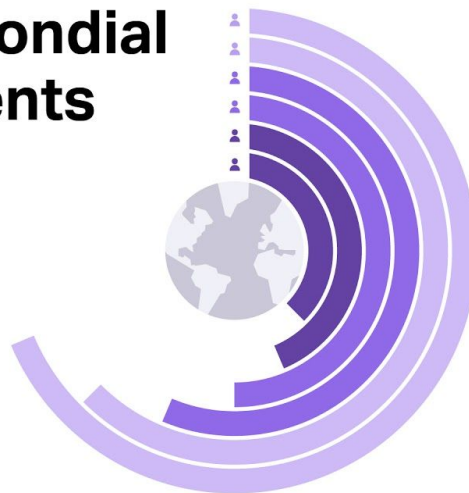


Rapport mondial sur les talents 2020



jfgagne

Le rapport se lit plus facilement [en ligne](#) pour visualiser les tableaux et graphiques. Si vous êtes sur un appareil mobile, lisez les [faits saillants ici](#) et une [définition des rôles techniques les plus importants en matière d'IA ici](#).

Structure du rapport

Considérant que ce rapport porte sur de nouveaux paramètres et un champ d'application beaucoup plus large que par les années précédentes, nous nous sommes attardés à présenter les données avec les points forts des tendances et des valeurs aberrantes (« outliers ») pour vous aider à commencer votre exploration de cette information. L'analyse exige une compréhension fine des enjeux sur le terrain que nous viserons à intégrer en complément par des articles et des collaborations ultérieurs. D'autres avenues d'analyse quantitative sont également exposées dans la conclusion pour nous - ainsi qu'à d'autres - permettre de pousser la réflexion encore plus loin.

Ce rapport est divisé en quatre sections principales. La première section offre une introduction et un résumé avec le classement des pays les plus performants en fonction des totaux de leurs bassins de talents. La seconde explique les différents rôles techniques spécialisés tout au long de la chaîne de valeur de la conversion en produit (« productisation » de l'IA, de la recherche jusqu'au déploiement). La troisième section explore les données arXiv, une plateforme de prépublication qui nous offre le meilleur moyen de se rapprocher d'un recensement de la recherche en IA. Nous y examinerons la localisation, les mouvements et le genre des talents. La quatrième section est celle où nous examinons nos premières estimations de l'offre et de la demande dans l'industrie pour tous les différents rôles abordés dans la première section. Nous aborderons la méthodologie tout au long de ce document, et avons inclus une annexe regroupant l'ensemble des notes méthodologiques, de même qu'une autre annexe présentant des voies futures de recherche pour ajouter de la granularité aux données présentées.

INTRODUCTION

Dans un récent article de la série « [Technology Quarterly](#) », *The Economist* avait sous-titré : « Après des années d'engouement, l'IA déçoit les attentes ». L'angle adopté abordait deux enjeux fondamentaux : l'accès aux données requises et le fait que les algorithmes n'étaient pas très intelligents, pour le moment. Cela n'est pas surprenant, car une grande partie des communications et contenus médiatiques portant sur l'intelligence artificielle (IA) et l'apprentissage machine (AM) promettait des algorithmes magiques en échange des bonnes données.

Cette équation ne se vérifie cependant qu'en partie, puisqu'il existe bel et bien un potentiel inexploité de l'IA qui peut se réaliser avec les données et les algorithmes actuellement disponibles – mais qui demande que l'on dispose en quantité suffisante d'une autre ressource tout aussi rare pour l'appliquer : le talent. Nos précédentes estimations des viviers de talents se sont concentrées sur les chercheurs en IA – ceux qui rendent les algorithmes plus intelligents et travaillent sur des ensembles de données toujours plus importants – et ce biais d'analyse a peut-être ajouté à l'illusion que c'était tout ce dont nous avions besoin.

Amener l'IA de la recherche vers des retombées dans le monde réel crée une longue chaîne de valeur qui dépend d'un ensemble de compétences et d'expériences. Il est courant de voir des personnes qui peuvent, et qui occupent de multiples rôles le long de la chaîne de valeur pour pallier à la rareté de ces expertises. Malgré ce chevauchement, nous croyons utile de classer et d'explorer ces rôles séparément pour mieux comprendre les outils nécessaires pour construire et faire fonctionner les solutions d'IA¹, et reconnaître les facettes de l'accessibilité à ce talent.

SOMMAIRE DÉCISIONNEL

De nouveaux paramètres, et une vision plus large, pour la mesure du talent

Cette année, avons ajouté des estimations de plusieurs autres rôles techniques spécialisés qui sont critiques dans la chaîne de valeur du développement d'un produit d'IA : ingénierie en AM, mise en œuvre technique et architecture des données. Nous avons estimé la taille du bassin de talents disponibles pour l'industrie en nous basant sur des données autodéclarées sur les médias sociaux, et la demande, grâce aux total mensuel des offres d'emploi pour les mêmes fonctions.

Nous avons également élargi notre observation du bassin de talents en IA pour inclure non seulement les chercheurs fondamentaux, mais aussi ceux et celles qui s'adonnent à du travail appliqué.

Pour ce faire, nous sommes passés de l'utilisation des conférences comme proxy, dont les places limitées ne permettent pas de saisir la croissance complète de l'écosystème, aux dépôts d'articles dans les domaines de l'IA et de l'AM sur arXiv (cs.AI, cs.LG et stat.AM). ArXiv se présente comme le lieu où les chercheurs prépublient leurs articles (c'est-à-dire avant l'examen par les pairs pour acceptation en vue d'une publication ou d'une conférence²) et c'est peut-être ce qui se rapproche le plus d'un recensement de la recherche en IA. Il offre également une vision beaucoup plus large de la croissance de l'IA en incluant des articles sur les méthodes appliquées.

¹ Nos catégories mettent l'accent sur les aspects des rôles dans *la construction* d'une solution d'IA, par opposition à *son exécution*, bien que les compétences nécessaires puissent couvrir les deux.

² La prépublication en IA est très bien acceptée, car il y est facile de tester les méthodes et de voir si elles sont reproductibles et utiles. Il est vrai que nous n'avons pas fait le tri en fonction de la popularité comme effet significatif, mais seulement pour observer où en est le volume de la recherche.

En voici brièvement les points importants

Le nombre total d'auteurs publiant chaque année sur arXiv a augmenté en moyenne de 52,69 % par an depuis 2007. L'année dernière, 58 000 auteurs ont été recensés et nous estimons qu'ils seront 86 000 d'ici la fin de l'année. Parmi les quatre rôles sectoriels que nous avons examinés pour le bassin de talents, nous estimons qu'il y a environ 477 000 personnes (voir la section « Les rôles techniques spécialisés de la chaîne de valeur de l'IA »).

Les classements nationaux figurent dans les deux graphiques ci-dessous, et doivent être considérés avec une note importante sur la méthodologie : nous avons attribué un emplacement basé sur le lieu du siège social de l'organisation à laquelle l'auteur est affilié, afin de mettre l'accent sur l'endroit où la propriété intellectuelle (PI) est détenue. Ce choix donne un poids important à des pays comme les États-Unis, dont les entreprises de Big Tech disposent de laboratoires de recherche en IA dans le monde entier.

Comme nous l'avons fait par le passé, nous poursuivons avec plusieurs analyses quantitatives du vivier de talents de recherche sur arXiv (pour plus de détails, voir la section « Les personnes qui rendent l'IA plus intelligente »).

Le talent reste mondial et mobile (du moins avant la pandémie)

Notre analyse montre qu'au moins pour les talents qui contribuent à la recherche, le bassin est véritablement mondial. Les collaborations se poursuivent dans le monde entier, à l'exception des pays du Sud global, dont les écosystèmes sont beaucoup moins avancés dans leur développement. L'Irlande est bonne première avec plus de 15 collaborations moyennes par auteur, alors que la plupart des pays se situent entre 4 et 6.

La migration est peut-être un paramètre obsolète avec la pandémie et le travail à distance, mais elle permet toutefois de montrer quels pays (ou entreprises) ont la plus grande force d'attraction par rapport à d'autres qui se trouvent en position de prestataires, et d'autres encore qui montrent une certaine insularité avec relativement peu de mouvements des ressources humaines. Bien sûr, les États-Unis ont de loin la plus grande force d'attraction. Toutefois, cela pourrait se voir en partie inversé par les nouvelles politiques de visas, qui créent une occasion importante pour leurs voisins, mais les nombreux laboratoires internationaux des grandes entreprises technologiques américaines pourraient atténuer en partie ce phénomène.

L'équilibre entre les genres a peu progressé

Le genre dans arXiv est encore moins équilibré que ce que nous avons vu en observant les conférences, et la proportion est plutôt stable depuis 2007, passant seulement de 12 % à 15 % aujourd'hui.³ Les ratios varient sensiblement d'un pays à l'autre, [mais la pandémie a plus fortement affecté la production des femmes chercheuses](#) et cette année pourrait donc être un pas en arrière au final. Un autre facteur réside dans celles et ceux qui contribuent à la publication, ce que nos ensembles de données sur l'autoreprésentation aident à éclairer.

Peu de personnes qui travaillent dans l'industrie semblent faire de la recherche fondamentale à plein temps

Sur les médias sociaux liés à l'emploi, nous constatons qu'il n'y a qu'environ que 4 100 personnes qui se présentent comme des chercheurs professionnels pour l'industrie, ce qui indique peut-être qu'un certain nombre d'auteurs travaillent principalement comme ingénieur et font de la recherche à temps partiel. Nous avons également constaté que cela est anecdotiquement vrai parmi nos pairs. Bien que cela puisse s'expliquer par une préférence pour le travail appliqué, nous avons constaté que la proportion de la demande pour des chercheurs purs dans l'industrie est également faible (environ 1 % des offres d'emploi)

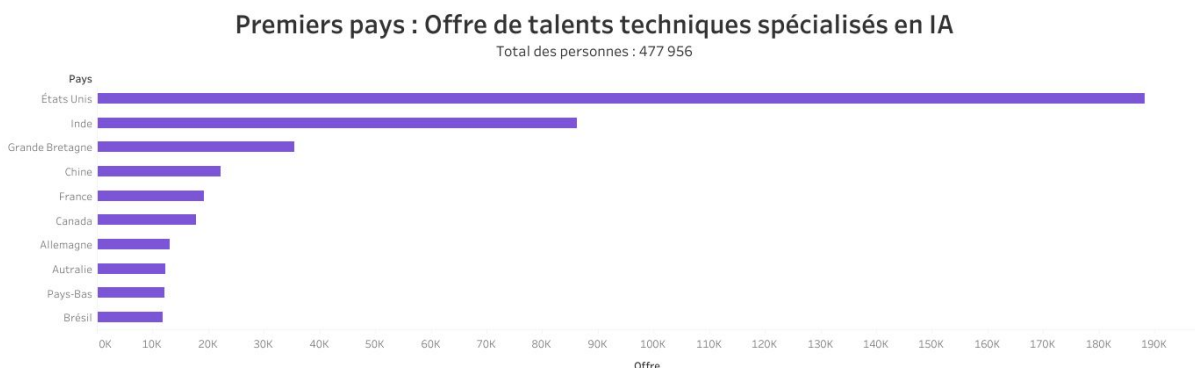
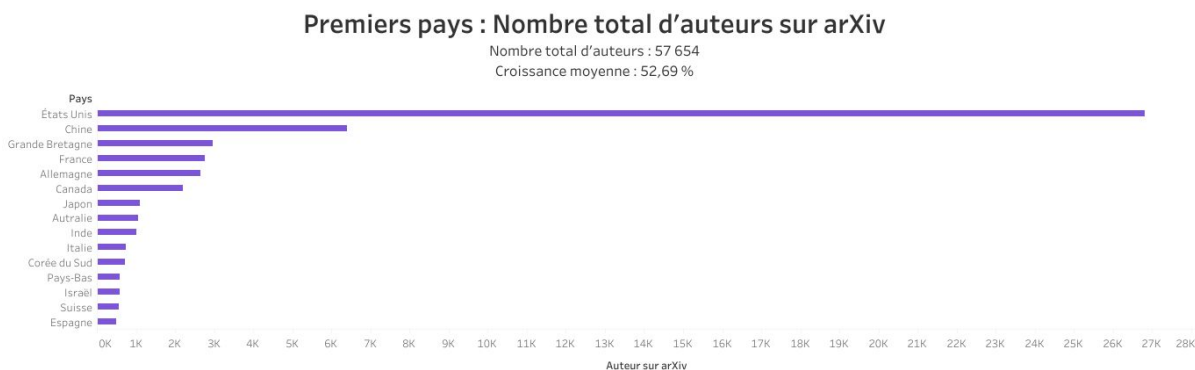
³ Pour les conférences, la situation n'a été que légèrement meilleure. Le rapport de l'année dernière montrait que la proportion d'articles écrits par des femmes publiés dans les conférences les plus importantes était de 18 %.

et que les travaux de recherche pure dans l'industrie ne sont pas largement disponibles.⁴ L'ajout de données sur les taux d'obtention de diplômes, les emplois dans les universités et les laboratoires de recherche privés, et la compréhension de la durée moyenne où un chercheur publie, ainsi que sa production, permettront de répondre plus clairement à la question qui se pose, à savoir comment l'industrie influence la production de la recherche.

Ce phénomène de partage du temps entre le travail appliqué et la recherche pourrait affecter davantage l'équilibre entre les genres dans les contributions à la recherche. [Certaines observations ont montré](#) que les femmes obtiennent moins souvent des emplois leur permettant de continuer à contribuer à la recherche. Avec le défi que représente l'obtention d'un emploi dans la recherche pure après l'obtention du diplôme, cela met encore plus de pression sur les cohortes entrantes pour qu'elles comprennent un nombre élevé de femmes afin d'affecter le ratio global des contributions des femmes à la recherche.

L'ouverture de nouveaux postes a été constante, mais a connu une forte baisse en 2020

Malheureusement, les données sur la demande sont limitées à ce stade, mais le suivi des offres d'emploi mensuelles nous permet d'en avoir une idée. Nous avons constaté que la proportion de rôles d'ingénierie spécialisée, de mise en œuvre technique et de recherche dans la demande correspond étroitement à la proportion de l'offre : environ 61 % pour les rôles de mise en œuvre qui développent le logiciel à partir des capacités de l'IA, 38 % pour les rôles d'ingénierie de l'IA qui développent les modules de base de l'IA, 1 % pour les chercheurs. Nous ne connaissons pas les chiffres globaux, mais les flux mensuels ont augmenté régulièrement en 2019, de 2 à 6 % environ pour les différents titres de poste. Il n'est pas surprenant que nous ayons également constaté une baisse de 20 à 30 % de la demande pour les titres de postes concernés en 2020, mais tant 2019 que 2020 affichent des valeurs aberrantes significatives pour les personnes qui entrent en scène et qui persistent pendant la pandémie.



⁴ Il est également possible que la majorité des ressources humaines restent dans la recherche en milieu universitaire et ne se créent tout simplement pas de profils d'emploi, et que les diplômes professionnels et la formation en ligne en IA représentent une part plus importante de l'offre, en partie parce qu'ils sont nettement plus rapides (brèves) que les filières de doctorat orientées vers la recherche.

