

APPEL A PROJETS RECHERCHE 2020

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

Titre	REseaux des acteurs POLitiques du Grand-Avignon 2020
Acronyme	RePoGA
Nom du/des porteur(s)	Jessica Sainty
Laboratoires associés	LBNC : Jessica Sainty, MCF Science politique Guillaume Marrel, PR Sociologie LIA : Richard Dufour, MCF Informatique
Budget demandé	7 488€
Résumé Max. 1 000 caractères espaces compris	<p>Dans le cadre des élections municipales françaises de mars 2020, le projet pluridisciplinaire RePoGA souhaite mobiliser des sources d'information disponibles en ligne pour comprendre les réseaux d'interaction des acteurs politiques d'un territoire et analyser les contenus qu'ils partagent sur les réseaux sociaux numériques. Le territoire retenu est celui des 16 communes du Grand-Avignon. RePoGA comprend également une dimension historique en proposant d'analyser et de comparer depuis 2008 les réseaux politiques, économiques, sociaux et culturels qui sont repérables dans la constitution des listes de candidat.e.s puis des exécutifs locaux.</p> <p>Les enjeux sont doubles dans ce projet : être capable d'identifier les relations existantes entre des personnalités politiques et d'en définir la nature ; représenter et étudier la propagation de l'information sur les réseaux sociaux, en s'intéressant ici à la manière dont les messages sont rediffusés sur la plateforme d'échange Twitter.</p>

1. Contexte, positionnement, objectif(s) / problématique et cohérence par rapport à la thématique « Culture, Patrimoines, Sociétés Numériques »

Le projet RePoGA répond principalement aux enjeux de régulation politique dans les sociétés numériques du cadre général de la FR Agorantic¹. Il s'agit non seulement de mobiliser des problématiques et méthodologies croisées en sciences politiques et informatiques sur un objectif de sociologie du personnel politique local, mais aussi d'exploiter la diversité des traces numériques de la vie politique contemporaine.

Dans le cadre des élections municipales françaises de mars 2020, le projet pluridisciplinaire RePoGA souhaite mobiliser des sources d'information disponibles en ligne pour extraire et comprendre les réseaux d'interaction des acteurs politiques d'un territoire donné et analyser les contenus qu'ils partagent sur les réseaux sociaux numériques. Le territoire retenu est celui des 16 communes de la communauté d'agglomération du Grand Avignon. L'étude inclut également une dimension historique : en partant des élections municipales de 2008, nous souhaitons à la fois visualiser les transformations de l'identité des acteurs mobilisés et de leurs coalitions, mais aussi pouvoir observer et comparer les contenus partagés par ces acteurs et la structure des relations sur les réseaux sociaux numériques.

Nous avons choisi les élections municipales de 2008 comme point de départ pour ce projet pour trois raisons :

- l'incomplétude des données disponibles dans le Répertoire National des Élus (RNE) de 2001

¹ Néanmoins, la thématique précise de ce projet n'a pas été financé par la FR Agorantic dans le cadre d'un précédent appel à projet.

à 2007 ;

- l'importance prise depuis la fin des années 2000 par la communication politique en ligne - particulièrement depuis la création de Twitter en 2006 ;
- l'importance prise par le niveau intercommunal dans les jeux politiques locaux à partir du milieu des années 2000 : après la loi Chevènement de 1999, le niveau intercommunal va progressivement se structurer et devenir un échelon incontournable (Desage, Guéranger, 2011).

La constitution de cette base de connaissances sur les acteurs politiques du territoire et leurs relations a pour objectif de répondre à une problématique ancienne de la sociologie du pouvoir local, à savoir de repérer au niveau local les acteurs "qui gouvernent ?" (Dahl, 1961). En proposant une cartographie élargie des personnes associées à la vie politique locale, candidat.e.s et élu.e.s, sur le temps long, il s'agit de déconstruire l'apparente stabilité des réseaux de pouvoir, au-delà des quelques personnalités de premier plan qui laissent dans l'ombre les élu.e.s et candidat.e.s, moins médiatiques ou occupant des postes de moindre importance. Cette cartographie des réseaux politiques territoriaux a pour objectif de renseigner les équilibres politiques et institutionnels locaux en retraçant la façon dont les personnes construisent leurs positions politiques en circulant entre les espaces municipaux et intercommunaux, mais aussi entre les identités partisanes, les groupes d'intérêts, les milieux économiques, académiques et culturels, les réseaux militants et associatifs issus de la société civile, etc. Il s'agit donc d'objectiver les circulations individuelles entre les champs politiques, économiques et associatifs, tout en soulignant les modalités de constitution et de reconversion des capitaux politiques et symboliques issus de ces différents champs.

