



**HAL**  
open science

# Analyser les débats publics autour d'une loi controversée. A propos d'une enquête sur l'Observatoire TransMedia

Valérie Devillard, Guillaume Le Saulnier

## ► To cite this version:

Valérie Devillard, Guillaume Le Saulnier. Analyser les débats publics autour d'une loi controversée. A propos d'une enquête sur l'Observatoire TransMedia. INA Editions. En quête d'archives. Bricolages méthodologiques en terrains médiatiques, , pp.175-182, 2018, Médias et Humanités, 978-2-86938-250-3. hal-04478552

**HAL Id: hal-04478552**

**<https://hal.science/hal-04478552>**

Submitted on 26 Feb 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.

## **Analyser les débats publics autour d'une loi controversée. À propos d'une enquête sur l'Observatoire Transmedia**

Valérie Devillard (Université Panthéon-Assas – Paris 2, IFP/CARISM)  
Guillaume Le Saulnier (Université de Reims Champagne-Ardenne, CÉREP)

Le foisonnement des sources médiatiques sur le web, leur volatilité, ou encore leur plus ou moins grande circulation elle-même dépendante du jeu des algorithmes, invitent les chercheurs·e·s en sciences sociales à une grande prudence méthodologique dans leur ambition de constituer un corpus « contrôlé » pour servir leur objet de recherche. Aussi saisissons-nous l'opportunité offerte d'expérimenter une archive de données, l'Observatoire Transmedia, élaboré conjointement par des chercheurs de l'Université Sorbonne-Nouvelle – Paris 3 et par le département « Recherche » de l'Ina.

De prime abord, nous espérons ainsi éviter hauts fonds et écueils en tout genre, que nous promettent l'analyse des arènes numériques autour de l'examen parlementaire d'une nouvelle loi contre la prostitution. Le tri et la sélection de milliers de documents issus du web, ou numérisés pour le web, s'avèrent à notre portée dans la mesure où nous accédons également à la manière dont les données ont été recueillies. Pour autant, nous devons résoudre bien des problèmes pour ne pas saborder notre recherche à l'épreuve du numérique.

En effet, le web s'apparente autant à un Eldorado qu'à un Triangle des Bermudes pour la recherche en sciences sociales, donnant matière à une « *fast research*<sup>1</sup> » souvent oublieuse des exigences scientifiques. Cette dérive est à la fois la cause et l'effet d'une carence de référentiels méthodologiques et déontologiques. Sur le web, le bricolage méthodologique demeure la règle. Il s'élabore dans un va-et-vient entre raisonnement sociologique et méthodes logicielles, à la condition expresse que l'usage des outils informatiques reste subordonné au regard sociologique<sup>2</sup>.

### Épisode 1 – L'exploration des féminismes en ligne

L'objet de notre recherche sur le web s'est construit en plusieurs étapes méthodologiques et a façonné en retour nos usages de l'Observatoire au fur et à mesure de leur progression. Au croisement des études sur les mouvements féministes et les usages militants voire activistes de l'internet<sup>3</sup>, il s'agit de mieux comprendre ce que fait le web aux débats publics sur une des questions les plus clivantes au sein des mouvements féministes : la prostitution. Le choix se porte sur l'examen (en première lecture) par l'Assemblée nationale, du 29 novembre au 4 décembre 2013, de la proposition de loi « renforçant la lutte contre le système prostitutionnel » ; celle-ci prône « le respect de la dignité de la personne, et [elle] entend participer à la politique de lutte contre les violences faites aux femmes et d'égalité entre les femmes et les hommes<sup>4</sup> ». Cette proposition de loi est l'aboutissement d'une croisade morale<sup>5</sup> relayée par les associations abolitionnistes. Ses partisan·e·s comme ses détracteur·trice·s se sont

---

<sup>1</sup> Latzko-Toth G., Proulx S., « Enjeux éthiques de la recherche sur le web », in Barats C. (dir.), *Manuel d'analyse du web en Sciences Humaines et Sociales*, Paris, Armand Colin, 2013, p. 32 sq.

<sup>2</sup> Demazière D., Brossaud C., « Introduction. Méthodes logicielles et réflexivité du sociologue », in Demazière D., Brossaud C., Trabal P., Van Meter K. (dir.), *Analyses textuelles en sociologie. Logiciels, méthodes, usages*, Rennes, PUR, 2006, p. 11 sq.