Les rôles techniques spécialisés de la chaîne de valeur de l'IA

Les méthodes de construction des produits d'IA ont évolué et les défis se sont développés à mesure que nous nous sommes attaqués à de nouveaux problèmes, mais nous voyons maintenant clair en ce qui a trait aux types d'expertise standard nécessaires pour faire passer l'IA du stade des idées théoriques aux produits concrets.

Chez Element AI, nous bénéficions des 20 ans d'expérience des fondateurs dans le développement de produits d'IA, et de leur connexion avec l'écosystème au sens large pour la vérification et la validation de ces catégories. Bien qu'il existe des différences et des nuances dans les détails des rôles selon les entreprises (et les zones géographiques), la série suivante de titres et de descriptions de postes couvre les rôles techniques spécialisés qui constituent la chaîne de valeur du développement d'un produit d'IA :

- Chercheur
- Ingénieur IA/AM
- Productisation IA/AM
- Ingénierie et architecture des données

Nous développons ci-dessous leurs rôles et qualifications respectifs.

Recherche

Les chercheurs continuent à développer rapidement les capacités d'IA. La course aux points de bascule de l'efficacité ou de l'efficience de la technologie sous-jacente est si importante que de nombreux chercheurs professionnels travaillent à plein temps dans l'industrie, ou du moins y sont affiliés – même si la recherche est ouverte en tant que norme commune. Cette ouverture est importante, car elle permet la collaboration entre les organisations (universitaires ou industrielles) afin de mieux définir les agendas de recherche et de développer de nouvelles méthodes.

La recherche progresse rapidement, en partie grâce à la prépublication sur arXiv, qui permet d'évaluer et d'exploiter rapidement de nouvelles méthodes. Toutefois, les meilleurs articles qui seront acceptés pour la publication et les conférences offrent aux découvertes dont ils rendent compte une plus grande portée et un examen plus approfondi en vue de leur amélioration, ce qui fait de la prépublication et de la publication officielle deux fonctions importantes du rôle du chercheur.

Même si le présent rapport et le marché mettent l'accent sur les autres rôles ci-dessous, il est crucial que le rôle du chercheur ne perde pas de son importance. Les organisations qui bénéficient de techniques de pointe ne peuvent espérer conserver longtemps leur avance dans ce domaine en évolution rapide sans réinvestir dans de bonnes recherches. En outre, les chercheurs ont souvent d'excellentes compétences qui sont transférables vers des rôles plus appliqués, ce qui fait que le rôle reste donc un excellent point de départ pour les meilleurs talents.⁵

Qualifications : Les qualifications requises pour ce poste comprennent un doctorat en apprentissage machine, en informatique, en intelligence artificielle ou dans un domaine quantitatif connexe; l'expérience et la maîtrise de la programmation scientifique et des bibliothèques pertinentes; l'expérience de la contribution aux communautés de recherche importantes pour faire progresser la qualité de la recherche.

⁵ Il persiste une question ouverte sur la durée de vie moyenne des publications des chercheurs, et sur le maintien du rythme de publication lorsque ceux-ci rejoignent l'industrie, car ils partagent ainsi plus de leur temps avec le travail appliqué. Il apparaît également dans les données ci-dessous que très peu d'auteurs sont des chercheurs professionnels à temps plein (4,6 %), et nous voyons beaucoup d'ingénieurs contribuer à la recherche à temps partiel. Une meilleure compréhension précise de ces chiffres devrait aider à déterminer un investissement équilibré dans les résultats de la recherche.

Ingénierie IA/AM

Ces ingénieurs représentent le pont essentiel entre la recherche fondamentale en IA et son application dans le monde réel. La différence significative avec les autres ingénieurs en logiciel est qu'ils « codent avec des données ». Une grande partie de leur travail consiste à utiliser les techniques les plus récentes d'apprentissage et à travailler avec des ensembles de données pour obtenir les techniques permettant d'apprendre quelque chose de spécifique et d'applicable à un problème.

Leurs titres les plus courants sont « scientifique de données », un exemple du théorème de Tesler sur la façon dont les techniques d'IA cessent d'être appelées IA lorsqu'elles deviennent standards. Ils peuvent également être nommés ingénieur IA/AM ou scientifique en recherche appliquée. Pour résoudre des problèmes pratiques, les meilleurs ingénieurs mettent en application leur profonde expertise technique, leur rigueur scientifique et leur créativité, développées au fil des années d'expérience. Cela exige d'eux de comprendre les besoins des utilisateurs tout en étant capables de mener les recherches nécessaires pour que les techniques disponibles pertinentes soient utilisables à l'échelle.

Eux aussi pré-publient souvent à l'avance leurs conclusions sur les méthodes appliquées sur arXiv, contribuant ainsi à un domaine très collaboratif et axé sur la science. C'est pourquoi nous avons pris en compte à la fois les chercheurs et les chercheurs appliqués dans notre évaluation d'arXiv (voir section suivante). Le travail appliqué prend de plus en plus d'importance en raison de la difficulté de faire fonctionner l'IA à l'échelle dans le monde réel, bien que la répartition précise ne soit pas claire.⁶ Il y a également beaucoup plus de personnes qui déclarent travailler dans le domaine de l'ingénierie (1 50 500) que de personnes qui publient sur arXiv (projection de 86 000 personnes d'ici fin 2020).

Les données autodéclarées et les données sur la demande provenant des médias sociaux reflètent plus fidèlement le marché de l'emploi. Elles montrent que la demande est plus forte pour les scientifiques de données et les développeurs en IA que pour les chercheurs, ce qui explique pourquoi il est possible de trouver des compétences de niveau « chercheur » tout au long de la chaîne de valeur en raison de la nécessité de les remplacer (voir la section « Le talent pour le développement »).

Qualifications : Un doctorat ou une maîtrise dans une discipline quantitative (par exemple, informatique, mathématiques, recherche opérationnelle, physique, génie électrique)⁷; une compréhension significative de la théorie sous-jacente de l'apprentissage profond et des domaines connexes de l'IA; une expérience des logiciels statistiques (par exemple, R, Python, MATLAB, pandas) et de la programmation scientifique; la capacité d'exécuter des modèles sur des données souvent bruyantes qui existent dans un contexte commercial; une expérience de la modification de modèles/techniques pour s'adapter aux limitations des données; des compétences dans la sélection des bons outils statistiques compte tenu d'un problème d'analyse des données.

Ingénierie et architecture des données

Étant donné la nature du travail des scientifiques de données qui est justement de coder avec les données, la mise à disposition d'un flux de données pour former un modèle d'IA et ensuite travailler à l'échelle est une part importante du travail. Les ingénieurs ou architectes de données construisent des structures de base de données modernes qui stockent des téraoctets de données industrielles utilisées

⁶ Cela fait en sorte que le classement par pays pour les publications sur arXiv s'avère utile pour les personnes qui cherchent à savoir s'il y a assez de talents pour ouvrir un laboratoire, ainsi que pour repérer qui est le plus proche des dernières avancées pour rester en tête. Les travaux futurs devraient faire la distinction entre les travaux appliqués et fondamentaux sur arXiv afin de déterminer la véritable répartition.

⁷ Il a été impressionnant de constater à quel point ceux qui évoluent dans des domaines connexes ont pu s'adapter avec les ressources en ligne disponibles.

pour développer des modèles d'IA. Ils construisent, testent et maintiennent des architectures optimales de pipeline de données, et s'assurent que l'architecture répondra aux besoins de l'entreprise. En outre, ils sont responsables des meilleures pratiques en termes d'organisation des données, de normes et de versions, et de la conformité aux réglementations internes et externes.

Ce qui est malheureusement caché, mais qui représente une majorité des heures de travail des gens, c'est tout le travail de préparation, « wrangling », de données qui se produit pour mettre en forme les données elles-mêmes. Le nettoyage, l'étiquetage et l'augmentation des données peuvent prendre jusqu'à [65 % des heures](#) d'un projet d'apprentissage machine [selon Cognilytica](#). Ce travail de préparation des données est très variable, car les ingénieurs en données peuvent également mettre au point des méthodes automatisées pour ces tâches, mais de nombreuses tâches de traitement des données nécessitent un certain degré d'expertise humaine, en particulier lorsqu'on travaille sur un nouveau domaine d'application.

Bien que notre rapport n'examine que les compétences de l'ingénieur de données en tant que rôle technique spécialisé nécessaire dans la chaîne de valeur, nous recommandons fortement que ce domaine ne soit pas négligé et qu'il soit considéré comme une main-d'œuvre valorisée, car l'IA fonctionne selon le principe « à mauvais intrant, mauvais extrant ».⁸

Qualifications : Un diplôme universitaire en informatique ou dans un autre domaine informatique est souvent une base nécessaire, de même qu'un ensemble d'aptitudes technologiques telles que la maîtrise de Spark, Cassandra, Hadoop, ainsi que des bases de données NoSQL et relationnelles. Le plus important demeure l'expérience des différentes composantes du rôle d'un ingénieur de données : la mise en œuvre de méthodes de transformation des données et le soutien de pipelines robustes en production; l'architecture des données et leur optimisation pour divers modèles de conception de logiciels; et l'expertise des méthodes qui assurent la qualité à grande échelle sur les principales plateformes infonuagiques.

Productisation IA/AM

La construction d'un produit autour d'une capacité d'IA exige une grande expertise logicielle traditionnelle, une bonne connaissance des méthodes d'IA et une compréhension approfondie du contexte de l'utilisateur final. Nous avons subdivisé cette catégorie en deux rôles distincts, l'un qui est plus axé sur l'évaluation de ce qui est une solution constructible possible et l'autre plus technique qui « exécute » la construction. Au fur et à mesure de la démocratisation des outils d'IA, les qualifications des développeurs d'IA convergeront vers celles d'un analyste de données et seront plus facilement complétées par des cours de formation en ligne largement disponibles.

Les **développeurs en IA** sont les personnes-clés pour la construction de l'environnement logiciel centré sur une capacité d'IA, ce qui en fait un produit pleinement fonctionnel. Ils participent à l'élaboration, l'architecture, la conception, le développement, les tests, le déploiement, l'exploitation, la maintenance et l'amélioration des modèles d'IA. Par la productisation des modèles d'IA, ils doivent également être en mesure d'aider à l'évaluation et à la sélection de la plateforme technologique, des cadres et de l'architecture de déploiement appropriés pour chaque problème à résoudre, ainsi que d'aider à maintenir les modèles d'IA déployés en production. Ce rôle s'étend à la conception, la mise en œuvre et l'exploitation d'API conviviales et évolutives, et même au travail sur l'UX et l'interface si l'équipe est réduite.

Qualifications : Un baccalauréat en informatique ou une expérience professionnelle équivalente qui leur aura donné la connaissance des cadres d'interface graphique Web et la maîtrise de

⁸ Consultez Samasource pour une approche permettant de valoriser de manière rentable le travail d'étiquetage des données.

quelques langages de programmation. Comme il s'agit de projets de développement novateurs, il sera favorable de cumuler au moins 5 ans d'expérience sur des projets à grande échelle (de préférence des solutions basées sur des microservices), et une expérience pertinente des outils de déploiement et d'automatisation basés sur des conteneurs.⁹

Les **analystes de données** sont des développeurs de solutions techniquement compétents. Leur rôle est d'analyser et de comprendre un problème et des ensembles de données pertinents afin d'évaluer la meilleure approche pour concevoir une solution. Bien qu'ils se concentrent sur l'utilisateur final, ils ont besoin d'un contexte technique pour pouvoir évaluer efficacement la faisabilité d'une solution (vérifier les distributions de données, identifier les lacunes des données, évaluer l'exactitude des étiquettes, la reproductibilité et la traçabilité autour de la collecte, de la transformation et des tâches analytiques associées) et communiquer le contexte complet à une équipe technique pour l'élaborer ou demander à l'entreprise de combler les lacunes.

Qualifications : Bien qu'aucun diplôme spécifique ne soit requis, ils devront avoir une expérience de la programmation, de la visualisation de données, des outils de statistiques et de préparation « wrangling » de données en rapport avec l'analyse des données et l'apprentissage machine (par exemple, scikit-learn, pandas DataFrame, NumPy, R, awk, RapidMiner, Tableau, D3.js), une expérience dans la mise au pas des données non structurées (par exemple, images, PDF) et des données structurées (par exemple, csv), et une large connaissance des techniques statistiques et d'apprentissage machine avec la capacité d'exécuter et de comparer des modèles existants.

Personnel de soutien

Les consultants, les concepteurs de solutions et les opérations du secteur de l'IA (par exemple, la comptabilité, le droit, etc.) jouent tous un rôle important dans l'adoption de l'IA et exigent une compréhension de la dynamique de cette nouvelle technologie. Nous n'avons pas inclus un examen approfondi de ce groupe, car une formation technique en IA n'est pas nécessaire pour remplir son rôle et qu'il est possible d'être mis à niveau rapidement. Nous avons toutefois pensé qu'il était suffisamment important de le mentionner pour souligner l'intérêt de la sensibilisation à l'IA afin de favoriser une meilleure compréhension de la nouvelle dynamique des logiciels d'IA et de comment nous nous y adaptons.

Bien que ce groupe soit relativement restreint et s'attache à faire comprendre aux dirigeants de l'organisation la valeur de l'IA et à les inciter à l'adopter, cela signifie qu'à terme, une grande partie de la main-d'œuvre sera prête à adopter cette nouvelle technologie et à collaborer avec succès.

