Beyond the Salary: Unpacking Job Attributes in French Ads with NLP

Alfonso AWADALLA - Sofia VACA

April 11, 2025

Overview

Research Question

- Goal: analyze the job attributes advertised by employers in France by exploiting the language used in job vacancies.
- What are the paid and non paid job attributes advertised by employers in job vacancies across France?
- Can variation in advertised attributes be explained by sector and location?

Motivation

- Growing interest in non wage amenities (flexibility, health, work-life balance).
- Employers use job ads as signals in imperfect labor markets.
- Need for systematic analysis using NLP.

Literature review

Non wage amenities:

- **Maestas et al. (2023)**: Workers often value flexibility, benefits, or work environment more than wage increases.
- Use of stated-preference experiments to estimate willingness to pay for amenities.

Non wage amenities in job ads with NLP:

- **Audoly et al. (2024)**: NLP applied to Norwegian job ads; nearly all mention non-wage amenities.
- Industry/occupation explain 30–40% of variation; firm-level effects increase this to 50–70%.
- **ILO (2025)**: Rule-based NLP in emerging economies. Taxonomy includes flexibility, work-life balance, and career development.

Literature review II

Specific non wage amenities in job ads with NLP:
Flexible Work:

- **Hansen et al. (2023)**: 99% accuracy identifying remote work in 250M job ads. Remote work tripled since 2019.

Training:

- **Adams-Prassl et al. (2022)**: 20% of U.S. ads mention on-the-job training, more common in concentrated markets.

Skills:

- **Deming
  Kahn (2017)**: Cognitive and social skill demands vary across firms; detected via job ad text.

Theoretical Motivation: Imperfect information and Matching

- **Marinescu and Wolthoff (2020)**:
- Job titles, more than wages, predict the number and quality of applicants.
- Job ad content functions as a signal under search frictions and imperfect information.
- **Sockin and Sojourner (2023)**: Jobseekers use external reviews (e.g., Glassdoor) to assess firms.
- Jobseekers actively seek employer insights.
- Limited transparency leads to asymmetric information, highlighting the role of job ads as structured signals.

Data JOCAS

- **Description:** The JOCAS database collects online job offers daily via web scraping. It provides full job descriptions, posting dates, durations, salaries, and ROME codes. This rich dataset enhances labour market analysis, though some sectors are overrepresented.
- **Variables of interest:** descriptions, ROME codes, locations
- **Producer:** DARES (Ministry of Labour and Employment)
- **Year:** 2020 (metropolitan France focus)
- **Limitations:** sampling bias, missing data

Job Classification

Mapped ROME codes to:

- FAP22 (22 occupational families)
- FAP87 (87 detailed categories)

Methodology:

We built on ILO (2025) and Audoly et al. (2024) to answer our main question: What are the paid and non paid job attributes advertised by employers in job vacancies across France? Methodology for using NLP techniques:

- Dictionary-based classification of amenities
- Topic modeling (LDA)
- Word embeddings
- Tailored to French job ad conventions

Metodology II: Logistic regression

To answer out second question can variation in advertised attributes be explained by sector and location?

- Estimate logistic regressions:

$$\text{logit}\left(\mathbb{P}(\text{AttributePresence}_i = 1)\right) = \alpha + \beta \cdot \text{Sector}_i + \gamma \cdot \text{Location}_i + \varepsilon_i \qquad (1)$$

- Outcome: Binary presence of attribute
- Explanatory variables: Sector, Location
- Goal: Descriptive mapping (not causal)

Sample of translated attributes

| Catégorie agrégée	Catégorie d'attribut	Exemples d'expressions
Attributs du lieu de travail	Bons collègues	collègues bienveillants, collègues sympathiques, ambiance agréable...
	Environnement social	bonne ambiance de travail, environnement de travail convivial, équipe soudée...
	Espace de bureau partagé	open space, bureaux en open-space, espace de travail partagé...
Régime de rémunération	Bonne rémunération des heures supplémentaires	heures supplémentaires payées, paiement des heures supp, majoration...
	Niveau de rémunération	salaire horaire, salaire mensuel, rémunération mensuelle...
	Prime à l'embauche	prime d'embauche, prime de bienvenue, prime à l'entrée...

Exploration of Job Ads
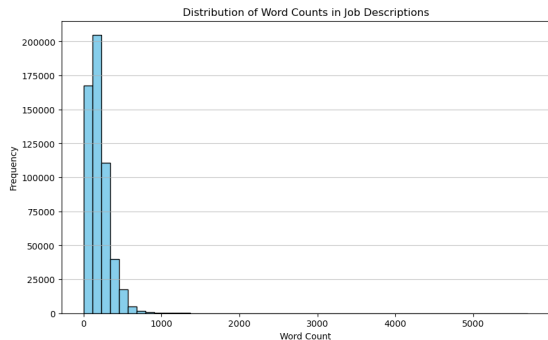
Figure 1 : Word count per ads
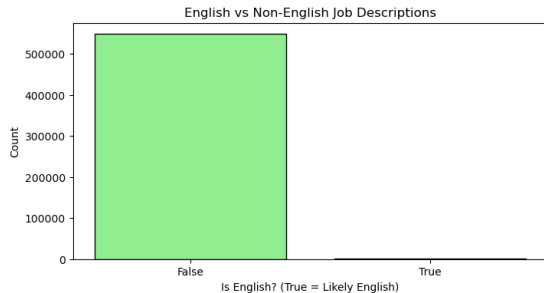

Figure 2 : English ads?

