extension) Pushshift.io
YouTube	Oui	Non	
Telegram	Oui	Non	Telethon Pushshift.io
VKontakte	Oui	Non	
WhatsApp	Oui - l'API WhatsApp business permet des communications automatisées depuis les entreprises vers les clients. Elle n'est généralement pas utilisée pour le type d'analyse dont nous discutons dans ce guide.	Non	Plusieurs outils tiers permettent une analyse statistique ou une visualisation des chats WhatsApp[7]. ChatAnalyzer WhatsApp Chat Analyzer Chatilyzer WhatsAnalyzer
VKontakte	Oui	Non	

Outils utiles pour la collecte de données dans un contexte d'intégrité électorale :

- Modules pour de l'API de Twitter :
  - Modules R : https://github.com/ropensci/rtweet, twitteR
  - Modules Python : python-twitter, tweepy
- CrowdTangle - CrowdTangle et l'extension CrowdTangle sont actuellement les meilleurs outils pour analyser Facebook et Instagram.
- Pushshift.io - site qui archive les données des plateformes de réseaux sociaux. Reddit, Gab, Twitter et Telegram. Le fondateur et opérateur de Pushshift, Jason Baumgartner, obtient ses données via un accès API, ce qui rend l'utilisation des données sûre d'un point de vue éthique.
  - Outil de streaming Reddit : https://github.com/pushshift/reddit_sse_stream

7 Il est important d'être clair sur les implications de l'utilisation de ces outils en matière de respect de la vie privée. En particulier, vous et votre équipe voulez vous assurer que les discussions privées ne sont pas visibles ou enregistrées par un tiers lorsque vous rencontrez des outils d'analyse WhatsApp.

- MIT Media Cloud - MIT Media Cloud est un outil d'agrégation d'actualités qui peut être utile pour explorer la couverture de sujets d'intérêt dans divers médias. Dans sa description de l'outil, le site Web indique : « Nous agrégeons les données de plus de 50 000 sources d'information du monde entier et dans plus de 20 langues, dont l'espagnol, le français, l'hindi, le chinois et le japonais. Nos outils aident à analyser, fournir et visualiser des informations sur les conversations médiatiques à trois niveaux principaux : les pics d'attention et de couverture des problèmes, l'analyse des réseaux et l'utilisation de groupes sémantiques. »

Analyse de données et de réseaux

Construire des représentations visuelles des réseaux sociaux et analyser les relations en leur sein est un processus connu sous le nom d'analyse des réseaux sociaux (SNA pour ses initiales en anglais). Ce processus est également communément appelé création d'une carte de réseau social ou « cartographie » d'un réseau social. Bien qu'il ne soit pas toujours nécessaire d'utiliser l'analyse des réseaux sociaux pour comprendre la sphère des médias en ligne autour d'une élection, cela peut être un moyen utile de générer des aperçus instructifs sur l'influence au sein d'une partie donnée d'une communauté de réseaux sociaux, et cela peut être un moyen utile pour visualiser cette communauté.

Fondamentalement, la création d'une carte de réseau social est un processus qui se compose de 5 étapes :

1. Collecte de données
2. Décider de la relation avec la carte
3. Élagage des données
4. Génération de la carte
5. Analyse de la carte

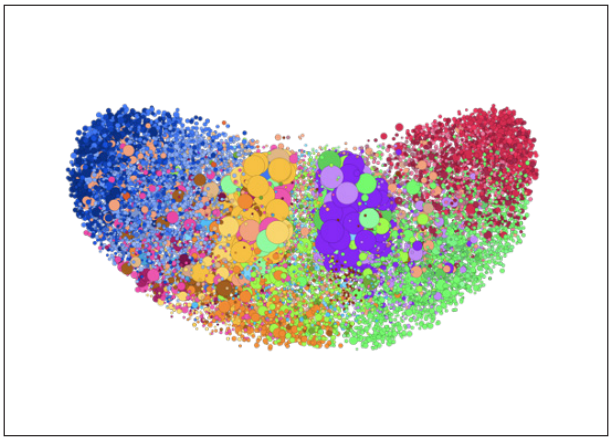
Nous aborderons chacune de ces étapes plus en détail ci-dessous.

Terminologie de base

Lors de la discussion et de l'analyse des cartes de réseaux sociaux, il y a quelques termes qu'il est utile de connaître pour aider à comprendre le vocabulaire et les pratiques typiques de la surveillance des réseaux sociaux. Les réseaux sont extrêmement utiles car ils peuvent représenter de nombreuses relations dans de nombreux contextes différents. Alors que les cartes des réseaux sociaux sont probablement les plus familières, les réseaux peuvent être utilisés pour modéliser la propagation de maladies et de virus, cartographier l'activité neuronale dans le cerveau ou représenter des itinéraires possibles pour se déplacer d'une ville à une autre. Bien que cette applicabilité des réseaux à tant de domaines différents soit utile, elle produit également le besoin d'un langage abstrait pour parler des réseaux quel que soit l'espace d'application. Cette section vous présente brièvement quelques termes de base utiles pour discuter des réseaux.

Les plus pertinents sont le graphe, le nœud et l'arête.

- Graphe - Graphe est la terminologie informatique pour un réseau composé de nœuds et d'arêtes. Il est important de connaître ce mot car vous pouvez le rencontrer dans les outils qui créent des cartes de réseau ou dans les discussions sur ces outils. Vous pouvez considérer que le mot graphe est à peu près synonyme de réseau.
- Nœud (ou sommet) - Les nœuds sont les éléments qui composent un réseau. Une chose essentielle à savoir sur les nœuds est que ce qu'un nœud représente est différent selon la carte/la visualisation que vous regardez. Un nœud dans la carte du réseau de sites Web pro-Kremlin de Lawrence Alexander représente un domaine, tandis que chaque nœud circulaire dans la carte du paysage politique américain sur Tweet 2018 de Graphika représente un compte Twitter.



Cette carte Graphika du spectre politique américain 2018 sur Twitter montre des nœuds. Chaque cercle de la carte est un nœud au-dessus représentant un compte Twitter individuel. Les nœuds et les arêtes sont les deux principaux éléments constitutifs des réseaux.

- Arête (ou arc) - Les arêtes sont les connexions entre les nœuds d'un réseau, le plus souvent représentées comme une simple ligne entre deux nœuds. Ces connexions peuvent représenter des choses différentes. Dans un modèle de contagion de maladie, les arêtes peuvent représenter la propagation d'un virus d'un hôte à un autre. Dans un graphe d'aéroports aux États-Unis, les arêtes entre deux nœuds (aéroports) peuvent représenter un vol direct disponible entre les deux aéroports. Les arêtes peuvent être dirigées ou non dirigées[8], et elles peuvent avoir une valeur numérique associée[9] (souvent appelé pondération).
    - Les arêtes des graphes de réseaux sociaux peuvent représenter l'une des choses suivantes : Relations de followers, retweets ou likes. La plupart des réseaux que vous voyez sur Twitter sont des réseaux d'abonnés (« followers »)[10] - dans lesquels l'arête signifie qu'un utilisateur en suit un autre - ou des réseaux de retweets.

Créer une carte de réseau social : Étapes impliquées

Dans cette section, nous allons définir et examiner les cinq étapes impliquées dans la création d'une carte de réseau social à partir de laquelle générer des informations.

1. Collecte de données - Examinée dans la section précédente, cette étape implique la collecte de données relatives à une élection locale via une API de réseaux sociaux ou un outil tiers. Une fois les données collectées, vous disposez d'une banque de données de base dont vous aurez besoin pour effectuer une visualisation du réseau social. Il est important de réaliser que seules des parties de ces données seront utilisées dans la génération de la carte - le même jeu de données de base peut être utilisé pour générer toutes sortes de cartes. Le choix du type de réseau/de relations qui vous intéresse le plus et des données les plus pertinentes est traité aux étapes 2 et 3.
2. Décider de la relation à cartographier - Dans une carte de réseau social, chaque « nœud » (c'est-à-dire un cercle sur la carte) représente probablement un compte Twitter ou une page Facebook. Cependant, ce sont les relations entre ces nœuds (relations de like, relations de suivi, retweets, etc.) qui donnent sa structure au réseau de nœuds pertinents. Vous pouvez considérer cette relation comme déterminant le squelette de la carte - à partir de ce squelette, nous extrayons les relations pertinentes concernant les nœuds de la carte. Dans la théorie des graphes - le domaine de l'informatique qui traite des réseaux et de la cartographie - ces connexions sont également appelées arêtes.

8 Par exemple, si le graphe modélisant la contagion virale discutée ci-dessus avait une direction associée à la propagation de l'infection, l'arête se déplacerait de l'infecteur à l'infecté dans ce cas et serait représentée avec une flèche.

9 Dans l'exemple de graphe d'aéroport ci-dessus, les valeurs des arêtes pourraient être la distance en milles entre les aéroports. Une autre valeur d'arête possible pour ce graphe serait le temps qu'il faut pour voler d'un aéroport à un autre.

10 Pour des exemples de réseaux de followers, consultez un réseau de comptes faisant la promotion de messages anti-vaccin dans Wired ici.

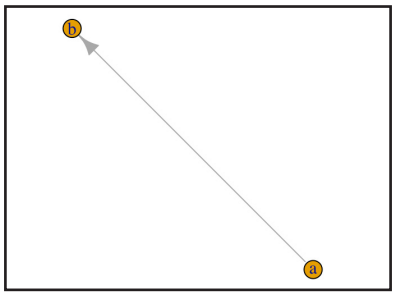
3. Élagage des données - Si vous collectez une grande quantité de données, vous vous retrouverez presque toujours avec plus de données que vous ne pouvez en cartographier. Affinio, NodeXL, Gephi et d'autres outils de génération de cartes ont tendance à mieux fonctionner dans une plage de quelques milliers de nœuds. La cartographie à plus grande échelle est généralement trop coûteuse en calculs pour être utile - pour cette raison, il est nécessaire de déterminer le réseau le plus pertinent pour vos questions. Le processus de suppression de données superflues est souvent appelé élagage en informatique - vous pouvez également entendre parler de réduction de réseau ou de réduction de dimensionnalité. Certains outils (par exemple Graphika) effectuent ce travail de réduction de réseau pour vous, mais vous pouvez également prendre vous-même des décisions de seuillage avec des outils tels que Gephi. Certains exemples incluent uniquement les nœuds de mappage qui ont utilisé un hashtag lié aux élections plus d'une fois, ou n'incluent que les nœuds qui ont une connexion à cinq nœuds ou plus dans le réseau.[11]
4. Création de la carte - Une fois que vous avez décidé de la relation à cartographier (arêtes) et des nœuds qui seront dans votre réseau grâce à la réduction du réseau, vous êtes prêt.e à créer votre carte. Généralement à ce stade, vous devrez mettre vos données pertinentes dans un format lisible pour l'outil que vous utilisez (par exemple un fichier CSV, un fichier .graphml ou un fichier .gexf pour Gephi), et les lire dans votre logiciel. Après cela, le gros du travail est fait pour vous. Il existe de nombreux tutoriels gratuits et utiles sur YouTube pour générer des cartes de réseau avec Gephi[12] et d'autres outils open source.
5. Analyse de la carte - Une fois votre carte générée, vous pouvez personnaliser les visuels et passer à l'analyse. Normalement, les mesures de la centralité sont essentielles pour comprendre l'influence dans un réseau. Analyser les réseaux sociaux (Borgatti, Everett et Johnson 2019) a un chapitre entier consacré aux différents types de centralité qui mérite d'être consulté.[13]

Exemples de types de réseaux

Comme mentionné ci-dessus, il existe plusieurs types de réseaux qui peuvent être générés à partir d'un ensemble de données donné. Le facteur déterminant pour le type de réseau que vous cartographiez réside dans la relation que vous avez choisi de cartographier. Sur Twitter, deux types courants de réseaux dirigés sont fréquemment utilisés pour analyser les données : les réseaux de suivi et les réseaux de retweets. Les deux types de réseaux peuvent être utiles pour analyser les données et l'influence.

