



Le projet #Idéo2017 : Quelles implications du/de la chercheur-e en tant qu'acteur-trice potentiel du changement social ? Exemplification à partir du discours politique numérique

Julien Longhi

► To cite this version:

Julien Longhi. Le projet #Idéo2017 : Quelles implications du/de la chercheur-e en tant qu'acteur-trice potentiel du changement social ? Exemplification à partir du discours politique numérique. Cahiers de Linguistique, 2019, 442. halshs-02153102

HAL Id: halshs-02153102

<https://shs.hal.science/halshs-02153102>

Submitted on 12 Jun 2019

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.

Le projet #Idéo2017 : Quelles implications du/de la chercheur-e en tant qu'acteur-trice potentiel du changement social ? Exemplification à partir du discours politique numérique

Julien LONGHI
Université de Cergy-Pontoise

Introduction

Le projet #Idéo2017 fait suite à plusieurs projets de recherche dont l'objectif était l'analyse linguistique des tweets politiques (voir le corpus *Polititweets* pour un aboutissement et une mise à disposition du corpus en 2014, et LONGHI 2017 pour une synthèse). En effet, les tweets politiques étant considérés comme un genre du discours politique (LONGHI 2013), leur analyse permettrait d'analyser le discours politique contemporain, en agrégeant une pluralité de formes d'expression (interviews de matinales, allocutions, meetings, etc.) reprises par une succession de tweets sur les comptes des femmes et hommes politiques (ou des partis, mouvements, etc.). Ces travaux ont eu l'intérêt de constituer une méthode d'analyse, et de caractériser certains aspects des tweets politiques, et du discours politique ; leur limite, comme souvent dans le domaine académique, est le décalage entre les contextes de production des discours (campagnes électorales, événements) et la diffusion des résultats de la recherche (colloque, articles).

Pour agir sur la compréhension des discours, et intervenir sur le changement social, la plateforme #Idéo2017¹, développée dans le cadre de l'élection présidentielle 2017, traitait les messages des candidats en constituant un corpus en quasi temps réel (mis à jour chaque 24 heures) à partir des tweets publiés dans leurs comptes officiels (à partir du premier septembre 2016). Par l'utilisation d'outils et de fonctionnalités issues de la linguistique outillée, ce traitement rendait compte des principales caractéristiques de ce corpus, et permettait notamment des comparaisons entre les différents candidats. Ce projet visait en outre à un outil et des méthodes d'analyse à la disposition des citoyens, et de les accompagner dans leur appréhension de ces outils. D'une certaine manière, ceci pouvait s'intégrer dans une démarche de science participative² :

Les sciences participatives correspondent à des dispositifs de recherche dans lesquels les acteurs de la société civile participent aux processus de production scientifique. Ces collaborations entre scientifiques et citoyens, entre laboratoires, amateurs et associations, se développent et se diversifient. Portées par l'avancée propre des sciences, par l'intérêt des objets et phénomènes étudiés, par la volonté d'agir des acteurs variés et par des expériences réussies, elles sont aussi stimulées par le potentiel offert par le numérique.

Cet article pose donc la question, de manière réflexive et rétrospective, de la dimension citoyenne, voire participative, de cette plateforme. Nous examinons ensuite les nouveaux formats d'expression pour les chercheurs, qui souhaitent être acteurs du changement social, et les perspectives que cela soulève pour de futures recherches.

¹ <http://ideo2017.ensea.fr/plateforme/>. Ce projet a reçu le soutien financier de la Fondation de l'université de Cergy-Pontoise, puis du consortium Corli pour la finalisation du corpus et son dépôt sur la plateforme Ortolang.

² <http://www.sciences-participatives.com/Mission>

#Idéo2017, une plateforme pour rendre les citoyens acteurs de l'analyse d'une campagne électorale

Dans le rapport « Les Sciences Participatives en France »³ coordonné par François Houllier, il est indiqué :

les sciences participatives se déclinent au pluriel. La grande diversité de disciplines, de sujets, d'acteurs, de méthodes et de finalités ne peut conduire à l'élaboration d'une définition unique et consensuelle, qui en réduirait le champ et la portée. Cette diversité est caractéristique des sciences participatives ; elle est aussi une composante intrinsèque de leur intérêt.

