

APPEL A PROJETS 2019

**FEDERATION DE RECHERCHE AGORANTIC
«CULTURE, PATRIMOINES, SOCIETES NUMERIQUES »**

1. Éléments de présentation du projet :

Titre : **Détournements du « genre » dans les représentations politiques des populismes**

Acronyme : **DéGeRPoP**

Nom du porteur : **Christèle LAGIER**

MCF de science politique – LBNC EA3788/FR Agor@nTIC Avignon Université/Cherpa IEP Aix en Provence - christele.lagier@univ-avignon.fr

Laboratoires associés : **LBNC – LIA – ICTT**

Co-porteurs :

Eric San Juan - MCF, Informatique, LIA, Avignon, FR Agor@nTIC

Vincent Labatut - MCF, Informatique, LIA, Avignon, FR Agor@nTIC

Juan Manuel Torres Moreno - MCF HDR, Informatique, LIA, Avignon, FR Agor@nTIC

Tania Jimenez - Ingénieur Recherche LIA, Avignon

Graham Ranger - Professeur d'Université – EA 4277 ICTT, FR Agor@nTIC

Partenaire extérieur :

Société Syllabs (Olivier Hamon)

2. Résumé du projet :

L'objectif du projet DéGeRPoP est d'analyser la contribution de la communauté collaborative Wikipédia à la construction des représentations numériques contemporaines du politique. Cette communauté est ici appréhendée comme une forme contemporaine et collaborative de "Leader d'opinion" opérant une médiation entre le domaine du savoir politique et les citoyens profanes. C'est ce laboratoire de pratiques d'engagement sociétal, datées, archivées et reliées au temps de l'actualité que les partenaires proposent d'analyser en interrogeant les usages détournés de la ressource "genre" par les leaders femmes des formations populistes d'extrême-droite. Par la confrontation au sein d'un corpus inédit de différentes catégories de discours, il s'agit d'interroger à partir d'une étude de cas, les chaînes de déplacements discursifs opérés par ces leaders et leur diffusion au sein de l'espace revendiqué comme neutralisé du

Wikipédia. Les collaborations envisagées s'inscrivent dans le développement d'outils exploratoires partagés favorisant la mise en œuvre d'interfaces publiques de valorisation des archives institutionnelles du Web.

3. Contexte, positionnement et objectif(s) de la proposition

L'objectif du projet **DéGeRPoP** est d'initier l'analyse de la contribution de la communauté Wikipédia à la construction des représentations numériques contemporaines du politique. Forte de près de 18 000 contributeurs actifs (pour plus de 3 millions d'utilisateurs dans le monde), cette communauté interroge certaines problématiques classiques de science politique au-delà de l'étude de sa seule gouvernance (Jacquemin, 2011). Communauté volontairement anonyme, on sait néanmoins des wikipédiens actifs qu'ils appartiennent à des groupes proches de ceux qu'ils référencent de manière régulière et disposent donc d'un capital relationnel et social significatif (Jankowski-Lorek et al, 2016). D'après les sources de la Wikimedia Foundation, le profil du contributeur type de Wikipédia est celui d'un homme blanc, âgé d'une trentaine d'années, informaticien averti et résidant aux Etats-Unis ou en Europe (Adams, Bruckner, 2015). Les travaux de science politique ont de leur côté démontré de longue date combien l'intérêt porté aux questions politiques est lui-même socialement et culturellement distribué (Gaxie, 1978), ce que les usages désormais incontournables du numérique tendent à renforcer (Boyadjian, 2016).

Sur la base de ces éléments, les partenaires formulent l'hypothèse que les contributeurs actifs du Wikipédia puissent être appréhendés comme une forme contemporaine et collaborative de "Leader d'opinion". Dans la suite des travaux de l'école de Columbia (Lazarsfeld P.F et alii, 1968) et face à l'écrasante fréquentation de ses pages, Wikipédia constitue un espace de médiation que l'on peut aujourd'hui difficilement ignorer pour comprendre la construction du rapport des citoyens ordinaires au politique. Dans ce domaine, comme dans d'autres, cette « *encyclopédie des ignorants* », comme d'aucuns la décrivent plutôt péjorativement (Barbe, Merzeaud, Shafer, 2015), joue en effet un rôle dans la diffusion des savoirs, de l'information et des représentations, paradoxalement assez peu étudié en science politique. L'historique du Wikipédia peut, de ce point de vue, être envisagé comme un laboratoire de pratiques sociétales, datées, archivées et qu'il est donc possible de relier au temps de l'actualité.

