

APPEL A PROJETS RECHERCHE 2025

FÉDÉRATION DE RECHERCHE AGORANTIC
«CULTURE, PATRIMOINES, SOCIÉTÉS NUMÉRIQUES »

Titre	Détection automatique de la controverse : étude sur l'impact des éoliennes terrestres.
Acronyme	CONTROV
Nom du/des porteur(s)	Éric TRIQUET (CNE) Juan-Manuel TORRES (LIA)
Coordonnées du/de la gestionnaire de laboratoire	Michèle MANEN michele.manen@univ-avignon.fr Laboratoire Informatique d'Avignon (LIA)
Laboratoires associés	Centre Norbert Elias (UMR 8562) Laboratoire Informatique d'Avignon (LIA)
Budget demandé	8000 €
Résumé Max. 1 000 caractères espaces compris	<p>Depuis la fin des années 1990, les éoliennes terrestres sont pointées par certains éleveurs comme affectant la santé de leurs élevages. D'autres infrastructures électriques ont depuis été mises en cause (lignes HT ou THT ; antennes 4G ou 5G). Toutefois aucune étude n'a permis d'apporter des preuves de causalité. Face aux incertitudes générées par ces innovations, les acteurs concernés expriment des désaccords et du désarroi, et une controverse socio-scientifique se fait jour. Le présent projet porte sur l'expression de cette controverse dans les médias de presse écrite et sur les réseaux sociaux numériques.</p> <ul style="list-style-type: none"> - identifier et caractériser la façon dont ces différents acteurs investissent le champ médiatique (médias traditionnels, médias numériques) pour y exprimer leurs points de vue et y développer leurs stratégies ; - repérer et analyser la façon dont les médias s'emparent de la controverse et introduisent une grille de lecture du problème, voire prennent parti dans la controverse. - Modéliser informatiquement la controverse de façon à pouvoir l'identifier et la classifier. Cette modélisation peut être discrète (modèles RI et graphes) ou bien avec des représentations denses (plongement de mots). - Évaluer empiriquement les modèles proposés au moyen des mesures classiques de précision, rappel et F-score.

1. Contexte, positionnement, objectif(s)

Dans la seconde moitié du 20^{ème} siècle, l'évolution des rapports entre science et société a conduit à l'émergence dans les sciences sociales de la notion de "controverse" pour analyser les enjeux et les conflits provoqués par la diffusion des innovations scientifiques et technologiques dans nos sociétés (Badouard & Mabi, 2015a-b). Aujourd'hui, nous assistons à une augmentation significative des controverses environnementales, non sans lien avec une évolution de la conscience environnementale dans la sphère publique. Elles concernent aussi bien des questions environnementales variées comme les pesticides et le nucléaire (Juan, 2011), le gaz de schiste (Mercado et al. 2014; Molinatti & Simonneau, 2015; Stein, 2018) ou les éoliennes. Face aux incertitudes générées par ces innovations, certains acteurs expriment des désaccords pour lesquels les connaissances produites par la recherche ne permettent pas toujours de trancher, c'est le cas de l'impact possible des éoliennes sur les élevages bovins.

Depuis la fin des années 1990, les éoliennes sont pointées du doigt par certains éleveurs comme affectant la santé de leurs élevages, notamment en élevage bovin-lait : mortalité, baisse de production laitière, diminution de l'appétit, etc. D'autres infrastructures électriques ont été contestées plus récemment et les oppositions sociales ont abouti à des décisions de justice "donnant raison" aux agriculteurs.

La question reste quasiment vierge d'études scientifiques rigoureuses et de résultats débouchant sur des leviers pour y faire face. Si plusieurs types d'infrastructures électriques sont suspectés, une hypothèse raisonnable semble être l'impact possible des champs électromagnétiques ou des courants induits dans l'élevage bovin.