Dans le cas qui sera décrit avec #Idéo2017, il n'y a pas eu une collaboration initiale avec les citoyens. Néanmoins, dans la mesure où ils sont les destinataires de la plateforme proposée, certains relais ont pu être trouvés : journalistes (partenariat avec la chaîne locale Vo TV/ Vo News pour connaître d'éventuels besoins sur le traitement de l'information), étudiants (tests de la plateforme), communautés sur les réseaux sociaux (par la communication faite autour du projet).

Faire de la science citoyenne

Pour caractériser notre projet du point de vue participatif, on peut se fonder sur les critères présentés dans le rapport sur les sciences participatives : objectifs ; formes d'engagement (bénévolat, rémunération, durée précise ou non, individuel ou collectif, etc.), rôles des différents participants, publics visés, échelle du projet, type de protocole.

On constate alors qu'il est peut-être plus « citoyen » que « participatif », mais la discussion de cette frontière est intéressante, à la fois pour réfléchir sur les différents paramètres, mais aussi penser de prochains projets. En effet, la citoyenneté est bien comprise ici comme forme de participation. Les différents choix qui ont été effectués, en termes d'ergonomie, de fonctionnalité, de communication, ont été faits dans le but que la plateforme soit utile pour les citoyens, et utilisée. Si certaines orientations ont été le fruit de discussions avec des personnalités extérieures au domaine académique, il n'y a pas eu de réel protocole d'association préalable

³ <https://inra-dam-front-resources-cdn.brainsonic.com/ressources/afile/321785-853d8-resource-rapport-impression-sciences-participatives-fevrier-2016.html>

de la sphère sociale. Si ceci peut sembler paradoxal, il faut noter que la participation des citoyens était alors entrevue dans l'usage qui serait fait de la plateforme (quels mots ou comptes étaient les plus utilisés ? quelle durée moyenne la plateforme était-elle utilisée par un internaute ? des difficultés étaient-elles signalées ?). De plus, #Idéo2017 était considéré comme un prototype opérationnel, et de prochains développements pourront associer des citoyens avant leur mise en œuvre effective.

On suivra notamment Sandra Laugier⁴ quand elle écrit :

La question de la participation citoyenne à la science est une question controversée car elle met en cause une frontière, voire une hiérarchie entre les « citoyens ordinaires » et les experts scientifiques, et donc le monopole de la production de la science par ses professionnels. Elle est, autant que la participation politique, une question de démocratie, et touche au caractère public et ouvert de la science comme bien commun.

L'objet de la plateforme (le discours politique) étant au cœur de la question démocratique, l'implication des « citoyens ordinaires » est fondamentale pour l'analyse du discours, et nous rejoignons encore Sandra Laugier quand elle précise :

il revient aux sciences humaines et sociales de s'engager dans un examen des conditions d'une véritable participation, à partir de la capacité des citoyennes à intervenir dans le traitement de questions qui les concernent et les affectent directement

La plateforme #Idéo2017 était conçue comme une interface entre les citoyens et les messages politiques. Elle devait générer une interactivité et une participation active sur l'interface, propices à la compréhension des résultats des analyses. Ceci rejoint le propos de Garric et Léglise (2012) sur le discours d'expert et d'expertise : elles montrent que les doutes portés parfois sur l'activité d'expertise nourrissent « un mouvement contestataire dans le contexte duquel sont apparus les termes de contre-expertise « pour désigner un usage croissant de l'expertise comme arme critique vis-à-vis du pouvoir politique et des institutions » (Mouchard, 2005) et d'expertise profane ou citoyenne qui, en désacralisant l'expertise (Blondiaux, 2008), remettent en cause les expertises savantes ou techniques autorisées ». Elles explicitent également le caractère problématique de l'interaction entre « profanes » et « experts » : « si la compétence des premiers à poser des questions pertinentes est parfois remise en cause par les experts, ceux-ci voient en retour critiquer leur exercice