<sup>3</sup> Granjon F., *L'Internet militant*, Paris, Apogée, 2001.

<sup>4</sup> <http://www.assemblee-nationale.fr/14/propositions/pion1437.asp>.

<sup>5</sup> Mathieu L. *La Fin du tapin. Sociologie de la croisade pour l'abolition de la prostitution*, Paris, Éditions François Bourrin, 2013.

livré·e·s à une bataille interprétative auprès des instances politiques et dans les arènes médiatiques. Une des caractéristiques de ce débat est d'avoir été construit en termes de rapports de genre. D'où notre attachement dans un premier temps à analyser, en collaboration avec l'historienne contemporaine Bibia Pavard, les différents prismes féministes favorables ou opposés à la loi, de l'abolitionnisme radical au positionnement « pro-droits ».

L'intérêt est d'explorer les usages politiques et genrés du web, recherches relativement peu développées en France et souffrant le plus souvent d'un double biais. Celles-ci tendent à aborder les relations entre genre, internet et politique de manière restrictive. Le spectre de l'activité politique se voit cantonné à des pratiques plutôt conventionnelles, reléguant au second plan les engagements publics sporadiques (pétition en ligne, par exemple) ou, au contraire, plus radicaux des activistes. À cette première lacune s'ajoute une seconde plus contraignante encore. Ces recherches autour du genre numérique privilégient les usages et les usager·e·s jugé·e·s « ordinaires » de l'internet, au risque de les évaluer à l'aune du modèle normatif des pratiques dominantes donc « masculines » du web.

La première perspective de recherche est de montrer comment les rapports sexués et de genre font trace sur l'internet dans la sphère des activistes féministes françaises. Puis, l'enquête dégage les différents modes d'intervention au sein des blogosphères féministes. Elle consiste également, après cette cartographie de la galaxie féministe sur le web (portails d'information spécialisés, médias alternatifs, sites des principaux groupes d'activistes), en une série d'entretiens approfondis auprès d'activistes féministes, ou de militant·e·s engagé·e·s dans le débat public autour de la prostitution. De cette première enquête émergent quelques résultats. En premier lieu, le web permet aux campagnes militantes de gagner en complexité. Le débat sur la prostitution a impliqué un large spectre de féministes exposant des positions diverses, les usages militants d'internet favorisant une plus large expression des féminismes, au-delà de la binarité « pour » ou « contre » la loi. On peut d'ailleurs dire que le débat public sur la prostitution donne à voir l'éclatement actuel de la galaxie féministe et ses contradictions internes. En deuxième lieu, l'internet ne vient modifier qu'à la marge les usages militants traditionnels des médias pour les féministes intellectuel·le·s, ou engagé·e·s dans des groupes, des associations, ou pour les féministes institutionnel·le·s, en reprenant la typologie de l'espace social et politique de la cause des femmes<sup>6</sup>. En revanche, il donne une nouvelle représentation aux engagé·e·s non affilié·e·s : « les citoyen·ne·s engagé·e·s » à travers les médias sociaux. En dernier lieu, cette pré-enquête permet d'effectuer un coup de sonde exploratoire dans les corpus numériques mobilisables sur le web pour analyser des débats publics relatifs à la prostitution. Leur analyse est aiguillée par une recension à l'aide de divers moteurs de recherche, fournissant un aperçu provisoire de la structuration des débats publics, des acteur/trice·s mobilisé·e·s, de leur argumentation et leurs stratégies de publicisation<sup>7</sup>.

La seconde perspective aborde les discussions au sein des arènes publiques en ligne à propos de cette même proposition de loi comme cadre fécond pour penser le numérique dans et par les marges. Les modes d'engagement et d'intervention, les espaces d'expression des représentant·e·s des travailleur/se·s du sexe ont été observé·e·s dans les arènes numériques, afin de mettre au jour leurs manières de déjouer la marginalité en formant des « contre-publics

---

<sup>6</sup> Bereni L., « Penser la transversalité des mobilisations féministes : l'espace de la cause des femmes », in Bard C. (dir.), *Les Féministes de la deuxième vague*, Rennes, PUR, 2012, p. 27 sq.

