GOUVERNEMENT
Liberté
Égalité
Fraternité

HUB FRANCE IA

FRANCE 2030

« REFCO-IA Phase 1 »

Premier levier des transitions numériques et écologiques, la formation des jeunes et des salariés permet de renforcer le capital humain indispensable au fonctionnement de nos entreprises et au-delà de toute la société. C'est aussi le meilleur moyen pour proposer des emplois durables et de tous niveaux de qualification sur l'ensemble du territoire.

C'est également une des conditions majeures pour la réussite du plan France 2030 : soutenir l'émergence de talents et accélérer l'adaptation des formations aux besoins de compétences des nouvelles filières et des métiers d'avenir. 2,5 milliards d'euros de France 2030 seront mobilisés sur le capital humain pour atteindre cette ambition.

L'appel à manifestation d'intérêt « Compétences et métiers d'avenir » s'inscrit dans ce cadre et vise à répondre aux besoins des entreprises en matière de formations et de compétences nouvelles pour les métiers d'avenir.

Dans le cadre de ce dispositif, la réalisation de diagnostics des besoins en compétences et en formations sont financés et diffusés.

Sommaire

1– Introduction et méthodologie de travail 3

2– L1.1 - Le Référentiel de compétences v0 5

3– L1.2 - Le questionnaire 5

4– L2 - Questionnaires et comptes-rendus : l'enquête 6

5– L3 - Le Référentiel de compétences v1 8

6– Impacts sur les formations 10

7– Conclusions et préconisations 13

8– Annexes 14

1. Introduction et méthodologie de travail

L'appel à manifestation « Compétences et métiers d'avenir » de France 2030

L'appel à manifestation d'intérêt (AMI) « Compétences et métiers d'avenir » s'inscrit dans le cadre du volet dirigé du PIA4 et vise à répondre aux besoins des entreprises en matière **de formations et de compétences nouvelles pour les métiers d'avenir.** Les projets soutenus portent notamment sur :

- La réalisation de diagnostics des besoins en compétences et en formations ;
- L'identification des initiatives et projets en rapport avec une stratégie ou plusieurs stratégies nationales ;
- Le financement des projets les plus adaptés qui auront été sélectionnés par une procédure exigeante.

Le projet REFCO-IA est l'initiative d'un consortium de onze acteurs (cf Annexe 1); il est piloté par le Hub France IA. Il s'inscrit dans une démarche de soutien de la compétitivité des entreprises françaises dont les évolutions technologiques entraînent de nouveaux besoins en compétences techniques. La pénurie de talents qui sévit dans les métiers de l'intelligence artificielle (IA) conduit les entreprises à adopter des stratégies d'upskilling[1] et reskilling[2], voire de reconversion, de leurs effectifs bien formés à leur métier. Les nouvelles technologies évoluant à une vitesse fulgurante, elles nécessitent un besoin quasi-constant d'actualisation des compétences. Sans une visibilité claire des compétences requises pour exercer des tâches et missions techniques en IA, les entreprises font face au manque d'une information cruciale. En conséquence, elles tentent d'établir des référentiels de compétences internes qui ne permettent pas d'avoir une vue d'ensemble nationale et multisectorielle.

Les objectifs de REFCO sont ainsi de : 1 / produire un diagnostic « universel » des besoins en compétences IA de l'ensemble des entreprises françaises ; 2/ évaluer les volumétries de collaborateurs devant acquérir ces compétences et 3/ produire un référentiel de compétences unifié, sans viser une branche professionnelle ou un secteur d'activité en particulier.

Il est précisé que lors de son dépôt, le projet REFCO-IA avait été imaginé en coordination avec le projet de diagnostic des formations IA ayant fait également l'objet d'un dépôt par l'Institut Polytechnique de Paris. Ce dernier n'ayant pas été retenu, notre consortium a produit ses livrables sans pouvoir s'appuyer sur un diagnostic exhaustif des formations IA existantes.

- Temporalité du projet

Le projet a débuté le 17 novembre 2022 pour finir au 25 mai 2023. Une phase d'enquête (questionnaire et entretiens) de 3 mois (de mi-décembre à mi-mars 2023) a été suivie d'une période de consolidation et de rédaction des livrables.

Méthodologie générale

- **Étapes du projet et indicateurs de performances**

Pour répondre aux objectifs du projet, le consortium s'est réuni à une fréquence bimensuelle durant la durée du projet afin d'en assurer le bon déroulement. Pour se faire, 4 tâches clefs ont été identifiées :

1. Faire un état des lieux des référentiels de compétences existants chez les partenaires ou en ligne et les agréger pour créer une version zéro (v0) du référentiel REFCO-IA ;
2. Établir un questionnaire basé sur le référentiel v0 ;
3. Mener une enquête terrain auprès des entreprises françaises au travers du questionnaire et d'entretiens. Dans notre proposition, nous avions fixé l'objectif de réaliser environ 160 questionnaires et 100 interviews répartis comme suit :

	Grandes entreprises	ETI	PME	Startups	Total
Questionnaire	37	38	35	53	163
Interviews	32	20	27	22	101
Total	**69**	**58**	**62**	**75**	**264**

4. Consolider les réponses recueillies durant la phase d'enquête et en déduire une mise à jour du référentiel REFCO-IA (v1) ;

Pour répondre à une demande exprimée par la CNIA à mi-parcours, nous avons ajouté la prise en compte du besoin d'identifier les besoins en formation au niveau national à partir des besoins en compétences exprimés par les entreprises. Comme indiqué dans la proposition, le projet REFCO vise un diagnostic de compétences et ne s'intéresse pas aux formations : nous avons donc proposé d'utiliser des « persona » pour obtenir une estimation de ces besoins de formation (voir plus loin). Vu les délais, il s'est avéré impossible de modifier le questionnaire pour mieux cerner ces besoins, l'estimation obtenue n'est donc qu'indicative.

- **Livrables attendus**

[1] *Upskilling : montée en compétence.*

[2] *Reskilling : acquisition de compétences nouvelles dans le cadre d'un changement de métier ou d'une montée en compétence transversale.*

Lors de la réponse à l'AMI, 4 livrables ont été identifiés. Parmi ceux-ci, les livrables L.1.1 Référentiel de compétences v0 et L.1.2 questionnaire fournissent les outils pour l'enquête.

L1.1 : Référentiel de compétences v0

Le livrable L1.1 a été produit à partir des référentiels existant chez les membres du consortium et sur France Compétences. Ces documents étant des référentiels métiers – et non des référentiels de compétences – l'ensemble des partenaires a dû fournir un travail d'analyse et de restructuration pour aboutir à la version 0 du référentiel (disponible en annexe 2). L'objet de ce référentiel a été limité aux compétences purement IA, en excluant les compétences plus générales (mathématiques, statistiques, ou *soft skills*) : un métier devra alors être caractérisé par ses compétences générales en plus des compétences IA requises.

L.1.2 : Questionnaire

Sur la base du livrable L.1.1, le consortium a élaboré un questionnaire en 4 parties, composé de questions fermées, sauf exception, pour permettre une consolidation aisée des réponses. Le questionnaire a été élaboré directement à partir du référentiel v0, en lui ajoutant simplement une partie préliminaire pour recueillir des informations générales sur l'entreprise répondante et une conclusion. Après un *benchmark* de comparaison de divers outils d'enquête du marché, nous avons décidé d'utiliser l'outil *Limesurvey*[3] qui seul nous permettait de mettre en œuvre la complexité du questionnaire tout en préservant un hébergement des données en Europe :

Critères	Typeform	Eval&Go	Limesurvey	EU Survey
Score	X	X	X	no
Sauts logiques selon réponses	X	X	X	X
Sauts logiques selon scores	X	no	X	no
Hébergement des données	USA	FR	DE	EU
Reporting	X	X	X	X
Nombre de questions par mois	1 000	Non limité	1 000	?
Choix	Non	Non	Oui	Non

Le questionnaire est disponible en ligne[4]. La phase d'enquête s'est déroulée sur 3 mois (du 17/12/2022 jusqu'au 17/03/2023), après la phase de création du référentiel v0 et du questionnaire.

L2 : Questionnaires et comptes-rendus

Le questionnaire a été envoyé par les partenaires à leurs réseaux de contacts (après dédoublonnage entre les listes de contacts pour éviter les contacts en doublon). Nous avons cherché à obtenir un échantillon représentatif des entreprises sollicitées (grands groupes, PME/ETI, startups), toutefois, le secteur d'activité du numérique est, bien naturellement sur-représenté.

Enfin, au vu de la difficulté à inciter les décideurs à répondre au questionnaire, nous avons décidé d'allonger la phase d'enquête de deux mois en proposant des interviews ciblées permettant d'accompagner le répondant : cette stratégie a permis d'augmenter le nombre de répondants, mais aussi de recueillir en direct des commentaires et suggestions très utiles pour réviser le référentiel v0.

L3 : Référentiel de compétences v1

Le questionnaire a été un support autoporteur pour les professionnels répondant en autonomie. Il a également servi de base d'échange lors des entretiens réalisés. Les éléments issus de l'enquête - questionnaire et entretiens confondus - nous ont permis d'améliorer le référentiel v0 pour produire un référentiel v1 avec une structure améliorée, mieux adaptée aux retours des professionnels et plus pérenne. Des travaux d'analyse (exploitant des technologies d'IA) ont permis d'établir la validité de ce référentiel.

