

Diffusion de l'Intelligence Artificielle en France

Philippe Aghion¹ Antonin Bergeaud² Simon Bunel³ Paul Delbouve⁴

10 juillet 2025

Résumé

À partir de plus de 120 millions d'offres d'emploi publiées en ligne en France depuis 2019, cette étude examine les caractéristiques de la diffusion de l'intelligence artificielle (IA) dans les entreprises et ses effets sur l'emploi. La diffusion de l'IA se révèle très hétérogène : elle est particulièrement avancée dans certains secteurs (informatique, finance, conseil) et au sein de grandes entreprises pionnières. Cette hétérogénéité s'observe aussi à l'échelle territoriale, les offres d'emploi en IA étant fortement concentrées dans les zones d'emploi accueillant des clusters de recherche ou des pôles industriels spécialisés (semi-conducteurs, assurances, aéronautique, etc.). Nous montrons que les entreprises les plus exposées à l'IA ont été significativement plus enclines à recruter des spécialistes de l'IA dès l'apparition de ChatGPT. En parallèle, on observe une baisse relative de la création d'emplois dans les professions les plus exposées à l'IA. Toutefois, les recrutements dans les métiers de l'IA s'accompagnent d'une hausse de la productivité du travail et du chiffre d'affaires, concomitante avec une hausse de l'emploi total dans les entreprises concernées, tirée par les professions fortement exposées, au détriment de celles qui le sont moins.

¹Collège de France, INSEAD, LSE, and CEPR. E-mail : philippe.aghion@insead.edu

²HEC Paris, CEP-LSE, and CEPR. E-mail : bergeaud@hec.fr

³Banque de France and ENS. E-mail : simon.bunel@banque-france.fr

⁴Collège de France and PSE. E-mail : paul.delbouve@finances.gouv.fr

1 Introduction

En novembre 2022, OpenAI annonce la sortie de ChatGPT, un outil accessible permettant d'échanger avec une intelligence artificielle (IA) générative capable a priori de comprendre et d'exécuter des ordres très différents et peu structurés. La nouveauté de ce produit et sa prise en main très intuitive ont facilité une diffusion extrêmement rapide : 1 million d'utilisateurs hebdomadaires atteints en 5 jours, 100 millions en un an et 400 millions en 2025. En moins de 3 ans de nombreuses mises à jour et de nombreux concurrents sont apparus, repoussant les capacités et la puissance de ces IA générative de manière parfois spectaculaire. Cette expansion rapide a généré de nombreuses interrogations et anxiétés concernant l'impact de l'IA sur l'emploi ainsi que des tentatives d'évaluation (voir par exemple [Eloundou et al., 2024](#); [Felten et al., 2023](#); [Gmyrek et al., 2023](#); [Aghion and Bouverot, 2024](#); [Pizzinelli et al., 2023](#); [Bergeaud, 2024a](#); [Aghion et al., 2025](#)).

Pour autant, l'année 2022 ne constitue pas le véritable point de départ de la diffusion de l'intelligence artificielle dans l'économie. Dès le début des années 2010, les avancées en apprentissage automatique, nourries par l'explosion des volumes de données numériques et par la montée en puissance des capacités de calcul, ont permis le développement d'applications concrètes d'IA dans des secteurs variés tels que la logistique, la finance, la santé, le commerce, l'industrie ou encore les ressources humaines. Moins visibles que les interfaces conversationnelles grand public comme ChatGPT, ces technologies ont néanmoins commencé à transformer en profondeur les modes de production, de décision et d'organisation du travail, avec des effets notables sur la productivité. Cette amélioration peut passer par plusieurs canaux. L'IA permet tout d'abord une prise de décision plus performante, fondée sur l'exploitation de données massives et actualisées ([Brynjolfsson and McElheran, 2016](#)). Elle facilite aussi l'automatisation de certaines tâches, y compris cognitives, renforçant ainsi la productivité horaire ([Acemoglu and Restrepo, 2018](#)). Elle peut en outre accélérer les processus d'innovation en appuyant les activités de recherche et développement ([Cockburn et al., 2018](#); [Aghion et al., 2017](#); [Agrawal et al., 2018](#)). Dans de nombreux cas, l'IA complète les compétences humaines en augmentant les capacités d'analyse, de création ou d'exécution des travailleurs ([Furman and Seamans, 2019](#)). Elle contribue également à réduire les coûts d'expérimentation, en substituant aux approches par essai-erreur des simulations et des prédictions plus fiables ([Agrawal et al., 2022](#)). Ainsi, loin de se limiter à la seule substitution de tâches, les effets de l'IA sur la productivité engagent donc des transformations plus diffuses et systémiques des modes de travail. Cet impact est particulièrement déterminant pour analyser les effets de l'IA sur l'emploi : comme le soulignent [Aghion et al. \(2019\)](#), l'adoption de ces technologies peut générer une création nette d'emplois si

les gains de productivité l'emportent sur les effets d'éviction. La littérature empirique récente sur l'impact de l'IA sur l'emploi et la productivité se concentre principalement sur des cas d'usage spécifiques, souvent étudiés dans des contextes expérimentaux ou semi-expérimentaux. Par exemple, [Brynjolfsson et al. \(2023\)](#) analyse les effets d'un assistant conversationnel dans des centres d'appels, [Peng et al. \(2023\)](#) mesure les gains de performance de programmeurs utilisant ChatGPT, tandis que [Noy and Zhang \(2023\)](#) s'intéresse à l'impact sur les consultants en stratégie. Ces études mettent en avant une augmentation de la productivité individuelle dans les professions concernées après quelques mois, dont la magnitude varie de 25% dans le cas des téléconseillers à 55% dans le cas des programmeurs informatiques. Toutefois, en raison de leur portée limitée à des contextes particuliers, elles peinent à éclairer les effets macroéconomiques potentiels de l'IA générative, lesquels dépendront en grande partie de la part des tâches concernées par de tels gains de productivité ([Acemoglu, 2025](#)).

C'est en ce sens que, pour anticiper les effets macroéconomiques de l'IA, une étape essentielle consiste à mesurer sa diffusion effective au sein des entreprises, à en identifier les éventuels freins, et à évaluer son impact sur l'emploi et la productivité. Alors que d'autres études sur l'impact économique de l'IA comme par exemple [Aghion et al. \(2025\)](#) reposent directement sur des enquêtes permettant de distinguer entre entreprises adoptant l'IA dans un intervalle de temps donné et un des entreprises semblables qui n'ont pas encore adopté l'IA jusque-là, nous considérons une mesure d'adoption de l'IA fondée sur l'analyse automatisée de plusieurs millions d'offres d'emplois à partir de la base de données JOCAS, développée par la Dares. Cela nous permet de développer une méthodologie originale permettant de classifier les offres d'emplois selon leur lien avec l'IA, qu'il soit direct (i.e. "interne") — par exemple, une annonce visant à recruter un ingénieur en vision par ordinateur — ou indirect (i.e. "externe") — telle une offre pour un analyste marketing amené à utiliser régulièrement des outils de traitement automatique du langage naturel. Notre classification permet également de distinguer entre les offres faisant référence à de l'IA dite "générative" et les offres se référant à des outils d'analyses de données plus classiques (Natural Language Processing, vision par ordinateur, Big Data, etc...).

La richesse granulaire de la base de données JOCAS et son exhaustivité, combinée à sa haute fréquence de mise à jour, nous permet de suivre en quasi temps réel la diffusion des technologies d'IA, métier par métier, secteur par secteur, et territoire par territoire. À notre connaissance, il s'agit du premier travail exploitant ce type de données à une telle échelle pour construire un tableau de bord interactif de suivi de la diffusion de l'IA via les offres d'emploi, mis à jour en continu.¹

¹Ce tableau de bord est accessible à l'adresse suivante : [XX](#).

Si ces données d'offres d'emploi liées à l'IA ne capturent qu'une partie du phénomène — en se limitant par nature aux flux d'embauches d'employés IA internes, et en laissant de côté les formes d'externalisation ou les usages intégrés sans embauche dédiée —, leur analyse permet néanmoins d'identifier un certain nombre de tendances dans les dynamiques d'adoption. Ainsi, nous n'observons pas de hausse tendancielle de la part des offres d'emploi en IA pouvant traduire une démocratisation de cette technologie et le recours à des logiciels clés en main ne nécessitant pas de recrutement pour les déployer. Cette tendance générale masque cependant une croissance très forte mais encore récente des offres liées à l'IA générative suivant le lancement de ChatGPT, signe d'un basculement technologique important. Notre approche permet également de décomposer les offres en IA par type d'emplois. Près de la moitié concernent des profils très spécialisés dédiés au développement de modèles d'intelligence artificielle, tandis que l'autre s'apparente à des métiers liés à l'informatique générale, au management ou aux fonctions support non techniques nécessaires pour faciliter le déploiement de cette technologie. S'agissant de la diffusion de l'IA au sein des entreprises, on distingue une croissance relativement stable du nombre de nouveaux utilisateurs de l'IA, mais là aussi une croissance exponentielle du nombre d'entreprises adoptant l'IA générative. Cette diffusion de l'IA est cependant marquée par une forte hétérogénéité, avec certains secteurs (comme l'informatique, la finance ou le conseil) et de grandes entreprises pionnières en avance dans l'usage de ces technologies. Enfin, cette hétérogénéité se retrouve également à l'échelle territoriale : les offres d'emploi en IA sont fortement concentrées dans des zones d'emploi abritant des clusters de recherche ou des pôles industriels spécialisés (semi-conducteurs, assurances, aéronautique, etc.), où les conditions d'adoption de l'IA sont les plus favorables.

Toujours à partir des données issues de la base de données JOCAS, nous analysons dans un second temps l'effet de l'irruption de l'IA générative sur les dynamiques de recrutement. Un premier résultat consiste à observer que les entreprises plus fortement exposées à l'intelligence artificielle — c'est-à-dire celles dont une part importante des recrutements avant 2022 concerne des métiers susceptibles d'être transformés par l'IA — sont significativement plus enclines à embaucher des spécialistes de l'IA à partir de novembre 2022, date à laquelle la sortie de ChatGPT marque le point de départ d'une diffusion massive de l'IA générative. Un second résultat est qu'après novembre 2022, les créations de nouveaux emplois dans les professions plus exposées à l'IA baissent relativement aux professions moins exposées à l'IA.

Ces résultats peuvent sembler a priori en contradiction avec les effets positifs mis en évidence notamment par [Aghion et al. \(2025\)](#).² Une première remarque essentielle tient

²Nous détaillons comment ils peuvent être réconciliés dans la section 5.

à la nature des variables étudiées. Les résultats précédemment mentionnés s'intéressent aux offres d'emploi (flux) publiées par les entreprises et renseigne donc sur la demande de travail, tandis que [Aghion et al. \(2025\)](#) s'appuient sur des données administratives permettant de mesurer les niveaux d'emploi (stock) effectifs dans les entreprises. En d'autres termes, les deux approches saisissent des dynamiques complémentaires : nous capturons des ajustements contemporains de demande liés à l'anticipation de transformations technologiques, tandis que l'approche par données administratives mesure les effets réalisés *ex post* sur l'emploi, notamment en réponse à un effet productivité. Une deuxième remarque tient à la définition même de l'adoption de l'IA considérée, également différente dans ce qu'elle capture. Notre mesure se concentre en grande partie sur les recrutements de profils spécialisés IA, qui signalent des démarches de développement interne de solutions d'IA. Les enquêtes mobilisées par [Aghion et al. \(2025\)](#) ou par [Calvino and Fontanelli \(2023\)](#) incluent à la fois les entreprises développant en interne et celles utilisant des solutions d'IA achetées sur le marché. Une dernière remarque provient du fait que notre analyse couvre une période marquée par l'émergence rapide des IA génératives, à partir de fin 2022. Ces technologies, par leur capacité à automatiser des tâches cognitives complexes, pourraient avoir des effets spécifiques sur la structure des emplois, encore mal captés par les enquêtes administratives disponibles, et qui pourraient devenir positif après quelques trimestres supplémentaires.

Nous apportons plusieurs éléments empiriques à l'appui de cette distinction entre adoption interne et externe, et entre IA générative et IA au sens large. Nous montrons d'abord que le recrutement de profils IA est associé à une augmentation du niveau d'emploi total dans les entreprises, conformément aux résultats de [Aghion et al. \(2025\)](#) avec une hétérogénéité importante selon l'exposition des métiers à l'IA : l'emploi augmente dans les professions fortement exposées, mais tend à diminuer dans les professions peu exposées. Dans un second exercice, nous exploitons les données d'enquête sur l'adoption déclarée de technologies d'IA pour distinguer les entreprises ayant développé leurs solutions en interne de celles ayant adopté des solutions externes. Cette distinction révèle des effets différenciés sur l'emploi : l'adoption par développement interne est associée à une hausse substantielle de l'emploi, notamment dans les professions exposées, tandis que l'adoption via des solutions externes a un effet plus modeste et statistiquement non significatif. Ce contraste s'interprète naturellement dans une logique de fonction de production : le développement au sein de l'entreprise d'une IA transforme profondément les processus internes, en modifiant les besoins en compétences et les chaînes de valeur. C'est celui que l'on capture principalement avec notre approche. A l'inverse, l'adoption externe accroît la productivité à la marge, sans reconfiguration organisationnelle majeure. Notons que la similitude des résultats obtenus en considérant le recrutement de profils IA via les données JOCAS d'un côté et le développement interne de solutions d'IA via

les données d'enquête de l'autre montre la complémentarité de ces deux approches.

Notre étude s'inscrit dans le prolongement d'une littérature empirique récente cherchant à documenter la diffusion de l'intelligence artificielle et ses effets sur l'emploi et la productivité. Les contributions existantes s'appuient principalement sur trois types de sources de données. Un premier ensemble de travaux mobilise les bases de brevets, qui permettent de suivre l'innovation dans les technologies liées à l'IA et d'identifier les tendances sectorielles et géographiques de sa diffusion (Webb, 2019; Alderucci et al., 2020; Damioli et al., 2021; Calvino et al., 2022, 2023; Bergeaud, 2024b). Toutefois, ces données présentent plusieurs limites. D'une part, le décalage temporel entre l'invention, le dépôt, et la publication d'un brevet rend difficile la détection en temps réel des évolutions récentes, en particulier dans le cas de l'IA générative, dont la dynamique est extrêmement rapide. D'autre part, il n'est pas certain que les brevets constituent un instrument de protection systématique pour les algorithmes ou modèles d'IA, notamment dans des environnements très concurrentiels où d'autres stratégies (secret industriel, publication ouverte, open source) peuvent être privilégiées. Enfin, l'utilisation des brevets comme mesure de la diffusion d'une technologie pose le problème habituel que l'acte d'invention ne coïncide pas nécessairement avec son adoption effective : une technologie brevetée peut être licenciée ou vendue, sans que l'on sache précisément si, où, et comment elle est réellement utilisée (Arora et al., 2004).