<sup>7</sup> Barbot J., « L'engagement dans l'arène médiatique. Les associations de lutte contre le sida », *Réseaux*, 95, 1999, p. 155-196.

subalternes<sup>8</sup> », et d'analyser leurs rhétoriques et les dynamiques de circulation, leurs soutiens et les rapprochements de positions. L'Observatoire Transmedia est, comme nous le verrons, un outil de recherche approprié pour observer les processus de médiatisation dans les arènes médiatiques *mainstream* ; il l'est moins, en revanche, pour examiner les médias alternatifs et les médias communautaires situés plus à la marge.

## Épisode 2 –

L'Observatoire Transmedia consiste en une plateforme de bases de données, créée pour analyser les flux médiatiques informationnels à l'aide de moteurs de recherche. Les flux ont été capturés de juillet 2011 à juin 2014. Ces données rassemblent : des retranscriptions des journaux et des émissions d'actualité de douze chaînes de télévision, de TF1 à France 24, et de neuf radios, de RTL à France Culture ; les éditions de six quotidiens nationaux (*lemonde.fr*, *liberation.fr*, *lefigaro.fr*, *lacroix.fr*, *lesechos.fr*, *leparisien.fr*) ; les dépêches de l'AFP ; ainsi que 1 300 sites web (sites de news, sites institutionnels ou partisans, blogs, etc.) et le réseau social Twitter. Le dépôt légal web de l'Ina a également été mis à disposition.

Sur cette plateforme, le flux est converti en stock, l'actualité se mue en archive. Cet archivage permet une analyse approfondie des débats publics autour de tel ou tel événement, en l'occurrence une proposition de loi suscitant de nombreuses controverses sur un thème au demeurant passionnel. Par son recours aux technologies informatiques et à la science algorithmique, la fouille de données actualise et exacerbe des problèmes inhérents aux travaux de recherche à partir d'archives. Au premier rang desquels la constitution de corpus, qui sera au cœur de notre propos. Comment constituer et stabiliser un corpus satisfaisant aux critères de représentativité, d'homogénéité et de pertinence, dans le cadre d'une recherche par fouille de données ? Comment maximiser ces critères au profit des données récoltées ?

L'enjeu réside d'emblée dans l'appropriation de la plateforme, par des chercheur·e·s novices en matière de fouille de données. Si le recours aux méthodes logicielles est de plus en plus fréquent dans les sciences sociales, la technique s'apparente souvent, par la complexité ou la fausse simplicité de ses procédures, à une « boîte noire<sup>9</sup> ». Ce point aveugle varie en fonction de la nature et du nombre des opérations déléguées à la technique, et de la réflexivité que l'on s'impose dans l'usage des outils informatiques. En l'occurrence, ce biais est corrigé par l'accompagnement dont les chercheur·e·s bénéficient dans l'appropriation de la plateforme OTMedia. Non seulement ses concepteur·trice·s revendiquent un dialogue continu entre ingénieur·e·s et chercheur·e·s, pour en améliorer le fonctionnement et les résultats, mais en outre l'automatisation des tâches reste (heureusement) inachevée, nécessitant des interventions humaines pour programmer certains calculs ou opérer certaines vérifications<sup>10</sup>.

La déjà un impératif catégorique, mais aussi un problème épineux pour les corpus prélevés dans les médias audiovisuels<sup>11</sup>. Ce problème devient particulièrement aigu sur le web. D'une part, parce que les sources et les documents disponibles en ligne sont, à proprement parler, incommensurables. D'autre part, à cause de la mobilité et, surtout, de la

---

<sup>8</sup> Fraser N., « Repenser la sphère publique : une contribution à la critique de la démocratie telle qu'elle existe réellement », *Hermès*, 31, 2001, p. 125-156.

<sup>9</sup> Demazière D., Brossaud C., *op. cit.*, p. 15.

<sup>10</sup> Entretien avec Marie-Luce Viaud (juillet 2016), chercheuse à l'Ina Recherche, coordinatrice du programme OTMedia.

<sup>11</sup> Bonville de J., *L'Analyse de contenu des médias. De la problématique au traitement statistique*, Paris, De Boeck & Larcier, 2006.

