



**HAL**  
open science

# Étudier et interpréter les tweets des voyageurs pour améliorer la communication de la SNCF. Retour critique sur la méthodologie et les outils d'analyse.

Agata Jackiewicz, Luce Lefevre

## ► To cite this version:

Agata Jackiewicz, Luce Lefevre. Étudier et interpréter les tweets des voyageurs pour améliorer la communication de la SNCF. Retour critique sur la méthodologie et les outils d'analyse.. Analyser les réseaux sociaux numériques. Questionner les méthodes, LERASS; CREM; MICA; Communications et Sociétés, Mar 2024, Toulouse, France. hal-04527445

**HAL Id: hal-04527445**

**<https://hal.science/hal-04527445v1>**

Submitted on 30 Mar 2024

**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.

**Etudier et interpréter les tweets des voyageurs pour améliorer la communication de la SNCF.  
Retour critique sur la méthodologie et les outils d'analyse.**

Agata Jackiewicz, ReSO, Université Paul-Valéry Montpellier,

Luce Lefeuvre, Direction Technologies, Innovation & Projets Groupe, SNCF.

[agata.jackiewicz@univ-montp3.fr](mailto:agata.jackiewicz@univ-montp3.fr), [luce.lefeuvre@sncf.fr](mailto:luce.lefeuvre@sncf.fr)

**Mots clés :** Information Voyageur, Twitter/X, situation perturbée, modèle, analyse linguistique

# Étudier et interpréter les tweets des voyageurs pour améliorer la communication de la SNCF.

## Retour critique sur la méthodologie et les outils d'analyse.

Pannes techniques, grèves, travaux, phénomènes météorologiques extrêmes ou, plus prosaïquement, « absence inopinée d'un agent », représentent autant de sources d'incidents qui affectent la circulation des voyageurs et provoquent leur insatisfaction. En situation perturbée, la communication à l'attention des usagers constitue un enjeu essentiel pour les acteurs des transports publics : panneaux d'affichage, annonces sonores, applications dédiées, information directe par les agents en gare ou à bord du train, ou encore réseaux sociaux, représentent autant de moyens pour transmettre l'information attendue.

Le programme FIRST, mené par SNCF, se donne pour objectif d'améliorer l'expérience des voyageurs, en travaillant notamment sur la mise en cohérence de l'information diffusée par les différents médias d'information, en tenant compte en particulier de la pertinence des motifs de perturbation communiqués par SNCF, et de l'expression de la situation vécue et des besoins par les voyageurs. Les observables sont les verbatims des usagers postés sur le réseau Twitter/X. Le corpus d'analyse est formé de tweets collectés de janvier à avril 2019, classés suivant le motif de perturbation auquel ils font référence (« accident de personne », « régulation du trafic »...).

 @SNCF ce matin j'ai entendu une nouvelle excuse pour un retard..."Sortie tardive du dépôt» !!! À priori c'est comme les "préparations plus longues» 😊 Chez moi on appelle ça une panne de réveil... 🤔 🙄

 @Co\_ParisChateau Pour info ce matin un coup la SNCF dit qu'un train est en panne et deux minutes après une autre personne dit dans le micro que c'est une panne de signalisation 😊 Faudrai que la SNCF soit cohérente dans ses propos 😊 Cc @LIGNEP\_SNCF

Les traitements visent en particulier l'identification et la caractérisation des cibles incriminées dans l'Information Voyageur (IV) et des qualités s'y rapportant, ainsi que la détection et la classification (suivant leur valeur, polarité et force) des attitudes, des jugements et des ressentis exprimés par les usagers. Sur le plan linguistique, il s'agit de cerner de quelle manière les différents types de composants et de signes (linguistiques, graphiques, technolangagiers...) contribuent, par leurs agencements spécifiques, à rendre compte de l'expérience et des ressentis des voyageurs.

Le projet a donné lieu à la conception d'une méthode d'analyse, basée sur un modèle de communication original et un ensemble de ressources linguistiques afférentes. Un tableau de codage impliquant plusieurs dizaines de critères (55 colonnes), appliqués de manière automatique à une sélection de 826 tweets, a permis de mettre en évidence - avec des degrés de certitude variables - plusieurs tendances significatives. L'étude a livré des critères d'acceptabilité des motifs diffusés, ainsi que des pistes pour une typologie des attentes relativement à l'information diffusée.