Une seconde source mobilisée dans la littérature repose sur les enquêtes menées auprès des entreprises, qui permettent d'observer directement les usages déclarés de l'intelligence artificielle, ainsi que ses effets perçus sur l'emploi, la productivité ou l'innovation. Calvino and Fontanelli (2023) et Czarnitzki et al. (2023) s'appuient sur ce type de données pour quantifier la diffusion de l'IA dans les organisations et en évaluer les implications économiques. Aghion et al. (2025) les mobilisent quant à eux pour analyser l'effet différencié de l'IA selon les cas d'usage, qu'il s'agisse de leur intégration aux processus de production, à l'organisation des processus administratifs, ou encore pour la sécurité informatique des entreprises. Ces enquêtes, souvent réalisées par des organismes statistiques ou des institutions internationales, tendent à s'enrichir au fil du temps, tant en termes de périmètre que de finesse des questions posées. Elles constituent une source précieuse pour l'analyse fine des pratiques d'adoption. Toutefois, elles restent encore principalement centrées sur les grandes entreprises, et peinent à capturer la diversité des usages dans les PME ou les très petites structures, limitant ainsi leur capacité à refléter pleinement la diffusion de l'IA dans l'ensemble du tissu productif.

Enfin, une troisième approche — à laquelle s'inscrit le présent article — consiste à exploiter les données issues des offres d'emploi mentionnant des compétences ou

des technologies liées à l'intelligence artificielle. Cette méthode, utilisée notamment par [Calvino et al. \(2024\)](#); [Babina et al. \(2024\)](#); [McElheran et al. \(2024\)](#); [Hampole et al. \(2025\)](#); [Acemoglu et al. \(2022\)](#), ne capture qu'une partie des usages effectifs de l'IA : elle ne permet pas, par exemple, de détecter l'utilisation ponctuelle d'un logiciel d'IA sans modification des profils recherchés, ni les cas d'externalisation des activités de R&D. Elle constitue donc une mesure partielle de l'adoption. Toutefois, cette approche présente l'avantage majeur de s'appuyer sur des données à haute fréquence, offrant une lecture quasi-contemporaine des évolutions du marché du travail en lien avec l'IA, là où les brevets ou les enquêtes souffrent d'un décalage temporel ou d'une couverture incomplète.

L'article est structuré de la façon suivante : la section 2 décrit les données mobilisées et la méthodologie utilisée pour identifier les offres d'emploi liées à l'IA, la section 3 fournit une image de la diffusion de l'IA en France sur la période 2019-2025, la section 4 analyse l'effet de la sortie de ChatGPT sur le *flux d'emplois*, la section 5 discute les effets de l'adoption de l'IA sur le *stock d'emplois* au niveau des entreprises. Enfin, la section 6 conclut.

2 Données

2.1 Base de données JOCAS

Notre analyse repose sur la base de données JOCAS, constituée par la Dares (Ministère du Travail) depuis 2019. Issue d'un processus de webscraping réalisé sur une quinzaine des principaux sites de recrutement en ligne, cette base regroupe à ce jour plus de 130 millions d'annonces. Chaque offre est enrichie d'un ensemble d'informations structurées ou semi-structurées : intitulé du poste, description, type de contrat, niveau de salaire, localisation géographique, etc. Grâce à son actualisation quotidienne, JOCAS fournit une photographie très récente du marché du travail en France et permet de détecter en temps réel l'émergence de nouvelles tendances, notamment celles liées à l'essor de l'IA générative.

L'exploitation de ces données présente toutefois plusieurs défis méthodologiques qui rendent nécessaire le recours à des techniques avancées de traitement automatique du langage naturel (NLP). Contrairement à des bases plus homogènes ou administratives, JOCAS agrège des annonces issues de sources multiples, rédigées de façon libre et hétérogène. Les intitulés de poste, les descriptions ou les compétences requises varient fortement en longueur, en vocabulaire et en précision, et peuvent omettre des éléments essentiels comme le niveau de qualification, la nature des technologies utilisées, ou les

compétences attendues. Cette variabilité rédactionnelle rend incontournable une étape rigoureuse de nettoyage, d'uniformisation et de classification des contenus textuels, préalable à toute analyse systématique de la présence de l'intelligence artificielle dans les offres.

Les analyses faites sur la diffusion de l'IA au sein des entreprises nécessitent un filtrage supplémentaire. En effet, une partie conséquente des offres sont postées par des entreprises d'intérim qui ne permettent pas d'identifier l'organisation ayant finalement recruté pour développer un outil d'IA. Les figures et tableaux présentés en section 4, n'incluent pas les entreprises correspondant aux "Activités liées à l'emploi" dans la nomenclature NAF ([code NAF 78](#)). Plus de détails sur la base de données sont donnés en Annexe A.1.

2.2 Identifier les offres d'emplois IA

Afin d'analyser finement l'adoption de l'intelligence artificielle dans les entreprises, nous cherchons à identifier, pour chaque offre d'emploi, à la fois la technologie IA mobilisée (IA générative, vision par ordinateur, traitement automatique du langage, etc.) et le niveau de technicité du poste proposé (profil ingénieur, fonctions support, management, etc.). Une simple recherche par mots-clés constitue une première approche naturelle, car elle présente l'avantage d'être transparente, reproductible, et peu coûteuse en calcul, comme le soulignent [Gentzkow et al. \(2019\)](#); [Bergeaud et al. \(2025\)](#). Elle permet souvent de détecter rapidement des tendances générales ou des signaux forts.

Cependant, cette méthode présente aussi des limites importantes. Elle peut aboutir à de nombreux faux positifs — par exemple lorsque des termes liés à l'IA apparaissent dans une description sans que le poste lui-même mobilise réellement ces technologies — ou à des faux négatifs, lorsque le lien à l'intelligence artificielle est implicite, mal orthographié, ou formulé de manière non standardisée. Comme l'indiquent [Gentzkow et al. \(2019\)](#), ces approches lexicales ont tendance à ignorer les subtilités du langage naturel et peinent à distinguer les contextes d'usage réel des simples références périphériques. Elles nécessitent donc d'être complétées par des méthodes plus avancées lorsque l'on souhaite classifier finement le contenu technique des annonces.

Dans ce contexte, nous avons recours à un modèle d'apprentissage supervisé, entraîné à partir d'un échantillon d'offres annotées manuellement par nos soins. Cette étape d'annotation experte, bien que coûteuse, constitue selon [Gentzkow et al. \(2019\)](#) un levier essentiel pour exploiter pleinement les capacités des modèles récents de traitement du langage naturel. Nous mobilisons ici un modèle de type *BERT*, dont l'architecture permet d'apprendre à partir du contexte linguistique complet d'un document. Ce modèle

dépasse ainsi les limites des méthodes lexicales en identifiant les structures de langage et les contextes d'usage typiques associés aux véritables usages de l'intelligence artificielle dans les offres d'emploi. Il en résulte une classification plus fine et plus robuste, y compris lorsque l'information est partielle, implicite ou formulée de manière non conventionnelle.

2.3 Classification

Notre procédure de classification repose sur deux étapes complémentaires visant à garantir l'exhaustivité de la détection et la précision de la catégorisation des offres liées à l'intelligence artificielle.

Nous appliquons d'abord un filtre lexical large, fondé sur une liste enrichie de mots-clés proposée par [Babina et al. \(2024\)](#) et complétée par des termes liés aux technologies émergentes de l'IA générative (*LLM, GAN, ChatGPT, transformers*, etc.). Cette première étape vise à maximiser le rappel, en identifiant toutes les offres potentiellement pertinentes pour l'analyse.

Le sous-ensemble ainsi identifié sert ensuite à entraîner un modèle BERT, selon une approche supervisée. Nous constituons à cet effet un jeu d'entraînement d'environ 1 000 offres annotées manuellement, représentant une diversité de technologies et de niveaux de technicité. Cette stratégie est cohérente avec les recommandations de [Bergeaud et al. \(2025\)](#), qui mettent en avant les performances des modèles de langage pré-entraînés comme *BERT* pour les tâches de classification de texte, dès lors qu'ils sont finement adaptés à un domaine spécifique avec un objectif bien défini.

Le modèle classe ainsi chaque offre selon deux dimensions principales :

Technologie IA dominante

1. **IA générative** : Génération automatisée de texte, d'images ou d'agents conversationnels.
2. **Vision par ordinateur** : Traitement d'images, reconnaissance d'objets, analyse vidéo.
3. **IA générale** : Autres usages de l'IA non générative, comme le NLP classique ou les systèmes prédictifs.
4. **Data** : Analyse de données “classique” sans recours à l'apprentissage automatique.

Type et technicité du poste

1. **Ingénieur** : Conception, entraînement et évaluation de modèles IA.
2. **Support** : Déploiement opérationnel, développement d'outils ou d'API IA.
3. **Management** : Encadrement d'équipes IA ou pilotage de projets techniques.
4. **Fonctions connexes** : Identification d'usages IA, gestion produit, acculturation interne.

Cette typologie s'appuie en partie sur la classification des compétences numériques élaborée par l'OCDE ([Squicciarini et al., 2021](#)). Les mots-clés associés à chaque classe, les exemples d'intitulés, la procédure d'annotation ainsi que les métriques de performance du modèle sont présentés en Annexe B.

3 Une photographie de la diffusion de l'IA en France

3.1 Evolution des offres IA dans le temps

Notre approche permet d'obtenir une vue d'ensemble des profils recherchés pour le développement et la mise en œuvre de l'intelligence artificielle dans les entreprises, avec les limites déjà évoquées. Nous identifions au total 176 392 offres comme faisant intervenir un des types de technologies présentées en Section 2.3 ce qui représente un peu plus d'une offre d'emploi sur 1000 (0.13%). Les Tableaux 1 et 2 présentent respectivement la répartition des offres d'emploi en fonction des types de technologies mobilisées et des catégories de poste proposées. L'IA générative y occupe encore une place modeste : elle ne représente qu'environ 2 offres sur 100 000, soit moins d'une offre IA sur 25. Toutefois, la Figure 1 montre une dynamique clairement haussière depuis le début de l'année 2023, suggérant une montée en puissance progressive de ces technologies dans les pratiques de recrutement.

Plus étonnamment, la Figure 1 révèle l'absence de tendance nettement croissante sur la période 2019–2025 en ce qui concerne les offres liées à l'IA. Après un pic entre 2020 et 2021, leur part relative se stabilise autour de 0,04% à 0,06% des annonces publiées. Ce constat peut s'expliquer par le fait que le recrutement de spécialistes de l'IA — notamment en IA générale — a déjà débuté avant 2019, et que la diffusion des compétences associées vers des fonctions connexes (support, management, etc.) progresse de manière lente et différée. Par ailleurs, l'intégration croissante de l'IA dans des outils standards peut ne pas se traduire immédiatement par des recrutements explicitement identifiés comme tels dans les offres. A ce titre, il est clair que notre approche sous-estime l'ampleur de la diffusion de l'IA dans les entreprises. Ces éléments suggèrent que la diffusion de l'IA dans l'économie française s'effectue plus par transformation progressive des

métiers existants que par une augmentation visible de la demande directe en profils spécialisés à court terme.

TABLE 1 – Répartition des offres d’emploi selon le type de technologie

	Vision par ordinateur	IA générale	IA générative	Total hors data	Total IA
Nombre d’offres	9 394	64 014	3 159	77 074	176 392
Part des offres (%)	(0,007%)	(0,05%)	(0,002%)	(0,06%)	(0,13%)

Notes : les données proviennent de JOCAS entre janvier 2019 et Janvier 2025. 130 millions d’offres d’emplois.

TABLE 2 – Répartition des offres d’emploi liées à l’IA selon le type de poste

	Ingénieur	Support	Management	Fonctions connexes
Nombre d’offres	54 090	70 486	30 513	21 303
Part des offres (%)	(0,041%)	(0,053%)	(0,023%)	(0,016%)

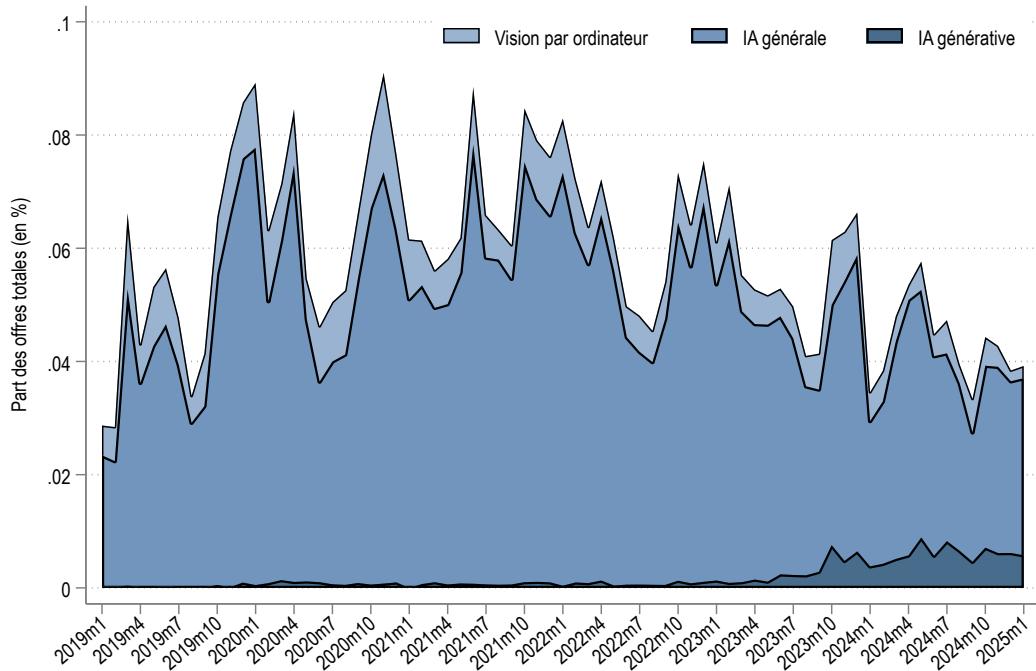
Notes : données extraites de JOCAS entre janvier 2019 et Janvier 2025.