L4 : Préconisations

Sur la base de l'analyse des résultats de l'enquête et de l'analyse du référentiel v1, nous avons élaboré des préconisations présentées en section 7.

Nous avons également établi des prévisions des besoins en formation sur la base de « persona » décrites en section 6.

Nous détaillons dans les sections suivantes les travaux réalisés et les résultats obtenus.

2. L1.1 - Le Référentiel de compétences v0

Le référentiel v0 (voir annexe 2) a été élaboré à partir de référentiels de compétences disponibles chez les partenaires (Numeum, La Poste, Société Générale, Business & Décision), surtout orientés métiers et non compétences, sauf celui de Business &

[3] https://www.limesurvey.org/fr/

[4] Disponible à l'URL : https://hfia.limesurvey.net/784422?lang=fr et en Annexe 3

Décision que nous avons donc majoritairement utilisé. Nous l'avons ensuite complété de compétences qui nous ont semblé manquer et que nous avons identifiées dans France Compétences (RNCP et RS pour Intelligence Artificielle et Data Science).

Il faut remarquer que ce rapport n'aborde pas la question des **pré-requis** et de la formation initiale des personnes demandant une formation dans un profil de métier. En particulier, certaines compétences du référentiel supposent des bases scientifiques solides et des compétences techniques en informatique et/ou en statistique de niveau ingénieur. Une culture de base en IA est indispensable, et pour des profils plus techniques des connaissances plus avancées en informatique (génie logiciel, architecture, infrastructure, calcul haute performances, …) ou pour les profils orientés données les outils de base de la statistique descriptive et de la statistique inférentielle. Enfin, ce référentiel de compétences doit être envisagé dans une **dynamique de formation**. En effet, les progrès scientifiques dans le domaine de l'IA vont nécessiter une adaptation continue aux nouvelles technologies. Les compétences proposées devraient être envisagées dans le long terme et fournir un socle solide pour permettre l'évolution des métiers.

Notre référentiel a donc été limité aux compétences purement IA, en excluant les compétences plus générales (informatique, mathématiques, statistiques, ou *soft skills*) : un métier devra être caractérisé sur la base de ses compétences générales en plus des compétences IA requises, décrites dans le référentiel.

Il est précisé que les compétences sont structurées en macro-compétences décomposées en sous-compétences.

- Exemple : la macro-compétence « savoir expliquer les bénéfices d'un projet IA aux interlocuteurs métier » est décomposée en 3 sous compétences : « connaître les bénéfices et les impacts de l'IA », « être capable d'identifier les sujets relevant de l'IA », « savoir scénariser des implémentations possibles de cas relevant de l'IA ».

Nous avons distingué deux types de compétences IA :

- **Compétences transverses** : elles regroupent les compétences nécessaires aux tâches préliminaires au projet (faisabilité du besoin métier), au pilotage du projet et à la conduite du changement pour accompagner le déploiement du projet. Elles représentent 10 macro-compétences et 36 sous-compétences ;
- **Compétences techniques** : elles ont été regroupées en 5 catégories (architecture, données, modélisation, industrialisation et transfert au métier), représentant 19 macro-compétences et 84 sous-compétences.

Le référentiel v0 comprend donc en tout 29 macro-compétences et 120 sous-compétences. Ce niveau de détail a été retenu pour permettre la réalisation d'une enquête exhaustive et précise sur le besoin en compétences IA des entreprises. L'inconvénient de cette approche est bien sûr la longueur du questionnaire qui en est résulté.

3. L1.2 - Le questionnaire

Le questionnaire (voir Annexe 3) utilisé pour l'enquête visant à recueillir les besoins en compétences IA des entreprises françaises est basé sur le Référentiel v0. **Tous les résultats présentés font donc référence au Référentiel v0**. Le questionnaire est organisé sur trois niveaux :

- Niveau 1 : Volumétrie de collaborateurs concernés pour chaque macro-compétence ;
- Niveau 2 : Niveau d'expertise attendu pour chaque sous-compétence ;
- Niveau 3 : Délais sous lequel chaque sous-compétence est attendue.

Les 4 parties du questionnaire sont les suivantes :

1. **Informations générales** : il s'agit de recueillir des informations générales sur l'entreprise répondante ; les questions sont fermées pour pouvoir facilement agréger les réponses :
   - Le type : start-up, PME/ETI, grand groupe ;
   - La taille : <5, 5 à 10, 10 à 50, 50 à 500, 500 à 1 000, > 1 000
   - Le secteur d'activité : Aéronautique & Défense, Tourisme, etc. ;
   - La maturité en matière d'IA : en quelques questions, il s'agit d'évaluer la maturité de l'entreprise, ce qui conditionnera l'apparition de certaines questions techniques.
2. **Compétences transverses** : ce sont celles du référentiel v0, complétées de questions, déclenchées par la réponse « oui » à la question « Avez-vous besoin de monter en compétence sur … », sur :
   - La volumétrie de personnes concernées ;
   - Le niveau d'expertise attendu pour chaque sous-compétence ;
   - Le délai sous lequel les sous-compétences sont attendues (avant ou après 1 an) ;
   - La source où rechercher les sous-compétences (en interne à l'organisation ou via un prestataire externe).
3. **Compétences techniques** : comme pour les compétences transverses ;
   Les informations sur les nombres de collaborateurs concernés permettent d'établir les statistiques de besoins par compétence présentées en Annexe 6.
4. **Conclusions** : l'entreprise est invitée à partager ses coordonnées pour recevoir le résultat de l'étude, sinon le questionnaire est anonyme. La plupart des entreprises ont donné leurs coordonnées.

4. L2 - Questionnaires et comptes-rendus : l'enquête

Le tableau ci-dessous présente le nombre d'entreprises ayant répondu à la phase d'enquête par typologie.

	Grandes entreprises	PME/ETI	Startups	Total
Questionnaire	20	35	40	95
Entretiens	22	12	14	48

La phase d'enquête étant le cœur du projet, le consortium a veillé à solliciter des profils pertinents au regard d'un diagnostic de compétences IA et de cibler des entreprises diversifiées, aussi bien fournisseurs qu'utilisateurs, de toutes tailles et tous secteurs. Environ 45% des entreprises sollicitées ont répondu à l'enquête, attestant de la qualité des réseaux du Consortium. L'ensemble des graphiques présentés dans cette section sont extraits des réponses aux entretiens et questionnaires **basés sur la version v0 du référentiel** tel que précisé plus haut.

Au cours de l'enquête, les questionnaires remplis (en direct ou pendant les entretiens) ont été intégrés dans l'outil *Limesurvey*. Des *reportings* réguliers ont été extraits automatiquement grâce aux fonctionnalités de l'outil et les réponses, ainsi agrégées, ont finalement été exportées dans un fichier Excel qui a permis de générer les graphiques présentés ici ou dans les annexes.

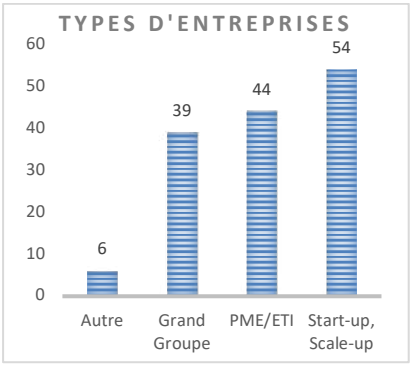
De plus, au cours des entretiens, les entreprises qui le souhaitaient ont pu nous communiquer des compétences / sous-compétences qui leur semblaient manquer ou nous faire des commentaires plus généraux (voir « questions finales » dans l'annexe 3). Ces éléments ont été consolidés dans le fichier Excel et seront utilisés pour la production du référentiel v1 (section suivante).

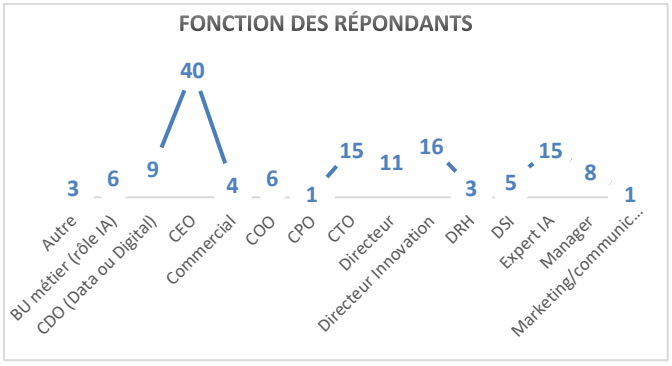
Répartition des questionnaires et entretiens au sein de l'enquête

La figure montre une répartition équilibrée entre questionnaires et entretiens, avec 67% pour les premiers et 33% pour les seconds. Ce résultat correspond aux objectifs fixés lors du début de l'enquête. La part importante d'entretiens a été un atout pour construire la version v1 du référentiel de compétences, puisqu'il a été possible d'interroger plus en profondeur l'interlocuteur dans l'entreprise.

Typologie des répondants : types d'entreprises et fonctions

La figure ci-dessous (gauche) montre un équilibre dans la typologie des répondants entre grands groupes, PME/ETI et start-ups/scale-ups, démontrant la bonne représentativité des besoins des entreprises issues de l'enquête. Le graphique de droite présente la diversité des fonctions des répondants avec une forte représentation des CDO[5] et CEO (pour les start-ups), ce qui provient de la nécessité d'avoir un bon niveau de compréhension de l'entreprise pour pouvoir répondre au questionnaire (et notamment concernant les besoins chiffrés).