Réseaux de suivi

Dans un « réseau de suivi », les nœuds sont des comptes Twitter et les connexions entre eux représentent les relations des followers entre eux. Normalement, ces connexions sont directionnelles (elles passent d'un nœud à un autre dans une direction donnée). Vous pouvez les considérer comme des lignes qui signifient « suit ». Par exemple, le graphe ci-dessous montre deux nœuds, A et B, et montre que « A suit B ».



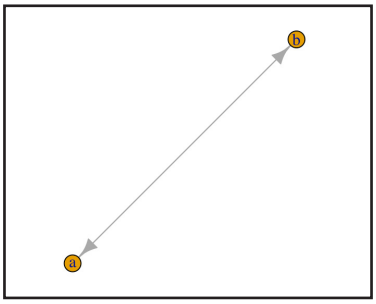
Représentation d'une relation de suivi à sens unique - dans ce diagramme, A suit B. Ceci est représenté par une seule ligne dirigée (ou flèche) de A à B.

11 Ce type d'élagage est connu sous le nom de réduction k-core.

12 Gephi propose également un tutoriel PDF gratuit et d'autres supports d'apprentissage sur son site officiel - gephi.org.

13 Chapitre 10, Centralité. Ce livre est un excellent texte pour apprendre les bases des réseaux et des méthodes d'analyse de réseau. Comme toujours, il existe également toutes sortes d'excellents tutoriels gratuits et open source en ligne.

Notez que dans ce graphe, il n'y a une relation de suivi que dans une seule direction - c'est-à-dire que A suit B, mais que B ne suit pas A. S'il y avait une relation de suivi mutuelle, dans laquelle A et B se suivaient tous les deux, une représentation du réseau montrerait des flèches dans les deux sens, comme dans la figure ci-dessous.



Représentation d'une relation de suivi mutuel : A suit B et B suit A.

Dans les réseaux générés sur les données Twitter autour d'une élection, le réseau pertinent contiendra généralement beaucoup plus de deux nœuds. Les exemples ci-dessus servent à donner une idée des éléments de base à partir desquels un réseau complexe se construit.

Avantages et inconvénients des réseaux de suivi

Les relations de suivi sur Twitter, en particulier, sont en quelque sorte des relations à long terme et permanentes - les utilisateurs ne se désabonnent pas souvent des autres utilisateurs. Pour cette raison, les réseaux de suivi représentent une vision à plus long terme que les réseaux de retweets en ce qui concerne la dynamique du réseau et les flux d'informations. D'un autre côté, les utilisateurs ont tendance à avoir plus d'un intérêt sur Twitter - cela ouvre la possibilité que certains des utilisateurs de votre réseau de followers ne soient pas aussi pertinents que vous le souhaiteriez pour le contenu qui vous intéresse. Les réseaux automatisés ont souvent des objectifs plus monothématiques, par exemple pour soutenir un parti, un candidat ou un problème particulier, mais leur contenu varie encore souvent. Il s'agit sans aucun doute d'un aspect à prendre en compte lors de l'analyse de l'automatisation de comptes ou d'autres formes d'activité coordonnée.

Il serait tout à fait possible, par exemple, dans une carte hypothétique du spectre politique américain, d'avoir une partie du réseau qui a tweeté principalement sur la culture pop et la musique, mais qui serait significativement connectée à un réseau de comptes principalement politiques. Ces avantages et inconvénients sont utiles à garder à l'esprit lorsque vous décidez si un réseau de followers est intéressant pour votre tâche.

Réseaux de retweets

Les réseaux de retweets sont un autre type de relation sur Twitter - les connexions entre les nœuds de ces réseaux représentent qui a retweeté qui dans les données collectées. À cet égard, ces cartes sont plus axées sur le contenu que les réseaux de followers. Une représentation des arêtes dans ce type de graphe est présentée ci-dessous.

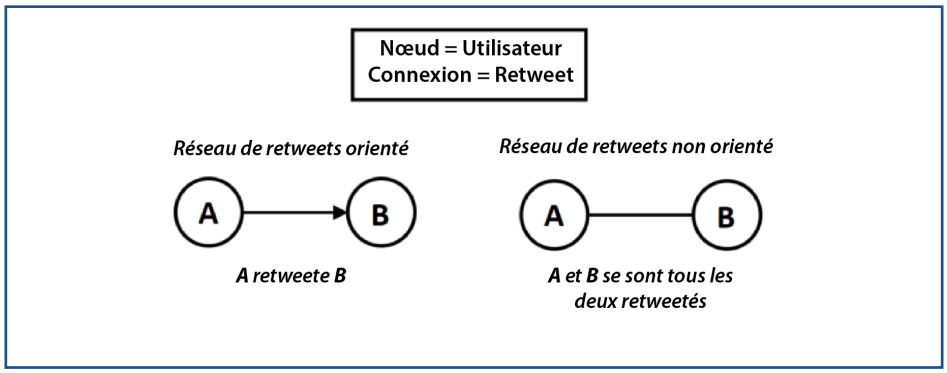
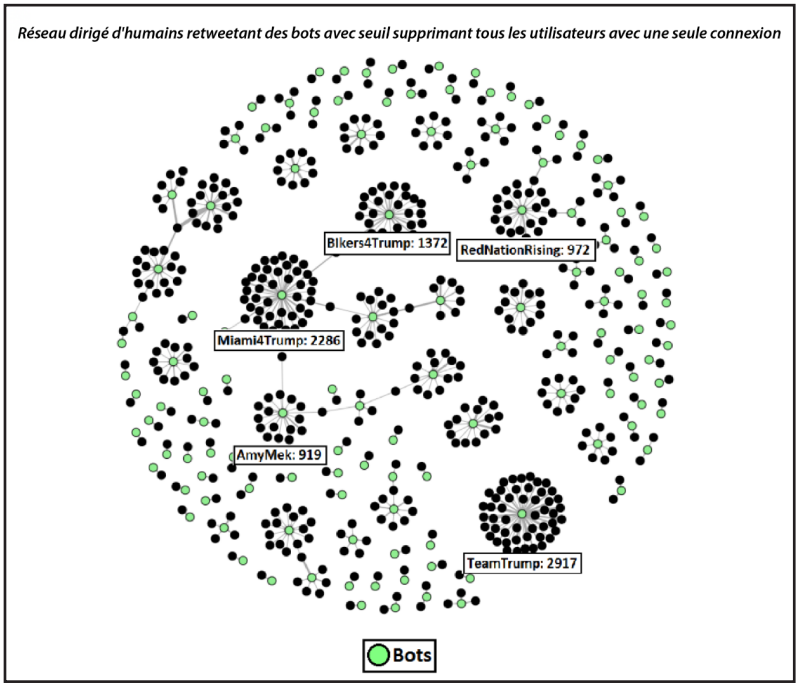


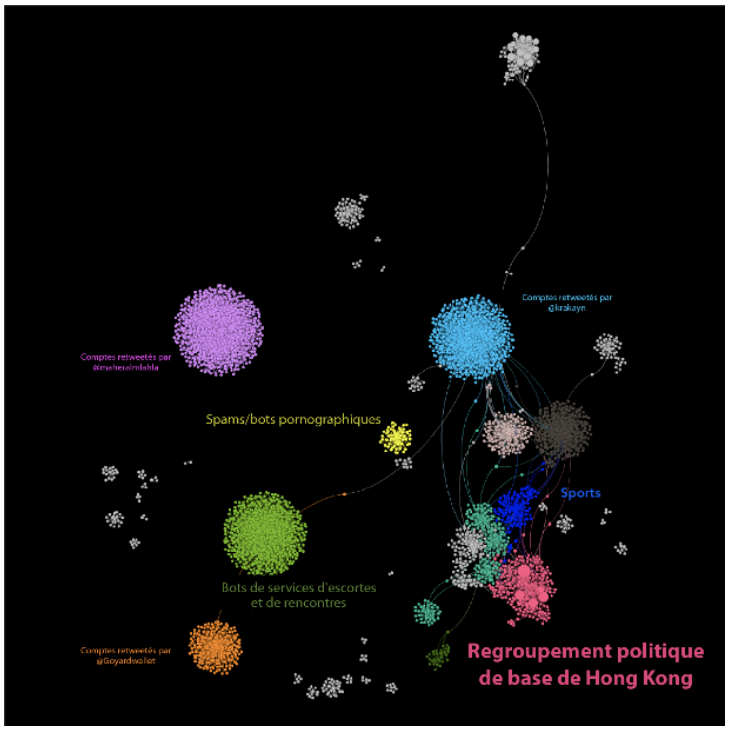
Diagramme illustrant les nœuds et les arêtes au sein d'un réseau de retweets. Source : Samuel Woolley et Douglas Guilbeault (2017). Computational Propaganda in the United States: Manufacturing Consensus Online. Disponible ici.)

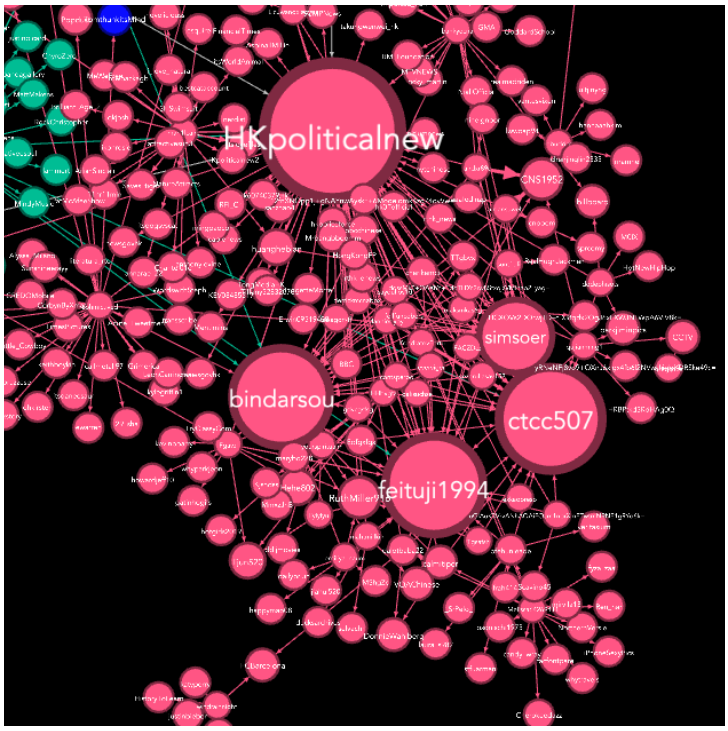
Avantages et inconvénients des réseaux de retweets

Ces réseaux sont un peu plus éphémères que les réseaux de followers, car les réseaux de retweets représentent un instantané dans le temps - qui retweetait qui dans un laps de temps donné. À cet égard, ils fournissent une description précise de la dynamique d'influence dans un court laps de temps - comme une campagne de hashtag dédiée ou les jours précédant une élection.



Un réseau de retweets généré à partir des données Twitter collectées autour de l'élection présidentielle américaine de 2016. Dans ce réseau, les nœuds représentant les comptes de bots sont verts et les comptes humains sont noirs. Ce réseau de retweets représente les retweets entre les utilisateurs et montre que les humains ont retweeté de manière significative le contenu des bots lors de l'élection présidentielle de 2016. (Source : Samuel Woolley et Douglas Guilbeault (2017). Computational Propaganda in the United States: Manufacturing Consensus Online. Disponible ici.)



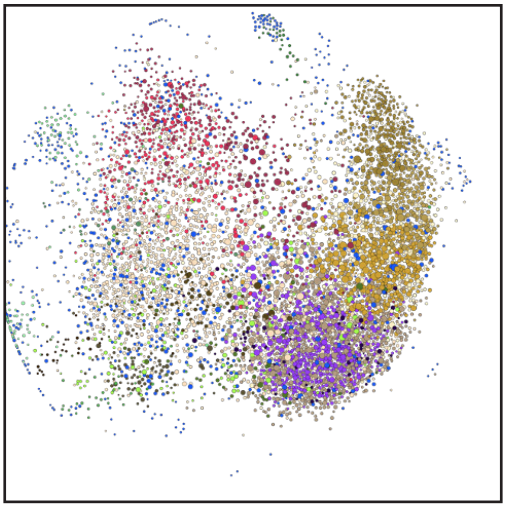


Images d'un réseau de retweets de comptes du gouvernement chinois diffusant de la désinformation pour discréditer les manifestations en faveur de la démocratie à Hong Kong de juin à août 2019. Ce réseau de retweets a été compilé à partir des archives des opérationsd'information de Twitter[14] . Ces images représentent l'ensemble du réseau de retweets (à gauche) et une vue agrandie du principal groupe politique ciblant Hong Kong (à droite). (Source : Rapport Digital Intelligence Lab de l'Institute for the Future, disponible ici.)

