

APPEL À PROJETS 2022 – DONNÉES & OPEN DATA

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

Titre	Base Révisée des Élus de France – Version 2
Acronyme	BRÉF 2
Nom du/des porteur(s)	Marrel Guillaume, Labatut Vincent, Volpi Emilie, Noémie Févrat
Laboratoires associés	LBNC, LIA
Budget demandé	8 044€
Résumé <i>Max. 1 000 caractères espaces compris</i>	La Base Révisée des Élus de France (BREF) intègre les données ouvertes du Répertoire National des Elus (RNE) et d'autres bases des assemblées politiques du territoire national. L'équipe qui la porte sollicite un financement pour le recrutement d'un ingénieur contractuel afin de finaliser, d'ici fin 2022, les intégrations de données inachevées et l'interface de consultation nécessaire au projet d'ouverture de la base aux communautés scientifiques intéressées.

1. Contexte, positionnement, objectif(s)

La BRÉF est une base de données relative au personnel politique élu en France, depuis le début de la Cinquième République, qui a vocation à faire autorité dans le monde académique.

Ce projet est né de la mise à disposition en OpenData des données des élus par le ministère de l'Intérieur. C'est un projet central à Agorantic qui permet à plusieurs projets d'obtenir des données pour ensuite les croiser.

La première version a été élaborée en 2020 dans le cadre de la thèse de Noémie Févrat hébergée au LBNC et financée par la FR Agorantic. Cette version se base principalement sur les données du Répertoire National des Elus (RNE), fournies par le ministère de l'Intérieur en juillet 2018. Ces données sont supposées couvrir la période 2001–2018, mais en pratique elles sont incomplètes et contiennent de nombreuses erreurs. Afin de pallier ces problèmes, un travail de croisement avec d'autres sources est nécessaire. Plusieurs sources ont été identifiées dans ce but, et la V1 de la BRÉF en intègre certaines : BD de l'Assemblée Nationale (2 549 mandats), BD du Sénat (3 272 mandats), BD du Parlement Européen (893 mandats). Ceci a permis de résoudre certaines erreurs, d'intégrer des mandats manquants, mais aussi d'étendre la couverture temporelle de la base..

La version 2 de la BRÉF est actuellement en cours d'élaboration. Ceci implique le passage à une technologie plus adaptée, reposant sur un système PostgreSQL, et l'intégration de sources supplémentaires : données du site collaboratif Politiquemania, obtenues à la suite d'une prise de contact avec son gestionnaire, données issues des archives nationales et saisies manuellement par la doctorante, mais aussi actualisation du RNE provenant du nouveau système mis en place récemment par le ministère de l'Intérieur (RNE2), base Sycomore de l'Assemblée Nationale, base historique du Sénat, et d'autres issues notamment de sources opendata (data.gouv.fr). La version 2 de la BREF a déjà été enrichie des mandats issus de Politiquemania, ainsi que de quelques mandats issues des archives départementales de la Côte d'Or et comprend à présent 1.384.133 positions électives occupées par 907.738 élus.

Ces deux dernières années, l'administratrice SI de la FR a beaucoup travaillé sur ces tâches et encadré 2 étudiants de L3 et 4 étudiants de M1 d'informatique, qui ont travaillé sur la BREF, ses intégrations et son interface. Aujourd'hui le projet BREF est un système d'information composé d'une base PostgreSQL, de fonctions PLpgSQL, d'une interface graphique Angular tournant sur un serveur NodeJS. Ce système doit être complété et renforcé rapidement et la charge de travail actuelle de l'administratrice ne permet pas d'être à la hauteur de nos ambitions de consolidation.

L'équipe de portage de la BREF sollicite donc la FR Agorantic dans le cadre de l'AAP 2022 Données & Open data afin de recruter un ingénieur d'étude contractuel. L'objectif du CDD recherche de 2 mois est

de finaliser toutes les intégrations de données dans la BREF qui ont été entreprises depuis 2 ans :

- intégrer l'ensemble des données de candidatures sur toutes les élections disponibles en open-data (<https://www.data.gouv.fr/fr/pages/donnees-des-elections/>)
- intégrer les 6 extractions OpenData du RNE, réalisées par le ministère de l'Intérieur tous les 3 mois, entre octobre 2018 et décembre 2020, dont nous avons récupéré une version augmentée des nuances politiques de chaque élu non disponible en ligne
- intégrer une nouvelle version historique du RNE2, en cours de récupération auprès du ministère
- intégrer les dernières données des bases Sycomore de l'Assemblée Nationale et la base historique du Sénat
- intégrer toutes les données restantes de la base de données Politiquemania qui n'ont pas encore été intégrées : intégrer les données candidatures des élections, comparer le découpage des territoires de Politiquemania, toutes les données de candidature, fonction et mandat qui sont incorporé dans du HTML et difficile à extraire
- intégrer les données de 4 fichiers Excel et 2 bases de données Access appartenant à Guillaume Marrel
- finaliser l'interface conviviale et sécurisée d'accès distant aux données
- intégrer les données de l'assemblée de Martinique, de l'assemblée de Corse, des maires d'arrondissements de Lyon, Marseille et Paris
- revoir tous les problèmes déjà identifiés : mandats sénatoriaux sans circonscription, utilisation du nom martial, chevauchement mandats conseiller municipal et EPCI