Typologie des répondants : Secteurs d'activité, fournisseurs / utilisateurs d'IA, maturité

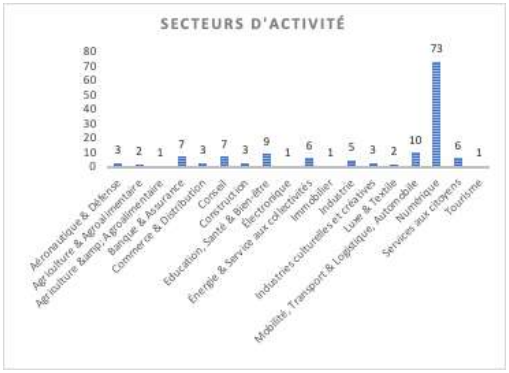
Les ambitions initiales de cibles d'entreprises répondant à notre questionnaire n'ont pas été atteintes, malgré plusieurs relances. Ceci s'explique par plusieurs difficultés que nous avons rencontrées :

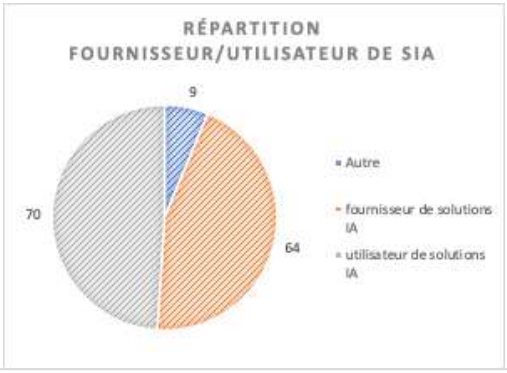
- Le questionnaire était long à compléter (environ trente minutes) et nécessitait pour le collaborateur y répondant un niveau de connaissance détaillé des besoins de l'entreprise ;
- De nombreuses entreprises sollicitées ne se sentent pas concernées par le projet car elles ont positionné leurs enjeux stratégiques à d'autres niveaux ;

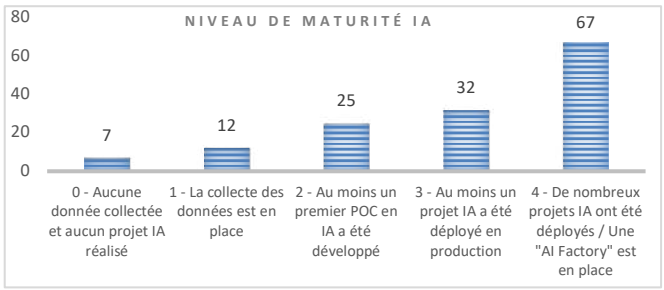
[5] *CDO : Chief Digital Officer, directeur de la transformation digitale.*

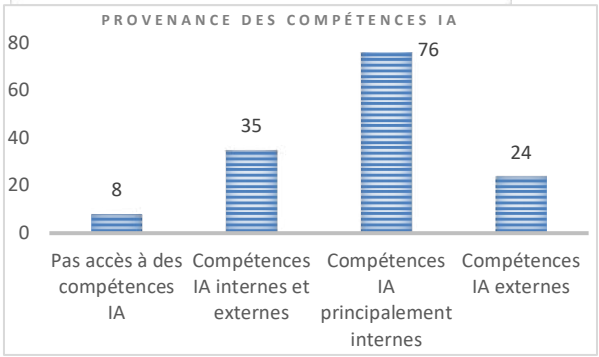
- De nombreuses entreprises sollicitées n'ont pas entamé ce type de réflexion en interne et se sentent démunies pour répondre.

Comme on le voit dans les figures ci-dessous, les entreprises du numérique représentent 51% des répondants, et sont à parts presque égales fournisseur d'IA et utilisateurs de solutions IA. Près de 70% ont mis en production au moins un projet IA, et sont donc matures en IA. Enfin, l'enquête a montré que la majorité des entreprises interrogées possèdent des compétences en IA avec une tendance marquée pour l'internalisation de ces compétences :









Les difficultés de recrutement indiquées précédemment tiennent dont très certainement au manque général de maturité IA des entreprises en France, ce qui a eu pour conséquence une **sur-représentation du secteur d'activité du numérique** parmi les répondants, beaucoup plus matures et très au fait des problématiques décrites dans le questionnaire. En effet, les représentants de ces entreprises, lors des entretiens, nous ont témoigné l'intérêt qu'ils portaient aux réflexions amenées par ce questionnaire. Ce dernier leur a permis de faire le point sur leurs besoins en compétences de manière détaillée, ce qui a été très apprécié. De même, interrogés sur le niveau de détail du référentiel, ils ont indiqué que la granularité était adaptée aux besoins du terrain. Le retour de toutes les personnes interrogées a été le même : **ce travail de référentiel est nécessaire et le niveau de détail proposé est pertinent et utile.**

<u>Conclusions générales de l'enquête</u>

Les résultats de l'enquête démontrent que les compétences transverses sont celles pour lesquelles les entreprises expriment le plus de besoins de montée en compétences (voir annexe 6). Parmi celles-ci, les 5 macro-compétences les plus mentionnées sont les suivantes :

- Connaître les enjeux de conformité et réglementaire de l'IA pour 48,25% ;
- Piloter des projets data et IA pour 42,66% ;
- Connaître les problématiques éthiques de l'IA pour 41,96% ;
- Savoir expliquer les bénéfices d'un projet IA aux interlocuteurs métier pour 38,46% ;
- Savoir mettre en place une approche sobre de l'IA pour 37,76%.

Pour les compétences techniques, les entreprises indiquent leurs priorités :

- Restituer aux interlocuteurs métiers les résultats des modèles pour 28,67% est la compétence la plus demandée ;
- Programmer les indicateurs nécessaires au fonctionnement des algorithmes pour 27,27% ;
- Concevoir et réaliser les infrastructures big data pour 26,57%
- Augmenter les données pour 26%
- Identifier et sélectionner des algorithmes de base à utiliser pour la modélisation pour 25%
- Maintenir les modèles et algorithmes en état de fonctionnement pour 25%

Notons, cependant, que ces pourcentages sont beaucoup plus faibles que pour les compétences transverses prioritaires (aux alentours de 50%), ceci reflète certainement la maturité des répondants : ils ont déjà mis en place les compétences nécessaires critiques en IA et ne font qu'envisager les formations complémentaires ou les nouveaux recrutements.

Cette liste met en exergue les besoins en compétences critiques actuels des entreprises, pour lesquels elles estiment ne pas avoir les ressources suffisantes en interne. Ces résultats sont cohérents avec les évolutions actuelles qui invitent les acteurs de l'IA à

respecter un cadre de plus en plus réglementé, respectueux de l'environnement et sensible aux questions éthiques.

Dans le même temps, le lien entre les équipes métier et les équipes techniques est un point clé pour assurer une réalisation et une industrialisation réussie d'un projet IA, ce qui explique la présence des macro-compétences transverses 2 et 4 de la liste ci-dessus. En complément de ce premier niveau d'analyse, nous voyons que la première macro-compétence technique la plus recherchée – qui arrive en 9ième position du classement global (voir annexe 7) – est « Restituer aux interlocuteurs métier les résultats des modèles ». Elle concerne 28,67% des répondants. Cela démontre encore une fois le niveau de criticité de cette capacité à faire le lien entre les différentes parties prenantes d'un projet IA et est à mettre en regard des déclarations des répondants (annexe 5) indiquant qu'ils ne disposent pas de profil conduite du changement ou AMOA.

Il est à noter que pour 100% des compétences transverses et techniques, les entreprises témoignent majoritairement d'un besoin de montée en compétence pour une population inférieure à 5 personnes, les personnes identifiées par les répondants ayant plutôt des profils d'experts.

Ces éléments démontrent la maturité des entreprises interrogées en ce qui concerne les compétences déjà détenues en interne. La plupart ayant indiqué que de nombreuses compétences étaient déjà maîtrisées par une population d'experts. Cela explique notamment le moindre besoin exprimé sur la partie correspondant au besoin de montée en compétence sur les macro-compétences techniques. Ce propos est à nuancer dans la mesure où l'enquête a interrogé essentiellement des entreprises matures, qui se sont déclarées capables de répondre à nos questions : **les entreprises non matures, au fur et à mesure de leur montée en maturité, devront monter en compétences sur les compétences techniques IA**.

5. L3 - Le Référentiel de compétences v1

Méthodologie de conception du livrable

Le référentiel v1 a été construit à partir des éléments suivants :

- Les résultats des entretiens, notamment de l'ensemble des commentaires des entreprises interrogées et des compétences manquantes ou redondantes dans la version v0 ayant servi de base au questionnaire,
- Une veille technologique et l'expertise du consortium sur les évolutions récentes en IA (IA générative notamment).

Ce référentiel adresse spécifiquement les compétences en intelligence artificielle, mais, pour mettre en œuvre des projets d'IA, il est nécessaire que les équipes intègrent des compétences permettant de concevoir et gérer l'infrastructure sous-jacente ainsi que l'industrialisation. Si des compétences autour de la *data* de manière générale sont absolument nécessaires, ce référentiel traite uniquement de ce qui est indispensable pour les projets d'intelligence artificielle et omet volontairement des sujets spécifiques à la *data* comme la gouvernance des données par exemple.

