

Exposition à l'intelligence artificielle générative et emploi : une application à la classification socio-professionnelle française

Antonin Bergeaud

Janvier 2024

L'ouverture progressive au public du générateur d'image *Dalle-2* au cours de l'année 2022 suivi de près par le lancement de *ChatGPT* en novembre de la même année ont considérablement relancé et amplifié les débats, pourtant déjà anciens, sur les effets potentiellement massifs que pourraient avoir l'intelligence artificielle (IA) sur l'emploi. Bien que des progrès remarquables aient été observés auparavant avec l'émergence de modèles d'IA capables de réalisations de plus en plus sophistiquées,¹ leur utilisation était limitée à un cercle relativement restreint de spécialiste. Ce qui distingue ChatGPT, c'est sa facilité d'utilisation, son coût d'accès quasiment nul et sa diffusion fulgurante². En outre, *ChatGPT*, ainsi que ses concurrents tels que *Midjourney*, *Bard*, *Claude* ou *Perplexity*, ont mis en lumière les capacités génératives impressionnantes de ces IA, tant dans le domaine du texte que dans celui des illustrations. Cette mise en évidence s'accompagne d'une amélioration fulgurante en termes de qualité et de performances, soulignant le potentiel considérable de ces technologies, y compris dans divers secteurs créatifs et à des niveaux professionnels. Dès lors il est apparu que de nombreuses tâches, y compris celles complexes et non répétitives jusqu'alors à l'abri de l'automatisation (Autor, 2015), pourraient être rapidement et avantageusement prises en charge par l'IA. Cela inclut également des tâches traditionnellement réservées à des travailleurs hautement qualifiés, démontrant ainsi l'étendue et la profondeur de l'impact potentiel de ces technologies sur le marché du travail.

¹ L'apparition des *transformers*, une architecture de *deep learning* en 2017 a permis aux langages de modèle de réaliser d'importants progrès

² OpenAI, l'entreprise développant ChatGPT a rapporté un record de diffusion avec un million d'utilisateurs enregistrés seulement cinq jours après le lancement public de ChatGPT, en outre depuis mars 2023, ChatGPT continue de connaître un succès massif, attirant environ 1.5 milliards de visites mensuelles

Ce que dit la littérature économique

La littérature économique s'est longuement penchée sur l'évaluation et la prévision de l'impact des technologies à haut potentiel d'automatisation, comme la robotique, le numérique et, plus récemment, l'intelligence artificielle, sur l'économie dans son ensemble et, en particulier, sur le marché du travail. En 2013, une étude marquante de Carl Benedikt Frey et Michael Osborne (publiée en 2017 : [Frey et Osborne, 2017](#)) a estimé que près de la moitié des emplois aux États-Unis étaient menacés à moyen terme par les technologies numériques. Cette analyse détaillée confrontait les capacités de l'informatique aux caractéristiques et compétences requises dans divers métiers. Néanmoins, cette perspective quelque peu alarmiste a été tempérée par des études ultérieures. [Arntz et al. \(2016\)](#), par exemple, ont souligné que se concentrer uniquement sur les professions néglige la diversité des tâches effectuées quotidiennement par un individu. Les recherches de David Autor, notamment dans [Autor et al. \(2003\)](#) et [Acemoglu et Autor \(2011\)](#), ont révélé une hétérogénéité significative dans la vulnérabilité à l'automatisation des tâches au sein d'une même profession. Certaines activités routinières peuvent effectivement être automatisées, tandis que d'autres, nécessitant des compétences sociales ou plus complexes, pourraient gagner en importance. Ainsi, l'automatisation pourrait transformer la manière dont une profession évolue, valorisant potentiellement certaines compétences dites « *soft* », comme le montrent [Aghion et al. \(2023\)](#). Toutefois, la disparition complète d'une profession n'est pas inéluctable, sauf si toutes ses tâches sont aisément automatisables. [Arntz et al. \(2016\)](#) estiment que seulement 9% des métiers sont menacés de disparition dans les pays de l'OCDE, et non 47%.

