

Journées d'études 16 & 17 janvier 2020

Sciences Po Paris, 27 rue Saint-Guillaume, 75007

QUELS OUTILS D'ANALYSE POUR LES GILETS JAUNES ?

Approches interdisciplinaires, mais aussi méthodes variées et innovantes tout au long de deux journées d'études

Trois temps d'échanges

- l'étude du mouvement des Gilets Jaunes à travers l'analyse des réseaux sociaux, les approches spatiales et l'ethnographie ;
- l'analyse des Grand et Vrai débats ;
- un retour critique sur les outils techniques et informatiques, et sur les sources de données

Jeudi 16 janvier

- Ethnographie
- Analyses spatiales & temporelles
- Qui dit quoi, comment ?
- Regard d'un garant du Grand débat national, Pascal Perrineau

Vendredi 17 janvier

- Groupes en ligne et réseaux sociaux numériques
- Grand débat et Vrai débat, comparaison
- Sources et critiques de sources

INSCRIPTION ET PROGRAMME SUR METSEM.HYPOTHESES.ORG

METSEM
SÉMINAIRE DE MÉTHODOLOGIE

TRAVAUX
ENQUÊTES
MATE-SHS
MÉTHODES
ANALYSES
Sciences Humaines
et Sociales

sag
Sociétés,
Acteurs,
Gouvernement
en Europe
UMR
7363

ined
INSTITUT NATIONAL
D'ÉTUDES DÉMOGRAPHIQUES

Centre
Emile Durkheim

SciencesPo

Comité scientifique

Diego Antolinos-Basso (Sciences Po, Centre de recherches politiques (CEVIPOF), CNRS & Sciences Po, médialab, Paris, France)

Stéphane Baciocchi (LaDéHiS, EHESS)

Frédéric Cassor (Sciences Po, Centre de recherches politiques (CEVIPOF), CNRS, Paris, France)

Flora Chanvril (Sciences Po, Centre de recherches politiques (CEVIPOF), CNRS, Paris, France)

Estelle Czerny (SAGE, Université de Strasbourg)

Guillaume Garcia (Sciences Po, Centre de données socio-politiques (CDSP), CNRS, Paris, France)

Bénédicte Garnier (Institut national d'études démographiques (INED))

Viviane Le Hay (CNRS, Centre Emile Durkheim, Sciences Po Bordeaux)

Victor Lepaux (SAGE, CNRS)

Jean-Baptiste Pressac (Centre de recherche bretonne et celtique, Université de Bretagne Occidentale)

Paul Salze (LIVE (laboratoire Image, Ville, Environnement), UMR 7362 CNRS/Université de Strasbourg)

Yannick Savina (Sciences Po, Observatoire sociologique du changement (OSC), CNRS, Paris, France)

Invités

Natalia La Valle Torres : référence sociologie à la bibliothèque de Sciences Po

Brice Le Gall : photographe, auteur de *Justice et Respect. Le soulèvement des Gilets Jaunes*, Édition Syllepse, décembre 2019

Pascal Perrineau : garant du Grand débat national

Jeudi 16 janvier 2020

Amphi Chapsal – 6è palier

9h Accueil des participants

9h45 Introduction

Pour le comité scientifique : Flora Chanvriil (Sciences Po, Centre de recherches politiques (CEVIPOF), CNRS, Paris, France)

10h Session 1 : Observer et compter (ethnographie, analyses spatiales et temporelles)

Responsables : Yannick Savina (Sciences Po, Observatoire sociologique du changement (OSC), CNRS, Paris, France)

- **La contestation sous main de justice : Élargir son champ de recherche pour appréhender certains profils de Gilets jaunes et le rapport d'une institution étatique à ces derniers** - Mattéo Giouse (Université Lyon 2 - Centre Max Weber)
- **Les Gilets jaunes et le comptage des manifestants** - Jérôme Collet (Sud énergie)
- **Incidence des manifestations sociales sur la délinquance enregistrée** - Ketty Attal-Toubert, Olivier Filatriau & Tiaray Razafindranovona (Service Statistique Ministériel de la sécurité Intérieure)

12h30 Déjeuner

14h Session 2 : Comment analyser qui a dit quoi et comment... dans le Grand ou le Vrai Débat

Responsable : Luc Rouban (Sciences Po, Centre de recherches politiques (CEVIPOF), CNRS, Paris, France)

- **Qui a participé sur GrandDebat.fr ? Prédiction du soutien aux Gilets jaunes chez les contributeurs à partir d'une deuxième plateforme** - Benjamin Monnery (EconomiX, CNRS & Université Paris Nanterre)
- **Sens et matérialités des contributions citoyennes. Analyse textométrique et discursive du cahier citoyen de Dole (Jura)** - Marion Bendinelli (Université Bourgogne Franche-Comté, EA 4661 ELLIADD, USR 3124 MSHE Claude-Nicolas Ledoux) & Anaïs Rico--Perrier (Université Bourgogne Franche-Comté)
- **Analyse des données issues du Vrai débat** - Jean-Claude Zancarini & Vincent Ventresque (Triangle, ENS Lyon)

15h30 Pause

15h45 Regard d'un garant du Grand débat

Pascal Perrineau (Sciences Po, Centre de recherches politiques (CEVIPOF), CNRS, Paris, France)

16h30 Fin de la journée

Vendredi 17 janvier 2020

Amphi Leroy Beaulieu Sorel – 10^è palier

8h30 Accueil des participants

9h Session 3 : Groupes en ligne et réseaux sociaux numériques

Responsables : Diego Antolinos-Basso (Sciences Po, Centre de recherches politiques (CEVIPOF), CNRS & Sciences Po, médialab, Paris, France) et Frédéric Cassor (Sciences Po, Centre de recherches politiques (CEVIPOF), CNRS, Paris, France)

- **Les déterminants de la mobilisation des “Gilets jaunes”** - Pierre C.Boyer (CREST, Ecole Polytechnique), Thomas Delemotte (CREST, ENSAE), Germain Gauthier (CREST, Ecole Polytechnique), Vincent Rollet (CREST, Ecole Polytechnique) & Benoît Schmutz (CREST, Ecole Polytechnique)
- **Ce que disent les Gilets jaunes, ce que l’on dit des Gilets jaunes** - Dominique Cardon, Jean-Philippe Cointet, Benjamin Ooghe-Tabanou & Guillaume Plique (Médialab, Sciences Po)
- **Les Gilets jaunes sous toutes les coutures. Une approche méthodologique mixte des liens entre populisme et précarité** - Stéphanie Abrial (Pacte-CNRS), Chloé Alexandre (Sciences Po Grenoble, UGA, Pacte-CNRS), Frédéric Gonthier (Sciences Po Grenoble, UGA, Pacte- CNRS) & Tristan Guerra (Sciences Po Grenoble, UGA, Pacte-CNRS)
- **“Gilets jaunes”, inégalités de richesse et populisme** - Laure Bokobza & Jonas Markgraf (Nuffield College, University of Oxford)
- **La place des Gilets jaunes sur YouTube** – Bilel Benbouzid, Emma Gauthier (UPEM-LISIS) & Hervé Guérin (IBM)

11h15 Pause

11h30 Session 4 : *Comment comparer et quelle différence entre Grand et Vrai débat ?*

Responsable : Bénédicte Garnier (Institut national d'études démographiques (INED))

- **Les thématiques du Grand Débat et du Vrai Débat : approche textométrique** - Lucie Loubère & Pascal Marchand (LERASS, Université Toulouse)
- **Se référer au passé, se projeter dans l'avenir. Analyses textuelles et fouille spatiale pour appréhender les horizons temporels des contributions au Grand et au Vrai Débat** - Gaëlle Delétraz, Grégoire Le Campion, Delphine Montagne, Julie Pierson & Olivier Pissoat (Pôle Analyses et Représentations des Données, UMR 5319 PASSAGES / CNRS, université de Pau et des Pays de l'Adour, Université Bordeaux Montaigne, Université Bordeaux et Ecole nationale d'architecture et de paysage de Bordeaux)
- **Analyser le "grand débat national" et le "vrai débat" en confrontant les outils de l'analyse des données textuelles et des algorithmes de type réseau de neurones** - Mathieu Brugidou, Philippe Suignard, Caroline Escoffier & Lou Charaudeau (EDF R&D)

13h Déjeuner

14h Session 5 : Sources et critiques des sources 1/2

Responsable : Guillaume Garcia (Sciences Po, Centre de données socio-politiques (CDSP), CNRS, Paris, France)

