



UNIVERSITÉ DE LILLE

FACULTÉ DE MÉDECINE HENRI WAREMBOURG

Année : 2023

THÈSE POUR LE DIPLÔME D'ÉTAT
DE DOCTEUR EN MÉDECINE

Scandale ORPEA : analyse des réactions sur Twitter

Présentée et soutenue publiquement le 12 octobre 2023 à 17h30

au Pôle Recherche

par Laetitia VERDIER

JURY

Président :

Monsieur le Professeur Éric BOULANGER

Assesseurs :

Monsieur le Professeur Jean-Baptiste BEUSCART

Monsieur le Docteur Antoine LAMER

Madame le Docteur Clémentine ROLLAND

Directeur de thèse :

Monsieur le Professeur Jean-Baptiste BEUSCART

Travail du Laboratoire ULR 2694 - METRICS

AVERTISSEMENT

La Faculté n'entend donner aucune approbation aux opinions émises dans les thèses : celles-ci sont propres à leurs auteurs.

Liste des abréviations

EHPAD	Établissement Hospitalier pour Personnes Âgées Dépendantes
NLP	Natural Language Processing
LDA	Latent Dirichlet Allocation
IGAS	Inspection Générale des Affaires Sociales
IGF	Inspection Générale des Finances

TABLE DES MATIERES

RESUME	1
INTRODUCTION	3
MATERIEL ET METHODES	6
1. Design de l'étude	6
2. Critères d'inclusion et d'exclusion des tweets dans l'étude	6
3. Extraction des données	7
4. Analyse des données	8
a. Fréquence des tweets	8
b. Nettoyage des données	8
c. Topic modeling	8
d. Catégories des comptes	9
e. Focus sur le compte ORPEA	10
RESULTATS	11
1. Corpus de tweets	11
2. Six clusters identifiés	13
a. Fonctionnement du système des EHPAD	15
b. Impact du scandale ORPEA sur le groupe Korian	16
c. Dénonciation de maltraitance des personnes âgées institutionnalisées suite au scandale	17
d. Médiatisation du scandale ORPEA	18
e. Influence politique et volonté d'action	18
f. Avis des proches et soignants, opinions personnelles sur un sujet sociétal	19
3. Profils des comptes s'exprimant sur le scandale	20
4. Focus sur le compte @ORPEA_	22
DISCUSSION	23
1. Résumé des principaux résultats	23
2. Conséquences du scandale	24
3. Forces et limites de l'étude	26
CONCLUSION	28
BIBLIOGRAPHIE	30
ANNEXES	35

RESUME

Contexte : Face au vieillissement de la population, de plus en plus de personnes âgées vivent institutionnalisées en EHPAD. Certaines familles, ayant à cœur d'offrir une bonne qualité de vie à leurs parents, se tournent vers des entreprises privées, dont ORPEA représente le leader mondial. En janvier 2022, celle-ci s'est retrouvée au cœur d'un scandale après la publication d'un livre qui dénonce les maltraitances subies par les personnes âgées au sein de l'entreprise. Le scandale a été largement suivi sur les réseaux sociaux, dont Twitter.

Objectif : Analyser les réactions de la population via Twitter en lien avec le scandale ORPEA.

Méthodes : Une méthode de *web scraping* a été appliquée pour extraire tous les tweets contenant au moins un mot-clé relatif au scandale. Ensuite, une analyse par *topic modeling* a permis d'identifier les principaux thèmes. En complément, une analyse des comptes tweetant sur le scandale a été menée, avec une attention spéciale portée sur le compte ORPEA.

Résultats : Parmi les 85 342 tweets uniques extraits, on retrouve un pic de publications à la sortie du livre, puis lors de l'ouverture d'une enquête par les autorités françaises, ainsi qu'au moment de la diffusion de l'émission *Cash Investigation* sur les EHPAD du groupe privé Korian. Les six principaux thèmes abordés portaient sur (1) le fonctionnement des EHPAD, (2) l'impact du scandale sur le groupe Korian, (3) la dénonciation des abus sur les personnes âgées, (4) la médiatisation du scandale, (5) la volonté d'une action politique, (6) les opinions et témoignages individuels. Les profils s'exprimant sur le sujet étaient surtout des comptes non vérifiés, ou des médias. Le compte ORPEA a présenté une longue période d'inactivité pendant le scandale, puis revient avec la volonté de fédérer résidents et employés, et de transformer l'image du groupe.

Discussion : Le scandale ORPEA, en faisant réagir la population sur le plan politique, économique et humain, s'inscrit dans un débat sociétal large. Les réseaux sociaux ont permis la médiatisation du scandale et de ses conséquences sur le groupe, ainsi que le développement d'organisations destinées aux aidants. Cependant, plus d'un an et demi après le scandale, peu de changements semblent avoir été opérés dans la gestion de la personne âgée institutionnalisée en France.

INTRODUCTION

La dépendance est définie, selon l'Académie de Médecine, comme le « besoin d'aide dans les actes essentiels de la vie quotidienne » (1). En France et dans le monde, le nombre de personnes âgées dépendantes ne cesse de croître. Ceci est dû à l'association de plusieurs facteurs : d'une part, une augmentation de l'espérance de vie, aujourd'hui estimée à la naissance à 85,2 ans chez la femme et 79,3 ans chez l'homme, contre respectivement 82,8 et 75,2 ans en 2000 (2). D'autre part, une augmentation de la proportion de personnes âgées, avec un étalement du sommet de la pyramide des âges, en lien avec le *papy-boom*, c'est-à-dire l'avènement dans le Grand Âge des générations nées après-guerre (3). Les personnes âgées de plus de 65 ans représenteraient, au 1^{er} janvier 2020, 20,5% de la population française selon l'INSEE et pourraient atteindre 28,7% d'ici à 2070 (3). Les projections de l'INSEE supposent alors un doublement du nombre de personnes âgées dépendantes d'ici à 2060 (de 1,15 million en 2010 à 2,3 millions en 2060) (4).

Du fait de difficultés de maintien à domicile en lien avec cette dépendance, une institutionnalisation en Établissement Hospitalier pour Personnes Âgées Dépendantes (EHPAD) doit parfois être envisagée. On dénombre actuellement 605 000 résidents en EHPAD en France (5). La qualité de vie en institution est un sujet de préoccupation important pour les familles, c'est pourquoi celles-ci se tournent régulièrement vers des établissements privés, qui représentent aujourd'hui la moitié des EHPAD en France, dont 24% sont à but lucratif (6).

L'entreprise ORPEA, leader mondial des EHPAD privés à but lucratif, avec 983 établissements dans le monde dans plus de vingt pays, et un chiffre d'affaires annuel estimé à 4,299 milliards d'euros en 2021 (7), s'est retrouvée au cœur d'un véritable scandale à la suite de la publication, le 26 janvier 2022, du livre *Les Fossoyeurs* (8) par Victor Castanet,

journaliste indépendant ; un ouvrage qui dénonce les maltraitances physiques et institutionnelles subies par les résidents du groupe ORPEA. Ce phénomène médiatique, dénommé le « scandale ORPEA », a été responsable de vives réactions à la fois individuelles et de la part des pouvoirs publics, avec notamment la réalisation d'enquêtes vérifiant les dires de l'auteur concernant la gestion des EHPAD privés et les conditions de prise en charge des personnes âgées institutionnalisées en France (9,10). Le groupe Korian, numéro un français des EHPAD (619 établissements en France, soit plus de 32 000 lits (11)), a par la suite été accusé lors de la diffusion le 1^{er} mars 2022 sur France 2 de l'émission télévisée *Cash Investigation, EHPAD : l'heure des comptes ?* (12).

A une époque de plus en plus connectée, le scandale a été largement suivi et relayé sur les réseaux sociaux, et notamment Twitter.