- **Architecture du livrable**

Tout comme le référentiel v0, le référentiel v1 est constitué de deux grandes parties : l'une sur les compétences transverses et l'autre sur les compétences techniques, cette dernière partie ayant été plus développée pour prendre en compte les résultats de l'enquête.

- La partie transverse comprend 10 macro-compétences et 40 sous-compétences. Dans cette partie, quelques sous-compétences ont été ajoutées et/ou réorganisées, mais de manière globale, la partie transverse diffère peu entre les référentiels v0 et v1.
- La partie technique comprend 24 macro-compétences et 158 sous-compétences. Elle a été largement réorganisée et de nombreuses compétences ont été ajoutées pour tenir compte des évolutions récentes de l'IA et des besoins à couvrir en matière de mise en œuvre, avec notamment :
  - La refonte de la partie Architecture : tendance vers le *cloud* et vers de nouveaux outils intégrant des services managés ;
  - La refonte de la partie Données : le référentiel v0 ne prenait pas assez en compte les nouvelles structures (*data lake*, *cloud*) et pratiques de gouvernance des données ;
  - Dans la partie Modélisation, nous avons ajouté une section pour prendre en compte les techniques d'IA spécialisées (correspondant aux modèles pré-entraînés de vision, parole et texte), et d'IA spécialisées (graphes, *reinforcement learning, federated learning*, IA symbolique et hybride, etc.), d'IA générative (suite à l'arrivée des *LLMs* et des agents conversationnels) ;
  - Nous avons largement revu la partie Industrialisation pour mieux prendre en compte la gestion du cycle de vie, la supervision et le *monitoring*.

Le référentiel v1 complet est disponible en annexe 4 (modifications par rapport au Référentiel v0 indiquées en bleu) et en ligne sous format excel à l'adresse suivante : https://www.hub-franceia.fr/refco/

Validation du référentiel : les persona

- **Méthodologie**

Le questionnaire portant sur les besoins de compétences, il n'est pas possible de traduire directement les résultats en besoins de personnes, une même personne ayant en général plusieurs compétences. Il a donc été décidé de construire des « Persona » – des profils de personnes – et de leur attribuer des compétences sur les versions v0 et v1 du référentiel. Le choix s'est porté sur des Persona correspondant à des métiers existant dans les entreprises sans les décomposer en sous-métiers et de rester sur une typologie minimale permettant de constituer des équipes en charge de la mise en œuvre de projets en intelligence artificielle. Notons que chaque entreprise a développé son propre référentiel de métiers de l'IA : nos Persona ne sont donc sans doute pas parfaitement alignées avec ces référentiels.

Ces Persona vont nous servir d'une part à analyser la validité du Référentiel v1 et d'autre part à produire des estimations de volumes de formations en IA.

- **Profils de Persona**

Les Persona de notre étude sont les suivants :

- Le *data scientist*, en charge des modèles de *machine learning* ;
- Le *data engineer*/développeur qui aide le *data scientist* sur la partie amont de collecte et de mise en qualité des données et sur la partie aval sur l'industrialisation des projets d'IA, ;
- Le chef de projet qui a une vue globale et coordonne l'ensemble des opérations ;
- Le *data architect* qui conçoit et met en œuvre l'architecture IT sous-jacente ;
- Le métier qui exprime des besoins et sera le principal utilisateur ;
- Le *ML Ops/Machine Learning engineer*, qui joue un rôle important dans l'industrialisation et effectue le lien entre le *data* engineer et le data scientist.

La définition des Persona a été faite à partir du Référentiel v1 en indiquant pour chaque Persona les compétences souhaitées (annexe 8). Le nombre de compétences attachées à chaque Persona est assez homogène (il était bien moins équilibré dans la version v0) : 69 compétences pour le *data scientist*, 68 compétences pour le *data engineer*, 39 compétences pour le chef de projet, 31 compétences pour le *data architect*, 21 pour les métiers et 54 pour le *ML Ops/ML engineer*. Cette répartition montre bien que dans ce référentiel l'accent est mis sur l'intelligence artificielle dans laquelle les *data scientists, data engineer* et *Machine Learning Engineer* ont un rôle fondamental.

- **Analyse de données des Persona**

A partir du référentiel et des compétences rattachées aux Persona, nous avons mis en œuvre des analyses de données (par l'*analyse factorielle* d'une part et *clustering hiérarchique* d'autre part) pour analyser la validité du référentiel (les analyses détaillées sont présentées en Annexes 9 et 10).

Analyse factorielle

Cette analyse (annexe 9) a pour objectif de comprendre l'interdépendance des diverses compétences en IA au sein des différentes fonctions liées à l'IA dans les entreprises françaises afin de confirmer la possibilité d'établir un référentiel cohérent de compétences. Pour atteindre cet objectif, une Analyse en Composantes Principales (ACP) a été utilisée : elle permet d'extraire les tendances prépondérantes de l'ensemble de données.

Les résultats de l'ACP ont révélé six composantes principales, chaque composante, interprétées comme un axe de l'ACP, représentant l'une des six fonctions / persona. Un aspect capital de ces résultats est l'orthogonalité des axes qui indique que **ces persona sont indépendantes** les unes des autres. Cela signifie qu'une force dans un domaine de compétence n'implique pas nécessairement une force ou une faiblesse dans un autre domaine de compétence.

Cette analyse factorielle a permis d'établir une compréhension claire de la structure des compétences en IA au sein des entreprises françaises. Les résultats montrent que les compétences sont bien distinctes et indépendantes les unes des autres et forment un référentiel. Cela suggère que les efforts pour renforcer les compétences dans une fonction spécifique n'affecteront pas nécessairement les compétences dans d'autres domaines.

Clustering Hiérarchique

En travaillant sur les Persona et les compétences (annexe 10), on montre (par une *heatmap*) que les compétences sont bien identifiées pour chaque Persona avec des blocs de compétences bien différenciés selon les Persona, seules quelques compétences communes apparaissant entre Persona voisines (*data engineer* et *data scientist*, *data engineer* et *data architect* ou encore *data scientist* et *machine learning engineer*).

- **Validation du Référentiel v1**

Les analyses ont ainsi montré que le Référentiel v1 permettait de décrire des fonctions à partir de leurs compétences : ces fonctions correspondent à des blocs de compétences bien différenciés, avec peu de recouvrement entre les fonctions.

Ceci implique :

- Les compétences sont bien distinctes et indépendantes les unes des autres et **forment un référentiel** ;
- Une **force dans un domaine de compétence**s n'implique pas nécessairement une force ou une faiblesse dans un autre domaine de compétence ;

- La **formation dans une fonction** spécifique n'affectera pas nécessairement les compétences dans d'autres domaines.

Ces études confirment donc la réussite du Référentiel v1 pour distinguer de manière précise et significative les différentes dimensions des compétences en IA, les Persona que nous avons définies sont bien indépendantes. Par ailleurs, les offres de formation devront tenir compte de ces conclusions.

6. Impacts sur les formations

Ce projet vise à identifier plus précisément les compétences nécessaires aux métiers de l'IA au sein des entreprises. Comme nous l'avons vu, nous avons identifié pour le Référentiel v1 34 macro-compétences et 198 sous-compétences (10 et 40 pour la partie transverse et 24 et 58 pour la partie technique. Le référentiel produit sera utile pour modifier / compléter les référentiels existants au sein des entreprises.

Cependant, ce projet ne traite pas du sujet des formations, et le questionnaire a été construit, sur la base du référentiel de compétences, pour donner une volumétrie globale des besoins pour chaque compétence et une évaluation temporelle de ces besoins. Les volumes indiqués par les répondants sont indiqués ci-dessous, le classement des macro-compétences dans ce tableau correspond à l'ordre indiqué dans le référentiel (Annexe 7).

A partir du référentiel de compétences v1 et des définitions des Persona, les organismes de formation (initiale et continue) pourraient identifier les blocs de compétences importants pour chaque Persona et estimer l'impact sur leurs formations, de manière que l'adaptation des formations au marché soit la meilleure possible.

L'analyse issue du projet ne permet pas d'identifier la typologie de nouvelles formations. Néanmoins, sur le constat que les besoins de montée en compétences des entreprises concernent principalement les compétences transverses, il est possible de supposer que les formations initiales issues de la technique offrent aujourd'hui un niveau de compétences suffisant, ce qui n'est *a priori* pas le cas pour les formations plus génériques. Ainsi, les écoles formant des étudiants à des métiers comme le commerce, le marketing ou encore les ressources humaines par exemple, pourraient inclure dans leur cursus des cours de base sur l'intelligence artificielle dans un but d'acculturer ces futurs professionnels aux outils technologiques. Pour pallier ce manque de connaissance sur les populations déjà actives, la formation continue est un levier intéressant pour les amener à un niveau de compétence supérieur, soit en autonomie, par la visualisation de MOOC[6], soit de manière supervisée, auprès d'organismes de formation certifiés ou par le biais de l'académie d'entreprise, dans le cas où une telle structure est bien présente.

