



HAL
open science

Les media sociaux lors de la crise sanitaire de Covid-19 Circulation de l'information et initiatives citoyennes

Camozzi Marie-Lys, Thubert Nathan, Julien Coche, Sandrine Bubendorff,
Robin Batard, Aurelie Montarnal, Caroline Rizza

► To cite this version:

Camozzi Marie-Lys, Thubert Nathan, Julien Coche, Sandrine Bubendorff, Robin Batard, et al..
Les media sociaux lors de la crise sanitaire de Covid-19 Circulation de l'information et initiatives
citoyennes. [Rapport de recherche] SID - Sociologie Information-Communication Design; SES - Dé-
partement Sciences Economiques et Sociales; IMT Mines Albi. 2020. hal-03041133

HAL Id: hal-03041133

<https://hal.science/hal-03041133v1>

Submitted on 4 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Les media sociaux lors de la crise sanitaire de Covid-19

Circulation de l'information et initiatives citoyennes

Marie-Lys CAMOZZI

IMT Mines Albi

Nathan THUBERT

IMT Mines Albi

Julien COCHE

IMT Mines Albi

julien.coche@mines-albi.fr

Sandrine BUBENDORFF

i3-SES, Télécom Paris, CNRS, Institut Polytechnique de Paris

sandrine.bubendorff@telecom-paris.fr

Robin BATARD

i3-SES, Télécom Paris, CNRS, Institut Polytechnique de Paris / IMT Mines Albi

robin.batard@telecom-paris.fr

Aurélie MONTARNAL

IMT Mines Albi

aurelie.montarnal@mines-albi.fr

Caroline RIZZA

i3-SES, Télécom Paris, CNRS, Institut Polytechnique de Paris

caroline.rizza@telecom-paris.fr

Working Paper 20-SES-01

Novembre, 2020

Pour citer ce papier / How to cite this paper: Camozzi M-L., Thubert N., Coche J., Bubendorff S., Batard R., Montarnal A., & Rizza C. (2020) Les media sociaux lors de la crise sanitaire de Covid-19 : circulation de l'information et initiatives citoyennes. i3 Working Papers Series, 20-SES-01

L'institut interdisciplinaire de l'innovation

(UMR 9217) a été créé en 2012. Il rassemble :

- les équipes de recherche des MINES ParisTech en économie (CERNA), gestion (CGS) et sociologie (CSI),
- celles du Département Sciences Économiques et Sociales (DSES) de Télécom Paris,
- ainsi que le Centre de recherche en gestion (CRG) de l'École polytechnique,

soit plus de 200 personnes dont une soixantaine d'enseignants chercheurs permanents.

L'institut développe une recherche de haut niveau conciliant excellence académique et pertinence pour les utilisateurs de recherche. Par ses activités de recherche et de formation, i3 participe à relever les grands défis de l'heure : la diffusion des technologies de l'information, la santé, l'innovation, l'énergie et le développement durable. Ces activités s'organisent autour de quatre axes :

- Transformations de l'entreprise innovante
- Théories et modèles de la conception
- Régulations de l'innovation
- Usages, participation et démocratisation de l'innovation

Pour plus d'information : <https://i3.cnrs.fr/>

Ce document de travail est destiné à stimuler la discussion au sein de la communauté scientifique et avec les utilisateurs de la recherche. Son contenu est susceptible d'avoir été soumis pour publication dans une revue académique. Il a été examiné par au moins un referee interne avant d'être publié. Les considérations exprimées dans ce document sont celles de leurs auteurs et ne sont pas forcément partagées par leurs institutions de rattachement ou les organismes qui ont financé la recherche.

The Interdisciplinary Institute of Innovation

(UMR 9217) was founded in 2012. It brings together:

- the MINES ParisTech economics, management and sociology research teams (from the CERNA, CGS and CSI),
- those of the Department of Economics and Social Sciences (DSES) at Télécom Paris,
- and the Management Research Center (CRG) at École Polytechnique,

meaning more than 200 people, including 60 permanent academic researchers.

i3 develops a high-level research, combining academic excellence and relevance for the end users of research. Through its teaching and research activities, i3 takes an active part in addressing the main current challenges: the diffusion of communication technologies, health, innovation, energy and sustainable development. These activities are organized around four main topics:

- Transformations of innovating firms
- Theories and models of design
- Regulations of innovation
- Uses, participation and democratization of innovation

For more information: <https://i3.cnrs.fr/>

This working paper is intended to stimulate discussion within the research community and among research users. Its content may have been submitted for publication in academic journals. It has been reviewed by at least one internal referee before publication. The views expressed in this paper are those of the author(s) and not necessarily those of the host institutions or funders.

RESUME

Le projet MESCOV « les media sociaux lors de la crise Covid-19 » traite des aspects création et circulation de l'information sur les media sociaux lors de la crise Covid-19, des initiatives citoyennes qui y ont émergé et des pratiques des professionnels de la gestion de crise associées (notamment Service d'Incendie et de Secours et Préfecture). Financé par le Comité analyse, recherche et expertise (CARE) du Ministère de l'Éducation Supérieure, de la Recherche et de l'Innovation, il est une spécialisation du projet ANR MACIV « Management of citizens and volunteers in crisis management : the social media contribution » (ANR n°ANR-17-CE39-0015) sur ces aspects. C'est un projet pluridisciplinaire qui mobilise à la fois les Sciences de l'Informatique et de la Donnée pour le module base de données et algorithme d'apprentissage automatique, et les Sciences Humaines et Sociales pour la partie documentation des mécanismes de création, de circulation et de vérification de l'information sur les media sociaux, l'émergence d'initiatives citoyennes et l'utilisation des media sociaux par les institutionnels. Ce working paper présente une synthèse des principaux résultats du projet. Une liste des livrables techniques et scientifiques est proposée à la fin du document

MOTS CLE

Media sociaux ; gestion de crise ; pandémie ; COVID-19 ; crisis informatics ; Wikipedia ; Twitter ; initiatives citoyennes ; reconnaissance d'entités nommées ; métamodèle

ABSTRACT

The MESCOV project "Social media at the time of Covid-19" aims to study the creation and circulation of information on social media, the citizens' initiatives which emerged on social media and the professional practices of crisis management (in particular SDIS and Prefecture), during the Covid-19 crisis. It is a specialization of the ANR MACIV Project « Management of citizens and volunteers in crisis management: the social media contribution » (ANR n°ANR-17-CE39-0015) on these aspects. It is a multidisciplinary project extending both Computer and Data Sciences regarding the dataset and the machine learning algorithm modules, and Social Sciences regarding the study of creation, circulation, and usage of social media by the institutions, the emergence of citizen initiatives, and the use of social media by institutions and practitioners. This working paper presents a summary of the main results of the project. A list of all technical and scientific deliverables can be found at the end of the document.

KEYWORDS

Social media; crisis management; pandemic; COVID-19; crisis informatics; Wikipedia; Twitter; grassroot initiatives; named entity recognition; metamodel

Table des matières

Introduction	5
I. L'apport d'un modèle d'apprentissage semi supervisé à la gestion de la crise sanitaire	6
Section 1 : État de l'art	6
Section 2 : Conceptualisation de l'information à extraire dans le cadre de MESCOV	12
Section 3 : Application d'un algorithme semi-supervisé de reconnaissance d'entités nommées relatives à la crise de Covid-19	15
II. Création et circulation de l'information sur les media sociaux et initiatives citoyennes lors de la crise Covid-19	22
Section 1 : Faire collectivement sens au temps de la crise de Covid-19.....	22
Section 2 : Analyse descriptive de la base de données de tweets en langue française	30
Section 3 : Les usages de Wikipédia lors de la pandémie	36
Section 4 : Un exemple d'organisation citoyenne sur les media sociaux en temps de crise - les <i>makers</i>	45
III. Perspectives	54
IV. Bibliographie	56
VII. Liste des tableaux	68
VIII. Table des figures.....	69
IX. Tableau des livrables	70

Introduction

Le projet MESCOV « les media sociaux lors de la crise Covid-19 » traite des aspects de création et de circulation de l'information sur les media sociaux lors de la crise Covid-19, des initiatives citoyennes qui y ont émergé et des pratiques des professionnels de la gestion de crise associées (notamment Service d'Incendie et de Secours et Préfecture). Il est une spécialisation du projet ANR MACIV « Management of citizens and volunteers in crisis management : the social media contribution » (ANR n°ANR-17-CE39-0015) sur ces aspects, qui étudie la complétude des enjeux associés aux media sociaux en gestion de crise : enjeu technologique, par les outils permettant une remontée automatisée des informations nécessaires aux acteurs institutionnels (Coche et al., 2020); enjeu institutionnel, par le statut de l'information issue des media sociaux et son intégration sur le terrain (Castagnino, 2019); enjeu citoyen, à la fois par les mécanismes de création et de circulation des informations sur les media sociaux (Bubendorff et al., 2019; Bubendorff & Rizza, 2020) et par l'intégration des initiatives citoyennes à la réponse à la crise (Batard, Rizza, et al., 2019; Rizza, 2020).

On a communément l'habitude de représenter les media sociaux comme une infrastructure qui permet d'une part aux institutionnels (ministères, préfectures, communes, services d'incendie et de secours) de communiquer vers les citoyens du haut vers le bas (« top-down ») et d'autre part d'améliorer l'état des lieux de l'évènement par une remontée des informations citoyennes du bas vers le haut (« bottom-up ») (Rizza, 2020). La littérature dans le domaine des « *crisis informatics* » (Palen et al., 2020) a mis en évidence les changements amenés par les media sociaux et comment les citoyens les ont utilisés pour progressivement communiquer sur un évènement, informer, et s'organiser pour aider.

Nous retenons la coexistence d'une dimension informationnelle et d'une dimension organisationnelle à l'apport des media sociaux en gestion de la crise : Informationnelle car l'ensemble du contenu publié constitue une source d'informations pertinentes pour évaluer ce qui se passe sur place et organisationnelle lorsqu'il s'agit de collaborer pour répondre à la crise.

MESCOV est un projet pluridisciplinaire, voire interdisciplinaire, qui mobilise à la fois les Sciences de l'Informatique et de la Donnée pour l'algorithme d'apprentissage automatique appliqué à la crise sanitaire de Covid-19 qu'il s'est proposé de développer, et les Sciences Humaines et Sociales pour la partie documentation des mécanismes de création, de circulation et de vérification de l'information sur les media sociaux ainsi que l'émergence d'initiatives citoyennes et l'utilisation des media sociaux par les institutionnels. Ce working paper présente une synthèse de l'état de l'art et des principaux résultats du projet. Une liste des livrables techniques et scientifiques est proposée à la fin du document.

I. L'apport d'un modèle d'apprentissage semi supervisé à la gestion de la crise sanitaire

Le projet ANR MACIV a notamment permis le développement d'un modèle d'apprentissage automatique permettant d'extraire les informations importantes pour la compréhension d'une situation de crise à partir de données brutes issues de Twitter (Coche et al., 2020, 2019). Pendant les six mois de travaux spécifiques sur MESCOV, il s'est agi d'entraîner cet algorithme d'apprentissage automatique sur les aspects de la crise sanitaire au moyen d'une base de données de tweets co-construite de manière interdisciplinaire (Section 2) ainsi que de créer un méta-modèle permettant de rendre compte de la situation sanitaire et faciliter les prises de décisions relatives à la crise de Covid-19 (Section 4).

Cette première partie est structurée comme il suit : Section 1 « état de l'art » ; Section 2 « Conceptualisation de l'information à extraire dans le cadre de MESCOV » ; Section 3 « Application d'un algorithme semi-supervisé de reconnaissances d'entités ».

Section 1 : État de l'art

1. Fournir une information actionnable aux gestionnaires de crise, à partir des réseaux sociaux

La littérature montre que les réseaux sociaux sont devenus, en l'espace de quelques années, un vecteur capital de communication et de diffusion de l'information. Déjà en 2012, (Cameron et al., 2012; Terpstra et al., 2012) montraient que les institutionnels de la gestion de crise s'intéressaient à l'extraction d'information à partir de données issues des réseaux sociaux, que ce soit sous la forme de rapports, d'un tri "pertinent" ou encore de résumés des informations recueillies. Si l'opportunité d'une nouvelle source d'information pour comprendre le théâtre de la crise et mieux y répondre est bien assimilée, il n'en reste pas moins que son exploitation peut devenir coûteuse en termes de ressources humaines, alors même que la situation requiert des gestionnaires des réactions rapides. C'est ainsi que s'est ouvert un nouveau pan de l'utilisation de l'intelligence artificielle (IA) en gestion de crise. Dans (Imran et al., 2018), les auteurs résument les usages communément faits des données issues des réseaux sociaux en gestion de crise, à partir d'une revue de la littérature : classifications binaires visant à définir si la donnée est pertinente ou non, segmentation des données permettant de regrouper des tweets sémantiquement proches et ainsi réduire le nombre d'information que l'humain devra *in fine* lire et interpréter, génération automatisée de résumés permettant une fois encore de réduire le volume d'informations à traiter humainement, la vérification de l'information, la reconnaissance d'entités nommées (Named Entity Recognition NER) permettant d'enrichir sémantiquement les mots d'un tweet et finalement l'exploitation de données multimédia telles que les images. En termes d'objectifs à atteindre, un en particulier reste récurrent et a longtemps été le principal lorsqu'il est question d'exploiter par l'IA des données issues des réseaux sociaux : la *situational awareness* - soit la compréhension de la situation de crise, définie dans (Vieweg et al., 2010). Le but en est, concrètement, de fournir aux cellules de crise une image commune de la situation de

crise (common operational picture - COP). À l'instar des communautés de la gestion de crise, telles qu'ISCRAM (Information Systems for Crisis Response and Management), l'engouement a été solide et nourri, de telle sorte que de nombreux exemples de systèmes informatiques ou d'IA peuvent être cités, tels que Tweedr (Ashktorab et al., 2014), Twitris (Jadhav et al., 2010), ESA (Cameron et al., 2012), AIDR (Imran et al., 2014) ou encore Twitcident (Abel et al., 2012).

Plus récemment, la question de l'*actionability* (dont il n'existe pas de traduction française fixée, et qui pourrait faire référence à *une capacité d'action*, soit le degré d'utilité de l'information pour prendre une décision) a été soulevée. En particulier, dans (Coche et al., 2019; Zade et al., 2018), les auteurs explicitent l'importance d'aller au-delà de la *situational awareness* et fournir, en sortie du module IA de traitement des données issues des réseaux sociaux, des informations ciblées sur les besoins précis des décisionnaires. C'est bien dans ce cadre d'extraction de l'information, grâce aux réseaux sociaux, que se situe cette étude.

2. Quels types d'informations extraire des réseaux sociaux pour les gestionnaires de crise ?

Le contexte de la crise est, comme nous l'avons dit, particulier : bien qu'inattendu et imprévisible, il faut savoir réagir vite et avec pertinence. Pour cela les informations dites *actionnables* à transmettre aux gestionnaires de la crise doivent être ciblées sur leurs besoins précis et les contraintes contextuelles doivent être prises en compte dans l'implémentation du module d'IA d'extraction de l'information.

2.1. Méta-modèles liés à la gestion de crise

L'implémentation d'une IA d'extraction de l'information s'appuie sur un modèle (plus spécifiquement méta-modèle) présentant les différents concepts à trouver dans le flux de données textuelles issu des réseaux sociaux. Dans le cadre de la gestion de crise, plusieurs méta-modèles/ontologies ont été développés et ceux de référence sont rappelés dans le récent article (Benaben et al., 2020). Le méta-modèle COSIMMA, dédié à la mise en place d'une réponse collaborative à une crise, et présenté dans ce même article, est particulièrement couvrant, avec un ensemble de classes dédiées à la description du contexte (concepts : *biens, population, site naturel, société civile, territoire, danger, risque intrinsèque*), aux acteurs (concepts : *acteurs, ressources sur site, service, service des acteurs, services de médiation*), aux objectifs (concepts : *risque émergent, effet, mission, évènement déclencheur, facteur de gravité, facteur de complexité*) et à la collaboration (dont les concepts reproduisent ceux de la norme BPMN2.0, formalisme de description de processus collaboratifs).

Dans (Kropczynski et al., 2018), la démarche menée est sensiblement différente : suite à des observations et entretiens avec les gestionnaires de crise aux États-Unis, les auteurs ont cherché à savoir à quelles questions les "preneurs d'appel" (*call-takers*) des services du 911 doivent chercher à répondre pour chaque intervention. Un cadre en 6 questions, les 6 *Ws* en a émergé : What, Where, Who, When, Why and Weapons.

De façon générale, il apparaît rapidement que les méta-modèles dédiés à la crise sont :

- Très génériques : ils sont conçus pour s'adapter à tout type de crise, mais rendent compte de concepts qui peuvent être donc très abstraits pour les gestionnaires de crise ;
- Adaptés à des crises liées à des événements précis déclencheurs ;

- Relatifs à des impacts et effets ciblés de la crise.

Et par conséquent, peu adaptés à une crise de l'ampleur de celle de la pandémie de Covid-19 dont l'on ne cherche plus le *pourquoi* que l'on ne peut pas traiter dans l'immédiat, mais bien les *effets*, multiples, que l'on souhaite mieux connaître et éventuellement éviter ou atténuer.

2.2. La modélisation de la crise du Covid-19

Dans le cadre du projet MESCOV, Marie-Lys Camozzi, dans son rapport de stage (Camozi, 2020), liste en date du 27 Juillet 2020 les différents modèles conceptuels existants, relatifs à la crise du Covid-19, et ce en utilisant la méthodologie suivante: requête ("model" OR "modélisation") AND "crisis" AND ("covid-19" or "coronavirus") sur le moteur de recherche Google Scholar. Cet état de l'art met en exergue six approches en particulier.

Dans (Diemer, 2020), les auteurs s'appliquent à modéliser la crise de Covid-19 en adoptant une vision systémique du problème et en y étudiant la dynamique du/des systèmes via la représentation de boucles causales en particulier adaptées à des "problèmes dynamiques survenant dans des systèmes sociaux, de gestion, économiques ou écologiques complexes". Une première modélisation générale de la crise de Covid-19 en est ainsi donnée établissant d'une part le lien entre crise économique et taux de confinement/déconfinement et l'état sanitaire, en termes de population (contaminée -asymptomatique, symptomatique, hospitalisée, morte, confinée ou non confinée), de rétroaction de la politique de tests massifs de la préparation en termes de moyens et des traitements. Des arbres causaux émergent ainsi entre ces différentes facettes de la crise. Dans un deuxième temps, les auteurs produisent une cartographie des acteurs de la crise et des controverses qui sont susceptibles de ralentir les délais de réponse. Celles-ci sont rassemblées selon sept dimensions : sanitaire, économique, écologique, sociale, culturelle, politique et technique. Finalement, une scénarisation est rendue possible et fait donc la démonstration de l'utilisation de ces modèles pour trouver les trajectoires probables du système étudié.

Dans (Dauvin et al., 2020), les auteurs font un état de l'impact économique de la pandémie de Covid-19 et du confinement en juin 2020. Les différents indicateurs retenus sont : l'impact sur le PIB français, les différents moyens de transport utilisés, la consommation d'électricité, l'impact économique par secteur d'activité par semaine de confinement, l'impact sur la consommation par secteur d'activité, les scénarios prévisionnels de croissance du PIB, l'impact sur l'emploi. Ces indicateurs permettent une étude détaillée en terme économique et ils ont été utilisés pour raffiner notre propre méta-modèle de la crise de Covid-19.

(Zhang et al., 2020) amène une vision encore différente de cette crise, puisque l'étude s'intéresse à la prise en charge psychologique des patients de Covid-19 en hôpital et en particulier à la réduction de l'anxiété en phase d'urgence puis ensuite en accompagnement.

(Chen et al., 2020) explicite une étude menée sur l'engagement citoyen pendant la crise de Covid-19 en se basant sur les données extraites du compte officiel Sina Weibo - un micro-blog - "Healthy China" (compte de la Commission Nationale de la Santé de Chine), entre le 14 janvier et le 5 mars 2020. En se basant sur cinq variables relatives à une publication (valence émotionnelle, popularité, richesse des types de media, niveau d'ouverture aux interactions et type de contenu) et une étude statistique, les auteurs proposent des recommandations pour les

agences gouvernementales pour rendre les interactions avec les citoyens plus performantes et maximiser leur engagement.

Dans (Fernandes, 2020), c'est à nouveau les effets de la crise de Covid-19 sur l'économie qui sont étudiés, prenant des cas spécifiques tels que l'arrêt des vols par compagnie aérienne, la fermeture de site de production de produits de luxe tels que ceux de Gucci et Hermes, la réinstauration des frontières en Europe... De manière générale, cet article permet de définir des points d'intérêts dans la modélisation de la crise en termes économiques.

Finalement, (Cortez & Johnston, 2020) investigate en particulier les pratiques managériales qui peuvent être conseillées dans un cadre instable où les stratégies à long terme ne peuvent plus être définies par les entreprises.

3. Utilisation de l'IA pour extraire de l'information issue d'un flux de tweets

3.1. Les verrous d'une IA pour extraire de l'information actionnable

La première partie de cet état de l'art sur l'utilisation de l'IA pour extraire de l'information des données issues des réseaux sociaux le montre : la classification ne suffit pas à réduire la ressource humaine liée à la lecture et à l'utilisation de l'information. (Coche et al., 2020) débat des verrous liés à l'utilisation de l'IA dans ce cadre de crise :

- Chaque crise est différente (et cette crise sanitaire dont l'ampleur est exceptionnelle renforce cette singularité) et il semble illusoire de mettre en place des modèles de classification purement supervisés qui seraient représentatifs et applicables pour toute crise ;
- Quand bien même ce serait le cas, on observe dans la littérature, d'une part, pléthore de jeux de données de tweets relatifs à des crises, et d'autre part, s'ils proposent une classification leur permettant d'être utilisés en tant que jeu d'entraînement d'algorithmes supervisés, chacun propose une classification (i.e. un méta-modèle) différente. Alors qu'un algorithme supervisé nécessiterait l'entraînement sur un nombre de tweets classifiés relatif au million, cette perspective n'est plus envisageable ;
- La plupart des études s'intéressant à la classification ou la segmentation des tweets et non du sens sémantique que convoient leurs phrases et mots. D'une façon ou d'une autre cela se traduit donc par un travail humain en fin de chaîne de traitement de l'information car le sens contextuel du message n'est pas relevé.

L'implémentation d'un algorithme d'apprentissage de type NER (Named Entity Recognition) s'impose donc comme un moyen d'adresser le dernier point en recherchant et analysant le sens sémantique contextuel des tweets et de leurs mots, tout en restant vigilants à ne pas réintroduire la contrainte d'une démarche uniquement supervisée.

3.2. La reconnaissance d'entités nommées et la propagation de labels

(Ritter et al., 2011) montre une tentative d'appliquer un algorithme NER sur un jeu de tweets, selon 10 classes de termes relatifs à des programmes télévisés. En utilisant un pipeline de traitement de la donnée incluant POS-Tagging, Chunking puis NER dans un cas de supervision distante (supervisé sur la base d'informations reconnues comme vraies, basées sur Freebase).