La controverse que le projet se propose d'analyser met en jeu de nombreux acteurs : des éleveurs, des géobiologues, des vétérinaires, mais aussi les services de l'État concernés et un ensemble d'acteurs liés à l'énergie : EDF et autres fournisseurs d'énergie, RTE et Enedis chargés respectivement du transport et de la distribution de l'électricité, France Énergie Éolienne (FEE) qui représente la filière industrielle des éoliennes.

Le présent projet s'inscrit dans un projet plus vaste et pluridisciplinaire : le projet SICECLAIR, financé par l'ADEME, dans lequel s'inscrit une thèse de doctorat préparée par Aurélie Deganello. Il intervient en appui d'un lot centré sur l'étude de l'expression de la controverse dans l'espace médiatique. Il vise quatre objectifs :

- identifier et caractériser la façon dont les différents acteurs engagés dans la controverse investissent le champ médiatique (médias traditionnels, médias numériques) pour y exprimer leurs points de vue et y développer leurs stratégies;
- repérer et analyser la façon dont les médias s'emparent de la controverse et introduisent une grille de lecture du problème, voire prennent parti dans la controverse ;
- établir des modèles informatifs de la controverse basés sur la polarité (à granularité variable) des textes contenant des potentiels débats ;
- définir un protocole d'évaluation des modèles proposés.

L'étude se focalisera dans un premier temps sur les discours produits au niveau de la presse écrite mais un prolongement envisagé concerne l'étude de la circulation de la controverse sur les réseaux sociaux numériques.

2. Questionnement scientifique

Ce projet aura pour objectif d'étudier la genèse et le développement de la controverse socio-scientifique sur les impacts des éoliennes (et autres installations électriques) sur la santé des bovins dans l'espace médiatique.

Le concept de controverse renvoie à des situations de frictions, voire de conflits, entre des acteurs d'origines et de statuts distincts. Les controverses socio-scientifiques peuvent avoir pour origine la diffusion d'innovations technologiques et scientifiques dans nos sociétés (nucléaire, éoliennes, antenne 4G, organismes génétiquement modifiés, nanotechnologies, etc.) Face aux incertitudes engendrées par ces innovations, des acteurs défendent des intérêts divergents et expriment des désaccords que l'état actuel de la connaissance scientifique ne permet pas toujours de trancher. Une controverse désigne ainsi l'expression d'un différend et, au-delà, l'affrontement entre différentes conceptions d'un même problème soutenues par des intérêts divergents (Badouard & Mabi 2015a-b).

Les études sur le traitement médiatique de l'incertitude scientifique montrent que les faits peuvent être définis, interprétés et négociés de différentes manières, et ce d'autant plus dans des situations où les enjeux scientifiques sont incertains et où la notion de risque est socialement aiguë.

L'étude des controverses a été particulièrement marquée par les travaux de Latour et Callon, qui se caractérisent par une approche symétrique des différents acteurs et de leurs arguments, conduisant à la réalisation de cartographies des controverses. D'autres approches, dites "sociotechniques", ont également été développées pour analyser les controverses, comme celle de Chateauraynaud (2011) qui donne la primauté à la portée des arguments et à l'étude de leurs trajectoires afin de rendre visible les véritables "prises" des acteurs sur les enjeux et leur capacité à influencer

les rapports de force. Les travaux de Pestre (2012) s'inscrivent dans une perspective historique et invitent à mettre en évidence les logiques de pouvoir et de domination qui encadrent le déploiement des sciences et des techniques afin de comprendre comment certaines innovations sont imposées par le marché sans être véritablement discutées collectivement. Enfin, les travaux sur les controverses en sciences de l'information et de la communication (SIC) s'inscrivent dans les traditions de ces recherches développées par d'autres disciplines comme la sociologie pragmatique, l'histoire des sciences ou les sciences du langage. Ils portent cependant un regard différent sur les dynamiques de ces situations conflictuelles et sur les formes et les enjeux de leur médiatisation (Le Marec & Babou 2015; Jeanneret, 1998; Cheveigné, 2003). Construites comme un champ interdisciplinaire (Ollivier, 2000), les SIC permettent d'appréhender la controverse socio-scientifique comme un objet politique, scientifique, éthique, économique, communicationnel. Elles se distinguent ainsi des autres approches par leur capacité à traiter des "phénomènes politiques et sémiotiques", en s'attachant aux processus et aux logiques sociales de construction de sens.