6 Massive Open Online Course

Transverse Technique (TR/TE)	Famille de compétences	Compétence	Rang	<5	5-10	10-50	50-500	>500
TR	Pilotage de projets IA	*Piloter des projets data et IA*	2	32	15	7	4	0
TR	Évaluer la faisabilité d'un besoin métier	*Savoir expliquer les bénéfices d'un projet IA aux interlocuteurs métier*	4	29	9	13	3	0
		Savoir identifier l'opportunité d'un projet IA	7	29	10	7	2	0
		Connaître les problématiques éthiques de l'IA	3	35	12	6	5	1
		Connaître les enjeux de conformité et réglementaire de l'IA	1	45	15	6	2	1
		Appliquer les bonnes pratiques pour gérer les risques IA	6	33	9	3	2	1
		Savoir mettre en place une approche sobre de l'IA	5	37	9	4	3	1
	Mettre en place la co	*Savoir identifier les enjeux stratégiques d'évolution*	11	18	14	1	3	1
TR		*Savoir concevoir un plan de changement*	23	17	3	2	2	1
		Savoir accompagner le changement induit par un projet IA	8	24	14	1	2	1
TE	Architecture	*Concevoir et réaliser les infrastructures big data*	12	21	10	3	2	0
		Spécifier et programmer les traitements des données dans le datalake	25	11	5	3	5	0
		Optimiser les traitements des données dans le datalake	26	11	6	1	6	0
TE	Données	*Auditer les sources de données disponibles*	20	17	3	4	3	0
		Auditer la qualité des données	21	16	4	2	5	1
		Préparer, nettoyer et programmer le recodage des données	27	13	5	1	4	0
		Augmenter les données	14	22	9	4	2	0
TE	Modélisation	*Réaliser une analyse exploratoire des données*	17	20	5	6	3	0
		Programmer les indicateurs nécessaires au fonctionnement des algorithmes	10	22	7	5	3	1
		Identifier et sélectionner les algorithmes de base à utiliser pour la modélisation	15	22	6	3	4	1
		Prototyper les modèles d'IA	24	13	7	4	2	0
		Tester et décider de rejeter/adopter les modèles en mesurant et évaluant leur pe	19	17	5	5	1	0
		Automatiser les modèles d'IA	22	15	6	4	2	0
TE	Industrialisation	*Optimiser les modèles d'IA*	18	18	6	4	2	0
		Industrialiser, sécuriser, rendre résilients et utilisables les modèles d'IA	13	21	9	6	2	
		Maintenir les modèles et algorithmes en état de fonctionnement	16	22	6	5	2	1
TE	Transfert au métier	*Restituer aux interlocuteurs métier les résultats des modèles*	9	26	4	7	4	0

Par ailleurs, un effort d'adaptation aux mutations conjoncturelles actuelles, dues à l'essor exponentiel des technologies d'IA génératives, est à prévoir dans les futurs cursus de formation. En effet, il ressort du projet que les formations devront intégrer des enseignements spécifiques pour suivre cette récente évolution. Notons cependant que l'enquête réalisée sur la base du Référentiel v0 ne contenait pas de questions spécifiques sur le sujet.

Plusieurs initiatives et bonnes pratiques existent au niveau international. La classification ESCO[7] recense et catégorise les aptitudes, compétences, certifications et professions ainsi que l'enseignement et la formation au sein de l'Union Européenne. Le DigitComp[8] (*Digital Competence Framework for Citizen*) fournit une compréhension de ce qu'est une compétence numérique, avec notamment plus de 250 compétences et attitudes aidant les citoyens à s'approprier les technologies digitales. Enfin, de manière plus spécifique, nous pouvons citer le *framework* NICE[9] (*National Initiative for Cybersecurity Education*) qui décrit les tâches liées à la cybersécurité, et les connaissances et compétences nécessaires pour effectuer ces tâches.

Le projet REFCO-IA a permis d'identifier les compétences nécessaires aux métiers de l'IA au sein des entreprises, aussi bien sur la formation continue que la formation initiale.

Suite à la demande de la CNIA, nous avons exploité les résultats de l'enquête pour essayer de produire une estimation indicative des besoins en formation (nous rappelons qu'aucune question n'était posée sur la nature des formations dans le questionnaire).

Nous avons redressé notre échantillon sur-représenté dans le secteur du numérique. La répartition des entreprises par secteur d'activité en France selon l'INSEE[10] et dans notre échantillon était la suivante :

	Population	Echantillon
Nombre d'entreprises	4 114 289	144
Numérique (Information et Communication)	153 727	73

[7] *Classification ESCO : https://esco.ec.europa.eu/fr*

[8] *https://publications.jrc.ec.europa.eu/repository/handle/JRC128415*
https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-181r1.pdf
https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-181r1.pdf

[10] *https://www.insee.fr/fr/statistiques/2498421*

Pourcentage	3,74%	50,69%

Nous obtenons alors les besoins estimés suivants :

	Data Scientist	Data Engineer/ Développeur	Chef de projet	IT/Data architect	Métier	ML Ops/ML Engineer
TOTAL DES BESOINS EN FORMATION	55 785	64 144	84 520	56 937	99 272	57 200
niveau	Bac+5	Bac+3	Bac+3	Bac+5	Bac+3	Bac+5

Pour compléter cette analyse, nous recommandons de projeter 4 types de formations continues au sein des entreprises en fonction des profils à former :

- **Acculturation** : à destination des métiers et des comités de direction, c'est une formation massive d'une durée variant de 1h à une demi-journée, voire une journée ;
- **Sensibilisation** : à destination des parties prenantes de projets IA – par exemple des acheteurs ou des managers – ce niveau de formation vise à leur permettre de comprendre suffisamment la technique pour être en mesure de prendre des décisions éclairées selon le contexte de leur entreprise. Cette formation peut durer entre 1 et 2 jours ;
- **Maîtrise** : à destination des chefs de projet qui sont les véritables chefs d'orchestre de projets IA et qui ont besoin de maîtriser plus en profondeur les aspects techniques, cette formation peut durer entre 1 et 2 semaines ;
- **Expertise** : enfin, pour les profils d'experts techniques, ce dernier niveau de formation vise à combler les manques des formations initiales, souvent théoriques, qui ne permettent pas de couvrir l'ensemble des besoins réels du terrain. Cette formation devrait durer entre 1 et 2 mois.

7. Conclusion et préconisations

L'analyse réalisée dans le cadre de ce projet nous a donc permis de **valider le référentiel de compétences IA** (v1 livrable L3) en termes de structure et de compétences représentées. Nous avons également démontré l'intérêt de **prévoir des cursus de formation bien distincts en fonction des profils à former** (voir §5. Validation du référentiel : les persona).

Il est nécessaire cependant de comprendre ces conclusions au vu de la sur-représentation des entreprises du numérique dans notre enquête. Cela s'explique par plusieurs facteurs :

- Les membres du comité de pilotage du projet sont tous issus, de près ou de loin, du secteur du numérique ;
- Les questions sur les compétences IA de notre enquête nécessitent un certain niveau de maturité en IA ;
- Les entreprises sollicitées n'ayant pas répondu ne se sont pas senties concernées par ces sujets car elles n'ont pas ce type de compétences en interne (ou – à la marge – elles, n'ont pas eu le temps de participer).

L'exercice demandé au travers de notre questionnaire (et des *interviews* basées sur ce même questionnaire) nécessite une réelle maturité des entreprises pour comprendre les compétences analysées, savoir se projeter sur des besoins futurs et identifier précisément les profils de collaborateurs à former. Nous avons constaté que sans un niveau de maturité en IA suffisant, il est impossible pour une entreprise de se positionner sur ces questions.

Notre enquête démontre, par le peu de mobilisation des entreprises hors du secteur du numérique, leur **manque de maturité en IA** : notons que notre échantillon de répondants, avec 50% d'entreprises du numérique correspond au taux d'adoption en IA indiqué par McKinsey[11] (au niveau mondial). Notre référentiel de compétences IA est bien compris par les entreprises déjà imprégnées par cette technologie, mais trop technique pour les entreprises d'autres horizons.

Nous retenons que le biais induit par la sur-représentation du secteur du numérique dans notre enquête peut expliquer que la hiérarchisation des attentes principales en termes de compétences soient orientées sur les compétences transverses telles que : le réglementaire, l'éthique, l'environnement et la gestion de projet (transfert entre la technique et les métiers notamment), les répondants à notre enquête disposant déjà de compétences d'experts sur la majorité des domaines techniques.

Les résultats de ce projet constituent la **phase préparatoire d'un long travail d'observation et d'analyse** d'une évolution à anticiper : la maturité IA des entreprises va augmenter, passant de 50% actuellement à sans doute 80% à 5 ans, voire davantage. Ces entreprises auront besoin d'acquérir les compétences techniques décrites dans le Référentiel IA.

L'évolution des compétences se poursuit à une vitesse fulgurante au fur et à mesure des avancées technologiques. Il nous semble donc nécessaire de **poursuivre ces travaux pour en tenir compte et pouvoir adapter les offres de formation aux besoins réels du terrain**.

Le consortium du projet REFCO-IA recommande donc de **réitérer cette enquête régulièrement** de manière à faire vivre et évoluer ce référentiel de compétences et permettre aux entreprises d'anticiper leurs besoins en fonction d'une vision globale

[11] https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review

partagée dépassant le seul cadre de leur cœur de métier et donc de bénéficier de l'expérience de projets développés dans d'autres secteurs et d'autres entreprises.

