

Estado da publicação: O preprint não foi publicado em outro meio.

Réseaux socio-numériques et frontières du lire/écrire dans la modération des commentaires

Aude Seurrat, María Inés Laitano

<https://doi.org/10.1590/SciELOPreprints.14015>

Submetido em: 2025-11-07

Postado em: 2025-11-12 (versão 1)

(AAAA-MM-DD)

Research Report

Social Media and Boundaries of Reading/Writing in Comment Moderation

Aude Seurrat¹ & María Inés Laitano²

1. *Professor in Information and Communication Sciences (PhD) Co-director of the Céditec Research Center. Université Paris Est Créteil, 61 Av. du Général de Gaulle Centre, 94000 Créteil, France. aude.seurrat@u-pec.fr
<https://orcid.org/0000-0001-6906-7471>.*
2. *Associate Professor in Information and Communication Sciences (PhD). Researcher at Paragraphe Lab. Université Paris 8, 2 rue de la Liberté, 93526 Saint-Denis, France. mlaitano02@univ-paris8.fr
<https://orcid.org/0000-0002-3794-7383>.*

Keywords

social media, content moderation, semiotics of screen writings, traces

Abstract

This article analyzes how social media shape the boundaries of reading and writing through the moderation of political comments. The authors compare practices across X, YouTube, Twitch, and Instagram, emphasizing the materiality of the platforms and their hybrid logics that combine algorithmic automation with human intervention. Moderation is approached as an active orchestration of reading and writing, influencing the visibility and legitimacy of public expression while modulating the accessibility and ordering of comments. The authors highlight the methodological challenges inherent in studying moderation practices and propose linking a semiotics of screen writings to interpret the materiality of traces left by moderation processes,

emphasizing their fragmentary nature. Three case studies illustrate how the disappearance, ranking, or modulation of comments question the visibility and legitimacy of political discourse online.

Rapport de recherche

Réseaux socio-numériques et frontières du lire/écrire dans la modération des commentaires

Aude Seurra¹ & María Inés Laitano²

3. *Professeure en sciences de l'information et de la communication, co-directrice du laboratoire Céditec. Université Paris Est Créteil. 61 Av. du Général de Gaulle Centre, 94000 Créteil, France. aude.seurra@u-pec.fr
<https://orcid.org/0000-0001-6906-7471>.*
4. *Maîtresse de conférences en Sciences de l'information et de la communication. Chercheuse au Laboratoire Paragraphe. Université Paris 8. 2 rue de la Liberté, 93526 Saint-Denis, France. mlaitano02@univ-paris8.fr
<https://orcid.org/0000-0002-3794-7383>.*

Mots clés

réseaux socio-numériques, modération de contenu, sémiotique des écrits d'écran, traces

Résumé

Cet article analyse comment les réseaux socio-numériques façonnent les frontières du lire-écrire à travers les pratiques de modération des commentaires politiques. Les auteurs s'appuient sur une comparaison de X, YouTube, Twitch et Instagram, mettant en avant la matérialité des dispositifs et leurs logiques hybrides, mêlant automatisation algorithmique et intervention humaine. La modération est envisagée

comme une orchestration active de la lecture et de l'écriture, jouant sur la visibilité et la légitimité des prises de parole, tout en modulant l'accès et l'ordre d'apparition des commentaires. Les auteurs soulignent que l'analyse des pratiques de modération se heurte à de nombreux obstacles méthodologiques et proposent d'articuler une sémiotique des écrits d'écran permettant d'interpréter la matérialité des traces laissées par les modérations, insistant sur leur aspect fragmentaire. Trois cas d'étude illustrent comment la disparition, la hiérarchisation ou la modulation des commentaires interrogent la visibilité et la légitimité des discours politiques en ligne.

Introduction

Cet article propose une réflexion sur les modalités selon lesquelles les réseaux socio-numériques tracent les frontières du lire/écrire dans les commentaires en ligne, à partir d'une analyse menée dans le cadre du projet ANR AMULEX¹. Le propos s'ancre dans l'étude des commentaires politiques générés sur quatre réseaux socio-numériques durant les élections présidentielles françaises de 2022, pour interroger les dynamiques discursives et les obstacles méthodologiques soulevés par l'analyse des pratiques de modération.

La recherche menée dans le cadre du projet AMULEX se consacre à l'analyse des commentaires politiques en ligne lors des élections présidentielles françaises de 2022. Le terrain de recherche couvre des corpus issus de X, YouTube, Twitch et Instagram, permettant d'interroger la matérialité complexe des pratiques de modération en prenant en compte la spécificité de chacun de ces dispositifs socio-techniques. Cette problématique invite à une analyse fine de la médiativité des dispositifs (Marion, 1997), soit la capacité de chaque réseau à singulariser ses cadres sociaux et techniques d'expression politique. Une attention particulière est accordée aux expressions ordinaires qui se présentent dans l'espace public numérique. Plus précisément, le projet s'intéresse aux réactions, dont le commentaire constitue la forme la plus élaborée, en tant que manifestation des expressions sur les réseaux sociaux numériques.

¹ Un projet ANR désigne un projet financé par l'Agence nationale de la recherche en France. L'acronyme AMULEX renvoie à Analyse multiplateforme de l'expression politique en ligne.

Les commentaires ont été collectés en 2022 dans des espaces numériques administrés par des partis politiques ou des médias. Sont ainsi recueillis non seulement les textes, mais aussi les images, les signes paralinguistiques et un ensemble de métadonnées, avec des variables qui diffèrent selon les dispositifs et les outils de recueil utilisés. L'ensemble de ces données est réuni dans une base de données.

La première tâche du projet AMULEX porte sur une analyse sémiotique comparative des dispositifs socio-techniques et de leurs espaces de commentaires. L'analyse vise à saisir les variations sémiotiques et discursives induites par l'architecture, la gouvernance et les espaces de publication de chacun de ces dispositifs socio-techniques. Dans les réseaux socio-numériques, les pratiques de modération ne se réduisent pas à la simple gestion des discours haineux ou de la désinformation (Badouard, 2020). Elles façonnent en profondeur l'espace du possible, règlent la visibilité des contributions, orchestrent l'expression politique et conditionnent la circulation des arguments.

L'article issu du projet repose sur l'analyse des discours d'escorte (Souchier et al., 2019) produits autour de ces quatre dispositifs numériques et articule cette réflexion à l'analyse d'extraits de notre corpus. Ce texte poursuit un double objectif. Premièrement, il s'agit de démontrer l'intérêt heuristique de concevoir la modération comme une orchestration de l'écriture et de la lecture, en s'appuyant sur une sémiotique des écrits d'écran dans la suite des travaux de Christin (1995), Jeanneret (2006, 2019) et Souchier (1996; 2019). Deuxièmement, il s'agit d'explicitier les difficultés méthodologiques spécifiques rencontrées dans le projet lorsqu'il s'agit d'observer et d'analyser ces formes plurielles de la modération, tout en proposant des pistes fondées sur l'interprétation de certaines traces (Jeanneret, 2019) repérées dans les corpus.