⁴ <http://www.allianceathena.fr/actualite/shs-et-sciences-participatives-les-citoyens-ordinaires-comme-ressource-de-la-connaissance>

de l'expertise, à travers la dénonciation, par les citoyens, de l'absence d'accord parmi les participants supposés produire des avis informés sur la question au cœur de la discussion ». L'enjeu dans #Idéo2017 était alors de fournir à l'utilisateur une certaine représentation de la matérialité discursive, et plusieurs points de vue sur son objet, mais sans aller jusqu'à l'interprétation qui peut relever du discours d'expert.

Nous le verrons, toute sa structure était donc pensée pour être compréhensible par le grand public.

Adapter son format de recherche aux enjeux citoyens

L'outil #Idéo2017 est réalisé via plusieurs étapes, comme indiqué dans la Figure 1 : (1) l'extraction de l'ensemble de tweets des candidats, (2) la mise en place d'une sauvegarde des tweets, (3) l'indexation des tweets pour faciliter la recherche dans l'ensemble de tweets, (4) l'application d'un ensemble d'analyses linguistiques sur les tweets, (5) la mise en place d'un moteur de recherche sur l'ensemble de tweets, et (6) l'affichage des résultats sur une page web.

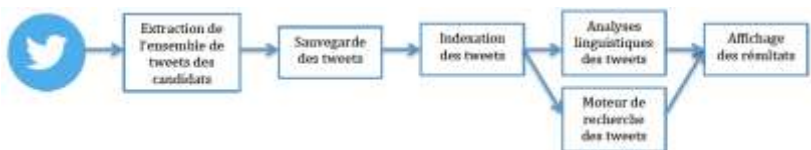


Figure 1 : Chaîne de traitement de l'outil #Idéo2017.

Le choix de l'affichage des résultats permet de concurrencer la démarche traditionnelle avec ce type d'outils d'analyse textuelle : extraction de corpus, mise en forme, formatage, balisage, puis usage d'un logiciel. Ici, tout ce travail est réalisé en amont, et l'utilisateur, en cliquant sur les fonctionnalités, a directement accès aux résultats. Comme indiqué sur l'image suivante, qui reproduit la page d'accueil de la plateforme, trois volets sont proposés à l'utilisateur :

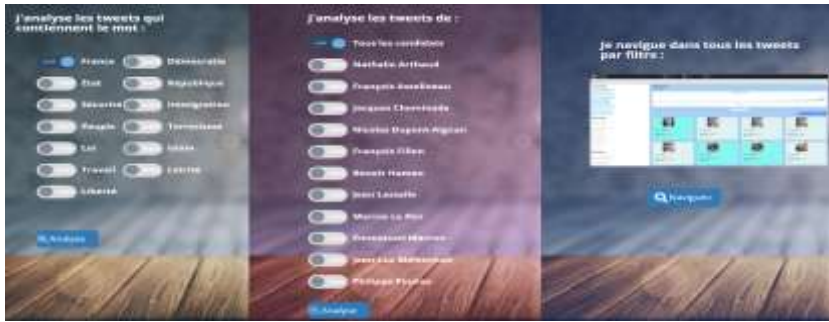


Image 1 : Page d'accueil de la plateforme #Idéo2017

L'analyse "*J'analyse les tweets qui contiennent le mot...*" permet à l'utilisateur de choisir un mot parmi les 13 mots qui sont souvent employés dans les débats politiques (ALDUY 2017). Cette entrée donne accès à quatre analyses possibles : l'usage de ce mot par les différents candidats (sur/sous-emploi et fréquences de la forme exacte), les mots associés à ce mot (analyse de similitudes, basée sur les cooccurrences entre les mots), l'emploi de ce mot et ses dérivés par les différents candidats (analyse basée sur les racines des mots : par exemple *islam*, *islamisme*, *islamiste*, seront regroupés sous la forme /islam/), et le nuage de mots. Ces analyses sont en fait des résultats produits grâce au code du logiciel d'analyse textuelle Iramuteq, issus de calculs qui portent dans le logiciel des noms plus techniques (l'*analyse de similitude* devient par exemple *les mots associés à ce mot*).

