

APPEL A PROJETS

AMORÇAGE 2019

FICHE BILAN ACTION

Adresse de dépôt : katel.lochet@univ-nantes.fr

NOM DU PROJET : GTBDML - Groupe de Travail Big Data et Machine Learning

1 – Bilan des actions réalisées

Action principale (Objectifs, programme, date, lieu, nombre de participants)

1- Participation à l'Ecole d'Hiver de la fédération de recherche TEEP

Une seule dépense a été faite à l'aide de support : le financement des frais des participations d'un participant à la 11^e Ecole d'Hiver de la fédération de recherche TEEP qui proposait une journée thématique sur les Big Data.

Lien vers l'évènement :

<https://sites.google.com/view/ecole-cnrs-tepp/editions-précédentes?authuser=0>

Intervenants et titre de la présentation en lien avec la thématique du groupe de travail :

- *Emmanuel Flachaire (Université Aix Marseille) : « Econométrie et Machine Learning »
- *Ewen Gallic (Université Aix Marseille) : « Etude de la démographie Française du XIX^e siècle à partir de données collaboratives de généalogie »

Participant financé :

Sylvie Blasco

Objectifs de la participation à cet évènement :

S'initier à la thématique et prendre des contacts de chercheurs travaillant sur la thématique.

Date et lieu :

du 18 au 22 mars 2019, Aussois, France.

Retours de cette action auprès des membres du groupe de travail :

Présentations sur le big data, transmission des supports des présentations aux membres du groupe big data et échanges avec des chercheurs-enseignant enseignant le big data dans les formations d'économie.

2 – Invitations de chercheurs et professionnels pour des présentations

Des invitations initialement prévues n'ont pas pu être concrétisées par des présentations d'intervenants extérieurs du fait d'incompatibilités d'emploi du temps (un des intervenants pressentis était en congé de recherche sur la période au Canada et n'a donc pas pu être invité).

En marge du groupe de travail, un membre du groupe, Salima Bouayad-Agha a contribué à l'invitation d'un intervenant extérieur, Pierre Blanc qui a donné une conférence le 11 mars 2020 sur l'intelligence artificielle et le management des ressources humaines

Autres résultats (Publications, numérique, actions de formation, médiation grand public, etc.)

-

2 – Bilan scientifique

1 – Réponse à un appel à projet régional de grande ampleur

En lien avec les thématiques abordées dans le groupe de travail, des membres économistes de de l'équipe ont répondu à l'appel à projet « Evaluation des programmes expérimentaux régionaux dans le cadre des pactes régionaux d'investissement dans les compétences », lancé par le Ministère du Travail, de l'Emploi, de la Formation Professionnelle et du Dialogue Social - DARES. Ce projet visait à faire une évaluation quantitative du programme « Mise en place d'une plateforme d'identification des besoins en compétences des entreprises » de la région Centre Val-de-Loire. Il impliquait l'utilisation des méthodes de big data et aurait demandé de travailler en collaboration avec des data scientists (non économistes) pour la mise en place de l'outil plateforme de la région et pour l'évaluation quantitative.

Calendrier et issue de la réponse à l'appel à projet :

L'élaboration et l'écriture du projet s'est faite de mars (date de lancement de l'appel à projet) à juin 2019 (date de dépôt des propositions).

Nous n'avons pas été retenus par le comité scientifique en charge de l'appel à projet et n'avons pas donc procéder à la réalisation de ce projet.

Dépenses faites à l'aide de la subvention :

Nous avons échangé par visioconférence entre les différents membres de l'équipes et n'avons pas mobilisé les fonds de ce support dans la préparation de la réponse à cet appel à projet.

Participants :

→ Participants du groupe de travail impliqués dans le projet : Sylvie Blasco, Ahmed Tritah

→ Autres participants impliqués dans le projet : Patrick Arni (Université de Bristol), Christelle Garrouste (Université de Créteil), Solenne Tanguy (Université du Mans), Jérémy Tanguy (Université Savoie Mont Blanc) + il était prévu le recrutement de deux contractuels data scientists.