La question devient encore plus complexe lorsqu'on considère les effets d'équilibre général (voir [Aghion et al., 2022a](#) pour une discussion). L'impact de ces technologies sur les travailleurs variera selon que la technologie agit comme un complément ou un substitut à certaines tâches. Par exemple, un journaliste pourrait déléguer certaines de ses tâches à un modèle de langage tel que ChatGPT, mais utiliser cet outil pour accroître sa productivité dans d'autres domaines, comme la rédaction d'articles de fond. Une productivité accrue peut avoir des effets positifs à l'échelle globale et, in fine, augmenter l'emploi. Cette approche, connue sous le nom de « *task-based model* », a été notamment adoptée par [Acemoglu et Restrepo \(2018, 2020\)](#) pour évaluer l'effet des robots sur l'emploi, en prenant en compte ces effets d'équilibre général. Cependant, cette question compliquée

peut conduire à des conclusions très variées concernant les effets agrégés, comme l'illustrent les résultats d'Aghion et al. (2022b).

En définitive, bien que ces études se concentrent sur des technologies existantes telles que les robots industriels ou les ordinateurs, dont les potentiels de substitution se limitent principalement aux tâches répétitives et peu créatives, elles procèdent généralement de la manière suivante : elles classent les différentes tâches selon le risque de substitution ou de complémentarité avec la technologie envisagée, en se basant sur une description détaillée de celles-ci. Cependant, cette méthode s'avère complexe à appliquer dans le cas de l'IA générative, tant ses potentialités sont difficiles à cerner et ses développements futurs à anticiper à court ou moyen terme.

Plusieurs études ont proposé des méthodes pour évaluer l'exposition des différents métiers à l'intelligence artificielle comme par exemple, Felten et al. (2021, 2023), Gmyrek et al. (2023), Webb (2020), Eloundou et al. (2023) et Pizzinelli et al. (2023). Webb (2020) a comparé les descriptions de métiers aux textes de brevets dans le domaine de l'IA, suggérant qu'une similarité dans le contenu pourrait indiquer un risque d'automatisation. Eloundou et al. (2023) se sont concentrés sur l'impact du modèle de langage GPT, estimant qu'un métier est exposé si la technologie peut réduire significativement le temps nécessaire pour certaines tâches. Felten et al. (2021) ont créé un indicateur, l'AI Occupational Exposure (AIOE), utilisé également par Pizzinelli et al. (2023), qui évalue l'exposition à l'IA en se basant sur les capacités de cette dernière. Gmyrek et al. (2023) ont employé GPT-4 pour prédire les tâches typiques de divers métiers et évaluer leur degré d'automatisation. Pizzinelli et al. (2023) ont combiné différentes mesures pour proposer une évaluation à deux dimensions : l'exposition AIOE et une mesure de complémentarité avec l'IA.

Il est important de noter que ces travaux partagent une limite importante mais difficile à contourner : la faisabilité de l'implémentation de la technologie. Même si l'IA peut techniquement remplacer de nombreuses tâches avec une plus grande précision et efficacité qu'un humain, cela ne garantit pas que les entreprises décident de l'adopter, en raison des coûts d'installation, de maintenance et d'utilisation qui peuvent être supérieurs au salaire d'un travailleur. Estimer ces coûts nécessite une analyse détaillée de la technologie et du secteur concerné. Svanberg et al. (2024) ont adopté cette approche pour les technologies de vision par ordinateur, concluant que leur implémentation n'est vraiment avantageuse que pour les très grandes entreprises, représentant une faible part de l'emploi

total. Selon eux, même avec une baisse significative des coûts, cette situation persistera probablement à l'avenir.³ Une autre limite de ces modèles est qu'il est difficile d'anticiper la manière dont les différents travailleurs vont évoluer avec l'IA et développer de nouvelles formes de complémentarité comme cela a pu être le cas avec les technologies passées comme l'ordinateur.