L'analyse "*J'analyse les tweets de [candidat]*" permet à l'utilisateur de choisir un candidat parmi les 11 candidats (ou le corpus global des 11 candidats) afin d'analyser ses tweets via les techniques suivantes : les mots les plus utilisés, les thématiques (issues de *la méthode Reneirt*), les relations entre les mots, le nuage de mots, les spécificités des différents candidats (possible si l'utilisateur a choisi d'analyser tous les candidats en même temps).

Enfin, le moteur de recherche permet à l'utilisateur de faire des recherches sur toute la base des tweets, grâce à un outil appelé *ElasticSearch*.

Nous ne détaillons pas plus les aspects techniques du développement de la plateforme, mais renvoyons à LONGHI et al. (2017) et MARINICA et al. (2018) pour les précisions. Notons simplement que les choix de fonctionnalités ont été guidées par l'usage qui pouvait en être fait par le grand public, sans que celui-ci doive recourir à des connaissances statistiques ou informatiques.

Certes, les calculs et le choix des méthodes ne sont pas neutres, et constituent un prisme spécifique sur le matériau. La mise à disposition de plusieurs fonctionnalités, ainsi qu'un outil de navigation, permet selon nous de minimiser ces concessions. Par exemple, l'objectif de la classification lexicale retenue pour faire ressortir les thématiques propres aux candidat.e.s est de « regrouper des « mondes lexicaux » et de mettre en évidence les thématiques générales du corpus », la méthode cherchant à « rendre compte de l'ordre interne d'un discours, à mettre en évidence ses mondes lexicaux »⁵. Elle découle de l'analyse factorielle des correspondances. Une autre fonctionnalité, l'Analyse de similitudes (ADS), établit un calcul sur la base d'« un indice de co-occurrence (combien de fois les éléments vont apparaître en même temps) » pour donner un résultat visuel « où la taille des mots est proportionnelle à la fréquence et où la taille des arêtes et proportionnelle à la force ». Ceci permet de rendre compte visuellement de la fréquence des mots en lien avec les associations spécifiques.

Nous allons par contre nous interroger plus précisément sur la manière dont le développement de cet outil donne à la recherche une portée sur le changement social. Nous entendons par changement social à la fois :

- Le rapport entre les citoyens et l'expertise scientifique, et l'appréhension de résultats issus du milieu académique ;
- La perception des discours politiques, et la prise en compte d'une vision de l'analyse du discours politique plus complexe que celle transmise par les grands médias ;
- La valorisation de la dimension critique de l'analyse du discours, qui se tourne vers les nouvelles technologies afin de s'adapter aux nouveaux formats de « consommation » de l'information.

⁵ <https://datahist.hypotheses.org/11>

La médiatisation de la recherche : vers des nouveaux formats d'expression des chercheurs

Pour accompagner ce projet, un site a été créé : <http://ideo2017.ensea.fr/>. Composé de différentes rubriques, il permettait de documenter le projet :



Image 2 : en-tête du site avec rubriques

Le chercheur blogueur

Mais cette partie vitrine était également complétée par une partie blog, dans laquelle des articles ou informations étaient publiées de manière très régulière :

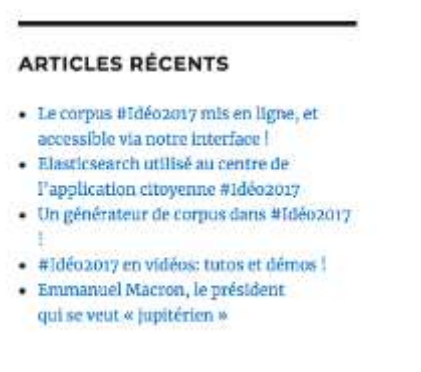


Image 3 : présentation des articles dans le blog

Il s'agissait à la fois d'articles informant de l'avancée du projet, l'introduction de nouvelles fonctionnalités, la création de supports de communication ou d'explications, ou encore la publication d'articles originaux de diffusion de la recherche.