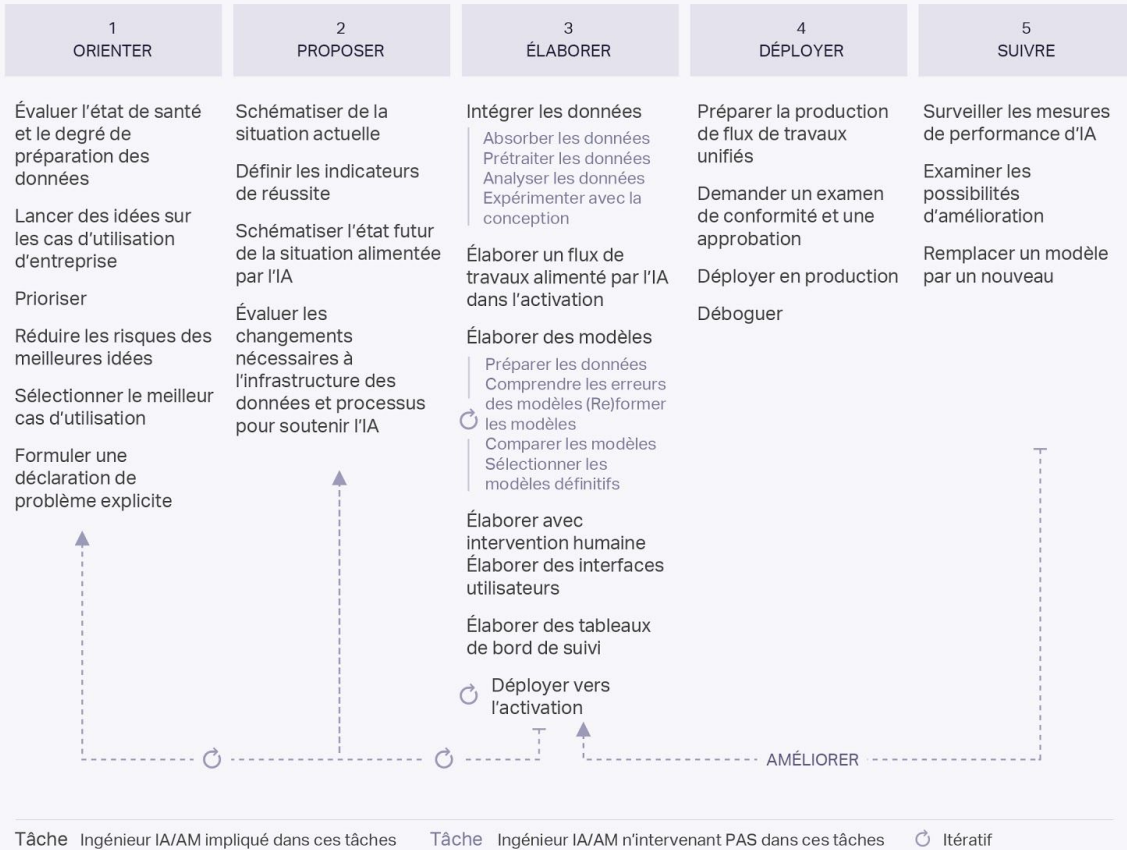
Limitation des typologies

Les rôles sont des catégories imparfaites, et les titres de postes associés peuvent souvent couvrir des emplois avec peu ou pas de travail en IA.¹⁰ Cependant, comme l'IA devient l'approche standard des logiciels, tous ces titres bénéficieront des compétences en IA – et le marché a beaucoup d'appétit pour le récompenser. En outre, cette nouvelle norme mettra encore plus l'accent sur l'ingénierie des données et leur préparation (« wrangling »), et où la gestion des données deviendra la principale entrée pour la construction de logiciels.

⁹ Cela signifie généralement que les rôles de développeurs d'IA sont occupés par des personnes ayant des qualifications de scientifique de données, ou alors qu'ils seront rapidement mis à niveau vers ce rôle. Voir la discussion dans « L'offre et la demande tout au long de la chaîne de valeur des talents » ci-dessous pour savoir comment cela a affecté les tendances en matière de talents.

¹⁰ Même s'il s'agissait d'une typologie parfaite, il existerait un problème de comparabilité internationale entre les autres typologies existantes qui sont en usage.

MISE EN PRODUCTION D'UN MODÈLE D'IA



Les ressources humaines qui rendent l'IA plus intelligente

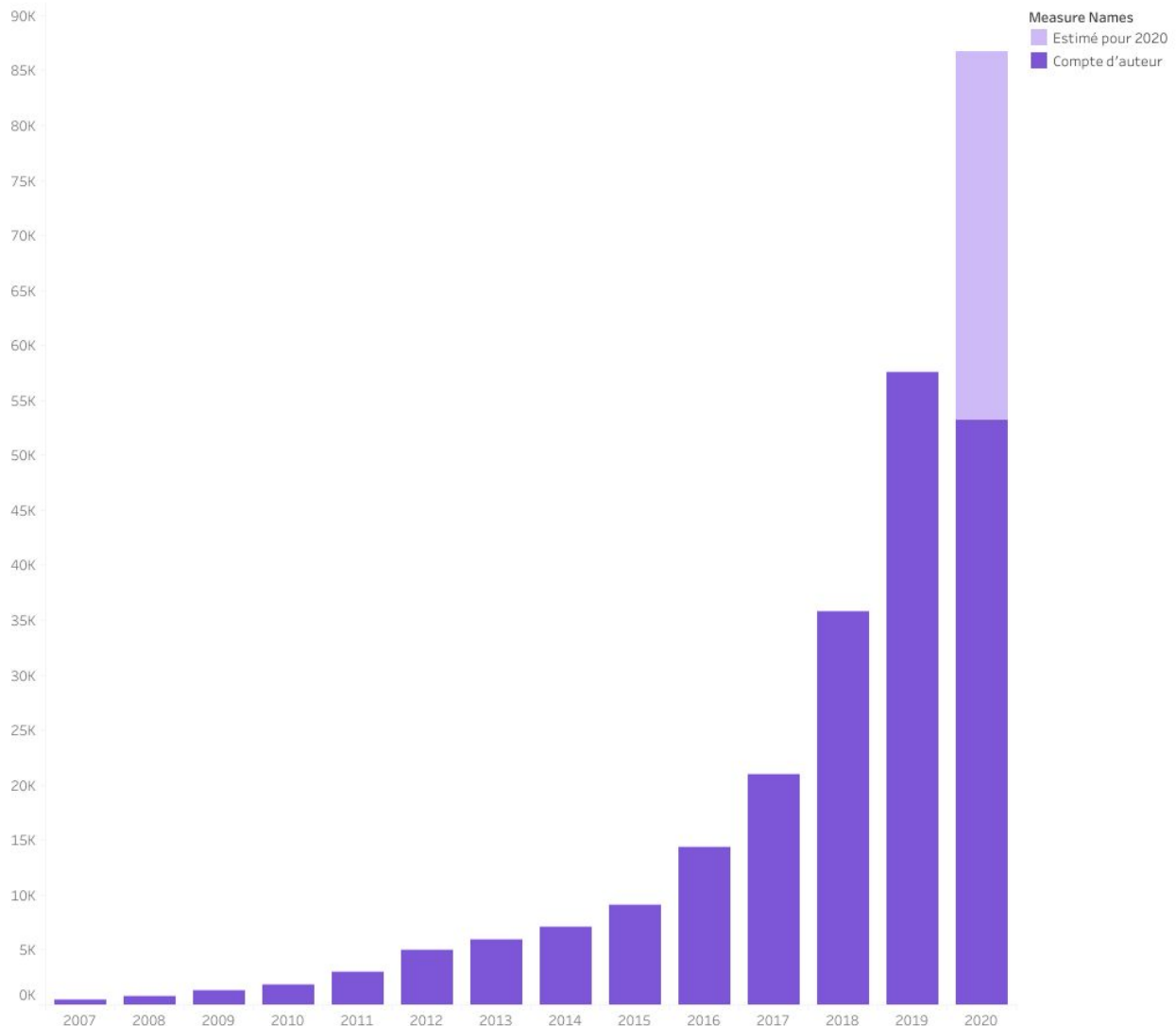
Données ArXiv couvrant les chercheurs appliqués et fondamentaux

L'année dernière, nous nous sommes concentrés sur les conférences scientifiques les plus prestigieuses autour de l'IA, dans le but de comprendre les recherches ayant le plus d'incidence. Cela a permis de compter 22 400 auteurs jusqu'en 2019. Cette année, nous avons décidé d'élargir notre regard sur la recherche en essayant de nous rapprocher le plus possible d'un « recensement de la recherche », le répertoire de prépublication d'arXiv. Cette approche nous a permis de mieux comprendre comment la recherche sur l'IA se développe au-delà de ce que nous avons vu par le passé : d'un petit cercle de chercheurs travaillant à rendre l'IA réelle, jusqu'à une constellation d'experts trouvant des moyens de mettre en œuvre l'IA dans différents domaines.

En d'autres termes, ce changement devrait nous permettre de nous concentrer moins sur la « recherche fondamentale en IA » et d'ajouter plus de « recherche appliquée en IA » dans notre champ d'action. Cette démarche a eu un certain effet que nous soulignerons tout au long du rapport, ainsi que dans la section détaillée sur la méthodologie.

Dans le graphique ci-dessous, nous avons appliqué la méthodologie aux années précédentes jusqu'en 2007 pour montrer la courbe de croissance mondiale avec cette nouvelle perspective.

Comptage annuel des auteurs d'IA sur ArXiv



Nombre total d'auteurs publiant sur arXiv chaque année. Comptage des auteurs basé sur les métadonnées arXiv au moment de la rédaction. Estimation basée sur la croissance moyenne pour toutes les années

Un changement significatif par rapport à la méthodologie de la conférence concerne les tendances de croissance. Alors que la croissance a augmenté de 36 % par rapport à 2015 et de 19 % en 2016 lors des conférences, nous constatons une croissance de 47,69 % et de 53,88 % respectivement pour les auteurs sur arXiv. Cela montre la part de l'activité qui se situe en dehors des canaux traditionnels, évalués par les pairs, ainsi que la limite des places de conférence pour montrer la croissance. L'une des principales raisons pour lesquelles il est acceptable d'échapper au monde de l'évaluation par les pairs est qu'il est facile de déterminer rapidement si quelque chose fonctionne ou non. Les conférences constituent toujours un filtre important pour les « meilleures idées » et influencent l'application des normes éthiques (voir les conférences [ICLR](#) et [NeurIPS](#) qui requièrent maintenant l'inclusion de considérations éthiques dans les articles soumis), mais le volume sur arXiv est indicatif du nombre de personnes qui s'engagent avec les outils d'IA et peuvent ainsi alimenter le vivier de talents.

Ces articles (environ 78 000 au total, soit [un peu moins de 3 % de l'ensemble des publications](#) scientifiques selon l'index HAI 2019 de Stanford) sont également susceptibles de couvrir des approches

appliquées plutôt que des découvertes fondamentalement nouvelles, rendant là encore le domaine d'autant plus accessible aux personnes moins formées.

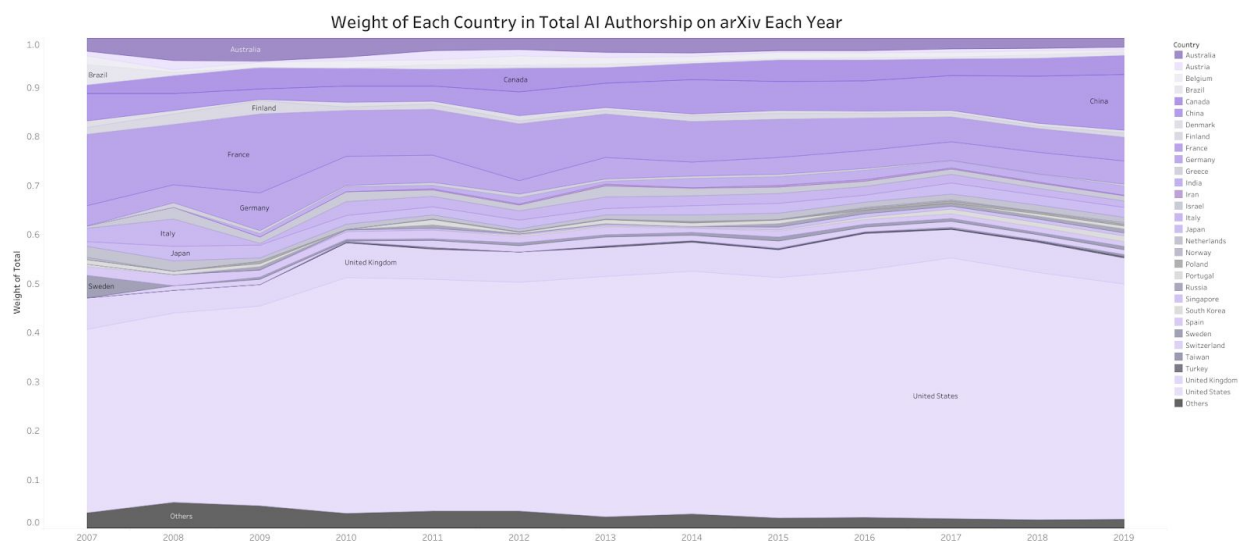
Mais cela ne doit pas être une raison pour abandonner la formation. Cette croissance est due à l'investissement soutenu dans l'enseignement supérieur, qui porte tranquillement ses fruits, car ces doctorats pluriannuels mettent du temps à se manifester. Il sera important de maintenir cette dynamique, car une partie de la croissance du nombre d'auteurs provient en fait d'experts d'autres domaines qui utilisent les méthodes d'AM et qui choisissent d'ajouter simplement l'AM comme catégorie de publication afin d'indiquer leur utilisation de techniques de pointe.¹¹ En outre, les titulaires d'un doctorat en IA ne resteront probablement pas tous dans le milieu universitaire et fourniront des talents de haut niveau dans l'industrie tout au long de la chaîne de valeur (ou peut-être qu'ils mettront leurs compétences au service d'un autre domaine de recherche et ne considéreront pas qu'il est important d'inclure l'AM comme catégorie de publication).

Dans l'ensemble, la croissance est remarquable puisque du très petit nombre (478) auteurs en 2007, nous sommes passés à un peu moins de 58 000 **au début de cette année et à une projection de 86 000 d'ici la fin 2020.**

Classement national

Les États-Unis restent dominants, et la Chine est dans une classe à part

Les classements nationaux se maintiennent, les États-Unis restant le plus grand acteur avec 47,89 % du nombre total d'auteurs. Cependant, il convient de faire preuve d'une grande prudence, car notre méthodologie est centrée sur un arXiv occidental et ne prend en compte que les articles publiés avec l'alphabet latin. Un poids supplémentaire est également accordé aux États-Unis en raison de la domination de leurs acteurs dans le domaine des hautes technologies, puisque nous avons attribué l'emplacement des auteurs de l'industrie au siège de leur affiliation, quel que soit le bureau local où ils travaillent. Bien que cela ne tienne pas compte de la valeur de la proximité géographique qui stimule les collaborations et la formation locales, la priorité est donnée à la localisation de la propriété intellectuelle (PI) qui est générée.



Comme nous ne disposons que d'une ventilation par pays pour cette année (celle-ci a été établie à partir d'un sous-ensemble de données sur la population), nous ne pouvons pas dire directement les

¹¹ Les éditeurs sur arXiv peuvent sélectionner plusieurs catégories pour publier.

changements de classement par rapport aux années précédentes. Cependant, par rapport aux années précédentes de publication de conférences, les principaux acteurs sont les habitués : la Chine se situe dans son propre classement, à environ 11,4 %; le Royaume-Uni (5,3 %), la France (4,9 %), l'Allemagne (4,7 %) et le Canada (3,9 %) dans leur propre groupe; puis on retrouve le Japon (2 %), l'Australie (1,9 %), l'Inde (1,8 %), l'Italie (1,3 %) et la Corée du Sud (1,3 %). Ces « 10 prochains » représentent 38,6 % du nombre total d'auteurs, et avec les États-Unis, ces 11 premiers représentent 86,5 %; des profils peu surprenants à voir, car tous ces pays sont des pays du « Global North » comprenant des secteurs technologiques importants. D'autres sur cette liste pourraient rattraper assez rapidement un certain investissement ciblé, grâce à l'aide d'un nombre relativement restreint de personnes dans leur ensemble.

Nombre moyen de collaborations internationales par auteur

Les collaborations se poursuivent au niveau mondial, les petits pays se distinguent

Cela ne veut pas dire que le travail n'est pas partagé au niveau mondial. En identifiant le pays de chaque auteur sur un document, nous avons constaté que les collaborations se poursuivent partout avec tout le monde, à la triste exception des pays du « Global South », qui sont largement absents de la carte des collaborations. Cette année, nous avons changé d'orientation, passant de la question des pays qui collaborent entre eux à celle du nombre moyen de collaborations par auteur dans un pays donné.