- **Derrière les mystères de l'infographie officielle des données issues du Débat national : les apports de trois approches qualitatives de réunions d'initiatives locales** - Cécile Blatrix (AgroParisTech/CESSP), Patrice Diatta (Sciences Po Rennes - Chaire TMAP), Cécile Ferrieux (AgroParisTech/UMR Territoires) & Zoé Kergomard (Institut historique allemand Paris)
- **L'argumentation dans la délibération entre cadre théorique, règles énoncées, mise en œuvre et auto-évaluation** - Bernard Reber (Sciences Po, Centre de recherches politiques (CEVIPOF), CNRS, Paris, France)
- **Recueil de la parole citoyenne et méthode d'analyse de small data territoriales : enjeux analytiques et éthiques** - Hervé Marchal (Université de Bourgogne, Centre Georges Chevrier UMR CNRS 7366), Elodie Mielczareck (Sémiologue indépendante), Louis Ollagnon (Co-fondateur de la société CartoDEBAT), Dominique Valck (Président du Conseil de développement durable du Grand Nancy et co-Président de la Coordination Nationale des Conseils de Développement) & Vladimir Latocha (membre du Conseil de développement durable de la Métropole du Grand Nancy, universitaire)

15h30 Pause

15h45 Session 6 : Sources et critiques des sources 2/2

Responsable : Viviane Le Hay (CNRS, Centre Emile Durkheim, Sciences Po Bordeaux) et Jean-Baptiste Pressac (Centre de recherche bretonne et celtique, Université de Bretagne Occidentale)

- **La complémentarité des outils informatiques dans le traitement des propositions en ligne du Grand Débat National** - Flavie Cernesson, Eric Chraïbi, Jacques Fize, Martin Lentschat, Lucile Sautot (UMR Territoire, Environnement, Télédétection et Information Spatiale, AgroParisTech, CIRAD, CNRS, IRSTEA) & Ludovic Journaux (Laboratoire d'Informatique de Bourgogne, Université de Bourgogne)
- **La reconnaissance optique de marques sur cases cochées au service du Grand débat national** - Diego Antolinos-Basso (Sciences Po, Centre de recherches politiques (CEVIPOF), CNRS & Sciences Po, médialab, Paris, France)
- **Grand débat et IA : quelle transparence pour le traitement des données ? Une tentative de rétro-ingénierie de la synthèse** - Aurélien Bellet (INRIA), Pascal Denis (INRIA), Rémi Gilleron (INRIA et Université de Lille), Mikaela Keller (INRIA et Université de Lille) & Nathalie Vauquier (INRIA)
- **Regard de diplomates sur le matériau documentaire du mouvement des Gilets jaunes : les données en bosse et les données en creux** - Marie-Anne Chabin (Université Paris 8, UFR MITSIC, laboratoire Paragraphe) & Edouard Vasseur (École nationale des chartes)

17h30 Fin de la journée

Session 1 :

**Observer et compter
(ethnographie, analyses spatiales et
temporelles)**

La contestation sous main de justice : Élargir son champ de recherche pour appréhender certains profils de Gilets jaunes et le rapport d'une institution étatique à ces derniers

Mattéo Giouse (Université Lyon 2 - Centre Max Weber)

Pour cette communication je souhaite aborder trois axes qui s'articulent entre eux. D'abord, le surgissement du mouvement des Gilets jaunes dans mon travail de thèse. J'ai commencé, à l'automne dernier, à travailler sur les procès de comparutions immédiates. Deux jours après l'appel national des Gilets jaunes du 17 novembre, des manifestants comparaissent devant la 14ème chambre du TGI de Lyon. Comment intégrer cette donnée nouvelle dans mes méthodes et dans mes pistes d'analyses ? Comment faire évoluer les outils ethnographiques au prisme de cet enjeu inattendu ? Le temps des procès permet aussi une certaine compréhension des profils des Gilets jaunes prompts à sortir du cadre de la légalité. Qui sont ces prévenus « Gilets jaunes » et en quoi diffèrent-ils des prévenus de comparutions immédiates que je qualifie d'« ordinaires » ? Par la visualisation graphique, entre autres, il est possible de saisir ces différences et de dresser des portraits de ces prévenus « Gilets jaunes ». Enfin, l'analyse fine des interactions entre les professionnels de justice et les prévenus, permet de saisir les rapports moraux, politiques et sociaux entre des acteurs d'une institution étatique et les personnes désignées comme sources de désordre. Comment dès lors comprendre ce mouvement social, non seulement en se concentrant sur ses membres, mais aussi en étudiant la réaction du pouvoir politique et judiciaire ?

Bourdieu, « La force du droit. Éléments pour sociologie du champ juridique », *Actes de la recherche en sciences sociales*, 64, p. 3-19, 1986.

Camaille, « Dévoiler la fonction politique de la justice pour favoriser son exercice critique », *Délibérée*, 1, p. 28-32, 2017.

Codaccioni, *Répression : l'Etat face aux contestations politiques*, Paris, Textuel (Petite encyclopédie critique), 90 p, 2019.

Israël, « Question(s) de méthodes. Se saisir du droit en sociologue », *Droit et Société*, 69_70, p. 381-395, 2008.

Rostaing, « L'ethnographie d'un lieu singulier est-elle une démarche comparative ? Réflexions à partir d'enquêtes en milieu carcéral », *Terrain & travaux*, 21, p. 37-54, 2012.

Les Gilets jaunes et le comptage des manifestants

Jérôme Collet (Sud énergie)

Quand les Gilets jaunes sont apparus, le consensus sur les méthodes de comptage des manifestants venait tout juste de s'installer : il a volé en éclats.

Nous montrerons comment ce consensus avait été construit, sur quelles hypothèses explicites et implicites il reposait. Nous constaterons que la variété de leurs formes de manifestations a obligé les Gilets jaunes à innover dans le comptage. Et finalement, nous constaterons qu'il est à peu près impossible de dire si les manifestations de Gilets jaunes sont quantitativement plus ou moins importantes que les manifestations contre, par exemple, la loi el-Khomri.

Café de la statistique du 10 novembre 2015 « Combien de manifestants ? »

https://www.sfds.asso.fr/sdoc-2181-d9dca8578dbcf6901a9cc1af43fd27eb-cr_2015_11_10_manifestants_version_finale_finale.pdf

« Comment est réalisé le comptage indépendant utilisé par « Le Monde » et d'autres médias ? » https://www.lemonde.fr/politique/article/2018/05/26/comment-est-realise-le-comptage-independant-utilise-par-le-monde-et-d-autres-medias_5305180_823448.html

« Pourquoi ce blog? Comment les comptages sont calculés? » <http://gjsciences.com/2019/01/17/pourquoi-ce-blog-comment-les-comptages-sont-calculés/>

Incidence des manifestations sociales sur la délinquance enregistrée

Ketty Attal-Toubert (Service Statistique Ministériel de la sécurité Intérieure)

Olivier Filatriau (Service Statistique Ministériel de la sécurité Intérieure)

Tiaray Razafindranovona (Service Statistique Ministériel de la sécurité Intérieure)

Les manifestations sociales débutées fin 2018 ont eu une incidence sur la délinquance enregistrée par les forces de sécurité. Le service statistique ministériel de la sécurité intérieure (SSMSI) a étudié ce phénomène à partir des données de crimes et délits issues des logiciels de rédaction des procédures de la police et de la gendarmerie nationales. Des premières statistiques par commune sur les sept derniers samedis de l'année ont été produites dans le cadre du bilan statistique annuel 2018 du SSMSI. Les infractions dont la hausse a été la plus forte correspondent à la « participation à un groupement en vue de préparer des violences contre personnes ou des destructions dégradations », aux « destructions et dégradations » et aux « violences à dépositaires de l'autorité ». La représentation cartographique de la variation communale du nombre de faits relatifs aux dégradations et aux violences et outrages à l'égard des personnes dépositaires de l'autorité publique constatés suggère une grande dispersion spatiale du mouvement le samedi 17 novembre 2018 puis une hausse marquée à Paris à partir du samedi 24 novembre 2018 suivie d'une diffusion dans plusieurs grandes villes, notamment Toulouse, Saint-Etienne et Marseille. Des travaux complémentaires intégrant les données de 2019 sont en cours, notamment à l'aide de méthodes d'analyse textuelle et en analysant le profil des personnes mises en cause, en particulier leur profil sociodémographique (publication prévue fin 2019).

Insécurité et délinquance en 2018 : premier bilan statistique - Eclairage complémentaire n°2 : Incidence des manifestations sociales de fin d'année sur la délinquance enregistrée en 2018 – (janvier 2019) - Karim Moussalam (administrateur de l'Insee) – Service statistique ministériel de la sécurité intérieure.