Détails du projet et lien avec la thématique du groupe de travail :

Les difficultés de recrutement que rencontrent les entreprises en France sont en grande partie due à une inadéquation des compétences offertes et demandées et à une pénurie de candidats aux profils souhaités (BMMO, 2019). La formation professionnelle apparaît comme l'outil adéquat pour remédier à ces problèmes, mais semble inopérante pour remédier à ces difficultés de recrutement. Cette efficacité limitée de la politique de formation peut venir d'une mauvaise appréciation des compétences demandées par les entreprises. La Région Centre Val-de-Loire propose une intervention qui s'attaque à cette question. Le dispositif comprend une charte pour coordonner les actions des acteurs de l'emploi, des clubs de compétences pour identifier et collecter les besoins des entreprises, et une plateforme pour harmoniser et diffuser ces informations. L'objectif du programme est de mieux repérer et traduire les besoins des entreprises pour avoir des formations qui répondent finement aux besoins réels locaux.

Dans ce projet nous proposons d'évaluer l'effet des clubs de compétences et de la plateforme sur les pratiques de recrutement et de gestion des compétences des entreprises, sur le marché de formation, sur les demandeurs d'emploi et enfin sur les appariements sur le marché du travail. Notre approche aurait mêlé approche expérimentale et approche quasi-expérimentale. Nous proposons un protocole d'évaluation qui s'attachait à l'identification d'effets causaux, à la détection d'hétérogénéités d'effets, notamment en fonction de la taille des entreprises et du degré de tension sur le marché du travail, et à la compréhension des mécanismes sous-jacents. Nous déclinions notre démarche en 3 axes. Le premier se concentrait sur l'évaluation de l'effet des clubs sur les vacances d'emploi, recrutements et appariements. Notre stratégie privilégiée reposait sur une expérience aléatoire contrôlée dans laquelle un groupe d'entreprise était aléatoirement sollicité pour participer aux clubs, tandis qu'un autre groupe, aussi aléatoirement constitué, ne se voyait pas proposer les clubs. Le deuxième axe s'intéressait à l'effet de la plateforme sur l'offre et la demande de formation. Nous voulions employer ici la méthode de différence de différence, en comparant les évolutions des variables de résultats dans les territoires-métiers concernés par la plateforme avec celles observées dans des territoires-métiers comparables mais non concernés par la plateforme. Dans le troisième axe nous aurions adopté une approche plus globale et revenions sur les sources des dysfonctionnements du marché du travail qui ont motivé le lancement de ce programme régional. Nous nous serions appuyés sur la temporalité de la mise en place des clubs et de la plateforme pour tenter de distinguer les effets propres des différentes composantes du programme et identifier quelles sont l'importance et la nature de l'inadéquation entre offre et demande sur le marché du travail. Le design du programme permettait en effet de distinguer deux causes potentiellement à l'origine des dysfonctionnements sur le marché du travail. La première est la mauvaise identification des besoins de compétences par les entreprises et les organismes de formation, due à des problèmes informationnels. Le second, plus classique, porte sur l'adéquation entre les compétences offertes par les travailleurs et chômeurs et celles demandées par les entreprises. Il est crucial de distinguer ces deux sources de dysfonctionnement car chacune requière une intervention et des ressources publiques différentes. Si le problème est lié aux frictions informationnelles, il convient de lever ces barrières par un accompagnement des acteurs dans la définition des besoins. En revanche, un déficit ou une inadéquation réels dans les compétences offertes nécessite des investissements en formation et une refonte des contenus pédagogiques des modules de formation.