3.2 Distribution par métiers

L’analyse des offres par machine-learning décrite en Section 2.3 permet d’identifier le type et niveau de technicité de l’offre selon les quatre classes détaillées en section 2. La figure 2 montre l’évolution de la répartition des offres selon ces catégories et indique qu’une moitié seulement des offres IA concernent des professions techniques chargées du développement de modèles et d’algorithmes d’intelligence artificielle. Ceux-ci sont secondés (pour environ 20% des offres) par des métiers plus proches de l’informatique classique et par des postes d’encadrement et de support non technique.

Il est toutefois techniquement possible d’analyser les offres à un niveau plus fin. La Dares a enrichi la base de données JOCAS avec le code ROME (Répertoire Opérationnel des Métiers et des Emplois, la classification utilisée par Pole Emploi) grâce à un algorithme prédictif à partir de la description de l’offre (Dares, 2023). Ceci permet notamment de regarder quels métiers sont les plus concernés par les offres que nous avons identifié comme IA. Le Tableau 3. Sans surprise le métier qui revient le plus fréquemment dans le cadre des offres IA est celui de développeur informatique avec près de 25 000

FIGURE 1 – Evolution des offres IA (en %)



offres (près de 1% des offres totales de développeur). On retrouve également différents métiers de la recherche et du développement ou lié à l'informatique. Ces métiers sont probablement liés à la mise en place de solution IA dans l'entreprise. Mais les métiers de chargé d'étude, responsable commercial ou responsable en organisation apparaissent aussi ce qui suggère une diffusion plus profonde dans les processus de production et de décision des entreprises.

L'algorithme d'allocation des métiers aux codes ROME est toutefois imparfait et l'analyse en détail de certaines offres fait apparaître des erreurs qui rendent difficile une analyse plus systématique.³ Dans l'attente de la mise en place d'un algorithme plus performant, nous proposons une approche différente consistant à analyser le titre des offres IA et à les regrouper en groupe de manière non supervisée (*k-mean clustering*). Les résultats sont présentés dans les Tableaux A4 et A5 et le détail de la procédure à l'Annexe A.4. Cet exercice peut être vu comme une extension des résultats présentés dans la Figure 2 mais de manière plus flexible.

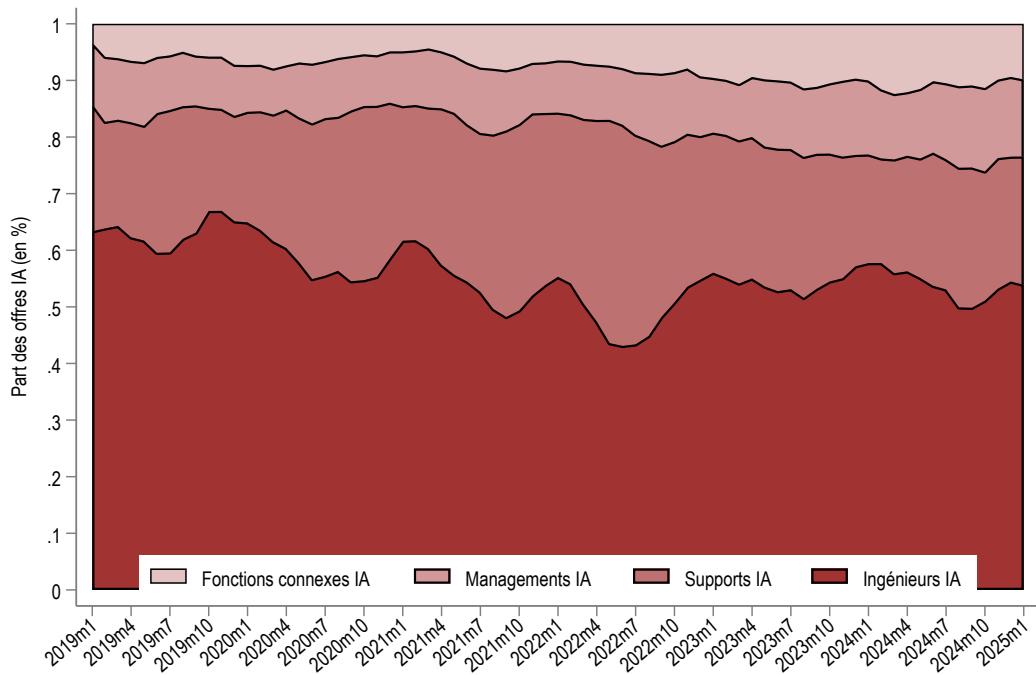
³Par exemple de nombreux métiers liés à la création de modèle d'IA sont identifiés au code ROME "mannequin" et certains métiers liés à la vision par ordinateur sont associés au code ROME "cadreur".

TABLE 3 – Top 10 métiers en nombre d'offres IA

Rang	Métier	Code ROME	Nombre d'offres IA
1.	Développeur informatique	M1805	24 931
2.	Chargé d'études socio-économiques	M1403	15 249
3.	Ingénieur R&D en industrie	H1206	6 104
4.	Expert systèmes et réseaux informatiques	M1802	2 233
5.	Ingénieure de recherche scientifique	K2402	2 194
6.	Directeur des systèmes d'information	M1803	1 494
7.	Consultant fonctionnel des systèmes d'information	M1806	1 421
8.	Responsable en organisation en entreprise	M1402	1 101
9.	Chercheur en sciences humaines et sociales	K2401	913
10.	Responsable commercial	M1707	814

Note : 10 code ROME classé par le nombre d'offres identifiés comme IA par la classification présentée en Section 2.3.

FIGURE 2 – Répartition des offres IA par type de métier



3.3 Distribution dans l'espace

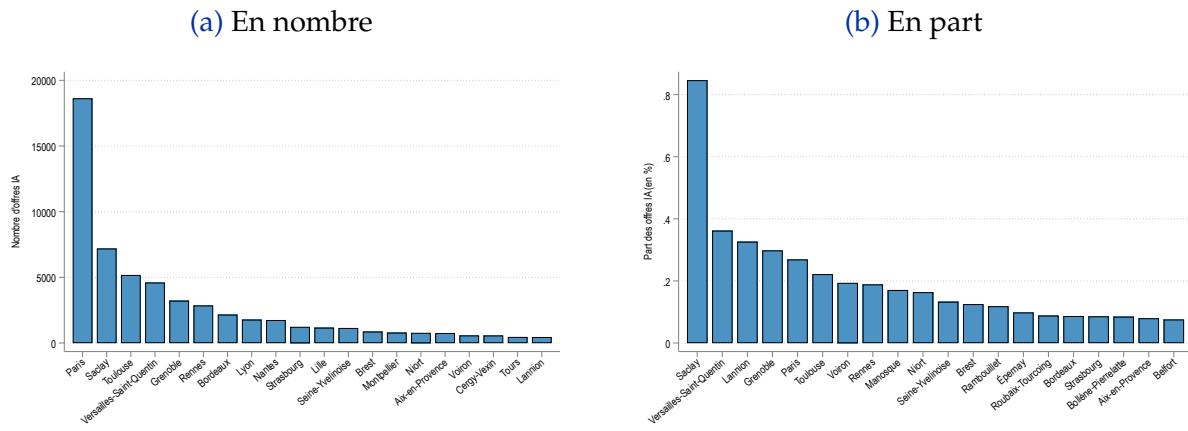
Une autre dimension importante de la base JOCAS est sa composante géographique. En effet, une large majorité des offres (82%, voir Annexe A.1) est associée à un code postal, ce qui permet une analyse fine de l'hétérogénéité territoriale du recrutement en intelligence artificielle. À partir de cette information, nous pouvons regrouper les offres au niveau des zones d'emploi.

Les figures 3a et 3b présentent ainsi les 20 zones d'emploi les plus dynamiques en matière de recrutement dans le domaine de l'IA, à la fois en volume absolu et en part relative des offres recensées. On y retrouve plusieurs pôles bien connus pour leur concentration d'acteurs de la recherche et de l'innovation, tels que Saclay, Paris ou Grenoble-Voiron, qui bénéficient d'un écosystème dense combinant entreprises technologiques, laboratoires publics et établissements d'enseignement supérieur.

Certaines zones apparaissent également du fait de leur spécialisation sectorielle, où les besoins en IA sont particulièrement marqués. C'est notamment le cas de Niort, pôle d'expertise en assurance et services financiers, de Toulouse, centre névralgique de l'industrie aéronautique et spatiale, ou encore de Lannion, historiquement tournée vers les télécommunications. Ces dynamiques régionales illustrent la manière dont

les spécificités économiques locales structurent la demande en compétences IA sur le territoire.

FIGURE 3 – Classement des zones d’emploi selon les offres d’emplois AI



Les offres d’emploi liées à l’intelligence artificielle présentent une concentration géographique plus marquée que l’ensemble des autres offres, comme l’illustre la Figure 4, qui représente la courbe de Lorenz de la part cumulée de l’emploi par zone d’emploi, classées selon le nombre total d’entrées dans la base JOCAS. Une telle concentration n’a rien d’étonnant s’agissant d’un domaine encore émergent, fortement intensif en R&D, et tiré par quelques pôles d’innovation majeurs. Ce phénomène est bien documenté dans la littérature sur la géographie de l’innovation, qui souligne la tendance des technologies de pointe à se développer dans un nombre restreint de clusters bénéficiant d’externalités locales et de main-d’œuvre spécialisée (Audretsch and Feldman, 1996).

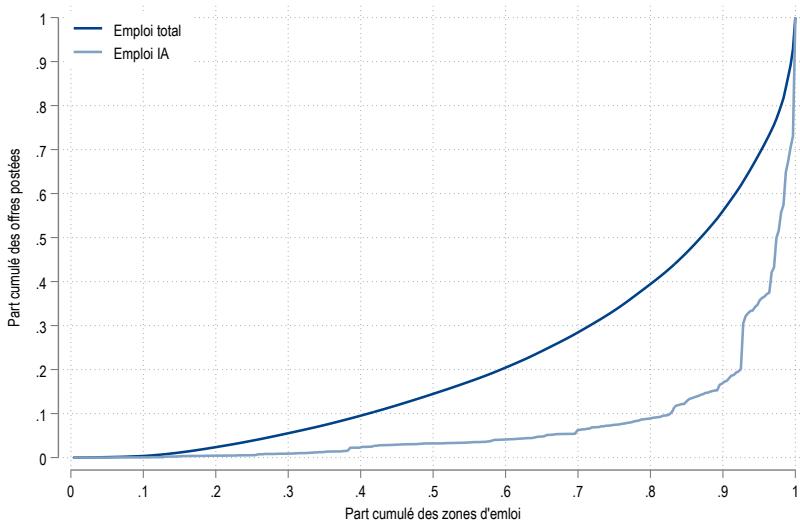
3.4 Diffusion de l’IA au sein des entreprises

Bien que le code administratif de l’entreprise publant l’offre ne soit pas disponible dans la base JOCAS, le nom de l’entreprise est en général précisé. Grâce à l’API SIRENE, nous avons pu identifier automatiquement 44% des entreprises à partir de leur nom,⁴ ce qui correspond à 77% des offres.

L’appariement avec le répertoire SIRENE fournit en outre des informations précieuses

⁴Des variations dans le nom renseigné, notamment par rapport à la dénomination sociale “officielle” de l’entreprise, peuvent entraîner des erreurs d’appariement. Afin de garantir la qualité de cette opération sur l’échantillon qui nous intéresse, nous avons procédé à une vérification manuelle sur 8942 entreprises, incluant l’ensemble des entreprises ayant publié au moins une offre identifiée comme liée à l’IA par notre algorithme ainsi que les noms d’entreprises ayant posté le plus d’offres jusqu’à obtenir 30% des offres

FIGURE 4 – Concentration des offres par zone d’emploi



Notes : Part cumulée des offres d’emplois dans JOCAS en par zone d’emploi. Les zones d’emplois sont classés de la plus petite à la plus grande à partir de la totalité des entrées dans JOCAS.

sur les entités identifiées : code secteur (NAF), catégorie de taille (TPE, PME, ETI, GE), statut juridique, etc.

Une rapide analyse sectorielle montre, comme attendu, que l’adoption de l’IA non générative s’est initialement concentrée dans l’industrie : entre 2019 et 2025, la part des entreprises ayant publié au moins une offre dans cette catégorie atteint 0,94% dans l’industrie contre 0,79% dans les services. À l’inverse, les offres liées à l’IA générative sont plus présentes dans les services (0,12%) que dans l’industrie (0,08%).

Bien que ces niveaux restent faibles, ils confirment l’idée selon laquelle l’IA générative, particulièrement adaptée à certaines tâches intellectuelles et créatives caractéristiques des services, y diffuse plus rapidement. Cette dynamique est renforcée après 2023, suivant la mise à disposition grand public de technologies comme ChatGPT. Le secteur tertiaire, bien que moins anticipateur que certaines entreprises industrielles en matière d’IA émergente, semble rattraper son retard grâce à l’accessibilité croissante de ces outils.

Par exemple, une enquête de Bpifrance Le Lab menée en 2024 auprès de 3 077 dirigeants d’entreprise indique que l’usage régulier ou occasionnel de l’IA générative concerne 24% des entreprises du secteur des services, contre seulement 12% dans l’industrie (Bpifrance, 2024). Cette tendance reflète une adoption tirée par les fonctions support (communication, marketing, rédaction) plus fréquemment internalisées dans les services, et donc directement concernées par les apports de l’IA générative.

En regardant plus en détail, on constate en revanche que l'adoption de l'intelligence artificielle est très hétérogène en fonction des secteurs et de la taille des entreprises. Parmi les grandes entreprises ("GE"), près de 6,2% des entreprises ont posté au moins une offre pour le développement de l'intelligence artificielle contre 0,6% des PME et 2,9% des ETI. Cette différence notable indique probablement une plus grande capacité à internaliser la production d'IA.

Concernant les secteurs, le Tableau 4 présente l'adoption de l'IA pour différents codes NAF, incluant les technologies génératives, à travers les offres d'emploi publiées entre 2019 et 2025 parmi les différents secteurs. Les secteurs les plus avancés sont sans surprise les secteurs producteurs des technologies de l'information et de la communication et celui des activités spécialisées scientifiques et techniques (comprenant le conseil et les bureaux d'études), où l'on observe à la fois une forte adoption de l'IA en général (jusqu'à 5,1% pour l'IT) et une présence notable d'offres associées à l'IA générative (0,72% pour la production de TIC, 0,30% pour l'enseignement). Ces résultats confirment que l'IA, dans ses différentes formes, reste largement concentrée dans des secteurs disposant de fortes capacités technologiques, organisationnelles et cognitives — une observation en ligne avec les travaux de [Bessen \(2018\)](#). Ces secteurs se distinguent non seulement par leur anticipation technologique, mais aussi par leur capacité à intégrer rapidement les innovations dans leurs processus, souvent à travers des métiers spécialisés (*data scientists*, ingénieurs IA, consultants IA), essentiels à la diffusion effective des outils. Le cas du conseil est emblématique de cette dynamique : en tant qu'activité transverse, il agit à la fois comme utilisateur direct et comme relais de diffusion dans d'autres secteurs.