Pour diffuser ces articles, le site *The Conversation*⁶ a notamment été utilisé, via la proposition de contributions régulières durant la campagne électorale (8 articles, plus un traduit par la déclinaison anglaise du site, à propos de la construction du storytelling chez Emmanuel Macron). Comme indiqué sur le site, *The Conversation* France est « un média en ligne d'information et d'analyse de l'actualité indépendant, qui publie des articles grand public écrits par les chercheurs et les universitaires ». Des journalistes travaillent avec les universités et les instituts de recherche afin de diffuser leur savoir vers le plus grand nombre. Le fondement que « l'accès à un journalisme d'expertise authentique et de qualité est un des piliers de la démocratie » nous semble entrer en cohérence avec le projet #Idéo2017. En outre la republication possible des articles sous Creative Commons permet de donner une grande visibilité aux articles, et aussi de pouvoir les reproduire sur le blog du projet.

Comme cela est indiqué sur cette image, le tableau de bord côté utilisateur permet de consulter ses articles, et de connaître le nombre de lecteurs, les commentaires, etc. :



Image 4 : tableau de bord côté « auteur » sur le site *The Conversation*

Par exemple, la publication de l'article suivant a permis de mettre en avant les analyses faisables avec #Idéo2017 :



avril 12, 2017
Le discours du FN sur l'islam à l'épreuve des tweets
 Julien Longhi, Université de Cergy-Pontoise
 La plateforme de recherche #Idéo2017 montre que si le mot « islam » est peu représentatif des tweets de Marine Le Pen, ses dérivés en -iste et -isme le sont beaucoup plus.

⁶ <https://theconversation.com/fr>

Image 5 : aperçu d’un article

On constate ainsi que plus de 11000 lecteurs ont vu cet article :



Image 6 : métriques et statistiques sur le tableau de bord de *The Conversation*

En regardant les choses en détails, on constate que près de 8000 de ces lecteurs ont consulté l’article via sa reprise sur le site Slate.fr :



Image 7 : répartition des lecteurs selon les différents sites

Le format blog, et la dimension « libre » de The Conversation, sont des points pertinents qui nous semblent à valoriser dans le cadre de tels projets en analyse du discours.

Faire des « relations presse » pour appuyer le développement d’un projet

Une autre manière de faire connaître le projet est de répondre aux sollicitations extérieures liées à tel ou tel aspect du projet.

Ainsi, 10 interviews ont été accordées pendant le projet, ce qui a permis la publication d'articles sur le projet dans des journaux tels que *Le Parisien*, *La Montagne*, *Ouest France* (voir la liste complète après la bibliographie).

Les sollicitations de la presse spécialisée ont également permis de faire connaître le projet pour sa dimension technologique. Cette reconnaissance est intéressante car elle apporte de la crédibilité à la dimension technique du projet. Ainsi, le site Elastic, qui met en lumière l'usage des outils de sa suite, tel qu'ElasticSearch, utilisé dans notre moteur de recherche, a réalisé un article⁷ sur le projet :



Image 8 : article sur le site d'Elastic à propos de #Idéo2017

Cette visibilité a eu un effet immédiat, puisque le journal en ligne *Le Monde Informatique*, a ensuite réalisé un article⁸ sur ce sujet. Ce site d'information est dédié aux informaticiens professionnels du monde informatique, et compte plus de 50 000 abonnés à sa newsletter. Ces deux articles liés à la presse spécialisée ont donc permis de mettre en lumière la robustesse technique du projet, et de faire connaître ces recherches auprès d'une plus large cible.