Concrètement, il s'agit d'abord de constituer une base de noms des personnalités politiques élues du territoire en commençant par les titulaires des mandats municipaux et intercommunaux, puis en élargissant à l'ensemble des autres mandats territoriaux et nationaux, à partir du RNE. Ce recensement sera effectué de manière rétrospective de 2008 à 2020. A cette base s'ajouteront les noms de l'ensemble des candidat.e.s aux diverses élections de la même période, à partir des dossiers de la préfecture et des archives du Ministère de l'Intérieur. Nous disposerons ainsi des données socio-biographiques de base pour l'ensemble des candidat.e.s et des élu.e.s du territoire sur les douze dernières années.

L'objectif est ensuite d'explorer les possibilités de fouille et d'extraction automatique d'informations en ligne sur ces listes de noms, afin d'enrichir les données biographiques et contextuelles. En particulier, il s'agira d'être capable de construire des réseaux d'interaction à partir des différentes personnalités politiques étudiées et de leur présence médiatique sur Internet. Deux sources seront alors privilégiées : les archives et publications périodiques de la presse quotidienne régionale (PQR) d'une part et, d'autre part, les comptes et messages des personnalités politiques observés et les comptes des institutions étudiées *via* le réseau social Twitter.

L'analyse de ces réseaux d'interaction s'appuiera sur les techniques de représentation des réseaux complexes, telles que des *graphes d'interaction* (Papegnies *et al.*, 2019), qui présentent de nombreux intérêts pour le projet : outre le fait d'être visualisables, exploitables et interprétables par un expert humain, de nombreuses mesures peuvent être extraites, rendant possible la fouille et l'extraction automatique d'informations dans des graphes de grande dimension, telles que la détection de communautés (Orman *et al.*, 2012). De même, il est possible d'intégrer des informations de nature diverse au sein d'une même représentation, comme les entités nommées (*i.e.* noms propres), le contenu textuel, les relations de proximité (*followers* sur Twitter), etc.

Les enjeux sont doubles dans ce projet :

- Être capable d'identifier les relations existantes entre des personnalités politiques et d'en définir la nature ;
- Représenter et étudier la propagation de l'information sur les réseaux sociaux, en s'intéressant à la rediffusion des messages sur la plateforme d'échange Twitter.

2. Méthodologie et résultats attendus

RePoGA repose sur l'utilisation de techniques quantitatives sur les bases de données déjà constituées, mais aussi sur des outils de fouille et d'extraction automatique de données issues des archives de la presse locale (PQR), et du réseau social Twitter, en partant du Répertoire National des Élus (RNE).

Tout d'abord, et d'un point de vue purement technique, il conviendra de nettoyer les différents documents collectés automatiquement pour la période 2008-2020, extraire les données socio-démographiques disponibles ainsi que les différentes entités nommées (noms propres) présentes. Sans détailler tous les outils, la collecte ne posera techniquement aucun problème, avec des outils éprouvés au LIA pour la collecte sur Internet, ainsi que l'API fournie par Twitter². L'extraction des entités nommées se fera au moyen d'outils existants, tels que LIA_NE : cet étiqueteur développé au LIA a fait ses preuves lors d'anciennes campagnes d'évaluation.

Ensuite, il s'agira d'étudier qualitativement les informations extraites (statistiques sur les entités nommées, relations existantes entre les personnes, présence des personnalités politiques sur Internet selon les périodes...). Ces premières analyses permettront de mieux appréhender les différents corpus collectés ainsi que de fournir des informations chiffrées. Enfin, nous nous concentrerons sur des problématiques de recherche informatiques plus ambitieuses et exploratoires, en particulier sur :