Ces dernières années, de plus en plus d'études utilisent l'analyse des réseaux sociaux dans le domaine médical. Par l'analyse du contenu des publications, elles ont permis par exemple d'étudier des comportements à risque, comme la consommation de tabac ou de cigarette électronique (13,14). D'autres études ont pu servir à mettre en évidence les principaux sujets abordés sur les réseaux chez les patients atteints de cancer (15,16), ou encore à comprendre les hésitations à la réalisation de vaccins, notamment contre le coronavirus (17), ou le vaccin anti-HPV pour la prévention du cancer du col de l'utérus (18). Ce type d'études a aussi permis la surveillance d'épidémies comme le virus Zika, en identifiant les symptômes développés et décrits dans les tweets (19), ou de s'intéresser aux contenus relatifs à la santé mentale lors de la pandémie à COVID-19 et les périodes de confinement (20–22). Ces méthodes ont encore été utilisées pour analyser l'opinion de la population, par exemple sur le développement de la télémédecine (23). Une utilisation à but pédagogique a aussi été envisagée pour évaluer les connaissances de la population médicale sur un sujet donné (24,25). L'utilisation des réseaux sociaux comme outils

d'entraide au sein de la communauté médicale a aussi fait l'objet d'études, avec l'utilisation de hashtags comme le *#doctoc* (26) ; ou encore comme outil de partage d'expériences (27). Pour le grand public, leur utilité comme vecteurs d'informations de santé publique, telles que la prise en charge de l'arrêt cardio-respiratoire, a été étudiée (28).

Un grand nombre de réactions a été publié sur Twitter à la suite du scandale ORPEA, ce qui donne une occasion unique d'explorer le ressenti de la population vis-à-vis de ce phénomène, et plus largement de la prise en soins en EHPAD. Les techniques modernes de *web scraping* offrent la possibilité de mener de telles analyses (29,30).

L'objectif de cette étude est donc de décrire les réactions d'une partie de la population concernant le scandale ORPEA, via les techniques d'extraction et d'analyse automatisées du langage appliquées sur Twitter, afin de comprendre les dynamiques de publications, les profils réagissant au scandale et les principales thématiques abordées.

MATERIEL ET METHODES

1. Design de l'étude

Il s'agit d'une étude observationnelle mixte (qualitative et quantitative), rétrospective, réalisée entre les mois de mai 2022 et de janvier 2023, avec l'appui de l'unité de recherche ULR 2694 – METRICS (CHU Lille, Université de Lille). Cette étude porte sur l'ensemble des tweets publiés sur le réseau social Twitter par des comptes utilisateurs réglés en « public », au sujet du scandale ORPEA.

L'étude n'a pas nécessité d'autorisation préalable de la CNIL car elle ne portait que sur les comptes Twitter réglés en mode public.

2. Critères d'inclusion et d'exclusion des tweets dans l'étude

Une recherche préliminaire de tweets concernant le scandale a été conduite sur Twitter, en soumettant des termes tels que « ORPEA », « #ORPEA », « scandale ORPEA » ou encore « EHPAD » au moteur de recherche Twitter. La lecture de ces tweets a permis d'identifier manuellement 8 mots-clés, se référant au champ lexical du scandale : ORPEA, #orpeagate, LesFossoyeurs, Victor Castanet, Laurent Garcia (ancien cadre infirmier chez ORPEA, considéré comme l'un des lanceurs d'alerte du scandale), Cash Investigation, associé au terme EHPAD afin de cibler l'émission dédiée (12), et Korian.

L'étude a ensuite été réalisée de manière automatisée avec les critères suivants :

Critères d'inclusion :

- Tweet publié sur le réseau social Twitter, par des comptes réglés en mode « public »
- Tweet contenant au moins un des 8 mots-clés identifiés
- Tweet publié entre le 1^{er} juin 2021 et le 1^{er} juin 2022

Critères d'exclusion :

- Tweet en langue étrangère

3. Extraction des données

L'extraction des tweets a été réalisée en juillet 2022, grâce à une technique de *web scraping*. C'est une méthode qui consiste à extraire les données d'un site internet via un programme informatique, afin d'en analyser le contenu. L'interface API (*Application Programming Interface*) pour Twitter a été utilisée pour extraire les données, en ayant recours à la bibliothèque Tweepy de Python.

Les données extraites des tweets contenaient les informations suivantes :

- Le contenu du tweet,
- La date de publication,
- Le nombre de « likes » attribué au tweet,
- Le nombre de « retweets »,
- Les « hashtags » (#...) contenus dans le tweet,
- Les mentions (@...) contenues dans le tweet.

Des informations sur les auteurs des tweets ont aussi été extraites :

- Le pseudonyme du compte Twitter,
- La description du profil,
- La date de création du profil,
- La ville associée au profil,
- Le nombre d'abonnés au compte,
- Le nombre d'abonnements du compte,
- Le nombre total de tweets du compte.

4. Analyse des données

a. Fréquence des tweets

Une frise chronologique a été réalisée pour représenter le nombre de publications de tweets en fonction du temps. Elle a été annotée manuellement, en reprenant les grandes étapes-clés du scandale ORPEA, selon une recherche documentaire sur des journaux d'actualité.

b. Nettoyage des données

Les méthodes de *Natural Language Processing* (NLP) ont été appliquées pour supprimer les doublons, enlever les URL (c'est-à-dire les liens hypertextes, documents associés, images, etc.) et retirer les *stop words* (ou *mots vides*, mots non significatifs présents dans un texte et inutiles à indexer ; par exemple : le, un, ou, ...).

Une *tokenisation* a ensuite été appliquée au contenu des tweets, c'est-à-dire que les mots ont été découpés en portions, puis *lemmatisés*, selon une méthode qui consiste à remettre tous les mots à leur forme de base, aussi appelée forme canonique (retirer les pluriels, mettre les verbes à l'infinitif, ...).

Ces deux étapes ont été réalisées afin d'obtenir un contenu homogène sur l'ensemble des tweets, permettant d'appliquer le *topic modeling*.

c. Topic modeling

Afin d'étudier le contenu des tweets relatifs au scandale ORPEA, une analyse de *topic modeling* a été menée, par *Latent Dirichlet Allocation* (LDA), méthode de NLP qui consiste à explorer le contenu des tweets en mettant en évidence les principaux sujets de discussion (31). Les mots contenus dans les tweets sont regroupés par *cluster*, chaque cluster contenant des données textuelles utilisant un vocabulaire proche voire similaire. Une liste

des mots les plus fréquents est alors attribuée à chaque cluster. La bibliothèque Gensim de Python a été utilisée pour réaliser le topic modeling sur le corpus de tweets, ainsi que l'outil *pyLDAvis* pour visualiser la répartition des clusters et de leurs principaux mots-clés.

Une fois chaque cluster identifié et après observation des mots le constituant, un *label* lui a été manuellement assigné, pour en résumer le contenu. L'attribution des labels a été réalisée par deux personnes indépendantes, avec discussion en cas de désaccord.

d. Catégories des comptes

Une des fonctionnalités de Twitter est la possibilité de « certifier » un compte (aussi appelé compte « vérifié »), ce qui atteste d'une certaine authenticité du compte et lui assure une plus grande visibilité par les autres utilisateurs. Cette certification est attribuée par Twitter, pour les « comptes d'intérêt public, notoires et authentiques [...] certifiés de manière indépendante en vérifiant leur conformité à certaines exigences » (32).

Grâce à l'extraction automatique des données concernant les comptes Twitter, les comptes certifiés ont pu être mis en évidence. Après observation de ces comptes, nous avons défini les grandes catégories de comptes certifiés :

- Association : compte officiel d'une association ou d'un cercle de réflexion,
- Enseignement supérieur : compte officiel d'une grande école ou d'une université,
- Entreprise : compte officiel d'une entreprise, compte d'un entrepreneur ou d'un dirigeant d'entreprise,
- État : compte officiel d'un ministre, d'un ministère, d'un membre actif du gouvernement, d'un ambassadeur, d'une agence publique de l'État,
- Étranger : compte identifié comme non français,
- Influenceur : compte d'une personnalité publique (n'étant ni identifiée comme média ni comme politique), d'un influenceur, d'un artiste, d'une célébrité,

- Média : compte officiel d'un journal ou d'une émission, d'un journaliste, d'un producteur,
- Médical : compte d'un médecin, personnel soignant, hôpital,
- Politique : compte d'un maire, député, sénateur, compte officiel d'un parti politique, compte d'un ancien membre actif d'un parti (hors gouvernement actif),
- Syndicat : compte officiel d'un syndicat.