Cette structuration sectorielle, marquée par une surreprésentation des services intensifs en capital humain, s'explique en grande partie par la nature des tâches réalisées. L'IA générative, en particulier, est fortement complémentaire d'activités impliquant la production de contenu, l'analyse textuelle ou la relation usager — typiquement rencontrées dans le tertiaire supérieur (enseignement, services administratifs, etc.). À l'inverse, les secteurs à forte composante physique ou routinière — comme l'industrie (0,96%IA, 0,09%IA générative), le transport ou la construction — présentent une adoption plus faible, en cohérence avec la typologie des tâches développée par [Acemoglu and Restrepo \(2022\)](#). L'adoption de l'IA y reste davantage tirée par l'automatisation de processus techniques ou logistiques, là où l'IA générative ne trouve encore que peu de cas d'usage opérationnels.

TABLE 4 – Entreprises ayant adopté l'IA et l'IA générative par secteur agrégé

Secteur	Code NAF	Nombre d'entreprises	Adopteurs IA	Adopteurs IA générative
Information et communication	58–63	20 512	5.09%	0.72%
Activités spécialisées	69–75	50 627	2.17%	0.22%
Enseignement	85	13 381	1.67%	0.30%
Activités financières	64–66, 89	14 932	1.52%	0.19%
Administration	84	4 402	1.23%	0.27%
Services administratifs et de soutien	77–82	35 200	1.06%	0.13%
Activités immobilières	68	25 135	0.72%	0.06%
Autres services	94–96	27 968	0.55%	0.04%
Électricité - Eau	35–39	3 454	0.55%	0.17%
Industrie	10–33	48 756	0.53%	0.06%
Commerce auto	45–47	85 082	0.30%	0.03%
Arts et spectacles	90–93	13 627	0.27%	0.02%
Santé humaine et action sociale	86–88	22 912	0.23%	0.01%
Transports et entreposage	49–53	27 955	0.16%	0.03%
Construction	41–43	64 782	0.10%	0.01%
Agriculture	01–03	17 428	0.09%	0.00%
Hébergement et restauration	55–56	52 138	0.08%	0.01%
Tous secteurs (non-missing uniquement)		528 291	0.79%	0.09%
Secteur manquant		702 773	0.15%	0.01%

4 Effet sur les flux d'emplois

Cette section propose d'analyser l'impact de l'irruption de ChatGPT au quatrième trimestre de 2022 sur la demande de travail, en particulier sur certaines catégories d'emplois particulièrement exposées à l'IA.

4.1 Exposition à l'IA et flux d'emploi au niveau métier

Dans un premier temps, nous construisons un panel cylindré au niveau *métier-trimestre* (code ROME k au trimestre t), comprenant 525 unités distinctes observées sur 20 trimestres, du premier trimestre 2020 au quatrième trimestre 2024. Pour chaque cellule (k, t) , nous observons le nombre total d'offres d'emploi publiées, noté $Y_{k,t}$, associé à la catégorie ROME k durant le trimestre t . Par construction, $Y_{k,t}$ est strictement positif sur l'ensemble du panel.

Nous estimons ensuite le modèle suivant par moindres carrés ordinaires :

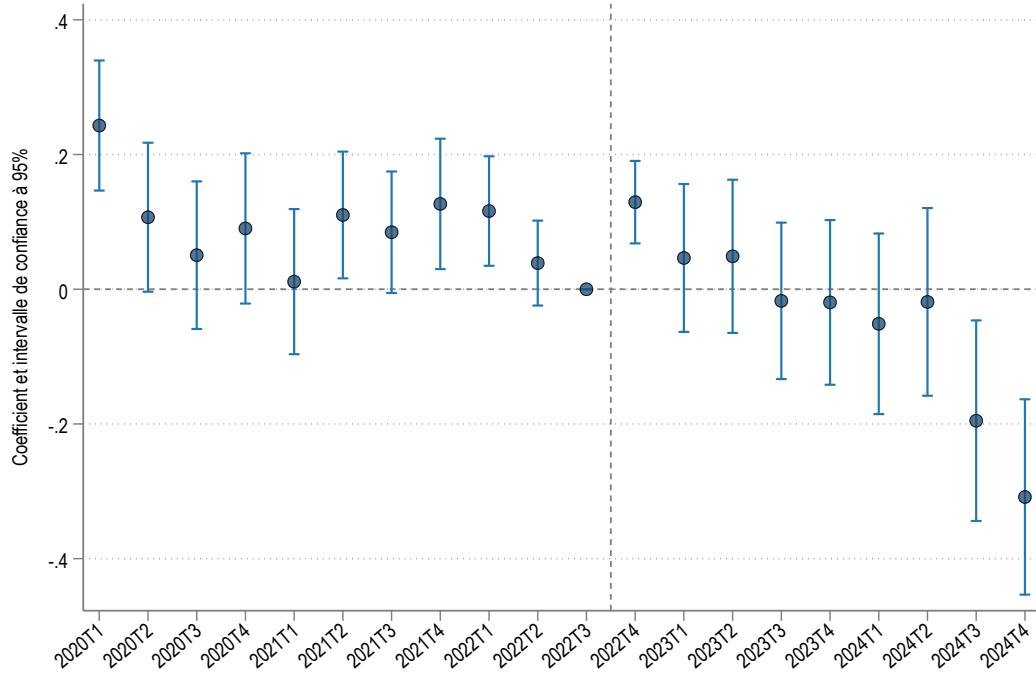
$$\log(Y_{k,t}) = \sum_{\tau=2020T1}^{2024T4} \gamma_\tau \mathbb{1}(t = \tau) \times \mathcal{E}_k + \mu_k + \delta_t + \varepsilon_{k,t} \quad (1)$$

où $\mathbb{1}(t = \tau)$ est une fonction indicatrice prenant la valeur 1 si $t = \tau$, et 0 sinon. Le terme \mathcal{E}_k mesure l'intensité d'exposition à l'IA générative de la profession k , selon la méthodologie proposée par [Bergeaud \(2024a\)](#). Les termes μ_k et δ_t représentent respectivement des effets fixes par métier (code ROME) et par trimestre, capturant les hétérogénéités temporelles communes et les effets spécifiques aux professions. L'ensemble des coefficients γ_τ capture l'effet différentiel de l'exposition à l'IA générative sur le volume d'offres d'emploi au cours du temps. À des fins d'identification, le trimestre 2022T3 — précédent immédiatement la mise à disposition publique de ChatGPT — est utilisé comme période de référence, et l'on impose $\gamma_{2022T3} = 0$.

Les résultats de cette estimation sont présentés en figure 5. Les valeurs rapportées correspondent aux estimations ponctuelles des coefficients γ_τ , accompagnées d'intervalles de confiance à 95%. Ces derniers sont construits à partir d'erreurs standards robustes à l'hétéroscédasticité, avec correction de la dépendance intra-groupe via un regroupement des erreurs standards par code ROME (*clustered standard errors*).

Les semblent indiquer une baisse significative des recrutements pour les professions les plus affectées par l'IA générative près d'un an et demi après l'irruption de cette technologie ayant popularisé l'usage de l'IA générative. Les coefficients pré-traitement sont légèrement positifs ce qui semble suggérer que le trimestre de référence correspond à un niveau relativement un peu bas pour les offres d'emplois de ces professions, toutefois la somme des coefficients pré-traitement est non significativement différente de 0.

FIGURE 5 – Estimation des coefficients du modèle (1)



Notes : Cette figure présente les coefficients estimés γ_τ du modèle (1), mesurant l'effet différencié de l'exposition à l'IA générative sur la dynamique des offres d'emploi, par trimestre. Le panel couvre 525 codes ROME (k) sur 20 trimestres, de 2020T1 à 2024T4. L'unité de dépendance utilisée pour le regroupement des erreurs standards est le code ROME (*clustering* par k). Le trimestre 2022T3 est pris comme période de base (valeur de référence normalisée à zéro). Les points représentent les estimations ponctuelles des coefficients d'interaction γ_τ , tandis que les barres verticales indiquent les intervalles de confiance bilatéraux à 95%.

4.2 Exposition à l'IA générative et dynamiques de recrutement au niveau entreprise

Une autre manière d'analyser l'effet différencié de l'arrivée de ChatGPT consiste à exploiter l'hétérogénéité d'exposition à l'IA générative au niveau des entreprises. Nous sélectionnons 7,186 entreprises,⁵ parmi lesquelles 3,304 ont publié au moins une offre associée à l'IA (au sens large), et 429 une offre en lien explicite avec l'IA générative.

Chaque entreprise i est caractérisée par une variable d'exposition \mathcal{F}_i à l'IA générative, construite à partir des offres postées avant 2022T3 selon la formule suivante :

$$\mathcal{F}_i = \frac{1}{N_i} \sum_k \mathcal{E}_k \times n_{k,i} \quad (2)$$

où \mathcal{E}_k représente l'intensité d'exposition du code ROME k à l'IA générative (voir

⁵[À compléter : critère de sélection, e.g. entreprises identifiables dans JOCAS et ayant publié au moins X offres entre 2019 et 2022.]

Bergeaud, 2024a), $n_{k,i}$ le nombre d'offres publiées par l'entreprise i avant 2022T3 pour ce métier, et $N_i = \sum_k n_{k,i}$ le total des offres publiées avant cette date. Cette mesure agrège donc le niveau d'exposition à l'IA générative des métiers recrutés par l'entreprise avant l'introduction publique de ChatGPT.

Nous estimons ensuite le modèle suivant par moindres carrés ordinaires :⁶

$$Y_{i,t}^{AI} = \sum_{\tau=2020T1}^{2024T4} \gamma_\tau \mathbb{1}(t = \tau) \times \mathcal{F}_i + \mu_i + \delta_t + \varepsilon_{i,t} \quad (3)$$

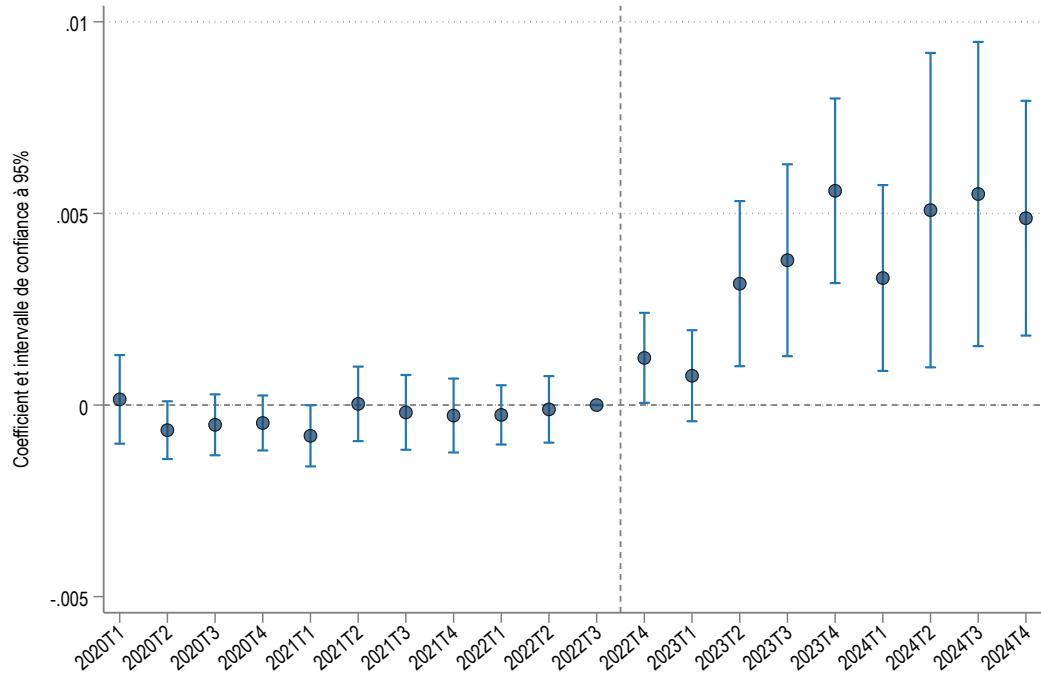
où $Y_{i,t}^{AI}$ est une variable valant 1 si une offre d'emploi associées à l'IA générative est publiées par l'entreprise i au trimestre t et 0 sinon. Comme précédemment, on impose $\gamma_{2022T3} = 0$ en tant que période de référence. Les effets fixes μ_i et δ_t capturent respectivement l'hétérogénéité inobservables entre entreprises et les chocs temporels communs. L'estimation est robuste à l'hétérosécédasticité, avec erreurs standards regroupées (*clustered*) par code NAF.

Les résultats sont présentés en figure 6, et suggèrent que les entreprises initialement les plus exposées à l'IA générative sont celles qui recrutent le plus activement dans ce domaine après fin 2022. Ce résultat semble suggérer une première vague de recrutement IA, principalement dans les grandes entreprises, afin de substituer ou de compléter les métiers existants particulièrement susceptibles d'être impacté par cette technologie.⁷

⁶Les résultats sont similaires avec un modèle Probit ou Logit, comme présenté Figure ??.

⁷Les résultats sont similaires avec un modèle de comptage de type Poisson, comme présentés en Annexe C.

FIGURE 6 – Effet de l'exposition IA sur le recrutement en IA générative



Notes : Cette figure illustre les coefficients estimés γ_τ du modèle (3), mesurant l'effet différencié de l'exposition initiale à l'IA générative (avant 2022T3) sur le nombre d'offres d'emploi liées à l'IA générative publiées par entreprise et par trimestre. L'échantillon comprend 7 186 entreprises identifiées dans JOCAS. La variable d'exposition \mathcal{F}_i est calculée comme une moyenne pondérée de l'exposition des codes ROME historiquement recrutés par l'entreprise (équation (2)). Le trimestre 2022T3 est pris comme période de référence ($\gamma_{2022T3} = 0$). Les intervalles de confiance à 95% sont construits à partir d'erreurs standards robustes regroupées par code NAF. Les résultats indiquent une corrélation positive entre exposition initiale et intensité de recrutement en IA générative après la diffusion publique de ChatGPT.