Approche retenue

Dans cette courte note, nous adoptons une démarche similaire à celle de [Gmyrek et al. \(2023\)](#) en nous basant sur les données de la base O*NET pour identifier les diverses tâches liées à 222 professions selon la nomenclature FAP (FAMILLES Professionnelles) de la DARES⁴. Ces tâches sont ensuite évaluées selon deux indices d'exposition à l'intelligence artificielle générative : le premier indice mesure la substituabilité (score1), tandis que le second évalue la complémentarité et le potentiel d'amélioration (score2).

Plus précisément, notre analyse part de la table « *Task Rating* » de la version 28.1 de la base de données O*NET, qui recense 16 937 tâches associées à 873 métiers⁵, avec une moyenne de 21 tâches par métier et leur degré d'importance pour chaque profession. Ces tâches ont été recensées grâce à des entretiens avec des professionnels du secteur ou des experts du marché du travail. Par exemple, la tâche suivante :

“Consult with product development personnel on product specifications, such as design, color, or packaging”

Est notée avec une importance de 3.65 sur 5 pour un directeur marketing. Pour chaque tâche, nous avons appliqué le modèle GPT-4 en utilisant une instruction spécifique (ou prompt) présentée dans le Tableau 1. L'objectif de cette démarche est de fournir au modèle de langage une compréhension des limites de l'IA générative et de ses capacités, et d'attribuer à chaque tâche un

³ L'étude ne prend toutefois pas en compte les gains de productivité qu'une telle technologie pourrait permettre d'obtenir et qui pourraient compenser son coût important.

⁴ Notre approche peut être étendue aux quelques 470 professions de la classification PCS (Professions et Catégories Socioprofessionnelles) de l'INSEE qui est donc plus fine. Toutefois la FAP possède l'avantage de proposer une catégorisation plus homogène avec des métiers plus distinguables

⁵ Comme expliqué ci-après, la classification utilisée par O*NET diffère de la classification française PCS et contient un niveau de détail un peu plus fin.

score, allant de 0 à 1, pour mesurer le potentiel de substituabilité, c'est-à-dire d'évaluer la probabilité que la tâche puisse être réalisée par l'IA dans un futur proche.

TABLE 1 : PROMPT UTILISE POUR NOTER LES TACHES AVEC LE MODELE GPT-4

<i>AI Capabilities Overview:</i> <i>Consider the full range of what Generative AI can currently achieve and what it is projected to do efficiently over a reasonable time horizon. This includes advancements in automation, data analysis, natural language processing, creative generation, and other emerging technologies. Take also into consideration what AI cannot do efficiently and the limitation of other related technologies such as robotics.</i> <i>Task Assessment:</i> <i>Based on AI Capabilities described above, analyze the task '{description}' in the following dimension:</i> <i>Substitution Potential (score1): Assess the extent to which AI and generative AI could potentially replace human labor in this task.</i> <i>Consider factors like technological feasibility, complexity, ethical considerations, and practical limitations.</i> <i>Instructions:</i> <i>Provide a score ranging from 0 to 1 with 2 decimals. No explanations are required. Only the scores should be given, representing the substitution potential of AI for the task respectively.</i> <i>Score:</i>

A noter que ce prompt considère non seulement les capacités génératives de l'IA mais également les potentiels liés à l'automatisation en général. Une tâche physique et relativement répétitive se verra donc attribuer un score élevé au même titre qu'une tâche nécessitant la génération de contenu relativement standardisé.