Réaliser des supports de diffusion

Pour finir, nous avons eu recours à l'aide d'étudiants de DUT métiers du multimédia et de l'internet, pour la réalisation de supports

⁷ <https://www.elastic.co/fr/blog/elasticsearch-utilise-au-centre-de-l-application-citoyenne>

⁸ <https://www.lemondeinformatique.fr/actualites/lire-ideo2017-analyse-les-tweets-politiques-a-travers-elasticsearch-69209.html>

de communication (affiches, logos, etc.). Un groupe d'étudiants a également réalisé des démonstrations en vidéos du prototype #Idéo2017, accessibles sur Youtube :

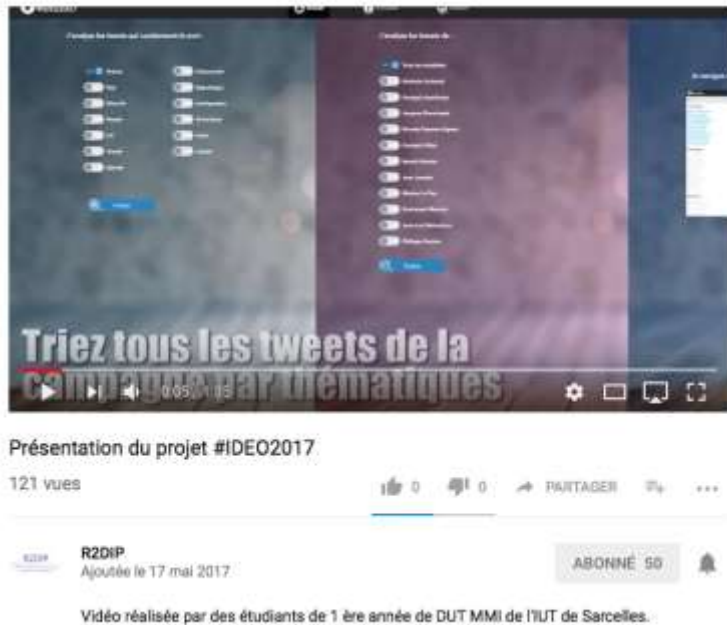


Image 9 : vidéo sur YouTube qui décrit l'usage de la plateforme

Outre l'intérêt pour la communication (certes limité, ici cette vidéo a eu 121 vues), ces supports permettent de montrer, en moins de deux minutes, la palette des fonctionnalités de la plateforme, et de les présenter à de potentiels utilisateurs.

Conclusion

L'expérience du projet #Idéo2017 a permis d'entrevoir plusieurs enjeux pour une recherche tournée vers les citoyens et le changement social. Le développement d'un outil utilisable par le grand public permet d'espérer une efficacité directe sur leur compréhension du discours politique. Pour le chercheur, cela suppose de convertir certaines dimensions, comme par exemple les termes techniques des niveaux d'analyse, ou encore l'ergonomie et l'interactivité. Ce souhait d'impacter le changement social doit aussi s'accompagner de

productions différentes de celle réalisée sur les supports de publication traditionnels : blogs, sites d'informations, vidéos Youtube, réseaux sociaux. C'est donc par une adaptation des formats et discours qui entourent le projet que peut s'effectuer l'implication des citoyens. Pour de prochains projets, comme évoqué au début de l'article, une dimension davantage participative pourrait être envisagée : pour faire émerger des questions de recherche pertinentes ; pour envisager le type de plateforme, l'ergonomie, l'affichage des résultats ; pour co-construire certains scénarios d'utilisation. C'est notamment l'un des enjeux du projet Doxavisu⁹, dont un des objectifs est de produire « un prototype déclinable plus largement à tous les contextes informationnels (et pas seulement politiques) ». Une dimension pédagogique de l'outil est également à l'étude, avec une mise à jour de la plateforme pour les élections européennes de 2019, et l'intégration de son usage dans un MOOC sur les usages du numérique.