Nous avons réalisé une analyse manuelle du pseudo de chaque profil certifié et de sa description, parfois complétée par une recherche documentaire plus approfondie sur Internet. Cette analyse a été menée par deux examinateurs indépendants. En cas de discordance malgré discussion, un troisième examinateur est intervenu pour décider de la catégorie finale. Cette démarche nous a permis d'attribuer une des catégories ci-dessus à chaque compte vérifié.

e. Focus sur le compte ORPEA

Une analyse ciblée sur le compte @ORPEA_, compte officiel de l'entreprise ORPEA, a été menée. Tous les tweets publiés par le compte entre novembre 2019 et novembre 2022 ont été extraits par méthode de *web scraping*. Les mots les plus utilisés dans les tweets du compte ont été représentés sous la forme d'un nuage de mots.

RESULTATS

1. Corpus de tweets

Un total de 197 315 tweets contenant au moins un des 8 mots-clés identifiés a été extrait de Twitter sur la période du 1^{er} juin 2021 au 1^{er} juin 2022. Après suppression des doublons, le corpus était constitué de 85 342 tweets uniques dont 41 348 contenaient le mot-clé ORPEA, 13 244 avec Korian, 8483 avec Victor Castanet, 753 avec lesfossoyeurs, 603 avec #Orpeagate, 461 avec Cash Investigation et EHPAD, 290 avec Laurent Garcia.

La **Figure 1** représente la répartition des tweets en fonction du temps, annotée des principaux événements en lien avec le scandale. 10 076 tweets ont été publiés autour de la sortie du livre *Les Fossoyeurs* le 26 janvier 2022. L'IGAS-IGF (Inspection Générale des Affaires Sociales et Inspection Générale des Finances) a été saisie début février 2022 par l'Etat, et on dénombre 12 466 tweets autour de l'ouverture de l'enquête. Un autre pic de publications semble être lié à la diffusion de l'émission *Cash Investigation* en lien avec les EHPAD du groupe Korian, le 1^{er} mars 2022.

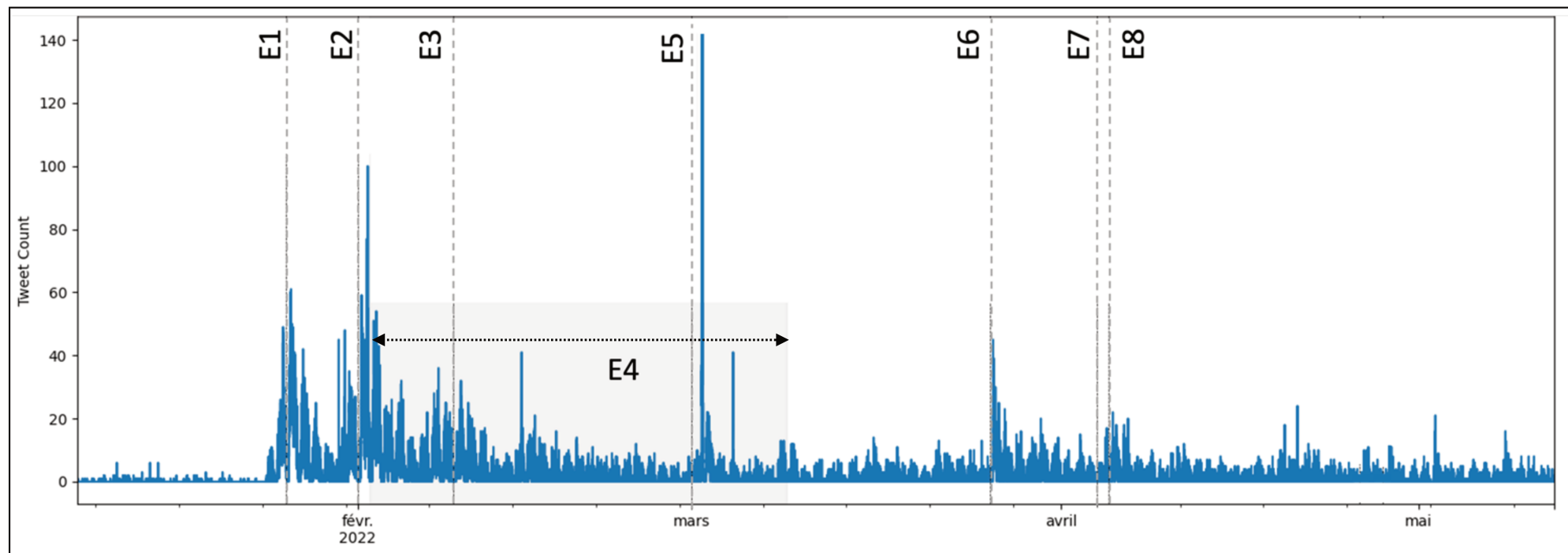


Figure 1. Frise chronologique du nombre de tweets publiés en fonction du temps, annotée des principaux événements liés au scandale.

E1 : publication du livre *Les Fossoyeurs* le 26 janvier 2022

E2 : l'État saisit l'IGAS et l'IGF le 1^{er} février

E3 : audition de Victor Castanet à l'Assemblée Nationale le 9 février

E4 : période d'auditions à l'Assemblée Nationale du 2 février au 9 mars

E5 : diffusion de l'émission *Cash Investigation* sur France 2 le 1^{er} mars

E6 : l'État porte plainte le 26 mars

E7 : Dépôt de plainte de 80 familles de résidents le 4 avril

E8 : publication du rapport de l'IGAS-IGF le 5 avril

2. Six clusters identifiés

L'analyse par topic modeling par la méthode de LDA a permis d'identifier 6 clusters de tweets. La **Figure 2** représente la répartition des différents clusters. On visualise un recouvrement entre les topics 2 et 3, et entre les topics 4 et 5. Les topics 1 et 6 semblent, au contraire, bien distincts des autres topics.

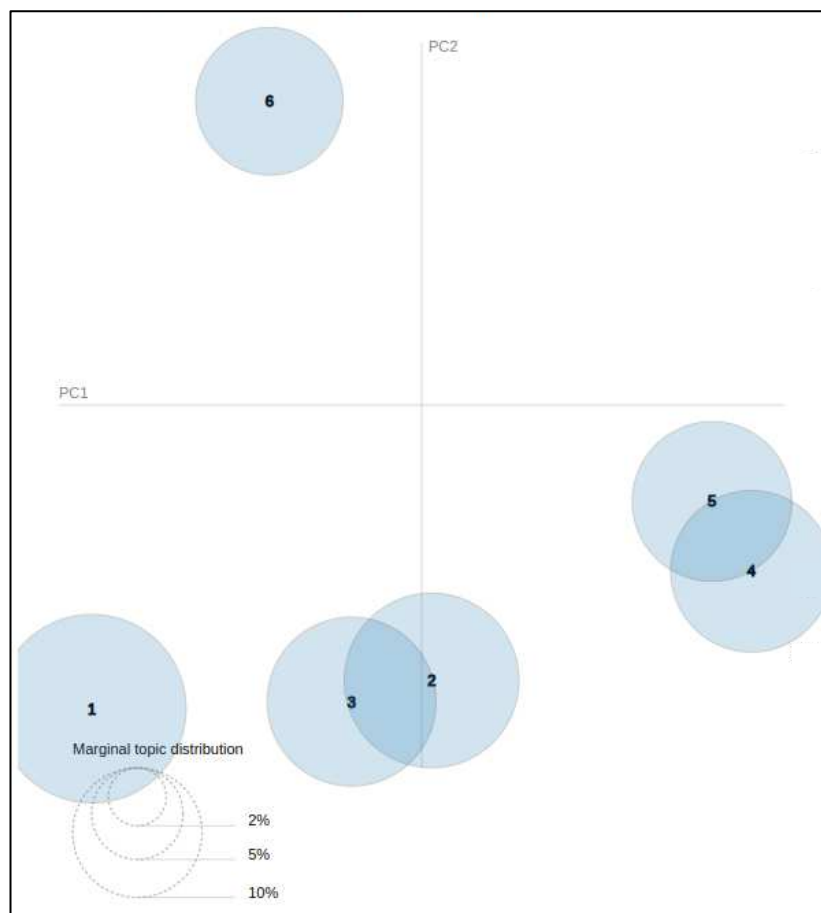


Figure 2. Visualisation de la répartition des différents topics (outil pyLDAvis).

La répartition de chaque cluster et l'assignation manuelle des différents labels sont représentés dans le **tableau 1**. La liste complète des mots contenus dans chaque topic est disponible en **ANNEXE I**.

Tableau 1. Répartition des clusters, assignation des labels, et mots les plus fréquemment identifiés par label, par ordre décroissant.