14 Plus précisément à partir de la divulgation d'août 2019 de tweets attribués au gouvernement chinois.

Réseaux de mentions

Les cartes de mentions sont un autre type de carte courant. L'idée est à peu près la même que les cartes de retweets - les arêtes représentent des « mentions » - les cas dans lesquels un utilisateur en mentionne un autre. À l'aide des données Gab assemblées par Pushshift.io, Graphika a généré une carte du réseau de mentions pour analyser les conversations et les communautés sur la plateforme sociale Gab - un réseau social très similaire à Twitter dans son utilisation et sa construction.



La carte de Graphika mentionne les utilisateurs de Gab. Cette carte a été compilée à partir de l'ensemble de données Gab accessibles au public de Pushshift.io, qui s'étend d'août 2016 à fin octobre 2018.

Suivre, retweeter et mentionner les réseaux ne sont que quelques-unes des options dont vous disposez pour générer des réseaux à partir des données Twitter - il existe certainement d'autres possibilités. Sur d'autres plateformes, comme Facebook ou Gab, d'autres relations peuvent être possibles. On peut imaginer une carte du réseau de pages publiques et des relations de like entre elles, par exemple.

Limites de la collecte de données et de la cartographie de réseaux

Comme pour tout outil, la cartographie réseau a certaines limites, tout comme le processus de collecte de données qui la précède. Les deux limites les plus importantes à garder à l'esprit sont l'échelle et le temps. Comme mentionné dans la section relative à la collecte de données, tous les outils disponibles pour générer des cartes de réseau ont actuellement des limites d'échelle, qu'elles soient gratuites ou payantes. En effet, à mesure que le nombre de nœuds dans un réseau augmente, le nombre de connexions possibles augmente de façon exponentielle[15]. Cette quantité de données est éprouvante, même pour les ordinateurs, et il est donc peu probable que vous trouviez un outil capable de générer une carte de plus de 15 000 nœuds environ[16]. Cependant, il serait erroné de considérer cette limitation exclusivement comme un problème : une partie importante de la science des données et de l'analyse des données consiste à choisir les parties d'un grand ensemble de données qui sont les plus susceptibles de fournir des informations utiles. En d'autres termes, analyser correctement les données implique de choisir les parties des données qui méritent le plus d'être étudiées - sous cet aspect, la science des données est autant un art qu'une science. Cette tâche, qui est un élément central de l'enquête, de l'analyse et de la compréhension des données, n'est pas toujours facilitée par le fait d'avoir plus de données.

La deuxième limitation que vous êtes susceptible de rencontrer avec la collecte de données et la cartographie

15 Pour ceux qui s'intéressent aux mathématiques, un réseau de n nœuds a n*(n-1)/2 connexions possibles.

16 Le logiciel de cartographie de Graphika est actuellement capable de cartographier à la plus grande échelle, permettant de visualiser jusqu'à 5,5 millions de nœuds

du réseau est une limitation de temps, en particulier lors de la collecte de données historiques. L'API de recherche Twitter standard permet uniquement la collecte de données historiques remontant à une semaine à partir du moment de la requête. Les données datant de plus de 7 à 9 jours doivent provenir d'une autre source, telle qu'un outil (payant) permettant une collecte historique plus étendue, ou l'achat de données auprès d'un fournisseur de données, tel que GNIP. Bien que l'analyse des données historiques puisse être utile dans certains cas, l'achat de données historiques peut rapidement devenir coûteux. Pour cette raison, la meilleure solution est toujours de collecter toutes les données pertinentes en temps réel. Si vous ou un membre de votre équipe souhaitez diffuser ou collecter des données autour d'une élection, il est toujours préférable de commencer à collecter dès que vous avez une idée claire de ce qui vous intéressera.

Réseaux fermés et cryptés

Un défi auquel sont confrontés les efforts de surveillance des médias ces dernières années est la popularité et l'adoption généralisée d'applications de messagerie cryptées telles que WhatsApp, Telegram et Signal. De plus en plus, les médias sur lesquels les citoyens comptent pour se tenir au courant des développements politiques pendant les élections et d'autres événements politiques majeurs se trouvent sur ces plateformes. Des pays comme le Brésil, l'Inde et le Mexique ont connu une nette augmentation des messages politiques sur WhatsApp, par exemple. Si l'utilisation de réseaux de messagerie cryptés est sans aucun doute avantageuse pour le respect de la vie privée, la sécurité et les droits numériques des citoyens, elle pose de nouveaux défis pour comprendre la diffusion d'informations politiques en ligne.

À l'heure actuelle, les meilleures méthodes pour les efforts de surveillance des médias des réseaux fermés reposent sur la vérification manuelle des faits - des équipes d'experts qui surveillent les canaux WhatsApp pertinents individuellement ou en groupe, vérifient les faits et diffusent les résultats de ces efforts publiquement. De tels efforts ont été couronnés de succès dans plusieurs cas tels que La Silla Vacía[17], détecteur de WhatsApp en Colombie, ainsi que le travail de Verificado[18] au Mexique et le Centre pour la démocratie et le développement au Nigéria.[19]

Un autre effort notable est celui du bot Cofacts à Taïwan. Johnson Liang et une équipe de développeurs travaillant avec le mouvement g0v ont conçu un moyen de combiner la vérification manuelle des faits et la distribution automatisée pour aider les citoyens taïwanais à vérifier si une information est fausse ou non. Les utilisateurs peuvent ajouter le bot Cofacts[20] sur LINE, une application de messagerie cryptée populaire utilisée à Taïwan et au Japon. Si un utilisateur voit une information suspecte, il peut coller le lien dans un chat vers le bot Cofacts. Si l'information n'a jamais été vue auparavant, une équipe de personnes vérifie l'information et télécharge un message en résumant la véracité dans une base de données centrale. Le bot transmet ensuite ce message à l'utilisateur d'origine et à tout autre utilisateur curieux de savoir si l'information est vraie ou non. Cofacts permet l'accès public aux données anonymisées qu'elle a recueillies sur Github, et permet également aux utilisateurs d'effectuer des recherches dans cette base de données via un site Web public. Meedan, une entreprise qui prend en charge la vérification des faits et d'autres recherches en ligne, a également publié un ensemble d'outils similaire sur Check qui aide divers utilisateurs à automatiser et à gérer collectivement les flux de travail pour le processus de vérification des faits sur WhatsApp et d'autres plateformes.

Certains outils existent, tels que Backup WhatsApp Chats, qui permettent aux utilisateurs d'exporter les conversations WhatsApp vers des fichiers CSV, mais un utilisateur doit toujours appartenir à un canal pour exporter les chats à partir de celui-ci. Telegram, une application de messagerie cryptée, possède une API qui permet aux utilisateurs d'accéder aux chaînes publiques par programmation. Cela permet une certaine surveillance des médias publics, bien que les données ne soient pas aussi riches que sur Twitter.

17 https://www.niemanlab.org/2017/03/to-slow-the-spread-of-false-stories-on-whatsapp-this-colombian-news-site-is-enlisting-its-own-readers/
18 https://www.niemanlab.org/2018/06/whatsapp-is-a-black-box-for-fake-news-verificado-2018-is-making-real-progress-fixing-that/
19 https://www.cddwestafrica.org/whatsapp-nigeria-2019-press-release/
20 Le nom chinois de ce bot est 真的假的, « vrai ou faux ».

Outils et modules utiles de visualisation de réseaux

Le tableau ci-dessous répertorie plusieurs outils et modules de code courants utilisés pour la visualisation et l'analyse des réseaux.

Outils de visualisation et d'analyse de réseaux			
Outils open source/ gratuits	Outils payants	Modules couramment utilisés (Python)	Modules couramment utilisés (R)
Gephi NodeXL	Graphika Affinio	networkx matplotlib igraph	igraph plotrix

Identifier les influenceurs, les groupes et les comptes

L'objectif de la collecte et de l'analyse des données est de comprendre comment l'information est distribuée au sein d'un réseau. Quelles sont les histoires qui suscitent le plus d'intérêt ? Quels sont les utilisateurs les plus influents ? Quels domaines d'actualité sont les plus fréquemment cités dans la conversation ? Avec un ensemble de données solide et les outils appropriés, vous pouvez commencer à répondre à ces questions avec spécificité et comprendre la dynamique de diffusion de l'information dans l'espace en ligne que vous observez.

Il existe deux manières de conceptualiser l'influence sur les réseaux sociaux : nous pouvons désigner ces méthodes comme étant basées sur le contenu ou comme étant basées sur les acteurs afin d'identifier l'influence. Nous allons les explorer toutes les deux ci-dessous.

Méthodes basées sur le contenu pour identifier l'influence

Les méthodes basées sur le contenu se concentrent sur les Tweets, les hashtags, les mots-clés ou les sites Web sous la forme d'URL ou de domaines[21]. Dans certains scénarios, des acteurs à surveiller sont connus - des médias ou des hommes politiques qui diffusent fréquemment de la désinformation, par exemple. D'un autre côté, il est assez courant de ne pas connaître les sources de désinformation, de discours de haine ou d'autres formes de contenu que vous recherchez.

Souvent, cela peut être très utile lors de l'analyse des conversations en ligne autour des élections ou d'autres discours politiques pour comprendre quel contenu suscite le plus d'intérêt - surtout lorsqu'il n'y a pas un acteur particulier que vous cherchez à analyser et que vous avez identifié dans les données. Regarder quelles URL ou quels Tweets suscitent le plus d'intérêt dans les données peut être un bon point de départ dans ce scénario.

Tweets

Lors de l'examen d'un Tweet, l'influence peut être estimée en examinant le nombre de retweets et/ou de likes qu'il a recueillis[22]. Que vous utilisiez un outil tiers ou que vous récupériez des données de l'API Twitter, vous devriez avoir un accès facile à ces données à tout moment. Après avoir déterminé quels tweets sont les plus influents, vous pouvez utiliser ces Tweets comme tremplin pour une enquête plus approfondie. Certaines questions méritent d'être étudiées :

21 Les URL font référence à un lien complet qui vous dirige vers une histoire ou un site particulier, tandis que les domaines font référence au site qui héberge ce contenu - cela correspond au texte précédant le domaine de premier niveau (abrégé en TLD - tel que .org, . com, .gov, etc.) et le TLD lui-même. Par exemple, les trois URL example-news-site.com/story1, example-news-site.com/story2 et example-news-site.com/story3 sont toutes hébergées sur le même domaine - example-news-site.com.

22 Une note intéressante à garder à l'esprit pour les Tweets est qu'ils ont tendance à avoir plus de likes que les retweets. Ceci est similaire au fait que la plupart des publications Facebook ont plus de likes que de partages. Si ces ratios sont anormaux, cela peut être un indicateur d'activité non authentique, mais ce n'est en aucun cas une garantie.

- Qui a produit le Tweet à l'origine ? Qui a retweeté la publication ? Comment se présentent les suivis de ces utilisateurs ? S'ils sont volumineux, il peut être intéressant de les analyser en réseau.
- Quels hashtags sont utilisés dans le Tweet ? Si l'un d'entre eux est distinctif ou n'est poussé que par un petit groupe d'utilisateurs, ces utilisateurs ont-ils quelque chose en commun ?
- Quelles URL sont présentes dans la publication ? Si l'URL est suspecte (récemment créée ou promouvant la désinformation), une enquête supplémentaire peut être utile- en vérifiant les registres d'enregistrement de domaines via une recherche Whois[23], ou en cherchant sur Twitter d'autres mentions intéressantes de l'URL, par exemple. Vous pouvez également utiliser l'extension de navigateur CrowdTangle pour voir si cette URL gagne en popularité sur Facebook, Instagram, Reddit ou ailleurs sur Twitter.