« mutabilité » des données numériques. La localisation d'un site (ou d'une page) web peut changer, s'il reçoit une nouvelle URL, de même que ses conditions d'accès, gratuites ou payantes. Les données numériques sont fondamentalement instables, car elles peuvent à tout moment être modifiées, complétées, amputées, voire supprimées. La blogosphère consiste en un vaste « cimetière<sup>12</sup> » de blogs, en ce qu'ils ne sont plus actualisés et peuvent s'évanouir du jour au lendemain. Le web constitue ainsi un matériau « quasi vivant<sup>13</sup> », une « cible mouvante<sup>14</sup> » en perpétuelle évolution. Dans ces conditions, il semble irréductible à tout archivage systématique. Ces limitations « prennent à contre-pied les procédures les plus stabilisées de la recherche en sciences sociales comme l'échantillonnage ou la vérification<sup>15</sup> ». Elles rendent pratiquement impossibles à la fois une recension exhaustive des documents sur un thème donné et la vérification des résultats par la reproduction à l'identique des corpus auxquels ils réfèrent.

La plateforme OTMedia aide à surmonter ces difficultés. Grâce à un archivage systématique et contrôlé, elle autorise la constitution de corpus représentatifs sur les années considérées, par une recension quasi exhaustive des ressources disponibles. Cette méthode neutralise les biais d'une recension au moyen des moteurs de recherche, généralistes ou spécialisés, ou d'une recension par « boule de neige », à savoir les logiques internes de tout dispositif d'indexation pour l'un, et les formes horizontales de référencement entre sites web, pour l'autre.

Après différents tests de requêtes, deux corpus sont ainsi générés, respectivement à partir de « prostitution AND "proposition de loi" » et « Strass AND prostitution ». Le premier corpus rassemble jusqu'à 1 720 documents, sur la période mars 2013 à mars 2014. Il concerne les débats publics relatifs à la mesure législative promue par le gouvernement socialiste, et leur réfraction dans les arènes mass-médiatiques et numériques. Le second corpus se compose de 484 documents, sur une période élargie de septembre 2011 à mai 2014. En interpénétration avec le premier, il se concentre sur les modes de présence et d'intervention du Syndicat du travail sexuel (Strass) dans les arènes médiatiques.

Le critère de l'homogénéité du corpus se confronte à la dimension transmédia des ressources disponibles. Celles-ci agrègent des dépêches d'agence, des articles de presse, des contenus audiovisuels, des écrits d'écran, émanant d'entreprises de presse, d'« infomédiaires<sup>16</sup> », de sites institutionnels, ou des auto-publications sur sites ou blogs. Un tel dispositif facilite l'analyse de la couverture médiatique d'un événement tous médias confondus, ainsi que des logiques complexes de diffusion et de « circulation circulaire de l'information<sup>17</sup> » entre agence de presse, médias traditionnels, sites web et médias sociaux. Mais cela suppose une double réduction, formelle et contextuelle : d'un côté, les contenus audiovisuels sont systématiquement retranscrits ; de l'autre, les documents sont arrachés à leur environnement, et se présentent, sur l'interface OTMedia, sous forme de vignettes unitaires, chacune donnant accès à leur contenu converti en texte brut. Les documents sont ainsi vidés de leur « énonciation éditoriale<sup>18</sup> ». Ce

---

<sup>12</sup> Cardon D., Delaunay-Téterel H., « La production de soi comme technique relationnelle. Un essai de typologie des blogs par leurs publics », *Réseaux*, 138, 2006, p. 19-20.

<sup>13</sup> Burnay N., Klein A., « Blogs et questionnements méthodologiques », in Klein A. (dir.), *Objectif blogs ! Explorations dynamiques de la blogosphère*, Paris, L'Harmattan, 2007, p. 26.

<sup>14</sup> Convert B., Demailly L., *Les Groupes professionnels et l'internet*, Paris, L'Harmattan, 2007, p. 15.

<sup>15</sup> Convert B., Demailly L., *op. cit.*, p. 13.

<sup>16</sup> Rebillard F., Smyrniaios N., « Les infomédiaires, au cœur de la filière de l'information en ligne. Les cas de Google, Wikio et Paperblog », *Réseaux*, 160-161, 2010, p. 163-194.

<sup>17</sup> Bourdieu P., *Sur la télévision*, Paris, Raisons d'agir, 1996, p. 22-29.