Dans le cas même d'un traitement supervisé (sur seulement quelques milliers de tweets néanmoins), les résultats du NER ne sont pas convaincants, et les auteurs concluent en soulignant la complexité de mettre en place un NER sur des textes si courts (limités à 140 caractères à cette époque). Twiner (Li et al., 2012) est un système proposant un NER appliqué aux tweets. Les auteurs amènent de très fortes contributions : (i) la capacité du système à segmenter les tokens des tweets afin de déterminer ses propres entités nommées puis (ii) l'enrichissement/vérification par un site tiers (Wikipédia) de ces entités nommées, permettant ainsi une approche entièrement non-supervisée. Néanmoins les résultats obtenus ne permettent pas l'implémentation au niveau de veille sensible attendue dans le cadre de la gestion de crise. Tweedr (Ashktorab et al., 2014) utilise des champs aléatoires conditionnels (Conditional Random Fields) et obtiennent un F1 score de 0.6 pour leur NER appliqué à des tweets. L'approche est elle aussi supervisée.

Supervisés et donc trop chronophages en termes d'entraînement et pas assez adaptables, ou non-supervisés et pas assez performants, de tels algorithmes ne sont pas applicable dans le cadre de notre étude. Des modèles semi-supervisés semblent donc plus opportuns : moins gourmands dans la phase de préparation des jeux de données d'entraînement, ils permettent aussi plus d'adaptabilité. Un pan de la recherche en traitement automatique du langage (TAL) s'intéresse à la propagation de labels, comme par exemple présenté dans (Le et al., 2013) dans le cadre d'un NER adapté à des textes en vietnamien. (Coche et al., 2020) présente cette approche en détail, en appliquant une démarche de propagation de labels adaptée au méta-modèle présenté dans (Benaben et al., 2020) (cf. partie 2.a). L'expérimentation est réalisée sur un jeu de données de 4183 tokens issus de tweets portant sur des situations de crise et un vocabulaire de 154 tokens indépendants, annotés manuellement. L'algorithme BERT (Devlin et al., 2018) est utilisé afin de vectoriser (512 dimensions) l'ensemble des tokens à labéliser, puis un rayon de propagation est utilisé pour "contaminer" les mots voisins des mots de vocabulaire connus. Finalement la Figure 1 illustre les résultats obtenus après réduction de dimensions par T-SNE. Cette technique a été créée, et est toujours en cours d'amélioration, dans le cadre du projet MACIV.

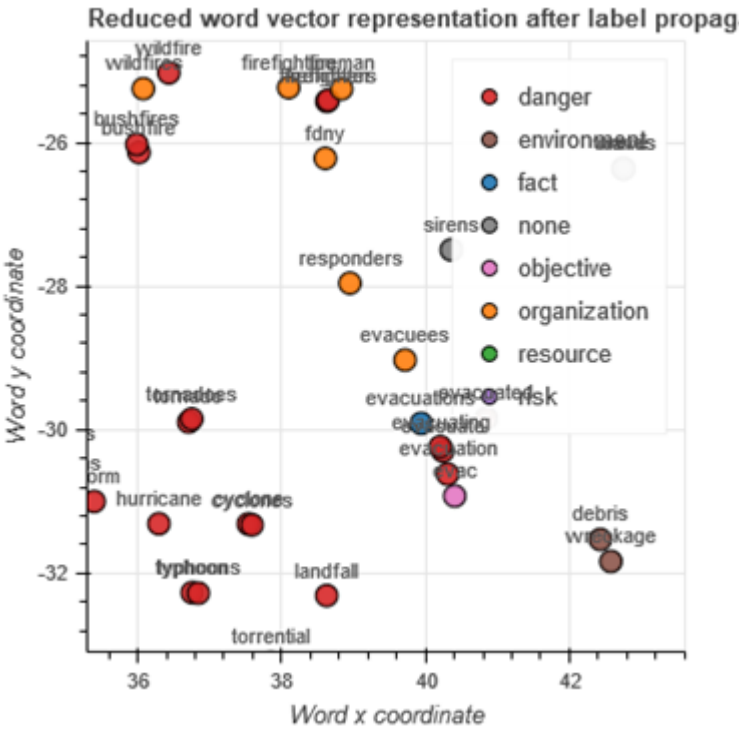


Figure 1 : Échantillon résultant de la propagation de labels après réduction de dimensions par T-SNE, proposé par (Coche et al., 2020).

L'algorithme développé par J. Coche est celui qui a été retenu dans le cadre de l'extraction d'information à partir de tweets, pour le projet MESCOV pour les raisons suivantes :

- L'algorithme est semi-supervisé et amène donc une réponse concrète aux deux premiers verrous soulevés dans cette partie (i.e. adaptable et peu gourmand en jeu d'entraînement) ;
- La vectorisation des mots dans leur contexte (en utilisant BERT dans la chaîne de traitement) permet de considérer les tokens dans leurs contextes d'utilisation (et donc de ne pas échouer dans des contre-sens syntactiques ou sémantiques) et supporte le troisième verrou qui avait été soulevé ;
- La propagation utilisée ne relève pas d'une approche "gloutonne", ainsi les mots non pertinents ne sont pas pris en compte. Cela évite en particulier le risque de faux positifs.

Section 2 : Conceptualisation de l'information à extraire dans le cadre de MESCOV

La section précédente présente un état de l'art des différents méta-modèles et ontologies faisant référence dans le cadre de la gestion de crise, et en particulier communément utilisés par le consortium du projet MESCOV.

Dans le contexte de la crise multi-aspects de Covid-19, nous les avons questionnés, et les avons entièrement remis en question pour répondre à la problématique suivante :

Quelles sont les informations nécessaires à la compréhension la plus omnisciente possible de la situation de crise de Covid-19 et en particulier de la période de confinement qui lui est associée, pour prendre des décisions à court et plus long termes ?

Le méta-modèle proposé ici s'inspire librement du méta-modèle COSIMMA proposé par (Benaben et al., 2020) et de ses catégories principales, et est explicité en détail dans le rapport de M-L Camozzi (Camozzi, 2020), comme repris ci-dessous. Dans le cadre du projet MESCOV, nous avons choisi de catégoriser les concepts d'intérêts selon quatre dimensions :

- Les concepts relatifs au **contexte** font référence à l'environnement de la crise et aux différents éléments qu'elle peut impacter, ou qui peuvent influencer eux-mêmes sur la crise ;
- Les concepts relatifs à la **crise** font eux référence aux faits avérés et potentiels qui ont un impact sur les éléments du contexte ;
- Les concepts relatifs aux **acteurs** considèrent les acteurs et leurs capacités. Ils peuvent utiliser des ressources afin d'agir sur les éléments du contexte et/ou directement sur la crise, l'atténuer ou même l'annuler ;
- Les concepts relatifs aux **ressources** se réfèrent aussi bien aux ressources quotidiennes essentielles que sanitaires et qui ont pu venir à manquer pendant la crise de Covid-19.

1. Le contexte

Dans le paquet **contexte**, on retrouve les concepts suivants :

- Les **entreprises** : impactées par l'aspect économique de la crise, elles fournissent et consomment des ressources et sont implantées dans des zones géographiques ;
- La **population** : c'est l'ensemble de la population qui peut être elle-même redécoupée en sous-parties de la population (vivant dans des zones géographiques précises par exemple, ceci permettant d'avoir une vision de la propagation de la pandémie). Une population peut être spécifiée par un état psychologique résultant de la crise ou encore contrainte selon les réglementations en vigueur. Une population peut avoir accès à l'éducation. De même, une population a besoin de ressources sanitaires (ex. masques) et essentielles. Elle travaille dans des entreprises et peut fournir des ressources (ex. cas d'une initiative citoyenne) et elle habite dans une zone géographique ;
- La **zone géographique** : est définie, tout comme la population, par une auto-agrégation de telle sorte que des zones géographiques peuvent être incluses les unes dans les autres. En particulier, les zones peuvent être utilisées pour étudier les différentes Zones de Défense et Sécurité ou en région, en fonction des rôles décisionnaires utilisateurs de l'outil ;

- Les **transports** : regroupent tout moyen de transport utilisé pour transporter des personnes ou des ressources. Un ensemble de transports constitue une chaîne logistique qui peut être impactée par la crise, ou fournir des ressources essentielles ou sanitaires ;
- Les **contraintes** : elles sont définies par les instances administratives (ex. gouvernements, préfectures) pour une zone géographique donnée. Elles peuvent impacter positivement ou négativement la population, les entreprises ou encore les transports.

2. Les acteurs

Dans le paquet **acteurs**, on retrouve les concepts suivants :

- Les **institutions administratives** : elles peuvent s'auto-agréger (ex. gouvernement constitué de ministères, ministère de l'Intérieur constitué de préfectures...) et disposent de ressources, capacités et compétences propres et édictent les contraintes ;
- Les **professionnels et institutions de santé** : regroupe d'un côté les institutions administratives de la santé, et de l'autre les services d'urgence et de soin. Elles disposent de ressources propres, mais nécessitent des ressources sanitaires.

3. Les ressources

Dans le paquet **ressources**, on retrouve les concepts suivants :

- Les **ressources essentielles** : sont les produits essentiels au quotidien pour la population ;
- Les **ressources sanitaires** : font référence aux ressources nécessaires pour les services d'urgence ainsi qu'à la population pour lutter contre la propagation du virus (ex. masques).

4. La crise

Dans le paquet **crise**, on retrouve les concepts suivants :

- Le **type de crise** : sanitaire, économique, écologique, logistique... ;
- Les **conséquences de la crise** : les faits avérés ou potentiels impactant les éléments du contexte.

La Figure 2 présente le diagramme de classe UML du méta-modèle proposé.

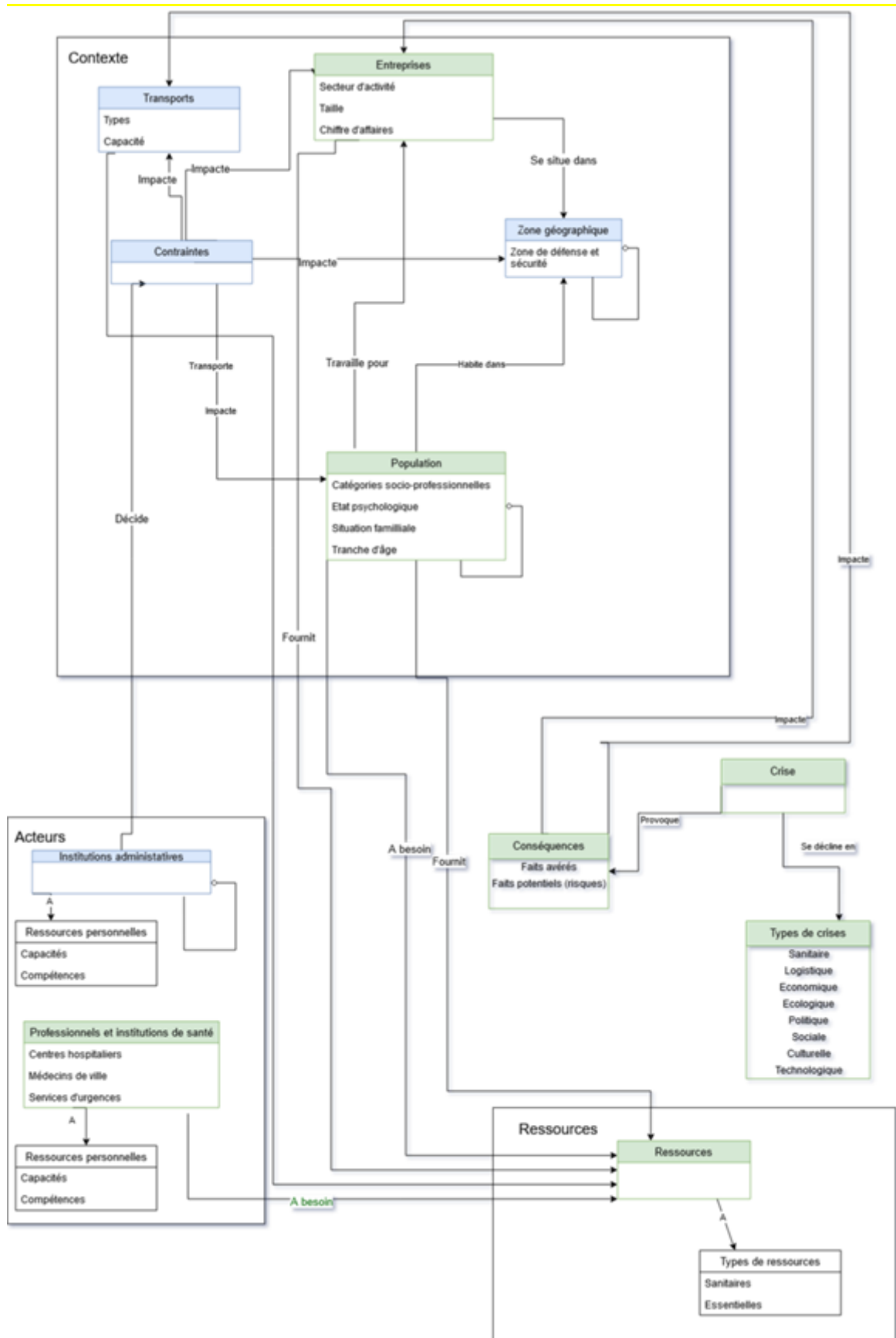


Figure 2 : Diagramme de classe UML du méta-modèle de la crise de Covid-19 (Camozzi, 2020).

Section 3 : Application d'un algorithme semi-supervisé de reconnaissance d'entités nommées relatives à la crise de Covid-19

La section précédente présente le méta-modèle choisi pour l'extraction d'information à partir de tweets, dans le cadre de la crise sanitaire de Covid-19. Les concepts étant ainsi connus, cette section s'intéresse maintenant à l'information que l'on a pu extraire automatiquement. Autrement dit, elle répond à la question suivante :

Comment extraire les informations utiles pour les décisionnaires à partir des tweets récoltés pendant et après le confinement dû à la crise de Covid-19 ?

L'implémentation technologique a eu lieu en plusieurs phases :

- Phase 1 : L'obtention du data-set de tweets relatifs à la crise ;
- Phase 2 : Une première analyse des sentiments sur l'ensemble du data-set ;
- Phase 3 : La reconnaissance d'entités nommées (semi-supervisée, incluant donc l'étape d'entraînement).

L'Obtention d'un data-set relatif à la crise de Covid-19

Une contrainte que nous nous sommes fixée dès l'acquisition des données a été la langue : il était primordial que le data-set ne soit constitué que d'une seule langue (pour permettre l'application des algorithmes d'apprentissage automatique) et si possible provenant de France, puisque le but était de comprendre la crise sur le territoire. La géolocalisation des tweets étant difficile du fait que très peu de tweets sont habituellement géo-tagués (de l'ordre de moins de 5% des tweets de manière générale), nous avons donc décidé que le data-set final serait en **langue française** (qui est l'une des métadonnées que l'API Twitter¹ nous permet de recueillir).

*Il est important de noter dès à présent que le RGPD et la politique de Twitter en matière de protection des données personnelles imposent des contraintes. Notamment, les data-sets ne peuvent être partagés que sous forme de fichier contenant des identifiants de tweets bruts. C'est au chercheur de réaliser une **hydratation**, c'est-à-dire d'utiliser l'API Twitter pour, à partir de l'identifiant de chaque tweet, retrouver l'ensemble des métadonnées qui lui sont associées (relatives au texte contenu, à l'utilisateur auteur ou encore à l'interaction générée autour du tweet), si le tweet existe encore, c'est-à-dire s'il n'a pas été supprimé par l'auteur entre temps.*

En mai 2020, lors du lancement du projet MESCOV, on recensait déjà plusieurs initiatives de partage de sets de tweets qui étaient donc accessibles et disponibles.

Parmi celles-ci, on trouve le data-set GeoCov19 (Qazi et al., 2020) qui est multilingue et basé sur de nombreux mots-clés relatifs aux mots *coronavirus* et *Covid-19* dans de nombreuses langues. Les tweets ont été récupérés entre le 1er Mars et le 1er Mai. Sur le principe c'est donc un data-set candidat tout à fait intéressant, puisqu'il englobe une bonne partie de la période de confinement en France - mais pas dans son ensemble. En revanche, la récupération n'est pas aisée car il faut télécharger l'ensemble des fichiers journaliers de tweets un à un. Plus important, l'étendue des mots-clés recherchés en font un data-set extrêmement volumineux (plus de 520 millions de tweets), et il faudrait l'hydrater dans son ensemble pour retrouver les tweets en langue

¹ <https://developer.twitter.com/en>

française. Compte tenu des limitations de l'API Twitter (900 requêtes/15 minutes), la durée du projet MESCOV ne permettrait pas de récupérer tous les tweets puis les trier et les analyser. Ce data-set n'était donc pas exploitable.

Les tweets ont été récupérés au cours du projet MESCOV entre la date du 22 mars (premier tweet du data-set en date du 24 juin - il a depuis été enrichi) et du 24 juin 2020. Finalement, une fois hydratés, nous avons pu filtrer les tweets dont la métadonnée *lang* (langue) était *fr* (pour les tweets en français). Plusieurs data-sets ont posé ce même problème. C'est finalement le data-set maintenu par le Georgia State University's Panacea Lab qui a été retenu (Banda et al., 2020). Il est réalisé en collaboration par plusieurs chercheurs répartis dans le monde et comporte environ 4,4 millions de tweets enregistrés par jour, contenant au moins un des mots clés suivants : *COVID19*, *CoronavirusPandemic*, *COVID-19*, *2019nCoV*, *CoronaOutbreak*, *coronavirus*, *WuhanVirus*, *covid19*, *coronaviruspandemic*, *covid-19*, *2019ncov*, *coronaoutbreak*, *wuhanvirus*.

Le data-set d'étude que nous avons ainsi produit est composé de 2 950 156 tweets, répartis quotidiennement de façon homogène.

Les deux parties suivantes sont le fruit du travail réalisé par Nathan Thubert lors de son stage (Thubert, 2020), inspiré par les travaux de thèse de Julien Coche, détaillés dans l'état de l'art.

1. Analyse de l'évolution des sentiments dans le data-set de tweets

Dans une première étape et pour prendre connaissance du data-set dans son ensemble, une analyse de sentiments a été réalisée.

Pour cela, une étape de **pré-traitement** a permis de *tokeniser* les tweets (c'est-à-dire les découper mot par mot) et d'enlever tous les tokens non pertinents (url, RT pour retweet, ponctuation, #, @, mots d'arrêt). Cette étape a été réalisée en utilisant la librairie Python de TAL multilingue NLTK. Un stemmer (SnowballStemmer²) est utilisé.

Ensuite, plusieurs **classifieurs** ont été confrontés. L'analyse de sentiment est généralement menée de façon supervisée et le data-set *French Twitter Sentiment Analysis* disponible sur Kaggle (*French Twitter Sentiment Analysis*, s. d.) est l'un des rares qui existent en français et qui peut donc servir de jeu d'entraînement. Ce data-set ne permet d'évaluer un sentiment que selon 2 labels : *positif* ou *négatif*. Après réglages des hyperparamètres via l'utilisation de GridSearch de la librairie Python Scikit-learn, la Figure 3 montre les résultats obtenus pour un classifieur de type (i) Naïve Bayes Gaussian puis (ii) Vecteur de Support - LinearSVC, (iii) un réseau de neurones multi-couche (Multi-layer perceptron - MLP) et enfin (iv) un réseau de neurone de convolution (CNN). Les résultats sont ici donnés pour un entraînement sur 160 000 tweets et un test sur 30 000 tweets, sur un ordinateur doté de 8Go de RAM et un processeur i5.

CLASSIFIEUR	TEMPS D'ENTRAINEMENT	TEMPS D'EXECUTION	PARAMETRES MODIFIABLES	VECTORISATION NECESSAIRE	PRECISION
NAIVE BAYES	66s	36s			73%
LINEARSVC	224s	59s	X	X	77,4%
MLP	337s	57s	X	X	78%
CNN	720s	61s	X	X	69%

² https://www.nltk.org/_modules/nltk/stem/snowball.html

Figure 3 : Confrontation de plusieurs classificateurs de sentiments, entraînés et testés sur le data-set French Twitter Sentiment Analysis, résultats issus de (Thubert, 2020).

Étant donné le nombre conséquent de tweets à classifier dans le data-set dédié au, le classifieur LinearSVC a été sélectionné comme le meilleur compromis précision /temps.

Ce classifieur de sentiments en français a donc été appliqué à l'ensemble du data-set d'étude et le sentiment moyen a été calculé pour chaque jour. La Figure 4 en illustre le résultat. 1 représente un sentiment positif, tandis que 0 un sentiment négatif. On voit que les tweets sont relativement équilibrés entre positifs et négatifs avec quelques fluctuations. Cependant, la précision de l'algorithme LinearSVC ne permet pas d'être certains que les petites variations d'un jour à l'autre se fassent ainsi. Nous n'avons d'ailleurs pas trouvé de lien entre les pics/creux de positivité et l'actualité pendant les jours concernés.

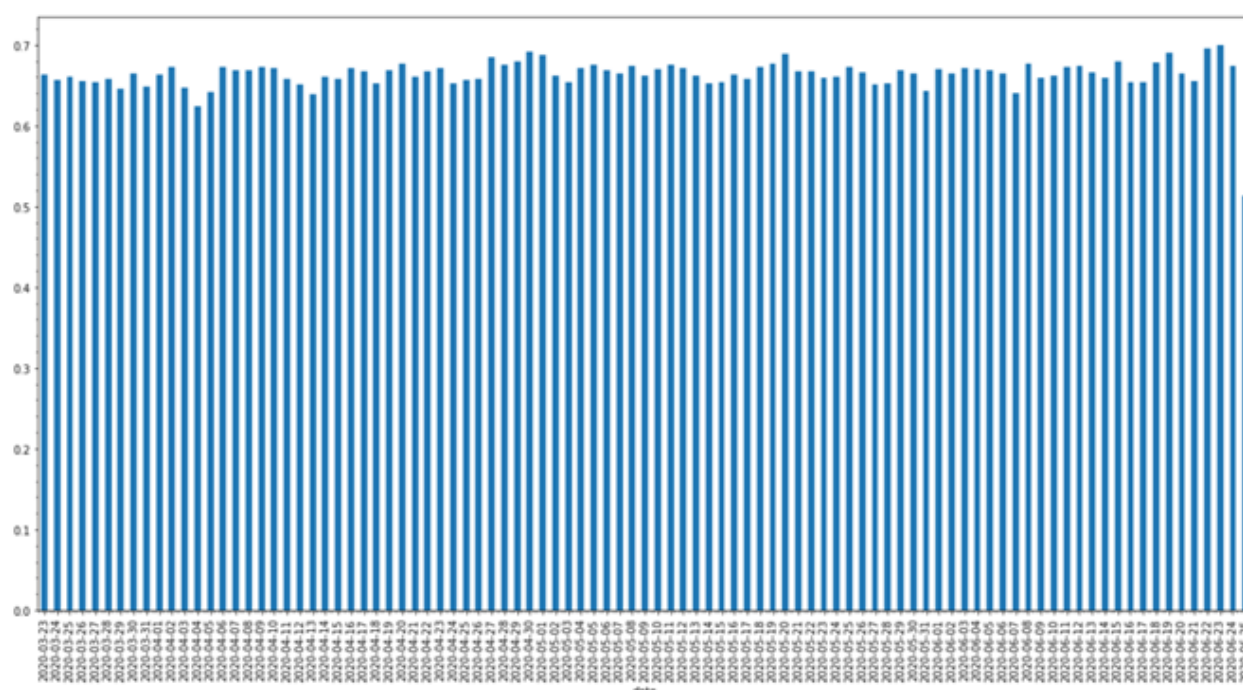


Figure 4 : Sentiment moyen quotidien calculé par LinearSVC sur le data-set d'étude dédié à la crise de Covid-19.