La première partie de l'article propose d'examiner l'intérêt heuristique que représente le fait de considérer la modération comme une orchestration des activités de lecture et d'écriture sur les réseaux socio-numériques, en mobilisant une sémiotique des écrits d'écran. Elle s'ouvre sur une analyse des formes et des enjeux contemporains de la modération, puis s'attache à la diversité des dispositifs, à la multiplicité des logiques en présence et aux dynamiques hybrides, humaines et non

humaines, qui façonnent l'expérience de la modération. Enfin, cette partie développe la proposition centrale de l'article : appréhender la modération comme un principe organisateur de l'écriture et de la lecture des commentaires en ligne.

La seconde partie est consacrée aux défis méthodologiques posés par l'observation et l'interprétation des formes de modération et des traces qui en subsistent dans les corpus. Elle présente tout d'abord une réflexion sur les difficultés spécifiques rencontrées au fil du projet, avant de s'attarder sur des exemples d'analyses de traces et sur les enjeux interprétatifs qui les accompagnent.

1. Formes et enjeux de la modération des commentaires sur les réseaux socio-numériques

1.1. Logique hybride et gouvernementalité du dispositif

Dans le cadre du projet ANR AMULEX, qui analyse les commentaires politiques en ligne produits lors de l'élection présidentielle française de 2022 sur quatre réseaux sociaux numériques (X, YouTube, Twitch et Instagram), nous proposons de saisir la modération comme une orchestration complexe des activités de lecture et d'écriture. La modération ne se limite pas à un simple filtrage de contenus problématiques, mais contribue à configurer les espaces discursifs numériques, affectant la visibilité, la légitimité et les formes d'expression possibles des internautes.

La régulation de ces espaces relève d'abord d'un cadre externe, juridique et politique, défini par des normes telles que les lois nationales ou les directives européennes tel que le Digital Services Act (DSA) ou le Règlement Général sur la Protection des Données (RGPD). Toutefois, comme le souligne Dominique Boullier ([2024](#)), ces régulations étatiques apparaissent limitées face à la puissance technique et économique des plateformes, qui redessinent les contours de ce qui est légitime et exprimable sur leurs dispositifs.

En parallèle, la modération, définie comme l'ensemble des opérations manuelles ou automatisées encadrant la publication et la visibilité des contenus, se déploie selon une logique hybride. Celle-ci combine l'automatisation algorithmique pour filtrer

massivement des contenus – discours haineux, désinformation, contenus violents – et l'intervention humaine dans les cas les plus complexes (Badouard, 2020; Gillespie, 2018). Par exemple, selon Romain Badouard (2020), 97% des discours haineux supprimés par Meta en 2022 l'ont été via détection algorithmique avant tout signalement manuel.

Dans le projet AMULEX, cette hybridation reflète les dynamiques propres à chaque dispositif étudié, posant des défis spécifiques quant à l'interprétation des commentaires et à l'observation des pratiques de modération. Notamment, l'impossible accès au fonctionnement de ces algorithmes, classés souvent comme « propriété intellectuelle », empêche d'analyser le fonctionnement spécifique de chaque réseau socio-numérique afin d'analyser son incidence sur les modalités de modération.

L'automatisation de la modération produit des effets paradoxaux et problématiques. Thibault Grison démontre dans sa thèse que cette délégation aux algorithmes engendre des régimes de censure insidieux et des discriminations structurelles (Grison, 2024). Certains groupes militants, notamment antiracistes, féministes ou LGBT+, sont sur-modérés car leurs usages de termes stigmatisants à des fins revendicatives sont mal interprétés par les algorithmes et peuvent conduire à l'invisibilisation ou au *shadowbanning*. Par ailleurs, la priorité économique portée à l'engagement pousse des plateformes à amplifier des contenus polémiques, parfois extrémistes, comme l'a révélé Twitter (maintenant X) en 2021 en favorisant certains discours de droite (Grison, 2024).

Ainsi, ces dispositifs numériques agissent comme de véritables « dispositifs de gouvernementalité » (Foucault, 2004), où la frontière du dicible est modulée de manière opacifiée à travers des scripts algorithmiques et la paramétrisation de filtres, masquant leurs choix politiques et sociaux derrière une prétendue neutralité technologique (Boullier, 2024; Grison et al., 2023). Le concept de gouvernementalité élaboré par Michel Foucault (2004) s'avère particulièrement heuristique pour penser la modération sur les réseaux sociaux numériques. En effet, la gouvernementalité ne désigne pas simplement un mode d'exercice vertical ou autoritaire du pouvoir, mais l'ensemble des procédés, techniques et savoirs qui visent à orienter, organiser et

configurer les conduites à distance, souvent sans coercition directe, en incitant, balisant et rendant certaines actions plus probables que d'autres.

Appliqué à la modération sur les réseaux socio-numériques, ce cadre permet d'analyser comment les architectures techniques, les algorithmes et les chartes de conduite agissent comme des dispositifs de pouvoir diffus, structurant de manière anticipative et granulaire les espaces du dicible, la visibilité des contenus et la circulation des discours. Ces dispositifs ne se contentent pas de censurer ou d'autoriser, ils orchestrent subtilement les flux, hiérarchisent les interactions et paramètrent les comportements attendus, déplaçant ainsi le centre du pouvoir de la contrainte punitive vers une modulation invisible qui conduit les conduites des usagers. Mobiliser la gouvernementalité foucauldienne éclaire ainsi la spécificité de la « gouvernance algorithmique » à l'œuvre dans la modération contemporaine, où réguler, normaliser et conduire s'articulent à travers des techniques de contrôle souvent imperceptibles mais décisives pour l'agencement du débat public en ligne.

Dans le cadre du projet AMULEX, cette problématique permet de comprendre comment les espaces de commentaires, en tant que traces écrites et visuelles, sont à la fois produits, encadrés et circonscrits selon des logiques socio-techniques spécifiques.

1.2. Diversité des dispositifs de modération

Le projet AMULEX met en lumière la multiplicité des dispositifs socio-techniques qui encadrent les commentaires en ligne lors des élections présidentielles françaises de 2022, avec une attention particulière sur les spécificités propres à chaque réseau socio-numérique étudié. L'étude comparative des espaces de commentaires sur X, YouTube, Twitch et Instagram révèle que les pratiques de modération diffèrent sensiblement en fonction de l'architecture, des règles internes, de l'offre de fonctionnalités pour l'écriture (longueur autorisée, inclusion de liens, images, emojis) et des pouvoirs d'intervention des acteurs humains et automatiques.