Avant d'aborder les résultats, il est important de souligner trois points clés. Premièrement, les données issues d'ONET reposent sur une classification professionnelle spécifique aux États-Unis, ainsi que sur des caractéristiques élaborées à partir de la réalité professionnelle américaine. Pour extrapoler ces résultats au contexte français, nous partons du principe que ces tâches sont transposables, c'est-à-dire qu'un même métier présentera une composition similaire de tâches en France et aux États-Unis, ou du moins que leur importance relative sera comparable. En outre, nous avons élaboré une table de correspondance pour convertir la classification d'ONET (SOC) en classification professionnelle et sociale (PCS) française, puis en classification FAP. Pour ce faire, nous utilisons la table de conversion du Bureau of Labor Statistics (BLS) permettant de passer de SOC à ISCO, la classification internationale de l'emploi établie par le Bureau International du Travail, puis nous appliquons les données de [Hensvik et al. \(2020\)](#) pour transposer ISCO en PCS. Chacune de ces étapes comporte malheureusement une part d'arbitraire inévitable, tant que des données comparables à celles d'O*NET ne seront pas disponibles en France. Deuxièmement, il est important de noter qu'aussi précise soit la description des tâches incluses dans la base O*NET, celle-ci ne peut pas rendre compte de l'importante diversité des réalités pour deux individus effectuant le même métier dans deux contextes différents. Il faut ainsi interpréter le résultat obtenu comme une mesure de l'exposition à l'IA générative pour un travailleur représentatif de la profession considérée. Enfin, l'exercice ne rend pas compte des éventuelles transformations endogènes des différentes tâches à la suite de l'arrivée d'une technologie comme l'IA ni plus globalement de l'apparition de nouvelles tâches, voire de nouveaux métiers.

Résultats

La procédure précédente permet d'attribuer à chaque tâche $t \in T$ une valeur $s_1(t, P)$ comprise entre 0 et 1 et qui dépend a priori du *prompt* P . Nous connaissant en outre grâce à la base O*NET l'importance relative, notée $p(t)$ de la tâche t pour un métier donné p . Nous pouvons ainsi agréger nos deux mesures au niveau de chacune des 222 familles d'emploi de la manière suivante :

$$s_1(p, P) = \sum_{t \in T} \frac{s_1(t, P)p(t)}{\sum p(t')}$$

En d'autres termes, nous évaluons la moyenne des scores individuels de chaque tâche, ajustée selon l'importance de ces tâches. Ce score est normalisé pour s'étendre entre 0 et 1, où la valeur 1 est attribuée à la profession la plus exposée à l'intelligence artificielle, et la valeur 0 à celle la moins exposée. Cet indicateur résume donc la vulnérabilité d'un emploi face à l'IA. Cependant, cette exposition peut être modulée, soit atténuée soit exacerbée, par deux mécanismes. D'une part, comme mentionné précédemment, une forte exposition à l'IA peut aussi indiquer un potentiel élevé de complémentarité, résultant en une collaboration entre l'humain et la nouvelle technologie, ce qui peut significativement améliorer les performances et la qualité du travail accompli. D'autre part, certaines professions peuvent être intrinsèquement plus exposées à l'IA que ce que les analyses basées sur les tâches suggèrent. Par exemple, un photographe pourrait voir plusieurs de ses tâches affectées par l'IA générative, mais des aspects tels que la mise en place du décor ou l'ajustement de l'éclairage, qui sont essentiels mais moins susceptibles d'être influencés par l'IA, ne sont pas au cœur de son activité. Or l'indicateur s_1 pourrait attribuer une valeur moyenne relativement faible en raison de l'importance que prennent ses tâches en termes de temps. Pour intégrer ces deux aspects, nous enrichissons notre indicateur avec l'indice agrégé C-AIOE, issu de [Pizzinelli et al. \(2023\)](#), calculé au niveau ISCO, qui associe un indice de complémentarité et un indice de substituabilité, tous deux basés sur les caractéristiques de la profession plutôt que sur les tâches spécifiques. Pour obtenir un indicateur global sans présupposer une forme fonctionnelle spécifique, nous multiplions simplement ces deux valeurs. Ainsi, une profession est considérée comme hautement exposée si elle a à la fois un score s_1 et un niveau C-AIOE élevé.