Cette hypothèse de travail nécessite cependant de pouvoir interroger l'exhaustivité et l'éventuelle partialité du Wikipédia en le recoupant avec d'autres ressources exogènes. C'est ce que propose de faire ce projet en se saisissant d'une question préoccupante pour les démocraties, celle de l'ascension apparemment irrésistible des populismes d'extrême-droite. L'entrée choisie pour l'étude est celle des usages de la ressource "féminité" ou "genre". Sur la base de l'expertise du porteur du projet concernant le terrain français (Marchand-Lagier, 2018, 2017), il s'agit de mettre à l'épreuve d'une étude de cas les usages instrumentaux et détournés de cette ressource par la leader du FN. Sur le modèle d'autres formations populistes d'extrême-droite, traditionnellement "masculinisées", qui maintenaient des positions plutôt conservatrices en matière de famille et sur les questions de genre, le FN/RN semblent investir celles-ci afin de présenter l'immigration comme une menace pour les droits des femmes européennes et/ou blanches notamment (Meret et Siim 2013). Les partenaires du projet envisagent

donc d'objectiver ces formes diffuses de racialisation des rapports sociaux de genre (Scrinzi, 2017) en mobilisant un corpus inédit de production de discours politiques et médiatiques produits à la faveur d'un fait d'actualité convoquant des lectures genrées.

Un travail qualitatif portant sur les périmètres des collectes ciblées est envisagé comme **tâche liminaire** sur la base de l'archive du Web que réalise la BNF dans le cadre de ses missions. Les travaux déjà conduits par d'autres collègues (notamment le laboratoire PACTE-Grenoble) sur les campagnes électorales Présidentielles 2012 et Européennes 2014 ont mis en évidence l'intérêt de l'exploration des métadonnées de collectes au format WAT supportant une représentation structurée de l'archive et constituant un premier niveau d'indexation des contenus archivés (Francony 2014 ; Francony, Stirling 2015). L'archive du Web permet d'envisager une rétro-analyse des principaux événements politiques depuis 2002 qu'il s'agirait de prolonger, dans le cadre de ce projet, lors de la séquence électorale européenne 2019. Scrutin de liste strictement paritaire, ces élections offriront en outre l'occasion d'objectiver la réalité de la place occupée par ces femmes dans le champ politique à partir d'un décompte précis des femmes engagées dans ces formations populistes, des positions occupées sur les listes, de l'étude des plateformes programmatiques et de la place des thématiques « genrées » mais également des supports de communication usant de stéréotypes féminisés. Les travaux scientifiques qui s'intéressent au rôle politique du genre au sein de ces organisations, populistes à l'extrême-droite sont en effet encore rares en science politique (Avanza 2009, Scrinzi 2017) et il y a dans ce projet matière à en objectiver les effets à la fois sur la transformation de l'offre politique et les représentations numériques attachées à celle-ci.

Les données Wikipédia seront de leur côté doublement exploitées pour renseigner non seulement les mutations de ces populismes (qui expliquent les difficultés à les décrire, les définir et les contenir) mais également les canaux par lesquels ils parviennent à transformer les termes du débat politique et social. Elles seront d'abord utilisées pour leur contenu informatif permettant de recenser les caractéristiques socio-démographiques de chacune des leaders identifiées. Elles seront ensuite envisagées dans leur construction même, celle-ci étant en principe neutralisée par les règles de « vigilance participative » (Cardon, Levrel, 2009) qui régissent la communauté. Cette entrée offre ainsi la possibilité de travailler sur les pages web des contributeurs régulièrement inscrits sur lesquelles il est possible de repérer leur capital social (Liu, Ram, 2018), la fréquence de leur contribution aux pages étudiées, leur diversité (Biuk-Aghai et al, 2013), les points de litiges ou encore leurs contributions aux différentes versions des notices aux dates des faits ou temporalités qui auront été préalablement sélectionnés.

4. Résultats attendus et caractère innovant de la recherche

Le projet vise, dans un premier temps, à résoudre concrètement une étude de cas à partir d'un fait d'actualité via son entrée Wikipédia (Agressions sexuelles du nouvel an 2016 en Allemagne (https://fr.wikipedia.org/wiki/Agressions_sexuelles_du_Nouvel_An_2016_en_Allemagne) confrontée aux entrées multiples à disposition dans le corpus numérique constitué :