Session 2 :

**Comment analyser qui a dit quoi et comment...
dans le Grand ou le Vrai Débat**

Qui a participé sur GrandDebat.fr ? Prédiction du soutien aux Gilets jaunes chez les contributeurs à partir d'une deuxième plateforme

Benjamin Monnery (EconomiX, CNRS & Université Paris Nanterre)

La plateforme gouvernementale GrandDebat.fr constitue un corpus exceptionnellement riche mais son caractère représentatif pose question. Au-delà de leur code postal, aucune information sociodémographique ou politique individuelle ne permet de connaître les participants au Grand Débat en ligne. Cet article propose une méthode simple pour prédire la position individuelle des répondants de la plateforme sur le mouvement des Gilets jaunes. Nous croisons pour cela les réponses à 14 questions à choix multiples (QCM) données sur GrandDebat.fr à celles de l'application Entendre la France, qui a récolté via Facebook les réponses de 16 000 individus aux mêmes QCM. L'application renseigne également les caractéristiques sociodémographiques des répondants (sexe, âge, niveau de diplôme, CSP, type de commune) et leur niveau de soutien déclaré au mouvement des Gilets jaunes. Les résultats économétriques montrent que, même si les contributeurs de la plateforme gouvernementale sont hétérogènes, leurs réponses au QCM tendent en moyenne à les rapprocher des personnes qui déclarent "ne pas soutenir" les Gilets jaunes. Ces résultats ne semblent pas pouvoir s'expliquer par des différences de composition sociodémographique des participants entre les deux plateformes, mais bien par un processus d'auto-sélection sur une dimension politique.

Algan, Beasley, Cohen, Foucault et Péron, « Qui sont les Gilets jaunes et leurs soutiens ? », *Observatoire du Bien-être du CEPREMAP et CEVIPOF*, 3, 2019.

Bennani, Gandré et Monnery, « Les déterminants locaux de la participation numérique au grand débat national : une analyse économétrique », *EconomiX Working Papers*, University of Paris Nanterre, EconomiX, 7, 2019.

Sens et matérialités des contributions citoyennes. Analyse textométrique et discursive du cahier citoyen de Dole (Jura)

Marion Bendinelli (Université Bourgogne Franche-Comté, EA 4661 ELLIADD, USR 3124

MSHE Claude-Nicolas Ledoux)

Anaïs Rico--Perrier (Université Bourgogne Franche-Comté)

À travers l'étude du cahier citoyen de Dole, nous montrons les enjeux méthodologiques et interprétatifs que ce support pose à l'analyste du discours pour une compréhension fine de sa complexité textuelle et discursive.

Le corpus, 63 contributions numérisées et transcrites, est établi grâce à des métadonnées descriptives et analytiques (XML-TEI et ad hoc) permettant de ne négliger aucun palier de la semiosis – unités lexicales, (péri)textuelles (paragraphe, séquences, éléments structurants), pragmatiques (actes de langage) et argumentatives (prémises, figures). Adoptant une approche textométrique (cadre méthodologique pour l'exploration de données textuelles) et le logiciel TXM, nous listons les spécificités discursives des contributions permettant d'établir des profils types interprétés dans le cadre de l'Analyse du discours. Il en ressort que, si le cahier dolois respecte le contrat de communication qui le sous-tend et sert de lieu où exprimer défiance et mépris à l'égard de la politique mise en oeuvre, il rompt cependant avec le registre des doléances (sont privilégiés verbes de souhait et tournures injonctives). Également, il tend à adopter les codes des écrits intimes (récits individualisés, registre émotionnel) ou des rapports formels (propos thématiques, registre argumentatif). Ce faisant, notre travail décrit les principaux rapports de force et positionnements idéologiques élaborés dans le cahier, interrogeant la définition même du genre cahier citoyen.

Cadres théorique et méthodologique

Adam, *La linguistique textuelle. Introduction à l'analyse textuelle des discours*, Paris, Armand Colin, 2005.

Adam et Viprey, *Corpus de textes, textes en corpus*, Corpus, 8, 2009.

Charaudeau, « De la "scène d'énonciation" au "contrat" et aller-retour », In Angermüller et Philippe (dir.), *Analyse du discours et dispositifs d'énonciation. Autour des travaux de Dominique Maingueneau*, Limoges, Lambert-Lucas, 109-116, 2015.

Habert, Nazarenko et Salem, *Les linguistiques de corpus*, Paris, Armand Colin, 1997.

Halliday et Matthiessen, *Introduction to Functional Grammar*, Londres, Arnold, 2004[1985].

Heiden, Magué et Pincemin, « TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement », In Bolasco (éd.), *Proc. of 10th International*

Conference on the Statistical Analysis of Textual Data - JADT 2010, Edizioni Universitarie di Lettere Economia Diritto, Roma, Italy, 1021-1032, 2010.

Née (dir.), *Méthodes et outils informatiques pour l'analyse des discours*, Rennes, Presses universitaires de Rennes, 2017.

Pincemin, « Sémantique interprétative et textométrie », *Texto!* [en ligne], 17(3), 2012.

Gilets jaunes, Expression citoyenne, Cahiers de doléances

Bronner, « Les démocraties et la méfiance », *Commentaire*, 166, 245-252, 2019.

Chartier, « Cultures, lumières, doléances. Les cahiers de 1789 », *Revue d'histoire moderne et contemporaine*, 68-93, 1954.

Gauchet, « Le révélateur des ronds-points. Notes sur une crise singulière », *Le Débat Dossier « La France des « Gilets jaunes »*, 204, 52-60, 2019.

Genestier, « Les « Gilets jaunes » : une question d'autonomie autant que d'automobile », *Le Débat Dossier « La France des « Gilets jaunes »*, 204, 16-34, 2019.

Gateau, *Les Cahiers de doléances. Une relecture culturelle*, Rennes, Presses Universitaires de Rennes, 2001.

Grunberg, « Les « Gilets jaunes » et la crise de la démocratie représentative », *Le Débat Dossier « La France des « Gilets jaunes »*, 204, 95-103, 2019.

Latour, « Les nouveaux cahiers de doléances. À la recherche de l'hétéronomie politique », *Esprit*, 3, 103-113, 2019.

Legris, « À quoi sert un débat en temps de crise ? », *Études*, 31-42, 2019.

Marchand, Sebbah, Renard, Cabanac, Thiong-Kay, Souillard et Loubère, « « Vrai débat » : Sortir du débat pour négocier », *Rapport scientifique préliminaire d'analyse du « Vrai Débat »*, Laboratoire d'Études et de Recherches Appliquées en Sciences Sociales, Université de Toulouse, 2019.

Mauger, « Compte-rendu de Patrick Farbiaz, Les Gilets jaunes. Documents et textes, Vulaines-sur-Seine, Éditions du Croquant, 2019 Joseph Confavreux (dir.), Le fond de l'air est jaune. Comprendre une révolte, Paris, Éditions du Seuil 2019 Sylvain Bourmeau (dir.), « Gilets jaunes » : hypothèses sur un mouvement, Paris, Éditions la Découverte, 2019 », *Savoir/Agir*, 109-117, 2019.

Pimenova, « Analyse de cahiers de doléances : l'exemple des cahiers de la noblesse », *Mélanges de l'École française de Rome. Italie et Méditerranée*, 103(1), 85-101, 1991.

Schneider, « Lexique du ressentiment », *Le Débat Dossier « La France des « Gilets jaunes »*, 204, 61-75, 2019.

Slatka, « Esquisse d'une théorie lexico-sémantique: pour une analyse d'un texte politique (Cahiers de doléances) », *Langages*, 23, 87-134, 1971.

Slatka, « L'acte de « demander » dans les « Cahiers de doléances » », *Langue française*, 9, 58-73, 1971.

Analyse des données issues du Vrai débat

Jean-Claude Zancarini (Triangle, ENS Lyon)

Vincent Ventresque (Triangle, ENS Lyon)

L'équipe de Triangle a répondu favorablement à la demande de traitement des données issues de la plateforme du Vrai débat. Nous avons donc analysé le corpus en partant de l'hypothèse qu'il s'agissait d'un réservoir d'intelligence collective et en tenant compte des caractéristiques du corpus : présence de votes non binaires (blanc/mitigé), possibilité d'argumenter et de voter pour les arguments.

Afin de dégager un socle programmatique des revendications, nous avons donc sélectionné les propositions les plus consensuelles (un millier environ), en combinant deux critères : a) un score (le nombre de votes favorables moins le nombre des votes défavorables et mitigés) ; b) un pourcentage significatif de votes favorables.