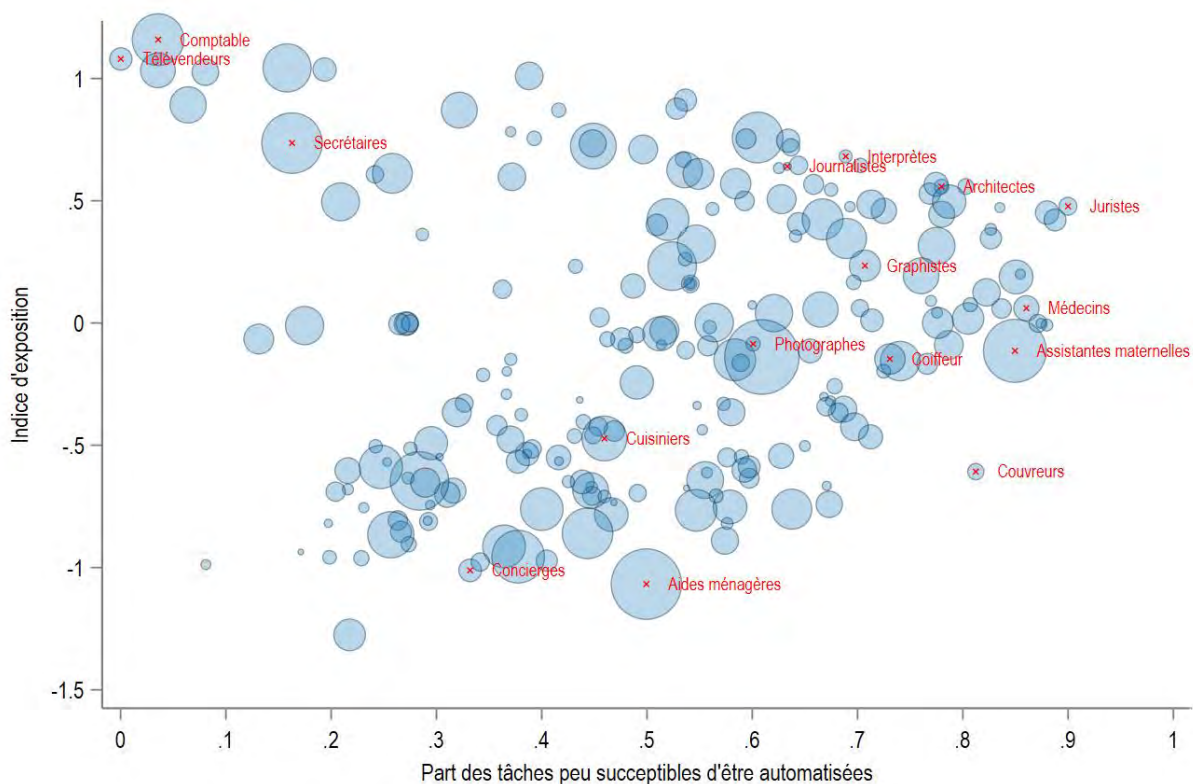
Cette première dimension reflète donc l'exposition théorique d'un métier à l'IA. Toutefois il est possible que cette exposition soit fortement freinée par un coût trop élevé à adopter l'IA. Collecter ou estimer des informations sur les prix est une opération particulièrement difficile comme discuté par [Svanberg et al. \(2024\)](#). Pour approcher cette idée, nous calculons pour chaque emploi une approximation de la part des tâches jugée difficile à automatiser que nous définissons comme les tâches ayant une valeur de $s_1(t, P)$ inférieure à 0.75 (correspondant à la médiane). Ainsi si un métier est par nature exposé à l'IA mais qu'il demeure une part importante de ces tâches qui ne sont pas facilement automatisable, il est probable que cette profession ne soit pas menacée par l'arrivée de ces nouvelles technologies, même si une partie de ces activités va vraisemblablement être bouleversée.

La Figure 1 dispose chaque FAP dans un graphique qui combine nos deux indicateurs. En ordonnées la mesure synthétique d'exposition obtenue en multipliant s_1 à C-AIOE et en abscisse

la part des tâches difficile à automatiser. Chaque cercle symbolise une profession, avec sa taille reflétant l'effectif total dans cette profession. Certaines professions illustratives ont été représentées en rouge.