Plusieurs questions structurent ce projet de recherche. Tout d'abord, quelle est l'origine de la controverse étudiée et quels sont les termes du débat? Quels sont les acteurs impliqués dans la controverse? Sur quelle légitimité se fonde leur intervention dans le débat? Quels sont leurs enjeux respectifs? Comment déterminent-ils leur capacité à se sentir concernés ou à s'impliquer? Quels sont les arguments avancés et comment s'opposent-ils? Quelles sont les visions du problème qui s'affrontent? Où se situent les zones de tension, les points d'achoppement ou au contraire, les espaces de rencontre et de négociation? Quelles grilles de lecture de la controverse sont activées par les médias? Comment ces derniers influencent-ils le débat? Prennent-ils parti? Comment les acteurs se saisissent-ils de ces espaces? Assiste-t-on à un déplacement des arènes de discussion? À un déplacement des enjeux? À quelles conditions une solution partagée est-elle envisageable? Comment contribuer à la pacification des tensions? Quels axes de réflexion collective proposer aux acteurs de la controverse, quelles pistes de compromis travailler avec eux? Au travers de quels dispositifs d'intermédiation?

La controverse, en tant que phénomène social et linguistique, consiste en une discussion ou un débat répété entre des individus ayant des positions opposées. Aujourd'hui, elle jouit d'une visibilité particulière grâce aux conditions idéales d'une société hyperconnectée, qui ont permis d'enregistrer et d'améliorer l'interaction des utilisateurs en ligne, souvent anonymes, ainsi que la création et la consommation de contenus inédits. L'analyse des propriétés et des caractéristiques de ce phénomène permet d'obtenir différentes informations sur le sujet qui fait l'objet de la controverse : une meilleure compréhension des raisons pour lesquelles un sujet est controversé, sa perception dans la communauté, l'équivalence du phénomène de controverse dans différents domaines et faciliter sa modélisation et le développement d'outils automatiques pour les utilisateurs, parmi d'autres aspects d'intérêt. Cependant, en raison de sa subtilité et de sa dépendance au contexte, sa définition et sa détection restent un problème ouvert (López López, 2020). Cela ouvre un nouveau champ de questions. Comment modéliser formellement la controverse de façon appropriée? Comment détecter des traces de controverses au sein d'un discours? Comment évaluer les éléments linguistiques présents (mots et leur catégorie grammaticale, syntagmes nominaux ou verbaux, tokens dans un sens large, etc) et leur contribution à la controverse?

3. Méthodologie

Pour étudier la genèse et le développement de la controverse étudiés, nous cherchons à repérer et à analyser :

- les représentations, discours et arguments qui circulent à propos de cette controverse (nature et origine du problème, bases scientifiques, enjeux économiques et politiques) ;
- les discours, positionnements et logiques d'acteurs qui se font jour au travers des débats médiatisés par la presse écrite et les réseaux sociaux numériques.

Les sciences humaines et sociales, et particulièrement la sociologie et les sciences de l'information et de la communication, ont grandement contribué à proposer des méthodes d'analyse susceptibles de donner une grille de lecture des phénomènes en jeu dans ce type de situation.

Les analyses seront fondées sur un corpus composé d'articles de presses, d'enregistrements audio et vidéo et de copies d'écran.

Du point de vue des sciences de l'information et de la communication trois approches sont convoquées :

- une analyse sémio-pragmatique : mise en évidence des types d'arguments mobilisés par les acteurs et leurs effets dans les interactions ;
- une analyse des aspects sémantico-rhétoriques : repérage des éléments narratifs, étude des éléments de langage et des procédés argumentatifs ;
- une analyse des jeux d'alliance et de pouvoir, des mécanismes d'influence à partir du modèle de l'analyse systémique et stratégique proposé par Crozier et Friedberg. Il sera nécessaire d'identifier et caractériser ce que ces

acteurs nomment les zones d'incertitude « cruciales » du problème puis de comprendre comment les acteurs s'en saisissent pour imposer leurs grilles de lectures des problèmes en jeu (apparaître comme maîtrisant les zones d'incertitude qui fondent la controverse).