Ces mêmes stratégies et principes s'appliquent également à Facebook, Twitter, Gab et d'autres plateformes de réseaux sociaux - l'analyse des interactions avec une publication est un moyen fiable de mesurer l'influence d'un message dans une communauté donnée.

Hashtags/Mots clés

Les hashtags sont naturellement l'une des principales entités d'intérêt lors de l'examen de Twitter et d'autres données sociales. La convention d'utiliser un hashtag pour mettre en évidence le sujet de ce qui est tweeté est une bénédiction pour les chercheurs - cela nous permet de collecter des données de conversation pertinentes pour un sujet d'intérêt avec une grande facilité. Une fois qu'un hashtag ou un ensemble de hashtags d'intérêt a été identifié, plusieurs possibilités d'enquête plus approfondie existent : analyser les hashtags co-occurrents les plus courants, diviser les citations de hashtag par heure ou analyser les URL co-occurrentes ne sont que quelques-unes des nombreuses options.

URL/domaines

Les URL et les domaines apparaissant dans les Tweets, le plus souvent liés à des sources d'actualités, sont une source extrêmement utile pour déterminer quels contenus, quelles publications et quels récits gagnent le plus en popularité au sein d'une communauté en ligne donnée.

Une première étape courante de l'analyse URL/domaine consiste à extraire des URL uniques d'un ensemble de données et à compter le nombre de fois où elles sont citées dans l'ensemble. Cette technique est simple et puissante, mais il est également facile de se tromper de plusieurs manières subtiles mais conséquentes. Pour vous assurer que votre analyse donne les informations les plus précises et les plus utiles, il convient de prêter attention à quelques-uns des problèmes ci-dessous.

- Résolution des URL raccourcies - Les URL dans les Tweets sont souvent raccourcies pour tenir dans les limites de caractères - l'utilisation d'un outil de résolution d'URL vous donnera l'URL complète vers laquelle pointe l'URL raccourcie. Certains outils, tels que URLex.org, offrent une API pour la résolution en masse et automatisée des URL. C'est le meilleur moyen de vous assurer qu'il ne vous manque aucune donnée dans les URL raccourcies.
- Normalisation des URL - Le même domaine peut être cité en utilisant différentes chaînes de texte[24]. Par exemple, des liens vers le New York Times peuvent apparaître dans le texte comme www.nytimes.com, nytimes.com, NYTimes.com, http://nytimes.com, nyti.ms, https://nytimes.com, http://www.nytimes.com, https://www.nytimes.com ou m.nytimes.com - ce ne sont là que quelques-unes des possibilités. Afin de vous assurer de ne pas sous-estimer l'influence d'une URL ou d'un domaine en comptant différentes chaînes, vous devez vous assurer de normaliser le format avant de compter. Il ne sera pas toujours possible de contrôler toutes les

23 Il existe de nombreuses bases de données Whois pour vérifier les détails de l'enregistrement ; l'une d'entre elles, fiable, est gérée par l'Internet Corporation for Assigned Names and Numbers (ICANN) https://lookup.icann.org/

24 Dans un contexte informatique, les données textuelles sont souvent appelées « chaînes ». C'est l'abréviation de chaînes de caractères, et c'est ainsi que les informaticiens se réfèrent souvent au texte du Tweet/post dans les ensembles de données des réseaux sociaux.

manières dont les URL renvoyant au même contenu peuvent se produire, mais un peu de réflexion en amont sur la normalisation peut grandement contribuer à améliorer vos analyses. Les éléments à prendre en compte lors de la normalisation sont (1) la sensibilité aux lettres majuscules et minuscules[25], (2) les préfixes (http://, https://, www., ww2., etc.) et (3) les sous-domaines, entre autres facteurs.

Après avoir analysé les URL les plus populaires d'un ensemble, vous disposez de plusieurs options pour une analyse plus approfondie. Si le domaine est une source d'actualités relativement nouvelle, vous pouvez utiliser des techniques de renseignement open source (OSINT), telles que la vérification des informations d'enregistrement du domaine pour commencer à avoir un aperçu des connexions du domaine. OSINT est un moyen utile de recueillir plus d'informations sur les comptes ou les sites Web d'intérêt. C'est un domaine en constante évolution, dans lequel des outils et des techniques apparaissent et disparaissent chaque jour. L'une des meilleures sources d'apprentissage de nouvelles compétences est Intel Techniques, mais il en existe de nombreuses autres. Bellingcat, un collectif de chercheurs qui documente souvent les opérations d'influence russes, a fait de l'utilisation d'OSINT une forme d'art pour livrer des articles journalistiques révolutionnaires, et gère un document Google public[26] avec un inventaire complet des outils d'enquête et de ceux d'OSINT. Lawrence Alexander, expert en mégadonnées basé au Royaume-Uni, a astucieusement utilisé les codes de Google Analytics pour cartographier un réseau de sites Web pro-Kremlin en 2015.

Après avoir repéré les sites web et les articles les plus influents d'un ensemble de données particulier, vous pouvez également procéder à une analyse de contenu afin d'analyser les récits qui apparaissent dans ces articles. Les techniques de traitement automatique du langage naturel (TALN) telles que l'utilisation de fréquences n-grammes pour analyser les phrases les plus courantes dans ces articles, tf-idf pour comparer les thèmes relatifs de différents articles, ou l'utilisation de méthodes d'analyse de contenu qualitative peuvent toutes conduire à des résultats instructifs une fois la recherche réduite à quelques articles et URL influents.

Basée sur le réseau

Une autre façon d'analyser l'influence dans un ensemble de données consiste à adopter une approche basée sur le réseau. Dans cette approche, vous utilisez vos données pour créer un réseau pertinent, comme indiqué dans la section Analyse des données et des réseaux ci-dessus - il peut s'agir de mentions, de followers, de retweets ou d'autres mesures basées sur le réseau.

Regroupement/Groupes

La détection des communautés est la base du travail le plus important qui compare les communautés (souvent appelées clusters dans le SNA). Bien qu'il existe des théories méthodologiques complexes qui sous-tendent différents algorithmes de regroupement de communautés, la plupart du travail est effectué pour vous par n'importe quel logiciel que vous utiliseriez. Gephi propose plusieurs options de mise en page et de regroupement de communautés (vous pouvez trouver plus de détails dans le tutoriel ici). Graphika fait également ce travail pour vous automatiquement. Bien entendu, une fois que votre logiciel de visualisation et d'analyse de réseaux détermine le nombre de communautés différentes de votre réseau, vous pouvez passer à une analyse qualitative pour déterminer ce que les membres d'un regroupement (cluster) donné ont en commun.

Souvent, les communautés et les groupes de comptes auront des caractéristiques communes perceptibles, telles que la promotion de sources d'information similaires ou l'appartenance à un parti similaire. C'est là que la compréhension du contexte politique global devient importante. Combiner l'analyse quantitative des données de la communauté avec l'analyse qualitative du contenu est le meilleur moyen de déterminer ce que les membres d'une communauté donnée ont en commun.

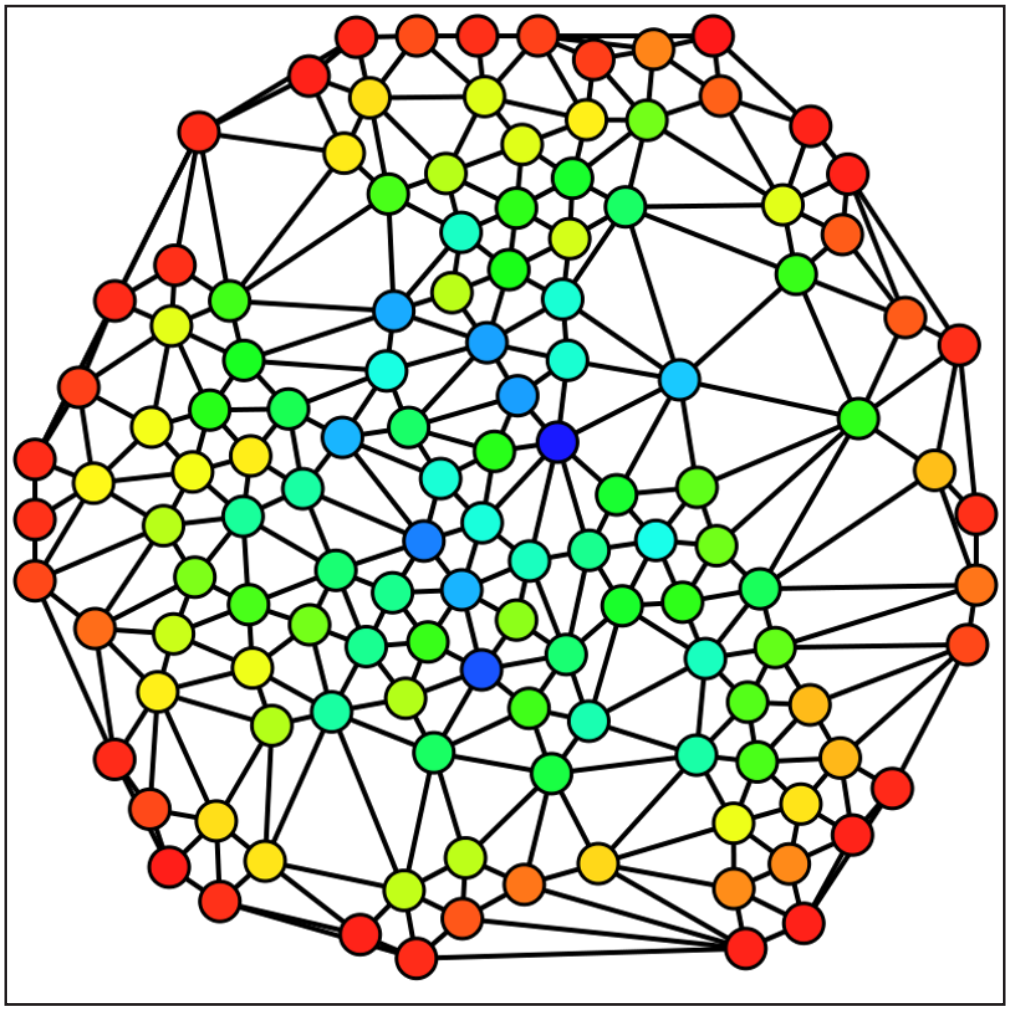
25 La sensibilité aux lettres majuscules et minuscules indique si les données de texte sont en minuscules ou en majuscules. Ces problèmes sont plus facilement traités en mettant en minuscules toutes les URL de l'ensemble avant de compter le nombre d'occurrences de chaque URL. Il est toutefois important de noter que si la plupart des URL complètes ne sont pas sensibles aux lettres majuscules et minuscules, de nombreuses URL raccourcies le sont. Il est donc recommandé que les chercheurs s'occupent d'abord de la résolution des URL avant de passer à des URL en minuscules/normalisantes.