Les métiers localisés dans le quadrant supérieur droit sont ceux qui font face à une forte exposition à l'IA, mais pour lesquelles il existe encore d'importants freins à l'automatisation et dont on peut donc attendre qu'ils subissent de fortes transformations qui pourraient conduire à un effet "superstar". Cela pourrait bénéficier à certains travailleurs mieux formés et préparés, capables d'exploiter les nouvelles technologies génératives, tandis que d'autres pourraient être confrontés à une concurrence accrue. Les professions situées dans le quadrant supérieur gauche sont particulièrement vulnérables face à l'IA. Elles se caractérisent par une exposition supérieure à la moyenne et à une part importante de leurs tâches qui pourraient se voir remplacer.

FIGURE 1 : POSITIONNEMENT DE CHAQUE FAP DANS UN REPERE EXPOSITION (AXE DES ORDONNEES) ET PART DES TÂCHES DIFFICILE A AUTOMATISER (AXE DES ABSCISSES). LA TAILLE DE CHAQUE CERCLE EST PROPORTIONNELLE A L'EMPLOI TOTAL DANS LA FAP EN FRANCE EN 2017

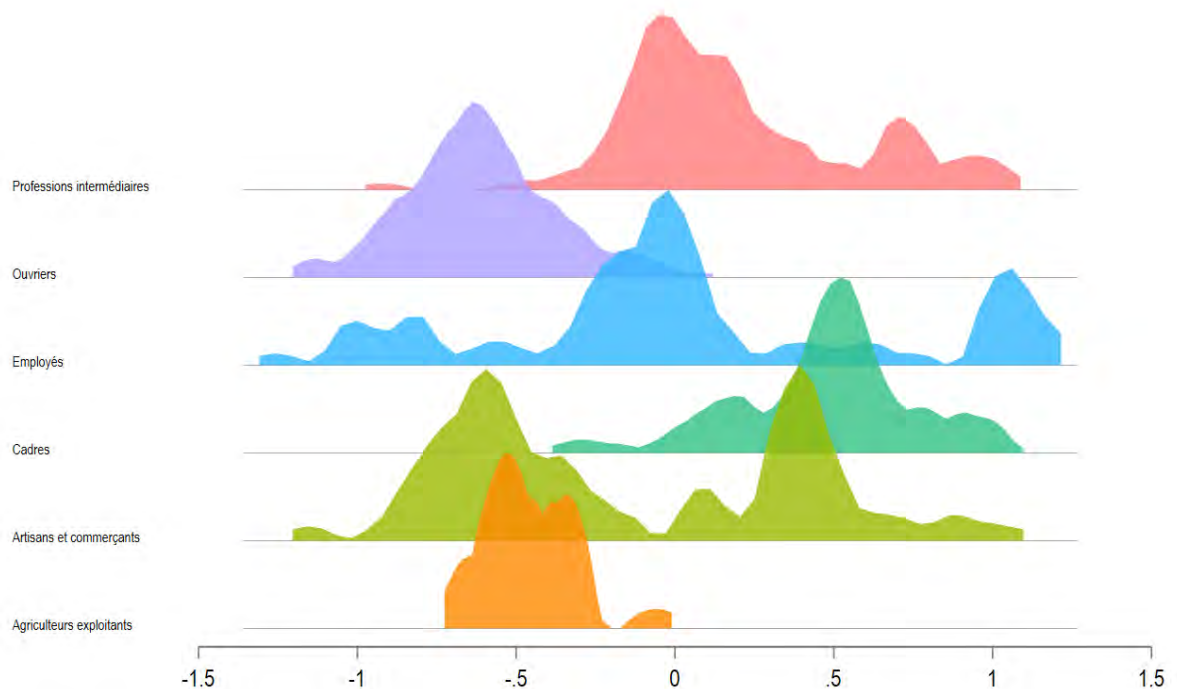


Les professions situées dans la moitié basse, et en particulier dans le cadran inférieur droit sont à ce jour encore préservées par l'arrivée de l'IA.