2. Reconnaissance d'entités nommées, basées sur le méta-modèle de la crise de Covid-19

Dans l'état de l'art, l'approche semi-supervisée proposée par Julien Coche (Coche et al., 2020) est détaillée. Nous allons ici préciser son application spécifique et nouvelle, d'une part à des tweets en langue française et d'autre part basée sur le méta-modèle de la crise de Covid-19 proposé en Section 2.

La première étape explicite la création du data-set d'entraînement puis la deuxième montre les résultats obtenus.

2.1. Création du jeu d'entraînement

Afin de générer un data-set d'entraînement, 2400 tweets ont été extraits de manière aléatoire, répartis dans le temps de manière homogène. Un serveur Label Studio³ a été déployé. Il s'agit d'une application web permettant de labéliser de manière efficace tout type de données, dans le texte. La capture d'écran en Figure 5 montre l'outil. 7321 tokens ont pu être labélisés manuellement (i.e. associés à une classe du méta-modèle proposé) et se répartissent de la façon suivante comme présenté en Figure 6.

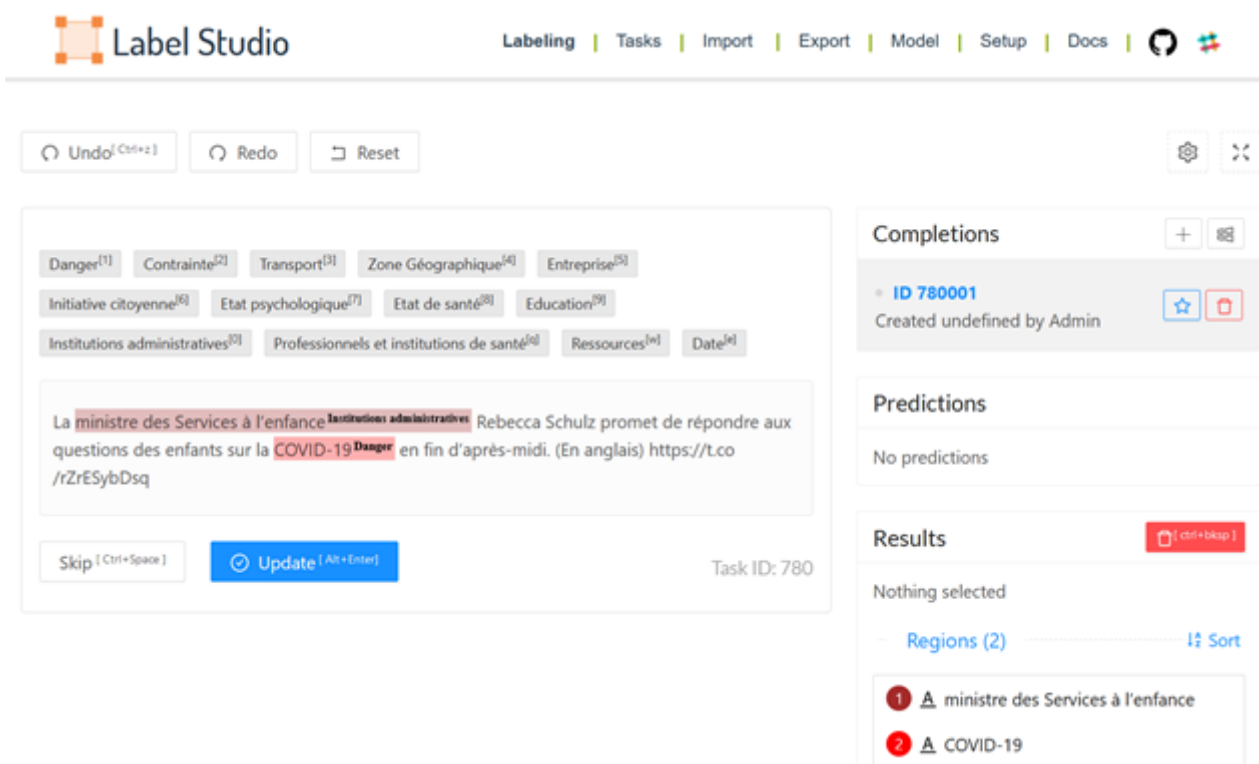


Figure 5 : capture d'écran du serveur Label Studio, pour générer le set d'entraînement de manière collaborative.

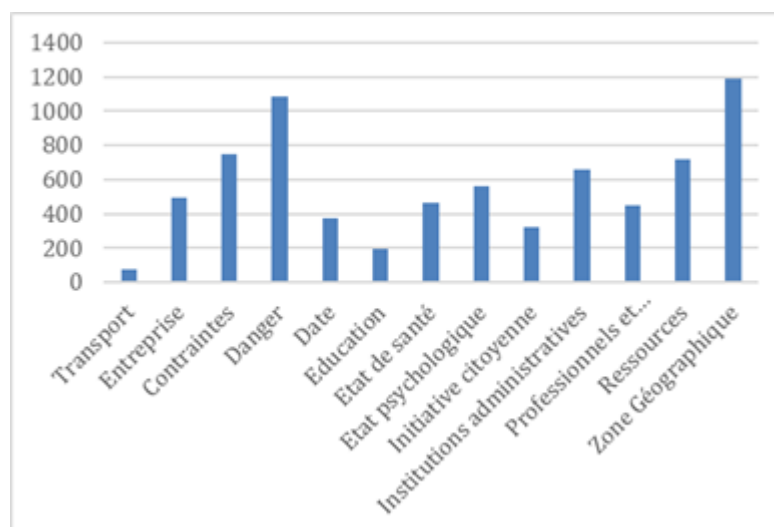


Figure 6 : Répartition des tokens labélisés dans le data-set d'entraînement.

³ <https://labelstud.io/>

2.2. Propagation des labels dans le data-set d'étude

Afin de propager ces labels, il nous faut, dans l'ordre, vectoriser les labels du jeu d'entraînement, prétraiter puis vectoriser les tokens du data-set d'étude (rappel : 2 956 156 tweets tokenisés) puis finalement propager les labels des tokens du data-set d'entraînement vers ceux du data-set d'étude à l'aide d'un rayon fixé à 0,55, ainsi que préconisé par (Coche et al., 2020). La Figure 7 récapitule ces étapes.

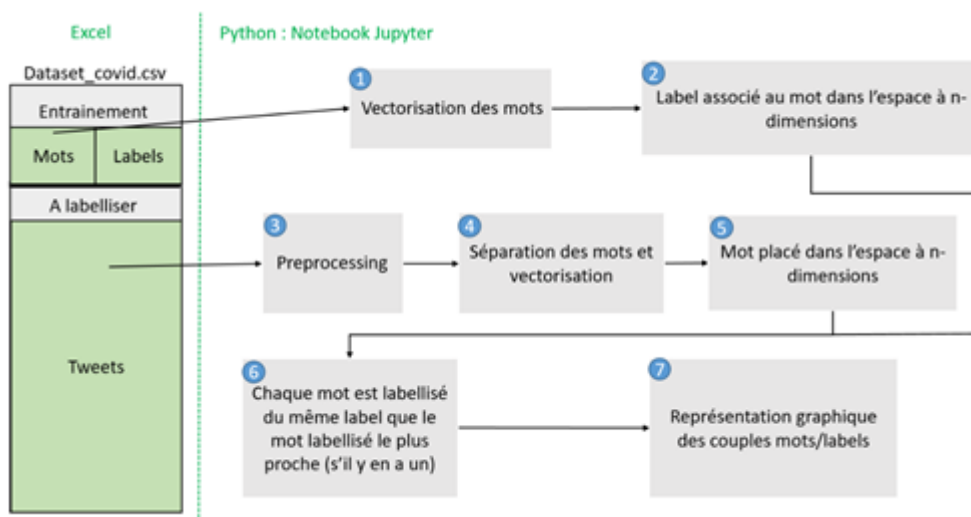


Figure 7 : Étapes de la propagation de labels (Thubert, 2020).

Sur la semaine du 23 au 30 mars, on obtient environ 23 000 entités nommées, dont on a un exemple ci-dessous, en Figure 8.

```
accompagnées education
mondialement ressources
lantenne contrainte
lagos zone géographique
parvis danger
américain zone géographique
prêter initiative citoyenne
castorama ressources
vieilli etat de santé
coupent etat psychologique
moyenorient education
denrée professionnels et institutions de santé
futé education
tournés ressources
trompé etat psychologique
```

Figure 8 : Exemples de labels propagés par l'algorithme. À gauche un token du data-set d'étude et à droite la classe qui lui a été associée.

Il est clair que la précision nécessite d'être améliorée. En particulier, le déséquilibre entre le nombre de tokens représentatifs de chaque classe dans le data-set d'entraînement est prégnant (cf. Figure 6). L'étude menée montre qu'en rééquilibrant le data-set d'entraînement et en ne sélectionnant qu'un nombre fixe de représentants de chaque classe, les résultats semblent meilleurs. Nous allons les discuter dans la sous-partie suivante. Afin de rendre compte d'un outil qui pourrait in fine permettre de trouver les informations liées à la situation de crise de manière aisée, une visualisation en 2D des vecteurs de chaque token (à l'origine, ils sont donc représentés à un point) est produite après une réduction de dimension par T-SNE. Celle-ci permet directement à l'utilisateur de choisir la classe qu'il/elle souhaite visualiser (cf. Figure 9).

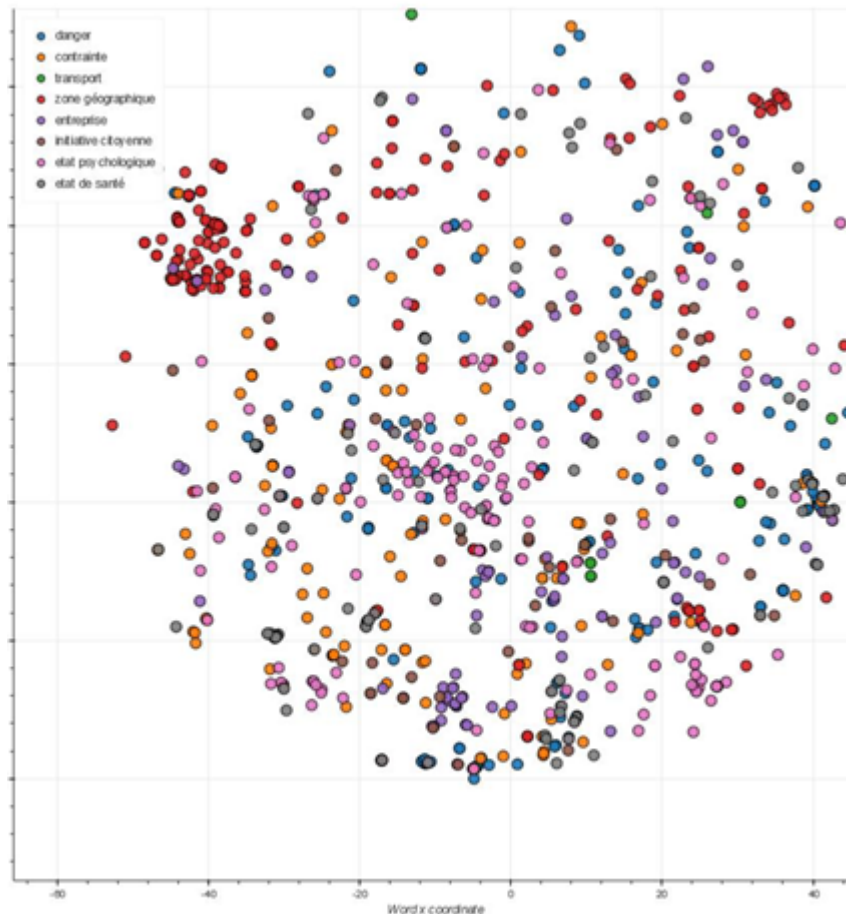


Figure 9 : Visualisation des tokens après propagation de labels dans le data-set d'étude.

3. Contributions, limites et perspectives de l'IA au service de l'extraction automatisée d'information

Cette étude technologique amène ses propres contributions :

- Un data-set a été créé et partagé. Ce data-set pourra être facilement utilisé dans le cadre de recherches liées à la crise de Covid-19 en France. Il pourra aussi être librement complété des périodes antérieures et postérieures à son acquisition ;
- Un méta-modèle adapté à cette crise sans précédent est proposé et tente d'une part de catégoriser, répertorier et lier les différents types d'information pertinents dans ce contexte particulier, et, d'autre part, de poser les bases d'une approche supportée par l'IA telle que le NER ou la propagation de labels ;
- Tester et éprouver le modèle de propagation de label pensé par Julien Coche, dans le cadre du projet MACIV, sur un contexte d'application différent et en langue française.

Il semble que les résultats finaux mitigés de la propagation de labels remettent en question la conceptualisation de l'information telle qu'elle est proposée dans le méta-modèle. Notamment, certaines classes semblent être plus "terre-à-terre" (ex. zones géographiques) tandis que d'autres semblent déjà relever de l'interprétation humaine (ex. danger), voire de la décision anticipée. Dans un second temps, les travaux menés actuellement par l'équipe MESCOV s'orientent vers une vision commune et visent un alignement des deux champs disciplinaires que

sont la science des données et les sciences humaines. Le méta-modèle a ainsi déjà été rediscuté et précisé. Notamment les classes ont été repensées et définies : certaines sont suffisamment "simples" pour être amenées de façon fiable par l'IA tandis que d'autres relèvent d'une interprétation humaine mieux cernée et guidée/supportée.

II. Création et circulation de l'information sur les media sociaux et initiatives citoyennes lors de la crise Covid-19

Dans son volet Sciences Humaines et Sociales, le projet MACIV a permis d'une part de documenter et d'étudier la création et la circulation de l'information sur différents types de media sociaux (notamment Wikipédia et Twitter) et les mécanismes de vérification par les citoyens (Bubendorff et al., à paraître, 2019) et d'autre part de mettre en évidence et de comprendre l'émergence et l'organisation d'initiatives citoyennes qui viennent complexifier et enrichir (si elles sont intégrées aux processus de gestion de crise) la réponse institutionnelle à la crise (Batard et al., 2018; Batard, Montarnal, et al., 2019; Batard, Rizza, et al., 2019; Rizza, 2020).

Le projet MESCOV a ainsi permis d'approfondir la connaissance sur les mécanismes de création, circulation et vérification de l'information lors de la crise de Covid-19 et d'étudier les « initiatives citoyennes » qui ont émergé lors du confinement et leur prise en considération par les institutions gestionnaires de crise. Plus spécifiquement, en mobilisant la base de tweets co-construite pour le projet il a également été possible de qualifier et décrire les informations qui ont circulé sur Twitter pendant le confinement et la période suivant sa levée (22 mars au 24 juin). Nous nous sommes également intéressés à la communauté experte Wikipédia et à certains des articles relatifs à la crise sanitaire ainsi qu'aux communautés citoyennes qualifiées de « makers » qui ont, lors du pic de crise, fourni notamment des masques aux institutions de santé et SDIS.

Cette partie est structurée comme il suit : Section 1 « Faire collectivement sens au temps de la crise de Covid-19 » ; Section 2 « Analyse descriptive de la base de données de Tweets en langue française » ; Section 3 « Les usages de Wikipédia lors de la pandémie » ; Section 4 « Un exemple d'organisation citoyenne sur les media sociaux en temps de crise : les makers ».

NB : Les éléments présentés dans cette partie font d'ores et déjà l'objet de publications soumises à la revue Communiquer : revue de communication sociale et publique⁴ et à la revue Journal of Contingencies and Crisis Management⁵.

Section 1 : Faire collectivement sens au temps de la crise de Covid-19

La pandémie de Covid-19 constitue une crise sanitaire majeure. En tant que crise, elle présente notamment un caractère inattendu et génère de l'incertitude. Elle met en tension ou menace des acquis sociaux, économiques et structurels qui garantissent le bon fonctionnement de la société. En d'autres termes, la pandémie constitue une situation d'exception qui, au-delà des aspects

⁴ Appel à communication sur la thématique « Crises sanitaires et communication : enjeux sociétaux et organisationnels » de la revue Communiquer : Revue de communication sociale et publique <https://journals.openedition.org/communiquer/6787>

⁵ Bubendorff, Rizza and Prieur, Construction and dissemination of information veracity on French social media during crises: comparison of Twitter and Wikipedia, à paraître.

sanitaires, mérite d'être étudiée et interrogée par des experts et chercheurs des sciences humaines et sociales, notamment en communication ou en gestion de crise. Cette étude s'inscrit ainsi dans le champ des « *crisis informatics* », champ pluridisciplinaire qui est apparu et s'est progressivement construit de manière simultanée à la mobilisation des technologies de l'information et de la communication en gestion de crise par les acteurs de la sphère publique (par exemple, États et citoyens) et privée. Plus précisément, les « *crisis informatics* » étudient comment les technologies numériques en réseau, en particulier les media sociaux à partir des années 2000, interagissent avec la gestion de crise, à partir d'une sensibilité relevant à la fois des sciences sociales et des sciences informatiques notamment par les sciences de la donnée (Palen et al., 2020). La littérature dans le domaine des « *crisis informatics* » a mis en évidence les changements amenés par les media sociaux et comment les citoyens les ont progressivement utilisés pour communiquer sur un évènement, s'informer, et s'organiser pour aider lors d'un évènement majeur ou d'une crise s'inscrivant plus longuement dans la durée. Les media sociaux deviennent alors des lieux où prendre des nouvelles de ses proches, trouver ou échanger des informations, des recommandations, des opinions, organiser une réponse à un ou plusieurs des aspects de la crise.

La pandémie de Covid-19 présente des spécificités en comparaison aux crises qualifiées de sécurité civile (telles que les crues, les feux de forêt, les attentats) dont la plus significative pourrait être sa cinétique lente qui l'inscrit dans le moyen-long terme. Elle suscite néanmoins depuis qu'elle a débuté une mobilisation forte des media sociaux. Ainsi, dans cet article, nous appréhendons les media sociaux comme des plateformes où du sens est collectivement créé en réponse à la pandémie (Starbird, 2020). Plus précisément, nous nous intéressons aux plateformes Wikipédia et Twitter ; nous analysons comment l'information y est traitée, négociée, organisée durant cette période.

Pour Wikipédia, en nous appuyant sur une méthodologie mixte d'entretiens semi-directifs et d'analyse ethnographique des pages de l'encyclopédie en lien avec la pandémie de Covid-19, nous montrons de quelles manières les savoir-faire en matière de recherche et de vérification de l'information des contributeurs à l'encyclopédie sont mobilisés en temps de crise pour proposer un sens à l'évènement en train de se dérouler.

Pour Twitter, en explorant la base de données de tweets présentée ci-dessus (cf. Partie I, Section 3 « Application d'un algorithme semi-supervisé de reconnaissance d'entités nommées relatives à la crise de Covid-19 »), nous analysons comment la parole se distribue sur ce medium lors de la pandémie et cherchons à montrer les mécanismes qui structurent les débats, les controverses lors de cette crise.

Le cas du traitement de l'information pendant la pandémie permet de mettre en avant les manières dont les initiatives citoyennes sont vectrices de sens lors des crises et peuvent être organisées de manière délibérative, sans intervention des organisations traditionnelles ou historiques. L'analyse de ce matériel permet non seulement de mettre en lumière la manière dont l'information émerge et circule sur les media sociaux, mais surtout d'observer comment ces informations sont débattues, sélectionnées, structurées, mises en forme (Keegan et al., 2013). Observer ce processus de légitimation de l'information révèle celui sous-jacent du « faire sens » (« sensemaking ») lors d'une crise. Cet aspect se décline sur les deux media sociaux auquel nous avons prêté attention dans le cadre de cette recherche. En effet, Twitter a été, comme lors d'autres crises (Bubendorff et al., 2019) particulièrement plébiscité lors de la pandémie afin de

débattre, d'échanger des informations et parfois même des discussions entre scientifiques. En outre, la spécificité de Wikipédia par rapport aux autres media sociaux réside dans la recherche d'un consensus entre ses différents membres, contributeurs. Contrairement à Twitter où les débats peuvent être infinis et les controverses alimentées sur le long terme, dans le cas de Wikipédia, l'objectif poursuivi par sa communauté de contributeurs est de proposer un état des savoirs à l'instant -t, même si cela suppose de faire coexister différentes « narrations » (Starbird et al., 2019) autour d'un même évènement. Dans le cas de la pandémie de Covid-19, cette manière de faire est particulièrement intéressante à prendre en compte dans le cas d'une crise émaillée de controverses et de débats auxquels n'échappent pas les échanges sur l'encyclopédie, ni les discussions sur Twitter.

1. Faire sens à la crise au moyen des media sociaux

Les media sociaux sont des plateformes ou applications Web 2.0 qui permettent à leurs utilisateurs de créer du contenu en ligne, de l'échanger, de le consommer et d'interagir avec d'autres utilisateurs ou leur environnement en temps réel (Kaplan & Haenlein, 2010; Luna & Pennock, 2018; Reuter et al., 2020). Ces dernières années, l'utilisation des media sociaux a considérablement augmenté ; ainsi en 2019 Facebook, Twitter et Whatsapp comptaient 5,7 milliards d'utilisateurs dans le monde (Étude Statistica 2019 citée par Reuter et al., 2020, p. 2).

D'une manière générale, les media sociaux permettent à leurs utilisateurs de communiquer et d'interagir de manières différentes et souvent combinées : la création et diffusion d'informations, la gestion des relations, la communication et l'expression de soi. Sur la base de ces activités, on peut ainsi distinguer (Reuter et al., 2012) :

- les wikis : pour la collecte d'informations et la création de connaissances selon une logique collaborative ;
- les blogs et microblogs : pour la publication d'informations et/ou l'expression de soi ;
- les sites de réseaux sociaux : pour la gestion des relations, l'expression de soi et la communication, ainsi que la collecte d'informations ; et
- les systèmes de partage et d'indexation de contenu : pour la création et l'échange d'informations multimédia (photos, vidéos).

Cette catégorisation permet ainsi de différencier les media sociaux selon leurs fonctions principales : Twitter en tant que microblog permet de diffuser ou de collecter de l'information ; Facebook est un réseau social qui permet l'interaction entre « amis » ou au sein d'une communauté « groupe » ; Wikipédia est une plateforme de création collaborative de connaissance (Kaplan & Haenlein, 2010). Cependant, un contexte de crise, d'urgence ou de controverse donne lieu à des utilisations spécifiques des media sociaux où souvent on observe une combinaison des fonctions principales d'une plateforme avec d'autres fonctions nécessaires dans ce contexte : par exemple, les pages de discussion de Wikipédia sont le lieu d'échanges au sein de la communauté des contributeurs au même titre qu'un réseau social (Bubendorff & Rizza, 2020). Il existe par ailleurs des plateformes spécialisées en gestion de crise : à titre d'exemple, Crisis mapper (déclinaison d'Open Street Map) permet de cartographier des zones touchées par un évènement majeur (Wendling et al., 2013).