Bibliographie

- ALDUY Cécile (2017), *Ce qu'ils disent vraiment. Décoder le discours des presidentialables*. Paris : Seuil.
- GARRIC Nathalie,, LEGLISE Isabelle (2012), Analyser le discours d'expert et d'expertise. In Isabelle Léglise et N. Garric, *Discours d'experts et d'expertise*, Peter Lang, 1-16, <halshs-00731499>.
- LONGHI Julien (2013), Essai de caractérisation du tweet politique. *L'information grammaticale*, 136, 25-32
- LONGHI Julien (2017), Le corpus Polititweets : enjeux institutionnels, juridiques, techniques et philologiques. In Ciara Wigwam et Gudrun Ledegen G, *Corpus de communication médiée par les réseaux : construction, structuration, analyse*. Paris : L'Harmattan, coll. « Humanités Numériques », 37-50
- LONGHI Julien, MARINICA Claudia, HASSINE Nader, ALKHOULI Abdulhafiz, BORZIC Boris (2017), The #Idéo2017 platform, *5th conference CMC and Social Media Corpora for the Humanities*, Bolzano, Italy, 3rd and 4th October 2017 - Conference proceedings, 46-51.
- MARINICA Claudia, LONGHI Julien, HASSINE Nader, ALKHOULI Abdulhafiz, BORZIC Boris (2018), #Idéo2017 : une plateforme citoyenne dédiée à l'analyse des tweets lors des événements politiques, *Revue des Nouvelles Technologies de l'Information*, vol.RNTI-E-34, 459-462.

Interviews accordées pour présenter le projet :

⁹ <http://www.dim-humanites-numeriques.fr/projets/doxavisu/>

- 18 janvier 2017 : Interview pour France culture, « Tweets politiques : prendre les candidats au mot », <https://www.franceculture.fr/politique/tweets-politiques-prendre-les-candidats-au-mot>
- 5 mars 2017 : Interview pour Le Parisien, « #Idéo2017, une plate-forme d'analyse des tweets politiques made in Val-d'Oise », <http://m.leparisien.fr/val-d-oise-95/ideo2017-une-plate-forme-d-analyse-des-tweets-politiques-made-in-val-d-oise-05-03-2017-6734049.php#xtor=AD-1481423553>
- 31 mars 2017 : Interview pour Bonjournalist.eu, « Quand les mots partent en campagne », <http://www.bonjournalist.eu/fr/quand-les-mots-partent-en-campagne/>
- 4 avril 2017 : Interview pour Neue Zürcher Zeitung (Nouvelle Gazette de Zurich), « Wie Le Pen und Macron ihr Publikum verführen », <https://www.nzz.ch/international/wahlen-in-frankreich/wie-le-pen-und-macron-ihr-publikum-verfuehren-stunde-der-populisten-ld.155287>
- 4 avril 2017 : Interview pour Radio RGB (radio locale de Cergy-Pontoise) sur le projet #Idéo2017
- 8 avril 2017 : Interview pour Ouest France, « Présidentielle. Comment les candidats tentent de convaincre en ligne », <http://www.ouest-france.fr/elections/presidentielle/presidentielle-comment-les-candidats-tentent-de-convaincre-en-ligne-4916271>
- 8 avril 2017 : Interview pour Bonjournalist.eu, « Worte wie Waffen », <http://www.bonjournalist.eu/de/worte-wie-waffen/>
- 20 avril 2017 : Interview pour la chaîne VoTv/VoNews, « #Idéo2017 : le site qui décortique les tweets des candidats à la présidentielle (vidéo) », <http://95.telif.tv/2017/04/20/ideo2017-le-site-qui-decortique-les-tweets-des-candidats-a-la-presidentielle-video/>
- 27 mai 2017 : Interview pour Le Journal du Centre, « Festival du Mot : quel impact de Twitter sur l'élection présidentielle ? », http://www.lejdc.fr/charite-sur-loire/loisirs/art-litterature/2017/05/27/festival-du-mot-quel-impact-de-twitter-sur-l-election-presidentielle_12419154.html
- 30 mai 2017: Interview pour le podcast Conversation Express du site TheConversation, « Construction et influence des discours politiques», <https://theconversation.com/podcast-construction-et-influence-des-discours-politiques-77754>

Démonstrations en vidéos du prototype #Idéo2017

Réalisées par des étudiants du DUT MMI de l'IUT de Cergy-Pontoise, site de Sarcelles

<https://www.youtube.com/watch?v=nIzW8ZF7miM>

<https://www.youtube.com/watch?v=3MzWiZ8tBns>

<https://www.youtube.com/watch?v=mexY3nPpO4k>