Pour interpréter la magnitude de l'indicateur d'exposition présenté en ordonnées de la Figure 1, il faut lire qu'une valeur négative correspond à un métier dont le niveau d'exposition CAIOE est inférieur à la moyenne. La valeur de CAIOE a été calculé de sorte qu'une valeur de 1 soit considérée comme particulièrement grande. Notre indicateur multiplie CAIOE par le score s_1 qui est lui-même compris entre 0 et 1. Aussi, même s'il n'existe pas de valeur seuil claire, nous pouvons considérer qu'une valeur supérieure à 0.5 correspond à une exposition forte.

La Figure 2 présente la distribution de l'indicateur d'exposition (l'ordonnée du graphique de la Figure 1) en fonction du niveau 1 de la PCS, c'est-à-dire des grandes catégories d'emploi. On peut y constater qu'il existe une hétérogénéité importante entre les différents métiers et que contrairement aux révolutions technologiques précédentes, ce sont les cadres c'est-à-dire les professions les plus qualifiées qui sont le plus exposées.

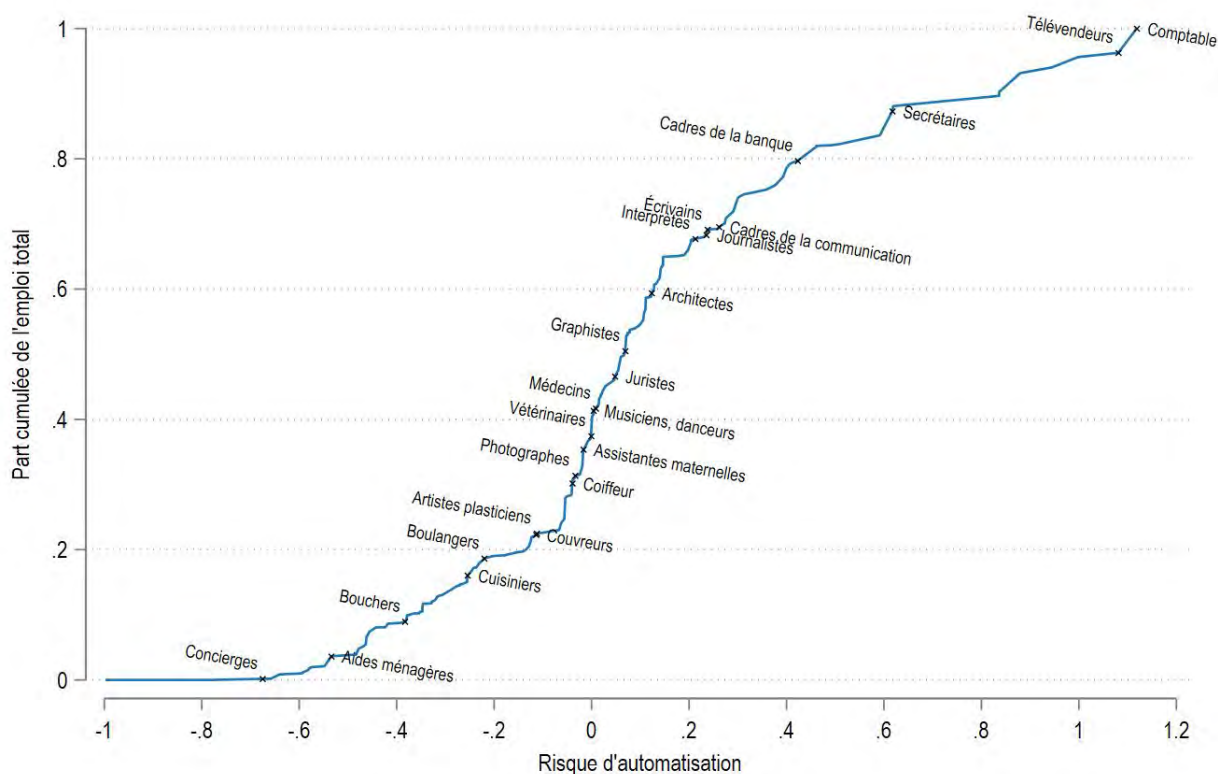
FIGURE 2: DISTRIBUTION DE L'INDICATEUR POUR DES GRANDES CATEGORIES DE TRAVAILLEURS