Les grands gagnants sont des petits pays pour la plupart, et une explication intuitive est que c'est par nécessité en raison du nombre réduit d'options locales. Bien que les États-Unis ne figurent pas en tête de liste des collaborations moyennes par auteur, avec une moyenne par auteur moyenne (4-6) similaire à celle d'autres pays ayant atteint la maturité en matière d'IA, ils restent de loin le pays où le nombre total de collaborations internationales est le plus élevé. La Chine et la Corée du Sud, notamment, se situent dans le bas de l'échelle avec une moyenne de moins de 3 par auteur.

À l'autre extrémité se trouve un cas extrême, l'Irlande, avec 15 par auteur. Les finalistes sont Singapour et la France, avec respectivement 8 et 7 auteurs. D'autres pays se situent entre Singapour et l'Irlande (par exemple la Hongrie (11,5), l'Estonie (9,5) ou la Belgique (8,5), mais ces pays ne doivent pas être considérés autant en raison de leur nombre de chercheurs nettement inférieur.

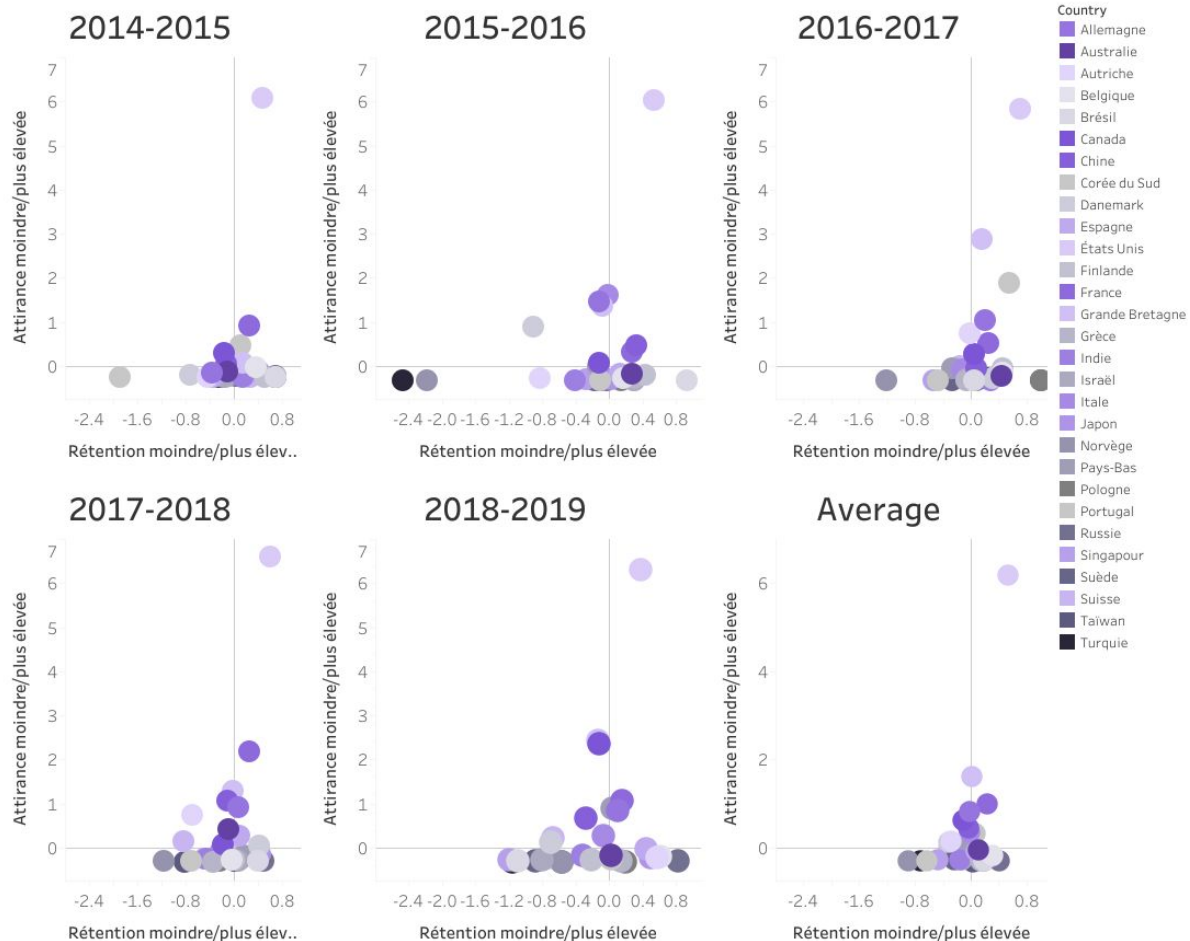
Moyenne des collaborations internationales par auteur



Attraction et rétention des talents

Plus de pays sont capables de résister à l'attraction des États-Unis qu'il y a 5 ans, et 2020 pourrait accélérer cela

En traçant les pays en fonction de leurs entrées et sorties de talents en IA, nous voyons que les pays appartiennent à l'un des quatre types suivants : pays invitants, pays producteurs, pays ancrés et pays plateformes.



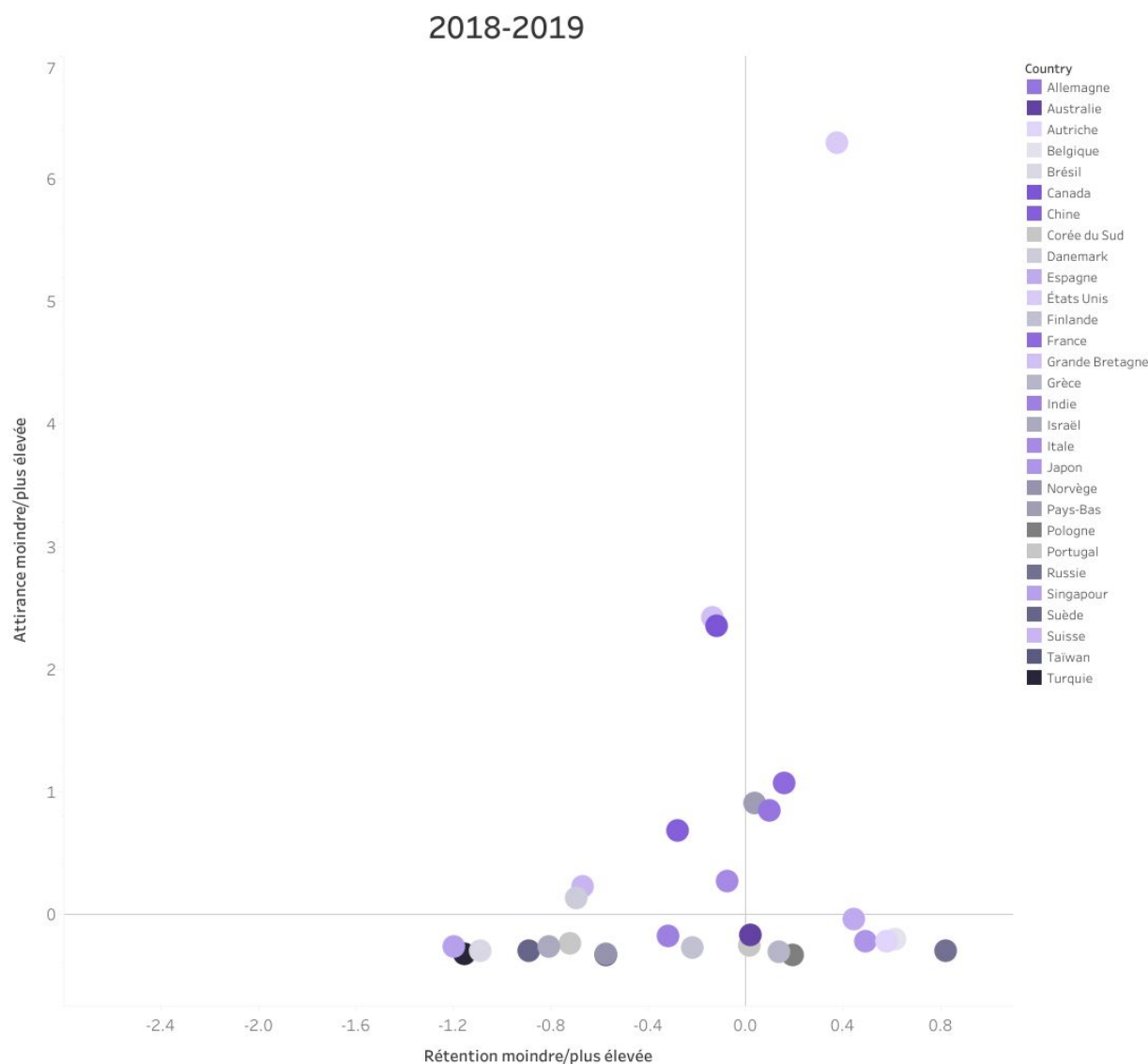
Comparaison de la capacité des pays à retenir et à attirer les talents. L'axe des X représente la capacité à attirer les talents par rapport à la mesure de la probabilité que les auteurs commencent dans un autre pays et finissent dans un pays cible. L'axe Y représente la capacité à retenir le talent en mesurant la probabilité qu'un auteur qui commence dans un pays y reste. Les deux axes sont des mesures de l'écart type par rapport à la moyenne de tous les pays.

En examinant l'emplacement des affiliations des auteurs d'une année sur l'autre, nous avons pu comparer les entrées et les sorties de talents dans chaque pays en pourcentage du bassin total de talents des auteurs du pays. Le flux de talents a été mesuré en comparant la probabilité qu'un auteur se déplace vers et à partir d'un pays par rapport à la probabilité moyenne pour l'ensemble des pays. Nous avons calculé les flux entrants et sortants moyens de tous les pays, puis nous avons examiné la distance de chaque pays en termes d'écart types par rapport aux flux entrants et sortants moyens pour obtenir un score normalisé.

La raison d'être d'un score normalisé dans ce cas est que les pays n'existent pas isolément. La capacité d'un pays à attirer les talents est liée à l'incapacité d'un autre à les retenir. Nous comparons à une moyenne afin de voir comment les efforts de chaque pays affectent leur poids sur le marché mondial des talents de la recherche.

La mesure « d'invitation » montre la probabilité qu'ont les auteurs de se retrouver dans un pays, ce qui représente l'attraction qu'un écosystème est capable d'exercer sur le talent, tandis que la mesure « de rétention » est celle de la probabilité qu'un auteur reste pendant un an.

Pour voir comment les pays se comparent dans cette dynamique de « répulsion-attraction », nous avons tracé les scores « d'invitation » sur l'axe des x et les scores « de rétention » sur l'axe des y. Ces valeurs nous ont permis de classer les pays dans les quatre groupes distincts décrits ci-dessous en fonction du quadrant du graphique dans lequel ils se trouvaient.



Comparaison de la capacité des pays à retenir et à attirer les talents. L'axe X représente la capacité à attirer les talents, mesurée par la probabilité que les auteurs commencent dans un autre pays et finissent dans le pays cible. L'axe Y représente la capacité à retenir le talent, mesurée par la chance qu'un auteur qui commence dans un pays y reste. Les deux axes sont des mesures de l'écart type par rapport à la moyenne de tous les pays.

Pays producteurs – Quadrant inférieur gauche

© 2020, Element AI Inc. Tous droits réservés. « Element AI », « EAI » et le logo d'EAI sont des marques de commerce d'Element AI, et peuvent être enregistrées ou en attente d'enregistrement.

Nous considérons l'Inde, Singapour et Israël comme des pays producteurs, parce qu'ils ont connu moins d'entrées et plus de sorties, en proportion du bassin de talents du pays, que la moyenne. Israël a connu quelques revers, passant de 0,3 dans sa meilleure année à -0,8 en 2018-2019 pour la rétention des talents, tout comme Singapour, qui est passée de 0,16 dans sa meilleure année à -1,19 en 2018-2019 pour la rétention.

Beaucoup de ceux qui publient sur arXiv développent un ensemble de compétences qui est souhaité tout au long de la chaîne de valeur, et donc ces pays qui ont une sortie plus importante voient leurs investissements en talents partir pour créer de la valeur dans d'autres pays. Singapour est un exemple important de ce danger, car cet exode a contribué à la croissance zéro de leurs données d'approvisionnement.

Pays ancrés – Quadrant inférieur droit

Le Japon, la Belgique et la Russie sont ce que nous appelons des pays ancrés. Les entrées et les sorties de talents ont été moins importantes que la moyenne, en proportion du bassin total de talents du pays. Elle signale la relative stabilité de leurs viviers de talents, mais peut-être aussi une insularité risquée.

Le Japon se classe peut-être au 7^e rang pour les chercheurs, mais il est l'un des pays où les mouvements d'entrée et de sortie sont les plus faibles, et se situe en dessous de la moyenne pour les collaborations internationales (28^e au total). Comme l'IA évolue rapidement, la connexion avec la communauté mondiale est importante pour se tenir au courant des dernières avancées.

La Serbie ne figure pas dans le graphique en raison de son nombre d'auteurs qui est trop faible, mais il convient de noter qu'elle a fait une grande amélioration, passant d'un score de -3,7 à +,98 « qui restent ». Ceci pourrait également signifier que cette section contient des pays qui font l'important premier pas pour conserver le talent qu'ils créent, sinon l'attirer de l'étranger. La Belgique et les Pays-Bas ont également montré que les efforts déployés pour influencer sur les flux de talents ont porté leurs fruits en termes de croissance soutenue. Ils sont passés d'écosystèmes en développement qui peinaient à retenir les talents, à une croissance régulière par rapport à leur population, malgré leur taille plus modeste que celle de leurs pays voisins (de -0,2 à 0,9 de taux d'invitation pour les Pays-Bas, et la Belgique est passée de 0,3 « qui restent » à 0,6, doublant ainsi sa capacité comparative à retenir les gens). Les Pays-Bas ont en fait suffisamment bien réussi pour être qualifiés de pays « invitants ».

Pays plateformes – Quadrant supérieur gauche

Plusieurs pays ont connu à la fois plus d'entrées et plus de sorties que la moyenne, en proportion du bassin total de talents du pays. Ces pays parviennent à attirer des travailleurs formés à l'étranger tout en enregistrant un nombre de départs nettement supérieur à la moyenne. Ces écosystèmes, que nous appelons les pays plateformes, sont le mieux illustrés par le Royaume-Uni, la Chine et le Canada. Ces pays sont connus pour leurs points chauds compétitifs en matière d'IA et présentent des options attrayantes pour les talents en termes de pairs et d'institutions de haut niveau. Toutefois, ces succès peuvent générer une pression concurrentielle sur leur capacité à retenir les talents, notamment en suscitant l'intérêt des laboratoires d'IA internationaux pour s'implanter localement.

Le Royaume-Uni et le Canada ont enregistré des gains importants. Le Royaume-Uni est passé d'un niveau historiquement bas de 0,05 « d'invitation » en 2014/2015 à 2,4 cette année, et le Canada est passé de 0,3 « d'invitation » en 2015/2016 à 2,3 (la plus grande partie de la hausse ayant eu lieu l'année dernière). La stratégie de la Suisse a réussi à inviter des talents, de -0,2 en 2014/2015 à 0,2 en 2018/2019, peut-être en raison de sa grande différence de revenus avec l'Italie juste au sud de la frontière, mais elle leur a aussi coûté de conserver leur pouvoir de rétention (0,26 à -0,66).