En complément de cette première approche nous avons mis en place un dispositif permettant aux chercheurs d'effectuer des "parcours de lecture outillée" : en utilisant les outils d'analyse statistique et textométrique, les thèmes suivants ont pu être étudiés : privatisation, mobilité (faits), logement, travail/revenu, citoyenneté, égalité/équité (en cours).

Nous présenterons les spécificités du corpus, les contraintes de la commande, et la démarche suivie.

Session 3 :

Groupes en ligne et réseaux sociaux numériques

Les déterminants de la mobilisation des “Gilets jaunes”

Pierre C. Boyer (CREST, Ecole Polytechnique)

Thomas Delemotte (CREST, ENSAE)

Germain Gauthier (CREST, Ecole Polytechnique)

Vincent Rollet (CREST, Ecole Polytechnique)

Benoît Schmutz (CREST, Ecole Polytechnique)

Cet article présente les résultats d'une étude sur les territoires dont sont originaires les « Gilets jaunes ». Dès le premier samedi de mobilisation le 17 novembre 2018, ce mouvement se démarque par son caractère local et sa couverture nationale. A partir de données inédites de la mobilisation sur Facebook, nous montrons une forte corrélation entre mobilisation online (sur Facebook) et offline (blocages des rond-points). Nous réalisons alors une cartographie fine et contrastée de la contestation. L'étude économétrique met en évidence le rôle de la mobilité pour expliquer les origines du mouvement, au travers notamment du passage des routes à 80km/h et des distances domicile-travail.

Acemoglu, Hassan et Tahoun, « The power of the street: Evidence from Egypt's Arab Spring », *Review of Financial Studies*, 31(1), 1-42, 2017.

Algan, Beasley, Cohen, Foucault et Péron, « Qui sont les Gilets jaunes et leurs soutiens ? », *Observatoire du Bien-être du CEPREMAP et CEVIPOF*, 3, 2019.

Arbia, « The configuration of spatial data in regional economics. In Spatial Data Configuration in Statistical Analysis of Regional Economic and Related Problems », *Advanced Studies in Theoretical and Applied Econometrics*, volume 14, Springer Netherlands, 1989.

Enikolopov, Makarin et Petrova, « Social media and protest participation: Evidence from Russia », Available at SSRN 2696236, 2017.

Jayet, « Les zones d'emploi et l'analyse locale des marchés du travail », *Économie et Statistique*, 182, 37-44, 1986.

Malgouyres, « The impact of Chinese import competition on the local structure of employment and wages: Evidence from France », *Journal of Regional Science*, 57(3), 411-441, 2017.

Openshaw et Taylor, « A million or so correlation coefficients: three experiments on the modifiable areal unit problem », *Statistical Applications in the Spatial Sciences*, 127-144. Pion, 1979.

Sebbah, Souillard, Thiong-Kay et Smyrniotis, « Les Gilets jaunes, des cadrages médiatiques aux paroles citoyennes », *Rapport de recherche préliminaire*, Laboratoire d'Études et de Recherches Appliquées en Sciences Sociales, Université de Toulouse, 2018.

Ce que disent les Gilets jaunes, ce que l'on dit des Gilets jaunes

Dominique Cardon (Médialab, Sciences Po)

Jean-Philippe Cointet (Médialab, Sciences Po)

Benjamin Ooghe-Tabanou (Médialab, Sciences Po)

Guillaume Plique (Médialab, Sciences Po)

Cette communication propose de confronter avec des méthodes numériques les formes d'expression des Gilets jaunes au traitement dont ils ont été l'objet dans les médias. Elle prendra appui sur deux corpus réunis par le Médialab. Le premier est une aspiration de grande ampleur des propos tenus dans les groupes publics des Gilets jaunes sur Facebook (aspiration réalisée avec l'outil CrowdTangle). Le second est une capture de l'ensemble des articles publiés par 420 médias français (réalisée avec l'outil Mediacloud en partenariat avec le MIT et le Berkman Klein Center de Harvard). À l'aide d'outils de traitement automatique de la langue, nous souhaitons mettre en œuvre un ensemble de méthodes permettant de comparer et de contraster les thématiques et les modes d'argumentation déployés dans ces deux corpus. Par ailleurs, l'analyse purement textuelle des expressions des Gilets jaunes et des récits journalistiques sera aussi croisée avec des éléments d'identification de la position des énonciateurs (Groupes Facebook de Gilets jaunes et position des médias dans la carte des médias) permettant d'introduire d'autres critères d'interprétation de ce mouvement social.

Les Gilets jaunes sous toutes les coutures. Une approche méthodologique mixte des liens entre populisme et précarité

Stéphanie Abrial (Pacte-CNRS)

Chloé Alexandre (Sciences Po Grenoble, UGA, Pacte-CNRS)

Frédéric Gonthier (Sciences Po Grenoble, UGA, Pacte- CNRS)

Tristan Guerra (Sciences Po Grenoble, UGA, Pacte-CNRS)

Le mouvement des Gilets jaunes (GJ) et le Grand Débat National ont suscité un fort intérêt auprès de la communauté scientifique. Cet intérêt est en grande partie lié à leur visibilité sur internet et, comme l'a souligné l'AMI lancé par l'ANR, aux perspectives d'exploitation innovante rendues possibles par un matériau qualitatif volumineux et directement accessible via les réseaux sociaux et autres forums online. De fait, les questionnements méthodologiques ouverts jusqu'à présent ont un peu passé sous silence l'intérêt des méthodes mixtes plus classiques, mêlant approches qualitative et quantitative.

Notre communication s'appuiera sur des données de première main concernant les caractéristiques sociales et les valeurs politiques des GJ. Ces données ont été collectées dans le cadre d'une enquête par questionnaire diffusée entre décembre 2019 et avril 2020 sur près de 300 groupes Facebook locaux de GJ (N=5321). Le questionnaire documente notamment les attitudes populistes (au sens de Mudde) et la vulnérabilité sociale des répondants sous forme de questions fermées permettant de reconstruire différents indices canoniques, mais aussi de questions ouvertes (QO) permettant de qualifier les raisons du soutien et les registres de discours utilisés par les GJ.

Examinant conjointement « ce qui est dit » et « qui le dit », notre propos se centrera sur l'analyse des liens entre attitudes populistes et précarité. Les deux dimensions seront saisies au travers de mesures quantitatives (i.e., l'indice de populisme d'Akkerman et al., et le score EPICES), d'un double codage inductif (NVivo) et déductif (par une mesure d'holistic grading inspirée d'Hawkins) des réponses aux QO et d'une analyse de discours globale (IRaMUteQ). Cette mise en perspective de différents types de données, d'échelles de mesure et de logiques d'analyse offrira un éclairage original sur le mouvement des GJ.

Akkerman, Mudde et Zaslove, « How Populist Are the People? Measuring Populist Attitudes in Voters », *Comparative Political Studies*, vol. 47, no 9, p. 1324-1353, 2014.

Alexandre, Gougou, Gonthier, Guerra et Persico, « Qui sont vraiment les « Gilets jaunes » ? Les résultats d'une étude sociologique », *Le Monde*, 26 janvier 2019.

Guerra, Alexandre et Gonthier, « Populist Attitudes Among the French Yellow Vests », *Populism*, vol. 2, no 2, 2019 (à paraître).

Hawkins, « Is Chavez Populist?: Measuring Populist Discourse in Comparative Perspective », *Comparative Political Studies*, 42(8), p.1040-1067, 2009.

Hawkins et Silva, « Textual Analysis – Big Data Approach » in Hawkins, Carlin, Littvay et Kaltwasser, *The Ideational Approach to Populism*, Routledge, p.27-48, 2018.

Mudde, « Populism: An Ideational Approach » in Kaltwasser, Taggart, Espejo et Ostiguy (eds.), *The Oxford Handbook of Populism*, Oxford, Oxford University Press, p. 27-47, 2017.

Mudde et Kaltwasser, « Populism » dans Michael Freeden, M. Stears et L. Tower Sargent (eds.), *The Oxford Handbook of Political Ideologies*, Oxford, Oxford University Press, p. 493-512, 2013.

“Gilets jaunes”, inégalités de richesse et populisme

Laure Bokobza (Nuffield College, University of Oxford)

Jonas Markgraf (Nuffield College, University of Oxford)

Ce projet analyse les déterminants géographiques, économiques et sociodémographiques du mouvement des “Gilets jaunes”¹. Notre hypothèse est que ces manifestations sont l’expression d’une insatisfaction et d’une défiance croissantes envers les élites politiques, mais également de griefs et d’anxiété quant aux perspectives économiques individuelles.