26 https://docs.google.com/document/u/1/d/1BfLPJpRtyq4RFtHJoNpvWQjmGnyVkfE2HYoICKOGguA/edit

Centralité

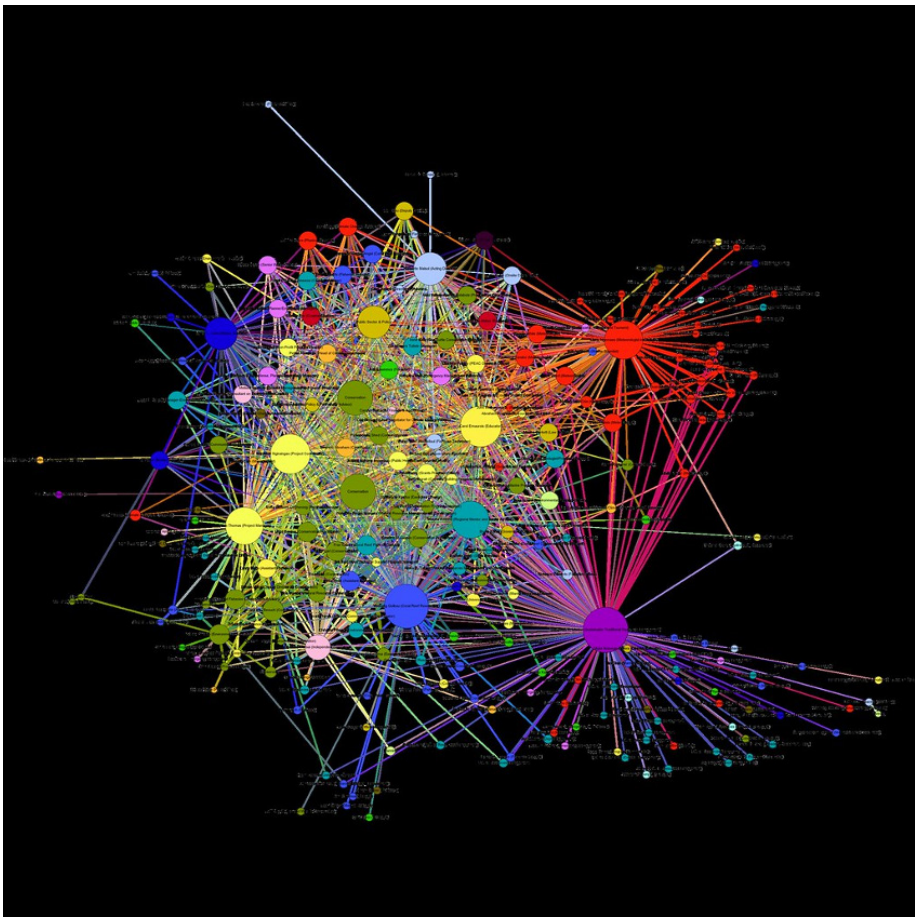
Dans la théorie des réseaux, en particulier telle qu'elle s'applique au SNA, la principale métrique pour analyser l'influence d'un même acteur ou d'une même groupe d'acteurs est la centralité. La centralité est une mesure du degré de connexion ou d'influence d'un nœud donné au sein d'un réseau. Il existe plusieurs méthodes différentes pour mesurer la centralité.

- Degré de centralité - Dans un graphe non orienté, le degré de centralité d'un nœud donné est le nombre de connexions directes qu'il a avec les autres nœuds qui l'entourent. Dans un graphe orienté, où la direction de l'arête est prise en compte, il existe deux types de centralité de degré - la centralité entrante ou le nombre de connexions entrantes d'un nœud - et la centralité sortante, ou le nombre de connexions sortantes d'un nœud. La centralité entrante dans ces scénarios est sans doute la plus pertinente - un nœud avec un nombre élevé de connexions entrantes a un potentiel d'influence élevé. Il s'agit d'un concept relativement intuitif - un compte avec 15 000 followers (un degré d'entrée de 15 000) a une plus grande capacité d'influencer directement qu'un compte avec 100 followers. L'importance du degré d'entrée/de sortie varie en fonction du réseau que l'on examine.
- Centralité d'intermédiarité - La centralité d'intermédiarité peut être considérée comme la mesure de la capacité d'un nœud à diffuser rapidement un message à d'autres nœuds. En d'autres termes, la centralité d'intermédiarité affecte des scores plus élevés aux nœuds susceptibles d'être des vecteurs de viralité ou de diffusion rapide de messages. Cette valeur est mesurée en comptant le nombre de fois qu'un nœud donné se trouve dans la partie du chemin le plus court d'un nœud à un autre.



La teinte (de rouge = 0 à bleu = max) indique la centralité d'intermédiarité de chaque nœud. (Claudio Rocchini) Creative Commons BY 2.5 https://en.wikipedia.org/wiki/Social_network_analysis#/media/File:Graph_betweenness.svg

- Centralité de vecteur propre - La centralité de vecteur propre est très utile pour comprendre l'influence au sein d'un réseau ou d'un sous-réseau donné. La question sous-jacente à la centralité de vecteur propre est essentiellement « Qui influence les influenceurs ? ». Avec la centralité de vecteur propre, l'influence dépend moins du nombre de liens qu'un nœud cible a avec d'autres nœuds que du nombre de connexions que le nœud cible a avec d'autres nœuds bien connectés. Vous pouvez considérer cette version de la centralité comme un type d'« influence indirecte » : quelle est la probabilité que les informations provenant d'un nœud source se propagent à d'autres nœuds influents ? Le moteur de recherche de Google utilise une forme de ce concept car il classe les pages (ou nœuds) en fonction du nombre d'autres pages qui les référencent. Les nœuds à forte centralité de vecteur propre peuvent jouer un rôle particulièrement puissant dans la diffusion de biens (tels que l'argent, l'information, les microbes, etc.) à travers un réseau.



Pacific RISA Social Network Analysis Project, carte centrée sur Palau ; tous les membres sont originaires de Palau ou y sont liés. Les nœuds sont dimensionnés selon la centralité de vecteur propre et colorés selon la profession. Creative Commons BY 2.5 https://www.flickr.com/photos/pacificrisa/11344578486

- Ponts - Bien qu'il ne s'agisse pas d'une forme de centralité, les ponts sont un phénomène de structure de réseau utile à garder à l'esprit. Un nœud ou un petit ensemble de nœuds agissent comme un pont lorsqu'ils connectent un regroupement dense à un autre regroupement dense. En l'absence de ponts, les regroupements d'intérêts divers (tels que les membres de deux partis politiques distincts ou les amateurs de politique de deux pays différents) n'auraient aucun moyen de se transmettre des informations ou de communiquer entre eux. À cet égard, les ponts sont des éléments importants du contenu viral sur les réseaux en ligne.
- Feuilles - Les feuilles sont des nœuds d'un graphe connectés à un réseau par une seule arête. Ces comptes méritent souvent d'être supprimés d'un réseau, en particulier des grands réseaux, dans lesquels il est très peu probable qu'ils aient une quelconque influence.

On trouvera en ligne plusieurs tutoriels et ressources sur différents types de centralité - une ressource utile est le livre Analyzing Social Networks (Borgatti et al., 2013), dont le chapitre 10 présente différents types de centralité et les façons dont ils peuvent être mesurés.

Des outils open source comme Gephi peuvent faire le calcul difficile des différentes formes de centralité pour vous en quelques secondes[27]. Le choix de la forme de centralité à utiliser est finalement un choix qui vous appartient, à vous et à votre équipe. Bien que consulter un expert en mégadonnées ou un théoricien des réseaux soit idéal pour faire ce choix, vous n'avez pas besoin de trop insister sur les détails méthodologiques ici - presque toutes les formes de centralité supposent que les nœuds les plus influents sont, dans un certain sens, les nœuds avec le plus grand nombre de liens dans un réseau ou un sous-réseau.

Une fois que vous avez déterminé la forme de centralité que vous souhaitez utiliser, vous pouvez même personnaliser les visuels de votre réseau pour refléter votre choix. Gephi vous permet d'utiliser la taille des nœuds ou des nuances de couleur pour mettre en évidence la centralité, par exemple.

Identifier les sources d'information

Pour les élections et l'analyse de l'information en particulier, la question du nombre de sources d'information citées dans les données est d'une importance capitale. Il s'agit d'un domaine dans lequel il est particulièrement utile de travailler avec des personnes qui ont une compréhension sociale et politique approfondie du pays ou de la région en question. Les sources d'information les plus fréquemment citées sont souvent bien connues de ces personnes. Il est important d'avoir une solide compréhension de l'environnement médiatique sous-jacent dans le pays, tel qu'exposé ci-dessus, y compris les parts de marché et l'influence respectives de la télévision, de la radio, des journaux et des médias en ligne.

Il existe plusieurs méthodes utiles pour identifier de nouvelles sources d'information dans une région ou un ensemble de données. Une option consiste à effectuer des recherches manuelles fréquentes sur Google sur les actualités et la politique de la région, en particulier dans plusieurs langues si la région est multilingue. Les sources d'actualités qui se disputent l'influence sont susceptibles de bénéficier d'une forte optimisation des moteurs de recherche (SEO pour ses initiales en anglais) pour apparaître dans les premiers résultats retenus par Google, et peuvent souvent faire apparaître des sources d'actualités récemment créées et inconnues du citoyen moyen. Une autre option, fortement recommandée, consiste à effectuer une analyse de citation de domaine et d'URL sur un ensemble de données. Cela peut prendre plusieurs formes - surveiller CrowdTangle pour trouver les principales URL dans un ensemble donné de pages politiques, ou extraire des URL/domaines uniques à partir des données de publication[28] dans un ensemble de données collectées avec un expert local et rechercher des domaines que vous n'avez jamais vus auparavant. Crowdtangle dispose d'un plugin gratuit pour Chrome et d'autres navigateurs qui vous permet de voir les partages d'un article donné, mais pour accéder à son tableau de bord plus avancé et à d'autres outils, les utilisateurs doivent obtenir une licence institutionnelle, souvent via Facebook, qui en est désormais propriétaire.

L'idée directrice ici est que la désinformation émane souvent de domaines nouvellement créés qui apparaissent rapidement autour des questions politiques et des élections, et disparaissent tout aussi rapidement. Ce fut le cas de streetnews[.]one, domaine qui a été créé et promu sur Gab avant les élections de 2018 aux États-Unis. Le domaine a été utilisé pour diffuser des contenus islamophobes et sensationnalistes avant de disparaître rapidement. Ce fut également le cas dans Endless Mayfly, une analyse de la désinformation iranienne menée par le Citizen Lab de l'Université de Toronto en collaboration avec des experts du secteur. Dans ce cas, l'équipe de Citizen Lab a inventé l'expression « désinformation éphémère », pour décrire les domaines de désinformation qui apparaissent et disparaissent rapidement et qui ciblent des problèmes, des élections et des campagnes clés.

27 Ou de minutes, selon la taille du réseau en question.

28 De préférence avec le nombre de citations comme approximation de la popularité de ces sources. D'autres méthodes, telles que l'extraction d'URL à partir des publications les plus retweetées ou les plus appréciées d'un ensemble de données, permettent également d'estimer l'influence d'un domaine particulier au sein d'un ensemble de données.

L'utilisation de listes assemblées de domaines de désinformation ou de sites Web de fausses informations à ce stade peut être extrêmement utile. Les chercheurs de Stanford ont élaboré une liste de plus de 600 domaines connus pour produire du faux contenu à la fin de 2018, ciblant principalement les États-Unis (Allcott et al, 2018)[29] . Vous devez garder à l'esprit deux mises en garde importantes si vous décidez d'utiliser de telles listes : (1) la désinformation et la sphère en ligne évoluent rapidement - les nouveaux domaines de désinformation qui sont apparus depuis le moment de la compilation de la liste seront absents de l'analyse ; et (2) toute liste utilisée doit être correctement vérifiée au regard de la rigueur méthodologique - l'utilisation de listes assemblées au hasard trouvées en ligne pourrait nuire à l'intégrité de la recherche.

Analyse des comptes et du contenu

Une fois que vous avez collecté des données et/ou construit un réseau de comptes pertinents qui vous intéressent, vous et votre équipe êtes prêts à commencer à analyser les comptes et le contenu de l'ensemble de données. Il s'agit probablement la partie du travail sur laquelle vous passerez le plus de temps, surtout si vous ne savez pas au départ ce que vous recherchez.

L'analyse du contenu est mieux réalisée en oscillant entre les méthodes qualitatives et quantitatives, en itérant souvent le processus plusieurs fois pour trouver des comptes ou des contenus d'intérêt spécifiques. Dans cette section, nous vous fournirons quelques conseils et outils pour vous aider à orienter ce travail.

Types de contenu

Il est utile, avant de se plonger dans les données, d'avoir une idée de certains des types de contenu que vous pouvez rechercher dans l'ensemble de données.