Le champ des « *crisis informatics* » (Palen et al., 2020) s'est largement penché sur l'utilisation des media sociaux dans leurs fonctions de « partage d'information » (ex. Twitter) ou de réseau social (ex. Facebook) en temps de crise (voir Reuter & Kaufhold, 2018). Ces études démontrent le rôle

de ces outils dans les processus de résilience collective et/ou individuelle (Dufty, 2012) et notamment dans l'élaboration de sens autour d'évènements majeurs (Starbird, 2020). La notion de « faire sens » collectivement a beaucoup été mobilisée par les sciences de gestion (Maurel, 2020; Weick et al., 2005) à partir d'une perspective organisationnelle. Pour (Boersma et al., 2020), il s'agit d'un processus de construction sociale qui vise de manière rétrospective à rationaliser l'action des individus en période de grande incertitude – c'est-à-dire marquée par de l'information, des compétences et des ressources incomplètes ainsi que des mesures publiques dont on ne connaît pas les effets. Les « crisis informatics » s'emparent également du concept de « sense making » et étudient les processus en ligne par lesquels les individus s'organisent lors d'une crise pour répondre aux incertitudes qu'elle génère. Comme le rappellent (Starbird et al., 2020), les épidémies – plus encore que d'autres types de crises – révèlent l'importance du « faire sens » collectif lors de périodes d'incertitude. Parce que ces crises viennent redéfinir – à plus ou moins long terme – les comportements sociaux, les manières d'être-ensemble, de travailler, elles sont source d'incertitude notamment pour les instances habituellement chargées d'y répondre. Pour les citoyens, l'une des manières de répondre à ces incertitudes consiste à agréger des informations (Heverin & Zach, 2012; Jurgens & Helsloot, 2018). Les media sociaux jouent un rôle central dans cette démarche, aussi bien pour les personnes directement impactées par les évènements que par une plus large communauté se sentant concernée. Eismann, Passega et Fischbach montrent que « partager et obtenir des informations factuelles sont les principales fonctions des media sociaux, et ce quel que soit le type de crise » (Eismann et al., 2016, p. 7). Ce faisant, chercher et partager de l'information sur les media sociaux lors d'une crise « permet de construire une image globale de ce qui est en train de se passer » (Heverin & Zach, 2012, p. 44). Cette utilisation des media sociaux permet de « faire face » aux évènements (C. Yates & Partridge, 2015) notamment en améliorant sa « connaissance de la situation » comprise comme le processus qui consiste idéalement à rassembler l'ensemble des connaissances accessibles sur une situation incertaine pour en tirer une image cohérente et l'évaluer (Vieweg et al., 2010). Starbird (2020) a par ailleurs d'ores et déjà souligné le caractère anxiogène de la pandémie et de la désinformation associée. L'incertitude liée à cette crise s'accompagne d'un phénomène « d'infodémie » telle que décrite par l'OMS⁶ où les media sociaux sont le support de débats, d'échanges et de diffusion d'informations, mais également de rumeurs ou de fausses informations. Les rumeurs doivent être comprises différemment de la « désinformation » : si les premières sont le plus souvent inhérentes à la crise et permettent aux individus de gérer l'incertitude qui lui est liée en l'absence d'informations fiables ou de communication institutionnelle, la seconde répond d'une organisation et d'une volonté de répandre un type de narration s'élevant contre le discours majoritairement relayé (voir par exemple Starbird et al., 2019). Les auteurs mettent le plus souvent ce type de contenu en lien avec les théories conspirationnistes (par exemple Tandoc et al., 2018), les controverses devenant structurantes dans une période d'incertitude exacerbée.

Dans ce contexte, l'analyse de Twitter permet de comprendre autour de quelles thématiques se structurent les discussions à propos de la pandémie, c'est-à-dire dans quels domaines cette crise vient impacter le quotidien des internautes francophones. De plus, l'observation et l'analyse des articles Wikipédia liés à la pandémie permettent de comprendre comment la plateforme constitue à la fois un lieu d'accès et d'élaboration de la connaissance par les processus de

⁶ Source : <https://www.who.int/fr/news-room/events/detail/2020/06/30/default-calendar/1st-who-infodemiology-conference>, consulté le 10.09.20

synthèse de l'information et d'interactions entre contributeurs, notamment en prêtant attention aux espaces discursifs de cette plateforme (Bubendorff et al., à paraître).

2. La mobilisation des media sociaux lors de crises

2.1. Les propriétés des espaces discursifs de Wikipédia en temps de crise

Les études au sujet de Wikipédia ont montré à de nombreuses reprises son rôle significatif dans l'accès à l'information et à la connaissance sur internet (Singer et al., 2017). Comme le souligne (Waller, 2011), le bon référencement de l'encyclopédie sur les moteurs de recherche contribue à sa popularité par un double mouvement. D'une part, ce référencement rend Wikipédia très accessible aux internautes ; de l'autre, il témoigne de la confiance et de la fiabilité conférée par les internautes à Wikipédia.

Les articles Wikipédia les plus consultés traitent de l'actualité ; néanmoins les recherches sur l'encyclopédie ont écarté les études spécifiques relatives aux crises à cause de l'opposition entre la cinétique rapide de la crise et le temps long dans lequel on inscrit traditionnellement le travail d'écriture d'un article encyclopédique (Keegan et al., 2013). Ainsi, les visites de pages liées aux événements d'actualité comme une crise atteignent des pics de consultations au moment même où ils se déroulent ou lors des commémorations relatives à l'évènement (Bubendorff et al., 2019). Ces périodes de consultations intenses constituent également de véritables appels d'air pour l'encyclopédie en termes d'arrivée de nouveaux contributeurs (Bubendorff & Rizza, 2020; Keegan, 2015). Ces pics de consultations et de collaborations en lien avec l'actualité intensifient en effet les échanges, les interactions, les débats et les controverses au sein des différentes pages de discussions constitutives de l'encyclopédie.

Les études sur l'utilisation de Wikipédia lors de crises se sont longtemps limitées à sa dimension collaborative d'élaboration de la connaissance, et certaines enquêtes se sont spécifiquement centrées sur l'utilisation d'autres applications basées sur la même technologie : le wiki (Palen & Liu, 2007). Cependant, en s'attachant à l'analyse de ses espaces discursifs, l'encyclopédie peut être comprise comme un réseau social numérique (Bubendorff & Rizza, 2020), c'est-à-dire une plateforme qui permet également et notamment la collecte d'informations, la communication entre contributeurs et la gestion de leurs interactions. L'encyclopédie présente ainsi les mêmes propriétés que les dispositifs numériques habituellement analysés pour leur contribution à la résilience des populations en favorisant « la diffusion et l'accès rapide » aux informations (Chan, 2014) et la sélection des informations dans un contexte de profusion d'informations contradictoires et de rumeurs (Starbird et al., 2020). L'observation des espaces de débats et d'échanges que propose l'encyclopédie en ligne permet d'accéder à la manière dont un sens est construit ou donné à un évènement. Ces pages sont associées à des articles ou à des profils de contributeurs et hébergent les interactions autour de l'actualité relatée. Elles donnent la possibilité aux chercheurs d'accéder aux processus de discussions délibératives au sens d'Habermas, (1987, 1995) (Bubendorff & Rizza, 2020).

2.2. L'instantanéité de Twitter : dispersion de l'information et structuration des échanges lors d'une crise

Twitter est également très réactif aux événements d'actualité et ses usages particulièrement actifs lors de crises. À ce titre, ce medium est souvent envisagé dans la littérature comme un nouvel outil d'échange d'information. Kwak et al. étudient par exemple les « caractéristiques

topologiques de Twitter et sa puissance en tant que nouveau moyen de partage de l'information » (Kwak et al., 2010). Dès 2010, Heverin & Zach (2012) et Vieweg et al. (2010) proposent une catégorisation des tweets contenant de l'information lors des crises :

- Tweets originaux ;
- Tweets de synthèse ;
- Tweets propagateurs (qui constituent près de 90% des échanges dans les tweets).

Outre l'analyse des messages postés, les auteurs se sont intéressés aux profils des internautes actifs sur le réseau lors des crises et plus spécifiquement à leur typologie. Les dynamiques d'abonnement aux comptes sont décrites comme caractéristique d'un usage informationnel. Ainsi comme le soulignent Kwak et al. : « 67.6% des utilisateurs ne sont suivis par aucun des comptes auxquels ils sont abonnés sur Twitter. Nous affirmons que pour ces utilisateurs, Twitter est une source d'information plutôt qu'un réseau social (Kwak et al., 2010, p. 593) ». Si les internautes massivement suivis ne suivent pas à leur tour l'ensemble de leurs abonnés, c'est bien qu'il constitue pour ces derniers un point d'accès à l'information. Les usages sur Twitter lors des crises prendraient alors une distribution verticale au sein des échanges informationnels. On retrouverait alors sur Twitter une distribution verticale des rôles au sein de ces échanges. En ce sens, il est lieu incontournable de l'espace public en temps de crise.

3. La question de l'autorité et de la légitimité sur les media sociaux

Malgré – ou en partie du fait de – leur forte mobilisation durant les événements d'actualité, les media sociaux sont souvent considérés comme des espaces publics où des rumeurs ou des fausses informations peuvent être répandues. La pandémie en est un exemple flagrant. Dans la littérature, de nombreuses recherches se sont penchées sur ces phénomènes, cherchant à les démontrer ou à les infirmer, notamment en regard de l'usage de Twitter. Alexander (2014) s'est intéressé à des cas où de fausses informations ont été partagées sur les media sociaux après des catastrophes ; Starbird étudie « les narrations alternatives autour d'évènements de crises induites par l'homme » (Starbird, 2017, p. 230) ; Mendoza et al démontrent que Twitter est un outil favorisant la transmission « de rumeurs sans fondements » (Mendoza et al., 2010). La littérature a également montré que les citoyens partageaient largement cette perception des media sociaux. (Spielhofer et al., 2019) ont par exemple mené une enquête sur la population britannique dont les résultats montrent qu'une majorité d'entre eux ne considèrent pas les media sociaux comme source d'informations fiables (66% de leur échantillon), notamment du fait de la mise en avant d'informations peu crédibles (60%) ou de rumeurs (64%). Malgré ces perceptions, la littérature et les études cas montrent comment des processus ou mécanismes structurels sont mis en œuvre par les utilisateurs des media sociaux afin de garantir la véracité de l'information diffusée.

3.1. Le référencement : processus de vérification de l'information

Dans le cas de Wikipédia, les mécanismes de vérification de l'information tels que le référencement sont identifiés et largement analysés par la littérature (Auray et al., 2009; Jullien, 2012; Sahut, 2016; D. Yates et al., 2010). Pour les contributeurs, l'objet même de Wikipédia est de fournir aux lecteurs une synthèse des informations et de la connaissance traitant d'un sujet, en relation ou non avec un événement d'actualité ou une crise (Bubendorff et al., 2019). Il s'agit pour eux de résumer et d'agréger les savoirs afin de les mettre à disposition du plus grand nombre. Pour garantir la fiabilité de la synthèse proposée par les contributeurs, ces derniers sont appelés à sourcer les informations qu'ils ajoutent sur les articles. Ce travail de « *sourcing* » est essentiel

sur Wikipédia et constitue « l'une des clés de voûte du modèle éditorial wikipédien » (Sahut, 2014, p. 109). Il assure le sérieux et la fiabilité du modèle collaboratif de l'encyclopédie en ligne et est l'objet de multiples recommandations sur les pages d'aides aux contributeurs. Des règles simples de référencement leurs sont proposées et les usages sur Wikipédia tendent à importer des critères d'autorité établis en dehors de l'encyclopédie (publications académiques, presse, etc. voir notamment Ford et al., 2013; Sahut, 2016). Ce mécanisme de recours à des sources traditionnelles sur Internet est déjà identifié en dehors des périodes de crise (Rebillard, 2006), mais la prédominance de citations renvoyant à des articles de presse sur Wikipédia a été mise en avant dans le cadre spécifique d'études relatives aux crises qualifiées de sécurité civile (Bubendorff et al., 2019). Le rythme de couverture et d'analyse au plus près de l'évènement proposé par la presse correspond alors à celui engagé par les contributeurs.

Lors de crises ou plus largement de controverses, ces principes à la base du fonctionnement de l'encyclopédie sont mis en tension. Les outils de régulation (Cardon & Levrel, 2009) permettant aux wikipédiens de garantir les processus de synthèse des informations habituellement usités par les wikipédiens sont particulièrement mobilisés lors d'une crise (Bubendorff et al., 2019) : des avertissements sous forme de bandeaux apposés en en-tête de la page permettent de prévenir le lecteur des limites de l'article (instabilité due à un évènement en cours ou à une controverse, manque de sources, etc.) ; les contributions des articles peuvent être temporairement limitées à certains utilisateurs ; des rappels à l'ordre et des menaces de sanction sont utilisés à l'égard de ceux ne respectant pas les règles éditoriales (vandalisme) (Cardon & Levrel, 2009).

Ces garde-fous proposés par la plateforme sont complétés par la spécialisation de certains contributeurs à la rédaction d'articles de crise : ils sont nombreux à intervenir dans plusieurs articles concernant la même thématique de crise (attentats, sanitaire, météorologique, voir Bubendorff et al., 2019). La présence de ces contributeurs amène à une correction plus rapide des erreurs (volontaires ou non) que lors de la création d'articles ordinaires (Keegan, 2015) ainsi qu'à un effort de retenue dans le rythme des publications (Bubendorff et al., à paraître). Cette gestion du rythme de contributions est couramment évoquée par les contributeurs les plus chevronnés, habitués aux débats et aux évènements exceptionnels venant perturber la routine wikipédienne. Dans les pages de discussions, ces contributeurs rappellent les règles éditoriales de l'encyclopédie, suppriment les informations non sourcées, et initient des discussions sur des caractéristiques propres aux articles de Wikipédia (respect de l'orthographe, longueur et découpage de l'article, caractère encyclopédique des informations présentes.).

Ainsi, les espaces discursifs de l'encyclopédie sont d'autant plus intéressants à prendre en compte qu'ils regroupent les négociations à propos de l'utilisation de sources. Ils hébergent également les réflexions menées par les contributeurs en période de crise et permettent d'accéder à la manière dont les éléments qui se déroulent sont appréhendés à l'aune des règles collaboratives mis en œuvre sur l'encyclopédie.

3.2. Les controverses et l'autorité dans les usages de Twitter

Les mêmes types de mécanismes sont à l'œuvre sur Twitter comme le montrent des recherches récentes sur ce medium.

Des mécanismes de vérification de l'information partagée sont également décrits sur ce medium social. Par exemple, Alloing et Vanderbiest (2018) montrent que les démentis de rumeurs sont beaucoup plus propagés que les messages qui les proposent. Ils évoquent des pratiques

« d'infomédiation sociale » qui permettent un filtrage collaboratif des informations et « une propagation rapide par l'usage des fonctionnalités proposées par Twitter ». Dans le corpus de messages étudié par les auteurs, 44% des tweets émis sont destinés à infirmer des rumeurs. Les auteurs ne nient ainsi pas la présence de rumeurs sur Twitter, mais ils montrent comment la vigilance de certains des utilisateurs permet de les contrecarrer rapidement.

Ainsi, les utilisateurs de Twitter relaient des informations crédibles par le biais des processus de citations d'autres messages (Kwak et al., 2010). Par exemple, (Thomson et al., 2012) montre que, lors d'un tremblement de terre au Japon, les sources diffusées sur Twitter proviennent principalement de sites d'information crédibles. Les mêmes conclusions sont présentées à la fois pour la France – où les principales sources diffusées sur Twitter au moment d'une crise proviennent des media traditionnels (Rieder & Smyrnaio, 2012) – et pour les États-Unis, où les tweets les plus diffusés au moment d'une crise proviennent des media locaux et des services institutionnels (Starbird & Palen, 2010). Ainsi, les études sur le contenu généré par les utilisateurs soulignent la place importante accordée aux media traditionnels et aux institutions officielles en raison de leur haut degré de crédibilité (Austin et al., 2012; Leskovec et al., 2009). Ces auteurs suggèrent que le contenu publié par les media traditionnels et les institutions officielles a une plus grande influence en termes de communication, elle est mesurée par le nombre de citations, c'est-à-dire sur Twitter par le nombre de personnes ayant relayé une information émanant de ce type de sources.

Section 2 : Analyse descriptive de la base de données de tweets en langue française

Nous avons procédé à une première analyse descriptive de la base de données. Ces analyses s'appuient sur le travail interdisciplinaire mené par l'ensemble des participants à l'enquête MESCOV. L'objectif de cette première phase de l'analyse est d'appréhender la composition du set de données de manière fine et de repérer ainsi des mécanismes dans la construction et la diffusion de l'information autour de la pandémie.

1. Méthodologie

Différentes mises à plat de la base de données permettent une première analyse des orientations de la conversation. Elles sont réalisées à partir de l'identification de marqueurs caractéristiques propres à Twitter. Ils permettent de tracer à grandes lignes le contenu des messages et des débats et ainsi d'offrir une première compréhension de la structuration des échanges lors de la pandémie.

À ce stade de l'analyse, plusieurs marqueurs caractéristiques ont été repérés et analysés :

- **Utilisateurs** : Afin d'isoler la parole des personnalités publiques et des institutions, nous avons observé les « comptes vérifiés ». Ce statut proposé par Twitter est symbolisé par une icône, accolée au pseudonyme de l'utilisateur. Cette certification permet selon la plateforme « d'assurer aux utilisateurs l'authenticité d'un compte d'intérêt public » (Source : Twitter). Il peut être obtenu à la suite d'une demande de certification émanant de l'utilisateur lui-même. Dans le cadre de cette recherche, cette distinction permet d'isoler les prises de parole publiques et officielles ainsi que celles de personnalités médiatiques bénéficiant d'une audience importante. La certification renseigne sur l'authenticité du compte, elle est par exemple utilisée par les instances gouvernementales pour attester que c'est bien ici une parole officielle qui est partagée et non des citoyens qui en auraient emprunté le nom. Elle est également le gage d'une audience importante et peut éclairer sur les mécanismes de diffusion de l'information.
- **Hashtags** : Afin d'analyser un nombre important de messages diffusés sur Twitter, les hashtags constituent une porte d'entrée intéressante puisqu'ils doivent être considérés comme des éléments permettant d'offrir une contextualisation d'un message court. Il est alors un « élément contextuel fort qui ré-articule le sens du message en lui-même » (Alloing et Vanderbiest, 2018). Leurs usages sont multiples, ils sont notamment identifiés comme des outils permettant de « documentaliser le web » ou d'« offrir une possibilité ludique de détournement » (Thiault, 2015, p. 69). Cette perspective est néanmoins moins présente lors de crise importante (Bubendorff et al., 2019), les hashtags humoristiques par exemple étant beaucoup moins utilisés que lors d'évènements d'actualité plus communs comme une controverse liée aux activités d'une entreprise ou d'un homme politique par exemple.

Afin de les comprendre et d'analyser le data-set, nous avons procédé à une classification des hashtags sous deux axes : d'une part, nous avons étudié les hashtags les plus présents ; de l'autre, nous avons procédé à un codage des hashtags en fonction des catégories utilisées pour l'entraînement de l'algorithme et s'appuyant sur le méta-modèle et sur nos connaissances de la gestion de crise à partir des observations et des échanges avec les partenaires institutionnels du projet (cf. Partie I, Section 3.2).

- **Retweets** : Cette fonctionnalité présente sur Twitter permet à un individu de partager un message à l'ensemble de ses abonnés. Les messages ayant eu le plus grand nombre de retweet sont alors analysés afin de comprendre quels types d'informations ont attiré l'attention des internautes sur ce medium durant la pandémie.

Afin de traiter les messages composant le data-set, nous nous sommes appuyés sur le logiciel Compass Mongo DB qui permet l'exploration de la base de données grâce à des requêtes spécifiques portant sur les marqueurs cités précédemment. Des extractions de ces requêtes ont ensuite été analysées en détail.

2. Les hashtags comme marqueurs de l'ampleur de la crise

À l'inverse d'autres crises que nous avons pu observer sur les media sociaux (Bubendorff et al., 2019), la pandémie ne donne pas lieu à une structuration forte des discussions. Le data-set regroupe 289 019 intitulés de hashtags différents très inégalement utilisés pour contextualiser les messages des internautes. La Figure 10 donne une première idée de cette répartition inégale. Elle représente le nombre de citations des vingt hashtags apparaissant le plus dans le data-set. Ainsi, le hashtag « coronavirus », le plus présent dans nos données, accompagne 207 305 tweets quand celui référant au « confinement » -- cinquième le plus présent de l'échantillon-- apparaît uniquement dans 41 268 tweets.

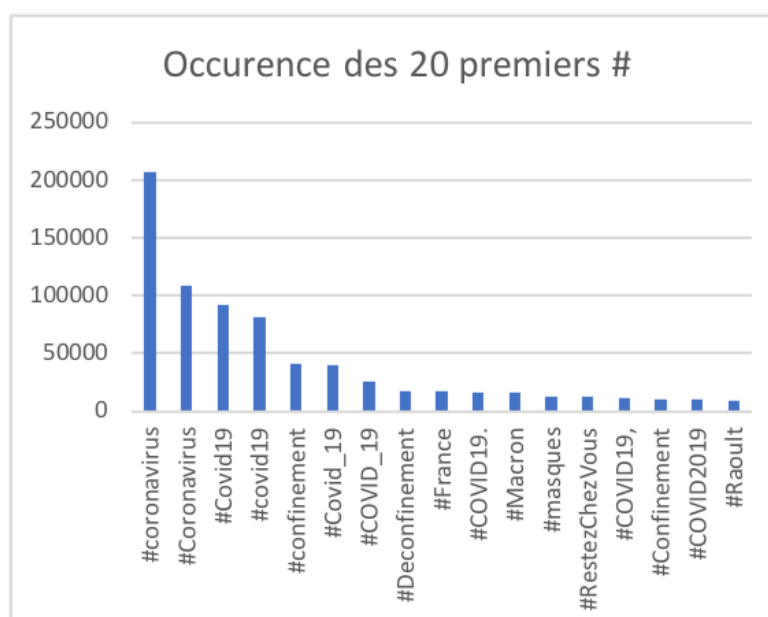


Figure 10 : Occurrence des vingt hashtags (#) les plus présents dans le set de données

Comme le montre la Figure 10, l'incertitude sur la dénomination stricte de la crise induit une grande disparité dans les hashtags. Il est courant de retrouver dans les données issues de Twitter un même phénomène orthographié de manières différentes dans les hashtags ; notamment lorsqu'ils qualifient un évènement récent encore non clairement établi. À ce phénomène bien identifié, s'ajoute ici la question même du nom de la crise, comme c'était déjà le cas pour Wikipédia. L'analyse et la classification montrent que 49 orthographes différentes coexistent dans la base de données pour qualifier le coronavirus et la maladie qu'il engendre (Covid-19).