Le conseiller immobilier est un intermédiaire entre une personne vendant un bien immobilier et les potentiels acquéreur. Il a en charge différentes tâches, dont trouver des logements à vendre et visiter ces derniers afin de connaître leur agencement, trouver des acheteurs et leur faire visiter les biens. La gestion des dossiers de vente se fait à l'aide d'un notaire.

Profil: – Motivation – Fibre commerciale – Organisation – Persévérance – Détermination

Conditions de travail: – Sans contrainte horaire – Sans hiérarchie – Sans lourdeurs administratives

Missions: – Prospecter des biens immobiliers, des terrains. – Prospecter des nouveaux clients. – Formaliser une demande de transaction. – Définir des actions de promotions commerciales. – Valoriser un bien immobilier. – Présenter un bien immobilier à un client. – Présenter les modalités d'acquisition d'un bien. – Conseiller un client – Analyser le marché immobilier

Topic Representation (LDA)
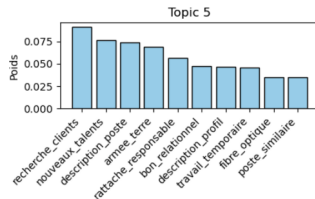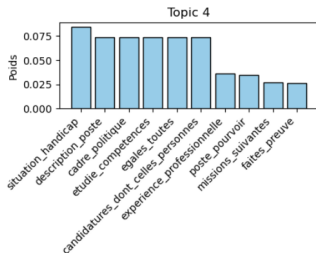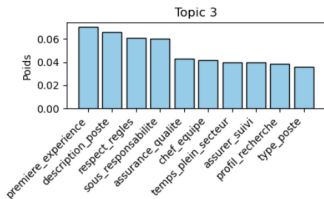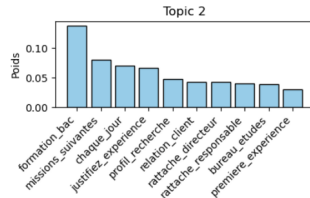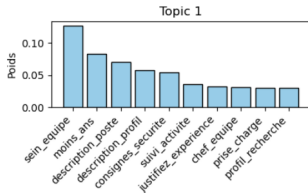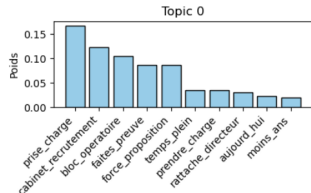
References I

- DARES. (2024). *La nomenclature des familles professionnelles 2021*.
  `https://dares.travail-emploi.gouv.fr/donnees/`
  `la-nomenclature-des-familles-professionnelles-2021`. Accessed March 2025.
- ILO. (2025). *Developing a New Method to Uncover Skills Trends in Emerging Economies Using Online Data and NLP Techniques*.
  ILO Methodological Brief No. 1. `https://doi.org/10.54394/HQQX3200`. Accessed April 2025.
- Audoly, R., Bhuller, M., & Reiremo, T. (2024). *The Pay and Non-Pay Content of Job Ads*.
  First version: July 2024. This version: September 2024.
  `https://arxiv.org/abs/2407.13204`. Accessed April 2025.
- ILO. (2025). *Measuring Quality of Employment in Emerging Economies: A Methodology for Assessing Job Amenities Using Online Data*.
  ILO Methodological Brief No. 2. `https://doi.org/10.2139/ssrn.3957002`. Accessed April 2025.

References II

- Deming, D., & Kahn, L. (2018). *Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals*.
  Journal of Labor Economics, 36(S1), S337–S397. `https://doi.org/10.1086/694106`. Accessed April 2025.

- Marinescu, I., & Wolthoff, R. (2020). *Opening the Black Box of the Matching Function: The Power of Words*.
  Journal of Labor Economics, 38(2), 535–568. `https://doi.org/10.1086/705880`. Accessed April 2025.

- Maestas, N., Mullen, K. J., Powell, D., von Wachter, T., & Wenger, J. B. (2023). *The Value of Working Conditions in the United States and the Implications for the Structure of Wages*.
  American Economic Review, 113(7), 2007–2047.
  `https://doi.org/10.1257/aer.20210415`. Accessed April 2025.

- Hansen, S., Lambert, P. J., Bloom, N., Davis, S. J., Sadun, R., & Taska, B. (2023). *Remote Work across Jobs, Companies, and Space*. NBER Working Paper No. 31007. `https://www.nber.org/papers/w31007`. Accessed April 2025.

- Adams-Prassl, A., Balgova, M., & Qian, M. (2020). *Flexible Work Arrangements in Low Wage Jobs: Evidence from Job Vacancy Data*. IZA Discussion Paper No. 13691. `https://www.iza.org/en/publications/dp/13691`. Accessed April 2025.

- Sockin, J., & Sojourner, A. (2023). *What's the Inside Scoop? Challenges in the Supply and Demand for Information on Employers*. *Journal of Labor Economics*, 41(4), 1041–1079. `https://doi.org/10.1086/722936`. Accessed April 2025.

- DARES. (2023). *Base JOCAS – Document méthodologique résumé*. `https://dares.travail-emploi.gouv.fr/publication/`. Accessed April 2025.