Le contrôle des contenus problématiques n'est donc pas uniforme, mais articulé selon des logiques hybrides où l'algorithme assure la détection en masse et les humains interviennent sur les cas ambigus, contestés ou soulevant des enjeux de

contextualisation. Cette hybridation nécessite une analyse sémiotique fine des traces laissées par la modération, que ce soit dans les contenus autorisés, supprimés, ou dans la visibilité fluctuante des commentaires (Badouard, 2020).

Dans son ouvrage *Custodians of the Internet*, Tarleton Gillespie (2018) propose une typologie des formes de modération déployées sur les plateformes numériques, qui dépasse largement la simple idée de suppression de contenus problématiques. Cette approche éclaire la complexité des interventions qui façonnent la visibilité et la reproductibilité des écrits d'écran dans les espaces socio-numériques, tels que ceux analysés dans le projet AMULEX.

Il distingue d'abord la modération réactive, basée sur la participation des utilisateurs par le biais de signalements de contenus problématiques. Ces signalements initient un processus collaboratif où une équipe ou un dispositif technique traite les contenus concernés. L'auteur souligne que ce mode de gestion, présenté comme participatif, masque souvent une centralisation du pouvoir décisionnel, lequel demeure largement discret et opaque.

Pour Gillespie, dans la modération distribuée, la gestion des contenus s'effectue au sein même de la communauté d'utilisateurs, grâce à des mécanismes d'évaluation collective, de votes ou de peer-review qui favorisent l'auto-organisation et structurent la visibilité des commentaires. Cette organisation est notamment adoptée par les plateformes qui misent sur la dynamique communautaire.

Gillespie insiste enfin sur le développement de la modération automatisée : l'usage croissant d'outils algorithmiques permet de filtrer, référencer ou déréférencer de façon proactive les contenus considérés comme sensibles ou illicites. Enfin, l'auteur développe la notion de modération hybride, qui désigne l'articulation de plusieurs modalités selon les contextes, les besoins et les enjeux propres à chaque plateforme : ce régime mixte associe pré- et post-modération, intervention humaine et automatisée, ainsi que formes communautaires de gestion des contenus, produisant ainsi une gouvernance originale, complexe, et spécifique à chaque dispositif socio-technique.

Au-delà de ces catégories, Gillespie insiste sur la diversité des actions qui constituent les pratiques de modération : suppression de contenus, mais aussi

invisibilisation discrète (shadowban), déréférencement, ajout d'avertissements, filtrage de comptes, ou encore gestion géographique de la visibilité des publications. Cette granularité d'actions éloigne la modération d'une simple binarité censure/non-censure et invite à considérer la modulation des espaces discursifs comme une orchestration sémiotique à multiples facettes.

S'ajoute à cette diversité des modalités de modération, l'évolution rapide de celles-ci. En témoignent les travaux de Van Dijck et Poell (2013), qui proposent d'analyser les plateformes comme des architectures évolutives intégrant des protocoles de gouvernance, où l'ajustement constant des scripts et des normes participe à l'élaboration des régimes de modération numérique. Par ailleurs, ils soulignent la montée en puissance de dispositifs de modération dits « ex ante » dans le sillage du Digital Services Act européen, introduisant une logique de proportionnalité et de responsabilisation.

Le projet AMULEX, par son analyse comparative des espaces de commentaires sur X, YouTube, Twitch et Instagram, contribue à mettre en lumière cette diversité des régimes de modération et invite à comprendre comment ces modalités spécifiques structurent le lire/écrire en ligne. Ces pratiques définissent des frontières de ce qui peut être lu et écrit, participant ainsi à la construction des débats politiques numériques et influent sur la dynamique argumentative et la circulation des opinions politiques.

1.3. La modération comme principe organisateur du lire/écrire des commentaires en ligne

Une des contributions du projet AMULEX consiste à articuler une sémiotique des écrits d'écran à l'analyse des processus de modération, en conceptualisant cette dernière comme une orchestration des activités simultanées d'écriture et de lecture des commentaires. Cette approche permet de comprendre la modération non pas comme un simple filtre de contenus, mais comme une intervention structurante de la co-construction des discours politiques dans l'espace numérique.

Cette perspective s'appuie sur des travaux en sciences de l'information et de la communication qui insistent sur la matérialité des dispositifs socio-techniques et

leurs configurations sémiotiques, à l'instar des analyses de Jeanneret (2014) et de Souchier et al. (2019). Elle articule trois niveaux : la prétention des plateformes, c'est-à-dire leurs discours d'escorte d'auto-présentation et leurs règles ; leur prédilection, soit les formes d'expression qu'elles favorisent ou restreignent ; et l'implication, qui se manifeste dans les usages effectifs et la normativité interne des interactions (Jeanneret, 2014).

La modération, envisagée comme un « régime d'écriture », encadre ce que les utilisateurs peuvent produire ou lire et influence la nature des commentaires exprimés, prescrivant ou contraignant les formes argumentatives et discursives en contexte électoral. Ainsi, la comparaison des dispositifs permet de discerner comment leurs architectures spécifiques orientent l'émergence ou la restriction de certaines prises de parole. Lecture et écriture sont deux moments indissociables de l'expérience numérique : qui peut exprimer un commentaire, et qui peut le lire, voir, partager. Les plateformes permettent filtration, masquage, épingleage, et la différence entre ce que voit l'auteur et ce que perçoit le public exemplifie la performativité du dispositif. La notion d'architexte « désigne les dispositifs d'écriture de l'écriture et permet ainsi de rendre compte des strates de signification qui président à toute production d'un contenu numérique. Nous entrerons dans la question du code et de sa capacité à normer la communication informatisée en prenant toutes distances à l'égard des discours reçus sur le numérique. L'analyse des outils de production du texte nous permettra en revanche de comprendre quelles sont les structures à l'intérieur desquelles se produisent les textes dans le cadre d'une économie capitaliste qui à la fois détermine de plus en plus la communication et impose ses modèles d'affaires » (Souchier et al., 2019, p. 159).

La théorie des écrits d'écran de Souchier, Jeanneret et Le Marec (2003) met l'accent sur la spécificité matérielle et sémiotique de la production numérique. Écrire sur écran, c'est nécessairement inscrire son message dans un dispositif qui balise ses formes et oriente sa réception. La perspective techno-sémiotique portée par les auteurs incite alors à ne jamais séparer strictement lecture et écriture : ils proposent la notion d'« écriture-lecture », soulignant que toute participation numérique combine une production de sens par l'écriture et une configuration pour autrui par la lecture organisée. Modérer la lecture (par exemple via un masquage ou une hiérarchisation

algorithmique) ou l'écriture (par la restriction d'accès ou la suspension de compte) relève du même geste de reconfiguration du champ de visibilité et du débat. Cette perspective met en avant le rôle actif des dispositifs : la plateforme ne se contente pas de médiatiser la parole, elle redéfinit ce qui est lisible, dicible, échangeable dans l'espace numérique.