Avant de discuter plus en détail les effets potentiels sur l'emploi, nous procérons à un test de sensibilité visant à s'assurer que le lien observé entre l'exposition initiale à l'IA générative (\mathcal{F}_i) et le recrutement effectif en IA générative ne résulte pas d'une simple mécanique de composition liée aux flux de recrutement.

En effet, la variable \mathcal{F}_i mesure l'exposition d'une entreprise i à l'IA générative à partir de la structure de ses recrutements avant l'arrivée de ChatGPT (équation (2)). Cependant, il est possible que ce lien soit mécaniquement induit par la structure de $s_{k,i}$ elle-même, indépendamment des valeurs réelles de \mathcal{E}_k . Par exemple, une entreprise ayant une distribution plus concentrée de ses recrutements sur quelques codes ROME pourrait mécaniquement avoir une valeur de \mathcal{F}_i plus élevée, même en l'absence d'un effet causal de l'IA générative. Plus généralement, si certaines structures de recrutement sont plus dynamiques après 2022T3 (quelle que soit leur exposition réelle à l'IA générative), alors l'effet estimé pourrait être un artefact de ces flux, et non de \mathcal{E}_k .

Pour évaluer ce risque, nous mettons en place un test placebo fondé sur une permu-

tation des valeurs d'exposition. Nous construisons 5 000 variables d'exposition contre-factuelles $\mathcal{F}_i^{(p)}$, définies comme :

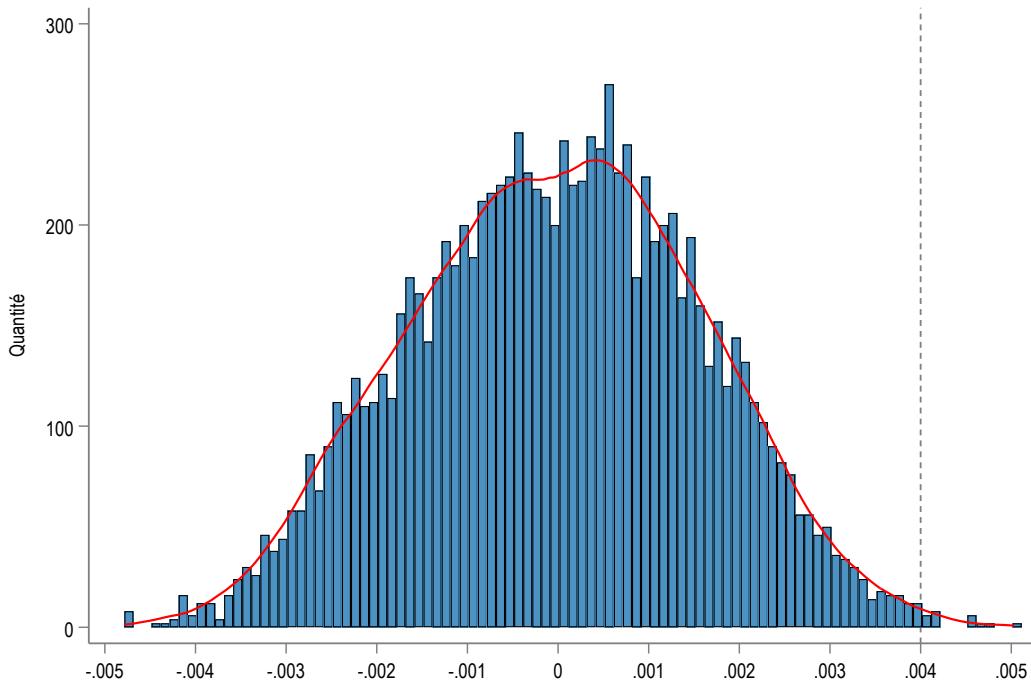
$$\mathcal{F}_i^{(p)} = \sum_k s_{k,i} \cdot \mathcal{E}_k^{(p)} \quad \text{avec } p = 1, \dots, 5000$$

où, pour chaque itération p , les valeurs de $\mathcal{E}_k^{(p)}$ sont obtenues par rééchantillonnage aléatoire (sans remise) des expositions réelles \mathcal{E}_k à travers l'ensemble des codes ROME. Autrement dit, on conserve la structure des flux de chaque entreprise ($s_{k,i}$) mais on lui associe une exposition tirée au hasard, indépendante des métiers effectivement recrutés. Cela permet de tester si le lien observé est spécifique aux vraies valeurs de \mathcal{E}_k , ou s'il pourrait émerger aléatoirement du seul fait de la structure de recrutement.

Si l'effet estimé — défini ici comme la moyenne des coefficients γ_τ du modèle (3) après la période de référence — obtenu avec la véritable variable d'exposition \mathcal{F}_i se situe en dehors de la distribution des effets obtenus avec les 5 000 variables placebo $\mathcal{F}_i^{(p)}$, cela constitue une indication indirecte mais robuste de la validité informative de notre mesure d'exposition. Cela suggère que le lien observé avec les comportements de recrutement en IA générative ne peut pas s'expliquer uniquement par la structure des parts $s_{k,i}$, c'est-à-dire la composition des flux de recrutement antérieurs.

La figure 7 illustre ce test sous la forme d'un histogramme représentant la distribution empirique des moyennes post-traitement des coefficients γ_τ estimés avec les expositions contrefactuelles $\mathcal{F}_i^{(p)}$. La valeur obtenue à partir de la véritable variable \mathcal{F}_i égale à 0.0037 est représentée par une ligne verticale, et se situe nettement dans la queue droite de la distribution. Cette configuration implique une faible probabilité qu'un effet aussi élevé soit obtenu aléatoirement en conservant uniquement la structure des parts $s_{k,i}$, renforçant ainsi l'interprétation causale de notre mesure d'exposition.

FIGURE 7 – Valeur moyenne des coefficients de régression randomisés



Notes : Cette figure montre la distribution empirique des moyennes des coefficients γ_τ postérieurs à 2022 T3 issus du modèle (3), obtenues à partir de 5 000 expositions contrefactuelles $\mathcal{F}_i^{(p)}$ construites par permutation aléatoire des \mathcal{E}_k . La droite verticale indique la moyenne observée avec la véritable variable \mathcal{F}_i .

5 Discussion : quel effet sur l'emploi ?

Les résultats présentés dans la section précédente suggèrent que, dans le cas de l'IA générative, les professions les plus exposées à cette technologie connaissent une diminution relative de la demande de travail, mesurée par le nombre d'offres d'emploi publiées, depuis la mise sur le marché de ChatGPT. Ce constat peut, à première vue, sembler en contradiction avec les travaux antérieurs de la littérature, et notamment ceux de [Aghion et al. \(2025\)](#) dans le contexte français, qui documentent une relation positive entre l'adoption de l'IA et l'emploi au sein des entreprises. Il convient néanmoins de souligner que les résultats présentés dans la section précédente portent sur la demande de travail, mesurée à travers le nombre d'offres d'emploi publiées, et non sur le stock d'emploi, comme c'est le cas dans [Aghion et al. \(2025\)](#).

Une première piste pour réconcilier ces résultats consiste à postuler une hétérogénéité des effets selon le type de technologies d'IA considérées. Les technologies d'IA générative, de par leur capacité à traiter des tâches cognitives complexes, pourraient engendrer des

effets différents de ceux obtenus pour des générations précédentes de technologies d'IA. Néanmoins, la diffusion très récente de ces technologies (novembre 2022 dans le cas de ChatGPT) limite à ce stade l'exploitation des données administratives pour en évaluer précisément les effets à moyen terme.

Au-delà des différences technologiques, une seconde piste pour réconcilier ces résultats consiste à observer que les chocs d'adoption de l'IA sont de nature différente. Dans la section précédente, l'adoption est mesurée à travers la publication d'une offre d'emploi requérant un profil fortement orienté IA, qu'il s'agisse d'un ingénieur chargé du développement de l'IA (comme c'est majoritairement le cas), d'un poste de support IA, de management IA, ou d'une fonction connexe (2). En revanche, dans [Aghion et al. \(2025\)](#), l'adoption correspond au fait qu'une entreprise déclare utiliser une technologie d'IA⁸ dans le cadre de l'Enquête sur les technologies de l'information et de la communication (Enquête TIC) en 2021, que cette technologie soit achetée sur le marché ou développée en interne.⁹

Dans cette section, nous nous intéressons donc aux effets de l'adoption de l'IA selon différentes définitions sur le stock d'emploi au niveau des entreprises, en mobilisant les données de la Base Tous Salariés (BTS) sur la période 2015–2023, afin d'éclairer le débat sur l'impact de l'IA sur l'emploi total.

Dans un premier temps, nous analysons l'effet associé à la publication d'une offre d'emploi IA. Une entreprise est considérée comme adoptante si elle a publié au moins une offre d'emploi IA entre janvier 2019 et décembre 2020.¹⁰ Afin de s'assurer que ce recrutement n'est pas anecdotique, nous retenons uniquement les entreprises pour lesquelles la part des offres d'emploi IA dans le total des offres publiées sur la période dépasse 0,09%.¹¹ Le groupe de contrôle est constitué des entreprises interrogées dans l'enquête TIC de 2019, appartenant à un secteur d'activité (NAF à 5 chiffres) dans lequel au moins une entreprise a publié une offre d'emploi IA en 2019 ou 2020. Cela permet de comparer des entreprises exposées à des secteurs d'activités similaires. Nous estimons

⁸Les technologies d'IA retenues sont les suivantes : "fouille de texte – text mining", "reconnaissance automatique de la parole", "génération automatique de langage naturel", "reconnaissance d'images, traitement d'images", "apprentissage automatique pour l'analyse de données", "logiciels d'automatisation des processus robotisés", "robots autonomes, véhicules autonomes, drones autonomes".

⁹Les caractéristiques des entreprises uniquement utilisatrices et celles développant l'IA a été précédemment étudié par [Calvino et al. \(2024\)](#). Ils montrent qui montre que les entreprises utilisatrices d'IA affichent une productivité plus élevée, qui semble s'expliquer par un phénomène d'auto-sélection des entreprises déjà plus productives et plus intensives en numérique, tandis que pour les entreprises développant l'IA, le lien positif entre usage de l'IA et productivité persiste au-delà des effets de sélection.

¹⁰Cette période correspond à l'intervalle entre deux éditions de l'enquête TIC intégrant des questions sur l'adoption de l'IA.

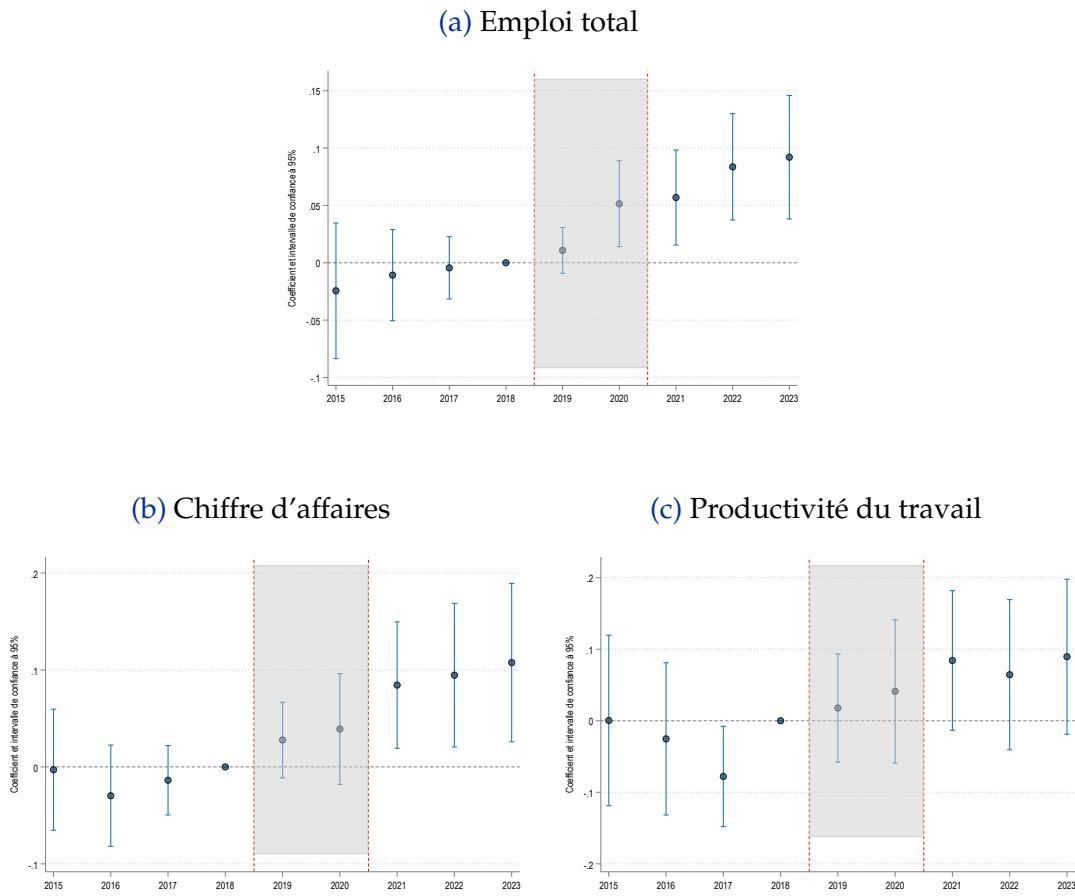
¹¹Ce seuil correspond au premier décile de la distribution du ratio entre offres IA et offres totales parmi les entreprises ayant publié au moins une offre IA sur la période.

ensuite les effets de cette adoption sur différentes variables d'emploi en adoptant une approche en différence-de-différences sur un panel cylindré d'entreprises entre 2015 et 2023. L'échantillon comprend 97 entreprises ayant publié une offre IA et 2 297 entreprises ne l'ayant pas fait. La spécification est la suivante :

$$\Delta \log(Y_{it}) = \alpha + \delta_t \cdot OffreIA_{it} + \mu_i + \lambda_{st} + \epsilon_{it}$$

où Y_{it} désigne la variable d'intérêt dans l'entreprise i à l'année t . La variable $OffreIA_{it}$ indique si l'entreprise a adopté l'IA via une offre d'emploi. Les termes μ_i et λ_{st} correspondent respectivement aux effets fixes entreprise et secteur x année.¹² Les résultats sont présentés dans la figure 8.