<sup>18</sup> Jeanneret Y., Souchier E., « L'énonciation éditoriale dans les écrits d'écran », *Communication & Langages*, 145, 2005, p. 3-15.

sacrifice de leur environnement contextuel autorise toutefois le façonnage de corpus cohérents à partir de ressources hétérogènes, et facilite le travail d'analyse sur des données transcrites plus aisément manipulables.

Une difficulté supplémentaire réside dans la pertinence du corpus au regard des résultats de la fouille. Le travail de recension des documents pertinents est parasité par un niveau élevé de « bruit », qui impose un nettoyage fastidieux. En effet, la requête « prostitution » draine nombre de documents relatifs à des scandales sexuels dans lesquels des personnalités publiques sont impliquées, et dont certaines retombées coïncident avec la période retenue : « l'affaire du Carlton de Lille » en matière de proxénétisme, le « Rubygate » en Italie et « l'affaire Zahia D. » dans le football français. De surcroît, les deux corpus se caractérisent par un taux de répliation très élevé. Cela signifie que la production journalistique se compose, dans une large mesure, de dépêches AFP reproduites mot pour mot ou à peine réécrites. Par exemple, dans le corpus de 1 720 documents, la dépêche AFP du 28 mars 2013 intitulée « Prostitution : le Sénat abroge le délit de racolage passif » est reprise jusqu'à 26 fois ; celle du 25 mars 2013, titrée « Pataquès chez les sénateurs PS sur l'abrogation du délit de racolage », est répliquée à 8 reprises. Le corpus illustre et confirme ainsi l'ampleur des logiques de retraitement à l'œuvre dans la presse en ligne en général, et sur les infomédiaires en particulier<sup>19</sup>.

Le principal écueil tient à l'orientation *mainstream* de la plateforme. L'Observatoire Transmedia a été conçu, initialement, pour étudier la campagne des élections présidentielles de 2012. Les flux informationnels capturés émanent logiquement de supports grand public. Sont ainsi ignorées toutes les arènes médiatiques davantage politisées ou militantes, notamment celles où se matérialisent des contre-discours et un cadrage « pro-droits » face à l'offensive menée par les mouvements abolitionnistes. Nous nous trouvons ainsi immergé·e·s dans un corpus fortement standardisé et dépolitisé ; et ce, alors que les débats publics autour de la prostitution sont structurés par une forte polarisation et un antagonisme chronique entre les mouvements abolitionnistes d'un côté, et celles et ceux qui revendiquent la liberté de recourir à la sexualité vénale de l'autre<sup>20</sup>. Pour corriger cette pente *mainstream* et pluraliser le corpus, des sources complémentaires sont prélevées dans la presse d'opinion (*L'Humanité*, *Mediapart*) et les médias communautaires (*Yagg*). Par leur positionnement éditorial, ces sources rétablissent en partie l'intensité des débats publics et la pluralité des points de vue en concurrence, à partir d'un corpus certes représentatif mais excessivement consensuel.

## Conclusion

L'importation des méthodes logicielles dans les sciences sociales présente bien des avantages : constitution et analyse de vastes corpus, synthèse et visualisation des résultats orientant leur interprétation, etc. Encore faut-il que leur usage soit toujours contrôlé par une réflexivité critique. En l'occurrence, la plateforme OTMedia permet de mener une analyse intensive des débats publics dans les arènes médiatiques. Son expérimentation stimule aussi notre imagination de chercheur·e·s, dans les possibilités offertes de penser les discours énoncés sur le problème public considéré, leur circulation d'une arène à l'autre, l'uniformité de la production journalistique au regard de la pluralité des positions en concurrence. Néanmoins, cette analyse demeure média-centrée, et il importe de conjuguer la fouille de données avec des méthodes d'enquête qualitatives, tels des entretiens compréhensifs ; et ce, en vue de remonter

---

<sup>19</sup> Rebillard F., « Du traitement de l'information à son retraitement. La publication de l'information journalistique sur l'internet », *Réseaux*, 137, 2006, p. 29-68.

<sup>20</sup> Mathieu L., *La Condition prostituée*, Paris, Textuel, 2007, p. 15-44.

en amont de l'archive, pour travailler au plus près des acteur/trice·s mobilisé·e·s, de leurs points de vue et de leurs stratégies de publicisation.