

GenderedNews – Offre de thèse (FR)

2020-05-15

L'Université Grenoble Alpes recrute un·e doctorant·e (3 ans) entièrement financé·e à partir d'octobre 2020 pour contribuer au projet GenderedNews.

Le projet GenderedNews

Le projet GenderedNews vise à proposer de nouvelles méthodes pour mesurer et expliquer le niveau de biais de genre dans les médias en France. Ces biais peuvent être définis comme le fait que les médias d'information tendent d'une part à surpondérer les hommes par rapport aux femmes en termes de mentions et de citations, et d'autre part à attribuer aux femmes un rôle social spécifique impliquant souvent, entre autres, l'anonymat, une capacité d'action réduite dans la société et la confusion entre cette action et leur état matrimonial ou familial. De nombreuses études empiriques ont prouvé l'existence de ces biais et ont permis de mieux les comprendre à l'échelle internationale.

Cependant, la recherche sur cette question est souvent basée sur des données limitées en volume et produites par des ONG, des administrations ou des organismes de réglementation des médias. Ces données sont généralement traitées par des analyses de contenu manuelles qui ne permettent pas de rendre compte systématiquement des évolutions des biais sexistes dans les médias à long terme et sur un grand nombre de sources, ni d'expliquer ces biais en termes de variables telles que le financement des médias, la taille des salles de rédaction ou d'autres variables organisationnelles.

Le projet GenderedNews vise à fournir et à analyser des sources de données importantes et stables dans le temps ainsi qu'à explorer de nouvelles méthodes pour documenter les préjugés sexistes dans les médias. Il est basé sur un programme de recherche collaboratif entre un sociologue des médias et un informaticien ayant des compétences en études des médias, études de genre, traitement du langage naturel et collecte de données numériques. Il a également une dimension de partenariat importante dans la mesure où plusieurs médias importants sont associés et fournissent un accès à leurs données.

Le doctorat : les biais de citation entre hommes et femmes dans l'information

GenderedNews se concentre sur deux types de biais et deux mesures différentes de ces biais. Les biais d'échantillonnage se produisent par la sélection d'un échantillon biaisé de personnes mentionnées dans les médias. Ils peuvent être étudiés en comptant simplement combien d'hommes et de femmes ont accès à la visibilité publique d'un côté et en étudiant les modèles de cadrage des hommes et des femmes représentés de l'autre côté. Les biais de citation proviennent de la sélection d'un échantillon biaisé de personnes qui, en plus d'être visibles, sont autorisées à exprimer leurs opinions dans les médias. Ils peuvent également être étudiés en utilisant les deux approches: compter combien d'un côté et analyser comment de l'autre.

Dans le cadre du projet, le/la doctorant·e contribuera plus spécifiquement à l'étude des biais de citation. Cela impliquera les tâches et opérations de recherche suivantes:

- Analyse des données : les grands ensembles de données déjà constitués seront utilisés pour mesurer les biais d'approvisionnement dans l'actualité. Le/la doctorant·e utilisera notamment l'exploration de texte et les méthodes de NLP pour a) identifier les entités nommées dans le corpus, b) identifier les modèles de citation des sources dans les textes et c) étiqueter les entités nommées en fonction de leur sexe (sur la base des prénoms et des termes spécifiques au genre). Ces tâches permettront ensuite des analyses multivariées pour expliquer le niveau de diversité de genre atteint dans les différents médias analysés en fonction de variables sociologiques (par exemple, il sera possible de mesurer l'effet d'une augmentation de la proportion de femmes dans les rédactions sur le biais de genre ou l'effet d'une diminution du nombre de journalistes).

- Entretiens : le/la doctorant•e engagé•e poursuivra également une recherche qualitative sur l'origine des éventuels biais de genre dans les organisations partenaires. Une première série d'enquêtes permettra de vérifier que les méthodes utilisées pour accéder au contenu des médias ne biaisent pas les résultats. Des entretiens seront menés avec les équipes techniques en charge du fonctionnement des API auprès des médias partenaires afin de mieux comprendre leur fonctionnement et d'identifier d'éventuels biais. Une deuxième série d'enquêtes se concentrera sur la compréhension de la hiérarchie de l'information dans ces médias et son influence possible sur les préjugés sexistes. Des entretiens seront menés avec des journalistes et des rédacteurs en chef pour comprendre ce qui rend un article d'actualité digne d'intérêt et comment les médias prennent en compte les questions de genre. Le but de ces interviews sera de comprendre comment les médias en tant qu'organisation contrôlent les préjugés sexistes dans leurs opérations quotidiennes.
- Sur le plan théorique, cette thèse répondra à certaines questions qui étaient auparavant laissées dans l'ombre. Il s'agit notamment de la définition et de la mesure opérationnelle de la « diversité » dans l'actualité en relation avec les questions de genre pour la communauté des sciences sociales ainsi que la question du biais et de l'équité dans les algorithmes d'apprentissage automatique pour la communauté NLP.

En plus de la thèse et des publications scientifiques, le/la doctorant•e sera également impliqué•e dans le projet GenderedNews à différents niveaux:

- Superviser la collecte et le traitement des données avec les médias partenaires;
- Contribuer à la production de nouveaux outils de mesure et de visualisation de la diversité des contenus et des sources dans les médias;
- Contribuer à la diffusion des résultats auprès du grand public sur un site Internet servant de plateforme pour promouvoir une approche plus diversifiée de l'actualité.

Environnement scientifique

La thèse sera menée en tant que projet de recherche conjoint de l'équipe Régulations du laboratoire de recherche Pacte (<https://www.pacte-grenoble.fr/page/regulations>) et de l'équipe Getalp du laboratoire LIG (<http://www.liglab.fr/en/research/research-areas-and-teams/getalp>). Le/la doctorant•e recruté•e sera principalement hébergé•e au laboratoire de recherche Pacte pour un soutien quotidien et en partie au LIG.

Le/la doctorant•e bénéficiera également du soutien du projet du Data Institute sur les médias sociaux et les sciences sociales (<https://data-institute.univ-grenoble-alpes.fr/research/data-science-social-media-and-social-sciences/>) et de la Chaire Algorithmic Society du MIAI qui encourage la recherche basée sur la collaboration des sciences sociales et de l'informatique pour contribuer à une meilleure compréhension du fonctionnement des algorithmes et à une évaluation critique des effets de l'IA dans la société (<https://algorithmicsociety.github.io/>).

Comment postuler ?

Les candidats doivent être titulaires d'un Master en sciences sociales ou en traitement du langage naturel (ou être sur le point d'en obtenir un). Ils devraient idéalement avoir une solide expertise dans les méthodes des sciences sociales et une bonne connaissance des méthodes d'exploration de texte et de traitement du langage. Ils doivent également avoir une très bonne connaissance des langues française et anglaise pour pouvoir traiter les données textuelles dans ces deux langues et mener des entretiens principalement en français.

Les candidatures sont attendues avant le 1er juillet 2020 et doivent être adressées à Gilles Bastin à gilles.bastin@iepg.fr et François Portet à Francois.Portet@imag.fr.

Les candidat•es doivent joindre :

- une lettre de candidature expliquant pourquoi ils/elles se considèrent capables de poursuivre ce projet de thèse,
- leur dernier diplôme,
- un CV.

Ils/elles peuvent également ajouter des lettres de recommandation.

Le comité de sélection informera les candidat•es de sa décision avant le 15 juillet 2020.

Si vous avez d'autres questions sur le poste et le projet, veuillez contacter Gilles Bastin et / ou François Portet.