Cluster	% des tweets	Label attribué	Exemples de mots contenus dans le topic
Topic 1	18,2%	Réactions sur le fonctionnement des EHPAD	Orpea, personne, age, public, plus, resident, grand, personnel, actionnaire, argent, payer
Topic 2	13,0%	Impact du scandale ORPEA sur le groupe Korian	Korian, Cash Investigation, compte, premier, heure, question, numero, soir, Elise Lucet
Topic 3	15,2%	Dénonciation de maltraitance dans les maisons de retraite suite au scandale	Orpea, groupe, livre, maltraitance, priver, action, enquête, dénoncer, famille, plainte
Topic 4	15,6%	Médiatisation du scandale ORPEA	Orpea, scandale, enquête, affaire, état, gouvernement, rapport, directeur
Topic 5	21,0%	Influence politique et volonté d'action	Victor Castanet, faire, avoir, pouvoir, falloir, Brigitte Bourguignon, plus, dire, Olivier Veran
Topic 6	17,0%	Avis des proches et soignants, opinions personnelles	Orpea, aller, bien, EHPAD, bon, Macron, vieux, place, éthique, bfmtv, retraite

a. Fonctionnement du système des EHPAD

Le premier topic est représenté par des mots relatifs au fonctionnement des EHPAD en France (personnel, résident, public, ORPEA) et notamment l'aspect financier (plus, argent, actionnaire, payer). Les tweets contenus dans cette partie questionnent de manière globale le fonctionnement des EHPAD privés :

« Affaire Orpea : faut-il remettre en question la gestion des Ehpads par le secteur privé lucratif ? [LIEN] [LIEN]. »

@sudouest

[LIEN] indique la présence d'un lien hypertexte dans le tweet, non reproduit ici par souci de lisibilité.

Certains tweets soulignent l'importance des enjeux financiers sous-jacents :

« Orpea a déclaré avoir signé un accord avec plusieurs banques afin d'obtenir un crédit-syndiqué sécurisé d'un montant de 1,733 milliard d'euros. [LIEN]. »

@lopinion_fr

D'autres tweets ont pour souhait de dénoncer la financiarisation des EHPAD au service d'objectifs de rentabilité ou de maximisation des profits, et au détriment des conditions de vie des personnes âgées :

*« Orpea maltraite les résidents
Orpea maltraite ses salariés
Orpea bichonne ses indicateurs de rentabilité pour prendre soin de ses profits !
#stopmarchandisationdu4eage
#ehpad publics 100%
#respectdesainés
#respectdelavie #albertville »*

@raftman63

« #Orpea, qui fait payer ses résidents entre 6500 et 12000 euros par mois, fait tout pour grignoter quelques euros et engranger toujours plus de profits sur le dos des personnes âgées. Nous exigeons des démissions ! [LIEN] »

@IndianaJones_2

b. Impact du scandale ORPEA sur le groupe Korian

Le deuxième topic est en lien avec la diffusion de l'émission Cash Investigation sur les EHPAD du groupe Korian.

« "Cash Investigation" : une journaliste embauchée sans expérience dans un Ehpad [LIEN]

Marie Maurice, journaliste pour Cash Investigation, a souhaité comprendre la réalité des Ehpad et pour cela, en a infiltré un... Dans l'enquête Ehpad : l'heure des comptes ... »

@FemmeActuRSS

Les auteurs des tweets soulèvent l'idée que les entreprises d'EHPAD privés utilisent toutes les mêmes procédés :

« @ElisaBraun @marcendeweld @VictorCastanet la manipulation obscure de l'argent semble être loi dans le monde des EHPAD... voir les révélations de @MRenahy sur DomusVI »

@_TFArm

« Ehpads privés : repas rationnés, effectifs insuffisants... Cash Investigation dénonce à son tour de grands manquements [LIEN] via @ladepechedumidi »

@AsaAsso

« @BrigBourguignon @olivierveran @IGF_finances A quand une enquête chez #Korian, #Domusvi, les autres groupes et les #Ehpad publics? Faut-il un livre sur chaque cas pour espérer que l'Etat agisse? Et une élection? Journalistes, assoc et lanceurs alerte dénoncent depuis des années. »

@collectif9471

« @HappyWandererss @cashinvestigati @korian Ce sont tous des pourris qui marchent mano a mano avec l'ARS, l'HAS et les politiques de tous poils. »

@ROBERTMASSEGLI1

c. Dénonciation de maltraitance des personnes âgées institutionnalisées suite au scandale

Dans ce topic, on retrouve des mots en lien avec la dénonciation des abus et de la maltraitance subis par les personnes âgées en EHPAD, avec des termes tels que : dénoncer, famille, cause, maltraitance, plainte.

A la différence du topic 1, la dénonciation se veut active et contraignante, avec par exemple le dépôt collectif de plaintes :

« Scandale des Ehpad : Orpea sous le coup d'une première salve de plaintes de familles @lemondefr @CgtSanteSocial CFDT et Fo ont déjà porté plainte pour discriminations syndicales et demande d'annulation des élections prof @ChaveDominique [LIEN] »

@gengfrancoise

« "Ehpad : 80 plaintes de familles déposées contre le groupe Orpea pour "mise en danger en danger d'autrui" et "homicide involontaire"" [LIEN] »

@123_INFO_FR

d. Médiatisation du scandale ORPEA

Le quatrième topic aborde la médiatisation des étapes du scandale en lui-même, en lien avec la publication du livre. On retrouve des termes comme : scandale, enquête, affaire, rapport, PDG, ministre, état, directeur.

Ce topic illustre l'utilisation des réseaux sociaux comme média à diffusion rapide, avec notamment l'utilisation de hashtags :

« Scandale #Orpea : le parquet de Nanterre a annoncé l'ouverture d'une enquête sur le groupe privé gestionnaire d'Ehpad, soupçonné de maltraitance institutionnelle et de détournements de fonds publics. »

@InfosFrancaises

« #Orpea : le gouvernement publie un rapport d'#enquête accablant [LIEN] ».

@CourrierPicard

e. Influence politique et volonté d'action

Le cinquième topic est le plus représenté dans le corpus (21%). L'évocation de Victor Castanet, auteur du livre, ainsi que de personnalités politiques (Brigitte Bourguignon, ministre déléguée chargée de l'Autonomie au moment du scandale ; Olivier Véran, ministre de la Santé à cette période) suggèrent l'influence de la politique sur le scandale et la gestion des EHPAD. La prédominance de verbes (falloir, avoir, pouvoir, faire, savoir) suppose une volonté d'action, de changement par les auteurs des tweets. Les thématiques sont variées, allant de la question de la place que notre société accorde aux personnes âgées à la maltraitance structurelle au service d'une financiarisation de la dépendance des aînés.

Ceci est illustré par les tweets ci-dessous :

« @BrigBourguignon @VictorCastanet @olivierveran Il est plus que temps que nos aînés, parents, amis, bénéficient d'un droit à vieillir dans la dignité »

@BrunoQuennesson

« @BrigBourguignon @VictorCastanet @olivierveran Mais ou vivez vous ma pauvre dame. Le système libéral que vous chérissez maltraite nos aînés pour du fric toujours + de fric. Vous n'allez bien sur rien faire car vous être lâche.....les lobbys n'auront aucun mal à vous faire taire ! A vomir »

@phgarcia69

« @qo_ideme6 @cavousf5 @VictorCastanet @MattBelliard Ce n'est pas que les politiques, c'est toute la société qui s'en fout des vieux ! On ne veut pas voir les vieux, on les cache dans des maisons de retraite et ils sont laissés à l'abandon... Nous sommes tous responsables collectivement il faudrait se remettre en question. »

@KTKrukowska

f. Avis des proches et soignants, opinions personnelles sur un sujet sociétal

Le sixième cluster relate les opinions de soignants ou d'aidants touchés par le scandale et exprimant leur avis sur le sujet, ou rapportant un témoignage. On identifie aussi des opinions de personnes concernées d'un peu plus loin. Ces tweets, plus personnels, relatent des émotions suscitées par le scandale :

« @SabDeliry @korian @EliseLucet @MarieMaurice Votre témoignage et celui de l'infirmier lanceur d'alerte étaient vraiment poignants! Je suis triste pour le personnel d'ephad qui vit la double peine: conditions de travail horribles de la part de la direction + en 1^{ère} ligne avec les patients qui ne peuvent s'en prendre qu'à eux. »