Pays invitants – Quadrant supérieur droit

Les États-Unis, la France et le Portugal ont tous connu plus d'entrées et moins de sorties que la moyenne, en proportion du bassin de talents du pays. Cela signifie que ces pays réussissent relativement mieux à la

fois à retenir les talents avec lesquels ils ont commencé et à attirer davantage de talents d'autres écosystèmes. Nous appelons ces écosystèmes des pays invitants.

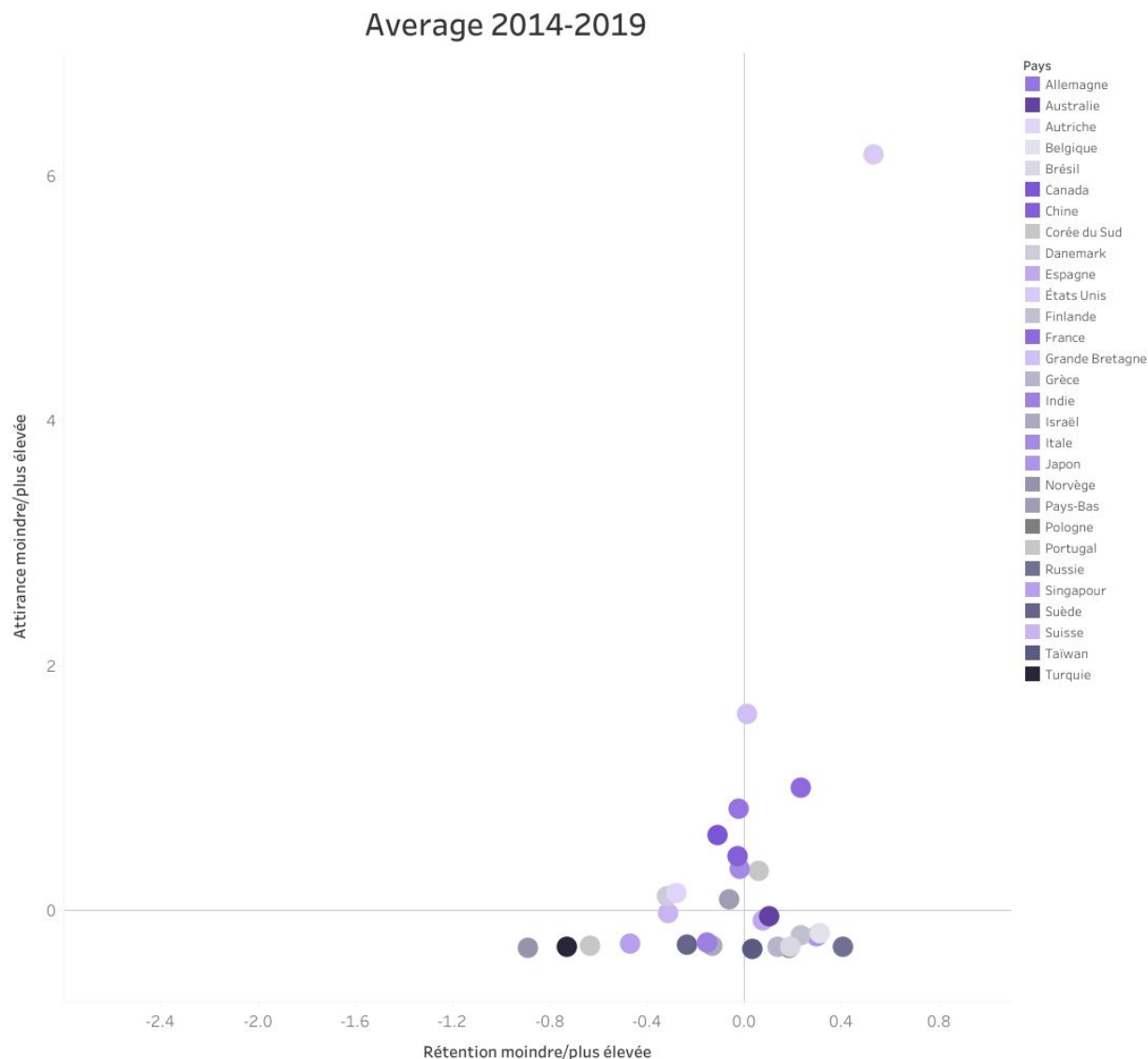
Les États-Unis dominent pour ce qui est d'attirer les talents dans la recherche et le milieu universitaire, et ont beaucoup d'emplois pour y maintenir les gens après l'obtention de leur diplôme.¹²

Selon une enquête du Center for Security and Emerging Technology (CSET), [38 % des titulaires d'un doctorat américain en intelligence artificielle ont effectué leur premier cycle d'études à l'étranger et 48 % sont nés en dehors des États-Unis](#). Cela est conforme à leurs conclusions selon lesquelles 55 % des titulaires d'un doctorat américain en STIM (sciences, technologies, ingénierie et mathématiques) sont des citoyens américains, tandis que les ressortissants chinois et indiens représentent respectivement 16 % et 6 %. Dans certaines universités, le pourcentage d'étudiants étrangers diplômés est sensiblement plus élevé : à la Tandon School of Engineering de l'université de New York, par exemple, [80 % des étudiants diplômés seraient venus de l'étranger](#). Le CSET a également constaté qu'entre 82 et 92 % des titulaires d'un doctorat américain en IA restent aux États-Unis pour y travailler au cours des cinq premières années suivant l'obtention de leur diplôme. Alors que la principale motivation pour venir était la qualité de l'éducation (82 %), seuls 42 % ont déclaré que les possibilités d'emploi étaient un facteur, cette situation peut donc être surtout un facteur de commodité.

Toutefois, les règles plus strictes proposées par la nouvelle politique américaine de visa H1B rendent plus difficile pour les étudiants de rester aux États-Unis après l'obtention de leur diplôme et limite même leur entrée en tant qu'étudiants étrangers avec certaines des plus strictes politiques de visa proposées. Cela représente une occasion pour les pays disposant d'un vaste marché du travail d'attirer et de retenir les talents, un aspect à ne pas perdre de vue dans les années à venir. De nombreux pays semblent être capables d'attirer et de retenir les talents, bien mieux qu'il y a cinq ans, et pourraient être avantageusement positionnés pour en tirer parti. Le Canada, en particulier, a montré qu'il en était capable. En 2017-2018, le Canada avait un score de 0,062 sur les invitations, pour 2018-2019 il affiche un score de 2,347, regagnant ainsi une grande partie des talents qu'il avait perdus les années précédentes au profit de la Silicon Valley.

Il n'en demeure pas moins que les États-Unis partent avec une grande avance. En moyenne, nos données montrent qu'il y a 4 % de probabilité que ceux qui font de la recherche en dehors des États-Unis finissent pourtant par y être publiés. En comparaison, il y a 1 % de probabilité d'une publication au Royaume-Uni.

¹² Comme notre étude définit le lieu d'implantation d'un chercheur en fonction du siège de la société où il travaille, ces résultats pourraient être considérés comme gonflant les chiffres américains. Cela s'explique par le fait que de nombreuses entreprises qui créent des laboratoires dans le monde entier ont leur siège aux États-Unis.



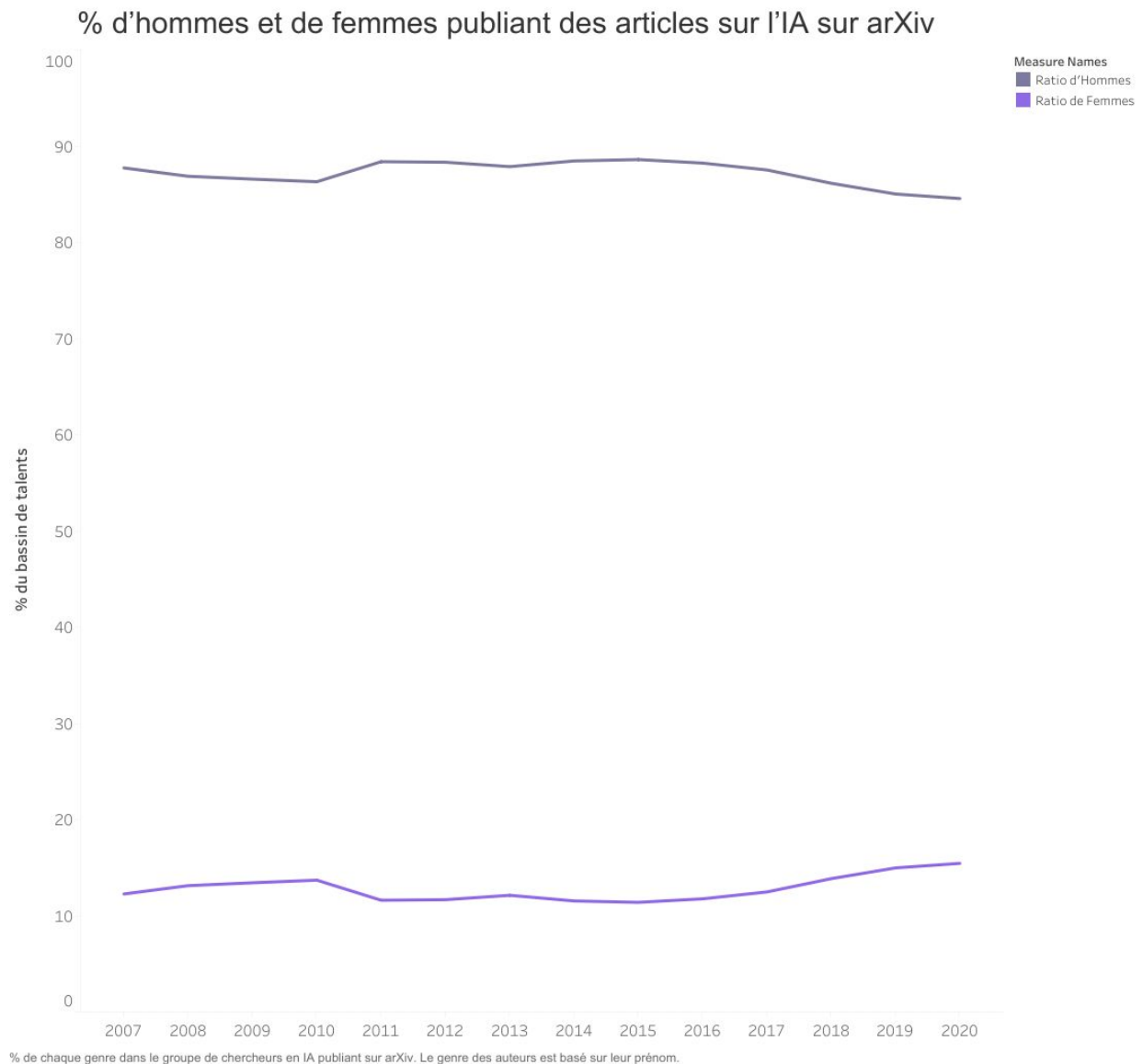
Comparaison de la capacité des pays à retenir et à attirer les talents. L'axe X représente la capacité à attirer les talents, mesurée par la probabilité que les auteurs commencent dans un autre pays et finissent dans le pays cible. L'axe Y représente la capacité à retenir le talent, mesurée par la chance qu'un auteur qui commence dans un pays y reste. Les deux axes sont des mesures de l'écart type par rapport à la moyenne de tous les pays.

Ces flux de talents devraient être surveillés attentivement au cours de l'année prochaine pour voir comment la pandémie a affecté la mobilité géographique. Les travailleurs de l'IA font montre d'un niveau de sophistication qui les rend capables d'exprimer leur talent par des connexions Internet, ce qui les rend théoriquement 100 % « mobiles ». Cette mobilité pourrait saper les avantages naturels des pays où se trouvent les meilleures écoles et les meilleurs emplois, et accroître la concurrence en permettant aux talents de mieux choisir leur combinaison préférée de rémunération (et d'impôts), de lieux (pour la vie et le travail), de champs d'application et de collègues. Toutefois, elle témoignera de l'importance (ou de l'absence) de la proximité avec d'autres chercheurs, par rapport au simple contact et à l'accès. Des études indiquent que cette proximité donne de meilleurs résultats, mais la question demeure de savoir si les chercheurs voient les choses de cette façon.

Équilibre des genres

Un lent mouvement sur 13 ans, mais de grandes variations d'un pays à l'autre

La mesure du genre a été basée sur les noms des auteurs. En utilisant les données du recensement américain, nous avons créé une liste de probabilités pour chaque nom et avons conservé ceux qui avaient une probabilité de plus de 95 % d'être d'un genre plutôt que de l'autre.¹³ Nous reconnaissons qu'il s'agit d'une mesure grossière en raison de l'ambiguïté de nombreux noms, de l'accent mis sur les États-Unis et, bien sûr, de la façon dont les gens peuvent ne pas s'identifier à l'un ou l'autre genre.