Nous avons ainsi pu repérer une grande dispersion des hashtags témoignant du foisonnement des sujets abordés en lien avec la pandémie et de son étendue géographique. Afin de donner

plus de lisibilité aux données étudiées ici, nous avons concentré notre analyse des hashtags sur ceux étant présents plus de 500 fois dans l'ensemble de la base de données.

2.1. Les contours géographiques de la pandémie

L'examen des hashtags révèle la grande disparité géographique des usages francophones de la plateforme. L'analyse proposée ici permet néanmoins d'amoindrir les difficultés techniques liées à l'utilisation des métadonnées et décrites plus haut. Les particularités linguistiques et les références culturelles propres à chaque pays ou aire géographique peuvent ainsi être repérées dans les hashtags et affiner la compréhension du set de données. Ici, la complémentarité entre différentes méthodologies et disciplines prend tout son sens : l'absence de métadonnées précises permettant de filtrer l'origine des messages (voir l. section 3) est alors compensée du fait de cette analyse. Si les tweets ne peuvent être catégorisés directement suivant leur pays d'émission, les hashtags nous permettent d'inscrire les messages au sein d'aires culturelles.

Ainsi, au sein des hashtags présents plus de 500 fois, près de 20% font référence à un terme géographique. Ces termes sont variés et peuvent couvrir des échelles très différentes : continentale (e.g. le hashtag « Europe »), nationale (#Gabon), régionale (#Alsace) ou communale (#Rennes). Ces termes sont autant de marqueurs des périmètres de la crise et témoignent des niveaux très variés de discussion. La pandémie, contrairement à d'autres crises, s'exprime à plusieurs échelles qui ne s'excluent pas : les informations qui transitent sur ce réseau peuvent tout aussi bien revêtir un caractère global que tout à fait localisé. Une analyse plus détaillée du texte contenu dans ces tweets accolés aux hashtags pourra être menée afin de saisir si la nature des informations échangées est impactée par l'échelle dans lequel le message s'inscrit. On peut ainsi faire l'hypothèse qu'un plus grand nombre d'informations pratiques pourraient être échangées grâce à des hashtags locaux lorsque des informations concernant les connaissances sur le virus et la maladie lui étant associées pourraient transiter via des hashtags à une échelle plus large comme celle d'un pays.

De plus, une grande partie des mots utilisés pour contextualiser les messages diffusés sur le réseau peuvent être rattachés à des entités géographiques spécifiques. Outre les termes faisant clairement référence à une entité géographique comme ceux que nous décrivions ci-dessus, d'autres caractérisent des thématiques de discussions que l'on peut géographiquement situer. C'est par exemple le cas des hashtags #Team221 ou Team222 qui regroupent des messages en lien avec l'actualité sénégalaise. Ces éléments donnent une idée de l'ampleur de la crise puisqu'elle constitue un sujet de débats et de discussions dans l'ensemble du monde francophone.

2.2. Les sujets en lien avec la pandémie à travers l'utilisation des hashtags

En nous appuyant sur les catégories établies pour l'algorithme d'apprentissage, nous avons procédé à la catégorisation des 600 hashtags les plus présents dans le data-set. À titre indicatif,

cet échantillon représente plus de 62% de l'ensemble des tweets récoltés. Le détail de cette catégorisation est consultable dans le Tableau 1.

Tableau 1 : Catégorisation des 500 # les plus présents dans le data-set.

Catégorie	Hashtags		Messages	
	Nombre de #	Pourcentage de la catégorie	Nombres de tweets	Pourcentage de tweets
Zone géographique	111	19%	192082	10%
Coronavirus	49	8%	998497	55%
Politique	52	9%	92219	5%
Contraintes	43	7%	120843	7%
Santé	37	6%	65738	4%
Informations	35	6%	45391	2%
Controverses	22	4%	41914	2%
Critiques	20	3%	31547	2%
Recommandations	19	3%	31578	2%
Solidarité / init. citoyennes	14	2%	18088	1%
Entreprises / Économie	37	6%	40120	2%
Crise	8	1%	12446	1%
Ressources	8	1%	22371	1%
Application « Stop-Covid »	2	0%	3901	0%
Éducation	7	1%	6607	0%
Sport	4	1%	3765	0%
Non pertinent	118	20%	104375	6%
Somme	586	100%	1 831 482	100%

Ce travail de catégorisation permet de mettre en avant les sujets déjà identifiés dans notre base de données. Ici, on peut alors voir que les hashtags les plus usités lors de notre période de collecte renvoient à des zones géographiques, mais également au virus lui-même ainsi qu'à la maladie qu'il provoque. Ainsi, au sein des 500 hashtags les plus représentés dans le data-set, la

catégorie zone géographique rassemble 111 hashtags différents et représente à elle seule 19% des hashtags présents dans l'échantillon. Une autre lecture de cette catégorisation permet d'éclairer sur la distribution des messages. Ainsi, la catégorie zone géographique est présente dans 192 082 tweets et représente 10% des messages étudiés ici. Malgré la prédominance nette de certaines des catégories, ce travail de catégorisation des mots-dièse permet de comprendre les pôles d'attentions autour desquels se structure la conversation. Ces thèmes de conversations pourront ainsi être recherchés plus en détail dans le contenu des messages (le texte) cette fois-ci et non dans les métadonnées qui l'accompagnent.

L'analyse des hashtags permet alors de conserver une trace des thématiques s'y trouvant et constituant un canevas pour les futures analyses menées sur cette base de données. En effet, on retrouve bien les sujets d'actualité en lien avec la pandémie dans les différentes catégories. Ainsi, les hashtags faisant référence aux contraintes liées à la pandémie sont également très présents (plus de 120 000 messages, soit 7%). Ici, la disparité géographique semble s'effacer totalement au profit d'une concentration des messages sur la situation française. Les hashtags qui forment cette catégorie sont uniquement en lien avec le confinement et le déconfinement décidé par le gouvernement français. Les autres hashtags qui apparaissent dans cette catégorie sont liés aux conditions de travail durant cette période particulière (#télétravail, #chomagepartiel).

De même, les différentes controverses qui ont émaillé l'émergence du virus peuvent également être repérées au travers des hashtags. La catégorie éponyme regroupe ainsi des hashtags faisant référence à des pistes de traitement ayant été discutées (#hydrochloroquine), les personnalités publiques scientifiques ou non, ayant pris part au débat sur la gestion de la crise (#Raoult).

3. Les prises de parole

Afin de comprendre qui diffuse de l'information sur Twitter lors de la pandémie, nous avons analysé les comptes ayant posté le plus de messages au sein de la base de données. Nous avons donné la priorité à l'analyse des comptes certifiés car ils sont ceux bénéficiant de la plus grande audience sur ce médium. Du fait de la grande disparité du nombre de messages postés par les différents comptes certifiés, nous avons concentré notre attention sur les 200 comptes les plus actifs au sein de l'échantillon. Pour l'ensemble de ces comptes, nous avons procédé à une classification permettant de définir la fonction des utilisateurs y étant liés. Nous avons ainsi différencié les messages émanant de :

- Media (1 dans le Tableau 2) : ici sont confondus tous les types de travail journalistique. Aucune différence n'est faite entre des media présents uniquement sur le web (comme le HuffingtonPost par exemple) et d'autres media historiques dont la présence en ligne est récente en regard de leur durée d'existence (comme Le Monde par exemple).
- Institution gouvernementale et intergouvernementale (2) : cette catégorie regroupe les institutions gouvernementales (Préfecture) et intergouvernementales (OMS), elle exclut les personnalités politiques disposant d'un compte en leur nom propre.
- Citoyens (3) : cette catégorie regroupe des personnalités publiques non affiliées à un parti politique.
- Journalistes (4) : cette catégorie regroupe les individus ayant un compte en leur nom propre, mais se déclarant comme journalistes.

Il est intéressant de souligner ici que la catégorie « personnalité politique » avait été créée, mais a été supprimée du fait de son absence.

Tableau 2 : Catégorisation des 200 utilisateurs les plus présents dans le data-set et volume de messages

	Comptes		Tweets	
1. Media	148	74%	28016	88,4%
2. Institutions	31	16%	1853	5,8%
3. Citoyens	2	1%	108	0,3%
4. Journalistes	18	9%	1700	5,4%

L'analyse de la répartition des messages confirme les conclusions apportées par la littérature. L'information partagée sur Twitter l'est le plus souvent par des sources fiables et de qualités. S'ils ne sont pas systématiquement parmi les messages les plus partagés ou commentés, les comptes officiels occupent cependant une place importante au sein de cet espace public en ligne. Ici, 74% des messages proviennent de media certifiés. En regardant plus en détail quel type de media prend la parole sur ce réseau, on constate que les plus actifs sont des media traditionnels ayant basculé une partie de leur activité en ligne.

Les prises de paroles témoignent comme c'était le cas des hashtags de la mobilisation de toute la francophonie autour des la pandémie. Ainsi des grands quotidiens français (Le monde, Le Parisien, Ouest France) côtoient des media belges (RTBF) ou québécois (RadioCanadaInfo).

De même, les organisations inter-gouvernementales ont été très présentes sur Twitter lors de cette pandémie. L'ONU s'est par exemple beaucoup servie de ce réseau pour partager des informations sur la situation sanitaire internationale et des recommandations, notamment via le compte de l'HCR (haut-commissariat des Nations unies pour les réfugiés en France).

Il est également intéressant de noter qu'au sein des comptes liés à des organisations politiques nationales ou internationales, on peut constater une mobilisation particulière des Ministères de la Santé de différents pays (France et Canada notamment). Plus spécifiquement pour les comptes français, l'observation des comptes les plus actifs laisse voir les territoires les plus touchés par la première vague de la pandémie : les préfets de la Guyane, du Bas-Rhin ou encore des Deux-Sèvres ont été particulièrement actifs sur la période de récolte des données, ces territoires ayant été parmi ceux connaissant le plus grand nombre de cas et d'hospitalisations lors de la première vague au printemps dernier.

La prise de parole de ces institutions via les media sociaux, les informations transmises directement aux citoyens semblent ainsi s'instituer de plus en plus nettement comme un canal de communication pour les autorités politiques françaises lors de périodes de crise. Une analyse plus poussée du contenu de ce message permettra de voir dans quelles mesures ces institutions ont investi la communication directe vers les citoyens. Elle permettra notamment de voir si ces messages consistent en des partages de recommandations ou si la pandémie a été le moment d'échanges moins formels entre agents du service public et citoyens.

Section 3 : Les usages de Wikipédia lors de la pandémie

1. Méthodologie

Afin de comprendre de quelle(s) manière(s) les contributeurs de Wikipédia contribuent à conférer du sens à la pandémie de Covid-19, nous avons mis en place une méthodologie mixte axée autour d'un travail d'ethnographie en ligne sur une sélection d'articles en lien avec la pandémie (voir Tableau 3) et avons mené des entretiens avec certains contributeurs de Wikipédia étant intervenus sur un ou plusieurs articles en lien avec les crises (voir Tableau 4).

Tableau 3 : Liste des pages Wikipédia observées

Titre de la page	Date de création	URL	Nombre de contributions	Nombre de contributeurs
Pandémie de Covid-19	19 janv. 2020	https://fr.wikipedia.org/wiki/Pandémie_de_Covid-19	1471	8485
Pandémie de Covid-19 en France	29 fév. 2020	https://fr.wikipedia.org/wiki/Pandémie_de_Covid-19_en_France	1041	7264
Désinformation sur la pandémie de Covid-19	26 mars 2020	https://fr.wikipedia.org/wiki/Désinformation_sur_la_pandémie_de_Covid-19	113	735
Test diagnostique du SARS-CoV-2	3 mars 2020	https://fr.wikipedia.org/wiki/Test_diagnostique_du_SARS-CoV-2	1114	71
Didier Raoult	10 déc. 2008	https://fr.wikipedia.org/wiki/Didier_Raoult	2136 (1175 en 2020)	381

En définissant et en appliquant un protocole spécifique, nous avons observé et collecté le contenu publié sur Wikipédia au moment de la pandémie. Pour chacun de ces articles, en nous servant des outils collaboratifs proposés par Wikipédia (archivage exhaustif de l'historique des pages) et par la fondation Wikimedia (notamment Xtools donnant accès à des données statistiques), nous avons analysé et comparé :

- Les historiques des articles : date de création, principales périodes et types de contribution ;
- La liste des contributeurs : nous avons extrait les listes complètes des contributeurs afin de comparer différents types d'articles (sanitaires, actualité) et de repérer des « catégories » de contributions sur l'encyclopédie. Nous avons par ailleurs analysé leurs autres contributions afin d'affiner cette typologie de contributeurs.

Nous avons également réalisé des observations directement sur ces articles. Elles nous permettent de comprendre spécifiquement comment les sources validant les informations sont commentées et débattues entre contributeurs. Ainsi, le caractère délibératif de Wikipédia s'incarne davantage dans les pages « annexes » aux articles : les pages de discussions et le « bistrot ». Les pages de discussions sont attachées à l'article principal (sous la forme d'un onglet) et rassemblent les discussions entre contributeurs. Ce forum est principalement utilisé pour

résoudre différents conflits d'édition (allant de la syntaxe de l'article aux types d'informations devant figurer dans l'article). Les pages « bistrot » rassemblent quotidiennement des discussions sur les actualités du monde et du site. Lorsque les évènements en cours sont particulièrement importants, comme c'est le cas pour la pandémie, des discussions à propos de leur traitement sur le site se retrouvent également dans cette section d'échanges du site. Ces échanges se poursuivant parfois directement sur les pages de discussions des utilisateurs, nous avons également prêté attention aux échanges ayant lieu sur les pages personnelles de certains des contributeurs les plus actifs.

Enfin, afin d'avoir une vue la plus globale possible de ces discussions, nous avons également analysé les pages de discussions liées à certaines des pages d'aide aux utilisateurs de Wikipédia :

- Observatoire des sources⁷ ;
- Citez vos sources !⁸ ;
- Sources fiables⁹.

Pour l'ensemble de ces pages, nous avons réalisé des captures d'écran afin de procéder à des analyses au moyen du logiciel Atlas.ti. Nous avons codé certaines parties des discussions menées entre les contributeurs afin de mettre en avant les processus de délibération à l'œuvre sur ces pages.

Ce corpus est enrichi d'un entretien avec un membre de la fondation Wikimedia, de la consultation et de l'analyse de documents de travaux élaborés par la fondation dans le cadre de la pandémie (Wikimédia Fondation, 2020) et d'entretiens avec certains des contributeurs d'articles relatifs à des crises de sécurité civile (attentat de Charlie Hebdo et du 13 novembre 2015 en France) ou à la pandémie de Covid-19 (Tableau 4). Ces entretiens ont été codés et analysés à l'aide des mêmes grilles d'analyse et logiciel que les ethnographies menées en ligne.

Tableau 4 : Sociographie des interviewés

Id.	Modalité de l'entretien	Genre	Âge	Profession	Nombre de contributions	Durée de contribution	Contribution à l'article			
							Pandémie de covid-19	Désinformation sur la pandémie de Covid-19	Attentat contre Charlie Hebdo	Attentats du 13 novembre 2015 en France
1	Appel	M	22	Étudiant	13028	NR	Non	Non	Oui	Oui
2	Appel visio	M	55	Journaliste	40954	10 ans ou +	Oui	Non	Oui	Oui
3	Entretien in situ	M	65	Ingénieur informatique	35218	10 ans ou +	Oui	Non	Oui	Oui
4	Appel visio	M	63	Retraité	27078	10 ans ou +	Non	Non	Non	Oui
5	Appel visio	M	58	Ingénieur informatique	38152	10 ans ou +	Oui	Oui	Oui	Non
6	Appel	M	NR	NR	89390	10 ans ou +	Oui	Oui	Non	Non

⁷ https://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Observatoire_des_sources

⁸ https://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Citez_vos_sources

⁹ https://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Sources_fiables

2. « Faire sens » : un processus en deux temps

Une première observation du traitement de la pandémie révèle qu'elle présente les mêmes caractéristiques immédiates au traitement d'une crise : les contributeurs de l'encyclopédie sont très nombreux et ils s'emparent très rapidement de cette thématique. Ainsi, le premier article sous le titre "*L'épidémie de pneumonie en Chine 2019-2020, également connue sous le nom de pneumonie de Wuhan*" est créé dès le 19 janvier 2020. Aujourd'hui appelé "Pandémie de Covid-19", il compte plus d'un millier de personnes – quand le nombre moyen de contributeurs se situe habituellement autour de 7 pour l'ensemble des articles de l'encyclopédie (Auray et al., 2009) (Voir Tableau 3). Cependant, la pandémie de Covid-19 se caractérise par une cinétique lente qui va fortement marquer la manière dont elle est traitée, c'est-à-dire la manière dont on va lui donner ou créer du sens dans cette période de forte incertitude. La cinétique lente de la pandémie permet une observation particulièrement détaillée de la construction du sens autour de cet événement incertain. Plus précisément, la réorganisation des manières de faire durant la pandémie est particulièrement observable.

Les résultats de la première section mettent en évidence la dynamique en deux temps du traitement de la pandémie sur Wikipédia : 1) un travail d'archivage des événements afin de suivre leur chronologie dans le temps et l'espace, et 2) un travail de documentation des analyses relatives à la pandémie face à sa complexification. La section 2 met en évidence l'adaptation des mécanismes habituels à l'œuvre dans Wikipédia : une mobilisation accrue des outils de vigilance fournis par la plateforme ainsi qu'une prise en charge inédite de la désinformation relative à la pandémie.

2.1. Archiver les événements

Dès l'apparition des premiers cas en Chine, la prise en charge du sujet est très rapidement mise en œuvre par les contributeurs francophones. L'article consacré à cette maladie est créé le 19 janvier 2020 sous le nom de « *L'épidémie de pneumonie en Chine 2019-2020, également connue sous le nom de pneumonie de Wuhan* » et celui propre à la situation en France le 29 février, avant même que l'État ne considère publiquement les faits comme préoccupants¹⁰ (cf. Tableau 3). Les questionnements provoqués par la situation internationale sont cependant abordés par les contributeurs de l'encyclopédie qui tentent alors d'y répondre. L'ampleur à venir de la crise est en quelque sorte anticipée par les internautes et son archivage au niveau local est lancé dès la fin du mois de février 2020 dans un article propre aux événements en France. Ce dernier n'est pas discuté ou remis en question alors qu'habituellement, pour respecter la finalité encyclopédique de Wikipédia, la création de ce type d'article est souvent source de débats entre contributeurs pour juger ou non de la pertinence et du caractère marquant d'un point de vue historique de l'évènement. À titre de comparaison, l'article concernant les attaques terroristes en Île-de-France en novembre 2015, créé le soir même des fusillades, avait été supprimé à plusieurs reprises avant d'être stabilisé. Dans le cas de la pandémie, dès le début de l'année 2020, et alors qu'elle n'est pas encore qualifiée comme telle par aucune instance officielle, l'article consacré à la pandémie en France entame un recensement des faits se déroulant sur le territoire national qui sera quotidiennement mis à jour (répertoire des cas importés, des hospitalisations, des prises de parole des membres du gouvernement). L'article se résume alors à un historique organisé

¹⁰ La semaine suivante, le Président de la République appelle les Français à continuer à vivre normalement

autour du rapatriement des français, de la propagation du virus, et des mesures sanitaires évoquant le confinement de certains ressortissants rapatriés sur le territoire national. Dès sa création, l'article est accompagné d'une carte nationale monitorant la présence du virus sur le territoire (voir Figure 13).

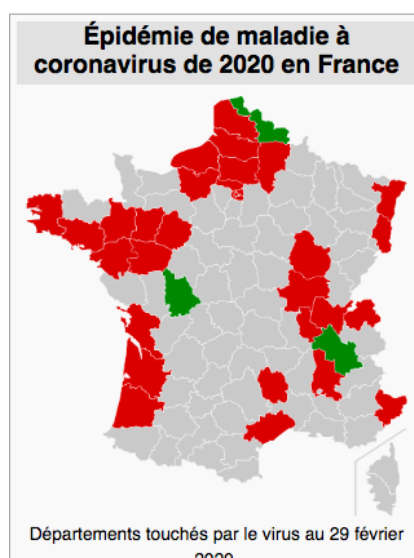


Figure 11 : Carte de la présence du virus sur le territoire, page Pandémie de Covid-19 en France¹¹

Première étape d'un processus de création de sens, cet article constitue à cette période pour les contributeurs un repère se voulant exhaustif et, permettant à tous les lecteurs d'accéder à l'évolution de la pandémie et des connaissances à son sujet.

2.2. Documenter la complexification de la pandémie

Cette synthèse de l'historique de la pandémie se complexifie moins d'un mois après le début des événements par le recensement des analyses qui se multiplient. Le choix du titre de l'article en est le premier exemple marquant. La page d'origine « *L'épidémie de pneumonie en Chine 2019-2020, également connue sous le nom de pneumonie de Wuhan*¹² » est ainsi renommée successivement en « *épidémie de Covid-19* » et en « *Pandémie de Covid-19* » dès début mars. L'analyse de la page de discussion montre que ce débat sur le titre, portant alors sur les contours de l'épidémie, émerge peu de temps après la création de la page : la première mention en faveur d'un changement a lieu dès le 26 janvier 2020. Un consensus est trouvé et aboutit au nom définitif de l'article le 2 mars 2020. Au fur et à mesure de la crise, ce type d'exemples se multiplie. Plusieurs points de l'actualité donnent lieu à des controverses et à des échanges fouillés entre contributeurs. Dans certains cas, les sujets de discussion suivent de près ceux fortement présents dans les media : la personnalité publique de Didier Raoult et la place dont il doit bénéficier dans le traitement de l'information, l'hydroxychloroquine et la manière dont cette molécule doit être décrite, ou non, comme un traitement au virus dans Wikipédia. En cela, l'encyclopédie n'échappe pas aux débats qui traversent la société dont elle participe.

¹¹ Source: https://fr.wikipedia.org/w/index.php?title=Pand%C3%A9mie_de_Covid-19_en_France&diff=167938285&oldid=167938239#Confinement

¹² Voir source : https://fr.wikipedia.org/w/index.php?title=Pand%C3%A9mie_de_Covid-19&dir=prev&action=history

Ces éléments sont illustratifs de comment un sens est créé dans la synthèse proposée sur les pages relatives à la pandémie. Cependant, d'autres discussions sont encore plus révélatrices de la volonté des contributeurs de réduire l'incertitude autour de la pandémie. Il s'agit notamment du choix de la présentation des données chiffrées liées à la pandémie et à sa progression. Une vaste discussion intitulée « *le principal critère à considérer toujours absent...*¹³ » est par exemple engagée à propos de l'article « *Pandémie de Covid-19* » sur la nécessité de fournir uniquement des chiffres « parlants ». Cette discussion se cristallise sur l'absence du taux d'incidence des différents pays parmi les chiffres proposés dans l'article alors que les seules données quotidiennement mises à jour par les rédacteurs sont celles des cas diagnostiqués et des décès. Le taux d'incidence est décrit par les contributeurs à cette discussion comme étant un « *critère primordial* » ou « *bien plus pertinent*¹⁴ » pour la compréhension de l'épidémie et de sa propagation. Les six contributeurs partageant tous ce point de vue, leurs échanges portent plutôt sur la définition de ce taux, et notamment sur le fait qu'il soit ou non un bon indicateur de la propagation de la maladie. Malgré le consensus de départ, la conversation se termine lorsque deux des contributeurs initient le débat sur la corrélation entre taux d'incidence et dangerosité de vivre dans un pays donné. Le débat s'oriente alors sur la stratégie de confinement des pays européens. Cela illustre l'ampleur de la portée des discussions relatives aux enjeux de la pandémie et au sens associé. Le taux d'incidence a été ajouté à la liste des données régulièrement mises à jour à la suite de cette discussion.