Nous lions cela à des variables politiques - soutien aux candidats “populistes” durant l’élection présidentielle de 2017 - ainsi qu’à des mesures d’inégalités de richesse, représentées par l’évolution récente des prix de l’immobilier. Nous nous appuyons sur les travaux liant prix de l’immobilier et préférences pour la redistribution, et avançons que les inégalités de richesse influencent aussi les préférences le long de la “seconde dimension” de la vie politique, via le soutien exprimé pour le populisme dans sa forme nationaliste. A partir des données de Facebook, du site des “Gilets jaunes” et des données du Ministère de l’Intérieur, nous étudions la géographie du mouvement en termes de mobilisation via les réseaux sociaux et d’activité réelle; afin d’établir un lien entre variables socio-économiques (faisant usage de la nouvelle base “Demande de valeurs foncières”) et politiques. Notre objectif est de mettre en lumière des tendances politico-géographiques aidant à comprendre les déterminants et implications de ce mouvement, et de le relier à d’autres formes de contestation en ces temps de montée du populisme.

Ahlquist et Ansell, « Taking Credit: Redistribution and Borrowing in an Age of Economic Polarization », *World Politics*, 69 (4), 640-675, 2017.

Ansell et Adler, « Brexit and the Politics of Housing in Britain », *The Political Quarterly*, 90 (S2), 105-116, 2019.

Boyer, Delemotte, Gauthier, Rollet et Schmutz, « Le Territoire des Gilets jaunes », *Technical Report Note IPP*, n39, Institut des Politiques Publiques (IPP), Paris, 2019.

Algan, Beasley, Cohen, Foucault et Madeleine, « Qui sont les Gilets jaunes et leurs soutiens ? », *Technical report*, Paris, 2019.

Gleizes et Grobon, « Le niveau de satisfaction dans la vie dépend peu du type de territoire de résidence », *Technical Report*, 139, Institut National de la Statistique et des Etudes Economiques (INSEE), Paris, 2019.

¹ Cette ébauche d’article s’ancre plus largement dans le cadre du projet WEALTHPOL - “Politics of Wealth Inequality in the Twenty-First Century”, financé par le European Research Council (ERC). Grant agreement number: 724949.

Grasso et Giugni, « Protest participation and economic crisis: The conditioning role of political opportunities », *European Journal of Political Research*, 55 (4), 663-680, 2016.

Kurer, Häusermann, Wüest et Enggist, « Economic grievances and political protest », *European Journal of Political Research*, 58 (3), 866-892, 2019.

Le Bras, *Gilets jaunes: Cartographie d'une mobilisation*, émission France Culture, url : www.franceculture.fr/emissions/dimanche-et-apres/gilets-jaunes-cartographie-dune-mobilisation

Nyrup, Ansell, Hjorth et Larsen, « Sheltering Populists? House Prices and the Support for Populist Parties », *University of Oxford*, 2018.

Piketty, « Brahmin Left vs Merchant Right: Rising Inequality & the Changing Structure of Political Conflict (Evidence from France, Britain and the US, 1948-2017) », *Technical Report*, N 2018/7, World Inequality Lab, Paris, mars 2018.

Sebbah, Souillard, Thiong-Kay et Smyrniotis, « Les Gilets jaunes, des cadrages médiatiques aux paroles citoyennes », *Rapport de Recherche Préliminaire*, Laboratoire d'Etudes et de Recherches Appliquées en Sciences Sociales, Université de Toulouse, Toulouse, février 2018.

Solt, « Economic Inequality and Democratic Political Engagement », *American Journal of Political Science*, 52 (1), 48-60, 2008.

La place des Gilets jaunes sur YouTube

Bilel Benbouzid (UPEM-LISIS)

Emma Gauthier (UPEM-LISIS)

Hervé Guérin (IBM)

Entre octobre 2018 et juin 2019, nous avons observé la sphère politique et médiatique française sur YouTube. Pour ce faire, nous avons construit un corpus de 1400 chaînes YouTube. Notre étude qui avait pour objectif de développer des méthodes pour se frayer un chemin dans YouTube a été l'occasion de recueillir des données sur la place du mouvement des Gilets jaunes sur la plateforme, que ce soit le traitement médiatique qui en est fait par les différentes chaînes ou encore les chaînes qui soutiennent le mouvement ou qui ont créé des chaînes spécifiques aux Gilets jaunes. Nous avons catégorisé les chaînes de notre corpus selon 12 catégories de « forme d'expression et de communication », dont une que nous avons appelé « Gilets jaunes », c'est-à-dire des chaînes dont l'essentiel du contenu consistait à soutenir le mouvement. Nous avons également recueilli un ensemble de données relatives aux contenus produits par ces chaînes (sous-titres des vidéos) et à leur audience (vues, likes, abonnés etc.). Par ailleurs, nous avons construit trois types de réseaux de chaînes : le réseau des chaînes amis ; le réseau des chaînes selon la recommandation algorithmique ; le réseau des chaînes liées par des commentateurs communs. Nous aimerions dans le cadre de cette table ronde discuter de la place des Gilets jaunes sur YouTube à partir des informations collectées. Nous observons que le réseau des chaînes Gilets jaunes est étroitement lié aux réseaux des chaînes qui appartiennent à la catégorie « Complotiste, confusionniste et pamphlétiste », que ce soit par le réseau de chaînes amis ou les publics de commentateurs. Ceci s'explique en partie par l'étiquetage Gilets jaunes de certaines chaînes très actives avant le mouvement et déjà fortement liées aux milieux « Complotiste, confusionniste et pamphlétiste ». Pour préciser cette explication, nous estimons le nombre de chaînes qui se sont créées avec et pour le mouvement et leur lien avec les chaînes de la catégorie Complotiste. L'analyse textuelle du contenu des vidéos (verbatim des discours dans les vidéos) montre que si les chaînes Gilets jaunes soient proches des chaînes de la catégorie complotiste dans les réseaux, elles ne traitent pas des mêmes contenus. Elles ne partagent pas avec les chaînes « Complotistes, confusionnistes et pamphlétistes » la spécificité des débats sur la « crise migratoire », le « grand remplacement », la sortie de l'Europe et d'autres sujets plus proches des chaînes d'extrême droite ou complotistes.

Munger et Phillips, « A Supply and Demand Framework for YouTube Politics », *Penn State Political Science*, 2019.

Ribeiro, Ottoni, West et al., « Auditing radicalization pathways on youtube », *arXiv preprint*, arXiv:1908.08313, 2019.

Orttung et Nelson, « Russia Today's strategy and effectiveness on YouTube », *Post-Soviet Affairs*, vol. 35, no 2, p. 77-92, 2019.

Session 4 :

**Comment comparer et quelle différence entre
Grand et Vrai débat ?**

Les thématiques du Grand Débat et du Vrai Débat : approche textométrique

Lucie Loubère (LERASS, Université Toulouse)

Pascal Marchand (LERASS, Université Toulouse)

Depuis le Traité de constitution européenne (Mange & Marchand, 2007), puis le Grand débat sur l'identité nationale (Marchand & Ratinaud, 2012), l'analyse textométrique des débats et controverses en ligne n'est pas une grande nouveauté en soi. En revanche, le développement des usages du Web et des réseaux sociaux accroît le nombre de participant-e-s, et donc le volume des données produites, posant de nouvelles questions méthodologiques. Aborder ces données textuelles massives demande donc de mobiliser des compétences techniques, analytiques et interprétatives.

L'objectif sera ici de montrer comment la textométrie permet de décrire, indépendamment, la structure du Grand débat national (GDN) et du Vrai débat (VD), mais également de les comparer formellement.

Ces deux corpus étant de volumes différents (1 932 884 contributions pour le GDN et 117 518 pour le VD), et utilisant des types de données propres (questionnaires pour le GDN et, votes et arguments pour le VD), ils nécessitent des traitements spécifiques de nettoyage et encodage. La première étape consiste à identifier les thématiques stables de chaque corpus, par classification lexicale (Reinert, 1983 disponible dans iramuteq: Ratinaud, 2014). Puis, sur l'ensemble des classes obtenues, nous avons calculé les distances inter-textuelles (Labbé & Labbé, 2003) pour comparer les thèmes des deux débats. Le résultat s'exprime sous forme d'un arbre identifiant les thèmes propres à un des débats et ceux communs aux deux corpus.

Labbé et Labbé, « La distance intertextuelle », *Corpus*, 2, 95-117, 2003.

Mange et Marchand, « Oui ou non à la Constitution européenne : l'éloquence du forum » *Mots : les langages du politique*, 83, 121-137, 2007.