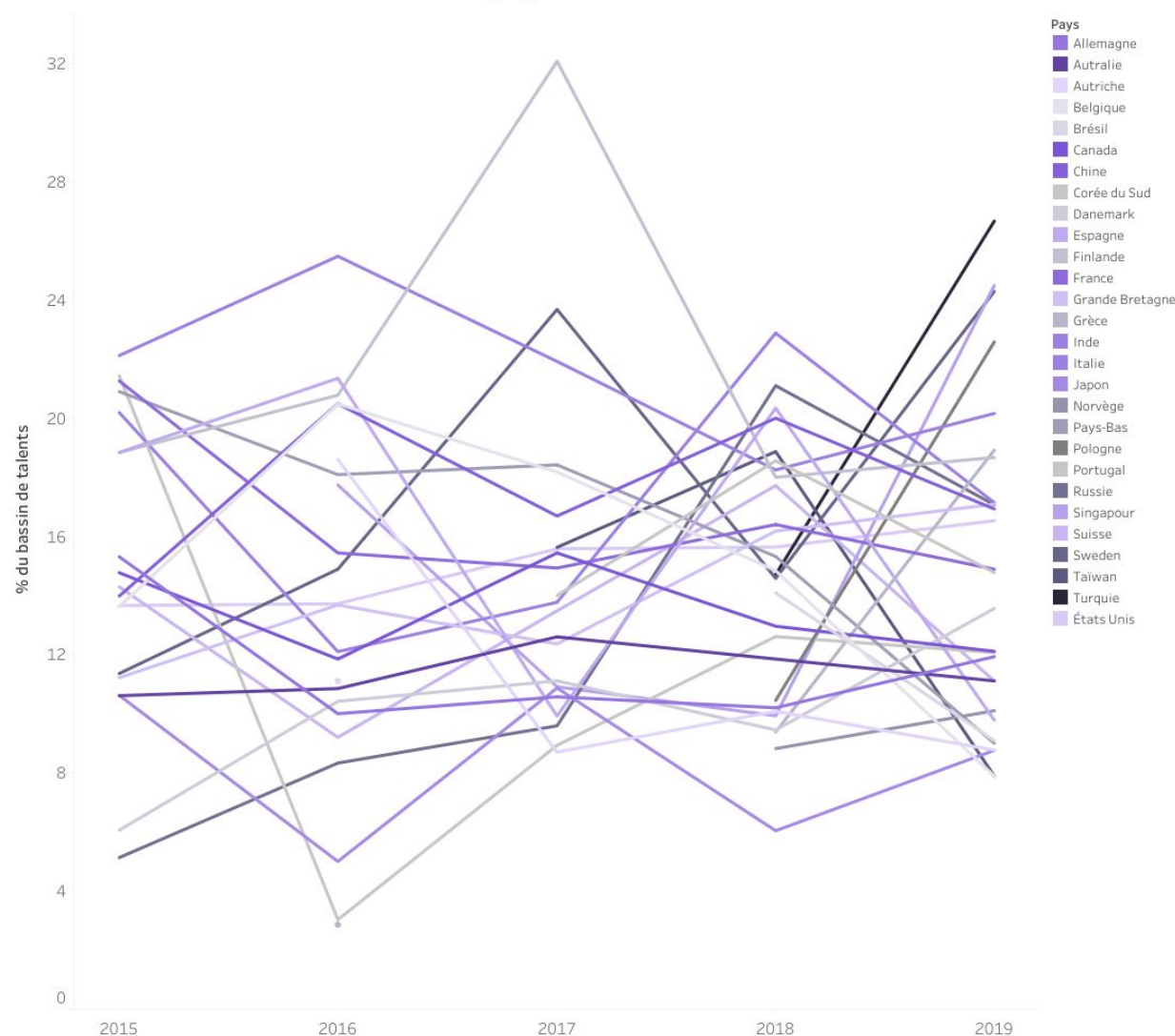
Lorsque l'on examine le ratio agrégé, on constate une légère convergence. Le pourcentage de femmes est passé de 12,26 % en 2007 à 15,44 % en 2020. Lentement, mais sûrement. Il est difficile d'évaluer la performance des pays dans ce domaine, car nous pouvons observer de grandes fluctuations du ratio global d'une année sur l'autre. Ces fluctuations sont probablement dues au fait que de nombreux auteurs

¹³ En nous basant sur la logique présentée dans cet article:
<https://journal.r-project.org/archive/2016-1/wais.pdf>

sont des étudiants, qui quittent ensuite la recherche et/ou s'installent dans d'autres pays après avoir obtenu leur diplôme. Ce mouvement de désertion de la publication est renforcé pour les femmes, qui ont tendance à obtenir des emplois offrant moins de possibilités de publication que les hommes. Par conséquent, le ratio d'une année donnée pour un pays dépend beaucoup des cohortes d'étudiants entrants, ce qui rend le résultat plus volatil.

Certains pays ont montré une croissance significative et persistante des données : Turquie (26,67 %), Singapour (24,49 %), Suède (24,29 %), Pologne (22,58 %), Grèce (18,92 %), Russie (17,07 %) et Danemark (13,56 %). Les principaux facteurs d'influence du ratio global restent toutefois les États-Unis (16,53 %) et la Chine (16,93 %) en raison de leur taille. Compte tenu de la reconnaissance dans le domaine de la valeur d'un leadership diversifié, les femmes seront très demandées, et les retenir aura probablement des retombées positives pour attirer davantage de femmes dans ce domaine dans un pays donné.

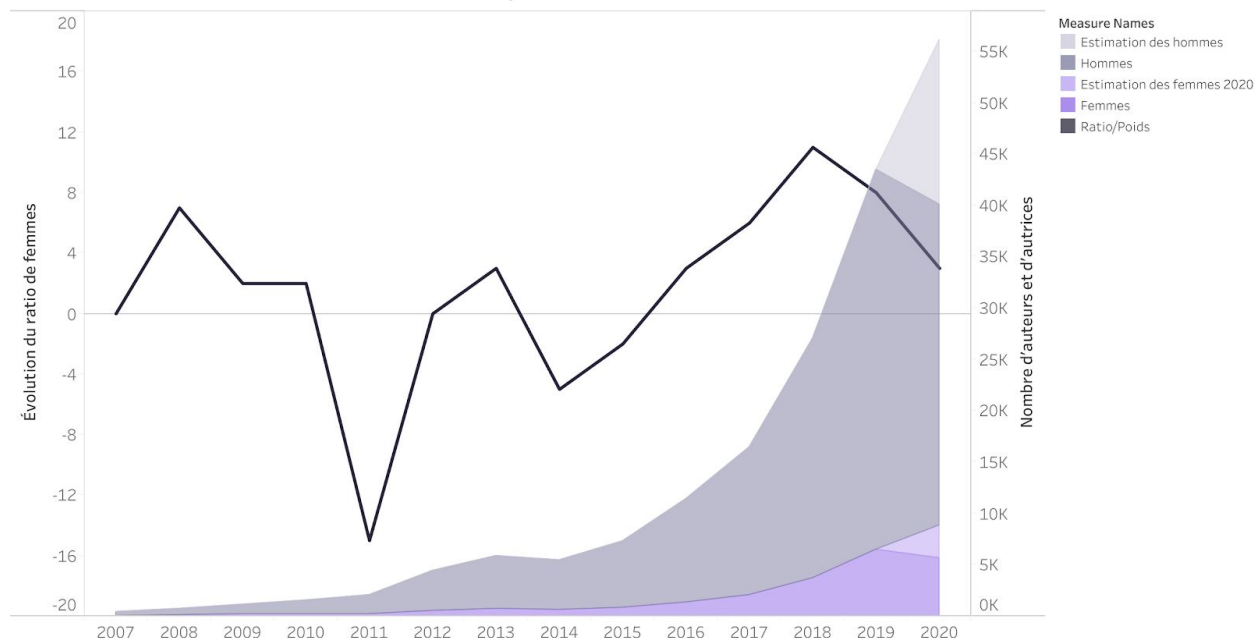
Ratio de femmes publiant des articles sur l'IA sur arXiv pour chaque pays



% de femmes dans le bassin de talents d'un pays d'une année à l'autre. Le genre de l'auteur est basé sur le prénom.

En observant le taux de croissance global, on peut également constater que certaines années, la croissance des auteurs a largement dépassé celle des autrices, ce qui a entraîné une baisse du ratio. Depuis 2007, le taux de croissance moyen des femmes est de 55,44 %, alors qu'il est de 51,21 % pour les hommes. Cela devrait suffire à combler l'écart au fil du temps, mais les années de croissance relative manquée retardent cette évolution. Si les taux de croissance moyens se maintiennent pendant le reste de l'année, le ratio continuera à s'équilibrer. Mais, il a été discuté que [les retombées de la pandémie ont une incidence disproportionnée sur les femmes](#). En ces temps incertains et face aux défis de la rentrée scolaire, on ne peut pas ignorer la possibilité d'un nombre plus faible d'autrices en 2020 par rapport à 2019, ce qui serait une première malheureuse.

Nombre d'hommes et de femmes publiant des articles sur l'IA sur arXiv



Nombre d'hommes et de femmes publiant, avec estimation du genre à partir du prénom, et ligne de tendance indiquant l'évolution annuelle en pourcentage de la proportion de femmes. Les données de 2020 sont en date d'août 2020. Estimations pour la fin 2020 basées sur la moyenne de la croissance pondérée des 5 dernières années. Exclut les noms pour lesquels l'algorithme avait moins de 95 % de confiance dans la classification.

Les personnes pour le construire

Offre disponible pour l'industrie – comment les talents de l'IA se modulent en fonction du marché commercial

En utilisant les rôles et les titres standards que nous avons catégorisés, nous avons recueilli des données sur l'autoreprésentation dans les médias sociaux. La collection était basée sur l'utilisation de mots-clés. Nous avons effectué une recherche à la fois par titre (par exemple « scientifique de données ») et par mots-clés pour leurs listes de compétences (par exemple « apprentissage machine » + « tensorflow » + « PhD »). La croissance représente les personnes qui ont un emploi en IA pour la première fois, sur une période d'environ un à trois ans.

Au total, nous avons compté **477 956 personnes** dans le monde entier dans les rôles ci-dessous. Comme les typologies de compétences varient d'un pays à l'autre, il est difficile de trouver des estimations antérieures du bassin total de talents pour comparer. Un point de données pourrait être l'estimation de Tencent pour 2017, soit 300 000 personnes, qui représentent *tous* les employés travaillant dans les

entreprises ou départements d'IA (y compris le personnel de soutien et administratif). En mettant l'accent sur le talent technique, notre mesure montre certainement une croissance significative en réponse à l'augmentation de la demande au cours des dernières années.

Ratios de l'ensemble du bassin de talents



Les proportions sont à peu près égales entre l'offre et la demande au niveau mondial (si l'on exclut l'ingénierie/architecture des données), généralement 61 % IA/production de données, 38 % ingénierie, et environ 1 % recherche.¹⁴ Il n'y a pas de pays où la part des chercheurs se rapproche de celle des autres rôles, le plus proche étant le Canada qui compte 2,36 % de chercheurs dans son réservoir de talents. Cependant, certains pays comptent plus d'ingénieurs que d'ingénieurs d'application : la France (61,20 % d'ingénieurs) et l'Allemagne (63,21 % d'ingénieurs). À l'autre bout des extrêmes, il y a la Chine avec 86,82 % d'ingénieurs d'application.

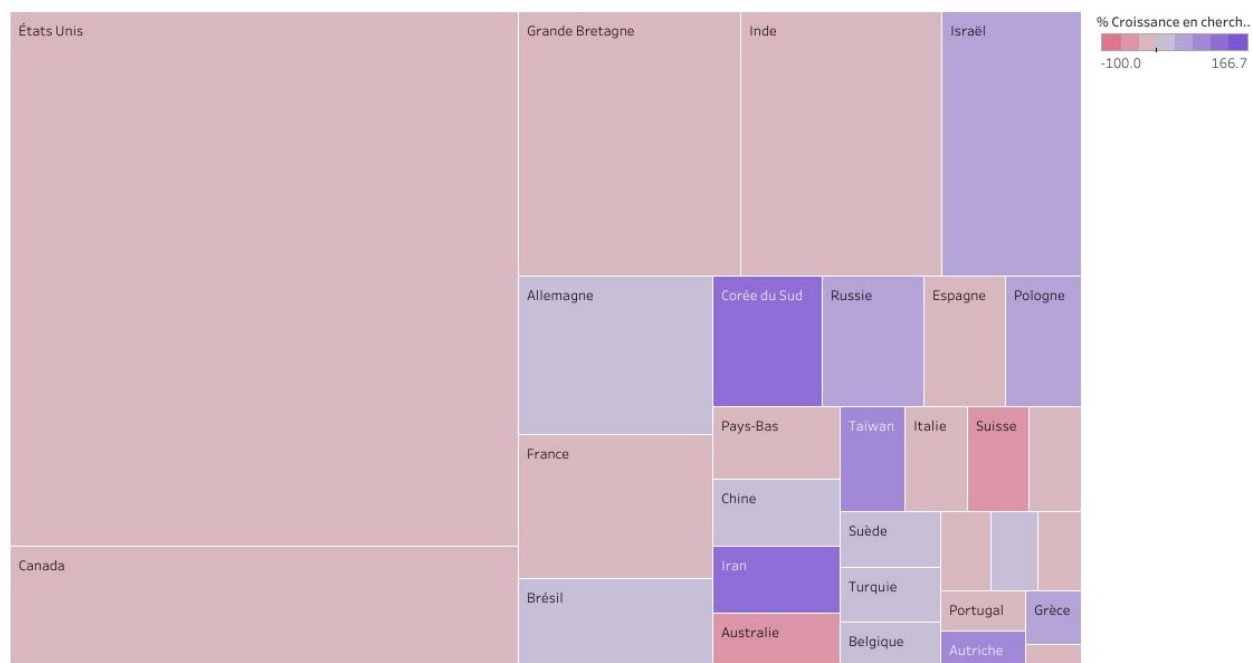
Le nombre total de chercheurs autodéclarés est de 4 149. Ce chiffre est bien inférieur au total des publications lors des conférences de l'année dernière (22 000) et au total sur arXiv (qui devrait être de 86 000 d'ici la fin de l'année). Cela nous indique que peu de chercheurs fondamentaux s'inscrivent sur les médias sociaux orientés vers les marchés de l'emploi de l'industrie, ou s'ils le font, se présentent peut-être comme ingénieurs pour obtenir les emplois appliqués qui sont disponibles. Bien que nous ne puissions pas dire avec précision combien de chercheurs continuent à effectuer du travail fondamental, les données sur la demande ci-dessous montrent que l'offre répond à la demande.

En d'autres termes, ces données représentent les chercheurs qui continuent à faire de la recherche professionnelle dans le secteur privé, ce qui montre la difficulté de continuer à publier de nouveaux travaux fondamentaux.¹⁵ Le Canada fait ici preuve d'un dynamisme remarquable, qui s'explique peut-être par le nombre d'entreprises internationales qui viennent mettre en place des laboratoires d'IA spécifiquement destinés à attirer les talents en recherche.

¹⁴ Alors que l'échelle relative aux autres rôles est similaire pour les chercheurs, la différence entre l'offre et la demande de chercheurs est beaucoup plus marquée. L'offre représente ici environ 0,87 % du total (là encore, hors ingénierie des données/architecture), alors que la demande représente 1,77 % du total. Voir la section suivante sur la demande pour en savoir plus.

¹⁵ L'enquête du Center for Security and Emerging Technology sur les docteurs en IA a montré que 73 % d'entre eux faisaient de la recherche appliquée dans le cadre de leur travail, contre 53 % pour la recherche fondamentale et 37 % pour les travaux d'ingénierie. Cependant, ils admettent qu'il s'agit peut-être d'un échantillon biaisé, 54 % des répondants étant issus du monde universitaire : « Une récente recherche de CSET analysant les parcours professionnels des diplômés de doctorat en IA des États-Unis issus des programmes les mieux classés entre 2014 et 2018 sur la base du codage du CV a révélé que 34 % d'entre eux travaillent dans le milieu universitaire et 60 % dans le secteur privé. »

Chercheur - 4 149 au total

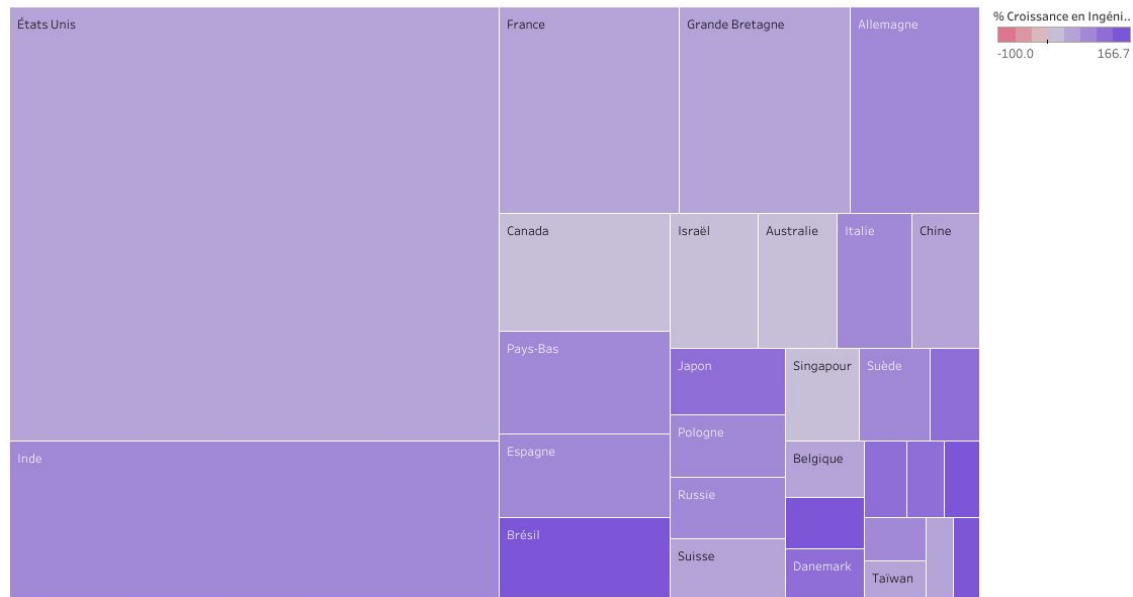


Certains pays à noter en termes de croissance des chercheurs sont la Corée du Sud avec une croissance de 133 %, Taïwan, l'Iran et l'Autriche avec des taux de croissance de 90-110 %. Israël, la Pologne, la Russie et la Grèce affichent également une croissance significative, entre 35 et 60 %. En revanche, l'Australie et la Suisse sont devenues négatives à -43 %. Des chiffres en baisses dans les pays en tête montrent potentiellement une tendance où les chercheurs bougent vers l'ingénierie et la mise en oeuvre, mais l'industrie de l'AI doit faire gaffe à ne pas sous-investir en recherche.