Loin d'être anecdotiques, ces exemples mettent en évidence les processus discursifs qui étayent la construction de l'information sur les articles de Wikipédia. Ainsi, la dénomination même de l'article ou le choix des indicateurs dépend de la signification accordée par les contributeurs aux termes et aux chiffres. Ils témoignent de la compréhension de l'évènement : a-t-il des frontières nationales ? Implique-t-il l'ensemble de l'humanité ? Comment est-il possible de comparer la situation des différents pays ? En échangeant leurs points de vue, en comparant des sources, les contributeurs aux articles de crises de l'encyclopédie ne se contentent pas d'agréger des informations, mais entament ainsi le processus d'élaboration de la connaissance sur les évènements. Comme l'illustre la discussion sur le taux d'incidence mentionnée ci-dessous (Figure 14), en parallèle de la construction de l'article même relatif à la pandémie, ces échanges contribuent à réduire l'incertitude, non seulement pour ceux qui l'élaborent, mais également pour les lecteurs de l'article. La Figure 14 permet de voir que c'est précisément ce processus qui est salué par les contributeurs.

Bravo et merci ! [\[modifier le code \]](#)

Je tenais à envoyer ce petit message car de nombreuses personnes de ma connaissance visitent cette page et semblent ravies de son existence...


Sinon, plus de 500 000 lecteurs en dix jours, ce n'est pas mal. Même l'article sur le **Mouvement des Gilets jaunes** -que j'avais initié- n'avait pas connu une aussi grande popularité (et surtout un tel suivi).

Sincèrement, même si certains passages sont très certainement perfectibles, au niveau des tableaux affichés sur la page, il ne faut pas faire la fine bouche. Ceux-ci sont nombreux, clairs (et faciles à comprendre) et, de plus, actualisés de façon constante.

Bravo à tous les contributeurs qui s'y sont impliqués ! Continuez ! [des questions ?](#) 13 avril 2020 à 11:14 (CEST)

+1 [discuter](#) 14 avril 2020 à 10:50 (CEST)

+1 [discuter](#) 14 avril 2020 à 19:22 (CEST)



¹³ Voir discussion archivée sur Wikipédia, source : https://fr.wikipedia.org/wiki/Discussion:Pandémie_de_Covid-19/Archive_2#Le_principal_critère_à_considerer_toujours_absent_..., consulté le 25.08.20.

¹⁴ Source : *ibid.*

Figure 12 : Capture d'écran, onglet discussion¹⁵

Dans l'exemple ci-dessus, un contributeur habitué des articles de crise prend la parole afin de remercier les rédacteurs (soulignement bleu) de l'article tout en insistant sur son utilité auprès d'un public de lecteurs, éloignés des processus de contributions sur Wikipédia (soulignement rouge). Il est remercié à son tour par d'autres contributeurs au moyen des « +1 » en réponse à son message. L'analyse de cette interaction met en avant la poursuite d'une création de sens autour d'un évènement par les contributeurs de l'encyclopédie, le motif des félicitations étant bien ici l'utilité de l'article face à l'incertitude de la situation.

3. Les spécificités du traitement de la pandémie sur Wikipédia

3.1. La mobilisation accrue des outils de vigilance

Les rédacteurs de Wikipédia sont habitués à construire la synthèse d'une information lors d'un évènement de crise (Bubendorff et al., 2019; Keegan, 2015). Cependant, l'analyse des pages relatives à la pandémie montre qu'elle met à l'épreuve ces manières de faire et met en tension les habitudes de contribution des rédacteurs.

Le processus d'élaboration des pages relatives à la pandémie présente les mêmes caractéristiques que celles des pages relatives aux crises de sécurité civile à la cinétique rapide. Cependant, l'ampleur de la pandémie (c'est-à-dire son temps long et la profusion des évènements et débats associés) complexifie le travail de synthèse des rédacteurs et entraîne un déploiement inédit des outils à disposition des contributeurs. Ainsi, cette difficulté est reconnue par l'encyclopédie et signifiée aux lecteurs par différents moyens. La multiplication des recours aux outils de régulation peut être envisagée comme la première notification de la difficulté pour les contributeurs à proposer une synthèse jugée satisfaisante des évènements. Des bandeaux d'avertissements sont incrustés en en-tête de plusieurs articles traitant de la pandémie. De même, tous sont au moins temporairement verrouillés pour y limiter le nombre de modifications. Certains contributeurs ont également multiplié leurs actions de modération afin de contribuer à la stabilité du contenu. Ces interactions visent notamment à rappeler aux rédacteurs des articles qu'il est nécessaire de référencer les informations ajoutées (Figure 15).

¹⁵ Source : https://fr.wikipedia.org/wiki/Discussion:Pandémie_de_Covid-19_en_France/Archive_2#Bravo_et_merci_!

Votre modification sur « Pandémie de Covid-19 » n'est pas vérifiable par une source : merci d'en ajouter une. [modifier le code]

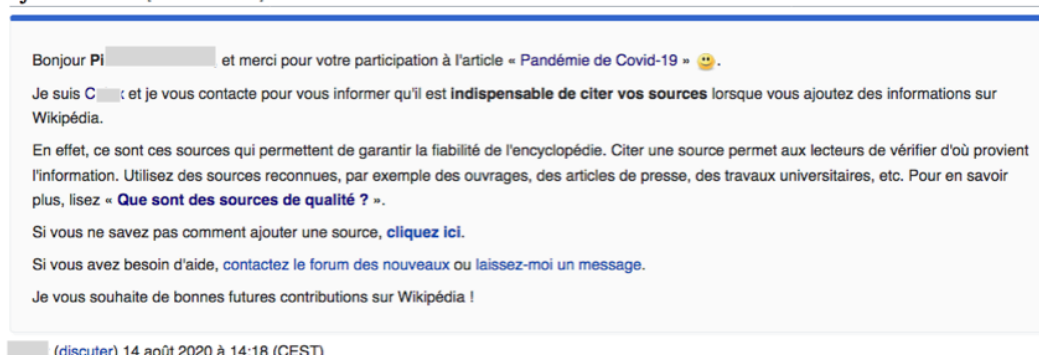


Figure 13 : Capture d'écran, onglet discussion : page personnelle d'un contributeur¹⁶

L'amplification des activités de modération de la part des contributeurs plus avertis peut être analysée comme le signe de leur volonté de ralentir le rythme des modifications dans le but de renforcer la fiabilité des articles. L'un des contributeurs interrogés explique vérifier quotidiennement les modifications réalisées sur les articles liés aux crises. En détaillant cette manière de faire, il souligne :

« Mais effectivement sur les articles liés aux actualités là beaucoup de contributeurs les suivent, de contributeurs expérimentés j'entends, et donc effectivement à ce niveau-là c'est... les théories du complot ont du mal à se répandre, c'est vrai que c'est une des forces de Wikipédia. » (id. 1). Ce choix est fait au détriment des mises à jour plus fréquentes des articles qui pourraient corriger la présence d'informations ou de chiffres obsolètes. Comme l'illustre le message présenté en Figure 15, certains contributeurs se lancent alors dans un travail de sensibilisation afin de stabiliser les informations et de construire la connaissance encyclopédique, faisant face par la même au rythme soutenu de sa création et de son évolution dans Wikipédia.

Ce processus perdure par ailleurs dans le temps dans le cas de la pandémie et de la désinformation associée.

3.2. Un traitement inédit de la désinformation relative à la pandémie vue par et depuis Wikipédia

La pandémie qui s'étend mondialement embarque avec elle son lot de rumeurs et de désinformations. Ce brouhaha informationnel impose aux rédacteurs un travail encore plus minutieux de triage et de filtrage. L'un des contributeurs interrogés souligne la difficulté de procéder à ce tri lors d'actualités « brûlantes » :

« On sait que sur des sujets aussi sensibles que ça, c'est-à-dire une pandémie, il faut faire très attention y compris aux chiffres. On sait que les chiffres issus de certains pays, je vais pas les mentionner, mais sont pas du tout fiables. C'est-à-dire que le nombre de morts, le taux de mortalité, le taux de morbidité sont en fait faux. Malheureusement c'est tout ce qu'on a. Et en

¹⁶ Source :

https://fr.wikipedia.org/wiki/Discussion_utilisateur:Pierre_VONESSA#Votre_modification_sur_%C2%AB_P_and%C3%A9mie_de_Covid-19_%C2%BB_n'est_pas_v%C3%A9rifiable_par_une_source:_merci_d'en_ajouter_une.

plus c'est une information... le fait que ce soit une information officielle est une information en soi. Donc on ne peut pas l'ignorer au motif qu'on a des doutes (...). On peut juste prendre des précautions en disant « tel pays d'Asie annonce qu'ils ont eu deux morts et qu'eux s'en sont très très bien sortis », donc on le dit, on le dit comme ça. (...) Dans ce genre de sujet il faut pas qu[e les lecteurs] soient trompés. C'est-à-dire qu'on essaie d'éviter qu'ils soient trompés donc on va dire les choses avec précaution. » (id. 5).

Comme le montre cet extrait, la difficulté pour les contributeurs réside dans l'apport d'informations de qualité au lecteur.

3.2.1. Un article dédié

Pour la première fois, un article uniquement consacré à la désinformation pendant un évènement est publié sur Wikipédia. Initialement publié sur le Wikipédia anglo-saxon puis traduit dans sa version francophone, il intègre les rumeurs et les campagnes de désinformation dans l'histoire de l'évènement, comme constitutives même de son historique. Alors que les principes fondateurs de Wikipédia rappellent qu'il ne s'agit pas de proposer des analyses originales des faits (le *travail inédit* est interdit sur Wikipédia) et que les contributeurs ne doivent pas se poser en juge de la véracité des faits, cet article liste l'ensemble des fausses informations ayant été « débunkées »¹⁷ par différents media ou analystes. Il propose une catégorisation du type de fausses informations et de rumeurs en fonction de leur origine ou de leur thématique. L'existence de l'article ne fait pas l'unanimité et est sujet à débat : sur la page de discussion associée, des contributeurs initient un vote afin de juger de sa pertinence, appelant à sa suppression en raison de l'absence de caractère encyclopédique de la démarche. Les seize contributeurs en faveur de son maintien soulignent dans leur argumentaire le travail de qualité de l'article : des sources « *fiabiles* » sont présentes, « *il sert l'intérêt de la communauté* », « *la désinformation en médecine est un souci majeur* »¹⁸. S'attaquant à la question de la désinformation, Wikipédia ré-affirme cette volonté d'aider à la construction rationnelle de sens autour d'un évènement. Comme le souligne le contributeur interrogé en mentionnant les chiffres officiels de certains pays, la désinformation fait partie intégrante de l'information et il s'agit de la prendre en compte, de la recenser pour *comprendre* les évènements dans leur globalité.

Au cours de la pandémie, Wikipédia a d'ailleurs été soupçonné par certains media et internautes d'être touché et traversé par cette campagne de désinformation. L'accusation concerne la suppression d'un article sur une personnalité publique, « Karine Lacombe ». Cette suppression a été perçue comme une volonté d'effacer certains aspects de la gestion de la pandémie, c'est-à-dire une volonté de travestir le déroulement des évènements. La suppression n'a cependant pas été attribuée à une entité bien définie, comme le rappelle le rapport publié par la fondation Wikimedia au sujet de la pandémie de Covid-19 : « *Sa biographie constituée d'une seule phrase a immédiatement été supprimée, ce qui a très vite été relayé sur les réseaux sociaux. Des internautes y ont vu l'intervention de Karine Lacombe, d'autres du gouvernement, voire un coup de la Banque mondiale parce que la directrice générale de Wikimedia Foundation y a travaillé.*

¹⁷ Anglicisme, démythifier, exposer ou discréditer des revendications supposées fausses, exagérées ou prétentieuses.

¹⁸ Source :

https://fr.wikipedia.org/wiki/Discussion:D%C3%A9sinformation_sur_la_pand%C3%A9mie_de_Covid-19/Suppression, consulté le 12.09.20.

Wikimédia France a immédiatement décidé de réagir en rappelant dans un billet de blog le fonctionnement de Wikipédia et apporter une réponse aux rumeurs. La suppression de l'article fait suite à l'application des critères de notoriété de la Wikipédia en Français ». Si l'encyclopédie se targue de ne mettre en avant que des informations vérifiables et de ne pas faire d'entorses à ses principes de fonctionnement, notre analyse montre que la pandémie fragilise quelque peu cette position.

3.2.2. Un travail supplémentaire sur les sources

Outre cette prise de parole de la fondation dans l'espace public, la création d'un article « Observatoire des sources » formalise des critères de légitimité jusqu'alors plus tacite entre les contributeurs.

Comme nous le soulignons plus haut, si les échanges à propos des sources et de leur légitimité sont nombreux, c'est justement parce qu'elles sont laissées ouvertes à discussions. Jusqu'à très récemment, aucun document ne statuait *in fine* sur la qualité des revues scientifiques, éditeurs ou journaux de presse. Les débats se tenaient autour de sources précises, entre contributeurs d'une page. La qualité des sources relevait, comme le rappelle la page d'aide « Citez vos sources », du « principe de bonne foi » et de bonne volonté de chaque contributeur. Dans cet article, la fiabilité de divers media, de revues académiques, d'émissions télévisées est discutée afin de proposer un état des lieux des sources pouvant être mobilisées dans Wikipédia. Par ailleurs, la page de discussion associée à l' « Observatoire des sources » comporte une section réservée aux « revues prédatrices¹⁹ » c'est-à-dire des revues qui ne soumettent pas leurs articles à la relecture des pairs ou bien dont la publication des articles est payante pour l'auteur. Cette discussion s'initie autour de la publication d'un canular dans « *Asian Journal of Medicine and Health* » et de l'inscription de cette revue dans la liste principale de l'article comme une revue non fiable. Rapidement, elle se poursuit autour du statut même des articles scientifiques et de leur caractère encyclopédique. La rapidité des publications scientifiques autour de la pandémie de Covid-19 conduit à interroger la fiabilité de ces publications au même titre que pouvaient l'être jusque-là les publications médiatiques.

La publication de l'article « Observatoire des sources » marque à notre sens l'un des plus importants bouleversements engendrés par la pandémie : l'extraordinaire des événements et leur complexité, sur la durée et sur l'impact, a rendu nécessaire la formalisation de critères de qualité à la fois plus fins et explicites. En effet, comme cela a été souligné, les contributeurs de Wikipédia perçoivent comme d'une grande fiabilité les sources émanant de la sphère académique. Cependant, la pandémie de Covid-19 a ébranlé cette autorité. La publication et la reprise massive dans les media d'articles scientifiques rapidement contredits ou non évalués par les pairs ont en quelque sorte remis en question cette certitude et entamé une discussion sur le statut de ces sources.

¹⁹ Source :

https://fr.wikipedia.org/wiki/Discussion_Wikip%C3%A9dia:Observatoire_des_sources#Revues_pr%C3%A9datrices

Section 4 : Un exemple d'organisation citoyenne sur les media sociaux en temps de crise - les makers

NB : Cette section présente un aspect des travaux de thèse en cours de Robin Batard qui portent d'une manière large sur les media sociaux comme vecteurs d'intégration des initiatives citoyennes lors d'évènements majeurs.

1. Présentation de l'étude

La crise liée à la pandémie de Covid-19 rencontrée en 2020 a également vu l'émergence de nombreuses initiatives citoyennes, dont l'action, l'organisation, et la communication ont été supportées par l'usage des media sociaux. Ces initiatives citoyennes ont émergé en ligne, et ont orienté leur activité sur la production de ressources, telles que des respirateurs, des masques chirurgicaux, ou encore des visières de protection. Ces réseaux citoyens peuvent se regrouper sous le terme générique de *makers*, et s'inscrivent dans la culture du *do-it-yourself*. Ainsi, l'activité de ces groupes pendant la crise a été marquée par la conception et la mise en place collective de solutions techniques et logistiques en vue de pallier le manque de ressources généré par la crise. Des ressources telles que des visières de protection ou des masques en tissu ont été confectionnées, et distribuées gratuitement auprès des systèmes de santé (hôpitaux, infirmeries, EHPAD), de commerces de première nécessité, ou de services des forces de l'ordre.

Entre mai et septembre 2020, un projet de suivi des initiatives citoyennes portées par des communautés indépendantes a été mené. Quatre entretiens semi-directifs ont été conduits avec des volontaires qui ont donné de leur temps pour participer à ces initiatives, et ainsi supporter la réponse citoyenne au manque de ressources, telles que des masques en tissu, ou des visières de protection. Les quatre personnes rencontrées au cours de cette étude se sont mobilisées au sein de deux réseaux de *makers* :

- Le mouvement Visière Solidaire s'est créé autour de la production de visières de protection, généralement fabriquées en impression 3D. Ce mouvement à l'échelle nationale est né sur Facebook, et s'est décliné au sein de nombreux départements, via des pages et des groupes de discussion locaux.
- Le réseau des élèves de l'École Arts et Métiers, soutenu par l'association de l'Union des Élèves, a mis en place des systèmes de production de visières, de conception de masques en tissu, et des réseaux logistiques pour supporter leur distribution. Ce réseau s'est formé à la suite d'une initiative d'un étudiant, partagée sur LinkedIn.

Les profils des quatre personnes rencontrées sont rassemblés au sein du Tableau 5.

Tableau 5 : Sociographie des enquêtés - réseaux de makers pendant la crise Covid-19

id	Statut	Réseau	Rôle
1	Étudiant M1	Visière Solidaire, Gard	Impression 3D de visières
2	Étudiant L3	Arts et Métiers (campus Aix-en-Provence)	Assemblage de visières, aide à la distribution et communication
3	Étudiant M2	Arts et Métiers (campus Metz)	Coordination logistique, développement de solutions techniques

4	Étudiant M1	Arts et Métiers (campus Châlons-en-Champagne)	Initiateur du mouvement, mise en place de guides techniques, coordination logistique, développement de solutions techniques
---	-------------	---	---

Les quatre entretiens ont été réalisés en visioconférence, et la grille d'entretien destinée à guider l'échange s'est axée autour de trois thématiques :

- Le profil des enquêtés, qui vise à comprendre le statut des enquêtés, et leur relation avec les media sociaux ou la technologie en général ;
- L'action, qui vise à la narration de l'initiative auquel l'enquêté a pris part, en identifiant son rôle au sein du réseau, et les schémas communicationnels ou logistiques qui ont été mis en place dans chaque initiative ;
- La perception de l'initiative, qui vise à une prise de recul de la part de l'enquêté sur les motivations de l'engagement, les bénéfices tirés de la mobilisation, ou encore la perception du rôle joué par les media sociaux au sein de cette action.

2. Profils des enquêtés : des étudiants technophiles

Il est important ici de noter la forte similarité entre les profils des quatre personnes rencontrées : tous sont des hommes, étudiants, entre 20 et 25 ans, qui présentent un intérêt marqué pour les nouvelles technologies, qu'elles soient numériques (media sociaux) ou physiques (impression 3D). Les enquêtés déclarent tous être actifs au quotidien sur les media sociaux (Facebook, Instagram, YouTube, LinkedIn sont parmi les plus mentionnés). L'un des enquêtés affirme même considérer d'une certaine manière son usage des media sociaux comme une addiction (id. 1). De plus, plusieurs d'entre eux occupent ou ont occupé par le passé un rôle au sein d'associations étudiantes, ou de fablabs de leur École, ce qui marque un intérêt déjà présent pour la conception et la gestion de projets en dehors de leurs études. De plus, l'un des enquêtés (id. 3) observe que parmi la soixantaine d'étudiants qui se sont mobilisés pour la production de visières sur le campus des Arts et Métiers de Metz, « c'était pas des gens qu'on avait l'habitude de voir dans l'associatif, et pas forcément non plus ceux qui sont dans les assos humanitaires, par exemple qui travaillent avec la Croix Rouge. Finalement, c'est plus des gens qui étaient assez fans de technologies, et qui pouvaient pas rester chez eux à rien faire. »

3. Des circonstances particulières, sources de motivation pour l'engagement

Il apparaît en effet que la situation particulière représentée par la crise Covid-19, à savoir le confinement généralisé à l'échelle nationale, a été propice à la mobilisation de certains enquêtés, qui affirment qu'ils ne se seraient probablement pas mobilisés dans d'autres circonstances.

D'une part, les circonstances du confinement ont conduit à une présence accrue en ligne : l'un des enquêtés, qui anime sa propre chaîne YouTube, a par exemple constaté une présence plus marquée de ses abonnés (id. 1). Cette présence active sur les media sociaux, forcée par les circonstances semble également avoir joué un rôle important dans la communication massive sur les initiatives. En effet, le medium social professionnel LinkedIn a été plébiscité par les étudiants des Arts et Métiers pour la diffusion de leurs initiatives et de leurs appels aux dons, préféré à Facebook pour l'intermédiaire qu'il apporte entre communautés citoyennes et réseaux professionnels et industriels.

D'autre part, les circonstances singulières du confinement ont permis l'engagement des personnes dans les initiatives. Les possibilités de loisirs étant restreintes, l'ennui ou la volonté de se sentir utile dans une période de grande nécessité, a pour beaucoup été un moteur de la mobilisation (id. 1 à 4). Ainsi, la participation à une initiative a simplement représenté pour certains l'opportunité de participer à une activité nouvelle, plutôt que de rester enfermé chez soi à ne rien faire : « *c'est une façon de s'évader de chez soi en y restant* » (id. 3). Dans le cas des enquêtés, leur statut d'étudiant leur permettait généralement de continuer à suivre leurs cours en ligne, mais également d'être plus facilement disponibles pour se consacrer à leur mobilisation.

Il est intéressant de noter que pour certains des enquêtés, leur participation relève simplement d'une opportunité saisie, et non d'une démarche active de recherche d'un moyen de contribuer (id. 1, id. 2). Dans tous les cas, leur engagement est généralement le fruit d'une simple volonté d'aider dans une situation où ils ont les moyens de le faire, en mettant à dispositions leurs compétences techniques ou logistiques. Dans une démarche purement altruiste, l'enquêté numéro 1 indique n'avoir cherché dans cette initiative aucune autre chose que l'opportunité de rendre un service : « *vraiment, c'était juste pour aider les gens. J'avais une seule motivation c'était de pouvoir fournir des protections à des gens qui en avaient besoin. Je cherchais vraiment rien d'autre que ça. J'étais pas forcément intéressé par découvrir des nouvelles personnes, par découvrir des makers, vraiment la seule motivation pour moi c'était de pouvoir aider.* »

Enfin, il ressort de la part des enquêtés que leur mobilisation est éphémère. Si, certes, en commençant à participer en mars, ils ne savaient pas encore pour combien de temps ils s'engageaient, il n'y avait pas pour les enquêtés de recherche d'un engagement durable au sein d'une communauté ou d'un projet. Comme nous allons le voir dans la section suivante, si certaines initiatives avaient le professionnalisme d'une entreprise, elles n'en avaient pas l'ambition : la mobilisation vient avant tout pour pallier un manque, et non pour faire un bénéfice, ou faire perdurer une structure (id. 4).