Pour mener à bien ces analyses (des coupures de presse écrites et des publications sur les réseaux sociaux numériques) l'aide des chercheurs du Laboratoire Informatique d'Avignon (LIA - UPR n°4128) apparaît essentielle.

Pour comprendre et analyser de quelle manière la controverse s'exprime dans l'espace médiatique, les approches informatiques semblent en effet pouvoir apporter des entrées méthodologiques pertinentes. Les situations de controverses, comme celle qui se trouve au cœur de ce projet, sont complexes à appréhender et les informations particulièrement difficiles à extraire, notamment lorsqu'elles se déploient en ligne.

L'approche informatique permettra ainsi :

- la collecte et l'assistance pour la constitution d'un corpus étendu ;
- l'analyse des contenus médiatiques (presse dans un premier temps, RSN dans un second temps) ;
- l'interprétation des traces langagières sur des corpus étendus.

Le développement de ce travail de modélisation permettant de rendre compte des modes d'expressions de la controverse nécessitera une étroite collaboration entre des compétences informatiques et le regard développé par les SHS dans l'analyse de situations controversées. En effet, l'analyse des contenus médiatiques ne pourra se faire sans l'élaboration d'indicateurs pertinents. D'un point de vue méthodologique, une première tâche sera consacrée à la constitution d'un corpus de contrôle, annoté, composé d'articles de presse. Les annotations donneront lieu à la définition d'occurrences et de cooccurrences, ainsi qu'à la définition des polarités observables dans les discours.

Les méthodes utilisées concernant à la fois des algorithmes de :

- Recherche d'Information (RI) (Gaussier, 2013) ;
- Traitement Automatique des Langues (TAL) (Torres, 2014) ;
- Représentations classiques et denses (plongements des mots) venant de l'apprentissage profond (DL) (Martin et al., 2020) ;
- Représentations sous forme de graphes des textes portant sur – ou convoquant – la controverse.

Nous proposons de combiner la puissance des méthodes DL et l'explicabilité et la simplicité des méthodes TAL, RI et graphes dans une approche hybride.

Enfin une évaluation des modèles s'avère indispensable pour mesurer les résultats produits par des traitements informatiques utilisés.

4. Résultats attendus et caractère innovant de la recherche

- L'identification d'éléments de communication susceptibles soit de générer/amplifier la controverse, soit de la faire évoluer vers une résolution.
- La réalisation d'une cartographie dynamique de la controverse incluant un repérage et une caractérisation des modes de rationalité en opposition dans la controverse et qui peuvent être mis en débat.
- Une connaissance plus générique des modes d'expression des controverses socio-scientifiques (émergence, fonctionnement, circulation médiatique) et des méthodologies pour leur étude.
- Un choix des modèles informatiques les plus appropriés pour traiter la controverse en français.
- Une étude empirique de l'évaluation des modèles proposés.
- La publication et la diffusion des résultats obtenus.

5. Dimension interdisciplinaire

Sciences de l'information et de la communication :

- l'informatique, en particulier la Recherche d'information et le Traitement Automatique des Langues ;
- l'apprentissage automatique (en particulier l'apprentissage profond) ;
- la modélisation mathématique au moyen des graphes et optimisation.

Les compétences des chercheurs en SIC concernent l'approche info-communicationnelle des controverses. Cette approche permet de se focaliser plus précisément sur les logiques propres aux médias relatives à la production de discours médiatiques à propos de science, mais aussi sur les logiques propres aux différents acteurs relatives à la dimension à la dimension communicationnelle, en prenant en compte les spécificités de la communication par le biais des médias de presse écrite et via les réseaux sociaux numériques.).