Grison (2024) analyse ainsi comment les chartes de modération ne se contentent pas de formuler des interdits (par exemple « proscrire le discours haineux »), mais organisent un contrôle graduel et procédural de la parole allant du filtrage algorithmique à la sanction humaine. La pluralité des pratiques de « faire taire » n'équivaut pas à une simple suppression du propos : elle recouvre l'amputation, le filtrage ou le remodelage de l'écrit, la charte faisant alors office d'arrangeur qui définit explicitement ou implicitement qui peut s'exprimer, selon quel mode et selon quelles conventions. Les discours d'escorte (règles, indications d'usage, chartes) jouent un rôle décisif : ils encadrent les permissions et prescrivent les conventions discursives propres à chaque plateforme. Pour comprendre la granularité des régulations, il faut s'intéresser à la matérialité des interventions (*shadowban*, épingleage, masquage, déréférencement automatique), aux stratégies de contournement (usage d'emojis, l33tspeak, ironie), et à la performativité communautaire (signalements massifs, mobilisation collective).

C'est dans ce cadre que les tableaux d'analyse comparative du projet AMULEX ont été élaborés. Chaque tableau opère une focale sur le positionnement des différents rôles dans la gestion de la lecture et de l'écriture de commentaires : l'entreprise (le réseau socio-numérique), le public, et le propriétaire de la publication (média ou parti politique). Pour chaque dispositif (Instagram, YouTube, Twitch, X/Twitter), il s'agit d'identifier les manières dont les différents rôles interviennent sur la lecture/écriture des commentaires, en décidant ce qui peut être lu, qui peut écrire et ce qui peut l'être.

Les tableaux 1 et 2 offrent une structuration des conditions et modalités permettant ou entravant la participation : l'écriture désigne la possibilité de prendre la parole, les restrictions d'accès (comme la publication par les seuls abonnés ou comptes vérifiés, la suspension du droit de commenter ou d'accéder à certains formats, etc.), tandis que la lecture renvoie à l'accès, la visibilité et l'ordre d'apparition des

commentaires pour chaque internaute (hiérarchisation par l'algorithme, masquage, suppression, etc.). Les tableaux proposés ci-dessous mettent en évidence ces processus : par exemple, sur Instagram, le propriétaire de la publication peut limiter l'accès aux followers et bloquer certains comptes, tandis que sur Twitch, la gestion des commentaires peut intégrer modération par bots, mode lent, ou restriction à certains types de comptes. Les entreprises de plateformes assurent en parallèle la gestion des signalements, la sanction automatisée et la modulation algorithmique de la visibilité. La capacité à commenter ou à lire est donc une variable dépendante de l'interaction entre logiques de l'interface, règles communautaires et intentions politiques des dispositifs.

L'analyse comparative met aussi en évidence que certains dispositifs d'intervention sont communs à tous les réseaux (masquage par le propriétaire, signalement par le public), tandis que d'autres relèvent de logiques d'architecture propres à chaque plateforme. Ainsi, sur Instagram et YouTube, plateformes de la même catégorie asynchrone « vidéo/commentaires », se déploient des options comme l'épinglage d'un commentaire, marquant une hiérarchisation explicite du discours, ou la désactivation des commentaires. Par contraste, sur Twitch, la diffusion en direct et la temporalité synchrone imposent un régime de modération orienté vers l'immédiateté : la plateforme ne peut pas trier les messages autrement que chronologiquement, et propose des outils automatisés (bots, mode lent) pour intervenir. Ces différences ne sont pas que techniques, elles traduisent des formes de gouvernementalité numérique à l'œuvre dans la modération des discours.

Tableau 1 : Types d'interventions sur la lecture des commentaires

Rôle	Instagram	YouTube	X/Twitter	Twitch
Propriétaire de la publication	Masquage automatique à partir de mots/expressions. Masquage ou suppression manuelle.			
	Épinglage			Masquage via bots.

		Approbation manuelle.		
Public	Signalement de commentaire. Contestation d'une décision.			
Entreprise RSN	Recommandation des commentaires. Gestion des signalements.		Gestion des signalements.	

Tableau 2 : Types d'interventions sur l'écriture des commentaires

Rôle	Instagram	YouTube	X/Twitter	Twitch
Propriétaire de la publication	Limitation aux abonnés/comptes suivis, comptes vérifiés. Signalement, restriction, bannissement ou blocage d'un compte.			
	Désactivation des commentaires.		Limitation aux comptes mentionnés.	Limitation aux abonnés depuis x temps Mode lent/émoticône. Empêchement temporaire d'un compte.
Public	Signalement de compte. Contestation d'une décision.			
Entreprise RSN	Recommandation ou invisibilisation de la publication. Sanctions (suspension, clôture, restriction, etc.). Gestion des signalements.			

L'approche techno-sémiotique permet ainsi de dépasser la binarité suppression/autorisation en analysant la manière dont le socle technique balise, module et hiérarchise effectivement la circulation des discours politiques ou militants sur les réseaux. Le rapport aux « actions » (masquage, signalement, épingleage, déréférencement, action communautaire) évolue selon la plateforme, le contexte et la dynamique propre à chaque communauté.

La question de “qui peut écrire” sur une plateforme dépend du système d'accès et des règles d'authentification : les administrateurs, les membres authentifiés ou parfois tout visiteur disposent de droits différenciés pour publier des contenus. Ce qui peut être écrit est conditionné par les scripts, les filtres et les formats imposés par l'architexte de la plateforme, qui encadrent à la fois la nature des contenus (texte, image, vidéo) et leur conformité aux règles communautaires ou légales (blocage de mots-clés, interdits juridiques). Qui peut lire les contenus est déterminé par des règles de visibilité, de confidentialité ou d'abonnement : un message peut être public, réservé à une communauté fermée ou visible seulement par certains groupes désignés par le modérateur ou l'algorithme. Ce qui peut être lu provient d'une sélection technique et algorithmique, qui filtre, hiérarchise, ou occulte certains messages : ainsi, la plateforme construit un espace du lisible par le masquage, l'épingleage ou la hiérarchisation algorithmique. Enfin, l'ensemble de ces médiations manifeste le pouvoir structurant de l'architexte, qui module simultanément la distribution sociale des voix et la perception collective du débat en ligne.

