

---

# Innovation et Crédit d'Impôt Recherche : Analyse Textuelle des Narratifs d'Entreprises (2017-2022)

Rim Bahroun\*<sup>†1</sup>, Nadine Levratto\*<sup>‡1</sup>, and Mathieu Bernard<sup>1</sup>

<sup>1</sup>Economix UMR 235, UPL, Univ Paris Nanterre, CNRS – Economix UMR 235, UPL, Univ Paris Nanterre, CNRS – France

## Résumé

Le Crédit d'Impôt Recherche (CIR) en France est un dispositif fiscal conçu pour encourager les investissements en recherche et développement. Depuis sa refonte en 2008, il représente un coût considérable pour les finances publiques, mais son efficacité reste débattue. Dans cette recherche, nous proposons de dépasser les limites des indicateurs de performances à l'innovation (Dépenses de RD et brevets) en extrayant des signaux pertinents à partir des narratifs d'entreprises bénéficiaires du CIR. Ces derniers nous permettent d'analyser l'évolution des déclarations sur l'innovation en fonction du temps et du destinataire (grand public ou financier). Pour ce faire, nous appliquons des techniques de traitement automatique du texte sur deux corpus : les rapports financiers des entreprises et les articles de journaux spécialisés, couvrant les années 2017 à 2022. Nous introduisons une nouvelle méthode d'incorporation des informations textuelles qui combine le comptage des termes avec des techniques sémantiques basées sur des modèles de langage pré-entraînés, tels que BERT. Cette méthode permet de calculer, pour chaque entreprise du panel, un " indice d'innovation " en fonction de l'année et du type de discours. Les résultats de l'analyse textuelle seront comparés, dans un futur travail, à une typologie obtenue à partir d'une analyse quantitative basée sur des indicateurs traditionnels de l'innovation, tels que les dépenses en R&D ou le nombre de brevets, afin de produire une vision complète de la manière dont les entreprises utilisent le CIR et des effets de ce dispositif sur l'innovation au sein du tissu productif.

**Mots-Clés:** Crédit Impôt Recherche, indice d'innovation, analyse textuelle, technique sémantique, modèles de langage préentraînés

---

\*Intervenant

<sup>†</sup>Auteur correspondant: rim.bahroun@cnrs.fr

<sup>‡</sup>Auteur correspondant: nadine.levratto@cnrs.fr