Informatique

Du point de vue informatique, un modèle qui pourrait donner des bons résultats est celui de la polarité (au niveau fin avec une échelle des polarités allant du négatif au positif) des phrases ou des morceaux de phrases (à tour de rôle des interlocuteurs) intervenant dans une controverse. En particulier, nous voulons étudier empiriquement l'utilisation de la catégorie grammaticale des termes (noms, verbes, adjectifs, etc.) ou des groupes nominaux ou verbaux lors de ces échanges informatifs. La statistique classique et la représentation au moyen des graphes et d'apprentissage profond seront d'une grande utilité dans cette étude. Notre étude présente un caractère novateur car à notre connaissance, très peu de travaux utilisent ces méthodes dans l'analyse des controverses, et aucun n'aborde l'analyse des controverses portant sur l'impact des éoliennes terrestres sur les élevages bovins.

6. Partenariats extérieurs envisagés

- INRAE – UMR 1114 EMMAH ; IDELE Institut de l'Élevage
- Universidad Veracruzana, Facultad de Matematicas. Mme Lorena AVENDAÑO est spécialiste en graphes et optimisation. Son expertise sera fondamentale pour la modélisation mathématique et informatique prévue.

7. Objectifs de pérennisation du projet

Ce projet s'inscrit déjà dans un appel plus vaste soutenu par l'ADEME. Il est possible que le temps consacré au travail sur la presse écrite ne permette pas d'explorer les réseaux sociaux, auquel cas celle-ci fera l'objet d'un nouveau projet déposé au prochain appel de la FR Agorantic.

8. Expression des besoins en assistance informatique

Nous aurons besoin de 2 stagiaires (L ou M) en informatique pour mettre en place les modèles ainsi que réaliser les tests et leurs évaluations quantitatives vis à vis des corpus de référence ou des annotations humaines. Nous aurons également besoin de visualisations adéquates permettant d'évaluer qualitativement les résultats. Il nous faut pouvoir disposer d'un ordinateur portable performant pour réaliser convenablement les expériences.

1. Évaluation du projet

SIC : - Céline PASCUAL: celine.PASCUAL@univ-amu.fr - Mikael CHAMBRU: mikael.chamburu@univ-grenoble-alpes.fr	Informatique / TAL : - Mathieu ROCHE : mathieu.roche@cirad.fr - Ismail BISKRI : ismail.biskri@uqtr.ca
--	---

2. Budget (€) prévisionnel *

	Brève description	Montant
Missions	Congrès CORIA, TALN ou JADT	3000 euros
Stages***, vacations	Développement en python et Tests des outils informatiques	2 stages x 2 mois à 1200 = 2400 euros
Prestations de service	Annotation des corpus	2600 euros
Budget demandé à Agorantic		8 000 euros
Co financement	Projet SICECLAIR (partie info-com) Ordinateur PC – Colloque SIC, Abonnements Factiva, Le Monde, Libération, Ouest France	3 000 euros
Budget total		11 000 euros