La question de la temporalité (commentaires en direct/streaming contre publication différée), la matérialité de la trace (effacement, annotation, balisage) et la performativité (notamment signalement, engouement ou boycott communautaire) sont également importantes pour comprendre les frontières mouvantes du lisible et du dicible. L'agencement des actions – du masquage au signalement, de l'épingleage au déréférencement – varie selon l'architecture technique, la logique algorithmique et la dynamique communautaire. Jamais espace neutre ni transparent, l'univers des commentaires porte la marque constante des stratégies de contournement et d'appropriation sémiotique, qui constituent un enjeu démocratique posé par la modération algorithmique contemporaine

La première partie de cet article a montré comment la modération, en organisant le lire/écrire sur les plateformes socio-numériques, restructure les frontières du dicible, module la visibilité des discours et conditionne l'expression politique en ligne. Cette orchestration sémiotique, analysée à travers les logiques des architextes, révèle que les permissions d'écriture ou de lecture sont toujours médiées, filtrées et orientées par les architectures techniques, les scripts et les règles communautaires. Toutefois, interroger les effets concrets de ces dispositifs nous a posé d'importants défis méthodologiques, notamment pour identifier, interpréter et documenter la diversité des traces laissées par la modération — qu'il s'agisse de suppressions, d'occultations ou de modifications de visibilité. Accéder à la matérialité réelle de la modération suppose alors de croiser une pluralité de méthodes, pour saisir la dynamique, souvent fragmentaire, des interventions humaines et algorithmiques qui reconfigurent l'espace public numérique. La seconde partie se consacre donc à l'exploration de ces défis méthodologiques et aux enjeux d'interprétation liés à l'identification et à l'analyse des traces de la modération en ligne.

2. Défis méthodologiques et enjeux d'interprétation des traces de la modération

Une des contributions du projet AMULEX consiste à articuler une sémiotique des écrits d'écran à l'analyse des pratiques de modération, en conceptualisant cette dernière comme une orchestration des activités d'écriture et de lecture des commentaires. L'étude des pratiques de modération à partir des traces collectées a cependant révélé une série de défis méthodologiques et d'enjeux d'interprétation exposés ci-après. Ils ne prétendent pas pour autant être exhaustifs : ils doivent être compris comme situés, liés à la fois aux dispositifs étudiés, aux traces collectées, aux conditions d'observation qu'ils rendent possibles, ainsi qu'au stade d'avancement du projet de recherche AMULEX . Cette reconnaissance de partialité constitue le point de départ de cette seconde partie, consacrée à l'examen des principaux obstacles méthodologiques rencontrés au fil du projet et aux questions interprétatives ouvertes par l'analyse des traces.

2.1. Les défis méthodologiques repérés au fil du projet

L'analyse des interventions effectives sur la lecture et l'écriture des commentaires met en évidence des défis méthodologiques qui ne sont pas uniformes : ils varient aussi en fonction des rôles qui organisent la modération —le propriétaire de la publication, l'entreprise du réseau socio-numérique et le public. Chacun d'eux a des possibilités d'action sur le lire/écrire et produit simultanément une part d'opacité spécifique pour l'analyse. Structurer les difficultés méthodologiques selon ces rôles permet de saisir comment l'accès et la traçabilité des opérations de modération se trouvent différemment affectés, ainsi que les différents moyens de contournement auxquels ont recours les chercheurs qui analysent la modération.

Le premier ensemble de défis concerne la modération effectuée par les propriétaires de publications (médias et partis politiques dans le projet AMULEX). L'accès impossible aux paramètres et aux tableaux de bord de modération de ces acteurs constitue un angle mort majeur. Ni les interfaces publiques des plateformes ni les outils de collecte existants ne permettent de retracer, par exemple, les choix opérés en matière de filtrage automatique ou de validation humaine des commentaires. À cela s'ajoute l'impossibilité d'identifier les commentaires supprimés ou les comptes bloqués par ces propriétaires, nous y reviendrons. Par ailleurs, les règles encadrant la pratique de ces acteurs ne font presque jamais l'objet d'une communication publique quant aux critères d'intervention ou aux seuils de tolérance. Enfin, la dimension collective de la modération est, elle aussi, difficile à documenter : identifier les différents comptes intervenant au nom du propriétaire de la publication (community managers, modérateurs externes, bots) relève souvent de l'hypothèse interprétative.

Des travaux antérieurs ont tenté de contourner ces obstacles par des enquêtes auprès des acteurs impliqués dans le rôle de propriétaire de la publication. Smyrnaiois & Marty (2017), par exemple, ont mené une enquête par entretiens et observation auprès d'un des trois principaux fournisseurs de services de modération aux médias français, afin de reconstituer les logiques de leur pratique de modération.

Un second ensemble de défis tient à la modération opérée par les entreprises de réseaux socio-numériques. L'analyste se heurte ici de nouveau à l'impossibilité d'accéder directement aux commentaires supprimés ou masqués ainsi qu'aux comptes bloqués ou suspendus, qu'il s'agisse d'actions algorithmiques, de décisions manuelles des employés de ces entreprises ou de suites données à des signalements d'utilisateurs. Dans la majorité des cas, ces suppressions entraînent la disparition des éléments concernés, rendant toute collecte ou observation *a posteriori* impossible. Les bases de données et rapports de transparence exigés par le *Digital Services Act* européen ne fournissent dans ce sens que des statistiques agrégées, sans donner accès aux contenus supprimés, aux profils suspendus ni aux justifications associées.

Différentes stratégies méthodologiques ont été expérimentées dans la littérature pour identifier des commentaires disparus. La comparaison de différentes versions d'un même ensemble de commentaires dans le temps permet notamment d'identifier indirectement les commentaires supprimés (Grison et al., 2023; Jiang et al., 2019). Dans le cas de commentaires en temps réel, sur Twitch par exemple, l'ethnographie des chats en direct permet de repérer les effacements au moment même de leur survenue (Thach et al., 2022).

Enfin, les défis méthodologiques liés à la modération par le public tiennent à l'opacité du processus de signalement des plateformes étudiées. Les signalements individuels ou collectifs ne laissent généralement aucune trace visible, ce qui oblige le chercheur à les inférer indirectement à partir d'indices contextuels, tels que la dynamique interactionnelle dans laquelle ils apparaissent, les caractéristiques des comptes impliqués ou les délais entre publication et disparition d'un message (Grison et al., 2023). Des approches ethnographiques et de design ont par ailleurs permis d'éclairer les styles de modération exercés par les publics eux-mêmes, en montrant qu'ils constituent parfois un travail collectif de cadrage du dicible, distinct des interventions des propriétaires ou des plateformes (Meunier et al., 2024).

Ainsi, les obstacles méthodologiques identifiés ne relèvent pas seulement de difficultés de collecte ou d'accès : ils rappellent que toute observation de la modération est située dans un régime de visibilité produit par les rôles qui l'organisent. Or, ces régimes de visibilité conditionnent également ce que le

chercheur peut interpréter. Le passage des traces observables à des hypothèses sur les opérations de modération exige dès lors un travail d'interprétation : il consiste à restituer les effets de ces interventions dans les rapports de pouvoir, de légitimité et d'intentionnalité qui les accompagnent. C'est à ce déplacement – du repérage méthodologique vers l'interprétation des traces – que la section suivante est consacrée.