FIGURE 8 – Effet du recrutement de profils IA au niveau des entreprises

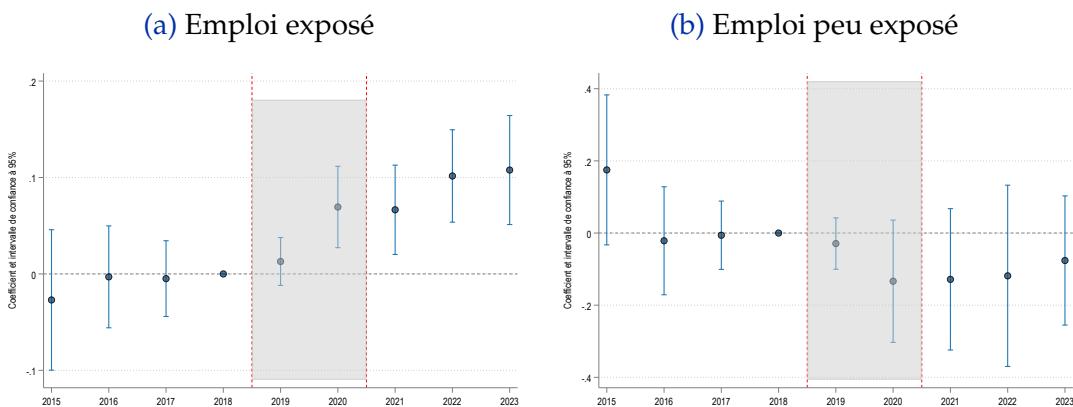


La figure 8a analyse l'effet du recrutement de profils spécialisés en IA sur l'emploi total. Alors que les deux groupes d'entreprises présentaient des trajectoires similaires

¹²On utilise le niveau "section" de la nomenclature NAF pour les effets fixes sectoriels.

avant 2019, celles ayant recruté des profils IA en 2019–2020 enregistrent par la suite une hausse plus marquée de leur emploi. Plus précisément, quatre ans après le recrutement, ces entreprises affichent une augmentation de l’emploi total supérieure de 10% à celle des entreprises comparables n’ayant pas procédé à de tels recrutements. Des résultats analogues sont obtenus lorsque l’on considère le chiffre d’affaires total avec un surcroît de chiffre d’affaires d’environ 10% après quatre ans pour les entreprises ayant recruté des profils IA (Figure 8b). Enfin, on observe des gains de productivité du travail plus marqués pour ces mêmes entreprises, bien que les estimations soient plus bruitées (Figure 8c). Dans l’ensemble, ces résultats suggèrent que le recrutement de profils IA contribue à une amélioration de la productivité, qui se traduit par une hausse du chiffre d’affaires et, en conséquence, par une augmentation de la demande de travail.¹³

FIGURE 9 – Hétérogénéité de l’effet du recrutement de profils IA



Toutefois, la réponse en termes de demande de travail pourrait être hétérogène selon le degré d’exposition des emplois à l’IA.¹⁴ Les résultats, présentés à la figure 9, indiquent que le recrutement de profils IA est associé à une augmentation de l’emploi dans les professions fortement exposées à l’IA au cours des années suivantes, avec une semi-élasticité estimée à 0,10, proche de celle observée pour l’emploi total (Figure 9a). En revanche, l’emploi dans les professions peu exposées tend à diminuer, avec une semi-élasticité estimée à -0,10, bien que les intervalles de confiance soient relativement larges (Figure 9b). Ces résultats suggèrent que le recrutement de profils IA conduit à une réorganisation de la main d’œuvre au sein des entreprises, orientée vers une

¹³Ce constat est cohérent avec les résultats de Aghion et al. (2022), qui mettent en évidence des effets positifs de l’adoption de technologies d’automatisation sur l’emploi dans le secteur manufacturier.

¹⁴Bergeaud (2024a) adapte au cas du marché du travail français l’approche de Gmyrek et al. (2023), en calculant pour chaque profession une « exposition globale à l’IA », définie comme une moyenne pondérée de l’exposition des tâches composant l’emploi selon la base de données O*NET, les pondérations reflétant l’importance relative de chaque tâche au sein de la profession.

intensification des profils fortement liés à l'IA.

Finalement, le recrutement de profils IA est positivement associé à l'emploi au niveau des entreprises. Ce résultat suggère que ces entreprises développent des produits intégrant l'IA, que ce soit pour un usage interne ou à des fins commerciales, ce qui les conduit à intensifier leur recours à des compétences spécialisées et, in fine, à accroître leur part de marché et leur niveau d'emploi. Ces effets apparaissent cohérents avec les résultats présentés par [Aghion et al. \(2025\)](#), qui documentent une hausse de l'emploi liée à l'adoption de l'IA au sens large, qu'il s'agisse de solutions développées en interne ou acquises sur le marché. Dans la suite de cette section, nous étudions justement l'impact sur l'emploi selon le type d'acquisition déclaré dans l'enquête TIC. L'analyse repose sur un panel cylindré d'entreprises observées entre 2015 et 2023, restreint à celles n'ayant pas encore adopté l'IA en 2018. On considère trois groupes : (i) les 610 entreprises n'ayant pas adopté l'IA entre 2018 et 2020, qui constituent le groupe de contrôle, (ii) les 39 entreprises ayant adopté une solution d'IA développée en interne, et (iii) les 185 entreprises ayant adopté une solution d'IA achetée sur le marché sur la même période.

Considérons d'abord comme groupe traité les entreprises ayant adopté une solution d'IA développée en interne entre 2018 et 2020. On considère alors la spécification suivante :

$$\Delta \log(Y_{it}) = \alpha + \delta_t \cdot AI_{it,int} + \mu_i + \lambda_{st} + \epsilon_{it}$$

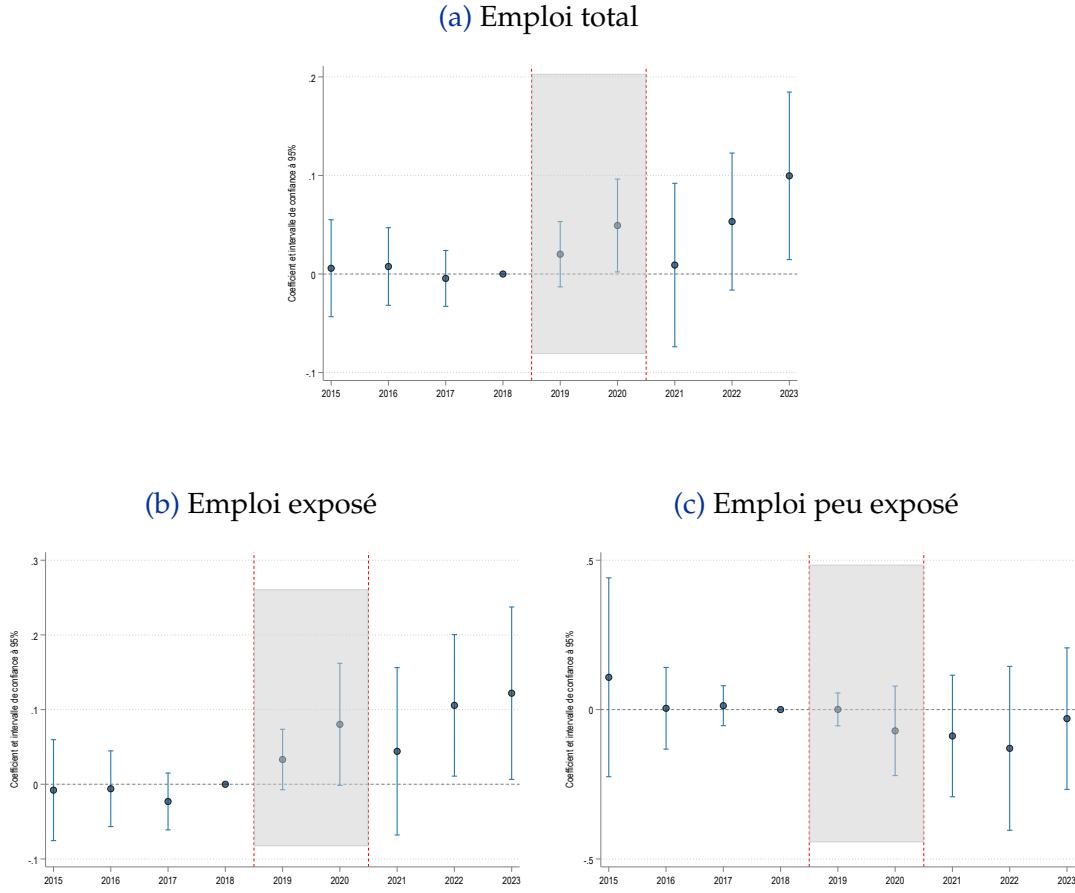
où Y_{it} désigne la variable d'intérêt dans l'entreprise i à l'année t . La variable $AI_{it,int}$ indique si l'entreprise a adopté l'IA après l'avoir développée en interne. Les termes μ_i et λ_{st} correspondent respectivement aux effets fixes entreprise et secteur x année.¹⁵

Les résultats sont présentés à la figure 10. Que l'on considère l'emploi total, l'emploi exposé à l'IA ou l'emploi peu exposé, les tendances sont similaires entre les entreprises ayant développé une solution d'IA en interne et celles du groupe de contrôle n'ayant pas adopté l'IA avant 2018. On observe par la suite une relation positive avec l'emploi total (figure 10a), avec une semi-élasticité de 0,10, ainsi que sur l'emploi exposé à l'IA, avec une semi-élasticité d'ordre comparable (figure 10b). En revanche, l'effet sur l'emploi peu exposé à l'IA est négatif, bien que les estimations soient très bruitées (figure 10c). Ces résultats sont proches de ceux obtenus dans le cas du recrutement de profils IA, ce qui semble cohérent : les entreprises recrutant ce type de profils le font pour développer des solutions internes d'IA, qu'elles adoptent par la suite.

Considérons maintenant comme groupe traité les entreprises ayant adopté une solution d'IA achetée sur le marché sans développement interne entre 2018 et 2020. On

¹⁵On utilise le niveau "section" de la nomenclature NAF pour les effets fixes sectoriels.

FIGURE 10 – Effet de l'adoption d'IA via le développement interne sur l'emploi



considère alors la spécification suivante :

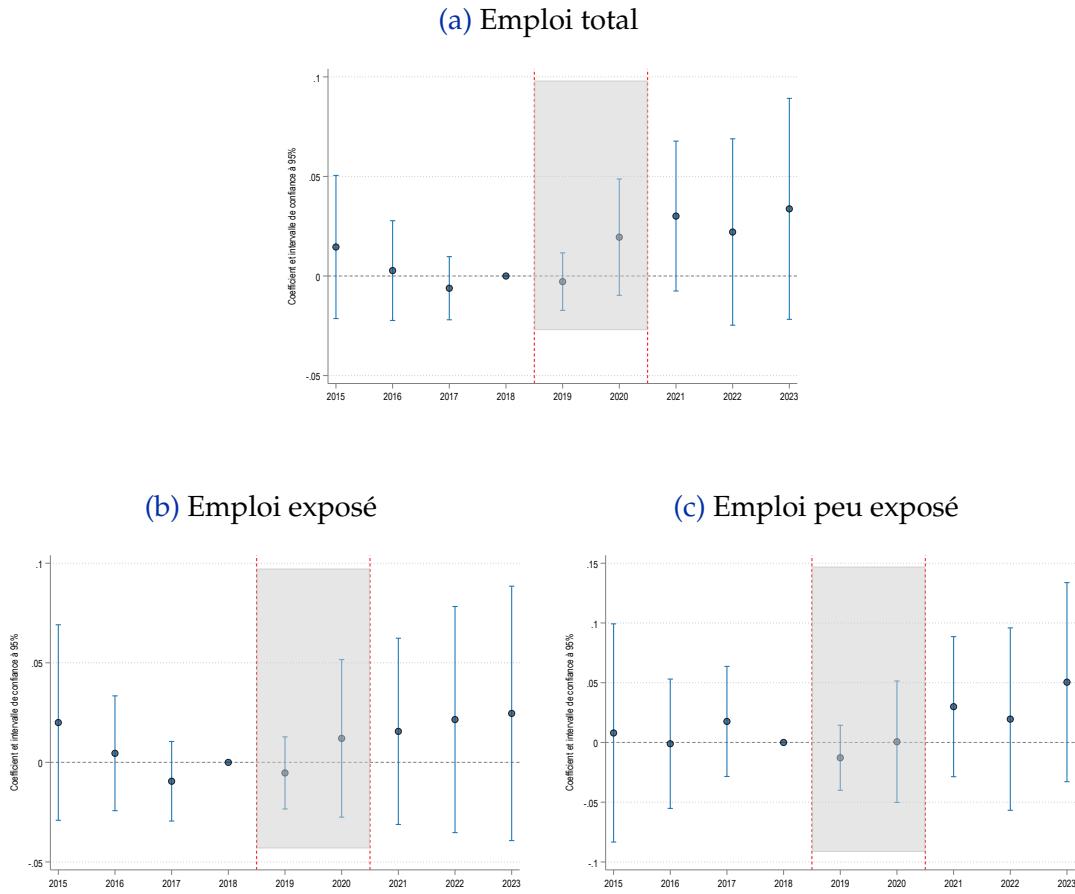
$$\Delta \log(Y_{it}) = \alpha + \delta_t \cdot AI_{it,ext} + \mu_i + \lambda_{st} + \varepsilon_{it}$$

où Y_{it} désigne la variable d'intérêt dans l'entreprise i à l'année t . La variable $AI_{it,ext}$ indique si l'entreprise a adopté l'IA sans l'avoir développée en interne. Les termes μ_i et λ_{st} correspondent respectivement aux effets fixes entreprise et secteur x année.¹⁶ Les résultats sont présentés à la figure 11.

Comme précédemment, on observe des tendances similaires entre les deux groupes avant 2018, quelle que soit la variable considérée. En revanche, les effets diffèrent sensiblement de ceux obtenus pour les entreprises recrutant des profils IA ou développant des solutions en interne. Si les entreprises adoptant une solution d'IA externe connaissent une augmentation de l'emploi légèrement supérieure à celle des entreprises n'adoptant pas l'IA, cet effet n'est pas statistiquement significatif (figure 11a). Par ailleurs, son

¹⁶On utilise le niveau "section" de la nomenclature NAF pour les effets fixes sectoriels.