@haricestldon

D'autres tweets réfèrent à l'expérience personnelle vécue des aidants, pour apporter des éléments d'amélioration :

« @tchenglaurence @SYNERPAcom @orpea @jdmatisse @FGzil @EspaceEthique @CPAE_ThinkTank Ds certains #ehpad l'épreuve du covid a peut-être renforcé ls liens avec ls familles. C'est loin d'être une généralité, par ex. ds l'ehpad de mon père un CVS a eût lieu au bout de 9 mois sans et ls familles n'en ont même ps été informées. On peut faire mieux non ? @CPAE_ThinkTank »

@Dantweet75

3. Profils des comptes s'exprimant sur le scandale

L'ensemble des tweets sur le scandale ORPEA a été publié par un total de 22 835 utilisateurs, dont 641 comptes vérifiés. Parmi les 85 342 tweets, une grande majorité (73 597) est publiée par des comptes non vérifiés. L'attribution manuelle d'une catégorie à chaque compte vérifié a mis en évidence la prédominance de la catégorie « média », avec 9945 tweets. On identifie aussi une proportion non négligeable de personnalités politiques (731 tweets) et 521 tweets publiés par des entreprises. La répartition des catégories de comptes vérifiés est représentée dans la **Figure 4**.

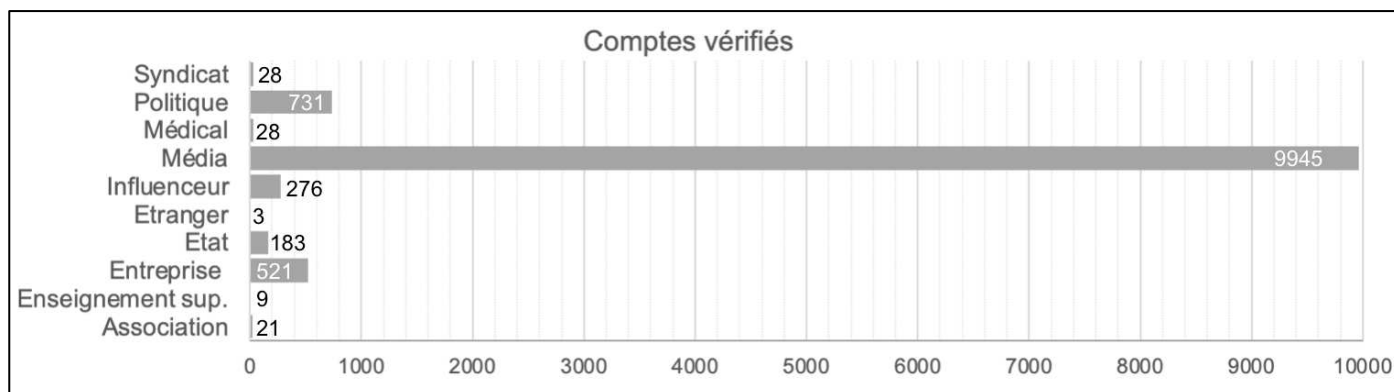


Figure 4. Répartition des tweets selon les catégories de comptes vérifiés.

De façon plus détaillée, voici les principaux comptes tweetant sur le sujet, parmi les catégories les plus représentées :

- Média : @LCP, @BFMTV, @YannLeGales (journaliste indépendant et ancien rédacteur chez *Le Figaro*), @franceinfo, @publicsenat, @lemondefr, @libe (compte officiel du journal *Libération*), @OuestFrance, @LCI.
- Politique : @Francois_Ruffin (député La France Insoumise), @BorisVallaud (député socialiste), @RachidTemal (sénateur), @agnesfirmin (Agnès Firmin-Le Bodo, députée lors du scandale et actuellement Ministre déléguée chargée de l'organisation territoriale et des professions de santé), @CarolineFiat54 (députée communiste et ancienne aide-soignante).
- Entreprise : @korian, @InvestirFr (branche économique du journal *Les Echos*), @ORPEA_, @GallantNicolas1 (analyste financier), @Novethic (filiale de la Caisse des dépôts et consignations, spécialisée dans la finance durable et l'économie socialement responsable).

4. Focus sur le compte @ORPEA_

Le compte Twitter officiel du groupe ORPEA a publié 389 tweets entre novembre 2019 et novembre 2022, avec un maximum de publications en juin 2022. Une période d'absence totale d'activité est observée entre avril 2021 et le 14 avril 2022. Les sujets principaux de ces tweets sont représentés dans le nuage de mots de la **Figure 5**, et semblent évoquer majoritairement l'envie de fédérer le groupe et les résidents (« nos », « nous », « résident », « personnel ») ainsi que la volonté d'aller de l'avant et de s'améliorer (« votez », « améliorer », « propositions », « unjouruneproposition », « nouvelle », ...). Aucun mot en lien avec le scandale en lui-même n'est identifié.



Figure 5. Nuage de mots représentant les mots les plus fréquemment retrouvés sur le compte Twitter @ORPEA_.

DISCUSSION

1. Résumé des principaux résultats

Avec plus de 85 000 tweets extraits, l'analyse semi-automatisée des réactions sur Twitter concernant le scandale ORPEA a permis de montrer l'importance de l'intérêt porté sur la problématique des conditions de vie des personnes âgées institutionnalisées.

Les tweets abordent des thématiques larges, qui confirment que le scandale ORPEA s'inscrit dans un débat de société profond, concernant la place de la personne âgée et la gestion du Grand Âge. Il y a un souhait de diffuser largement l'information et de la relayer avec une certaine efficacité, montrant qu'il s'agit d'un sujet contemporain qui intéresse une large portion de la société. Globalement, l'analyse des tweets suppose qu'un lien est fait entre la financiarisation des EHPAD et les mauvais traitements apportés aux sujets âgés dépendants. Cette situation est souvent dénoncée et considérée comme inacceptable par les auteurs des tweets, qui appellent alors à un changement par l'action politique, voire juridique en lien avec le dépôt de plaintes par exemple.

L'analyse des comptes réagissant sur le sujet a mis en évidence une forte prédominance de comptes non-vérifiés, représentant les citoyens « lambdas ». C'est tout un chacun en tant qu'individu qui semble vouloir s'exprimer sur la question. Parmi les comptes vérifiés, on identifie un nombre conséquent de médias ayant une diffusion à grande échelle (*Le Monde, Libération, Le Figaro, etc.*), témoignant de l'importance accordée à ce scandale dans les médias nationaux français et dépassant le simple « buzz » des réseaux sociaux. D'importantes personnalités politiques à rayonnement national (anciens ministres, députés, sénateurs) ont exprimé leur avis, confirmant l'ampleur du scandale notamment dans sa dimension politique. Il faut souligner que leur prise de parole a pu être amplifiée, le scandale

s'étant déroulé en pleine période électorale, lors de la campagne présidentielle 2022. Enfin, un certain nombre d'entreprises, notamment du domaine de la finance, se sont exprimées, ce qui montre que la problématique des EHPAD, telle qu'elle est actuellement envisagée, n'est pas limitée au domaine médico-sanitaire, mais représente bien un enjeu économique majeur.

En revanche, l'entreprise ORPEA semble s'être peu exprimée sur Twitter pendant la période du scandale, et aborde ensuite des thématiques différentes du reste des publications, plutôt centrées sur la volonté de rassembler et fédérer au sein de l'entreprise, et l'envie d'aller de l'avant. Elle revient sur Twitter le 14 avril 2022 au décours du début de la guerre en Ukraine, après un silence long de 12 mois, pour apporter son soutien au peuple ukrainien. L'entreprise ORPEA travaille ensuite à un changement d'image, avec notamment l'annonce de la création d'un groupe éthique au sein de l'entreprise et l'avènement des États Généraux d'ORPEA, moments consacrés à la transformation du groupe.

2. Conséquences du scandale

La médiatisation du scandale ORPEA sur les réseaux sociaux a permis de faciliter la prise de parole pour dénoncer la maltraitance subie par certaines personnes âgées institutionnalisées, comme en témoigne le topic 3. Ceci a conduit à la création, par quelques utilisateurs Twitter, de l'Observatoire du Grand Âge (33), association qui a pour but de réfléchir sur la question des EHPAD en France, permettre de dénoncer les abus et recueillir les témoignages de familles et aidants. Ces événements ont aussi mis en lumière des comptes Twitter comme le Cercle des Proches Aidants en EHPAD (@CPAE_ThinkTank), collectif de familles engagées dans la lutte contre la maltraitance dans les maisons de retraite.