- Qui en parle ? Distinguer les échos médiatiques (1. quels termes utilisés pour décrire les faits) des prises de positions politiques sur le sujet (2. quels termes utilisés pour analyser les faits)
- Parmi ces prises de position politiques isoler les prises de positions de Marine Le Pen (pour le cas français), puis hiérarchiser les formats de prise de position (tweet le 1.01.16, communiqué, tribune dans l'opinion le 13.01.16...), leur succession (initiative ou réponse à d'autres commentaires) selon la date pour repérer les glissements dans les associations de termes :
 - ❖ sur l'identité du locuteur : "politique">"femme">"libre",>"libertés",>"mère">"européenne">"Française"
 - ❖ sur des références féministes : "Simone de Beauvoir",>"Elisabeth Badinter">"droits des femmes"
 - ❖ sur l'événement : "crise migratoire",>"migrants",>"émeutes sexuelles"
 - ❖ sur la reprise d'éléments rhétoriques de l'extrême-droite : "immigration massive",>"identité nationale">"islamisation">
 - ❖ sur des enjeux de politique nationale : "situation économique">"cohésion sociale">"sécurité">"laïcité"

Pour répondre à ce défi, le projet bénéficie de l'expérience acquise par le LIA dans l'analyse des archives lors des campagnes NIST/TREC en *knowledge base acceleration* et *temporal summarization* ainsi que CLEF/INEX, de même que les travaux de Syllabs en R&D réalisés notamment avec la BNF et le projet RetroNews. Les travaux investis dans le cadre de récents ateliers CLEF tels que la contextualisation de tweets par recherche et extraction d'informations du Wikipédia menés par le LIA seront aussi mobilisés.

Plusieurs verrous technologiques et scientifiques ont été identifiés pour expliquer la sous-utilisation de ce corpus en sciences sociales et nécessitent des développements. Le soutien apporté par la FR Agor@nTIC permettrait d'initier le financement d'une première sous-tâche.

T1 - Extraction d'information et enrichissement : identification des déclarations de personnalités politiques reportées dans le texte du wikipédia, contextualisation de celles-ci en intégrant une analyse de la source citée, les pages de discussions du Wikipédia, et l'exploration des enrichissements des données multimédia utilisables (images) par des approches d'apprentissage automatique. Ceci inclut une analyse de l'intensité des débats associés à un passage précis (analyse de virulence et d'émotions) et l'identification des thèmes relatifs à l'étude des usages de la ressource "féminité".

L'objectif est la publication d'un article collectif des co-porteurs du projet dans *Big data et Society* : <http://bigdatasoc.blogspot.com/p/big-data-and-society.html>.

5. Dimension interdisciplinaire (champs disciplinaires associés) et cohérence par rapport à la thématique « Culture, Patrimoines, Sociétés Numériques »

Le porteur du projet a participé à 2 ANR reliées directement à son champ de spécialité : radicalisation et rapports ordinaires au politique. Spécialiste de l'analyse du vote FN, elle bénéficie en outre d'une expérience en matière de coordination de projets de recherche

interdisciplinaires, notamment avec les chercheurs du LIA, sur les thèmes de la fouille du web et de l'open data (et bien que ce projet n'ait pas fait l'objet de financements Agor@nTIC les années antérieures). Le nouveau master "Innovation, entreprise et société" de l'Université d'Avignon, parcours type "Gouvernance numérique" qui porte sur ces enjeux stratégiques émergents est adossé aux deux laboratoires, LBNC et LIA. Ce dernier est reconnu pour ces recherches en traitement automatiquement de la langue naturelle et ses applications à l'analyse des réseaux sociaux (Bellot et al. 2016). L'association du laboratoire ICTT offre l'expertise nécessaire en matière d'analyse linguistique du corpus étudié.

6. Partenariats extérieurs envisagés

Le LIA collabore régulièrement avec la société Syllabs (spécialisée en web mining, textmining et génération automatique de texte) partenaire du projet et dernièrement dans le cadre de l'ANR-14-CE24-0022 GAFES sur le programme : *Le numérique au service des arts, du patrimoine, des industries culturelles et éditoriales (2014 - 2018)*. Ces deux partenaires ont lancé et animé l'atelier MC2 (Contextualisation Culturelle de microblogs) lors des conférences CLEF 2016 - 2018 (Cossu et al. 2018).

7. Budget prévisionnel (dont financements complémentaires)

| Budget sollicité auprès de la FR Agora@nTIC pour le projet DéGeRPoP | |
|--|--------------|
| | € |
| Frais de personnel (rémunération de stagiaires) | 3 000 |
| Achat de matériel informatique | 1 800 |
| Missions : déplacements Avignon - Paris et frais conférence | 1 200 |
| Total | 6 000 |

| Co-financements | |
|------------------------|-------------|
| | € |
| LIA | 1200 |
| LBNC | 1200 |
| Total | 2400 |