FIGURE 11 – Effet de l'adoption d'IA via le développement externe sur l'emploi



ampleur reste modeste, avec une semi-élasticité de 0,03 après quatre ans, contre 0,10 dans le cas de l'adoption par développement interne. En ce qui concerne l'hétérogénéité des effets selon l'exposition des emplois à l'IA, les résultats s'écartent nettement des cas précédents : les effets sont faiblement positifs, mais non significativement différents de zéro, tant pour les emplois exposés (figure 11b) que pour les emplois peu exposés (figure 11c), avec des semi-élasticités similaires pour les deux catégories. Ces résultats suggèrent que l'adoption d'IA via une solution externe génère un effet de productivité modéré, qui bénéficie à l'ensemble des emplois, sans entraîner de réallocation marquée entre types de postes. Autrement dit, l'effet positif reste diffus et ne s'accompagne pas de déplacements d'emploi significatifs.

En résumé, la relation entre adoption de l'IA et emploi est plus faible, et non significative, lorsqu'il s'agit d'une adoption sans développement interne, contrairement au cas du développement interne où l'effet est significatif et plus marqué. Pour les emplois exposés à l'IA, l'effet est fortement positif dans le cas du développement interne

(semi-élasticité de +0,10), contre un effet légèrement positif mais non significatif dans le cas de l'adoption externe (+0,03). Pour les emplois peu exposés, l'effet est négatif mais bruité en cas de développement interne, tandis qu'il est faiblement positif et non significatif en cas d'adoption externe (semi-élasticité de +0,04). Ces différences mettent en évidence la complémentarité entre deux approches empiriques : d'une part, celle fondée sur les offres d'emploi liées à l'IA, utilisée dans cet article ainsi que dans de nombreux travaux récents [Calvino et al. \(2024\)](#); [Babina et al. \(2024\)](#); [McElheran et al. \(2024\)](#); [Hampole et al. \(2025\)](#); [Acemoglu et al. \(2022\)](#), et d'autre part, celle reposant sur les enquêtes auprès des entreprises pour mesurer leur adoption de l'IA, comme dans [Aghion et al. \(2025\)](#) ou [Calvino and Fontanelli \(2023\)](#). Les effets agrégés dépendront par ailleurs de la structure de l'adoption au sein du tissu productif, en particulier de la proportion relative d'entreprises développant en interne par rapport à celles se limitant à une adoption externe, ces dernières étant majoritaires en 2020.

6 Conclusion

Dans cet article, nous introduisons une nouvelle mesure de la diffusion de l'intelligence artificielle dans l'économie française à partir de l'analyse automatisée d'offres d'emploi sur la période allant de 2019 à 2024.

Nos résultats révèlent que si la part des offres d'emploi liée à l'IA au sens large ne connaît pas de progression marquée sur la période considérée, l'essor rapide et récent des annonces liées à l'IA générative suggère un tournant technologique. Les offres d'emploi concernent avant tout des profils techniques spécialisés mais, dans une moindre mesure, également des profils de support à l'IA. Ces offres sont par ailleurs concentrées dans certains secteurs et territoires, notamment ceux dotés de clusters technologiques.

Par ailleurs, nous mettons en évidence que, depuis la sortie de ChatGPT en novembre 2022, les entreprises qui avaient une part plus importante de profils exposés à l'IA dans leur main d'œuvre recrutent davantage de spécialistes dans ce domaine par la suite. Au niveau du marché du travail, la sortie de ChatGPT semble s'accompagner d'une baisse relative des offres d'emploi dans les professions les plus exposées à l'IA par rapport aux moins exposées.

Bien que nos résultats puissent sembler différer de ceux d'études antérieures ([Aghion et al., 2025](#); [Babina et al., 2024](#); [Hampole et al., 2025](#)), nous montrons qu'ils sont en réalité cohérents lorsqu'on considère le stock d'emploi total - plutôt que flux entrant - et que l'on distingue les modalités d'adoption de l'IA. Le recrutement de profils IA est associé à une hausse de la productivité du travail, du chiffre d'affaires et de l'emploi

total, mais avec une forte hétérogénéité : l'emploi progresse dans les métiers fortement exposés à l'IA et diminue dans les moins exposés. Les entreprises qui recrutent des profils en IA deviennent plus productives, ce qui entraîne une hausse de la demande à laquelle elles font face, et les conduit à augmenter leur niveau d'emploi pour y répondre. Des résultats similaires sont observés en mobilisant des données d'enquête sur les entreprises ayant développé leurs propres solutions d'IA, soulignant la convergence entre approches par des données exhaustives d'offres d'emploi et des données d'enquête. En revanche, l'adoption via des solutions externes a un effet plus limité, sans impact statistiquement significatif sur l'emploi. Alors que le développement de l'IA au sein des entreprises transforme profondément les processus internes et modifie les besoins de compétences, l'adoption externe accroît la productivité à la marge, sans reconfiguration organisationnelle majeure.

Enfin, en mettant à disposition un tableau de bord interactif de suivi des offres d'emploi liées à l'IA, cette recherche espère fournir un outil utile pour les décideurs publics et les chercheurs.¹⁷ Elle ouvre aussi la voie à de futurs travaux sur les effets de l'adoption de l'IA, notamment générative, que les données d'enquête pourraient avoir davantage de difficulté à capturer à court terme.

¹⁷Ce tableau de bord est accessible à l'adresse suivante : [XX](#).

Références

- Acemoglu, Daron**, "The simple macroeconomics of AI," *Economic Policy*, 2025, 40 (121), 13–58.
- **and Pascual Restrepo**, "Artificial intelligence, automation, and work," in "The economics of artificial intelligence : An agenda," University of Chicago Press, 2018, pp. 197–236.
- **and —**, "Tasks, automation, and the rise in US wage inequality," *Econometrica*, 2022, 90 (5), 1973–2016.
- , **David Autor, Jonathon Hazell, and Pascual Restrepo**, "Artificial intelligence and jobs : Evidence from online vacancies," *Journal of Labor Economics*, 2022, 40 (S1), S293–S340.
- Aghion, Philippe and Anne Bouverot**, *IA : Notre ambition pour la France*, Odile Jacob, 2024.
- , **Benjamin F Jones, and Charles I Jones**, *Artificial intelligence and economic growth*, Vol. 23928, National Bureau of Economic Research Cambridge, MA, 2017.
- , **Céline Antonin, and Simon Bunel**, "Artificial intelligence, growth and employment : The role of policy," *Economie et Statistique/Economics and Statistics*, 2019, (510-511-512), 150–164.
- , —, —, **and Xavier Jaravel**, "Modern manufacturing capital, labor demand, and product market dynamics : Evidence from France," 2022.
- , **Simon Bunel, Xavier Jaravel, Thomas Mikaelson, Alexandra Roulet, and Jakob Sogaard**, "How Different Uses of AI Shape Labor Demand : Evidence from France," 2025. Forthcoming AEA Paper and Proceedings.
- Agrawal, Ajay, John McHale, and Alexander Oettl**, "Finding needles in haystacks : Artificial intelligence and recombinant growth," in "The economics of artificial intelligence : An agenda," University of Chicago Press, 2018, pp. 149–174.
- , **Joshua Gans, and Avi Goldfarb**, *Prediction machines, updated and expanded : The simple economics of artificial intelligence*, Harvard Business Press, 2022.
- Alderucci, Dean, Lee Branstetter, Eduard Hovy, Andrew Runge, and Nikolas Zolas**, "Quantifying the impact of AI on productivity and labor demand : Evidence from US census microdata," in "Allied social science associations—ASSA 2020 annual meeting" 2020.
- Arora, Ashish, Andrea Fosfuri, and Alfonso Gambardella**, *Markets for technology : The economics of innovation and corporate strategy*, MIT press, 2004.

- Audretsch, David B and Maryann P Feldman**, "R&D spillovers and the geography of innovation and production," *The American economic review*, 1996, 86 (3), 630–640.
- Babina, Tania, Anastassia Fedyk, Alex He, and James Hodson**, "Artificial intelligence, firm growth, and product innovation," *Journal of Financial Economics*, 2024, 151, 103745.
- Bergeaud, Antonin**, "Exposition à l'intelligence artificielle générative et emploi : une application à la classification socio-professionnelle française," *document de travail*, 2024.
- , "The past, present and future of European productivity," in "ECB Forum on Central Banking" 2024, pp. 1–3.
- , **Adam B Jaffe, and Dimitris Papanikolaou**, "Natural Language Processing and Innovation Research," Working Paper w33821, National Bureau of Economic Research 2025.
- Bessen, James**, "AI and Jobs : The Role of Demand," Working Paper 24235, National Bureau of Economic Research January 2018.
- Bpifrance**, "L'IA générative dans les TPE et les PME," Technical Report, Bpifrance Le Lab 2024.
- Brynjolfsson, Erik and Kristina McElheran**, "The rapid adoption of data-driven decision-making," *American Economic Review*, 2016, 106 (5), 133–139.
- , **Danielle Li, and Lindsey R Raymond**, "Generative AI at work," Technical Report, National Bureau of Economic Research 2023.
- Calvino, Flavio and Luca Fontanelli**, "A portrait of AI adopters across countries : Firm characteristics, assets' complementarities and productivity," 2023.
- , **Chiara Criscuolo, Hélène Dernis, and Lea Samek**, "What technologies are at the core of AI? : An exploration based on patent data," Technical Report, OECD Publishing 2023.
- , **Hélène Dernis, Lea Samek, and Antonio Ughi**, "A sectoral taxonomy of AI intensity," OECD Artificial Intelligence Papers 30, OECD Publishing December 2024.
- , **Lea Samek, Mariagrazia Squicciarini, and Cody Morris**, "Identifying and characterising AI adopters : A novel approach based on big data," 2022.
- Cockburn, Iain M, Rebecca Henderson, Scott Stern et al.**, *The impact of artificial intelligence on innovation*, Vol. 24449, National bureau of economic research Cambridge, MA, USA, 2018.
- Czarnitzki, Dirk, Gastón P Fernández, and Christian Rammer**, "Artificial intelligence and firm-level productivity," *Journal of Economic Behavior & Organization*, 2023, 211, 188–205.

- Damioli, G, V Van Roy, and D Vertesy**, "The impact of artificial intelligence on labour productivity," *EURASIAN BUSINESS REVIEW*, 2021, 11, 1–25.
- Dares**, "Base JOCAS : Document méthodologique résumé," 2023. Accessed May 2025 : <https://data.progedo.fr/documents/adisp/lil-1586/Jocas-Methodologie,resumee.pdf>.
- Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock**, "GPTs are GPTs : Labor market impact potential of LLMs," *Science*, 2024, 384 (6702), 1306–1308.
- Felten, Ed, Manav Raj, and Robert Seamans**, "How will language modelers like ChatGPT affect occupations and industries?," *arXiv preprint arXiv:2303.01157*, 2023.
- Furman, Jason and Robert Seamans**, "AI and the Economy," *Innovation policy and the economy*, 2019, 19 (1), 161–191.
- Gentzkow, Matthew, Bryan Kelly, and Matt Taddy**, "Text as data," *Journal of Economic Literature*, 2019, 57 (3), 535–574.
- Gmyrek, Pawel, Janine Berg, and David Bescond**, "Generative AI and jobs : A global analysis of potential effects on job quantity and quality," *ILO working paper*, 2023, 96.
- Hampole, Menaka, Dimitris Papanikolaou, Lawrence D.W. Schmidt, and Bryan Seegmiller**, "Artificial Intelligence and the Labor Market," Working Paper 33509, National Bureau of Economic Research February 2025.
- McElheran, Kristina, J Frank Li, Erik Brynjolfsson, Zachary Kroff, Emin Dinlersoz, Lucia Foster, and Nikolas Zolas**, "AI adoption in America : Who, what, and where," *Journal of Economics & Management Strategy*, 2024, 33 (2), 375–415.
- Noy, Shakked and Whitney Zhang**, "Experimental evidence on the productivity effects of generative artificial intelligence," *Science*, 2023, 381 (6654), 187–192.
- Peng, Sida, Eirini Kalliamvakou, Peter Cihon, and Mert Demirer**, "The impact of ai on developer productivity : Evidence from github copilot," *arXiv preprint arXiv:2302.06590*, 2023.
- Pizzinelli, Carlo, Augustus J Panton, Ms Marina Mendes Tavares, Mauro Cazzaniga, and Longji Li**, *Labor market exposure to AI : Cross-country differences and distributional implications*, International Monetary Fund, 2023.
- Squicciarini, Mariagrazia, Lea Samek, and Emile Cammeraat**, "The human capital behind AI : Jobs and skills demand from online job postings," 09 2021.
- Webb, Michael**, "The impact of artificial intelligence on the labor market," *Available at SSRN 3482150*, 2019.

A Annexe : Description détaillées des données

A.1 Base de données Jocas

La base JOCAS (Job Offers Collection and Analysis System) est une base de données développée par la Dares à partir de la collecte quotidienne, depuis fin 2018, d'annonces d'offres d'emploi publiées en ligne sur un panel d'une quinzaine de sites de recrutement en France. Elle repose sur un processus automatisé de webscraping et, dans certains cas, sur l'utilisation d'API, permettant de constituer un corpus exhaustif d'offres couvrant une large variété de métiers, de qualifications et de localisations. Chaque offre est structurée autour d'un socle minimal d'informations (intitulé du poste, localisation, nom de l'employeur) et est enrichie via un algorithme de classification supervisée (SVM) permettant d'assigner un code ROME à partir de son contenu textuel, avec une erreur de classement inférieure à 5%. L'ensemble des informations disponibles dans JOCAS est affichée dans le Tableau A1. Des procédures de déduplication par URL et similarité textuelle garantissent l'unicité des offres dans la base finale.

A.2 Description des technologies

Les offres d'emploi peuvent être catégorisées selon la nature des technologies mobilisées. Quatre grandes classes sont identifiées dans ce cadre deux classes spécifiques faisant référence à des cas d'usages particulier : l'IA générative, vision par ordinateur ("CV"), et deux classes génériques ne mentionnant pas de cas d'applications particulier : General AI, et Sciences des données ("Data").

La catégorie Generative AI regroupe les technologies de génération de contenus (textes, images, vidéos) s'appuyant sur des modèles de type Large Language Models (LLMs), Generative adversarial Networks (GANs) ou modèles de diffusion. Elle couvre notamment le développement d'agents conversationnels et d'outils génératifs.

Computer Vision désigne l'ensemble des usages d'IA liés à l'analyse d'images ou de vidéos : détection d'objets, segmentation, classification visuelle ou traitement de flux vidéo.

La classe General AI rassemble les cas d'usage prédictifs fondés sur des méthodes classiques de machine learning, hors génération. Elle inclut par exemple l'analyse de séries temporelles, la classification ou la modélisation de texte non génératif.