Ensuite, comme le montrent le topic 1 et les nombreux comptes d'entreprises de finance, la question économique de la gestion des EHPAD semble être au premier plan. La publication du livre et la cascade médiatique qui a suivi ont entraîné, dès le lendemain, une chute des actions en bourse du groupe ORPEA de plus de 16% (34). Par ailleurs, une amende de 55,8 millions d'euros a été réclamée par l'État, initialement contestée puis finalement acceptée en novembre 2022 par le groupe (35).

Ensuite, le topic 5, majoritaire, témoigne de l'intérêt porté sur la gestion politique des établissements. La population semble être dans l'attente d'un changement dans la gestion du Grand Âge. Après le scandale, de nombreuses enquêtes ont été menées pour vérifier les faits et les accusations, notamment par l'IGAS-IGF, autorités françaises. Celles-ci ont confirmé les dires, mais peu de réponses ont été apportées de la part du gouvernement, hormis une augmentation du contrôle des maisons de retraite (l'État ayant ordonné le contrôle des 7500 EHPAD français sous deux ans, contre tous les vingt ans auparavant (36)). On peut ajouter que le projet de Loi Grand Âge ne cesse d'être reporté depuis de nombreuses années (37).

En regard, il convient de souligner l'effort politique discret mais gigantesque qui a été réalisé pour sauver les investissements des acteurs financiers français impliqués, et ce au plus haut niveau de l'État. En effet, selon *Médiapart*, le groupe ORPEA se serait vu injecter 1,35 milliard d'euros de fonds propres par la Caisse des Dépôts et Consignations (38). En parallèle, des négociations de crédits auraient eu lieu avec les banques françaises pour réduire la dette d'ORPEA d'un total de 3,8 milliards d'euros, tout en diminuant le poids des actionnaires et investisseurs étrangers, le tout sous la surveillance lointaine de l'État (38).

3. Forces et limites de l'étude

Les principales forces de cette étude sont d'abord représentées par le grand nombre de tweets extraits, avec plus de 85 000 tweets sur le sujet. Ce nombre important est obtenu grâce aux méthodes de *web scraping*, qui permettent d'extraire tous les tweets contenant au moins un mot-clé choisi, sans exception. Sans ces techniques, les études utilisent plutôt un échantillon de tweets, manuellement sélectionné (39).

La méthode de cette étude se démarque par son originalité, avec l'utilisation de techniques automatisées d'analyse du langage, novatrices dans le domaine de la gériatrie. En effet, il existe peu d'études utilisant les techniques de NLP actuellement au sujet des personnes âgées, d'autant plus sur les réseaux sociaux. Quelques auteurs s'y sont intéressés lors de la pandémie à COVID-19, par exemple l'équipe de *Xiang et al.* qui a analysé les tweets relatifs aux personnes âgées pendant cette période à la recherche de contenu d'âgisme et de discrimination envers les aînés (40). Une autre étude a analysé les sentiments des aidants de patients institutionnalisés atteints de troubles cognitifs pendant le confinement dû à la COVID-19 (41), en ne se limitant ainsi qu'à une partie de la population. Notre étude, par son grand nombre de tweets et sa sélection large, s'intéresse à une part plus globale des utilisateurs.

Cependant, cette étude présente aussi quelques limites. Premièrement, elle ne concerne que le réseau social Twitter et donc ne reflète qu'une partie de la population générale. En effet, les utilisateurs Twitter sont majoritairement représentés par des sujets jeunes. Aux Etats-Unis, l'âge médian des utilisateurs est de 40 ans (42). Les plus de 65 ans, qui représentent environ 20% de la population, ne sont que 8% des utilisateurs Twitter (42). En France, une personne âgée de plus de 65 ans sur 10 est inscrite sur le réseau (43). De plus, les données extraites sur Twitter ne permettent pas d'identifier dans notre étude le profil

détaillé des comptes tweetant sur le sujet. Ainsi, nous ne disposons pas d'informations sur le sexe ni l'âge des utilisateurs.

Certaines études utilisant le topic modeling ont aussi pu s'intéresser à plusieurs réseaux sociaux. Par exemple, l'étude de *Li et al.* comparait des notes données aux EHPAD sur différents réseaux (et non des données textuelles comme dans notre étude) (44). Il pourrait être utile de s'intéresser aux autres réseaux sociaux, car les codes d'utilisation de chaque réseau et les profils d'utilisateurs sont différents. Cela rendrait les résultats plus représentatifs de la population générale.

Ensuite, certaines étapes nécessitent encore le recours à une analyse manuelle, et donc à risque de subjectivité, notamment lors de l'attribution des catégories de comptes vérifiés, ou le choix des labels. Cette subjectivité est toutefois limitée par la double relecture, avec parfois recours à une tierce personne en cas de désaccord.

Enfin, il reste à rappeler que l'analyse des réseaux sociaux ne permet pas de donner la parole aux principaux intéressés, à savoir les personnes âgées institutionnalisées.

CONCLUSION

Cette étude a permis, par une méthode originale, de montrer l'intérêt marqué de la population sur le scandale ORPEA, en relevant le caractère inacceptable des mauvais traitements imposés aux personnes âgées dépendantes, avec la mise en évidence d'un lien entre cette maltraitance et la financiarisation des EHPAD. Les auteurs des tweets sont représentés en grande partie par des utilisateurs ordinaires, ainsi que par des médias et des personnalités politiques importantes. Les tweets expriment également le souhait d'améliorer le fonctionnement des EHPAD. Ce scandale semble donc s'inscrire dans un débat sociétal plus large, à savoir la place de la personne âgée institutionnalisée en France.

Notre étude montre ainsi que les méthodes de *web scraping* et d'analyse automatisée des contenus, doublées d'une validation humaine, permettent de décrire et percevoir l'opinion, de « prendre le pouls », d'une partie de la population sur une thématique d'intérêt majeur en gériatrie.

Une partie de ce travail a été valorisée par la publication dans le journal *Studies in Health Technology and Informatics* de l'article suivant : *Publication Dynamics on Social Media During the Orpea Nursing Homes Scandal: A Twitter Analysis (45)*.

BIBLIOGRAPHIE

1. Dictionnaire de l'Académie Nationale de Médecine [Internet]. [cité 17 juill 2023]. Disponible sur: <http://51.68.80.15/search/results?titre=d%C3%A9pendance>
2. Espérance de vie - Mortalité – Tableaux de l'économie française | Insee [Internet]. [cité 17 juill 2023]. Disponible sur: <https://www.insee.fr/fr/statistiques/3676610?sommaire=3696937>
3. Population par âge – Tableaux de l'économie française | Insee [Internet]. [cité 27 janv 2023]. Disponible sur: <https://www.insee.fr/fr/statistiques/4277619?sommaire=4318291>
4. Rahola A. Synthèse du débat national sur la dépendance.
5. Des résidents de plus en plus âgés et dépendants dans les établissements d'hébergement pour personnes âgées | Direction de la recherche, des études, de l'évaluation et des statistiques [Internet]. [cité 27 janv 2023]. Disponible sur: <https://drees.solidarites-sante.gouv.fr/publications-communique-de-presse/etudes-et-resultats/des-residents-de-plus-en-plus-ages-et>
6. Infographie : L'hébergement des personnes âgées en établissement - Les chiffres clés | Direction de la recherche, des études, de l'évaluation et des statistiques [Internet]. [cité 8 sept 2022]. Disponible sur: <https://drees.solidarites-sante.gouv.fr/infographie-video/infographie-lhebergement-des-personnes-agees-en-etablissement-les-chiffres-cles>
7. Chiffres clés | ORPEA Groupe [Internet]. [cité 29 janv 2023]. Disponible sur: <https://www.orpea-groupe.com/le-groupe/qui-sommes-nous/chiffres-cles>
8. Castanet, Victor. Les Fossoyeurs : Révélations sur le système qui maltraite nos aînés. Paris: Fayard; 2022.
9. « Les Fossoyeurs », un livre qui ouvre le débat sur la gestion et le contrôle des maisons de retraite. Le Monde.fr [Internet]. 24 janv 2022 [cité 8 sept 2022]; Disponible sur: https://www.lemonde.fr/idees/article/2022/01/24/les-fossoyeurs-un-livre-qui-ouvre-le-debat-sur-la-gestion-et-le-contrôle-des-maisons-de-retraite_6110753_3232.html
10. IGAS-IGF. Mission sur la gestion des établissements d'hébergement de personnes âgées dépendantes (EHPAD) du groupe Orpea. 2022 mars. Report No.: 2022-M-012-06 et 2022-016R.
11. Clariane en France | Clariane [Internet]. [cité 1 août 2023]. Disponible sur: <https://www.clariane.com/nos-implantations/clariane-en-France>