Enfin la Figure 3 présente la distribution cumulative de l'emploi en fonction d'un indicateur synthétique de risque d'automatisation définie comme le produit de l'indicateur d'exposition en

ordonnée de la Figure 1 et de la part des tâches qui ne sont pas difficile à automatiser. On peut notamment y constater que 20% de l'emploi environ possède un indicateur agrégé supérieur à 0.5 ce qui signifie qu'elles ont à la fois une exposition forte, voir très forte et sont en même temps peu protégé de l'automatisation par la prévalence de certaines tâches difficile à effectuer par une machine, même intelligente.

FIGURE 3: DISTRIBUTION CUMULATIVE DE L'EMPLOI EN FONCTION DU RISQUE D'AUTOMATISATION PAR L'IA



Conclusion

Cette courte note n'a pas pour prétention de donner une idée précise du nombre de métiers qui sont à risque d'être automatisé avec l'arrivée de l'IA, mais de permettre de mesurer l'importante hétérogénéité des niveaux possible d'exposition et l'impact important que cette technologie a et va continuer d'avoir sur certaines professions y compris les plus qualifiées. Une approche plus complète nécessiterait toutefois de prendre en compte l'intégralité de la chaîne de valeur associée à un secteur et un métier et d'estimer les effets de réallocation potentiellement importants pouvant conduire à une forte augmentation de l'emploi dans certains métiers. Enfin une extension utile à

cette note pourrait être d'évaluer la robustesse de l'indicateur $s_1(P)$ en fonction du *prompt* P en modifiant la définition de ce qu'est et de ce que peut et ne peut pas faire une IA.

Bibliographie

Acemoglu, D., & Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American economic review*, 108(6), 1488-1542.

Acemoglu, D., & Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets. *Journal of political economy*, 128(6), 2188-2244.

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics* (Vol. 4, pp. 1043-1171). Elsevier.

Aghion, P., Antonin, C., Bunel, S., & Jaravel, X. (2022a). The effects of automation on labor demand: A survey of the recent literature. CEPR Discussion Paper #16868

Aghion, P., Antonin, C., Bunel, S., & Jaravel, X. (2022b). What are the labor and product market effects of automation? New evidence from France. *Unpublished manuscript*

Aghion, P., Bergeaud, A., Blundell, R. W., & Griffith, R. (2023). Social skills and the individual wage growth of less educated workers. CEPR Discussion Paper #18456

Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of economic perspectives*, 29(3), 3-30.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, 118(4), 1279-1333.

Arntz, M., Gregory, T., & Zierahn, U. (2016). The risk of automation for jobs in OECD countries: A comparative analysis. OECD working paper

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023). Gpts are gpts: An early look at the labor market impact potential of large language models. *arXiv preprint arXiv:2303.10130*.

Felten, E., Raj, M., & Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, 42(12), 2195-2217.

Felten, E., Raj, M., & Seamans, R. (2023). How will Language Modelers like ChatGPT Affect Occupations and Industries?. *arXiv preprint arXiv:2303.01157*.

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation?. *Technological forecasting and social change*, 114, 254-280.

Gmyrek, P., Berg, J., & Bescond, D. (2023). Generative AI and Jobs: A global analysis of potential effects on job quantity and quality. *ILO Working Paper*, 96.

Hensvik, L., Le Barbanchon, T., & Rathelot, R. (2020). Which jobs are done from home? Evidence from the American Time Use Survey. IZA Discussion Paper No. 13138

Pizzinelli, C., Panton, A., Mendes Tavares, M., Cazzaniga, M., & Li, L. (2023). Labor Market Exposure to AI: Cross-country Differences and Distributional Implications.

Svanberg, M., Li, W., Fleming, M., Goehring, B. & Thompson, N. (2024). Beyond AI Exposure: Which Tasks are Cost-Effective to Automate with Computer Vision? Working paper FutureTech MIT

Webb, M. (2020). The impact of artificial intelligence on the labor market. *Unpublished manuscript*