Pour évaluer les besoins de compétences et évaluer comment le dispositif de la région aurait pu influencer sur la définition et satisfaction de ces besoins, nous envisageons d'exploiter les données disponibles sur internet (web scrapping) et de collecter et traiter en masse les cv et offres d'emploi postées sur internet. Ceci nous aurait permis d'étudier la correspondance (et l'évolution de cette correspondance) entre des compétences offertes telles que décrites par

les travailleurs et candidats sur leur CV et les compétences demandées telles que décrites par les entreprises dans leurs offres d'emploi. Par le biais d'analyses sémantiques du contenu des offres et des CV nous aurions pu identifier des termes et l'occurrence des termes qui désignent des compétences. Cette approche de web scrapping est mobilisée par la Région pour alimenter la plateforme sur la nature des compétences demandées. Nous nous serions positionné en complément de ce travail de la Région. Nous aurions donc été ici amenés à mener un travail pluridisciplinaire en collaboration avec des data scientist.

Le budget prévu du projet était de 182 251 euros pour 2 années.

2- Réponse à un appel à projet sur les effets du chômage et du travail sur la santé mentale

En marge du projet du groupe de travail, un membre de l'équipe (Sylvie Blasco) a répondu à un appel à projet intitulé « Santé mentale, expériences du travail, du chômage et de la précarité » lancé par la Drees et la Dares ». Ce projet n'était pas initialement en lien avec la thématique du groupe de travail, mais il s'est avéré que pour la bonne réalisation du projet des techniques de machine Learning vont devoir être employées, m'amenant à présenter ce projet dans ce bilan scientifique.

Calendrier et issue de la réponse à l'appel à projet :

L'élaboration et l'écriture du projet s'est faite d'avril (date de lancement de l'appel à projet) à juillet 2019 (date de dépôt des propositions).

Le projet a été retenu par le comité scientifique en charge de l'appel à projet (10 projets sur 39 soumis ont été retenus) et un contrat avec la Dares a été signé pour 2 ans. Le projet est en cours de réalisation.

Dépenses faites à l'aide de la subvention :

Nous n'avons pas mobilisé les fonds de ce support dans la préparation de la réponse à cet appel à projet car cela ne nous semblait pas en lien direct avec les objectifs de la subvention.

Participants :

→ Participants du groupe de travail impliqués dans le projet : Sylvie Blasco

→ Autres participants impliqués dans le projet : Julie Rochut (Cnav, URV) et Bénédicte Rouland (Université de Nantes)

Détails du projet et lien avec la thématique du groupe de travail :

Les innovations organisationnelles et managériales récentes peuvent se traduire par une intensification du travail et par plus d'autonomie dans le travail. Nous mobiliserons les volets employés et employeurs des enquêtes Conditions de Travail – Risques PsychoSociaux pour étudier empiriquement les effets théoriquement ambigus de ces deux évolutions sur la santé mentale des travailleurs. Nous offrirons une vue d'ensemble des liens entre intensification, autonomie et santé et relierons, grâce aux méthodes de décomposition, les inégalités d'exposition à certaines pratiques RH aux inégalités de santé. Nous regarderons ensuite les effets de deux sources possibles d'intensification du travail : le recours à la flexibilité interne ou externe pour faire face aux fluctuations de la demande et l'automatisation. Nous considérerons aussi les pratiques de management à forte implication en nous concentrant sur l'autonomie liée au lieu et temps de travail. Pour ces analyses, nous identifierons des relations causales en nous appuyant sur la structure longitudinale des données et en recourant aux méthodes d'appariement et de variables instrumentales.

Pour définir les objets d'étude (l'automatisation, l'intensité du travail et l'autonomie dans le travail) à partir des données d'enquêtes mobilisées, nous mettrons en œuvre les méthodes de machine learning.

Nous répliquerons notamment les études qui mobilisent ces techniques pour définir le risque d'automatisation.

Le montant du financement demandé et obtenu est de 36 813 euros pour 2 années.

Une première version du travail a été présentée au séminaire interne du Gains auprès des membres du groupe de travail et autres membres du laboratoire.

3 – Perspectives

Perspectives à l'issue du financement au vu des indicateurs de réussite initialement exposés dans le dossier de réponse à l'AAP

Indiquez spécifiquement les prochains AAP auxquels vous envisagez de répondre suite à cette action.

-