Nous avons constaté que le taux de croissance spécifique des développeurs d'IA (-68,8 %) a été l'exact inverse de celui des scientifiques de données (102,7 %). Cependant, les développeurs en IA et les scientifiques de données de nos jours ont souvent les mêmes compétences et rien n'empêche les développeurs en IA de simplement changer leur titre pour profiter de la rémunération plus élevée du scientifique des données qui joue un rôle d'ingénieur (environ 30 000 \$/an de plus en moyenne selon Glassdoor). Cette aptitude à changer de titre pourrait expliquer pourquoi ces deux rôles évoluent dans deux directions aussi diamétralement opposées. Cela n'est pas vrai pour un analyste de données, également dans la catégorie de la productisation, qui est beaucoup moins susceptible d'avoir ou d'avoir besoin de la même profondeur d'expertise technique pour son rôle. Cette catégorie pourrait aussi voir une croissance négative à cause de la tendance au développement de compétences pour devenir scientifique des données.

L'une des raisons pour lesquelles les compétences sont similaires est que les développeurs d'IA ont souvent besoin d'une bonne compréhension de l'ingénierie sous-jacente afin d'adapter les nouveaux outils d'IA aux logiciels. Le défi des problèmes à résoudre dans le développement de logiciels d'IA est la même raison pour laquelle l'IA a été largement inaccessible aux petites équipes qui n'ont pas les moyens de s'offrir à la fois des ingénieurs en AM et des développeurs spécialisés en IA. L'objectif de démocratiser l'IA avec des outils standards prêts à l'emploi pourrait permettre au développeur moyen ayant suivi une formation en ligne à l'IA d'intégrer l'IA dans son logiciel. Il s'agit d'une bonne chose pour les petites équipes, car cela leur permet de consacrer leur budget à la mise en oeuvre plutôt qu'à l'ingénierie, tout en étant capables de modifier les méthodes prêtes à l'emploi avec un certain degré de personnalisation.

Ingénieur IA/AM - 150 559 au total



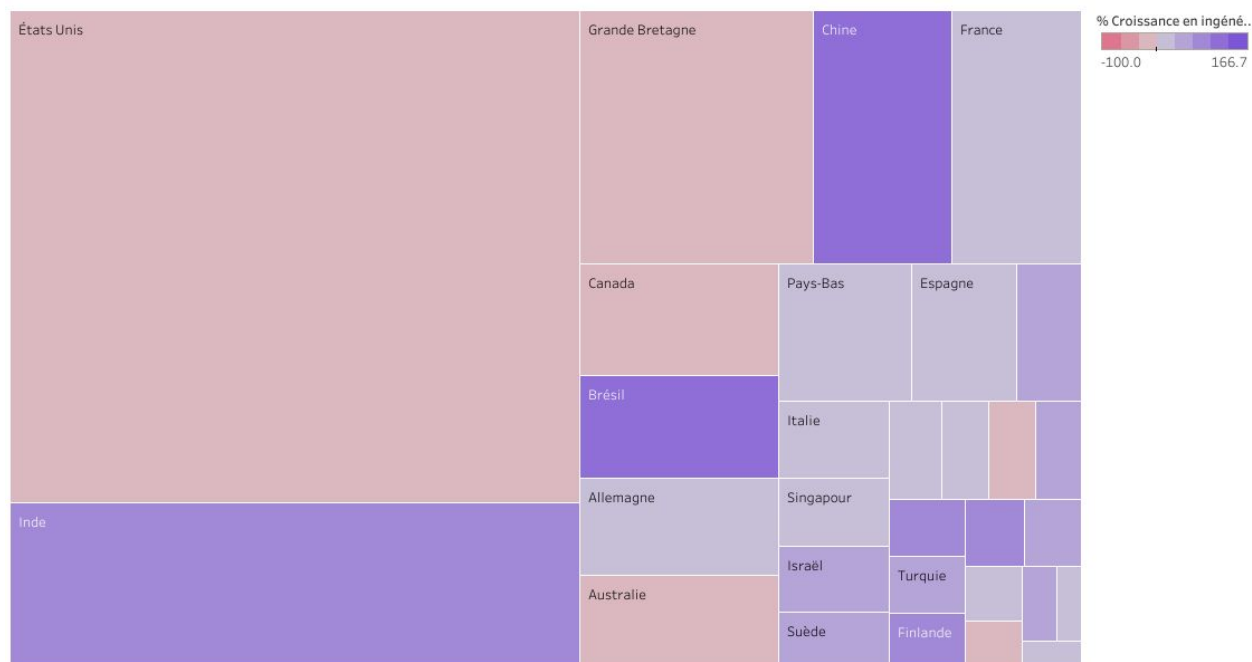
Productisation IA/AM - 250 500 au total



Une tendance à surveiller est celle des écosystèmes émergents comprenant un nombre croissant d'ingénieurs de données. Cela peut être le signe d'organisations qui ont mis au point des solutions d'IA, mais ayant ensuite rencontré le défi de les faire fonctionner à l'échelle sur des flux de données en direct. Ce n'est pas un défaut, car nous l'avons vu comme un signe de maturation vers des systèmes d'IA réellement intégrés. Ces pays sont le Brésil (110 %), la Finlande (87 %), l'Inde (83 %) et la Corée du Sud (94 %). Certains écosystèmes voient leur nombre d'ingénieurs des données diminuer, malgré une croissance de la catégorie en général. Ceci est potentiellement un signe de la délocalisation du travail ou

des employés, en plus d'une consolidation (voire même une automatisation) du travail dans [certaines compagnies qui utilisent l'IA pour compléter le travail](#).

Ingénierie et architecture des données - 72 748 au total

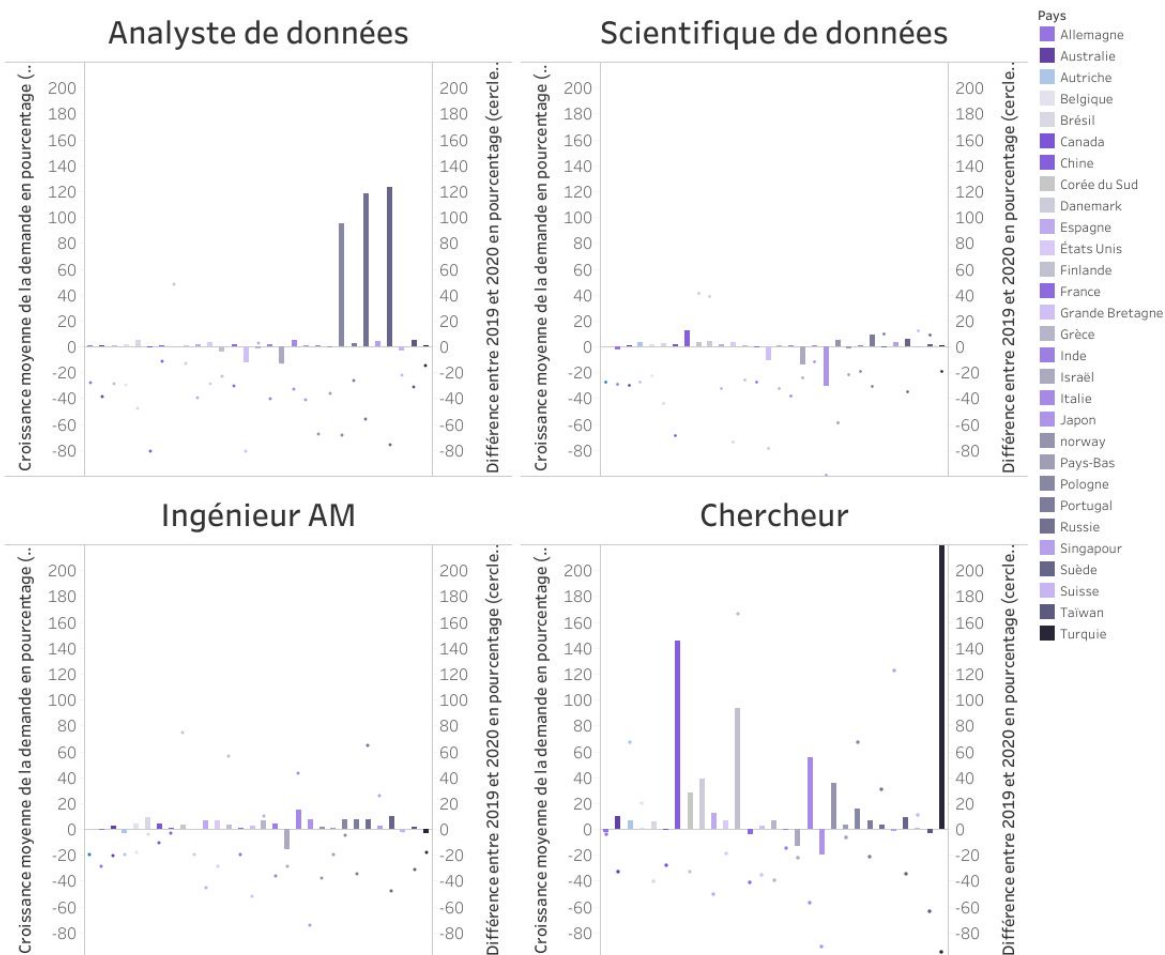


Demandes de l'industrie pour des talents

Une demande stable avant 2020, qui a chuté de 20 à 30 % pendant la pandémie. Quelques cas aberrants se sont distingués.

Pour mesurer un indicateur de la demande, nous avons recherché sur les sites d'agrégateurs d'emplois les titres de postes pertinents (par exemple « scientifique des données », « ingénieur en apprentissage machine ») sur une base mensuelle pour collecter le nombre d'offres d'emploi. Nous avons ensuite comparé la variation d'un mois à l'autre pour mesurer la croissance de la demande pour chaque titre de poste. Malheureusement, nous ne connaissons que la croissance nette, et nous ne savons donc pas combien d'offres d'emploi ont été fermées par rapport à celles qui ont été rouvertes, ni si les emplois non fermés ont en fait été pourvus, mais non retirés de la liste.¹⁶

¹⁶ Il convient de noter que les données relatives à la demande ne sont pas séparées dans les catégories que nous avons définies dans la section « Rôles techniques spécialisés », et ne concernent que le titre de poste spécifique mentionné dans chaque graphique, car nous ne disposons pas des chiffres de 2019



Croissance moyenne de la demande pour le titre du poste, basée sur la variance moyenne du nombre d'offres d'emploi avec des mots-clés apparentés d'un mois à l'autre.
La différence entre 2019-2020 est la différence entre la moyenne mensuelle de 2019 et la moyenne mensuelle de 2020.

Les tendances générales montrent que les taux de croissance médians pour les rôles d'analyste de données, de scientifique des données et d'ingénieur AM ont tous atteint une certaine stabilité en 2019, entre 1,24 % et 3,28 %. Les rôles de chercheurs en AM ont quant à eux un taux de croissance médian bien plus élevé, de 6,28 %. Si l'on ajoute à cela que les chercheurs représentent 1,77 % de la demande mesurée ici, alors qu'ils ne représentent comparativement que 1,02 % de l'offre de talents disponible pour l'industrie, on peut s'attendre à ce que la concurrence soit encore plus serrée pour les talents en recherche et que davantage de chercheurs soient attirés hors du milieu universitaire.

La Pologne, la Russie et la Suède sont des cas aberrants notables qui sont en plein essor (entre 95 % et 125 % de croissance) en ce qui concerne la croissance des analystes de données. Nous pouvons également observer des valeurs aberrantes importantes dans la demande de chercheurs en AM avec la Turquie (225 %), la Chine (145 %), la Finlande (166 %) et une poignée d'écosystèmes qui recherchent encore activement des chercheurs : L'Italie, le Danemark, la Norvège, la Corée du Sud, la Pologne et l'Espagne ont tous connu une croissance comprise entre 10 et 55 %.

pour tous les titres afin d'évaluer les taux de croissance. Il est également important de noter que « ingénieurs en apprentissage machine » et « scientifiques des données » correspondraient tous deux à la catégorie de rôle Ingénierie en IA/AM, et que nos données sur la demande ne couvrent pas la catégorie de rôle Ingénierie et architecture des données.

Ces chiffres peuvent signifier des changements de stratégie au palier local ainsi que l'arrivée de nouveaux venus sur la scène. L'explosion de la demande de chercheurs semble en partie due au fait que la Chine cherche à étendre son succès en invitant davantage de chercheurs universitaires (voir plus haut dans la section « Attraction et rétention des talents ») à d'autres chercheurs industriels, en profitant des nombreux ressortissants chinois aux États-Unis qui se heurtent soudainement à des problèmes de visa.

Ratio moyen de la demande pour chaque titre de poste

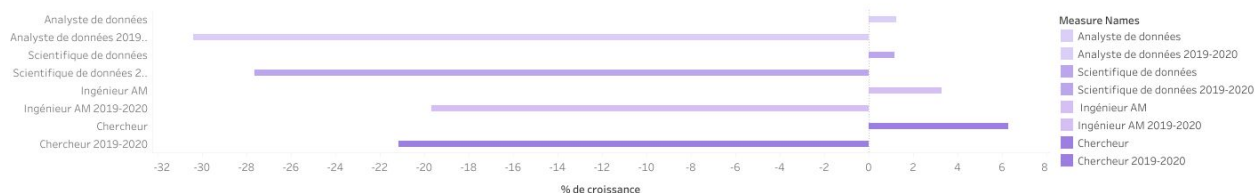


La différence de demande entre les scientifiques de données et les ingénieurs en AM montre qu'ils sont pratiquement les mêmes proportionnellement en termes d'offres d'emploi. Nous considérons qu'ils requièrent les mêmes compétences vu que le travail le plus important en science des données est en fait du travail en AM, donc les deux parties de notre catégorie de rôle d'ingénierie en IA/AM. Cependant, les offres d'emploi mensuelles pour « Ingénieur en AM » ont augmenté deux fois plus vite (3,28 %) que celles pour « Scientifiques des données » (1,17 %), ce qui indique que de nombreux pays ont l'intention d'accélérer l'innovation en AM et que les compétences en IA ne sont peut-être pas encore normalisées.