2.2. Analyse des traces et enjeux interprétatifs

Lorsqu'un commentaire est supprimé ou rendu invisible sur les réseaux sociaux numériques, il peut en rester des traces : notifications, ou simples vides ou incohérences perceptibles dans les enchaînements d'interactions. Pour Yves Jeanneret dans son dernier ouvrage *La fabrique de la trace* (2019), la trace est objet d'interprétation, jamais preuve auto-suffisante. Elle existe socialement parce qu'un ou des acteurs (chercheur, utilisateur, institution) choisissent d'y prêter attention, de la qualifier comme signifiant quelque chose. Dans ses travaux, Jeanneret insiste justement sur le caractère non-immédiat, non-transparent de la trace, qui n'est jamais donnée "en soi" mais toujours fabriquée, historicisée, investie par des schèmes sociaux et sémiotiques. Pour l'auteur, « l'évidence des traces masque la complexité des processus qui permettent de les produire, de les interpréter et de les utiliser. » (Jeanneret, 2019, p. 18)

Jeanneret évoque l'importance de la distance (l'inscription écrite, la médiation technique) qui fait que la trace n'est pas une empreinte brute du réel, mais le reste médiatisé, fragmentaire, sujet à l'interprétation, "énigme" pour le chercheur comme l'exprime l'approche peircienne : « Dans ce paradigme, la notion de trace se rapproche de l'idée d'énigme. C'est le manque de documentation explicite qui justifie le recours à la trace. » (Jeanneret, 2011, paragr. 28). Il cite à ce sujet l'idée d'un « paradigme indiciel » chez Carlo Ginzburg qui met l'accent sur le caractère éminemment interprétatif du travail de l'historien. Dans ce paradigme, la notion de trace rejoint celle d'énigme. C'est le manque de documentation explicite qui justifie le recours à la trace. Or, c'est bien ce manque de documentation que nous avons pointé ci-dessus dans le cadre du projet AMULEX.

Nous sommes dès lors à la recherche de ces indices fragmentaires que nous pouvons choisir d'interpréter comme traces — non dans une logique positiviste (appréhender "ce qui s'est vraiment passé") mais bien selon un schème socio-sémiotique (comment cette suppression s'inscrit-elle dans la dynamique sociale, informationnelle, politique de la plateforme ?). L'activité de qualification de la trace, pour Jeanneret, consiste alors à construire un regard analytique sur ces zones d'incertitude, sur ces "vides" signifiants, sur ces éléments qui font énigme par leur incomplétude. L'analyste dotera les traces d'un potentiel de sens particulier, que Jeanneret propose de spécifier « comme la capacité dans le présent de faire référence à un passé absent mais postulé » (Jeanneret, 2011, paragr. 4). Cette prise de conscience du décalage temporel dans l'analyse des traces revêt, pour l'étude de la modération en ligne notamment, un intérêt particulier, comme nous le verrons ci-dessous.

Nous présentons dans ce qui suit trois cas d'analyse de traces, sans prétention à l'exhaustivité mais dans la perspective d'éclairer notre conceptualisation de la modération comme orchestration des activités d'écriture et de lecture des commentaires. Concevoir la modération comme une intervention située sur la lecture/écriture d'autrui implique d'élaborer des hypothèses non seulement sur l'existence ou non d'une action de modération, mais également sur les rapports de pouvoir, de légitimité et d'intentionnalité qui l'accompagnent : par qui cette intervention a-t-elle été effectuée ? à l'encontre de quel type d'énonciateur ? sur quel fondement ? On verra que ces questions demeurent aussi difficiles à trancher que la mise en évidence empirique de la modération elle-même.

Notre premier cas concerne la disparition de certains commentaires, identifiés par la comparaison entre des fils YouTube collectés durant la campagne présidentielle de 2022 et ces mêmes fils tels qu'ils apparaissent encore en 2026. L'hypothèse de travail repose sur un critère minimal : lorsque les comptes auteurs de commentaires sont toujours actifs en 2026, il devient plausible d'attribuer la disparition à un tiers autre que l'auteur initial. C'est la situation du commentaire suivant, écrit pendant la campagne présidentielle de 2022, sous une rediffusion d'un entretien télévisé avec le candidat Éric Zemmour, et publié sur YouTube par la chaîne officielle du candidat :

Je ne comprends pas comment on peut inviter un nazi, un manipulateur, un menteur, falsificateur de la vérité dans une émission de grande audience.

C'est un signe net et clair que notre société entre dans une décadence jamais connue auparavant.

Zemmour, lorsque l'on vous confronte à vos contradictions, vous ne faites que boter en touche et embrayer sur un autre sujet. Mais ce n'est pas très correct et les gens incorrects finissent tôt ou tard par se casser les dents.

Une première hypothèse serait que la suppression provienne du propriétaire de la chaîne, dans une logique de protection de l'image du candidat ou de filtrage politique. Mais il est également envisageable que la suppression ait été effectuée par YouTube lui-même au titre de la violation de ses règles communautaires (discours haineux et harcèlement ciblé). Toutefois, faute de charte de modération publique émise par la chaîne, et du fait de l'opacité procédurale de retrait de la plateforme, ces pistes demeurent spéculatives : on ne peut que postuler l'une ou l'autre de ces formes d'intervention. La trace observable n'a ainsi valeur ni de preuve ni de démonstration, mais bien d'indice à interpréter, d'énigme sémiotique née du manque de documentation explicite.

Notre deuxième cas porte sur l'opacité du tri algorithmique des commentaires, et plus précisément sur les critères qui définissent les commentaires dits « les plus populaires ». En 2026, nous avons comparé des fils de commentaires publiés sur YouTube, au sein d'une même chaîne, sous des vidéos du même programme radiophonique, et consultés depuis un navigateur sans historique activé, pour trois candidats de la campagne présidentielle de 2022. Cette comparaison révèle que, dans l'émission consacrée au premier candidat, le commentaire mis en avant correspond à celui ayant recueilli le plus grand nombre de *likes* ; dans celle consacrée au deuxième candidat, il s'agit du deuxième commentaire le plus *liké* ; enfin, dans l'émission consacrée au troisième candidat, le commentaire placé en tête n'apparaît qu'au vingt-septième rang des plus *likés*. Ce dernier adopte par ailleurs une tonalité ironique et, comme plusieurs autres réactions associées à cette vidéo, prend pour cible le travail des journalistes ayant mené l'entretien. On peut alors supposer que la popularité des commentaires repose certes sur des indicateurs quantitatifs produits par le public, mais qu'elle intègre aussi des

paramètres plus subtils, liés au contenu des contributions et à leur positionnement vis-à-vis de la publication principale. Ainsi, nous sommes de nouveau face à une trace-énigme, dont l'analyse ne consiste pas tant à restituer un mécanisme de tri algorithmique qu'à explorer la multiplicité de ce qui échappe à la visibilité.