12. Cash Investigation Ehpac : l'heure des comptes ? [Internet]. 2022 [cité 29 janv 2023]. Disponible sur: <https://www.france.tv/france-2/cash-investigation/3103799-ehpac-l-heure-des-comptes.html>
13. Sun L, Tao C, Xie Z, Li D. Promotion of Disposable Electronic Cigarette Flavors and Topics on Twitter. *Int J Environ Res Public Health*. déc 2020;17(24):9221.
14. Westmaas JL, Masters M, Bandi P, Majmundar A, Asare S, Diver WR. COVID-19 and Tweets About Quitting Cigarette Smoking: Topic Model Analysis of Twitter Posts 2018-2020. *Jmir Infodemiology*. 16 mai 2022;2(1):e36215.
15. Sutton J, Vos SC, Olson MK, Woods C, Cohen E, Gibson CB, et al. Lung Cancer Messages on Twitter: Content Analysis and Evaluation. *J Am Coll Radiol JACR*. janv 2018;15(1 Pt B):210-7.
16. Nzali MDT, Bringay S, Lavergne C, Mollevi C, Opitz T. What Patients Can Tell Us: Topic Analysis for Social Media on Breast Cancer. *JMIR Med Inform*. 31 juill 2017;5(3):e7779.
17. Lyu JC, Han EL, Luli GK. COVID-19 Vaccine–Related Discussion on Twitter: Topic Modeling and Sentiment Analysis. *J Med Internet Res*. 29 juin 2021;23(6):e24435.
18. Surian D, Nguyen DQ, Kennedy G, Johnson M, Coiera E, Dunn AG. Characterizing Twitter Discussions About HPV Vaccines Using Topic Modeling and Community Detection. *J Med Internet Res*. 29 août 2016;18(8):e6045.
19. Farhadloo M, Winneg K, Chan MPS, Hall Jamieson K, Albarracin D. Associations of Topics of Discussion on Twitter With Survey Measures of Attitudes, Knowledge, and Behaviors Related to Zika: Probabilistic Study in the United States. *JMIR Public Health Surveill*. 9 févr 2018;4(1):e16.
20. Low DM, Rumker L, Talkar T, Torous J, Cecchi G, Ghosh SS. Natural Language Processing Reveals Vulnerable Mental Health Support Groups and Heightened Health Anxiety on Reddit During COVID-19: Observational Study. *J Med Internet Res*. 12 oct 2020;22(10):e22635.
21. Pellert M, Lasser J, Metzler H, Garcia D. Dashboard of Sentiment in Austrian Social Media During COVID-19. *Front Big Data*. 26 oct 2020;3:32.
22. Xue J, Chen J, Hu R, Chen C, Zheng C, Liu X, et al. Twitter discussions and emotions about COVID-19 pandemic: a machine learning approach [Internet]. *arXiv*; 2020 [cité 5 juill 2023]. Disponible sur: <http://arxiv.org/abs/2005.12830>
23. Pool J, Namvar M, Akhlaghpour S, Fatehi F. Exploring public opinion about

- telehealth during COVID-19 by social media analytics. *J Telemed Telecare*. déc 2022;28(10):718-25.
24. McKendrick DR, Cumming GP, Lee AJ. Increased Use of Twitter at a Medical Conference: A Report and a Review of the Educational Opportunities. *J Med Internet Res*. 11 déc 2012;14(6):e176.
25. Kitching F, Winbolt M, MacPhail A, Ibrahim JE. Web-based social media for professional medical education: Perspectives of senior stakeholders in the nursing home sector. *Nurse Educ Today*. déc 2015;35(12):1192-8.
26. Salles A. Analyse du Discours médical sur Twitter : étude d'un corpus de tweets émis par des médecins généralistes entre juin 2012 et mars 2017 et contenant le #DocToC. 2018.
27. Rossetini G, Peressutti V, Visintini E, Fontanini R, Caruzzo D, Longhini J, et al. Italian nurses' experiences of the COVID-19 pandemic through social media: A longitudinal mixed methods study of Internet posts. *Digit Health*. 1 janv 2022;8:20552076221129080.
28. Bosley JC, Zhao NW, Hill S, Shofer FS, Asch DA, Becker LB, et al. Decoding twitter: Surveillance and trends for cardiac arrest and resuscitation communication. *Resuscitation*. 1 févr 2013;84(2):206-12.
29. Krotov V, Silva L. Legality and Ethics of Web Scraping. In 2018.
30. Smith E, Michalski S, Knauth KHK, Kaspar K, Reiter N, Peters J. Large-Scale Web Scraping for Problem Gambling Research: A Case Study of COVID-19 Lockdown Effects in Germany. *J Gamb Stud*. sept 2023;39(3):1487-504.
31. Blei DM. Probabilistic topic models. *Commun ACM*. avr 2012;55(4):77-84.
32. Exigences de certification Twitter : comment obtenir la coche bleue [Internet]. [cité 21 juin 2023]. Disponible sur: <https://help.twitter.com/fr/managing-your-account/about-twitter-verified-accounts>
33. L'Observatoire du Grand Âge [Internet]. [cité 5 juill 2023]. L'Observatoire du Grand Âge. Disponible sur: <https://ogra.fr/>
34. Ehpads : le groupe Orpea chute en Bourse, confronté à la parution d'un livre dénonçant de graves défaillances dans ses maisons de retraite. *Le Monde.fr* [Internet]. 24 janv 2022 [cité 3 juill 2023]; Disponible sur: https://www.lemonde.fr/societe/article/2022/01/24/maisons-de-retraite-un-livre-denonce-de-graves-defaillances-dans-les-ehpad-du-groupe-orpea-qui-denonce-des-accusations-mensongeres_6110796_3224.html

35. Scandale des maisons de retraite Orpea : le groupe d'Ehpad accepte finalement de rembourser les 55 millions d'aides publiques - ladepeche.fr [Internet]. [cité 16 août 2023]. Disponible sur: <https://www.ladepeche.fr/2022/11/22/scandale-orpea-le-groupe-dehpad-accepte-finalement-de-rembourser-les-55-millions-daides-publiques-quil-a-touchees-10819347.php#:~:text=Des%20aides%2Dsoignants%20qui%20n,sans%20en%20avoir%20la%20qualification.>
36. gouvernement.fr [Internet]. [cité 17 août 2023]. Renforcement des contrôles et de la transparence dans les Ehpad. Disponible sur: <http://www.gouvernement.fr/actualite/renforcement-des-contrôles-et-de-la-transparence-dans-les-ehpad>
37. « Bien vieillir » : le texte de loi sur le grand âge à nouveau reporté - Le Parisien [Internet]. [cité 16 août 2023]. Disponible sur: <https://www.leparisien.fr/politique/bien-vieillir-le-texte-de-loi-sur-le-grand-age-a-nouveau-reporte-18-07-2023-WQLNCS4O75H23IMLUBCYFP5JSQ.php>
38. Thépot M. Mediapart. [cité 16 août 2023]. Petits arrangements entre amis pour se répartir les milliards d'Orpea. Disponible sur: <https://www.mediapart.fr/journal/economie-et-social/270623/petits-arrangements-entre-amis-pour-se-repartir-les-milliards-d-orpea>
39. Sedrak MS, Salgia MM, Decat Bergerot C, Ashing-Giwa K, Cotta BN, Adashek JJ, et al. Examining Public Communication About Kidney Cancer on Twitter. JCO Clin Cancer Inform. mars 2019;3:1-6.
40. Xiang X, Lu X, Halavanau A, Xue J, Sun Y, Lai PHL, et al. Modern Senicide in the Face of a Pandemic: An Examination of Public Discourse and Sentiment About Older Adults and COVID-19 Using Machine Learning. J Gerontol B Psychol Sci Soc Sci. 14 mars 2021;76(4):e190-200.
41. Yoon S, Broadwell P, Alcantara C, Davis N, Lee H, Bristol A, et al. Analyzing Topics and Sentiments from Twitter to Gain Insights to Refine Interventions for Family Caregivers of Persons with Alzheimer's Disease and Related Dementias (ADRD) During COVID-19 Pandemic. Stud Health Technol Inform. 14 janv 2022;289:170-3.
42. How Twitter Users Compare to the General Public | Pew Research Center [Internet]. [cité 5 juill 2023]. Disponible sur: <https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/>
43. Part des internautes français inscrits sur Twitter par âge 2019 | Statista [Internet]. [cité 6 sept 2023]. Disponible sur: <https://fr.statista.com/statistiques/996674/part-personnes-inscrites-twitter-france-par-age/>