- Extraction d'information dans les graphes d'interaction : Il s'agira ici de construire des réseaux d'interaction entre différentes entités nommées à partir de diverses sources d'information. De nombreux travaux existent sur cette problématique, comme par exemple l'analyse des influenceurs sur Twitter (Cha *et al.*, 2010 ; Bakshy *et al.*, 2011) ou encore la détection de communautés dans un contexte politique (Ozer *et al.*, 2016). Il s'agira d'aller plus loin dans les représentations ici, en prenant pour référence le RNE pour identifier clairement les "réseaux" existants. Puis de les enrichir au moyen des sources de données complémentaires (PQR, Twitter). Au final, nous pourrions être capable de fournir un "résumé", une "carte d'identité", pour chaque personnalité politique visée, au moyen des noeuds proches, des communautés automatiquement identifiées, de la distance avec les autres personnalités.
- La diffusion d'information dans les réseaux sociaux : Ce travail se focalisera sur la problématique de la diffusion d'une information à travers un réseau social, en prenant le cas particulier ici d'un tweet sur la plateforme Twitter. De nombreux travaux se sont intéressés à l'étude de la dynamique des graphes d'interaction, par exemple pour comprendre la création/destruction de liens entre suiveurs (*followers*) (Kwak *et al.*, 2012, Xu *et al.*, 2013) ou encore la prédiction de messages populaires (Hong *et al.*, 2011). Dans ces derniers travaux, les chercheurs s'intéressent au mécanisme de relais de messages. En effet, en considérant ici un message court (un *tweet*), il est possible de savoir qui a relayé ce message (un *retweet*³). Il apparaît néanmoins impossible de connaître le réseau à travers lequel celui-ci a été relayé. Prenons l'exemple où un message de l'utilisateur A a été relayé deux fois, par les utilisateurs B et C. Or C en a eu connaissance via B (car il est ami avec lui) et non par A : il est impossible d'avoir accès à cette information, on ne peut obtenir que le nombre de fois où un message a été *retweeté*. Or cette information peut être essentielle dans un contexte d'analyse de la diffusion d'information dans les réseaux sociaux, et par extension les réseaux d'influence (ici les comptes influenceurs), ce que nous cherchons à définir, explorer et comprendre dans ce projet. A notre connaissance, aucun travail ne s'est penché sur ce

² <https://developer.twitter.com/>

³ <https://developer.twitter.com/en/docs/tweets/post-and-engage/api-reference/get-statuses-retweeters-ids>

problème. Nous pourrions par exemple nous appuyer sur la connaissance des listes d'amis ainsi que des informations d'horodatage des *retweets*, toutes deux rendues disponibles par l'API Twitter : si un message a été retweeté postérieurement à celui d'un ami, nous pouvons considérer que c'est cet ami qui nous a fait découvrir le *tweet*. Au final, nous pourrions connaître le cheminement des messages, mais aussi les influenceurs, qui seraient ici les comptes qui ont permis de relayer de nombreux messages.

Au final, à partir des données issues de l'exploitation de ces trois sources (RNE, PQR, Twitter), nous gardons pour objectif principal un croisement des sources d'information pour caractériser les circulations des individus au sein des réseaux politiques locaux, à la fois dans le temps et entre les espaces municipaux ou intercommunaux.

Dissémination et valorisations :

- Publications : Dissémination scientifique par la publication dans des conférences majeures de chaque domaine, avec *a minima* une publication commune permettant de valoriser le volet interdisciplinaire du projet⁴. En science politique, les résultats seront diffusés dans le cadres des manifestations et publications du projet CREMI (cf. partenariats).
- Open-access et protection des données : Les chercheurs impliqués veilleront également à diffuser le plus largement possible les résultats et les données, tout en garantissant le cas échéant la protection des informations à caractère personnel.
- Valorisation grand public : Il est également prévu une dissémination grand public, lors des manifestations ponctuelles (fête de la science ou café des sciences par exemple) et plus régulièrement dans le cadre des partenariats noués avec les médias locaux pour la campagne municipale 2020.

3. Dimension interdisciplinaire (champs disciplinaires associés)

Outre les aspects et enjeux propres à chaque discipline de recherche (informatique et science politique) précédemment évoqués, nos intérêts se retrouveront sur des questions communes :

- La constitution des réseaux politiques locaux, en se focalisant sur la façon dont les élu.e.s circulent entre les espaces municipaux et intercommunaux, mais aussi au sein de différents univers partisans, à travers le traitement automatique de grandes masses de données;
- L'enracinement de ces réseaux politiques locaux dans les espaces municipaux et intercommunaux, que les différentes sources d'information collectées devraient pouvoir faire émerger, au moyen des graphes d'interaction et de leur analyse;
- Les mécanismes de renouvellement des acteurs politiques locaux, que les travaux que nous souhaitons mener en informatique pourraient permettre d'appréhender, en particulier sur l'aspect temporel de la collecte;
- Les effets des changements des règles du jeu politique : nouveau mode de désignation des élus intercommunaux depuis 2014, restriction du cumul des mandats depuis 2017;
- Les effets des transformations politiques nationales induits par l'apparition de La République en marche (LREM), les restructurations des autres partis politiques (PS, LR, RN, EELV, LFI notamment), l'irruption de mouvements citoyens comme les Gilets Jaunes sur la scène politique : ces transformations devraient apparaître dans l'analyse des communautés au moyen des graphes d'interaction.