Notre dernier cas concerne l'activation du mode lent² par la chaîne Twitch d'un groupe médiatique lors d'une interview avec le candidat E. Zemmour. Les interpellations du public adressées aux modérateurs, fréquentes sur Twitch (Thach et al., 2022), nous ont permis d'identifier trois messages qui font explicitement référence au dispositif. Le premier, posté environ onze minutes après le lancement du stream et avant l'arrivée du candidat, critique une application jugée trop stricte trop tôt : « slow mode d'une minute c'est un peu long pour le before, non ? ». Le second, en plein cœur de l'émission, alors que les messages relatifs à la religion et les prises de position favorables au candidat s'enchaînent rapidement, réclame à l'inverse davantage de modération : « Y a moyen de mettre le chat en mode slow ? ». Le troisième, en fin d'émission, dresse un bilan : « @ les modos : le slow mode n'était pas assez lent, le chat allait globalement trop vite et certains messages étaient "spammés" toutes les 2 minutes ». La mise en place du mode lent est donc ici observable mais son intentionnalité demeure incertaine : vise-t-elle à limiter l'effet de saturation produit par les supporters du candidat, à améliorer la lisibilité des échanges, à canaliser les sujets sensibles, ou relève-t-elle d'une forme de prévention ? Ici encore, la trace ne permet pas de conclure, elle invite à émettre des hypothèses sur les logiques de modération en jeu.

Ces trois cas montrent une modération qui ne se laisse pas saisir comme phénomène passé, mais uniquement comme forme indicielle, au sens où Jeanneret l'entend : un reste, une incomplétude qui demande à être interprétée. La trace ne prouve pas, elle problématise. Ce que l'on observe n'est pas la modération elle-même, mais l'espace d'incertitude qu'elle laisse derrière elle.

² Le mode lent ou *slow mode* est une fonctionnalité de modération de Twitch qui permet de limiter la fréquence d'envoi de messages par le public de la chaîne. Chaque spectateur doit ainsi attendre un certain délai, fixé par le streamer ou par les modérateurs, avant d'envoyer un nouveau message.

Conclusion

Les médias numériques, loin de se réduire à des espaces neutres d'expression ou de simple circulation de contenus, s'imposent comme des dispositifs techniques et sémiotiques qui reconfigurent sans cesse les frontières du débat public. Les logiques de tri et de modération, souvent invisibles et inaccessibles, dessinent des territoires informationnels où la visibilité, les logiques d'écriture et de lecture deviennent soumises à des mécanismes algorithmiques et des normes éditoriales propres à chaque espace socio-numérique. Face à cette opacité, l'analyste est contraint à une posture réflexive : il doit comparer les espaces discursifs pour inférer les principes de fonctionnement, et reconnaître combien les frontières numériques se révèlent tout autant sociales, techniques que symboliques.

L'articulation entre la sémiotique des écrits d'écran et le concept foucauldien de gouvernementalité ouvre une perspective heuristique pour analyser la modération sur les réseaux socionumériques. Là où la sémiotique des écrits d'écran, en s'attachant à la matérialité des dispositifs, éclaire la manière dont les formes, scripts techniques et logiques d'interface organisent l'expérience du lire/écrire en ligne, l'approche de la gouvernementalité permet de situer cette orchestration dans un vaste ensemble de dispositifs de pouvoir qui visent à conduire, canaliser et normaliser les expressions en ligne. En conjuguant ces approches, on comprend que les cadres techniques et sémiotiques ne sont pas de simples environnements neutres, mais constituent des régimes de visibilité, de légitimité et d'agencement des voix par l'infrastructure même de la communication numérique.

Dans ce contexte, la matérialité des dispositifs numériques joue un rôle central. Les interfaces, les formats d'écriture (caractères, hashtags, mentions), les règles d'accès et de modération, les architextes, sont autant de frontières qui rendent l'accès à la compréhension des mécanismes internes difficile. Cette matérialité, loin d'être passive, agit comme une force structurante qui oriente, filtre et normalise les discours, tout en offrant paradoxalement aux usagers et aux chercheurs un terrain riche pour l'expérimentation méthodologique, la créativité tactique et le contournement des cadres préétablis. Ainsi, l'étude des espaces numériques ne

peut se concevoir sans une prise en compte centrale des effets de frontière qui marquent chaque tentative d'investigation.

L'examen de nos trois cas d'étude a permis d'interroger la nature et la portée des traces laissées en ligne, tout en renouvelant le regard porté sur les processus de la modération. Chaque situation met en lumière le processus interprétatif auquel est confronté le chercheur : les indices, vides et discordances observables ne sont jamais des preuves brutes ou autosuffisantes, mais des énigmes à déchiffrer, selon la logique du « paradigme indiciel » développée par Jeanneret (2019). Ainsi, la disparition d'un commentaire polémique, l'opacité d'un tri algorithmique ou l'activation ambiguë d'un dispositif de modération ponctuel obligent à naviguer dans une zone d'incertitude, où seules des hypothèses peuvent être formulées sur les intentions, les logiques d'action, ou les commanditaires effectifs de la modération. Loin d'offrir un accès direct à la « réalité » sociotechnique, les traces accessibles ne prennent sens qu'au croisement d'un dispositif matériel, de pratiques sociales et d'un effort herméneutique : elles sont « fabrique » dans leur production comme dans leur interprétation

Une autre difficulté, qui tient également à la matérialité et à l'opacité des réseaux socio-numériques a été rencontrée dans le cadre du projet AMULEX, lorsque nous avons cherché à contacter les personnes³ ayant écrit des commentaires de notre corpus : impossibilité d'envoyer des messages directs, limitation algorithmique des interactions, absence de notification en cas de blocage ou de suppression, configuration de comptes nécessitant des abonnements premium ou l'activation de paramètres spécifiques. Ces obstacles techniques — souvent invisibles pour un usage ordinaire — deviennent de véritables contraintes pour le chercheur, qui doit inventer en permanence des tactiques de contournement (variation des formats de messages, alternance des plateformes, création de contenus audiovisuels pour maximiser la crédibilité de sa démarche scientifique). À cela s'ajoutent des difficultés méthodologiques : comptes désactivés ou inactifs, temporalité décalée entre le moment du commentaire analysé et la prise de contact, absence d'informations sur l'auteur dans les interfaces, risques d'être perçu comme spam ou bot. Ces limites soulignent combien l'enquête sur les réseaux numériques requiert une adaptabilité

³ Travail effectué au sein du projet AMULEX par Gocke Tuncel et Ibtihaje Chawki.

méthodologique, une réflexivité continue et l'acceptation de la dimension expérimentale, bricolée, et souvent fragile du terrain numérique contemporain.

Authorship contribution

Aude Seurrat: Conceptualization; Methodology; Writing – Original Draft; Project Administration.

María Inés Laitano: Conceptualization; Methodology; Data Curation; Writing – Original Draft.

Conflict of interest

Les autrices déclarent qu'elles n'ont aucun conflit d'intérêt concernant la recherche, la rédaction ou la publication du présent article.