En examinant les données de 2020 jusqu'à présent, nous pouvons voir les conséquences de la COVID-19. Presque tous les écosystèmes ont connu un ralentissement drastique de la demande de talents en IA. En comparant les moyennes annuelles pour 2019 et 2020 jusqu'en août, on constate que les offres d'emploi pour les analystes de données ont ralenti de 30 %, les scientifiques de données de 27 %, les ingénieurs en AM de 20 % et les chercheurs de 21 %.

Certains écosystèmes ont pris le ralentissement comme une possibilité de se développer, avec des pays qui persistent dans leur nouvelle croissance, ou même qui sautent pour la première fois dans le domaine, pendant la pandémie. La Corée du Sud a augmenté ses embauches dans tous les domaines sauf celui des chercheurs en 2020. Singapour a fait de même et a augmenté progressivement ses embauches pour tous les postes et a même doublé le nombre de chercheurs (122 % de croissance en 2020), se rendant peut-être compte que l'offre ne s'accroissait pas sur l'île.

Croissance médiane de la demande



Croissance moyenne de la demande pour le titre de poste, basée sur la variance moyenne du nombre d'offres d'emploi avec des mots-clés apparentés d'un mois à l'autre. La différence entre 2019-2020 est la différence entre la moyenne mensuelle de 2019 et la moyenne mensuelle de 2020.

Les écosystèmes bien établis ont ralenti la demande de manière plus significative que les écosystèmes en développement qui font du rattrapage. Les premiers résultats du dernier sondage mondial sur l'IA de McKinsey, qui sera publié en novembre, montrent que les répondants d'une majorité d'organisations sont plus susceptibles d'augmenter leurs investissements en matière d'IA au cours des trois prochaines années que de les diminuer. Nous devons voir si cela va relancer la demande de talents ou si une plus grande partie de l'investissement sera consacrée aux infrastructures dans le cadre de la reprise.

Pour les grands acteurs qui veulent contrôler leur pile technologique et leurs données afin de mettre au point de nouvelles applications et solutions, il sera toujours important d'embaucher et de développer des talents en sciences des données. Mais à mesure qu'ils normalisent leurs propres outils, ils augmentent la demande pour des rôles de mise en œuvre plus importants. En outre, à mesure que des capacités prêtes à l'emploi seront disponibles, les personnes qui se trouvent plus en aval de la chaîne de valeur et qui travaillent à la mise en œuvre de modèles de travail seront essentielles et devraient s'attendre à une forte demande pour leurs compétences.

Conclusion : Comblé le fossé des talents

La question de savoir si les louanges du plein potentiel de l'IA sont trop chantées est [une autre discussion](#), mais nous pouvons dire que le succès pratique de l'IA n'est pas seulement une formule réunissant des experts de haut niveau et un accès aux bonnes données. L'industrie de l'IA s'est d'abord concentrée sur les experts de très haut niveau, car eux seuls pouvaient administrer les nouvelles techniques émergentes avec le savoir-faire adéquat pour les appliquer à des domaines nouveaux. Il est maintenant reconnu que la dynamique de cette nouvelle technologie nécessite plus que des ingénieurs et des personnes capables d'élaborer de beaux modèles pour la déployer efficacement.

L'IA est une nouvelle génération de logiciel qui s'adapte aux données qui lui sont fournies; elle est codée avec des données plutôt qu'avec des règles logiques. Les logiciels traditionnels sont statiques en comparaison, et l'IA a besoin d'un nouvel écosystème de soutien et d'infrastructure qui ne soit pas seulement élaboré, mais aussi géré une fois qu'il est déployé. Pour que l'IA fonctionne à l'échelle, il faut de nombreux nouveaux talents pour l'ingénierie, l'élaboration d'infrastructures, le développement de nouveaux modèles commerciaux et le suivi des objectifs.

La Commission européenne a récemment publié [un sondage](#) sur l'adoption de l'IA par les entreprises en Europe. L'obstacle numéro un identifié dans le sondage était de pouvoir embaucher du personnel ayant les bonnes compétences. Notre rapport mondial sur les talents en matière d'IA n'est qu'un premier aperçu de ces rôles professionnels et de la manière dont l'inadéquation entre l'offre et la demande peut perdurer. Il reste beaucoup à faire pour dresser un tableau complet de la situation. Nous suggérons un certain nombre d'idées dans l'annexe « Pour la recherche future » afin d'ajouter de la profondeur et d'aider à créer une vision plus claire des personnes qui ont le talent nécessaire pour fournir l'IA, et où elles peuvent être trouvées.

Au fur et à mesure que l'IA se développe, elle deviendra de plus en plus prévalente. Nous verrons de nouveaux rôles spécialisés émerger pour gérer la nouvelle dynamique de l'IA, mais à terme, chacun devra mettre à jour ses compétences numériques pour collaborer avec cette nouvelle technologie. Nous avons déjà vu que la plupart des gens peuvent saisir le concept d'un algorithme de recommandation alimenté par l'IA et ajuster leur comportement pour influencer ses résultats. Cependant, les gens ont un choix très limité et ne disposent que d'outils rudimentaires pour manipuler un algorithme selon leurs besoins. Lorsque les différents outils et compétences seront normalisés tout au long de la chaîne de valeur, cela augmentera considérablement le choix et l'accès à la technologie d'IA et engendrera nettement plus d'innovation que ce que nous avons déjà vu avec les logiciels d'IA.

Pour y parvenir, nous devons relever le défi de combler le fossé entre la preuve de concept en laboratoire et le déploiement dans le monde réel. Les chercheurs et les ingénieurs jouent actuellement un rôle important pour aider à combler ce fossé, mais ils ne peuvent le faire seuls. Ils doivent, ainsi que les institutions qui les forment, se concentrer sur la normalisation de leurs outils et de leurs processus afin que d'autres puissent plus facilement collaborer avec d'autres en aval de la chaîne de valeur.

Annexe 1 : Pour la recherche future

Recherche d'ensembles de compétences plutôt que de titres

Une mise à jour essentielle pour mieux comprendre l'offre et la demande consistera à adopter une approche basée sur les compétences plutôt que sur les titres. Les titres sont susceptibles de changer à mesure que les rôles se normalisent et que de nouveaux rôles apparaissent. Le suivi des combinaisons de compétences permettra d'identifier plus rapidement la manière dont ces rôles prennent forme. De plus, l'obtention des données nécessaires à cette fin ajoutera probablement une plus grande granularité qui permettra le type d'analyse que nous faisons sur les données de recherche, comme la mesure du mouvement géographique et du genre.

Mieux comprendre la relation entre l'industrie et le monde universitaire

Il est difficile d'évaluer avec précision les conséquences et les bienfaits de l'industrie sur la recherche. Nous savons qu'elle trop recruter, laissant les établissements à sec et incapables de former de nouvelles générations. Mais elle peut aussi générer beaucoup de travail productif et utile en donnant aux chercheurs l'accès à des données et des problèmes du monde réel. Pour mieux répondre à cette question, quelques ajouts simples seraient les suivants :

- Durée moyenne des activités de publication d'un chercheur
- Si les affiliations sont des entreprises ou des universités
- Séparer la recherche en recherche appliquée et fondamentale et, si possible, en sous-catégories de chacune d'entre elles
- Meilleure compréhension des mouvements par genre, par les professeurs invités et par les chercheurs dans les laboratoires privés

Compréhension empirique des différentes régions

L'intégration des données d'enquête provenant de différentes localités serait un complément important pour comprendre les histoires qui se cachent derrière les données, ou lorsque les données présentent un angle mort. La Chine est probablement très sous-représentée dans ce rapport, tout comme d'autres pays qui publient dans des alphabets non latins ou qui sont simplement à un stade de maturité plus précoce. Même dans les écosystèmes développés, les données régionales aideront à comprendre la répartition des laboratoires privés et le poids des différents centres.

Annexe 2 : Méthodologie et mises en garde

Typologie et médias sociaux

Nous avons consulté des experts de l'industrie au sein d'Element AI et à l'extérieur sur leur expérience des projets liés à l'IA. À partir de ces conversations, nous avons décomposé les types d'expertise nécessaires pour faire passer l'IA des idées théoriques aux produits concrets, ce qui nous a conduits à une série de titres et de descriptions de postes qui forment notre hypothèse de travail sur l'expertise en IA. Si la réalité sur le terrain tend à avoir des frontières plus floues, nous pensons que ces catégories représentent bien les différentes compétences nécessaires. En utilisant les titres des postes, nous avons ensuite recueilli des données sur l'autoreprésentation dans les médias sociaux. La collection était basée sur des mots-clés trouvés dans les titres de postes (par exemple « scientifique de données » comme titre de poste), dans les listes de compétences (par exemple « apprentissage machine ») et dans le savoir-faire technologique (par exemple « tensorflow »).

Demande d'emploi

Pour mesurer un indicateur de la demande, nous avons recherché sur les sites d'agrégateurs d'emplois différents titres de postes (par exemple « scientifique de données », « ingénieur en apprentissage machine ») sur une base mensuelle pour collecter le nombre d'offres d'emploi. Nous avons ensuite comparé la variation d'un mois à l'autre pour mesurer la croissance de la demande pour chaque titre de poste.

arXiv et données démographiques

L'année dernière, nous avons concentré la recherche sur les conférences scientifiques les plus prestigieuses autour de l'IA, dans le but de comprendre les recherches les plus percutantes. Cette année,

nous avons élargi notre regard sur la recherche en essayant de nous rapprocher le plus possible d'un « recensement de la recherche » qui nous permettrait de mieux comprendre comment se déroule la recherche sur l'IA : d'un petit cercle de chercheurs travaillant à rendre l'IA réelle, à une constellation d'experts trouvant des moyens de mettre en œuvre l'IA dans différents domaines. En d'autres termes, cette nouvelle méthodologie devrait nous permettre de ne pas nous concentrer uniquement sur la « recherche fondamentale en IA » et d'en élargir le champ d'application pour englober davantage de recherche appliquée en IA.

La comparaison du rapport de l'année dernière avec celui de cette année met en évidence le volume de la recherche universitaire effectuée autour de l'IA et de ses applications qui ne se retrouve pas dans les conférences. Dans notre nouvelle approche, l'industrie a moins d'influence sur les chiffres globaux.

Nous avons téléchargé tous les documents des dépôts cs.AI, cs.LG et stat.ML, car ce sont ceux qui ont la relation la plus directe avec l'IA et qui ont le moins de probabilités de contenir des documents sans rapport avec l'IA. Dans ces documents, nous avons conservé tous ceux qui contenaient un fichier lisible au format Latex, PDF, Tex ou autre, en excluant les autres qui ne contenaient que des fichiers HTML ou d'autres fichiers de langage de codage. À partir de là, en raison de limitations techniques, nous n'avons conservé que les documents en caractères latins, en excluant certains documents comme ceux écrits en caractères chinois.

Avec cet échantillon, nous avons créé une heuristique pour extraire les noms et affiliations des auteurs de chaque article. L'heuristique a été créée en passant en revue plus de 3 500 articles et en créant une liste de toutes les affiliations possibles trouvées dans les documents. De cette liste d'affiliations potentielles, nous avons extrait les affiliations des journaux avec un script regex. Cela nous a laissé 35 418 articles qui contenaient des affiliations reconnues par notre script. À partir de la liste d'auteurs/d'affiliations créée, nous avons ensuite défini l'affiliation la plus probable statistiquement pour chaque auteur pour chaque année en fonction des affiliations récurrentes dans chaque article que l'auteur avait. Cela nous a permis de connaître les affiliations probables de chaque auteur pour chaque année de publication. Cette méthodologie, testée sur un échantillon de 200 articles, a donné un taux de réussite de 98,7 %.

La localisation de chaque affiliation a été établie à la main et repose sur la même logique que les rapports précédents : les systèmes universitaires (par exemple, le réseau de l'Université de Californie [UC] ou le réseau de l'UParis) ont été simplifiés (c'est-à-dire que toutes les universités UC sont comptées comme étant simplement en Californie), et toutes les entreprises sont comptées comme étant situées à leur siège (par exemple, tous les laboratoires Google sont comptés comme étant en Californie). Cette logique repose sur deux considérations : la première étant de nature pragmatique dans le sens où, par exemple, Google possède des laboratoires dans le monde entier, mais les auteurs ont rarement précisé dans lequel ils travaillent; et la seconde est plus pour répondre à un point sous-jacent du rapport où nous postulons que la plupart des laboratoires ne conservent pas la valeur de leurs recherches dans le pays d'accueil du laboratoire, mais les ramènent au siège de l'organisation.

La mesure du genre a été basée sur les noms des auteurs. En utilisant les données du recensement américain, nous avons créé une liste de probabilités pour chaque nom et avons conservé ceux qui avaient une probabilité de plus de 95 % d'être d'un genre plutôt que de l'autre (sur la base de la logique de ce document : <https://cran.r-project.org/web/packages/gender/gender.pdf>).

L'afflux de talents a été mesuré en comparant la probabilité qu'un auteur se déplace vers et depuis un pays, par rapport à la probabilité moyenne. L'invitation est la mesure du risque que les auteurs se retrouvent dans un pays, tandis que le séjour/le départ est la mesure du risque qu'un autre déménage ou ne déménage pas au cours d'une année. La moyenne a été calculée en faisant la moyenne des probabilités et en la comparant à la moyenne de toutes les probabilités moyennes.

Remerciements

Écrit avec Simon Hudson et Yoan Mantha.

Recherche et visualisation des données par Yoan Mantha.

Remerciements spéciaux aux collègues chez Element AI et ailleurs qui ont fourni leur support et leurs précieux commentaires :

Alexandra Mousavizadeh of [Tortoise Media](#)

Bruno Lanvin of the [Portulans Institute](#)

Helen Mullings, Paul Walsh and Andrew Fyffe of [Quantum Black](#)

Michael Chui and David DeLallo of [McKinsey Global Institute](#)

Time Davies and Nicolás Grossman of the [Global Data Barometer](#)

et

Frédérique Bouchard, Valérie Bécaert, Simon Bélanger, Annabelle Martin, David Bédard, Catherine Lefebvre, Jean-Philippe Reid, Adam Salvail, Lara O'Donnell, Pierre-Luc Beaubien, Benoit Hamelin, Julien-Pier Boisvert, Christian Jauvin

Merci à l'équipe qui a supporté la distribution de ce rapport : Kayla Gillis, Morgan Guegan, Guillaume Gagnon, Kevin Clark, Julien Desrosiers, Marie-Calude Savard, and Robyn Crump

Traduction en français par Melissa Guay and Guillaume Gagnon