Notons tout de même que cette remarque ne s'applique pas à l'ensemble des initiatives observées pendant la crise : des communautés se sont structurées en associations, et des entreprises ont vu le jour, construites autour du besoin en ressources. Ces structures n'ont cependant pas été étudiées plus en détail.

4. Le professionnalisme des initiatives citoyennes

Au cours de cette enquête, le caractère le plus marquant qui ressort des entretiens est le professionnalisme des réseaux qui ont été mis en place, en regard du peu de ressources et de structure préalable. En effet, les récits des enquêtés mettent en avant des initiatives citoyennes organisées et professionnelles, tant d'un point de vue technique que communicationnel ou logistique.

4.1. Réseau Visière Solidaire

Comme nous l'explique l'enquêté numéro 1, si le mouvement Visière Solidaire a bénéficié de son nom générique diffusé à l'échelle nationale et relayé par la presse, il a su se décliner à échelle locale via des groupes Facebook pour chaque département mobilisé. Au sein de chaque département, et en particulier avec l'exemple du Gard, un groupe central a été mis en place pour récupérer les matières premières (filaments plastiques), et les redistribuer au sein des différents secteurs locaux dans le département. Pour chaque secteur local, une personne était chargée de

la collecte des visières à intervalle fixe, et des données de production et de stock étaient tenues à jour dans des documents partagés. Chaque soir, des statistiques de production étaient publiées sur les pages Facebook. Ainsi, l'enquêteur numéro 1 nous explique « *Vraiment il y avait une hiérarchie dans le groupe, comme on pourrait trouver dans une entreprise. Clairement, ça m'a aidé, c'était plus simple, parce que j'avais pas envie de devoir gérer des gens, pas trop le temps de m'occuper de l'administratif, moi ça me convenait de juste fabriquer et donner ma production.* » De plus, au sein de son secteur, les 6 acteurs se sont réparti la production des visières en fonction de leur capacité de production et des modèles de visières préférés par les bénéficiaires. Enfin, du point de vue communicationnel, le mouvement s'est unifié autour d'un logo, d'une identité de marque improvisée et facilement reconnaissable entre les acteurs. Comme reconnu par d'autres enquêteurs (id. 3, id. 4), si la profusion d'initiatives au cours de la crise a rendu difficile leur unification autour d'un système logistique central, Visière Solidaire est l'un des mouvements qui s'est le mieux implanté, grâce à une diffusion nationale via des groupes Facebook déclinés à l'échelle locale.

4.2. Réseau Arts et Métiers

L'action menée par les étudiants des Arts et Métiers peut se présenter en deux composantes : la mise en place d'un réseau de production de visières en impression 3D, et la coordination de bénévoles locaux pour la fabrication de masques en tissu.

Le réseau de production de visières a vu le jour à la suite d'une initiative de l'enquêteur numéro 4, qui a communiqué sur LinkedIn au sujet d'une première production personnelle. Face à un succès grandissant, une cagnotte en ligne a été ouverte et diffusée, pour récolter des dons et permettre un investissement conséquent dans des matières premières, pour supporter la mise en place d'un réseau de production. Initié sur le campus de Metz, ce projet a consisté à une production permanente de visières assurée par quatre imprimantes 3D. Tout au long de l'opération, environ 25 étudiants se sont relayés pour lancer des productions et effectuer des opérations de maintenance. Des comparatifs entre modèles de visières et paramètres techniques ont permis de concevoir les modèles de visière performants, pour assurer une cadence de production optimale. Les données techniques permettant ces résultats ont été consignées dans des documents diffusés à destination d'autres groupes de makers. Une collecte des visières par drive a été mise en place sur le campus, permettant aux bénéficiaires qui le pouvaient de récupérer leurs visières tout en respectant les recommandations de distanciation physique. En 6 semaines, 10 000 visières ont été produites par ce réseau étudiant. De plus, cette mise en place s'est appuyée sur le développement de guides de bonnes pratiques réalisés par l'enquêteur numéro 4 : une fiche technique pour permettre à un novice de lancer une production de visières efficace en très peu de temps, une fiche de bonnes pratiques sanitaires, ou encore une fiche de conseils pour mettre en place une distribution logistique. Ainsi, ces guides ont permis la diffusion des bonnes pratiques développées dans ce réseau local, pour en faire bénéficier d'autres initiatives en France.

En parallèle de ce réseau de production de visières, un groupe mené, entre autres, par les enquêteurs 3 et 4, a conduit un projet de conception et de mise au point d'outil de découpage de tissu semi-industriel, en vue d'améliorer les cadences de production de masques en tissu. En moins de deux semaines, à partir d'outillage élémentaire et de pièces fabriquées en impression 3D, le projet est né : « *pour quarante euros, on est capables de transformer une table lambda en table de découpe industrielle* » (id. 4). Cette initiative secondaire a été organisée avec le soutien

de la mairie de Metz, et a contribué à fournir à des associations de bénévoles locaux des matières premières, des procédures, des outils et des bonnes pratiques pour assurer une production de masques plus efficace et plus responsable. Le projet a été accueilli au sein d'un gymnase aménagé spécifiquement en respect des recommandations sanitaires, tout en cherchant à maintenir un climat convivial recherché par les bénévoles. L'ensemble de cette activité a été supportée financièrement par les dons récoltés sur la cagnotte en ligne mise en place par l'enquêteur numéro 4, ainsi que par des partenariats établis avec certains fournisseurs. Pour soutenir la production, des imprimantes 3D ont été dédiées à la fabrication de pièces de rechange pour les machines à coudre utilisées par les bénévoles. Tout au long de ce projet, les enquêteurs 3 et 4 expliquent leurs ambitions, les difficultés rencontrées, et détaillent certains points d'attention dans leur démarche. Conscients de l'ampleur de leur action, les enquêteurs ont veillé à assurer un système logistique contrôlé, reposant, entre autres, sur une communication mesurée et illustrée sur LinkedIn, un suivi des recommandations sanitaires, ou encore un suivi des enjeux juridiques représentés par une telle distribution.

5. Le support des media sociaux : des outils adoptés pour leurs fonctionnalités

Si ces initiatives se sont avérées faire preuve d'un grand professionnalisme sur le plan logistique et technique, elles se sont également appuyées sur la diversité d'outils numériques à disposition des citoyens. En effet, au sein de ces réseaux citoyens organisés spontanément, les media sociaux sont la toile de fond qui permettent l'émergence, la communication, et l'organisation des initiatives citoyennes.

5.1. Divers media pour divers usages

Des groupes WhatsApp et Messenger ont généralement été utilisés pour communiquer au quotidien entre les membres d'une initiative.

Facebook a été utilisé à l'échelle locale, pour la création de pages en vue de donner de la visibilité aux mouvements. En particulier dans le cadre du mouvement Visière Solidaire, l'initiative s'est diffusée massivement via Facebook, et les pages dédiées pour chaque département y ont joué un rôle de vitrine et de point de contact vers les réseaux citoyens. Ainsi, cela a permis, entre autres, le recrutement de nouveaux volontaires, la communication au public, et la récupération de commandes de visières de la part des établissements dans le besoin (id. 1). De plus, Facebook est reconnu par les enquêteurs comme étant le réseau social le plus intergénérationnel. Les pages Facebook ont ainsi permis la diffusion d'initiatives portées par des technophiles auprès de volontaires aux profils divers, qui ont pu apporter leur soutien, qu'il soit financier, communicationnel, ou encore technique.

LinkedIn, pour sa part, a été privilégié par les enquêteurs numéro 3 et 4 pour sa position stratégique entre les communautés étudiantes et les réseaux industriels. Il ressort de l'étude que la forte cohésion au sein des étudiants et des anciens élèves de l'école des Arts et Métiers est également présente sur les réseaux sociaux, et en particulier LinkedIn. Ce réseau professionnel dense, et couvrant d'une large diversité de domaines industriels, a permis la diffusion massive de l'initiative, et notamment de la cagnotte en ligne, sans laquelle les productions de visières et de masques n'auraient pas été possibles. Ainsi, « *Sur Facebook, les gens vont aider parce qu'ils connaissent la personne. Sur LinkedIn, les gens vont aider pour l'action en elle-même.* » (id. 4) De plus, l'enquêteur numéro 4 indique apprécier sur LinkedIn la possibilité de suivre l'impact des

publications, c'est-à-dire non seulement le nombre de réactions, mais également le nombre de visionnages, permettant un meilleur suivi de la communication.

L'application Discord, initialement destinée à la discussion entre joueurs de jeux vidéo, a vu son nombre de téléchargements augmenter de plus de 60% au cours de la pandémie (source : Apptopia). Pendant la crise, des serveurs Discord ont été créés pour rassembler les communautés de *makers* à l'échelle nationale et locale, et les fonctionnalités de l'outil ont été utilisées par les communautés pour s'organiser, échanger sur les bonnes pratiques de production, comparer des modèles avec des communautés à l'international, et centraliser des informations sur les normes à suivre. Ainsi, des salons de discussion sur Discord ont permis des échanges sur des points spécifiques : suivi des commandes, discussions entre membres, résolution de problèmes techniques, ou encore partage de fichiers 3D.

Au même titre que Discord, l'application Slack a été utilisée par les étudiants des Arts et Métiers pour la communication entre les différents pôles du projet, à savoir : la conception de l'outil de découpage, la production des visières, la préparation de communications, ou encore le démarchage de nouveaux partenaires ou bénéficiaires. En effet, l'outil est apprécié pour sa capacité à structurer les échanges selon des fils de discussion thématiques pour une même équipe de projet.

Des plateformes de contenu telles que YouTube ont servi à la diffusion de messages, ou l'appel à la mobilisation, grâce au soutien d'acteurs bénéficiant d'une visibilité sur la plateforme. En effet, la consommation de contenu sur YouTube étant aujourd'hui tournée autour de créateurs de contenu (chaînes), ces acteurs ont mis à profit leur notoriété pour lancer des appels à contribution, transmettre des bonnes pratiques, documenter les initiatives ou guider les potentiels volontaires vers des moyens de contribuer. Les enquêtés ont mentionné en particulier l'influence à l'échelle française de chaînes YouTube dans le domaine de la conception 3D, de la vulgarisation technique, ou du *making* en général, telles que Heliox (205k abonnés au 01/04/2020), ou Monsieur Bidouille (144k abonnés au 01/04/2020).

Enfin, d'autres outils ont été utilisés par les communautés citoyennes pour supporter leur organisation logistique : des formulaires en ligne (JotForm, Google Form) pour récupérer les commandes des bénéficiaires, des cagnottes basées sur le concept de *crowdfunding* (Lydia) pour supporter financièrement l'investissement en matières premières, des applications de suivi d'avancement de tâches (Trello) pour assurer la gestion de projet, ou encore des outils de partage et d'édition en ligne de documents et tableurs (Google Drive, Google Sheets) pour recenser les volontaires, rassembler les commandes, suivre les productions et les livraisons, faire un état des stocks, ou encore tenir des comptes.

Enfin, des moyens de communication historiques (téléphone, e-mail) ont également été employés, généralement pour effectuer une première prise de contact avec les bénéficiaires, ou pour initier certaines démarches auprès de nouveaux partenaires.

5.2. La place des media sociaux dans ces initiatives

À travers les initiatives étudiées ici, les media sociaux occupent une place centrale comme outils de communication, de coordination, d'organisation, ou encore de suivi des actions menées.

Cependant, il est intéressant de noter que ces outils n'occupent pas une place centrale au sein des récits des enquêtés, qui décrivent généralement leurs actions sans spécifier les outils qui leur ont servi, et ont permis de supporter leurs actions. Si les noms de certains media sociaux apparaissent parfois spontanément, c'est avant tout l'objectif qui est décrit, tel que « lister les commandes », « planifier la production », « suivre les stocks », ou tout simplement « communiquer ». Au cours des entretiens, ce sont souvent des questions spécifiques qui ont amené les enquêtés à détailler leur usage de chaque application numérique et les raisons de cet usage plutôt qu'un autre, et ce même si l'objet de l'étude était initialement présenté en début d'entretien, à savoir « le rôle joué par les media sociaux dans les initiatives citoyennes pendant la crise Covid-19 ». Ceci met en avant plusieurs aspects sur ces initiatives.

D'une part, cela illustre le fait qu'au-delà de l'altruisme sous-jacent à l'engagement citoyen, les enquêtés ont avant tout agi par intérêt technophile, ou par intérêt pour l'action elle-même, sa mise en place, et les défis qu'elle représente, et non par intérêt spécifique pour les media sociaux.

D'autre part, cela met en avant le fait que les media sociaux ont, certes, joué un rôle essentiel tout au long des actions mises en place, mais ils ne représentent pas une fin en soi. Ils ont servi au cours de chaque phase des projets menés, et ce de diverses manières, mais ils ont avant tout été utilisés comme outils de support des actions, et non comme terrain d'action. Ainsi, ces outils se sont imposés naturellement comme des moyens de supporter les objectifs des acteurs, mais les raisons de l'usage de telle ou telle application ne sont pas au cœur de la perception qu'ont les acteurs de leurs initiatives.

5.3. Et sans les media sociaux ?

Lorsqu'ils sont interrogés sur la possibilité qu'une telle action se déroule sans l'usage des media sociaux, les enquêtés mettent en avant les nombreux avantages que ces plateformes ont apportés : « *Sans ces outils, mettre en place un tel réseau ça aurait été compliqué. Surtout pour gérer la distance et les volumes à produire* » (id. 1). Cependant, ils n'excluent pas le fait qu'une telle action aurait pu tout de même prendre place, et qu'ainsi, les media sociaux ont avant tout joué un rôle de support à l'action physique et logistique. Même s'il a été constaté un gros manque de coordination des initiatives à l'échelle nationale (id. 4), les enquêtés sont en effet convaincus que les media sociaux ont contribué à donner des idées à d'autres personnes, diffuser des messages, prendre contact, partager des modèles techniques ou logistiques, et ainsi faire grandir les réseaux ou émerger de nouvelles initiatives (id. 1 à 4). Une comparaison intéressante est proposée par un enquêté entre le mode de fonctionnement des media sociaux, et l'idéologie portée par la communauté des *makers* : « *les makers c'est l'open source, partage des données, de l'information, c'est la base d'un réseau social [numérique], l'échange d'informations, c'est normal qu'il y ait eu un très fort écho vis-à-vis de ça.* » (id. 4).

Enfin, à l'échelle locale, certains enquêtés mettent en avant le fait que les initiatives auraient tout de même vu le jour, mais avec beaucoup moins d'uniformité. Ils estiment que les modes d'action ne se seraient pas diffusés au-delà de l'échelle municipale, et se seraient concentrés sur des groupes d'amis, ou des communautés connectées géographiquement. Dans ce cas, « *les fab*

*labs*²⁰ auraient certainement constitué des points clés de rassemblement à l'échelle locale » (id. 4), autour d'interlocuteurs physiques, et de compétences techniques identifiées.

6. Que retenir de cette étude ?

Si la cinétique, l'ampleur et la temporalité de la crise liée à la pandémie de Covid-19 en font une crise singulière, celle-ci a tout de même permis de mettre en avant certaines facettes et rôles variés joués par les media sociaux dans la gestion de crise. Massivement reconnus comme vecteurs de circulation de l'information, permettant d'acquérir une meilleure conscience de la situation, ou encore de faire sens des événements, les media sociaux s'avèrent également être des outils pertinents et particulièrement performants pour permettre aux citoyens de développer des initiatives au-delà de la toile. L'étude des initiatives citoyennes au cours de la crise Covid-19 met en avant la pluralité des facettes des media sociaux, souvent sous-estimée en cellule de crise : les media sociaux jouent le rôle de canaux de circulation de l'information, d'outils de construction de la connaissance, mais font également leurs preuves en tant que plateformes d'organisation de la réponse, permettant ainsi de supporter l'action, qu'elle soit physique ou numérique. Cette diversité d'applications des media sociaux fait écho à la diversité de profils de volontaires identifiés par Batard et al. (2019), et dont les contributions sont supportées par leurs usages des media sociaux.

Les récits des enquêtés mettent en avant le fait que les media sociaux ont joué des rôles essentiels et divers, en tant que supports d'une action qui se déroule au-delà des plateformes numériques. On observe alors ici une différence notable avec des engagements citoyens numériques tels que la documentation de situations de crise sur Wikipédia, telle que présentée précédemment dans ce rapport. En effet, si les ambitions de ces acteurs sur Wikipédia peuvent trouver leurs origines au-delà du medium social (vérification d'informations, contribution à la connaissance), l'action elle-même prend place sur le medium social lui-même, ce qui n'est pas le cas des *makers* étudiés ici.

De plus, le professionnalisme démontré par les réseaux citoyens pour mettre en place des réseaux logistiques performants et pallier le manque de ressources au cours de la crise est à souligner. Ces considérations confirment des constats réalisés par plusieurs auteurs : en dépit de l'image propagée par certains films « catastrophes », et couramment répandue, de nombreux cas montrent que la plupart des citoyens réagissent de manière constructive, rationnelle et organisée face à une situation de crise, et cherchent alors à répondre à leurs besoins via une mobilisation efficace des ressources en leur possession (Perry & Lindell, 2003; Scanlon et al., 2014).

Enfin, même si l'engagement volontaire des enquêtés est resté éphémère (tous ont mis fin à leur implication dans le projet au cours de l'été 2020), il apparaît que de tels réseaux citoyens ne disparaissent pas totalement après leur formation. En effet, au cours de cette initiative, des

²⁰ Comprendre ici « *makerspace* », qui désigne des lieux communautaires mettant à disposition du matériel technique pour concevoir, prototyper et fabriquer des objets manufacturés. Si le terme « *fab lab* » est aujourd'hui communément employé de manière générique, l'appellation est initialement réservée à un type spécifique de *makerspace* qui s'assure de suivre une charte écrite par le M.I.T.

communautés ont structuré leur action, avec l'ambition de faire perdurer ce qui a été mis en place. À titre d'exemple, le mouvement Visière solidaire a conduit à la création de plusieurs associations à l'échelle locale. Pour le Gard, l'association Make Cœur du Gard a été créée en vue de pérenniser l'action qui a été mise en place, et de proposer le soutien à l'échelle locale et la mise en place d'un réseau de *makers* similaire aux réseaux de *fab labs*. Ainsi, tous les enquêtés affirment que si leur implication a pris fin, ils laissent derrière eux une contribution à quelque chose de plus grand, participant ainsi au développement de la communauté, que Dufty (2012) présente comme un pilier de la résilience communautaire face aux crises.

III. Perspectives

La pandémie de Covid-19 a été associée dès ses débuts à l'idée d'une « infodémie », où rumeurs et fausses informations viennent combler un vide informationnel, ou sont le pendant d'une masse d'informations parfois contradictoires émanant des institutions, personnes publiques, professionnels ou simples citoyens. Nous avons montré que, comme toute crise, elle présente un caractère inattendu, et elle génère de l'incertitude et de l'anxiété (Starbird, 2020). Elle suscite depuis qu'elle a débuté une mobilisation forte des media sociaux. Les citoyens s'y réunissent, communiquent et s'organisent pour trouver du sens et combler ce vide informationnel. En particulier, des communautés citoyennes expertes émergent, s'organisent et structurent une tendance, un débat, portent une voix. À titre d'exemple, le travail de synthèse effectué par la communauté de Wikipédia lors d'un événement majeur illustre l'importance de ces communautés expertes lorsqu'il s'agit de « faire sens collectivement » (Bubendorff et al., à paraître; Bubendorff & Rizza, 2020). De leur côté, les institutions publiques composent désormais avec ces experts citoyens qui s'imposent dans l'espace public : la Fondation Wikimedia France a été auditionnée par le comité d'experts du CSA sur la désinformation en ligne pendant la crise sanitaire (Wikimedia Fondation, 2020).

Sur la base des travaux développés dans le cadre de l'ANR MACIV « Management of citizens and volunteers in crisis management : the social media contribution » (ANR-17-CE39-0015) qui étudiait la complétude des enjeux associés aux media sociaux en gestion de crise : enjeux technologique - extraction automatisée de l'information décisionnelle (Coche et al., 2020) ; institutionnel - statut de l'information issue des media sociaux (Castagnino, 2019) ; et citoyen - dimension organisationnelle des initiatives citoyennes (Batard, Rizza, et al., 2019; Bubendorff et al., 2019; Rizza, 2020) et de sa spécialisation sur la crise sanitaire dont les résultats ont été présentés dans ce working paper, les deux équipes de recherche IMT Mines Albi et I3-Télécom Paris ont déposé une demande de financement dans l'appel à projets générique de l'ANR 2021. Ainsi, le projet « gestion de l'information en situation d'urgence : media sociaux et communautés citoyennes expertes » se positionne sur cet enjeu à la fois technologique et social de la gestion de l'information qui émerge sur les media sociaux lors d'une situation d'urgence : peut-on repérer automatiquement cette information décisionnelle ? Comment faciliter sa gestion (prise en compte, intégration dans les processus opérationnels et décisionnels) par les acteurs de la gestion de crise ? Il traite des aspects de création de l'information par les communautés expertes citoyennes sur les media sociaux lors d'une situation d'urgence et de leur intégration par les gestionnaires de crise. Il formule trois hypothèses relevant à la fois des sciences informatiques et de la donnée et des sciences humaines. Il offre ainsi une approche interdisciplinaire intégrée par la co-construction des approches et l'élaboration d'outils épistémologiques : l'intégration des dimensions sociales dans le méta-modèle sur lequel repose l'IA d'extraction d'information décisionnelle ; l'enrichissement du travail d'ethnographie en ligne et de traitement du contenu publié au moyen des outils informatiques développés. Enfin, il propose un mode de « recherche-accompagnement » qui garantit la co-construction des approches entre chercheurs et praticiens, et s'intéresse à la modélisation avec les partenaires institutionnels du projet, et les gestionnaires de crise.