Annexes

Annexe 1 - Consortium....15
Annexe 2 – Référentiel v0....16
Annexe 3 – Questionnaire....20
Annexe 4 – Référentiel v1....29
Annexe 5 – Résultats de l'enquête : les ressources....34
Annexe 6 – Résultats de l'enquête : besoins en compétences....36
Annexe 7 – Tableau récapitulatif des besoins de montée en compétences....46
Annexe 8 – Définition des persona....48
Annexe 9 – Analyse factorielle....52
Annexe 10 – Clustering hiérarchique....53

Annexe 1 - Consortium

Organismes de formation ou d'accompagnement (universités, écoles, lycées, CFA, organismes privés, Pôle emploi, etc.).
Business & Decision (Université)
ESSEC
EPITA IA Institut
LePont
Société Informatique de France
JustAI

Donneurs d'ordre publics dans l'achat de formation continue des chercheurs d'emploi (conseils régionaux, Pôle emploi, OPCO)

Employeurs ou leurs représentants (entreprises, groupements d'employeurs, organisations professionnelles, syndicats, fédérations professionnelles, , etc.)
Hub France IA
Numeum
Datavaloris
La Poste
Société Générale

AUTRES PARTICIPANTS AU PROJET

(Collectivités territoriales, équipes de recherche, entreprises…)

BNPP	Stellantis
CIGREF	ClusterAI
Ekitia	Digital Aquitaine
Gendarmerie Nationale	GrandENov+
Renault	L'Oréal
SNCF	

Annexe 2 – Référentiel v0

RÉFÉRENTIEL DE COMPÉTENCES INTELLIGENCE ARTIFICIELLE v0

1	**COMPÉTENCES TRANSVERSES**
1.1	**PILOTAGE DE PROJET IA**
1.1.1	Savoir réaliser un plan projet
1.1.2	Savoir utiliser les méthodes agiles
1.1.3	Connaitre les méthodes de gestion DataOps et DevOps
1.1.4	Savoir encadrer une équipe projet IA
1.1.5	Savoir garantir le respect du calendrier
1.1.6	Savoir accompagner les équipes métiers
1.2	**ÉVALUER LA FAISABILITÉ D'UN BESOIN MÉTIER**
1.2.1	*Savoir expliquer les bénéfices d'un projet IA aux interlocuteurs métier*
1.2.1.a	Connaître les bénéfices et les impacts de l'IA
1.2.1.b	Etre capable d'identifier les sujets relevant de l'IA
1.2.1.c	Savoir scénariser des implémentations possibles de cas relevant de l'IA
1.2.2	*Savoir identifier l'opportunité d'un projet IA*
1.2.2.a	Savoir expliquer les problématiques typiques en IA
1.2.2.b	Savoir évaluer la complexité d'une problématique IA
1.2.2.c	Etre capable d'évaluer la faisabilité d'un projet
1.2.2.d	Savoir identifier les algorithmes utilisables
1.2.3	*Connaître les problématiques éthiques de l'IA*
1.2.3.a	Savoir mettre en place une veille sur l'éthique de l'IA
1.2.3.b	Savoir intégrer une démarche éthique dans les projets IA
1.2.3.c	Savoir mettre en oeuvre des procédures de mesure et contrôle des impacts de la
1.2.3.d	Savoir mettre en place des procédures de protection des données personnelles
1.2.4	*Connaitre les enjeux de conformité et réglementaires de l'IA*
1.2.4.a	Savoir mettre en place une veille sur la réglementation de l'IA
1.2.4.b	Savoir évaluer les impacts d'un projet IA au vu de la réglementation
1.2.4.c	Savoir mettre en place le RGPD
1.2.5	*Appliquer les bonnes pratiques pour gérer les risques IA*
1.2.5.a	Savoir mettre en place une méthodologie de mesure des risques
1.2.5.b	Savoir mettre en place une gouvernance de contrôle des risques
1.2.6	*Savoir mettre en place une approche sobre de l'IA*
1.2.6.a	Savoir mesurer les dépenses énergétiques de l'IA
1.2.6.b	Savoir proposer des algorithmes IA plus frugaux
1.2.6.c	Savoir évaluer et mettre en place des best practices d'utilisation frugale de l'IA
1.3	**SAVOIR METTRE EN PLACE LA CONDUITE DU CHANGEMENT**
1.3.1	*Savoir identifier les enjeux stratégiques d'évolution*
1.3.1.a	Savoir identifier le contexte, les enjeux, les objectifs et les contraintes de l'entreprise
1.3.1.b	Savoir identifier les différents processus métier potentiellement impactés par l' IA

1.3.1.c	Savoir produire une étude de cadrage pour accompagner le changement par des projets IA
1.3.1.d	Savoir mettre en oeuvre une approche éthique pour mesurer l'impact des changements
1.3.2	*Savoir concevoir un plan de changement*
1.3.2.a	Savoir identifier les parties prenantes (MOA , MOE) et les canaux de communication
1.3.2.b	Savoir identifier les actions à mener pour optimiser les processus internes, suite au projet IA
1.3.2.c	Savoir produire le plan de changement
1.3.3	*Savoir accompagner le changement induit par un projet IA*
1.3.3.a	Savoir élaborer une stratégie de communication auprès des acteurs du projet IA
1.3.3.b	Savoir impliquer les directions métiers
1.3.3.c	Savoir traduire les besoins en termes de transfert de compétences
1.3.3.d	Savoir élaborer des plans de formation pour accompagner le changement

2	**COMPÉTENCES TECHNIQUES**
2.1	**ARCHITECTURE**
2.1.1	*Concevoir et réaliser les infrastructures big data*
2.1.1.a	Savoir identifier la bonne infrastructure & les différents types d'infrastructure
2.1.1.b	Maîtriser le fonctionnement de l'écosystème Hadoop
2.1.1.c	Maîtriser le fonctionnement de l'écosystème Spark
2.1.1.d	Maîtriser le fonctionnement des bases de données NoSql
2.1.1.e	Maitriser le fonctionnement d'une infrastructure cloud majeure
2.1.2	*Spécifier et programmer les traitements des données dans le datalake*
2.1.2.a	Savoir programmer en Java
2.1.2.b	Savoir programmer en Scala
2.1.2.c	Savoir programmer avec Hive
2.1.2.d	Savoir utiliser le logiciel Kafka
2.1.3	*Optimiser les traitements des données dans le datalake*
2.1.3.a	Mettre en oeuvre la mesure de la performance des traitements sur l'infrastructure cible
2.1.3.b	Savoir paralléliser les traitements sur infrastructure Spark
2.1.3.c	Pouvoir identifier les goulots d'étranglement et les traitements critiques
2.1.3.d	Savoir programmer en Scala avec Spark ML
2.1.3.e	Savoir programmer en PySpark avec Spark ML
2.2.	**DONNÉES**
2.2.1	*Auditer les sources de données disponibles*
2.2.1.a	Etre capable de reconnaitre le caractère éthique ou non éthique des données
2.2.1.b	Pouvoir identifier les données atypiques et/ou aberrantes
2.2.1.c	Pouvoir identifier et gérer les données manquantes
2.2.1.d	Etre capable d'appliquer les règles du RGPD
2.2.2	*Auditer la qualité des données*
2.2.2.a	Maîtriser la statistique descriptive
2.2.2.b	Savoir utiliser la librairie Numpy en langage Python
2.2.2.c	Savoir utiliser la librairie Pandas en langage Python
2.2.2.d	Savoir réaliser un audit de qualité de données
2.2.3	*Préparer, nettoyer et programmer le recodage des données*
2.2.3.a	Savoir programmer le recodage des données

2.2.3.b	Savoir traiter les valeurs manquantes
2.2.3.c	Savoir traiter les valeurs aberrantes
2.2.3.d	Maîtriser le fonctionnement des bases de données SQL
2.2.3.e	Pouvoir créer des jeux de données de test
2.2.4	*Augmenter les données*
2.2.4.a	Connaître les sources de données open-source
2.2.4.b	Savoir générer de nouvelles variables (feature engineering)
2.2.4.c	Savoir utiliser un simulateur / jumeau numérique
2.2.4.d	Savoir utiliser un modèle génératif (GAN)
2.3	**MODÉLISATION**
2.3.1	*Réaliser une analyse exploratoire des données*
2.3.1.a	Maîtriser les principales méthodes de réduction de dimension non supervisées
2.3.1.b	Maîtriser les principales méthodes de clustering et de typologies non supervisées
2.3.1.c	Maîtriser l'analyse statistique des corrélations
2.3.1.d	Savoir créer un jeu de données pertinent pour la cross-validation
2.3.1.e	Savoir éviter un biais dans les données d'apprentissage
2.3.1.f	Maîtriser les différentes méthodes de test d'hypothèses
2.3.2	*Concevoir des modèles analytiques d'IA et de machine-learning*
2.3.2.1	***Programmer les indicateurs nécessaires au fonctionnement des algorithmes***
2.3.2.1.a	Savoir utiliser la méthode MLOps pour mener à bien un projet
2.3.2.1.b	Savoir mettre en ≈ìuvre une stratégie d'inférence
2.3.2.1.c	Etre capable de mettre en ≈ìuvre du machine learning non supervisé
2.3.2.1.d	Etre capable de mettre en ≈ìuvre du machine learning supervisé
2.3.2.1.e	Savoir programmer les indicateurs nécessaires
2.3.2.2	***Identifier et sélectionner des algorithmes de base à utiliser pour la modélisation***
2.3.2.2.a	Connaitre les différentes formes d'apprentissage machine
2.3.2.2.b	Savoir entraîner les algorithmes d'apprentissage
2.3.2.2.c	Connaître les principaux algorithmes d'IA du domaine public
2.3.2.2.d	Savoir utiliser les algorithmes du domaine public pour créer un modèle
2.3.2.3	***Savoir utiliser les modèles pré-entraînés du domaine public pour créer un modèle (IA générative, vision, texte, parole)***
2.3.2.3.a	Savoir développer un modèle d'IA symbolique
2.3.2.3.b	Savoir développer un modèle multiagent
2.3.2.4	***Prototyper les modèles d'IA***
2.3.2.4.a	Savoir utiliser les principales librairies d'IA en Python
2.3.2.4.b	Etre capable d'assembler la chaîne de traitement des différents modèles
2.3.2.4.c	Savoir optimiser le compromis explicabilité & traçabilité / précision des algorithmes
2.3.2.4.d	Savoir documenter les modèles avec Jupyter Notebook
2.3.2.4.e	Savoir implémenter des pistes d'auditabilité et de traçabilité
2.3.3	*Tester et implémenter des modèles analytiques d'IA*
2.3.3.1	***Tester et décider de rejeter/adopter les modèles en mesurant & évaluant leur pertinence sur la base de leur précision, leur robustesse, leur absence de biais***
2.3.3.1.a	Savoir mesurer la performance des modèles
2.3.3.1.b	Savoir créer les indicateurs pour valider ou rejeter les modèles testés
2.3.3.1.c	Savoir calculer et mesurer la variance des résidus d'un modèle quantitatif
2.3.3.1.d	Savoir produire la matrice de confusion d'un modèle de classification