Marchand et Ratinaud, *Être français aujourd'hui. Les mots du « grand débat » sur l'identité nationale*, Paris : Les Liens qui Libèrent, 2012.

Ratinaud, *IRaMuTeQ : Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires* (Version 0.7 alpha 2) [Windows, GNU/Linux, Mac OS X], 2014.

Reinert, « Une méthode de classification descendante hiérarchique : application à l'analyse lexicale par contexte », *Les Cahiers de l'analyse des données*, VIII(2), 187-198, 1983.

Se référer au passé, se projeter dans l'avenir. Analyses textuelles et fouille spatiale pour appréhender les horizons temporels des contributions au Grand et au Vrai Débat

Gaëlle Delétraz (Pôle Analyses et Représentations des Données, UMR 5319 PASSAGES / CNRS, université de Pau et des Pays de l'Adour, Université Bordeaux Montaigne, Université Bordeaux et Ecole nationale d'architecture et de paysage de Bordeaux)

Grégoire Le Campion (Pôle Analyses et Représentations des Données, UMR 5319 PASSAGES / CNRS, université de Pau et des Pays de l'Adour, Université Bordeaux Montaigne, Université Bordeaux et Ecole nationale d'architecture et de paysage de Bordeaux)

Delphine Montagne (Pôle Analyses et Représentations des Données, UMR 5319 PASSAGES / CNRS, université de Pau et des Pays de l'Adour, Université Bordeaux Montaigne, Université Bordeaux et Ecole nationale d'architecture et de paysage de Bordeaux)

Julie Pierson (Pôle Analyses et Représentations des Données, UMR 5319 PASSAGES / CNRS, université de Pau et des Pays de l'Adour, Université Bordeaux Montaigne, Université Bordeaux et Ecole nationale d'architecture et de paysage de Bordeaux)

Olivier Pissoat (Pôle Analyses et Représentations des Données, UMR 5319 PASSAGES / CNRS, université de Pau et des Pays de l'Adour, Université Bordeaux Montaigne, Université Bordeaux et Ecole nationale d'architecture et de paysage de Bordeaux)

Orienté vers l'action et le présent, certaines contributions au Grand et au Vrai Débat se réfèrent malgré tout au passé et à l'avenir, voire à un « futur » un peu plus lointain. Nous formulons l'hypothèse que les quatre thèmes proposés pour « découper » les problématiques sociales et environnementales dans le Grand renvoient à des horizons temporels différents et que les arguments mobilisés sur ce point varient.

Après une analyse des biais et limites du corpus, nous réaliserons une série d'analyses textuelles complémentaires axées sur l'utilisation des conjuguaisons, des adverbes temporels et la réalisation de thésaurus intégrant divers marqueurs chronologiques. Notre objectif est de caractériser quels horizons temporels sont mobilisés dans les contributions.

Dans un premier temps, un panorama général sera proposé. Nous établirons ensuite comment se traduisent les variations de projections temporelles entre le Grand et le Vrai débat ainsi que les différences selon les 4 grands thèmes proposés par le dispositif : Transition écologique, Fiscalité et dépenses publiques, Démocratie et citoyenneté et Organisation de l'État et des services publics. Nous explorerons également le corpus à un

niveau de granularité plus fin afin de repérer certains sujets plus propices aux projections sur le temps long.

Une fois identifiés ces sujets, nous travaillerons leur composante spatiale à travers une fouille statistique spatiale afin d'en relever les éventuelles spécificités géographiques et/ou par typologie d'espaces (urbain/péri-urbain/rural). Pour finir, nous proposerons un rendu graphique, cartographique ou autre de nos résultats.

Projet de recherche : Cette proposition s'insère dans une réflexion plus large menée dans le cadre du projet de recherche PaSciFic « Projections d'Adaptation à travers la Science-Fiction ».

Baschet, « L'histoire face au présent perpétuel - Quelques remarques sur la relation passé/futur », *Les usages politiques du passé*, Editions de l'Ecole des Hautes Etudes en Sciences Sociales, pp. 55-74, 2001.

Barreau, Ducos et Teillant, « La science-fiction, du miroir de nos sociétés à la réflexion prospective », *La note d'analyse – Développement durable*, Centre d'analyse stratégique, n°311, 11 p, 2012.

Gendron, Audet et Girard, « Thinking post-ecological societies with science-fiction », *Les Cahiers de la CRSDD – Coll. Recherche*, n°9, 35 p, 2012.

Hartog, *Régimes d'historicité. Présentisme et expériences du temps*, Paris, Seuil, Coll. « La Librairie du XXIe siècle », 262 p, 2003.

Rumpala, *Hors des décombres du monde, Écologie, science-fiction et éthique du futur*, Champ Vallon, 288 p, 2018.

Analyser le “grand débat national” et le “vrai débat” en confrontant les outils de l’analyse des données textuelles et des algorithmes de type réseau de neurones

Mathieu Brugidou (EDF R&D)

Philippe Suignard (EDF R&D)

Caroline Escoffier (EDF R&D)

Lou Charaudeau (EDF R&D)

Dans cette contribution nous nous proposons d’analyser les énoncés issus de deux plateformes de « civic tech » (Benvegny, 2011 ; Mabi, 2014) - plateforme gouvernementale, le « Grand Débat National » et, sa riposte politique et algorithmique proposée par un collectif de Gilets jaunes, le « Vrai Débat » -, en confrontant deux familles d’algorithmes dédiées à l’analyse de textes. Nous nous proposons de mettre en œuvre, d’une part, des approches éprouvées en analyses des données textuelles (Méthode Alceste sous Iramuteq) qui ont montré récemment leurs intérêts pour l’analyse de très grand corpus (Sebbah et al., 2019 ; Brugidou, 2011) et, d’autre part, des méthodes nouvelles issues du croisement des mondes informatiques, de l’intelligence artificielle et du traitement automatique des langues (Cointet et Parasio, 2018). Nous nous intéresserons plus particulièrement à des familles d’algorithmes dites de « plongement de mots ». Nous chercherons à savoir dans quelles mesures, et selon quelles stratégies d’analyse mais aussi quelles « épistémologies embarquées » (Demazière et al., 2006 : 178), ces méthodes s’ajustent ou non aux formats d’énonciation (questions ouvertes vs échange d’arguments) et aux formats démocratiques (participatif ou délibératif) proposés par ces plateformes. Nous nous interrogerons en outre sur les solutions méthodologiques permettant de qualifier les propriétés sociales des locuteurs sur lesquelles on n’a que peu d’information directe. Enfin, nous tenterons de présenter quelques questions de recherche au croisement de la sociologie politique de l’opinion publique mais aussi des data science, qu’une telle confrontation ouvre.

Benvegna, *La politique des netroots La démocratie à l'épreuve d'outils informatiques de débat public*, CSI, Ecole des Mines, Directeur de thèse : M. Callon, 2011.

Benvegna et Brugidou, « Prendre la parole sur Internet : des dispositifs socio-techniques aux grammaires de la discussion », *Réseaux*, 26 (150), pp.51-82, 2008.

Brugidou, « Le Grenelle de l'environnement : corpus et dispositif d'écriture », *Corpus*, 1 (10), pp.155-178, 2011.

Cardon, *A quoi rêvent les algorithmes. Nos vies à l'heure des big data*, Seuil, la république des idées, 2011.

Demazière, Brossaud, Trabal et Van Meter, *Analyses textuelles en sociologie. Logiciels, méthodes, usages*, Rennes: Presses Universitaires de Rennes, 2006.

Mabi, *Le débat CNDP et ses publics à l'épreuve du numérique : entre espoirs d'inclusion et contournement de la critique sociale*, thèse en science de l'information et de la communication, Sous la direction de Laurence Monnoyer-Smith et de Serge Bouchardon, 2014.

Sebbah, Souillard, Thiong-Kay et Smyrniaios, « Les Gilets jaunes, des cadrages médiatiques aux paroles citoyennes », *Rapport de recherche préliminaire*, Laboratoire d'Études et de Recherches Appliquées en Sciences Sociales, Université de Toulouse, 2019.

Vinck, « Les digital humanities comme promesse pour et par les sciences humaines », in Audétat (éd.), *Sciences et technologies émergentes : pourquoi tant de promesses ?*, Paris : Éditions Hermann, 131-145, 2015.