⁴ Par exemple : <https://www.icwsm.org/2020/index.html>

4. Partenariats extérieurs envisagés

Ce projet s'inscrit dans le cadre d'une enquête nationale de science politique dans 15 grandes villes françaises à l'occasion des élections municipales et intercommunales de 2020, portée par le Collectif de Recherche sur les Élections Municipales et Intercommunales (CREMI). Le collectif associe des politistes, des sociologues et des géographes, spécialistes des questions locales, avec l'objectif de coordonner les recherches qualitatives et quantitatives, réalisées dans ces villes lors de ces élections. Ce collectif est financé en partie par des partenariats avec les médias nationaux (Le Monde, France Inter) et avec des associations d'élus locaux (ADCF et ADGCF). Nous prévoyons également de développer un partenariat avec un média local (La Provence, Vaucluse Matin ou France Bleu Vaucluse) pour diminuer le coût d'accès à ces ressources, fondé sur la diffusion des résultats des recherches auprès de notre partenaire.

Budget (€)		
	Brève description	Montant
Missions	Coordination RePoGA-CREMI	500
	Conférence internationale	1 500
Documentation	Abonnements PQR	300
Stages**	M2 informatique 6 mois	3 480
Post-doc	IGE informatique 1.5 mois ETP (INM388)	3708
Budget total		9 488
Cofinancements le cas échéant	CREMI	2 000
Budget demandé à Agorantic		7 488

5. Bibliographie

- Bakshy E., Hofman J. M., Mason, W. A., & Watts, D. J. (2011). Everyone's an influencer: quantifying influence on twitter. In Proceedings of the fourth ACM international conference on Web search and data mining (pp. 65-74).
- Cha M., Haddadi H., Benevenuto F., Gummadi K. P. (2010). Measuring user influence in twitter: The million follower fallacy. In fourth international AAAI conference on weblogs and social media.
- Dahl R. (1961), *Who governs? Democracy and power in an American city*, Yale University Press.
- Desage F., Guéranger D. (2011), *La politique confisquée. Sociologie des réformes et des institutions intercommunales*, Éditions du Croquant
- Hong L., Dan O., Davison B. D. (2011). Predicting popular messages in twitter. In Proceedings of the 20th international conference companion on World wide web (pp. 57-58).
- Kwak H., Moon S., Lee W. (2012). More of a receiver than a giver: why do people unfollow in Twitter?. In Sixth International AAAI Conference on Weblogs and Social Media.
- Papegnies E., Labatut V., Dufour R., Linarès G. (2019). Conversational networks for automatic online moderation. IEEE Transactions on Computational Social Systems, 6(1), 38-55.
- Orman G. K., Labatut V., Cherifi H. (2012). Comparative evaluation of community detection algorithms: a topological approach. Journal of Statistical Mechanics: Theory and Experiment, 2012(08), P08001.
- Ozer M., Kim N., Davulcu H. (2016). Community detection in political Twitter networks using Nonnegative Matrix Factorization methods. In Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (pp. 81-88).
- Xu B., Huang Y., Kwak H., Contractor N. (2013). Structures of broken ties: exploring unfollow behavior on twitter. In Proceedings of the 2013 conference on Computer supported cooperative work (pp. 871-876).

6. Annexe - Planning général estimé du projet RePoGA

	Mois 1	Mois 2	Mois 3 (élections municipales : 15 et 22 mars 2020)	Mois 4	Mois 5	Mois 6	Mois 7	Mois 8	Mois 9	Mois 10	Mois 11	Mois 12
Collecte des données (PQR, données préfecture et ministère, Twitter, etc.)												
Nettoyage des données												
Extraction des entités nommées												
Création des graphes d'interaction												
Extraction d'information des graphes												
Analyse de la propagation sur Twitter												
Analyse des réseaux politiques locaux												
Publication commune et participation à des congrès												
Réunions	Réu. 1	Réu. 2		Réu. 3		Réu. 4			Réu. 5		Réu. 6	
Discussions sur le futur du projet												
Dissémination dans des conférences grand public												
Compte-rendu final												