2.3.3.1.e	Maîtriser les méthodes d'optimisation de complexité et de cross-validation des modèles
2.3.3.1.f	Savoir exploiter les résultats d'une stratégie inférentielle de cross-validation
2.3.3.1.g	Savoir s'assurer du respect des hypothèses de fonctionnement des modèles
2.3.3.1.h	Savoir s'assurer de l'absence de biais dans les résultats du modèle
2.3.3.1.i	Savoir mesurer la robustesse statistique des modèles
2.3.3.2	***Automatiser les modèles d'IA***
2.3.3.2.a	Savoir choisir les hyper-paramètres pour faciliter l'automatisation
2.3.3.2.b	Savoir automatiser l'optimisation des hyper paramètres
2.4	**INDUSTRIALISATION**
2.4.1	*Optimiser les modèles de datascience*
2.4.1.a	Savoir documenter les modèles (par exemple avec Jupyter Notebook)
2.4.1.b	Savoir construire le workflow (graphique) d'un modèle
2.4.1.c	Etre capable de produire du code industrialisable (par exemple en utilisant des notebooks) et gérer les versions de modèles (par exemple avec Git/GitHub)
2.4.2	*Industrialiser, sécuriser, rendre résilients et utilisables les modèles de datascience*
2.4.2.a	Savoir spécifier les conditions d'utilisation des modèles d'IA
2.4.2.b	Maîtriser la méthode devops en lien avec la méthode MLOps (dataops)
2.4.2.c	Savoir orchestrer un ensemble d'algorithmes
2.4.2.d	Savoir intégrer les résultats dans les applications digitales de l'entreprise
2.4.3	*Maintenir les modèles et algorithmes en état de fonctionnement*
2.4.3.a	Maîtriser les méthodes statistiques de monitoring des modèles de Machine Learning
2.4.3.b	Savoir identifier les dérives des modèles
2.4.3.c	Savoir configurer les seuils d'alertes en cas de dysfonctionnement des modèles
2.5	**TRANSFERT AU MÉTIER**
2.5.1	*Restituer aux interlocuteurs métiers les résultats des modèles*
2.5.1.a	Savoir présenter des visualisations des données compréhensibles par les métiers
2.5.1.b	Savoir illustrer les résultats du modèle par des graphiques compréhensibles par les métiers
2.5.1.c	Savoir sélectionner les observations illustrant le comportement typique des modèles
2.5.1.d	Savoir interpréter les résultats des différents modèles et en expliquer les résultats
2.5.1.e	Savoir utiliser des modèles explicatifs pour pouvoir rendre traçables / explicables les modèles
2.5.1.f	Savoir expliquer les risques des modèles
2.5.1.g	Savoir présenter les biais des modèles
2.5.1.h	Savoir mesurer les écarts/non conformités et les communiquer
2.5.1.i	Savoir mesurer les impacts des résultats

Annexe 3 – Questionnaire

Nous précisons les réponses possibles aux questions posées.

BESOINS EN COMPÉTENCES INTELLIGENCE ARTIFICIELLE ET DATA	
Ce questionnaire s'inscrit dans le cadre du projet REFCO-IA soutenu par la Banque des Territoires dans le cadre de l'AMI Compétences et Métiers d'Avenir (volet diagnostic). L'objectif du projet, d'une durée de 6 mois, est la réalisation d'un référentiel de compétences IA qui reflète les besoins actuels des entreprises. *Ce questionnaire vise à recueillir vos besoins en compétences Data et Intelligence Artificielle. Répondez aux questions de votre mieux, et laissez les questions sans réponse si vous considérez qu'elles ne vous concernent pas. Remplir le questionnaire vous prendra entre 20 et 30 minutes.* *Merci de votre participation ! L'équipe du projet REFCO-IA.* *Le consortium du projet REFCO-IA est composé des organisations suivantes :* *Pilote : Hub France IACo-pilote : Numeum* *Partenaires : Datavaloris, LePont, La Poste Groupe, SIF, IA Institut, Business & Decision, Société Générale, Just AI, ESSEC-METALAB*	
GRILLE D'INTERVIEW	
QUESTIONS	**RÉPONSE**
QUESTIONS GÉNÉRALES	
DONNÉES DE L'ENQUÊTE	
Partenaire en charge de l'enquête	Hub France IA; DataValoris; LePont; La Poste; SIF; IA Institut; Business & Decision; Société Générale; JustAI; ESSEC; Numeum; Je ne sais pas
Format de l'enquête	Questionnaire; Interview
Date de l'enquête	
PRÉSENTATION DE L'ENTREPRISE INTERVIEWÉE	
Nom de votre entreprise	
Type d'entreprise	Start-up/Scale-up; PME/ETI; Grand Groupe
Taille de l'entreprise	< 5; 5-10; 10-50; 50-500; 500-1 000; > 1 000
Secteurs d'activité (Cocher la case correspondante)	Aéronautique & Défense Agriculture & Agroalimentaire Banque & Assurance Commerce & Distribution Construction Electronique Energie & Service aux collectivités Industries culturelles et créatives Industrie Mobilité, Transport & Logistique, Automobile Numérique Proximité, Economie sociale & Sécurité civile Education, Santé & Bien-être Services aux citoyens Luxe & Textile Tourisme
MATURITÉ IA DE L'ENTREPRISE INTERVIEWÉE	
Votre entreprise est-elle considérée principalement comme un fournisseur ou un utilisateur d'IA ?	fournisseur de solutions IA; utilisateur de solutions IA
Quelle est votre fonction au sein de votre entreprise ?	CEO; CTO; CDO (Data ou Digital); Directeur Innovation; DSI; DRH; BU métier; Autre (préciser)
Dans quel département au sein de votre entreprise excercez-vous	
Comment considérez-vous le niveau de maturité IA de votre département (ou de votre entreprise) ?	0 - Aucune donnée collectée et aucun projet IA réalisé; 1 - La collecte des données est en place; 2 - Au moins un premier POC en IA a été développé; 3 - Au moins un projet IA a été déployé en production; 4 - De nombreux projets IA ont été déployés / Une "AI Factory" est en place

Disposez-vous actuellement de talents ou d'expertise IA au sein de votre département (ou de votre entreprise) ? Si oui cette expertise est-elle internalisée ou externalisée ?		oui, cette compétence est principalement en interne; oui, cette compétence est à la fois interne et externe; oui, mais l'entreprise s'appuie sur des compétences externes; non, l'entreprise n'a pas accès à des compétences IA internes ou externes
Vous avez déclaré avoir des ressources IA en interne, quels sont les profils de ces ressources ?	Métier	Concerné; non concerné
	Données	Concerné; non concerné
	Data science	Concerné; non concerné
	IT/Ops	Concerné; non concerné
	Formation	Concerné; non concerné
	Chef de projet	Concerné; non concerné
	Conduite du changement	Concerné; non concerné
	AMOA	Concerné; non concerné
	Autre	Concerné; non concerné
Vous avez déclaré avoir des ressources IA en interne. Nous souhaitons maintenant en savoir plus sur ces ressources. Pour chaque profil de ressource identifié, quel est le niveau de formation ?	Métier	1 - Notions d'IA; 2 - Intermédiaire; 3 - Confirmé; 4 - Expert
	Données	1 - Notions d'IA; 2 - Intermédiaire; 3 - Confirmé; 4 - Expert
	Data science	1 - Notions d'IA; 2 - Intermédiaire; 3 - Confirmé; 4 - Expert
	IT/Ops	1 - Notions d'IA; 2 - Intermédiaire; 3 - Confirmé; 4 - Expert
	Formation	1 - Notions d'IA; 2 - Intermédiaire; 3 - Confirmé; 4 - Expert
	Chef de projet	1 - Notions d'IA; 2 - Intermédiaire; 3 - Confirmé; 4 - Expert
	Conduite du changement	1 - Notions d'IA; 2 - Intermédiaire; 3 - Confirmé; 4 - Expert
	AMOA	1 - Notions d'IA; 2 - Intermédiaire; 3 - Confirmé; 4 - Expert
	Autre	1 - Notions d'IA; 2 - Intermédiaire; 3 - Confirmé; 4 - Expert
Vous avez déclaré avoir des ressources IA en interne. Nous souhaitons maintenant en savoir plus sur ces ressources. Pour chaque profil de ressource identifié, quelle est la volumétrie concernée ?	Métier	< 5; 5-10; 10-50; 50-500; >500
	Données	< 5; 5-10; 10-50; 50-500; >500
	Data science	< 5; 5-10; 10-50; 50-500; >500
	IT/Ops	< 5; 5-10; 10-50; 50-500; >500
	Formation	< 5; 5-10; 10-50; 50-500; >500
	Chef de projet	< 5; 5-10; 10-50; 50-500; >500
	Conduite du changement	< 5; 5-10; 10-50; 50-500; >500
	AMOA	< 5; 5-10; 10-50; 50-500; >500
	Autre	< 5; 5-10; 10-50; 50-500; >500
Vous avez déclaré avoir des ressources IA en interne. Nous souhaitons maintenant en savoir plus sur ces ressources. Vos collaborateurs ont-ils suivi une ou des formations en IA ?		Oui; non
Vous avez déclaré que vos collaborateurs ont suivi une ou plusieurs formations. Quel(s) étai(ent) le(s) niveau(x) de cette (ces) formation(s) en IA ?		1 - Notions d'IA; 2 - Intermédiaire; 3 - Confirmé; 4 - Expert
Vous avez déclaré que vos collaborateurs ont suivi une ou plusieurs formations. Combien de collaborateurs au total ont suivi ces formations à l'IA ?		< 5; 5-10; 10-50; 50-500; >500
Vous avez déclaré vous appuyer sur des ressources IA externes, d'où proviennent ces ressources ?		Intégrateur / ESN; Editeur de solutions IA; Laboratoire de recherche
Vous avez déclaré vous appuyer sur des ressources IA externes, quels sont les profils de ces ressources ?	Données	Concerné; non concerné
	Data science	Concerné; non concerné
	IT/Ops	Concerné; non concerné
	Formation	Concerné; non concerné
	Chef de projet	Concerné; non concerné
	Conduite du changement	Concerné; non concerné
	AMOA	Concerné; non concerné
	Autre	Concerné; non concerné