Session 5 :

Sources et critiques des sources 1/2

*Derrière les mystères de l'infographie officielle des données issues du
Débat national : les apports de trois approches qualitatives de réunions
d'initiatives locales*

Cécile Blatrix (AgroParisTech/CESSP)

Patrice Diatta (Sciences Po Rennes - Chaire TMAP)

Cécile Ferrieux (AgroParisTech/UMR Territoires)

Zoé Kergomard (Institut historique allemand Paris)

Cette proposition de communication s'appuie sur un travail collectif d'enquête sur le Grand Débat National. Saisissant l'opportunité d'un dispositif inédit de débat public, nous avons mis en place dans l'urgence un protocole d'enquête, et conduit cette enquête avec deux objectifs initiaux : comprendre, d'une part, ce que ce dispositif révélait des institutions en charge du débat public, et d'autre part, ce qu'il révélait de l'expression politique et de l'action publique.

Le protocole d'enquête mis en place repose sur deux phases distinctes : une première phase d'observation des réunions d'initiative locale (RIL) dans trois départements (Ain, Ille-et-Vilaine – Puy de Dôme) suivie d'une campagne d'entretiens, principalement auprès des organisateurs de ces réunions.

Nous voudrions revenir sur trois aspects. Le premier concerne plutôt notre démarche d'enquête tandis que le second porte davantage sur les méthodes d'enquête à proprement parler ; le dernier point élargit le questionnement aux enjeux liés à l'analyse/évaluation d'un débat public.

- Le choix initial de se concentrer sur trois départements s'est avéré riche en enseignements et nous a amené à considérer plus sérieusement, la comparaison entre des territoires dont les configurations sont différentes. Par exemple, le positionnement des Gilets jaunes diffère ainsi sensiblement selon les départements considérés. Cette re-contextualisation territoriale permet de faire un pas de côté et de mieux saisir les modes d'appropriation du GDN par les acteurs locaux et les enjeux politiques afférents. Elle nous a permis de consolider notre grille d'analyse initiale et nous a conforté dans la conviction de la nécessité d'une approche qualitative tenant compte du contexte territorial dans ses différentes dimensions.

- Les premières restitutions de l'enquête (colloque) ont été l'occasion d'amorcer de nouvelles collaborations avec des chercheur-se-s étudiant également le Grand Débat. La confrontation de nos méthodes avec par exemple le cadre mis à disposition des chercheur-se-s participant à l'Observatoire des débats, permet, outre d'élargir l'analyse comparée à d'autres départements (Ile-de-France), de questionner les techniques d'enquête, du questionnaire de l'Observatoire jusqu'à l'observation plus fine des réunions et aux entretiens permettant de

comprendre les logiques d'appropriation des acteurs.

- Enfin, l'analyse de l'ensemble des contributions au GDN s'est traduite par un ensemble d'opérations de synthèse et de documents officiels. Le croisement du résultat de nos observations avec les analyses officielles du GDN, permet de questionner les méthodologies de construction du « bilan » qui en a été dressé officiellement.

Faure, « Les passions de l'élu local, du notable au médiateur », Centre d'histoire de Sciences Po, « *Histoire@Politique* » *Politique, culture, société*, n° 25 Vol. 1, p.197 à 211, 2015.

Juhem et Sedel éd., *Agir par la parole. Porte-paroles et asymétries de l'espace public*, Rennes, Presses universitaires de Rennes, 2016.

Sa Vilas Boas, « Ecrire la comparaison lorsque les données sont asymétriques. Une analyse de l'engagement dans les dispositifs participatifs brésiliens », *Revue internationale de politique comparée*, n°1 vol. 19, p.61 à 74, 2012.

Suaud, « Le mythe de la base », *Actes de la recherche en sciences sociales*, vol. 52, n° 52-53, pp. 56-79, 1984.

L'argumentation dans la délibération entre cadre théorique, règles énoncées, mise en œuvre et auto-évaluation

Bernard Reber (Sciences Po, Centre de recherches politiques (CEVIPOF), CNRS, Paris, France)

Si les buts et la justification des méthodes employées pour le Grand débat national n'ont pas été explicités, c'est souvent l'ombre de la théorie de la démocratie délibérative qui a plané sur cette expérience socio-politique inédite. On peut en trouver certaines traces disparates dans les principes des garants et sur la plate-forme du GDN, dans la charte des Réunions d'initiative locale, ainsi que dans les règles énoncées par les organisateurs des Conférences de citoyens régionales. Même si la théorie de la démocratie délibérative n'est pas unifiée et connaît des écarts interprétatifs importants entre ses expressions philosophiques et des travaux empiriques, la version d'inspiration habermassienne polarise son attention sur l'exigence argumentative. C'est la force du meilleur argument qui devrait l'emporter et permettre de faire changer d'avis.

Entre philosophie, enquête empirique (3 questionnaires) et observations de terrain (prises de notes des échanges langagiers), cette analyse se concentre sur ces deux éléments centraux de la théorie et des consignes ou des règles du GDN : l'exigence argumentative et le changement d'avis par la force d'un argument qui viendrait d'autrui.

Cette étude s'appuie sur plusieurs « prises » pour les confronter sur ces points particuliers de l'argumentation et du changement d'avis : 1) les questionnaires Observatoire-Cevipof (2 pages) distribués lors de réunions d'initiatives locales (RIL) auxquelles j'ai pu assister, 2) ceux distribués lors de 2 Conférences de citoyens régionales (6 pages de l'Observatoire-Cevipof + celui des organisateurs Votre Evaluation du Grand Débat), 3) les consignes (charte, règles énoncées par les organisateurs), 4) les notes des échanges langagiers lors des mêmes RIL et autour de deux tables du CCR.

Reber, *Precautionary Principle, Pluralism, Deliberation. Science and Ethics*, Londres, ISTE-international-New York, Wiley, 2016, (247 pages). Publié également en français, *La délibération des meilleurs des mondes. Entre précaution et pluralisme*, 2017 (311 pages).

Reber, *La démocratie génétiquement modifiée. Sociologies éthiques de l'évaluation des technologies controversées*, collection Bioéthique critique, Québec, Presses de l'Université de Laval, 2011, (264 pages).

Reber, « Vertus et limites de la démocratie délibérative », *Archives de philosophie*, Tome 74, 2011.

Reber, « Entre participation et délibération, le débat public et ses analyses sont-ils hybrides du point de vue des théories politiques ? », *Klesis, Revue philosophique, Philosophie et sociologie*, N° 6/1, pp. 46-78, 2007.

Reber, « Technologies et débat démocratique en Europe. De la participation à l'évaluation pluraliste », *Revue Française de Science Politique*, vol. 55, N° 5-6, pp. 811-833, 2005.

*Recueil de la parole citoyenne et méthode d'analyse de small data
territoriales : enjeux analytiques et éthiques*

Hervé Marchal (Université de Bourgogne, Centre Georges Chevrier UMR CNRS 7366)

Elodie Mielczareck (Sémiologue indépendante)

Louis Ollagnon (Co- fondateur de la société CartoDEBAT)

Dominique Valck (Président du Conseil de développement durable du Grand Nancy et co-
Président de la Coordination Nationale des Conseils de Développement)

Vladimir Latocha (membre du Conseil de développement durable de la Métropole du Grand
Nancy, universitaire)

Dans le cadre du Grand débat national, nous avons mis en place une méthodologie de recueil et d'analyse de la parole citoyenne qui se veut éthique et transparente dans le sens où elle est reproductible et où elle explicite ses fondements et autres postulats de base.

Le corpus des données étudiées se distingue de l'ensemble des données nationales. En premier lieu, il est rattaché à un territoire, le bassin de vie de Nancy. En deuxième lieu, il s'agit d'un ensemble de verbatims issus de multiples sources : débats en présentiel, débats en ligne, micro-trottoirs (en collaboration de cinq radios associatives locales) et cahiers de doléances collectées auprès de mairies. En troisième lieu, le rôle d'organisateur du débat local assumé par le Conseil de développement durable de la métropole du Grand Nancy a permis de réunir dans plusieurs ateliers une large diversité d'habitants et en particulier des groupes de Gilets jaunes.

Nous souhaiterions rendre compte des enjeux méthodologiques, sociologiques et éthiques soulevés lors de la phase de recueil et inhérents à l'analyse de ces débats en traitant les points suivants :

- l'ingénierie du débat : circulation de la parole individuelle et collective, travail de synthèse ;
- la représentativité sociologique : légitimité de la prise de parole et typologisation des participants ;
- le traitement numérique et la restitution : méthodologie d'annotation des verbatims et réflexivité analytique des présentations.