Enfin, la catégorie Data regroupe les activités d'analyse de données sans recours explicite à des méthodes d'apprentissage automatique. Elle inclut l'extraction, la visualisation et l'interprétation de données à des fins descriptives ou décisionnelles.

TABLE A1 – Principales catégories de variables dans la base JOCAS

Catégorie	Description
Identifiants	Identifiant unique de l'offre et lien vers l'annonce d'origine.
Métadonnées	Informations sur le site source, la date de collecte, et le statut du scraping.
Dates	Dates d'apparition, de disparition et de publication de l'offre.
Métier et qualification	Intitulé du poste, code ROME, niveau de qualification.
Contrat	Type, durée et calendrier du contrat proposé.
Temps de travail	Régime horaire (temps plein/partiel), nombre d'heures hebdomadaires.
Localisation	Lieu du poste (libellé, code postal, département).
Salaire	Informations sur la rémunération (montants et fréquence).
Entreprise	Nom et identifiant SIREN de l'employeur.
Secteur d'activité	Classification selon les nomenclatures sectorielles de l'INSEE.
Télétravail	Présence ou non de télétravail et modalités associées.
Expérience requise	Niveau d'expérience exigé pour le poste.
Niveau d'éducation	Diplôme(s) et domaine(s) d'études requis.
Descriptions textuelles	Contenu libre de l'annonce : missions, profil recherché, présentation de l'employeur.

A.3 Description des classes

L'analyse des offres d'emploi permet de structurer les professions en lien avec l'intelligence artificielle selon quatre grandes classes fonctionnelles : Engineer, Support, Management et Fonction connexes. Chacune d'elles se distingue par la nature des tâches associées et les compétences mobilisées.

La catégorie **Engineer** regroupe les offres spécialisées dans le développement des modèles et d'algorithmes d'IA. Les professionnels de cette classe sont responsables de

Variable	Nombre d'offres valides	Pourcentage des offres
Description	109557979	97%
Code ROME	111130825	98%
Code postal	92718440	82%
Nom de l'entreprise	88240580	78%
Type de contrat	101686273	90%
Salaire	31962737	28%

TABLE A2 – Part des offres de JOCAS présentant des informations complètes

la conception, de l'entraînement, de l'évaluation et de l'amélioration des algorithmes d'apprentissage automatique. Les offres de cette catégorie mentionnent des compétences et des technologies spécifiques au machine-learning (SVM, K-means, CNN, RNN, deep learning, etc.) ainsi que des bibliothèques de développement spécialisées telles que TensorFlow, PyTorch, Keras ou scikit-learn.

La classe **Support** désigne les fonctions techniques impliquées dans l'intégration concrète des solutions d'IA dans les outils et systèmes de production. Les tâches incluent le développement de pipelines de données, d'APIs, de plateformes ou d'applications. Les compétences typiquement associées relèvent du développement logiciel (frontend, backend, full stack), et informatique "classique" avec des libraires et langages comme Java, C#, HTML, CSS, Angular, Docker, Kubernetes auxquels s'ajoutent des librairie de mise en production spécifiques à l'IA comme LangChain ou OLLaMa.

La catégorie **Management** renvoie aux postes de direction ou de coordination dans les projets d'IA. Elle regroupe des individus responsables du pilotage des équipes, de la planification stratégique et du suivi opérationnel. Ces fonctions requièrent une double compétence : d'une part, une compréhension technique des systèmes d'IA, et d'autre part, des aptitudes en gestion de projet, en encadrement d'équipe et en leadership.

Enfin, la classe **Fonction connexe** englobe les rôles plus transversaux et non liés au développement technique de l'IA mais pouvant l'utiliser ou la promouvoir au sein de l'organisation. Ces postes participent à l'identification des cas d'usage de l'IA, à l'accompagnement du changement, et à l'acculturation des organisations à ces technologies. Ces professionnels mobilisent des compétences en communication, en transformation numérique et en gestion de projet, avec un accent particulier sur l'acculturation à l'IA.

TABLE A3 – Typologie des classes d’emplois en IA, mots-clés associés et exemples de postes

Classe	Compétences / Mots-clés	Exemples de postes
Engineer	TensorFlow, PyTorch, scikit-learn, Keras, Hugging Face, machine learning, deep learning, transformers, SVM, CNN, RNN, LLM, réseaux de neurones	Data Scientist, AI Engineer, Machine Learning Engineer
Support	Java, CSS, HTML, Angular, Docker, Kubernetes, Langchain, OLLaMa, frontend, backend, SQL, full stack	Software Engineer, Data Engineer, Développeur full stack, Développeur web
Management	Machine learning, deep learning, gestion de projet, leadership, coordination	AI Project Manager, Head of Data Science, Technical Team Lead
Fonctions connexes	Communication, transformation numérique, AI literacy, gestion de projet, usage de l’IA	AI Consultant, Innovation Manager, Change Management Lead

A.4 Analyse sémantique du titre des offres

Le Tableau A4 nous donne un premier aperçu des intitulés les plus fréquent des métiers associés au déploiement de l’intelligence artificielle via l’identification du code ROME. Cependant, le seul recours à cette nomenclature ne permet pas de saisir toute la nuance des différents métiers associés à l’intelligence artificielle. En effet, l’algorithme développée par la Dares permettant d’identifier automatiquement le code ROME peut comporter des erreurs ou être et être mal aligné avec le contenu réel de l’offre. Par ailleurs, la nomenclature ROME, trop agrégée, ne permet pas de distinguer les spécialisations fines entre métiers de l’IA. Ainsi, pour faire émerger les titres d’offres les plus souvent associées aux métiers de l’IA, chaque intitulé d’offre est vectorisé à l’aide de la pondération TF-IDF (Term Frequency-Inverse Document Frequency), puis nous appliquons un algorithme de *clustering* (K-means) pour faire émerger des groupes de métiers sémantiquement proches. Les Tableau et montrent les cluster et les intitulés associés pour l’IA et l’IA générative.

TABLE A4 – Thèmes identifiés par clustering dans les postes IA

Cluster	Titre possible	Mots-clés dominants
0	Data Scientist	data scientist, scientist, data, lead data
1	Responsable IA / ingénieur IA	responsable, ia, data, ingénieur ia
2	Analyste / Manager data	data, analyst, data analyst, manager
3	Chef de projet data	chef projet, chef, projet, projet data
4	Ingénieur développement logiciel	ingénieur développement, développement logiciel
5	Vision par ordinateur	ingénieur vision, vision ordinateur, computer vision
6	Développeur Python / Big Data	développeur python, développeur big data
7	Data Engineer	data engineer, engineer, data, lead data
8	Ingénieur Big Data	engineer, ingénieur data, ingénieur big_data, big_data
9	Développement back-end	développeur, back end, développeur java

B Procédure de classification

Afin de traiter un volume important d'offres, une étape de pré-filtrage automatisée a été mise en place. Elle repose sur une liste étendue de mots-clés inspirée des travaux de [Calvino et al. \(2024\)](#) et [Babina et al. \(2024\)](#), permettant d'isoler environ 350 000 offres potentiellement liées à l'intelligence artificielle. Ce filtrage initial est indispensable pour réduire le champ des offres à analyser et garantir la cohérence des données soumises au modèle de classification.

Le cœur de la procédure repose par la suite sur une classification via un modèle d'apprentissage profond de type BERT, adapté au texte en français et comportant 150 millions de paramètres. Ce modèle a été entraîné pour distinguer automatiquement les offres d'emploi selon deux axes majeurs : la technologie employée (par exemple IA générale, IA générative, vision par ordinateur, données) et la catégorie de poste (management, support, ingénieur ou fonctions non techniques). Cette seconde étape permet d'identifier environ 150 000 offres directement reliées au développement d'une technologie d'IA.

Le jeu de données d'entraînement comprend un total de 750 offres annotées manuellement, auxquelles s'ajoutent 300 offres destinées à l'évaluation des performances du modèle. Ces données d'entraînement ont été obtenues pour moitié par une sélection aléatoire d'environ parmi les offres préfiltrées, les données restantes ayant été filtrées

TABLE A5 – Thèmes identifiés par clustering dans les postes IA générative

Cluster	Titre suggéré	Mots-clés dominants
0	Ingénieur IA	ingénieur, ia, ingénieur ia, intelligence_artificielle
1	IA générative	ia générative, ia, générative, ingénieur
2	Développeur Python / IA	développeur, python, développeur python, développeur ia
3	Chef de projet innovation	chef projet, projet, chef, projet innovation
4	AI Engineer (en anglais)	engineer, intelligence_artificielle, data, ia
5	Data Scientist	data scientist, scientist, data, scientist ia
6	Consultant / Analyste data	consultant, analyst, data analyst, data
7	Lead data / santé	lead data, lead, scientist données, données santé
8	IA/MLOps	leader, mlops, machine_learning, ia générative
9	Analyste-développeur IA	analyste, analyste développeur, développeur, développeur cognitive

sur la base de mots-clés spécifiques tels que « generative AI », « LLM », « PyTorch », et d'autres termes pertinents, afin d'enrichir le corpus avec des exemples représentatifs des nouvelles technologies d'IA.

Les tableaux ci-dessous présentent la répartition des offres dans le jeu d'entraînement selon deux dimensions complémentaires. Le tableau B1 classe les offres selon la catégorie professionnelle : ingénieur, fonctions associées, management, et support et par technologie dominante, couvrant notamment la vision par ordinateur, les sciences des données hors machine-learning, l'IA générale, l'IA générative, ainsi que des postes relevant du développement informatique et d'autres non liées au numérique afin de servir de contre-exemple.

TABLE B1 – Répartition des offres d'emploi en IA dans le jeu d'entraînement

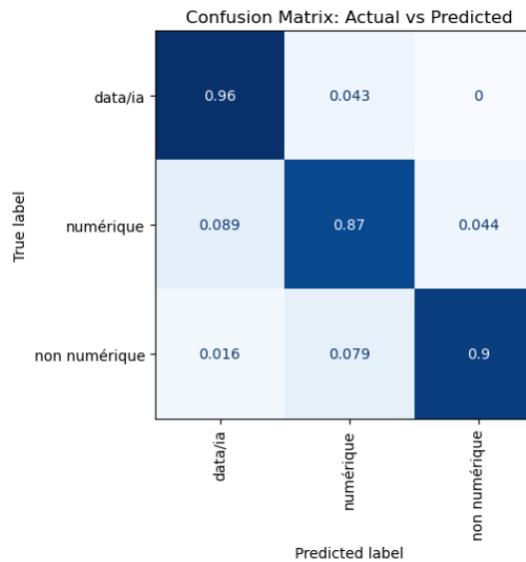
	Par catégorie de poste				Par technologie			
	Ingénieur	Fonctions connexes	Management	Support	Vision par ordinateur	Data	IA générale	IA générative
Nombre d'offres	246	89	131	245	102	173	221	215

Note : Le jeu d'entraînement comporte des annonces annotées manuellement pour l'entraînement du modèle de classification. Certaines annonces sont associées à plusieurs catégories.

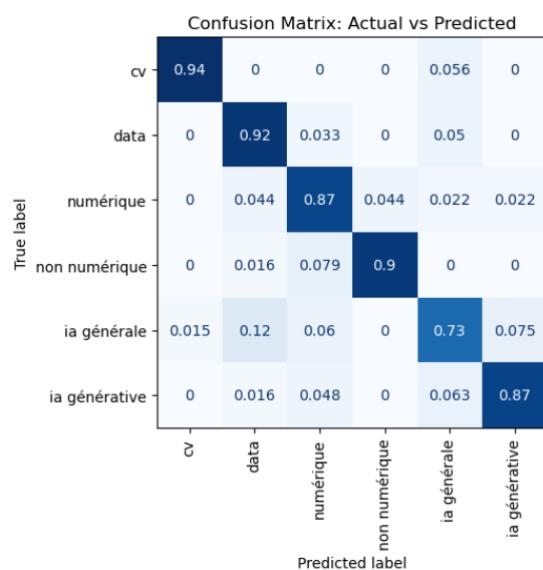
La combinaison de ce pré-filtrage par mots-clés et de la classification fine réalisée par le modèle BERT permet de disposer d'un outil performant pour segmenter précisément

le marché de l'emploi en IA. Les figures B1a à B1c présentent les performances du modèle sur les différentes tâches de classification (identification des offres IA, détection de la technologie et du type d'emploi), sous la forme de matrices de confusion. Ces matrices permettent de visualiser, pour chaque classe, le nombre de prédictions correctes et incorrectes, en comparant les classes prédites aux classes réelles..

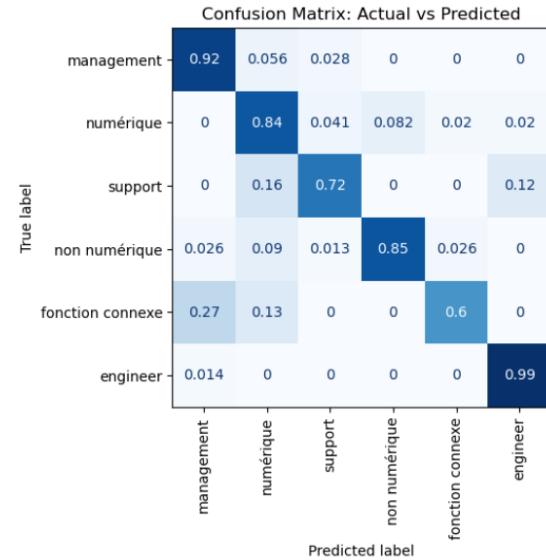
FIGURE B1 – Matrices de confusion pour les trois tâches de classification du modèle



(a) Détection des offres IA



(b) Type de technologie IA



(c) Type d'emploi IA

C Tests de robustesse

Le tableau B2 présente les résultats de régressions estimant l'effet de l'arrivée de ChatGPT sur le recrutement en IA générative, tel que discuté dans la section 4. Plusieurs modèles sont mobilisés à des fins de robustesse : modèles linéaires (OLS), modèles non linéaires (probit, logit) et un modèle de Poisson. Les coefficients reportés correspondent à l'effet moyen estimé du traitement, avec les erreurs standards entre parenthèses.

TABLE B2 – Effet estimé du traitement selon différents modèles

Modèle	Effet estimé (erreur standard)
Probit	0.06** (0.022)
Logit	0.046* (0.018)
OLS (binaire)	0.004*** (0.0006)
OLS (continu)	0.01*** (0.002)
Poisson	0.99* (0.44)

Notes : Chaque cellule présente le coefficient estimé de l'effet du traitement, suivi de l'erreur standard entre parenthèses. Significativité : * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.