Le fait de devoir répondre à une problématique concrète et aiguë d'amélioration de la communication entre fournisseur de services et clients amène à une étude opératoire des écrits de Twitter, dans toutes leurs dimensions, ce qui fait en même temps progresser la connaissance des usages de ce réseau social. Par ailleurs, dans le contexte de l'écoute client, d'autres études internes au groupe SNCF sont menées à partir des verbatims de Twitter/X, ce qui pose la question de la réutilisation de la méthode établie.

L'objet de cette communication est de présenter un retour critique sur l'ensemble du parcours d'étude, afin d'éclairer les apports et les limites des traitements manuels et automatiques proposés, compte tenu de plusieurs types de facteurs linguistiques et extra-linguistiques : sous-détermination (situation vs information sur la situation), temporalité, registres et modes d'écriture, rapports texte-image, ambiguïtés, implicite, humour, ironie.

## Références

- Caillat D., Jackiewicz, A., Lefevre, L. (2021). « Quelle perception de l'Information Voyageurs en situation perturbée ? Le ressenti des voyageurs à la lumière de la linguistique de corpus et de l'analyse du discours », *3èmes Rencontres Francophones Transport Mobilité, session De la collecte à l'analyse des données textuelles pour l'étude des mobilités quotidiennes*, 2- 4 juin 2021, Marne-la-Vallée (France).
- Denis, J., & Pontille, D. (2010). « Information voyageurs : inventer une signalétique collective à l'ère du 2.0. », *Ville, Rail & Transport*, n° 494, pp. 40-46, [En ligne] URL : <https://halshs.archives-ouvertes.fr/halshs-00533475/>
- Jackiewicz, A. (2016). « Matérialité linguistique des controverses sociétales. Rapports intersubjectifs et interdiscursifs dans des tweets polémiques », actes du *5<sup>e</sup> Congrès Mondial de Linguistique Française* [En ligne], SHS Web of Conferences, vol. 27, DOI 10.1051/shsconf/20162702008.
- Jackiewicz, A. (2017). « 24 heures dans la vie d'une controverse. Rapports texte – image dans des tweets polémiques ». A. Mercier & N. Pignard-Cheynel (dirs) *#info : Commenter et partager l'actualité sur Twitter et Facebook*, Éditions de la Maison des sciences de l'homme, 2017, 269-293.
- Jackiewicz, A., Caillat D., Lefevre, L. (2022). « LES eKiPes SnCF et rTE mEttENT toUt eN oEUvrE pOuR ReTAbLir LeLEcTrIcITÉ. Quand la matérialité des tweets informe sur la réception de l'information voyageur », *8e colloque international et pluridisciplinaire Ci-Dit, Matérialités du Discours Rapporté*, Luxembourg, Juin 2022.
- Marty, J. M., Wenzek, G., Schmitt, E., & Coulmance, J. (2015), « Analyse d'opinions de tweets par réseaux de neurones convolutionnels », DEFT 2015 [En ligne], Actes de l'atelier *Défi Fouille de Texte*, URL : [https://deft.limsi.fr/actes/2015/pdf/marty-2015deft\\_article.pdf](https://deft.limsi.fr/actes/2015/pdf/marty-2015deft_article.pdf).
- Paroubek, P., Grouin, C., Bellot, P., Claveau, V., Eshkol-Taravella, I., et al. (2018). « Recherche d'information et analyse de sentiments dans des tweets concernant les transports en Île de France », DEFT 2018 [En ligne], Actes de l'atelier Défi Fouille de Texte, in TALN 2018, Actes de la conférence Traitement Automatique des Langues, URL : [https://deft.limsi.fr/actes/Actes\\_TALN2018\\_vol2.pdf](https://deft.limsi.fr/actes/Actes_TALN2018_vol2.pdf)
- Pinède, N., Masso, u L., Mpondo-Dicka, P. (dir.), (2023). *Présence numérique des organisations*, ISTE Éditions.
- Vidak, M., & Jackiewicz, A. (2016). « Les outils multimodaux de Twitter comme moyens d'expression des émotions et des prises de position », *Cahiers de praxématique* [En ligne] n° 66, URL : <http://journals.openedition.org/praxematique/4247>.