- Désinformation/Més-information - Le contenu politique faux et trompeur est l'une des formes de contenu les plus pernicieuses que vous devrez surveiller dans votre ensemble de données. Bien que le contenu erroné soit désigné par plusieurs noms, y compris ceux de propagande informatique et de fausses nouvelles, ce type de contenu est le plus souvent appelé désinformation ou més-information. La différence technique entre ces deux termes réside dans l'intention sous-jacente au faux contenu[30]. Les définitions que nous utiliserons pour différencier ces deux mots sont tirées du rapport Lexicon of Lies de Data and Society (Jack 2017), qui explore un ensemble de mots utiles à comprendre et à utiliser pour discuter des faux contenus en ligne.
    - Désinformation - Faux contenu diffusé avec l'intention délibérée de tromper. Les acteurs des États-nations mus par des motivations politiques ou financières sont les plus susceptibles de diffuser sciemment de la désinformation, telle que définie de cette manière.
    - Més-information - Fausse information qui est propagée et distribuée sans intention de nuire. Si un compte fait la promotion d'une histoire sans avoir l'intention de tromper les utilisateurs, cela est considéré comme de la més-information.
    - Mal-information - Certains auteurs ont également mis en évidence le phénomène de « mal-information » - diffusion d'informations vraies ou pour la plupart véridiques avec l'intention de nuire. Claire Wardle et Hossein Derakhshan définissent la mal-information comme « une information basée sur la réalité, utilisée pour nuire à une personne, une organisation ou un pays » (Wardle et Derakhshan 2017).
- Discours de haine - le discours de haine est un discours qui diabolise les personnes d'une race, d'une ethnie, d'un sexe, d'une orientation sexuelle ou d'une religion cible. Souvent, le discours de haine incite les autres à harceler, dénigrer ou même se livrer à des actes de violence contre le groupe social ciblé.

29 Vous pouvez trouver ce document ici.

30 Bien que ce soit le cas, il est souvent impossible de connaître l'intention derrière le faux contenu qui se propage en ligne. Pour cette raison, vous entendrez probablement ces mots utilisés de manière interchangeable à certains moments.

Analyse linguistique

Une façon d'analyser le contenu de manière rigoureuse est de procéder à une analyse linguistique. L'analyse linguistique peut vous aider à comprendre les principaux langages de messagerie, les principaux thèmes de contenu, les récits poussés et développés, les nouveaux mots-clés et le jargon, et si des discours de haine ou d'autres contenus dangereux se produisent dans la conversation.

Mots-clés et lexiques

Tous les messages ou tweets d'un ensemble de données particulier ne sont pas pertinents pour les questions spécifiques qui vous intéressent. Pour cette raison, il peut être utile de compiler une liste de mots-clés pertinents susceptibles de contenir des informations pertinentes pour votre demande. Travailler avec un expert en la matière - quelqu'un qui a une compréhension approfondie de la langue et de la politique de la région d'intérêt - est le meilleur moyen d'assurer la qualité d'une liste de mots clés. Ces listes de mots clés sont aussi parfois appelées « lexiques » si elles se rapportent toutes à un thème commun. Par exemple, des lexiques du discours de haine dans différentes langues qui sont liés à différents contextes politiques sont souvent utilisés pour analyser le contenu politique. Même une session rapide d'une heure avec un expert en la matière peut grandement contribuer à introduire une certaine rigueur dans la compilation des listes de mots-clés pour garantir la qualité. Vous pouvez également consulter la littérature académique pertinente et utiliser des listes de mots clés précédemment compilées si elles sont de haute qualité et pertinentes pour votre contexte.

Exemples de lexiques

Il existe de nombreux exemples de lexiques disponibles en ligne et utiles pour l'analyse linguistique et l'extraction de messages pertinents. L'équipe Genre, Femmes et Démocratie du NDI a compilé un lexique du discours de haine relatif aux langues indonésienne, kenyane et serbe (Zeiter et al., 2019). PeaceTech Lab, une organisation à but non lucratif dédiée à l'utilisation de la technologie pour promouvoir la paix dans les pays en développement, propose également plusieurs lexiques de discours de haine accessibles au public et disponibles gratuitement sur son site Web. Ces lexiques sont multilingues, ce qui peut être extrêmement utile pour surveiller des conversations dans plusieurs endroits ou dans des régions à forte diversité linguistique.

Hatebase.org propose également des listes de mots-clés multilingues pour les discours de haine. Les chercheurs Roya Pakzad et Nilhoufar Salehi ont utilisé une liste compilée à partir de ce site Web pour leur étude de la propagande informatique ciblant les musulmans américains lors des élections de 2018 aux États-Unis (2019). Une liste similaire a été utilisée pour une analyse quantitative de l'islamophobie sur Gab à l'approche des mêmes élections (Woolley, Pakzad & Monaco, 2019).

Analyse linguistique qualitative

Une fois que vous avez extrait un ensemble pertinent de posts, vous pouvez vous plonger dans l'analyse linguistique. Les méthodes qualitatives sont celles qui analysent le contenu et les thèmes des messages des posts de votre ensemble de données. Ce type de travail est plus efficace lorsqu'il est effectué par des humains et par un expert ou une équipe d'experts connaissant bien le contexte politique et linguistique de la région en question.

Que vous utilisiez des méthodes qualitatives, quantitatives ou les deux lors de votre analyse, il est très important que vos méthodes soient cohérentes et systématiques. Utiliser les mêmes méthodes pour analyser chaque publication ou partie de votre ensemble de données est le meilleur moyen de se prémunir contre l'introduction de biais dans vos résultats.

Analyse narrative

L'analyse narrative est un moyen intensif d'analyser le contenu linguistique d'un ensemble de données, mais elle peut

être très intéressante. L'analyse narrative, qui consiste à analyser la manière dont certains médias font référence à un sujet donné, en particulier dans le temps, peut mettre en lumière le positionnement et l'influence des médias sur l'opinion publique.

Codage qualitatif

Une autre méthode qualitative utile pour comprendre quels types de messages se produisent dans un ensemble de données particulier est appelée codage qualitatif. Il s'agit d'une équipe d'experts attribuant indépendamment l'une des catégories prédéfinies (ou « codes ») aux messages un par un. Une fois les catégories attribuées à chaque tweet, une analyse quantitative peut être effectuée.

Par exemple, dans une étude hypothétique sur une élection présidentielle entre deux candidats A et B dans le pays imaginaire de Qumar[31], on peut imaginer cinq catégories possibles pour les messages collectés autour de l'élection :

1. Pro-candidat A
2. Pro-candidat B
3. Anti-candidat A
4. Anti-candidat B
5. Neutre

Une fois qu'une équipe d'experts a codé tous les messages de l'ensemble de données, nous pouvons entreprendre plusieurs analyses quantitatives dans la prochaine étape :

- Répartition des posts dans chaque catégorie - Nous pourrions analyser le nombre de posts publiés dans chaque catégorie pour examiner s'il y avait un plus grand soutien ou une plus grande opposition en ligne au candidat A ou au candidat B.
- Utilisation de mots-clés par catégorie - Nous pourrions également examiner les messages de chaque catégorie pour y identifier les mots-clés pertinents, pour voir si les partisans ou l'opposition de l'un ou l'autre des candidats utilisent certains mots.
- Contenu de bots dans chaque catégorie - Une autre possibilité consiste à examiner les posts de chaque catégorie pour y déceler le contenu provenant de bots afin de déterminer si l'un ou l'autre des candidats recueille un plus grand soutien ou une plus grande opposition de la part d'agents automatisés en ligne.

Ce ne sont là que quelques exemples d'analyses qui peuvent être entreprises après un codage qualitatif de haute qualité.

Analyse linguistique quantitative

L'analyse linguistique quantitative est également connue sous le nom de traitement automatique du langage naturel (TALN) ou linguistique informatique. Les méthodes utilisées dans l'analyse linguistique quantitative découlent de l'idée que nous pouvons obtenir certaines informations sur les types et les fréquences de certains messages lorsque nous considérons la langue d'un point de vue statistique. En termes plus concrets, si nous agrégeons des mots ou des ensembles de mots provenant de publications pertinentes dans notre ensemble de données, nous pouvons examiner quels types de thèmes émergent. Cette sous-section vous présentera certaines des bases de l'extraction du nombre de mots à partir d'un ensemble de données pour examiner des thèmes communs en ligne.

Bien que les techniques de cette section puissent être mises en œuvre avec n'importe quel langage de programmation, il convient de noter que Python et R sont fournis avec plusieurs modules qui simplifient le processus. Ces modules, tels que le Natural Language Toolkit (NLTK) de Python ou le module tm (tm pour text mining) de R, disposent d'une documentation complète sous forme de livres, de guides en ligne et de tutoriels qui peuvent vous enseigner les principes de base en quelques heures. Nous vous recommandons fortement de prendre le temps de vous familiariser

31 Qumar est un pays fictif qui apparaît dans la série dramatique politique américaine The West Wing.

avec l'un de ces modules, surtout si vous connaissez déjà quelques notions de Python ou R. Quelques heures supplémentaires au départ vous feront gagner beaucoup de temps à vous et à votre équipe par la suite !

N-grammes

Les N-grammes sont la pierre angulaire de presque toutes les analyses linguistiques quantitatives du contenu des réseaux sociaux. Un n-gramme est une suite de mots de longueur n. Trois types de n-grammes en particulier sont les plus courants pour l'analyse linguistique quantitative de textes :

- Unigrammes - Un seul mot peut également être envisagé comme une séquence d'un (1) mot. Ce type de n-gramme est connu sous le nom d'unigramme.
- Bigrammes - Deux mots apparaissant l'un à côté de l'autre dans une phrase forment un bigramme. En d'autres termes, un bigramme est une séquence de 2 mots qui apparaît dans un texte.
- Trigrammes - À ce stade, vous avez probablement deviné qu'un trigramme est un ensemble de trois mots qui apparaît dans une phrase. Chaque séquence de trois (3) mots dans une seule phrase ou texte forme un trigramme distinct.

Il est plus courant de travailler avec ces trois types de n-grammes. Cela est largement dû au fait que l'utilisation de valeurs plus élevées pour n est « coûteuse en calcul » - vous risquez de ralentir les performances de votre ordinateur et de ne pas en avoir pour votre argent. Pour les séquences de mots supérieures à 3, il est d'usage de se référer simplement à ceux-ci par le nombre en question suivi du mot « -gramme ». Par exemple, des 4-grammes seraient une séquence de 4 mots, des 5-grammes seraient une séquence de 5 mots, etc.

L'idée maîtresse des n-grammes est que vous disposez non seulement du nombre de mots, mais aussi d'une petite partie du contexte dans lequel un mot apparaît. Le fait de savoir que vous avez 16 occurrences du mot « roi » vous indique que vous avez un sujet fréquent dans les mains, mais pas beaucoup plus que cela. Si vous étendez votre analyse aux 4-grammes et constatez que vous avez 15 occurrences de « à bas le roi » et une seule occurrence de « longue vie au roi », vous pouvez dire avec certitude que la plupart des messages liés au « roi » dans votre ensemble de données ne sont pas des messages de soutien.

Une fois que votre équipe est à l'aise avec la technique d'extraction de n-grammes et de comptage des fréquences de n-grammes dans un texte, vous pouvez appliquer cette technique à différentes parties de votre ensemble de données. Par exemple, comparer des n-grammes entre des comptes ou des pages qui soutiennent différents candidats/partis politiques peut être instructif. La comparaison des fréquences de n-grammes de différents médias produisant des articles pertinents pour la même élection peut donner un aperçu de l'objectif principal de chaque média. Les comparaisons de fréquence de n-grammes entre les historiques de publication de deux comptes ou pages distincts peuvent révéler les sujets favoris de leurs échanges de messages. Ce ne sont là que quelques exemples de possibilités qui pourraient vous aider à améliorer votre recherche.

Autres techniques linguistiques quantitatives

L'assemblage de fréquences de n-grammes est une technique qui peut être appliquée dans un premier temps à d'autres techniques de TALN qui peuvent aider à analyser un ensemble de données plus en détail. Par exemple, après avoir extrait les fréquences de mots ou les fréquences de n-grammes, une technique statistique connue sous le nom de term frequency-inverse document frequency (tf-idf) peut être utilisée pour comparer les thèmes des messages de différents documents les uns par rapport aux autres. Ces « documents » pourraient être des collections d'articles de différents organes d'information menant à l'élection ou des histoires de publication pour différents comptes d'intérêt, par exemple. Tf-idf est un moyen de base pour déterminer quels thèmes uniques distinguent un document d'un autre. Cette technique peut être utilisée pour analyser quels comptes font le plus souvent la promotion d'un parti donné, ou ce qui distingue les messages d'une communauté de tous les autres que vous examinez.