44. Li Y, Cai X, Wang M. Social media ratings of nursing homes associated with experience of care and “Nursing Home Compare” quality measures. *BMC Health Serv Res.* déc 2019;19(1):260.
45. Fruchart M, Verdier L, Beuscart JB, Lamer A. Publication Dynamics on Social Media During the Orpea Nursing Homes Scandal: A Twitter Analysis. *Stud Health Technol Inform.* 18 mai 2023;302:502-3.

ANNEXES

ANNEXE I. Liste des mots contenus dans chaque cluster par fréquence d'apparition

Topic 1 :

Orpea	Payer	Hôpital
Personne	Ainé	Ancien
Age	Repas	Economie
Public	Profit	Travailler
Plus	Soignant	Million
Resident	Société	Rationner
Grand	Euro	Prix
Faire	Service	Moyen
Devoir	Fin	Patient
Jour	Etablissement	Maltraiter
Personnel	Mère	Honte
Mois	Lucratif	Entreprise
Comment	Moins	Laisser
Actionnaire	Humain	Cher
Argent	Parent	Secteur
Vie	Privé	Aide
Soin	Santé	

Topic 2 :

Korian	Mardi	Etablissement
Cash_investigation	Face	Diffuser
Cashinvestigati	Dernier	Défendre
Compte	Commission	Embaucher
Premier	Directrice	Sophie_boissard
Heure	Fort	Cadre
Question	Prochain	Sujet
Cashinvestigation	Réponse	Equipe
Numéro	Sénat	Débat
Soir	Point	Article
Mars	Regarde	Bon_bouille
Elis_lucet	Vidéo	Mission
Eliselucet	Audition	Valeur
Social	Direct	Demain
Député	Matin	Quel
Moment	Suivre	Répondre
Emission	Reportage	

Topic 3 :

Orpea	Plainte	Proposer
Groupe	Gestion	Sortie
Livre	Etablissement	Auteur
Maltraitance	Révélation	Recevoir
Priver	Témoignage	Dysfonctionnement
Mettre	Pratique	Denon_tour
Fossoyeur	Système	Direction
Famille	Million_euro	Chute
Action	Grave	Dérive
Enquête	Français	Franceinter
Victor_castanet	Vendre	Denonce
Journaliste	Directan	Révéler
Dénoncer	Manquement	Près
Bourse	Tour	Xavier_bertrand
Cause	Résident	
Lesfossoyeurs	Déposer	
Suite	Maltraitanc	
Patron	Semaine	

Topic 4 :

Orpea	Financier	Delit_initier
Scandale	Public	Annonce
Enquête	Santé	Ouvrir
Directeur	Dg	Limoger
Affaire	Viser	Rendre
Maison	Inspection	Secret
Nouveau	Révélation	Annoncer
Etat	Lancer	Mediapart
Gouvernement	Actualité	Parision
Rapport	E	Dotation
Général	Plainte	Saisir_justice
Ancien	Publier	Brigitt_bourguignon
Retraire	Porte	Soupcon_maltraintac
Ministre	Source	Administratif
Pdg	Justice	Remettre
Dirigeant	Scandal	Lemondefr
Franceinfo	Convoquer	

Topic 5 :

Victorcastanet	Passer	Attendre
Faire	Fait	Probleme
Avoir	Merci	Connaitre
Pouvoir	Déjà	Mal
Falloir	Donc	Cavousf5
Brigbourguignon	Jamais	An
Plus	Ars	Mattbelliard
Dire	Année	Parle
Autre	Peu	Beaucoup
Meme	Encore	Parler
Voir	Contrôle	Croire
Etre	Chose	Vrai
Bien	Monde	Travail
Très	Temps	Fois
Savoir	Pourquoi	Changer
Olivierveran	Cpae_thinktank	Juste
Alors	Garcialaurent16	

Topic 6 :

Orpea	Responsable
Aller	Pouvoir
Bien	Plutôt
Ephad	Vite
Bon	Oublier
Macron	Orpeagat
Vieux	Voter
Faire	Prison
An	Sortir
Petit	Enfant
Place	Trouver
Mieux	Temps
Bfmtv	Vraiment
Voir	Monsieur
Ethique	Venir
Retraite	Français
Devoir	Developpement_ethique
Meme	Père
Finir	Politique
Claudemalhuret	Trop
Coup	Histoire
Toujours	Jlmelenchon
Prendre	Covid
Espérer	Envoyer

AUTEURE : Nom : VERDIER

Prénom : Laetitia

Date de soutenance : 12 octobre 2023

Titre de la thèse : Scandale ORPEA – analyse des réactions sur Twitter

Thèse - Médecine - Lille 2023

Cadre de classement : Gériatrie

DES + FST/option : Gériatrie

Mots-clés : web scraping, Twitter, EHPAD, maltraitance, personnes âgées

RESUME :

Contexte : Face au vieillissement de la population, de plus en plus de personnes âgées vivent institutionnalisées en EHPAD. Certaines familles, ayant à cœur d'offrir une bonne qualité de vie à leurs parents, se tournent vers des entreprises privées, dont ORPEA représente le leader mondial. En janvier 2022, celle-ci s'est retrouvée au cœur d'un scandale après la publication d'un livre qui dénonce les maltraitances subies par les personnes âgées au sein de l'entreprise. Le scandale a été largement suivi sur les réseaux sociaux, dont Twitter.

Objectif : Analyser les réactions de la population via Twitter en lien avec le scandale ORPEA.

Méthodes : Une méthode de *web scraping* a été appliquée pour extraire tous les tweets contenant au moins un mot-clé relatif au scandale. Ensuite, une analyse par *topic modeling* a permis d'identifier les principaux thèmes. En complément, une analyse des comptes tweetant sur le scandale a été menée, avec une attention spéciale portée sur le compte ORPEA.

Résultats : Parmi les 85 342 tweets uniques extraits, un pic de publications est identifié à la sortie du livre, puis lors de l'ouverture d'une enquête par les autorités françaises, ainsi qu'au moment de la diffusion de l'émission *Cash Investigation* sur les EHPAD du groupe privé Korian. Les six principaux thèmes abordés portaient sur (1) le fonctionnement des EHPAD, (2) l'impact du scandale sur le groupe Korian, (3) la dénonciation des abus sur les personnes âgées, (4) la médiatisation du scandale, (5) la volonté d'une action politique, (6) les opinions et témoignages individuels. Les profils s'exprimant sur le sujet étaient surtout des comptes non vérifiés, ou des médias. Le compte ORPEA a présenté une longue période d'inactivité pendant le scandale, puis revient avec la volonté de fédérer résidents et employés, et de transformer l'image du groupe.

Discussion : Le scandale ORPEA, en faisant réagir la population sur le plan politique, économique et humain, s'inscrit dans un débat sociétal large. Les réseaux sociaux ont permis la médiatisation du scandale et de ses conséquences sur le groupe, ainsi que le développement d'organisations destinées aux aidants. Cependant, plus d'un an et demi après le scandale, peu de changements semblent avoir été opérés dans la gestion de la personne âgée institutionnalisée en France.

Composition du Jury :

Président : M. Le Professeur Éric BOULANGER

Assesseurs : M. Le Professeur Jean-Baptiste BEUSCART, M. Le Docteur Antoine LAMER, Mme Le Docteur Clémentine ROLLAND.

Directeur de thèse : M. Le Professeur Jean-Baptiste BEUSCART