Data Availability Statement

Les données analysées pour la production de cet article ne peuvent pas être rendues publiques en raison de contraintes imposées par le Règlement européen sur la protection des données (RGPD), ainsi que des risques de ré-identification. Un accès restreint peut toutefois être accordé aux chercheurs sur demande motivée, sous réserve de validation par les responsables du projet et du respect des conditions d'utilisation scientifique. Des informations complémentaires sont disponibles dans le Plan de gestion des données du projet AMULEX : <https://hal.science/hal-04693656/document>

Déclaration d'utilisation de l'IA

Absence d'usage.

Déclaration des sources de financement

Cette recherche a été financée par l'Agence Nationale de la Recherche (ANR) au titre du projet "ANR-23-CE41-0010-01".

Remerciements

Les autrices remercient Stéphanie Wojcik et Orélie Desfriches Doria pour la gestion du projet AMULEX et l'acquisition du financement.

Lien preprint

<https://preprints.scielo.org/index.php/scielo/preprint/view/14015>

Bibliographie

Badouard, R. (2020). *Les nouvelles lois du Web : Modération et censure*. La République des Idées. Seuil.

Boullier, D. (2024). *Social Media Reset : Réinventer l'infrastructure numérique des propagations pour couper les chaînes de contagion*. SciencesPo.

Christin, A.-M. (1995). *L'Image écrite ou La déraison graphique*. Flammarion.

Dijck, J. van, & Poell, T. (2013). Understanding Social Media Logic. *Media and Communication*, 1(1), 2-14. <https://doi.org/10.17645/mac.v1i1.70>

Foucault, M. (2004). *La naissance de la biopolitique. Cours au college de france (1978-1979)*. Gallimard / Le Seuil.

Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. Yale University Press.
<https://doi.org/10.12987/9780300235029>

Grison, T. (2024). *Faire dire, faire taire : La modération (algorithmique) des réseaux sociaux numériques au prisme des sexualités* [Phdthesis, Sorbonne Université]. <https://hal.science/tel-05007981>

Grison, T., Julliard, V., Alié, F., & Ecrement, V. (2023). La modération abusive sur Twitter: Étude de cas sur l'invisibilisation des contenus LGBT et TDS en ligne. *Réseaux*, 237(1), 119-149. <https://doi.org/10.3917/res.237.0119>

Jeanneret, Y. (2006). Désigner, entre sémiotique et logistique. In I. Timimi & S. Kovacs (Éds.), *Indice, index, indexation* (p. 17-36). Association des professionnels de l'information et de la documentation.

Jeanneret, Y. (2011). Complexité de la notion de trace : De la traque au tracé. In B. Galinon-Melenec (Éd.), *L'Homme trace : Perspectives anthropologiques des traces contemporaines* (p. 59-86). CNRS Éditions.
<https://doi.org/10.4000/books.editions-cnrs.16683>

Jeanneret, Y. (2014). *Critique de la trivialité : Les médiations de la communication, enjeu de pouvoir*. Éditions Non standard.

Jeanneret, Y. (2019). *La fabrique de la trace*. Paris : ISTE.

Jiang, S., Robertson, R. E., & Wilson, C. (2019). Bias Misperceived: The Role of Partisanship and Misinformation in YouTube Comment Moderation.

Proceedings of the International AAAI Conference on Web and Social Media,
13, 278-289. <https://doi.org/10.1609/icwsm.v13i01.3229>

Lire, écrire, récrire : Objets, signes et pratiques des médias informatisés. (2003). In E. Souchier, Y. Jeanneret, & J. Le Marec (Éds.), *Lire, écrire, récrire : Objets, signes et pratiques des médias informatisés*. Éditions de la Bibliothèque publique d'information. <https://books.openedition.org/bibpompidou/394>

Marion, P. (1997). Narratologie médiatique et médiagenie des récits. *Recherches en Communication*, 7. <https://doi.org/10.14428/rec.v7i7.46413>

Meunier, A., Crosset, V., & Ricci, D. (2024). *Styles de Modération* (E. A. | médialab S. Po, Éd.). <https://hal.science/hal-04834555>

Smyrnaiois, N., & Marty, E. (2017). Profession « nettoyeur du net ». De la modération des commentaires sur les sites d'information français. *Réseaux*, 205(5), 57-90. <https://doi.org/10.3917/res.205.0057>

Souchier, E. (1996). L'écrit d'écran, pratiques d'écriture & informatique. *Communication et langages*, 107(1), 105-119. <https://doi.org/10.3406/colan.1996.2662>

Souchier, E., Candel, É., & Gomez-Mejia, G. (2019). *Le numérique comme écriture. Théories et méthodes d'analyse*. Armand Colin; Cairn.info. <https://www.cairn.info/le-numerique-comme-ecriture--9782200618582.htm>

Thach, H., Mayworm, S., Delmonaco, D., & Haimson, O. (2022). (In)visible moderation: A digital ethnography of marginalized users and content

moderation on Twitch and Reddit. *New Media & Society*.

<https://doi.org/10.1177/14614448221109804>

Este preprint foi submetido sob as seguintes condições:

- Os autores declaram que os necessários Termos de Consentimento Livre e Esclarecido de participantes ou pacientes na pesquisa foram obtidos e estão descritos no manuscrito, quando aplicável.
- Os autores declaram que a elaboração do manuscrito seguiu as normas éticas de comunicação científica.
- Os autores declaram que estão cientes que são os únicos responsáveis pelo conteúdo do preprint e que o depósito no SciELO Preprints não significa nenhum compromisso de parte do SciELO, exceto sua preservação e disseminação.
- Os autores declaram que os dados, aplicativos e outros conteúdos subjacentes ao manuscrito estão referenciados.
- O manuscrito depositado está no formato PDF.
- Os autores declaram que a pesquisa que deu origem ao manuscrito seguiu as boas práticas éticas e que as necessárias aprovações de comitês de ética de pesquisa, quando aplicável, estão descritas no manuscrito.
- Os autores declaram que uma vez que um manuscrito é postado no servidor SciELO Preprints, o mesmo só poderá ser retirado mediante pedido à Secretaria Editorial do SciELO Preprints, que afixará um aviso de retratação no seu lugar.
- Os autores concordam que o manuscrito aprovado será disponibilizado sob licença [Creative Commons CC-BY](#).
- O autor submissor declara que as contribuições de todos os autores e declaração de conflito de interesses estão incluídas de maneira explícita e em seções específicas do manuscrito.
- Os autores declaram que o manuscrito não foi depositado e/ou disponibilizado previamente em outro servidor de preprints ou publicado em um periódico.
- Caso o manuscrito esteja em processo de avaliação ou sendo preparado para publicação mas ainda não publicado por um periódico, os autores declaram que receberam autorização do periódico para realizar este depósito.
- O autor submissor declara que todos os autores do manuscrito concordam com a submissão ao SciELO Preprints.