Une introduction rapide sur la façon d'utiliser tf-idf pour l'analyse linguistique[32] et le regroupement de contenu peut être trouvé dans Mining the Social Web (Russell et Klassen 2018). Cette discussion vaut la peine d'être consultée - elle comprend des exemples de code en Python et est particulièrement facile à suivre.

Détection automatisée des discours de haine

Pour analyser le discours de haine dans un ensemble de données, vous devrez probablement d'abord compiler une liste de termes de discours de haine sensibles qui sont pertinents pour la région que vous examinez. Ces termes doivent être examinés en détail avec un expert en la matière et dans toutes les langues susceptibles d'être utilisées dans la région en question, méthodologie décrite dans le rapport du NDI « Tweets That Chill ». (Zeiter et al., 2019)

Une heuristique pour la détection automatique des discours de haine sans compilation de mots clés utilise l'API Perspectives de Google Jigsaw, un outil open source introduit en 2017. L'API Perspectives ne prend actuellement en charge que les commentaires en anglais[33]. Elle prend une chaîne de texte comme entrée et génère un score de toxicité pour ce texte. Plus le score est élevé, plus l'énoncé est jugé « toxique » par les modèles d'apprentissage automatique de l'API Perspectives. Un exemple d'utilisation de cette API peut être trouvé dans l'ensemble de données Gab de Pushshift.io. En plus de contenir les messages de Gab, Pushshift.io a exécuté chaque message via l'API Perspectives et a enregistré le score de toxicité qu'il a reçu.

L'API Perspectives est actuellement l'un des seuls outils accessibles au public qui attribue des scores de toxicité à la langue. Il faut beaucoup de temps et d'efforts pour saisir les nuances de contexte et d'intention dans le langage humain, et pour cette raison, les outils d'analyse des sentiments et de détection de discours de haine en sont encore à leurs balbutiements. Outre la difficulté générale inhérente au problème, la spécificité diverse du contexte de chaque langue et du contexte social dans lequel se produit le discours de haine aggrave la difficulté de produire des outils automatisés de détection de discours de haine rigoureux et fiables. Pour ces raisons, travailler main dans la main avec des experts en la matière pour compiler des mots-clés et des lexiques pertinents reste la meilleure méthode pour analyser les discours de haine dans un ensemble de données de réseaux sociaux.

Bots

On a beaucoup écrit ces dernières années sur l'influence des bots sur les réseaux sociaux. Le Computational Propaganda Project (ComProp) de l'Université de Washington et l'Oxford Internet Institute (OII) ont fait un travail de pionniers sur les robots, explorant leur influence dans la diffusion et la promotion de la propagande informatique dans les pays du monde entier.

Les bots, en termes simples, sont des programmes informatiques qui contrôlent les profils sur les sites de réseaux sociaux, se faisant souvent passer pour de vraies personnes et interagissant avec d'autres humains en ligne. Dans le cas des réseaux sociaux, la plupart des bots sont contrôlés par des programmes informatiques qui contrôlent les comptes via l'API[34]. Les plus simples de ces bots ont souvent des signes révélateurs qui les trahissent comme étant automatisés - tweeter à la même minute de chaque heure ou ne pas avoir de photo de profil, par exemple. Ces

32 Ce texte présente également certaines techniques pertinentes pour améliorer votre analyse, telles que la suppression des mots vides et l'indexation par radicaux.

33 L'anglais a bénéficié de manière disproportionnée de l'attention des linguistes, à la fois informatiques et non informatiques, tout au long des 20e et 21e siècles. Les langues qui n'ont pas bénéficié d'une grande attention et de recherches sont souvent appelées langues à faibles ressources en linguistique et en TALN. On ne saurait trop insister sur l'importance de produire des recherches sur ces langues. Si vous ou votre équipe entreprenez des recherches linguistiques ou sur les discours de haine dans le cadre de vos efforts de surveillance des médias, il peut être utile de contacter des linguistes professionnels de l'industrie ou du monde universitaire pour voir si votre recherche pourrait contribuer à apporter des informations précieuses dans le domaine.

34 Dans la section ci-dessus relative à la collecte de données, nous avons exploré les différences entre les API et le web scraping - cette différence est également utile pour comprendre comment certains bots fonctionnent sur les réseaux sociaux et sur le Web en général sans utiliser d'API. Les bots peuvent être programmés pour interagir avec les pages Web de tous les jours - en fait, c'est assez simple à faire pour les programmeurs. Par exemple, les bots peuvent aller d'un site à l'autre et analyser le contenu de pages Web. Ces types de bots sont souvent appelés robots d'exploration ou araignées, et c'est d'ailleurs de cette manière que Google collecte les données sur les pages web pour les mettre dans ses moteurs de recherche. De même que tous les bots ne sont pas des bots de réseaux sociaux, tous les bots ne sont pas mauvais, et en fait certains d'entre eux sont nécessaires au fonctionnement quotidien de l'Internet tel que nous le connaissons.

données sont souvent utilisées dans les algorithmes d'apprentissage automatique pour les outils qui distinguent les bots des humains en ligne.

Outils de détection de bots

Bien qu'il n'y ait le plus souvent aucun moyen infaillible de dire avec certitude si un compte est un bot ou non, il existe des options utiles à la disposition du chercheur curieux pour détecter les bots en ligne. L'une des meilleures options actuellement disponibles est Botometer, un outil qui utilise l'apprentissage automatique pour attribuer un score de probabilité de bot aux comptes étudiés. Botometer a une longue histoire de recherche universitaire et de développement à l'Université de l'Indiana et est libre d'utilisation. Un certain nombre d'autres classificateurs existent également (voir tableau ci-dessous).

Outils de détection de bots

Nom de l'outil	Peut-il être mis en œuvre par du code pour la classification des lots de comptes ?	Plateforme[35]	Extensions/site Web disponibles ?
Botometer	Oui (Python)	Twitter	Les comptes individuels peuvent être vérifiés sur le site Web
Tweetbotornot	Oui (R)	Twitter	Seule l'implémentation du code est disponible.
Botcheck.me	Non	Twitter	Extension Firefox, mises à jour pour la dernière fois en 2019.
Botsentinel	Non	Twitter	Application Android et extension Chrome/Firefox
Pegabot	Non	Twitter	Site Web

Conclusion

En fin de compte, toutes ces techniques et tous ces outils devraient aider à former l'approche de l'utilisateur en matière de collecte et d'analyse des données. Les utilisateurs doivent également envisager de créer un flux de travail utilisant ces techniques, un système d'archivage et de tri des informations collectées, et éventuellement de les signaler aux plateformes ou à d'autres organismes nationaux qui peuvent agir sur les données collectées. Selon le sujet, ils doivent considérer un certain nombre des différents outils et techniques décrits ci-dessus. Qu'il s'agisse de rechercher de la désinformation lors d'une élection ou des discours de haine et de la propagande informatique dans le discours politique traditionnel en ligne, ces ressources peuvent être appliquées de différentes manières. Les équipes et les chercheurs doivent également tenir compte des ressources dont ils disposent -humaines, financières et techniques- dans la construction de leur projet.

Bon nombre des outils mentionnés sont open source, mais d'autres sont coûteux et souvent complexes à appliquer sans une connaissance approfondie des outils et des méthodes impliqués dans leur développement. Il vaut souvent la peine d'envisager des solutions plus simples qui peuvent être appliquées par des chercheurs moins expérimentés, ou de travailler en collaboration avec des équipes dotées de compétences différentes pour découvrir différentes informations dans le même ensemble de données. Envisagez de créer des partenariats entre des chercheurs locaux

35 Les outils open source de détection de bots n'ont été développés que sur Twitter - en grande partie parce que l'API de Twitter offre un ensemble riche d'informations sur les utilisateurs (métadonnées publiques), qui sont utiles en tant que fonctionnalités pour construire des modèles d'apprentissage automatique qui catégorisent les comptes comme étant des bots ou des humains. L'analyse de l'activité des bots sur d'autres plateformes repose en grande partie sur une analyse et une enquête manuelles, telles que la détection d'activités surhumaines (100 posts/min, la publication de messages à intervalles réguliers, etc.). Par exemple, dans une étude sur l'islamophobie précédant les élections de 2018 aux États-Unis sur la plateforme de réseaux sociaux Gab, nous avons détecté la présence d'un bot de désinformation en analysant des posts identiques d'un utilisateur se produisant à de courts intervalles.

et des experts techniques internationaux pour obtenir de nouveaux types d'informations, et travaillez à la création de méthodes de collaboration en ligne entre les communautés, les pays et les régions.

La Design 4 Democracy Coalition est une façon de trouver des partenaires avec lesquels collaborer, dans votre contexte local et international, et de nouer des liens avec des entreprises de réseaux sociaux pour des rapports et des recherches. La D4D Coalition est un groupe d'ONG internationales comprenant le NDI, l'International Republican Institute, l'International Foundation for Election Systems, International IDEA et des organisations nationales de la société civile du monde entier s'engageant avec des entreprises technologiques telles que Facebook, Microsoft et Twitter pour les encourager à concevoir leurs systèmes, leur modération de contenu et leurs politiques selon des principes démocratiques. La recherche et le suivi en ligne sont devenus des éléments cruciaux des élections et des démocraties dans le monde entier. Ce guide a été rédigé pour donner aux groupes qui travaillent à soutenir ces idées par le biais de politiques et de systèmes techniques les outils, les méthodes et les capacités nécessaires pour faire ce travail et contribuer à mieux informer la société.

Voir ci-dessous pour d'autres références, des outils open source et des exemples de code, et une procédure pas à pas pour utiliser l'API de Twitter.

Annexe I : Exemple de code API - Collecte de données à partir des API de recherche et de diffusion de Twitter avec le module Rtweet

Dans cette section, nous allons brièvement vous présenter du code que vous pouvez utiliser pour collecter des données Twitter historiques et en direct. Avec seulement quelques lignes de code, vous pouvez facilement collecter des données Twitter pertinentes pour votre contexte électoral. Après avoir collecté les données cibles, vous pouvez exporter les données souhaitées au format CSV, procéder à leur analyse dans Excel ou Google Sheets, ou les transmettre à un expert en mégadonnées dédié de votre équipe pour rechercher des informations plus approfondies. Pouvoir collecter des données par vous-même en temps réel, même si vous ne disposez peut-être pas d'expert en mégadonnées dédié dans votre équipe au moment de la collecte, est extrêmement utile. De manière contre-intuitive, les données deviennent souvent plus précieuses au fil du temps, car elles permettent de capturer la désinformation, les bots et les acteurs malveillants dont les actions et les messages peuvent être supprimés ultérieurement de la plateforme pour violation des conditions de service.

Étape 1 : Télécharger R, RStudio et Rtweet

Dans ce guide, nous vous apprendrons à extraire des données de base de l'API Twitter à l'aide du langage de programmation R. Il existe un module spécifique dans R, appelé rtweet, qui rend l'extraction de données de Twitter extrêmement facile.

Vous devrez télécharger et installer quelques éléments pour être opérationnel.

- Rstudio : https://www.rstudio.com/products/rstudio/download/
- R : https://www.r-project.org/

Étape 2 : Demander une application Twitter, récupérer vos clés d'application et créer votre jeton

Afin de collecter des données sur Twitter, vous utilisez une application - il s'agit d'un moyen sécurisé d'utiliser votre compte pour interagir avec la plateforme via du code. Par le passé, les applications Twitter étaient gratuites et aucune approbation n'était nécessaire. De nos jours, vous devez déposer une demande, ce que vous pouvez faire ici.