IV. Bibliographie

- Abel, F., Hauff, C., Houben, G.-J., Stronkman, R., & Tao, K. (2012). Semantics+ filtering+ search= twitcident. Exploring information in social web streams. *Proceedings of the 23rd ACM conference on Hypertext and social media*, 285–294.
- Alexander, D. E. (2014). Social Media in Disaster Risk Reduction and Crisis Management. *Science and Engineering Ethics*, 20(3), 717-733. <https://doi.org/10.1007/s11948-013-9502-z>
- Alloing, C., & Vanderbiest, N. (2018). La fabrique des rumeurs numériques. Comment la fausse information circule sur Twitter? *Le Temps des médias*, 1(30), 105-123.
- Ashktorab, Z., Brown, C., Nandi, M., & Culotta, A. (2014). Tweedr : Mining twitter to inform disaster response. *ISCRAM*, 269–272.
- Auray, N., Hurault-Plantet, M., Poudat, C., & Jacquemin, B. (2009). La négociation des points de vue, une cartographie sociale des conflits et des querelles dans le Wikipédia francophone. *Réseaux*, 154, 15-50. <https://doi.org/10.3917/res.154.0015>
- Austin, L., Liu, B. F., & Jin, Y. (2012). How Audiences Seek Out Crisis Information : Exploring the Social-Mediated Crisis Communication Model. *Journal of Applied Communication Research*, 40(2), 188-207. <https://doi.org/10.1080/00909882.2012.654498>
- Banda, J. M., Tekumalla, R., Wang, G., Yu, J., Liu, T., Ding, Y., Artemova, K., Tutubalina, E., & Chowell, G. (2020). *A large-scale COVID-19 Twitter chatter dataset for open scientific research—An international collaboration* [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.4287527>
- Batard, R., Montarnal, A., Benaben, F., Rizza, C., Prieur, C., & Tapia, A. (2019, mai 19). Integrating citizen initiatives in a technological platform for collaborative crisis

management. *Proceedings of the 16th International Conference on Information Systems for Crisis Response And Management*. ISCRAM 2019 - 16th International Conference on Information Systems for Crisis Response and Management, València, Spain. <https://hal.telecom-paristech.fr/hal-02173972>

Batard, R., Rizza, C., Montarnal, A., & Benaben, F. (2018). Ethical, legal and social considerations surrounding the use of Facebook groups during Hurricane Irma in Cuba. *ISCRAM 2018-15th International conference on Information Systems for Crisis Response and Management*, p-268.

Batard, R., Rizza, C., Montarnal, A., Benaben, F., & Prieur, C. (2019). Review-based taxonomy of post-impact volunteerism types to improve citizen integration into crisis response. *Proceedings of the 16th International Conference on Information Systems for Crisis Response And Management*, 1114-1125.

Benaben, F., Fertier, A., Montarnal, A., Mu, W., Jiang, Z., Truptil, S., Barthe-Delanoë, A.-M., Lauras, M., Mace-Ramete, G., Wang, T., Bidoux, L., & Lamothe, J. (2020). An AI framework and a metamodel for collaborative situations: Application to crisis management contexts. *Journal of Contingencies and Crisis Management*, 28(3), 291-306. <https://doi.org/10.1111/1468-5973.12310>

Boersma, K., Clark, N., Kyratsis, Y., Rollo, A., Falagara, S., Alani, H., de Vries M, Robert, L., & Berg, R. (2020). *Recommendations for governance and policies in the n-COV-2019 response (D1.1)*. https://www.heros-project.eu/wp-content/uploads/D1_1_Recommendations_for_governance_and_policies_in_the_COVID_19_response_v2_0.pdf

Bubendorff, S., & Rizza, C. (2020, avril 28). Wikipédia au temps de la crise covid-19. *I'MTech*. <https://blogrecherche.wp.imt.fr/2020/04/28/wikipedia-crise-covid-19/>

- Bubendorff, S., Rizza, C., & Prieur, C. (à paraître). Construction and dissemination of information veracity on French social media during crises : Comparison of Twitter and Wikipedia. *Journal of Contingencies and Crisis Management*.
- Bubendorff, S., Rizza, C., & Prieur, C. (2019). *Réseaux sociaux numériques et spatio-temporalité de l'information en gestion de crise* (Working Paper N° 19-SES-01; I3 Working Paper). Institut Interdisciplinaire de l'Innovation. <https://hal.telecom-paristech.fr/hal-02175338>
- Cameron, M. A., Power, R., Robinson, B., & Yin, J. (2012). Emergency situation awareness from twitter for crisis management. *Proceedings of the 21st International Conference on World Wide Web*, 695–698.
- Camozzi, M.-L. (2020). *Méta-modélisation d'une situation de crise type COVID et évaluation d'un modèle prédictif de classification des données issues des réseaux sociaux*. IMT Mines Albi; Rapport de stage Master 1.
- Cardon, D., & Levrel, J. (2009). La vigilance participative. Une interprétation de la gouvernance de Wikipédia. *Réseaux*, n° 154(2), 51-89.
- Castagnino, F. (2019). What can we learn from a crisis management exercise ? Trusting social media in a french firefighters' department. In Z. Franco, J. J. González, & J. H. Canós (Éds.), *Proceedings of the 16th International Conference on Information Systems for Crisis Response And Management*. Iscram. http://idl.iscram.org/files/florentcastagnino/2019/1967_FlorentCastagnino2019.pdf
- Chan, J. C. (2014). The role of social media in crisis preparedness, response and recovery. *citeseer*.
- Chen, Q., Min, C., Zhang, W., Wang, G., Ma, X., & Evans, R. (2020). Unpacking the black box : How to promote citizen engagement through government social media during the COVID-19 crisis. *Computers in Human Behavior*, 106380.

- Coche, J., Montarnal, A., Tapia, A., & Benaben, F. (2019). Actionable Collaborative Common Operational Picture in Crisis Situation : A Comprehensive Architecture Powered with Social Media Data. In L. M. Camarinha-Matos, H. Afsarmanesh, & D. Antonelli (Éds.), *Collaborative Networks and Digital Transformation* (p. 151-162). Springer International Publishing. https://doi.org/10.1007/978-3-030-28464-0_14
- Coche, J., Montarnal, A., Tapia, A., & Benaben, F. (2020). Automatic Information Retrieval from Tweets : A Semantic Clustering Approach. *ISCRAM 2020 - 17th International conference on Information Systems for Crisis Response and Management*, p.134-141. <https://hal.telecom-paris.fr/hal-02926851>
- Cortez, R. M., & Johnston, W. J. (2020). The Coronavirus crisis in B2B settings : Crisis uniqueness and managerial implications based on social exchange theory. *Industrial Marketing Management*, 88, 125–135.
- Dauvin, M., Ducoudré, B., Heyer, É., Madec, P., Plane, M., Sampognaro, R., & Timbeau, X. (2020). Évaluation au 26 juin 2020 de l’impact économique de la pandémie de COVID-19 et des mesures du confinement et du déconfinement en France. *Revue de l’OFCE*, 166, 111.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert : Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Diemer, A. (2020). Modéliser le COVID 19 Défis et perspectives. *Revue Francophone du Développement Durable*, 15, 72.
- Dufty, N. (2012). Using social media to build community disaster resilience. *The Australian Journal of Emergency Management (peer reviewed)*, 27(1). https://works.bepress.com/neil_dufty/8/

- Eismann, K., Possega, O., & Fischbach, K. (2016). *Collective Behaviour, social media, and disasters : A systematic literature review*. Twenty-Fourth European Conference on Information Systems (ECIS), Istanbul, Turkey.
- Fernandes, N. (2020). Economic effects of coronavirus outbreak (COVID-19) on the world economy. *Available at SSRN 3557504*.
- Ford, H., Sen, S., Musicant, D. R., & Miller, N. (2013). Getting to the source : Where does Wikipedia get its information from? *Proceedings of the 9th International Symposium on Open Collaboration*, 1–10. <https://doi.org/10.1145/2491055.2491064>
- French Twitter Sentiment Analysis*. (s. d.). Consulté 25 novembre 2020, à l'adresse <https://kaggle.com/hbaflast/french-twitter-sentiment-analysis>
- Habermas, J. (1987). *Théorie de l'agir communicationnel. Tome 2, Critique de la raison fonctionnaliste*. Fayard.
- Habermas, J. (1995). *Justification and application : Remarks on discourse ethics*. Polity Press, 1995.
- Heverin, T., & Zach, L. (2012). Use of microblogging for collective sense-making during violent crises : A study of three campus shootings. *Journal of the American Society for Information Science and Technology*, 63(1), 34-47. <https://doi.org/10.1002/asi.21685>
- Imran, M., Castillo, C., Diaz, F., & Vieweg, S. (2018). Processing Social Media Messages in Mass Emergency : Survey Summary. *Companion of the The Web Conference 2018 on The Web Conference 2018 - WWW '18*, 507-511. <https://doi.org/10.1145/3184558.3186242>
- Imran, M., Castillo, C., Lucas, J., Meier, P., & Vieweg, S. (2014). AIDR : Artificial intelligence for disaster response. *Proceedings of the 23rd International Conference on World Wide Web*, 159–162.

- Jadhav, A. S., Purohit, H., Kapanipathi, P., Anantharam, P., Ranabahu, A. H., Nguyen, V., Mendes, P. N., Smith, A. G., Cooney, M., & Sheth, A. P. (2010). *Twitris 2.0 : Semantically empowered system for understanding perceptions from social data*.
- Jullien, N. (2012). What We Know About Wikipedia : A Review of the Literature Analyzing the Project(s). *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2053597>
- Jurgens, M., & Helsloot, I. (2018). The effect of social media on the dynamics of (self) resilience during disasters : A literature review. *Journal of Contingencies and Crisis Management*, 26(1), 79-88. <https://doi.org/10.1111/1468-5973.12212>
- Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite ! The challenges and opportunities of Social Media. *Business Horizons*, 53(1), 59-68. <https://doi.org/10.1016/j.bushor.2009.09.003>
- Keegan, B. C. (2015). Emergent Social Roles in Wikipedia's Breaking News Collaborations. In E. Bertino & S. A. Matei (Éds.), *Roles, Trust, and Reputation in Social Media Knowledge Markets : Theory and Methods* (p. 57-79). Springer International Publishing. https://doi.org/10.1007/978-3-319-05467-4_4
- Keegan, B. C., Gergle, D., & Contractor, N. (2013). Hot off the wiki : Structure and dynamics of wikipedia's coverage of breaking news events. *American Behavioral Scientist*, 57(5), 595-622.
- Kropczynski, J., Grace, R., Coche, J., Jalse, S., Obeysekare, E., Montarnal, A., Benaben, F., & Tapia, A. (2018). Identifying actionable information on social media for emergency dispatch. *Proceedings of the ISCRAM Asia Pacific*.

- Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a Social Network or a News Media? *Proceedings of the 19th International Conference on World Wide Web*, 591–600. <https://doi.org/10.1145/1772690.1772751>
- Le, H. T., Sam, R. C., Nguyen, H. C., & Nguyen, T. T. (2013). Named entity recognition in Vietnamese text using label propagation. *2013 International Conference on Soft Computing and Pattern Recognition (SoCPaR)*, 366–370.
- Leskovec, J., Backstrom, L., & Kleinberg, J. (2009). Meme-tracking and the dynamics of the news cycle. *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '09*, 497. <https://doi.org/10.1145/1557019.1557077>
- Li, C., Weng, J., He, Q., Yao, Y., Datta, A., Sun, A., & Lee, B.-S. (2012). Twiner : Named entity recognition in targeted twitter stream. *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*, 721–730.
- Luna, S., & Pennock, M. J. (2018). Social media applications and emergency management : A literature review and research agenda. *International Journal of Disaster Risk Reduction*, 28, 565-577. <https://doi.org/10.1016/j.ijdrr.2018.01.006>
- Maurel, D. (2020). Sense-making : Un modèle de construction de la réalité et d’appréhension de l’information par les individus et les groupes. *Études de communication*, 35, 31-46. <https://doi.org/10.4000/edc.2306>
- Mendoza, M., Poblete, B., & Castillo, C. (2010). Twitter Under Crisis : Can We Trust What We RT? *Proceedings of the First Workshop on Social Media Analytics*, 71–79. <https://doi.org/10.1145/1964858.1964869>
- Palen, L., Anderson, J., Bica, M., Castillos, C., Crowley, J., Díaz, P., Finn, M., Grace, R., Hughes, A., Imran, M., Kogan, M., Lalone, N., Mitra, P., Norris, W., Pine, K., Purohit, H., Reuter,

- C., Rizza, C., Denis, L. S., ... Wilson, T. (2020). *Crisis Informatics : Human-Centered Research on Tech & Crises*. <https://hal.archives-ouvertes.fr/hal-02781763>
- Palen, L., & Liu, S. B. (2007). Citizen communications in crisis : Anticipating a future of ICT-supported public participation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '07*, 727-736. <https://doi.org/10.1145/1240624.1240736>
- Perry, R. W., & Lindell, M. K. (2003). Understanding Citizen Response to Disasters with Implications for Terrorism. *Journal of Contingencies and Crisis Management*, 11(2), 49-60. <https://doi.org/10.1111/1468-5973.1102001>
- Qazi, U., Imran, M., & Ofli, F. (2020). GeoCoV19 : A Dataset of Hundreds of Millions of Multilingual COVID-19 Tweets with Location Information. *arXiv:2005.11177 [cs]*. <http://arxiv.org/abs/2005.11177>
- Rebillard, F. (2006). L'information journalistique sur l'internet, entre diffusion mass-médiatique et circulation réticulaire de l'actualité. *Document numérique et société*.
- Reuter, C., & Kaufhold, M.-A. (2018). Fifteen years of social media in emergencies : A retrospective review and future directions for crisis Informatics. *Journal of Contingencies and Crisis Management*, 26(1), 41-57. <https://doi.org/10.1111/1468-5973.12196>
- Reuter, C., Kaufhold, M.-A., Spahr, F., Spielhofer, T., & Hahne, A. S. (2020). Emergency service staff and social media – A comparative empirical study of the attitude by emergency services staff in Europe in 2014 and 2017. *International Journal of Disaster Risk Reduction*, 46, 101516. <https://doi.org/10.1016/j.ijdrr.2020.101516>
- Reuter, C., Marx, A., & Pipek, V. (2012). Crisis Management 2.0 : Towards a Systematization of Social Software Use in Crisis Situations. *International Journal of Information Systems for*

Crisis Response and Management (IJISCRAM), 4(1), 1-16.
<https://doi.org/10.4018/jiscrm.2012010101>

Rieder, B., & Smyrnaio, N. (2012). Pluralisme et infomédiation sociale de l'actualité : Le cas de Twitter, Pluralism and the Social Infomediaion of News: the case of Twitter. *Réseaux*, 176, 105-139. <https://doi.org/10.3917/res.176.0105>

Ritter, A., Clark, S., & Etzioni, O. (2011). Named entity recognition in tweets : An experimental study. *Proceedings of the 2011 conference on empirical methods in natural language processing*, 1524–1534.

Rizza, C. (2020, juillet 20). *Gestion de crise : Mieux intégrer la réponse des citoyens*. The Conversation. <http://theconversation.com/gestion-de-crise-mieux-integrer-la-reponse-des-citoyens-141741>

Sahut, G. (2014). « Citez vos sources » : Archéologie d'une règle au cœur du savoir wikipédien (2002-2008). *Études de communication. langages, information, médiations*, 42, 97-110. <https://doi.org/10.4000/edc.5721>

Sahut, G. (2016). L'autorité importée dans Wikipédia : La question de la qualité des sources citées. *Quaderni*, 91, 93-105. <https://doi.org/10.4000/quaderni.1015>

Scanlon, J., Helsloot, I., & Groenendaal, J. (2014). Putting It All Together : Integrating Ordinary People Into Emergency Response. *International Journal of Mass Emergencies and Disasters*, 32(1), 43-63.

Singer, P., Lemmerich, F., West, R., Zia, L., Wulczyn, E., Strohmaier, M., & Leskovec, J. (2017). Why We Read Wikipedia. *Proceedings of the 26th International Conference on World Wide Web - WWW '17*, 1591-1600. <https://doi.org/10.1145/3038912.3052716>

- Spielhofer, T., Hahne, A. S., Reuter, C., Kaufhold, M.-A., & Schmid, S. (2019). Social media use in emergencies of citizens in the united Kingdom. *Proceedings of the 16th ISCRAM Conference*.
- Starbird, K. (2017). Examining the Alternative Media Ecosystem Through the Production of Alternative Narratives of Mass Shooting Events on Twitter. *Eleventh International AAAI Conference on Web and Social Media*. Eleventh International AAAI Conference on Web and Social Media. <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15603>
- Starbird, K. (2020, mars 19). *How a Crisis Researcher Makes Sense of Covid-19 Misinformation*. Medium. <https://onezero.medium.com/reflecting-on-the-covid-19-infodemic-as-a-crisis-informatics-researcher-ce0656fa4d0a>
- Starbird, K., Arif, A., & Wilson, T. (2019). Disinformation as Collaborative Work : Surfacing the Participatory Nature of Strategic Information Operations. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 127:1–127:26. <https://doi.org/10.1145/3359229>
- Starbird, K., & Palen, L. (2010). Pass it on ? : Retweeting in mass emergency. *Proceedings of the 7th International ISCRAM Conference*.
- Starbird, K., Spiro, E. S., & Koltai, K. (2020, juin 25). *Misinformation, Crisis, and Public Health—Reviewing the Literature, VVI.0*. MediaWell, Social Science Research Council. <http://doi.org/10.35650/MD.2063.d.2020>
- Tandoc, E. C., Lim, Z. W., & Ling, R. (2018). Defining “Fake News” : A typology of scholarly definitions. *Digital Journalism*, 6(2), 137-153. <https://doi.org/10.1080/21670811.2017.1360143>

- Terpstra, T., Stronkman, R., de Vries, A., & Paradies, G. L. (2012). Towards a realtime Twitter analysis during crises for operational crisis management. *ISCRAM*.
- Thiault, F. (2015). Le produsage des hashstags sur Twitter, une pratique affiliative. *Questions de communication*, 28, 65-79. <https://doi.org/10.4000/questionsdecommunication.10043>
- Thomson, R., Ito, N., Suda, H., Lin, F., Liu, Y., Hayasaka, R., Isochi, R., & Wang, Z. (2012). Trusting Tweets : The Fukushima disaster and information source credibility on twitter. *Proceedings of the 9th international ISCRAM Conference*.
- Thubert, N. (2020). *Application d'outils de Machine Learning pour extraire des informations relatives au COVID-19 dans les médias sociaux*. IMT Mines Albi; Rapport de stage Master 1.
- Vieweg, S., Hughes, A. L., Starbird, K., & Palen, L. (2010). Microblogging during two natural hazards events : What twitter may contribute to situational awareness. *Proceedings of the SIGCHI conference on human factors in computing systems*, 1079–1088.
- Waller, V. (2011). The search queries that took Australian Internet users to Wikipedia. *Information Research*, 2(16). <http://www.informationr.net/ir/16-2/paper476.html>
- Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the Process of Sensemaking. *Organization Science*, 16(4), 409-421. <https://doi.org/10.1287/orsc.1050.0133>
- Wendling, C., Radish, J., & Jacobzone, S. (2013). *The Use of Social Media in Risk and Crisis Communication* (Working Paper N° 24; OECD Working Papers on Public Governance). OECD Publishing. <https://doi.org/10.1787/5k3v01fskp9s-en>
- Wikimedia Foundation. (2020). *Covid-19 & désinformation* [Audition par le comité d'experts sur la désinformation en ligne du CSA].

- Yates, C., & Partridge, H. (2015). Citizens and social media in times of natural disaster : Exploring information experience. *Information research*, 20(1).
- Yates, D., Wagner, C., & Majchrzak, A. (2010). Factors affecting shapers of organizational wikis. *Journal of the American Society for Information Science and Technology*, 61(3), 543-554. <https://doi.org/10.1002/asi.21266>
- Zade, H., Shah, K., Rangarajan, V., Kshirsagar, P., Imran, M., & Starbird, K. (2018). From situational awareness to actionability : Towards improving the utility of social media data for crisis response. *Proceedings of the ACM on human-computer interaction*, 2(CSCW), 1–18.
- Zhang, J., Wu, W., Zhao, X., & Zhang, W. (2020). Recommended psychological crisis intervention response to the 2019 novel coronavirus pneumonia outbreak in China : A model of West China Hospital. *Precision Clinical Medicine*, 3(1), 3–8.

VII. Liste des tableaux

Tableau 1 : Catégorisation des 500 # les plus présents dans le data-set.	33
Tableau 2 : Catégorisation des 200 utilisateurs les plus présents dans le data-set et volume de messages.....	35
Tableau 3 : Liste des pages Wikipédia observées	36
Tableau 4 : Sociographie des interviewés	37
Tableau 5 : Sociographie des enquêtés - réseaux de makers pendant la crise Covid-19	45

VIII. Table des figures

Figure 1 : Échantillon résultant de la propagation de labels après réduction de dimensions par T-SNE, proposé par (Coche et al., 2020).....	11
Figure 2 : Diagramme de classe UML du méta-modèle de la crise de Covid-19 (Camozzi, 2020).	14
Figure 3 : Confrontation de plusieurs classificateurs de sentiments, entraînés et testés sur le data-set French Twitter Sentiment Analysis, résultats issus de (Thubert, 2020).	17
Figure 4 : Sentiment moyen quotidien calculé par LinearSVC sur le data-set d'étude dédié à la crise de Covid-19.....	17
Figure 5 : capture d'écran du serveur Label Studio, pour générer le set d'entraînement de manière collaborative.....	18
Figure 6 : Répartition des tokens labélisés dans le data-set d'entraînement.....	18
Figure 7 : Étapes de la propagation de labels (Thubert, 2020).....	19
Figure 8 : Exemples de labels propagés par l'algorithme. À gauche un token du data-set d'étude et à droite la classe qui lui a été associée.	19
Figure 9 : Visualisation des tokens après propagation de labels dans le data-set d'étude.	20
Figure 10 : Occurrence des vingt hashtags (#) les plus présents dans le set de données.....	31
Figure 13 : Carte de la présence du virus sur le territoire, page Pandémie de Covid-19 en France	39
Figure 14 : Capture d'écran, onglet discussion.....	41
Figure 15 : Capture d'écran, onglet discussion : page personnelle d'un contributeur	42

IX. Tableau des livrables

Titre du livrable	Type de livrable	Interne/Externe	Contribution
<i>Les media sociaux lors de la crise sanitaire Covid-19, Circulation de l'information et initiatives citoyennes</i>	Working paper et rapport final de projet	Externe	Ce working paper fait état de l'ensemble de l'étude réalisée dans le cadre du projet MESCOV, selon ses deux volets technologiques et sociologiques. Les contributions sont multiples et montrent l'importance de mener ce type d'étude <i>réellement</i> interdisciplinaire, font état des avancées en réponse aux enjeux informationnels d'une crise telle que celle de Covid-19, et dessinent des perspectives d'amélioration d'ores et déjà en cours.
<i>Set de tweets francophones relatifs à la crise de Covid-19 à des fins de recherche</i>	Data-set (fichier CSV)	Externe, sous licence Creative Commons	Le data-set comprend 2.950.156 tweets en français, dont l'acquisition est détaillée dans la Section 3 de la première partie de ce rapport.
<i>Rapport de stage de Marie-Lys Camozzi</i>	Rapport de stage	Interne	Référence (Camozzi, 2020)
<i>Rapport de stage de Nathan Thubert</i>	Rapport de stage	Interne	Référence (Thubert, 2020)
<i>Faire collectivement sens en temps de crise : l'utilisation de Wikipédia lors de la pandémie de Covid-19</i> (Bubendorff, Rizza)	Article de revue	Externe	Cet article interroge la mobilisation des contributeurs de Wikipédia pendant la pandémie. Il montre comment la construction de sens suppose une réorganisation différente des manières de faire en fonction des stades successifs de la propagation du virus et analyse les tensions qui traversent ces manières de faire. Ces analyses permettent de voir de quelles manières le climat d'incertitude révèle des débats sous-jacents dans l'encyclopédie en regard du travail de référencement. La section 2 de la seconde partie s'appuie largement sur cet article.