2. Questionnement scientifique en rapport avec l'intitulé de l'appel

Les questionnements scientifiques relatifs à la constitution d'une telle base de données et à son ouverture aux chercheurs en sciences sociales du politique sont innombrables et les usages possibles difficiles à anticiper. Évoquons seulement les motivations documentaires de la centralisation, de la mise à jour, de l'enrichissement continu et du contrôle d'une base recensant l'identité des élu.e.s de France : la mémoire et l'analyse sociologique des représentants élus du peuple, au plan local et national, participent des missions d'introspection académique importantes dans toutes les démocraties pluralistes. L'ouverture à terme de ces données à l'ensemble des scientifiques qui motiveront leurs recherches prolonge cette entreprise d'amélioration de l'accessibilité à la connaissance de la représentation démocratique. La BREF pourra à terme servir à des comparaisons internationales. Elle sert d'ores et déjà à l'étude des pratiques et usages du cumul des mandats, de la longévité politique, des marchés publics et de l'e-reputation sur Wikipedia.

3. Méthodologie

L'ingénieur de recherche mobilisera les méthodologies adaptées aux défis techniques des tâches de développement listées ci-dessus.

4. Résultats attendus et caractère innovant de la recherche

Le but original de la BRÉF est de rassembler des données longitudinales exploitables dans le cadre de la thèse de Noémie Févrat. Cependant, la ressource constituée est susceptible d'être utilisée à de nombreuses autres fins.

Dans sa V2, la BRÉF a d'ores et déjà été complétée de manière à pouvoir être mise en correspondance avec la FOPPA (French Open Public Procurement Awards database), la BD constituée dans le contexte du projet ANR DeCoMaP (Détection de Corruption dans les Marchés Publics) porté par Avignon Université. L'enjeu est ici d'identifier individuellement les élus impliqués dans le passage de marchés publics, afin de pouvoir exploiter cette information autrement absente des bases décrivant les marchés.

Ce type d'interconnexion permet d'enrichir les deux bases, et donc de démultiplier les possibilités d'analyse.

Il faut souligner que plusieurs chercheurs extérieurs à la FR ont manifesté leur désir d'avoir accès à la BRÉF, alors même que l'existence de la base n'avait pas encore été rendue publique, ce qui révèle un intérêt certain de la part de la communauté scientifique. Cet intérêt est notamment renforcé par le fait

que le Bureau des élections du Ministère de l'Intérieur renvoie désormais vers nous toute demande académique d'accès au RNE en dehors des données publiées en open data.

5. Dimension interdisciplinaire

Le projet est nativement interdisciplinaire et mobilise des chercheurs en science politique et science informatique. La dimension d'open-science prévue à terme renvoie à toutes les sciences sociales et computationnelles.

6. Partenariats extérieurs envisagés

La BREF liée au RNE repose sur un partenariat avec le Bureau des Elections du Ministère de l'Intérieur. L'intégration à venir de l'ensemble des bases académiques éparses relatives aux personnels politiques en France ouvre d'innombrables projets de partenariats en histoire politique contemporaine, science politique et sociologie politique, en France et à l'étranger.

7. Valorisation (si prévue) : déclaration d'invention permettant de valoriser un savoir-faire, une base de données ou un logiciel

La base de données est accessible en ligne sur demande à l'équipe, via le client PgAdmin qui demande tout de même un peu de connaissance technique. Une interface graphique permettant un accès beaucoup plus convivial et simple aux données est en cours de développement en ce moment même. La demande de financement a pour but de finaliser cette interface, en plus de consolider les données de la base.

	Budget (€)*	
	Brève description	Montant
Missions	Mission de 2 mois d'un ingénieur informaticien en CDD recherche à l'indice 554	8044€
Consommables, petits matériels**		
Organisation de réunions		
Stages***, prestations	ou prestation informatique	idem
Budget total		8044€
Co financements le cas échéant		
Budget demandé à Agorantic		
Recettes extérieures		

* Veuillez modifier les catégories de dépenses si besoin – ajoutez/supprimez des lignes à votre convenance

** Petit matériel ne dépassant pas les 600€

***Gratification de stage obligatoire au-delà de 2 mois – prévoir environ 580€ par mois