Session 6 :

Sources et critiques des sources 2/2

La complémentarité des outils informatiques dans le traitement des propositions en ligne du Grand Débat National

Flavie Cernesson (UMR Territoire, Environnement, Télédétection et Information Spatiale, AgroParisTech, CIRAD, CNRS, IRSTEA)

Eric Chraïbi (UMR Territoire, Environnement, Télédétection et Information Spatiale, AgroParisTech, CIRAD, CNRS, IRSTEA)

Jacques Fize (UMR Territoire, Environnement, Télédétection et Information Spatiale, AgroParisTech, CIRAD, CNRS, IRSTEA)

Martin Lentschat (UMR Territoire, Environnement, Télédétection et Information Spatiale, AgroParisTech, CIRAD, CNRS, IRSTEA)

Lucile Sautot (UMR Territoire, Environnement, Télédétection et Information Spatiale, AgroParisTech, CIRAD, CNRS, IRSTEA)

Ludovic Journaux (Laboratoire d'Informatique de Bourgogne, Université de Bourgogne)

Les contributions du Grand Débat National récoltées en ligne représentent un défi scientifique au vu des enjeux sociétaux. En effet, les contributions sont trop nombreuses et complexes pour être analysées une par une ou par une seule analyse automatique (153809 propositions sur la seule transition écologique). Un collectif pluridisciplinaire composé de géographes, géomaticiens, informaticiens et spécialistes de l'environnement ont testé leurs approches et outils pour décrypter ces questions. L'objectif est ici de proposer différentes approches méthodologiques et montrer les potentialités et contraintes des techniques mobilisées en prenant comme exemple les contributions en ligne sur la transition écologique. Notre dispositif intègre :

- Le stockage des contributions et leur enrichissement par des données exogènes (toponymes, taille de la population, annotations, limites administratives, etc...);
- La fouille des contributions permettant l'extraction automatique de toponymes, de mots clés et de motifs fréquents verbes / mots clés ;
- La visualisation des résultats via des graphiques ou des cartes.

Les verrous sont nombreux : choix de l'échelle spatiale, paramétrage des algorithmes, portée explicative des résultats.

Les approches mises en œuvre ne se substituent en aucun cas à l'analyse et à l'interprétation de spécialistes en sciences politiques, l'objectif est ici qu'elles les complètent et les supportent, en offrant des outils intéressants d'exploration et d'analyse des données.

Akerlof, Maibach, Fitzgerald, Cedeno et Neuman, « Do people “personally experience” global warming, and if so how, and does it matter? », *Global Environmental Change*, vol. 23, no 1, p. 81–91, 2013.

Baeza-Yates et Sáez-Trumper, « Wisdom of the crowd or wisdom of a few? an analysis of users’ content generation », *CoRR*, vol. abs/1604.03044, 2016.

Bennani, Gandré et Monnery, « Les déterminants locaux de la participation numérique au grand débat national : une analyse économétrique », *Document de travail*, 2019.

Borick et Rabe, « Weather or not? examining the impact of meteorological conditions on public opinion regarding global warming », *Weather, Climate, and Society*, vol. 6, no 3, p. 413–424, 2014.

Dumont, « Si les Gilets jaunes m’étaient contés... », *Population & Avenir*, no 742, p. 3-3, 2019.

Egan et Mullin, « Turning personal experience into political attitudes: The effect of local weather on americans’ perceptions about global warming », *The Journal of Politics*, vol. 74, no 3, p. 796–809, 2012.

Morstatter et Liu, « Discovering, assessing, and mitigating data bias in social media », *Online Social Networks and Media*, vol. 1, p. 1–13, 2017.

Tufekci, « Big questions for social media big data: Representativeness, validity and other methodological pitfalls », *CoRR*, vol. abs/1403.7400, 2014.

La reconnaissance optique de marques sur cases cochées au service du Grand débat national

Diego Antolinos-Basso (Sciences Po, Centre de recherches politiques (CEVIPOF), CNRS & Sciences Po, médialab, Paris, France)

L'initiative du "Grand débat national" lancée par Emmanuel Macron le 5 décembre 2018 donne lieu à une série de Réunions d'initiatives locales entre janvier et mars 2019. Durant cette période, au Cevipof (Sciences Po), un dispositif d'enquête d'ampleur s'organise pour permettre d'observer ces réunions. La campagne permettra à près de 50 personnes d'observer un échantillon total de 195 RIL partout en France, et de distribuer et administrer un questionnaire au format papier auprès des participants à ces réunions.

Nous proposons de présenter l'assemblage d'outils qui a permis de numériser les réponses à ce questionnaire et d'en consolider les données. Nous présenterons la chaîne méthodologique qui a permis de passer de réponses au format papier A4 à un fichier tableur unique au format CSV, en passant par le développement d'une application open-source d'Optical Mark Recognition, ou reconnaissance optique de marques sur cases cochées.

Bergeron, « Optical mark recognition », *Postgraduate Medicine*, 104:2, 23-25, 1998.

D'Elia, *RescueOMR: batch Optical Mark Recognition without foresight*, Github, 2017.

Blin, Dealberto et Guillot, « Un public éloigné des traits sociologiques des Gilets jaunes », *Libération*, 2019.

*Grand débat et IA : quelle transparence pour le traitement des données
? Une tentative de rétro-ingénierie de la synthèse*

Aurélien Bellet (INRIA)

Pascal Denis (INRIA)

Rémi Gilleron (INRIA et Université de Lille)

Mikaela Keller (INRIA et Université de Lille)

Nathalie Vauquier (INRIA)

La consultation en ligne du "Grand débat national" a donné lieu à des centaines de milliers de contributions de personnes anonymes, sous la forme de contenu textuel. À cette forme inédite de consultation, il fallait trouver une forme de synthèse qui en restitue la diversité tout en faisant émerger les principales tendances. Le gouvernement a choisi de faire appel à une entreprise privée pour produire une taxonomie question par question des contributions des répondants. Cependant, la méthode suivie pour produire cette taxonomie n'a pas été rendue publique, ce qui pose un problème évident de reproductibilité, limite la mise en regard avec d'autres études et ouvre la porte à des soupçons de manipulation. De surcroît, l'assignation des contributions aux catégories n'est pas non plus publique.

Par des approches classiques (bag-of-words, word embeddings, distance euclidienne), nous tentons de faire la rétro-ingénierie de la taxonomie obtenue et d'en interroger la pertinence. Par cette étude nous entendons mettre en avant le fait qu'il n'y a pas une seule bonne méthode pour faire "parler" les données, et que les choix méthodologiques sont essentiels. Nous étudions également quelle aurait pu être une forme plus transparente de restitution, une forme où au minimum le répondant aurait pu vérifier l'interprétation donnée à sa contribution.

Conseil Scientifique de la Société Informatique de France (SIF) et Bellet, « Grand débat et IA : quelle transparence pour les données ? », *Libération*, avril 2019.

Regard de diplomatistes sur le matériau documentaire du mouvement des Gilets jaunes : les données en bosse et les données en creux

Marie-Anne Chabin (Université Paris 8, UFR MITSIC, laboratoire Paragraphe)
Edouard Vasseur (École nationale des chartes)

De cahiers de doléances en contributions sur la plate-forme du Grand débat national, le mouvement des Gilets jaunes a généré la production d'un riche matériau documentaire. La diplomatique, science développée au XVII^e siècle pour identifier, au milieu des documents authentiques, les faux documents, et qui, plus largement a pour vocation d'identifier les objets d'information fiables pour la preuve et la connaissance, constitue un des outils offrant une méthodologie permettant d'approcher de manière originale ce matériau.

La présente communication a pour but de passer au crible de la diplomatique ce matériau documentaire produit à l'occasion du mouvement des Gilets jaunes, ainsi que les documents périphériques dont il a provoqué la création (presse, commentaires). Elle vise également à analyser leur contexte de création, leur forme et leur mode de diffusion. Elle montrera que, au-delà la prise en compte des seuls mots, de nombreuses informations peuvent être tirées de cette analyse, informations qui pourront s'avérer indispensables au traitement automatisé de ces matériaux.

Chabin, « Peut-on parler de diplomatique numérique ? », in Frey et Treleani (dir.), *Vers un nouvel archiviste numérique*, L'Harmattan, 2013.

Duranti, *Diplomatics: New Uses for an Old Science*, Scarecrow Press, 186 pages, 1998.

Macneil, *Trusting Records: Legal, Historical and Diplomatic Perspectives*, Dordrecht, Boston, London: Kluwer Academic Publishers, 163 pages, 2000.

Angio-Barros, Lhoumeau et Vasseur, « L'archivage électronique dans l'administration publique : expériences croisées de trois services d'archives ministériels dans le contexte réglementaire des archives publiques », in *Quand l'archivage devient électronique*, actes du colloque tenu au ministère des Affaires étrangères et aux Archives nationales, 5-6 février 2013.