Les directeurs d'unité ont été informés du dépôt de ce projet x

Bibliographie

- Allard-Huver F. (2021). Ce que les SIC font aux controverses environnementales, ce que les controverses environnementales font aux SIC. *Revue Française des sciences de l'information et de la communication*.
- ANSES (2021). Imputabilité à un champ d'éoliennes d'effets rapportés dans deux élevages bovins. Avis ANSES – rapport d'expertise. Saisine 2019-SA-0096 Imputabilité éoliennes, 219 p.
- Badouard R., Mabi C. (2015a). Le débat public à l'épreuve des controverses, *Hermès, La Revue*, 71, pp. 145-151.
- Badouard R., Mabi C. (2015b). Controverses et débat public : nouvelles perspectives de recherche, *Hermès*, 73, pp. 225-231.
- Benslimane, S., Azé, J., Bringay, S., Servajean, M., Mollevi, C. (2021). Controversy Detection: A Text and Graph Neural Network Based Approach. *Web Information Systems Engineering – WISE'21*. LNCS 13080 Springer.
- Carlino V. et Stein M. (2019). *Les paroles militantes dans les controverses environnementales*, Nancy, PUN-Editions.
- Catellani A. et Al. (2019). Les recherches en communication environnementale. État des lieux, *Communication Information médias théories pratiques*, 2, 36, <https://doi.org/10/ggq49b>
- Chateauraynaud F. (2011). *Argumenter dans un champ de force. Essai de balistique sociologique*, Paris, Petra.
- Comby J.-B. (2015). *La Question climatique. Genèse et dépolitisation d'un problème public*, Paris, Liber.
- De Livois A. (2018). La santé des hommes et des animaux face aux infrasons produits par les éoliennes. Colloque Santé et Infrasons Eoliennes (16/11/2018). (<http://docs.wind-watch.org/colloque-sante%E2%80%93et-infrasons-eoliennes.pdf>)
- Gaussier, E; Amini, R (2016), Recherche d'Information - applications, modèles et algorithmes, Eyrolles.
- Guillemaud M. (2021). Des plaintes contre quatre ministres pour « nuisances » d'un parc éolien. *La France Agricole* (publié le 26/01/2021) (<https://www.lafranceagricole.fr/elevage/article/758573/des-plaintes-contre-quatre-ministres-pour-nuisances-dun-parc-olien>).
- Jackiewicz, A (2017), Outils notionnels pour l'analyse des controverses, *Questions de communication*, 31, <https://doi.org/10.4000/questionsdecommunication.11084>
- Juan S. (2011). Environnement et controverses, *Sociologie économique*, pp. 13-64.
- Kazimierz Z, Radoslaw N, Adam W, Adam J, (2018) « Computing controversy: Formal model and algorithms for detecting controversy on Wikipedia and in search queries », *Information Processing & Management*, 54:1, 14-36.
- Kiran G, Gianmarco De Francisci Morales, Aristides G, and Mathioudakis M. (2018), Quantifying Controversy on Social Media. *ACM Trans. on Social Computing*, 1:1, 1–27 <https://doi.org/10.1145/3140565>
- Lemieux C. (2007). À quoi set l'analyse des controverses ? *Mil neuf cent*, 25, pp. 191-12.
- Le Marec J., Babou I. (2015). La dimension communicationnelle de controverses. *Hermès, La Revue*, 73, pp. 111-121.
- López López E. (2020). Un Método para la Detección de Controversia en Textos y su Aplicación al Caso de Comentarios sobre Fármacos en Foros de Salud. Master Thesis, UNED (España). Escuela Técnica Superior de Ingeniería Informática. Departamento de Lenguajes y Sistemas Informáticos.
- Martin, L, Muller, B, Ortiz S, Pedro J, Dupont, Y, Romary, L, De la Clergerie, E, Seddah, D and Sagot, B (2020), CamemBERT: a Tasty French Language Model, 58th Annual Meeting of the Association for Computational Linguistics, 2020. github.com/huggingface/transformers
- Mercado T. et al. (2014). The fracking debate in the media : Therole of citizen platforms as sources of information, *ESSACHESS – Journal for Communication Studies*, 7, 1, 13, pp. 45 62.
- Pascual-Espuny C. (2014). La société civile, de l'alerte à la controverse médiatisée, *Communication et organisation*, 45, 115 126.
- Rigalma K., Duvaux-Ponter C., Gallouin F., Roussel S. (2009). Les courants électriques parasites en élevage. *Inra Prod. Anim.* 22(4), 291-302.
- Molinatti, G., and L. Simonneau. (2015). "A Socioenvironmental Shale Gas Controversy: Scientists' Public Communication, Social Responsibility and Collective Versus Individual Positions." *Science Communication* 37: 190–216.
- Torres,-Moreno, J-M (2014), *Automatic Text Summarization*, London, Wiley
- Triquet E., Molinatti G. (2018). La controverse à propos du statut phylogénétique de Toumai : médiatisations et réputations. *Communication et Langages*, 198, pp. 3-19.