Une fois que vous avez récupéré votre application, vous devez vous connecter à Twitter sur un navigateur et vous rendre sur https://apps.twitter.com. Vous pouvez récupérer les clés de consommation et d'accès de votre application ici - il y aura quatre clés au total. Ces clés sont simplement des chaînes de texte que votre application utilisera pour vérifier qu'elle demande bien des données au nom d'une personne - dans ce cas, vous. Cette utilisation de clés pour l'authentification est un processus appelé oauth, qui signifie autorisation ouverte. Ce procédé a été développé pour permettre aux applications d'effectuer des actions en ligne pour le compte d'utilisateurs sans transférer leur mot de passe et leur nom d'utilisateur pour chaque action.

Une fois que vous avez vos clés et le nom de votre application, vous pouvez utiliser le code qui suit dans la capture d'écran ci-dessous pour créer votre « jeton » (token) d'autorisation. Une fois ce jeton créé, vous êtes prêt.e à commencer à utiliser l'API Twitter.

```
library(rtweet)
app<-"<app name here>"
consumer_key<-"<consumer key here>"
consumer_key_secret<-"<consumer key secret here>"
access_token<-"<access token here>"
access_token_secret<-"<access token secret here>"
create_token(app,consumer_key, consumer_key_secret, access_token, access_token_secret)
```

Les lignes de code R ci-dessus sont des lignes initiales que vous pouvez utiliser pour authentifier votre application Twitter et vous connecter à l'API Twitter. Une fois que vous avez demandé l'accès à une application Twitter et que vous avez été approuvé.e, vous pouvez récupérer

vos jetons de consommateur et d'accès sur Twitter. Vous utiliserez ces jetons pour authentifier votre application (indiquée dans le code ci-dessus). Une fois que votre application est authentifiée (c'est-à-dire que Twitter sait que l'application extrait des données pour vous et non pour quelqu'un d'autre), vous êtes prêt.e à commencer à interagir et à extraire des données de l'API Twitter.

Professeur à l'Université du Missouri et développeur de rtweet, le Dr. Mike Kearney détaille également le processus d'authentification étape par étape sur le site officiel de rtweet ici.

Collecte de données historiques :

Avec rtweet, vous pouvez facilement rassembler les Tweets utilisant un ou plusieurs hashtags d'intérêt à partir de l'API de recherche de Twitter. Cette API est historique, ce qui signifie que vous allez collecter les données historiques des 7 à 9 derniers jours qui correspondent à votre requête. Les requêtes sont simplement des critères qui vous intéressent dans les Tweets que vous collectez. Les requêtes peuvent contenir des hashtags, des mots-clés, des noms de compte, une URL, une combinaison de ceux-ci ou les quatre. Dans cette section, nous nous concentrerons sur les hashtags, mais le processus et la syntaxe sont exactement les mêmes pour les autres entités.

La fonction principale que vous utiliserez pour rechercher des tweets historiques s'appelle search_tweets().

```
my_data<-search_tweets("#ExampleHashtag", n=50000, retryonratelimit=TRUE)
```

Cette ligne de code R interroge l'API de recherche Twitter pour un maximum de 50 000 Tweets utilisant #ExempleHashtag au cours des 7 à 9 derniers jours, et enregistre les résultats dans une trame de données appelée « my_data ». Si le hashtag a été utilisé moins de 50 000 fois au cours de cette période, un plus petit nombre de Tweets sera renvoyé. Changer la valeur d'entrée pour « n » peut augmenter ou diminuer le nombre maximum de Tweets que vous extrairez.

```
multiple_hashtags<-search_tweets("#ExampleHashtag1 OR #ExampleHashtag2", n=50000, retryonratelimit=TRUE)
```

Cette ligne de code R est similaire à la requête ci-dessus, mais renvoie jusqu'à 50 000 Tweets contenant #ExempleHashtag1 ou #ExempleHashtag2. Cette syntaxe peut être utilisée pour interroger l'API de recherche de Twitter pour plusieurs hashtags. Dans cet exemple de code, les résultats sont enregistrés dans une trame de données appelée « multiple_hashtags ». Les hashtags multiples sont séparés par ce qu'on appelle des « opérateurs booléens » : il s'agit simplement de « ET » ou « OU ». Utilisez « AND » [ET] si vous souhaitez uniquement des Tweets contenant les deux hashtags ; « OR » [OU] renverra les tweets contenant soit l'un soit l'autre.

Streaming de données Twitter en temps réel :

Vous avez également la possibilité de diffuser des données Twitter en temps réel. Cette requête vous oblige à spécifier la durée pendant laquelle vous souhaitez diffuser des Tweets en secondes, ainsi que les entités que vous souhaitez diffuser (hashtags, URL, mots-clés, @-mentions, etc.).

La syntaxe pour le streaming des tweets dans rtweet est légèrement différente de celle utilisée pour interroger l'API de recherche. Lorsque vous diffuserez des Tweets, vous souhaiterez utiliser la fonction stream_tweets(). Les différents éléments de requête seront séparés par des virgules.

```
my_streamed_data<-stream_tweets(q="#ExampleHashtag1",timeout=60)
```

```
my_streamed_data_w_multiple_hashtags<-stream_tweets(q="#ExampleHashtag1,#ExampleHashtag2",timeout=60)
```

Les figures ci-dessus montrent comment utiliser la fonction stream_tweets() de rtweet pour collecter des Tweets en temps réel. Le chiffre du haut collecte tous les Tweets utilisant #ExempleHashtag1 pendant une fenêtre de streaming de 60 secondes. La figure du bas fait de même, mais collecte également les Tweets contenant #ExempleHashtag2.

Écriture des résultats dans un fichier de sortie CSV :

Un format de fichier couramment utilisé pour l'analyse des données est le fichier CSV, qui signifie « valeurs séparées par des virgules ». Dans un fichier CSV, chaque ligne représente un seul rang de la feuille de calcul[36], et chaque entité entre virgules représente une cellule - ou plus précisément, une valeur[37] au sein d'une cellule.

Un fichier CSV est essentiellement une feuille de calcul dans un format lisible par machine. Une fois que vous avez un fichier CSV des Tweets que vous avez collectés, vous pouvez le transmettre à un spécialiste des données pour une analyse plus approfondie, ou le charger dans un tableur tel que Microsoft Excel ou Google Sheets pour effectuer vous-même certaines analyses.

```
write_as_csv(my_data,"my_data_as_a_csv_file.csv")
```

Cette ligne de code R utilise le module rtweet pour écrire une trame de données de Tweets appelée « my_data » dans un fichier CSV appelé « my_data_as_a_csv_file.csv ». Après avoir exporté vos Tweets dans un fichier CSV, vos données peuvent être facilement partagées ou importées dans Microsoft Excel ou Google Sheets.

Réflexions finales

Pouvoir extraire des données de Twitter et les écrire dans un fichier CSV est une compétence inestimable. Même sans expérience en programmation, n'importe qui peut apprendre le processus que nous décrivons ici en moins de 2 heures, et probablement encore plus rapidement. Une fois que vous avez la possibilité de récupérer et de stocker des CSV de données Twitter pertinentes, vous et votre équipe êtes en mesure d'analyser des données précieuses maintenant et à l'avenir.

Si vous stockez ces données à long terme, vous pouvez également envisager de compresser le fichier dans un format .zip (ou un autre format, tel qu'un Tarball). Cela peut réduire la quantité d'espace de stockage nécessaire pour stocker le fichier, et faciliter la diffusion du fichier à d'autres personnes et dispositifs.

Annexe II : Outils OSINT

L'intelligence open source, communément appelée OSINT, est l'art d'enquêter sur une question en utilisant uniquement des informations et des données accessibles au public (ou « open source »). Dans le contexte de la désinformation et de la surveillance des réseaux sociaux, OSINT peut souvent fournir des détails supplémentaires sur des acteurs malveillants ou suspects en ligne, y compris de faux comptes ou des sites Web de désinformation. Vous trouverez ci-dessous une liste de ressources précieuses pour l'apprentissage des techniques OSINT.

- Liste publique des outils de vérification et d'OSINT du journaliste de BuzzFeed Craig Silverman : https://docs.google.com/document/d/1ZJbIUk5L8fe3VKK9CLVNMj9qOFdXG-RhQT6pyEgsS4I/edit
- Site et livre de Michael Bazzell :
  - https://inteltechniques.com
  - Open-Source Intelligence Techniques
- Boîte à outils d'enquête en ligne de Bellingcat : https://docs.google.com/document/d/1BfLPJpRtyq4RFtHJoNpvWQjmGnyVkfE2HYoICKOGguA/edit

36 Ces lignes sont parfois désignées par d'autres termes utilisés par les spécialistes des données : enregistrement, instance et observation sont d'autres termes que vous êtes susceptible d'entendre dans ce contexte, qui sont tous synonymes de « rang » lorsqu'il s'agit d'un csv.

37 Les valeurs d'un fichier csv ou d'une feuille de calcul peuvent également parfois être appelées champs.

- Comprop Navigator, publié par le projet Computational Propaganda de l'Oxford Internet Institute, compile des méthodes et des outils OSINT liés à la désinformation et d'autres recherches en ligne. https://navigator.oii.ox.ac.uk/
- First Draft Guide on NewsGathering and Monitoring on the Social Web https://firstdraftnews.org/wp-content/uploads/2019/10/Newsgathering_and_Monitoring_Digital_AW3.pdf?x36710
- Fighting Disinformation Online: A Database of Web Tools, hébergée par la Rand Corporation. https://www.rand.org/research/projects/truth-decay/fighting-disinformation.html
- Verification Handbook for Disinformation and Media Manipulation https://datajournalism.com/read/handbook/verification-3/

Références :

Allcott, H., Gentzkow, M., & Yu, C. (2018). Trends in the Diffusion of Misinformation on Social Media. https://web.stanford.edu/~gentzkow/research/fake-news-trends.pdf

Borgatti, S., Everett, G. et Johnson, J. (2013). Analyzing Social Networks.

Democracy Reporting International. (Octobre, 2019) Guide for Civil Society on Monitoring Social Media During Elections. https://www.ndi.org/sites/default/files/social-media-DEF.pdf

Jack, C. 2017. Lexicon of Lies. Data & Society. https://datasociety.net/output/lexicon-of-lies/

Monaco, N. (2019). Welcome to the Party: A Data Analysis of Chinese Information Operations. Extrait de https://medium.com/digintel/welcome-to-the-party-a-data-analysis-of-chinese-information-operations-6d48ee186939

National Democratic Institute. (mai 2019) Disinformation and Electoral Integrity: A Guidance Document for NDI Elections Programs. https://www.ndi.org/publications/disinformation-and-electoral-integrity-guidance-document-ndi-elections-programs

National Democratic Institute. (décembre 2018). Supporting Information Integrity and Civil Political Discourse. https://www.ndi.org/publications/supporting-information-integrity-and-civil-political-discourse

Pakzad, R., & Salehi, Ni. (2019). Anti-Muslim Americans: Computational Propaganda in the United States. Institute from the Future. Extrait de http://www.iftf.org/fileadmin/user_upload/downloads/ourwork/IFTF_Anti-Muslim_comp.prop_W_05.07.19.pdf

Russell, M., & Klassen, Mikhail. (2018). Mining the social web. Sebastopol, CA: O'Reilly Media.

Wardle, C., & Derakhshan, H. 2017. Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking. Conseil de l'Europe. https://rm.coe.int/information-disorder-toward-an-interdisciplinaire-framework-for-researc/168076277c

Woolley, S., Pakzad, R. et Monaco, N. (2019). Incubating Hate: Islamophobia and Gab. German Marshall Fund. https://www.gmfus.org/news/incubating-hate-islamophobia-and-gab

Woolley, S., & Howard, P. 2017. Computational Propaganda Worldwide: Executive Summary. Documents de travail. 2017.11. Oxford, Royaume-Uni : Project on Computational Propaganda. http://blogs.oii.ox.ac.uk/physicalbots/wp-content/uploads/sites/89/2017/06/Casestudies-ExecutiveSummary.pdf

Zeiter, K., Pepera, S., Middlehurst, M., Ruths, D. (2019). Tweets That Chill: Analyzing Online Violence Against Women in Politics. National Democratic Institute. https://www.ndi.org/tweets-that-chill

NATIONAL
DEMOCRATIC
INSTITUTE

NDI.ORG