Dans la suite du questionnaire, le référentiel V0 a été complété de questions sur les nombres de collaborateurs concernés par les différentes sous-compétences, ces questions sont déclenchées par la réponse « oui » à la question « Avez-vous besoin de monter en compétence sur … » :

Précisez le nombre de collaborateurs concernés.		
Niveau d'expertise	*Besoin dans le temps*	*Interne / externe*
1 - Notions d'IA; 2 - Intermédiaire; 3 - Confirmé; 4 - Expert	< 1 an; > 1 an	Interne; externe

Ces informations sur les nombres de collaborateurs concernés permettent d'établir les statistiques de besoins par compétences présentées en Annexe 6.

BESOIN EN COMPÉTENCES TRANSVERSES
PILOTAGE DE PROJET IA
Avez-vous besoin de monter en compétence sur "Piloter des projets Data & IA" ?
Précisez le nombre de collaborateurs concernés.
Niveau d'expertise - Besoin dans le temps - Interne/externe
Savoir réaliser un plan projet
Savoir utiliser les méthodes agiles
Connaitre les méthodes de gestion DataOps et DevOps
Savoir encadrer une équipe projet IA
Savoir garantir le respect du calendrier
Savoir accompagner les équipes métiers
ÉVALUER LA FAISABILITÉ D'UN BESOIN MÉTIER
Savoir expliquer les bénéfices d'un projet IA aux interlocuteurs métier
Avez-vous besoin de monter en compétence sur "Savoir expliquer les bénéfices d'un projet IA aux interlocuteurs métier" ?
Précisez le nombre de collaborateurs concernés.
Niveau d'expertise - Besoin dans le temps - Interne/externe
Connaître les bénéfices et les impacts de l'IA
Etre capable d'identifier les sujets relevant de l'IA
Savoir scénariser des implémentations possibles de cas relevant de l'IA
Savoir identifier l'opportunité d'un projet IA
Avez-vous besoin de monter en compétence sur "Savoir identifier l'opportunité d'un projet IA" ?
Précisez le nombre de collaborateurs concernés.
Niveau d'expertise - Besoin dans le temps - Interne/externe
Savoir expliquer les problématiques typiques en IA
Savoir évaluer la complexité d'une problématique IA
Etre capable d'évaluer la faisabilité d'un projet
Savoir identifier les algorithmes utilisables
Connaître les problématiques éthiques de l'IA
Avez-vous besoin de monter en compétence sur "Connaître les problématiques éthiques de l'IA" ?
Précisez le nombre de collaborateurs concernés.
Niveau d'expertise - Besoin dans le temps - Interne/externe
Savoir mettre en place une veille sur l'éthique de l'IA
Savoir intégrer une démarche éthique dans les projets IA
Savoir mettre en oeuvre des procédures de mesure et contrôle des impacts de la solution IA
Savoir mettre en place des procédures de protection des données personnelles
Connaitre les enjeux de conformité et réglementaires de l'IA
Avez-vous besoin de monter en compétence sur "Connaitre les enjeux de conformité et réglementaire de l'IA" ?
Précisez le nombre de collaborateurs concernés.
Niveau d'expertise - Besoin dans le temps - Interne/externe
Savoir mettre en place une veille sur la réglementation de l'IA

Savoir évaluer les impacts d'un projet IA au vu de la réglementation	
Savoir mettre en place le RGPD	
Appliquer les bonnes pratiques pour gérer les risques IA	
Avez-vous besoin de monter en compétence sur "Appliquer les bonnes pratiques pour gérer les risques IA" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Savoir mettre en place une méthodologie de mesure des risques	
Savoir mettre en place une gouvernance de contrôle des risques	
Savoir mettre en place une approche sobre de l'IA	
Avez-vous besoin de monter en compétence sur "Savoir mettre en place une approche sobre de l'IA" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Savoir mesurer les dépenses énergétiques de l'IA	
Savoir proposer des algorithmes IA plus frugaux	
Savoir évaluer et mettre en place des best practices d'utilisation frugale de l'IA	
SAVOIR METTRE EN PLACE LA CONDUITE DU CHANGEMENT	
Savoir identifier les enjeux stratégiques d'évolution	
Avez-vous besoin de monter en compétence sur "Savoir identifier les enjeux stratégiques d'évolution" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Savoir identifier le contexte, les enjeux, les objectifs et les contraintes de l'entreprise	
Savoir identifier les différents processus métier potentiellement impactés par l' IA	
Savoir produire une étude de cadrage pour accompagner le changement par des projets IA	
Savoir mettre en oeuvre une approche éthique pour mesurer l'impact des changements	
Savoir concevoir un plan de changement	
Avez-vous besoin de monter en compétence sur "Savoir concevoir un plan de changement" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Savoir identifier les parties prenantes (MOA , MOE) et les canaux de communication	
Savoir identifier les actions à mener pour optimiser les processus internes, suite au projet IA	
Savoir produire le plan de changement	
Savoir accompagner le changement induit par un projet IA	
Avez-vous besoin de monter en compétence sur "Savoir accompagner le changement induit par un projet IA" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Savoir élaborer une stratégie de communication auprès des acteurs du projet IA	
Savoir impliquer les directions métiers	
Savoir traduire les besoins en termes de transfert de compétences	
Savoir élaborer des plans de formation pour accompagner le changement	

BESOIN EN COMPÉTENCES TECHNIQUES	
ARCHITECTURE	
Concevoir et réaliser les infrastructures big data	
Avez-vous besoin de monter en compétence sur "Concevoir et réaliser les infrastructures big data" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Savoir identifier la bonne infrastructure & les différents types d'infrastructure	
Maîtriser le fonctionnement de l'écosystème Hadoop	
Maîtriser le fonctionnement de l'écosystème Spark	
Maîtriser le fonctionnement des bases de données NoSql	
Maitriser le fonctionnement d'une infrastructure cloud majeure	
Spécifier et programmer les traitements des données dans le datalake	
Avez-vous besoin de monter en compétence sur "Spécifier et programmer les traitements des données dans le datalake" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Savoir programmer en Java	
Savoir programmer en Scala	
Savoir programmer avec Hive	
Savoir utiliser le logiciel Kafka	
Optimiser les traitements des données dans le datalake	
Avez-vous besoin de monter en compétence sur "Optimiser les traitements des données dans le datalake" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Mettre en oeuvre la mesure de la performance des traitements sur l'infrastructure cible	
Savoir paralléliser les traitements sur infrastructure Spark	
Pouvoir identifier les goulots d'étranglement et les traitements critiques	
Savoir programmer en Scala avec Spark ML	
Savoir programmer en PySpark avec Spark ML	
DONNÉES	
Auditer les sources de données disponibles	
Avez-vous besoin de monter en compétence sur "Auditer les sources de données disponibles" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Etre capable de reconnaitre le caractère éthique ou non éthique des données disponibles	
Pouvoir identifier les données atypiques et/ou aberrantes	
Pouvoir identifier et gérer les données manquantes	
Etre capable d'appliquer les règles du RGPD	
Auditer la qualité des données	
Avez-vous besoin de monter en compétence sur "Auditer la qualité des données" ?	
Précisez le nombre de collaborateurs concernés.	

Niveau d'expertise - Besoin dans le temps - Interne/externe
Maîtriser la statistique descriptive
Savoir utiliser la librairie Numpy en langage Python
Savoir utiliser la librairie Pandas en langage Python
Savoir réaliser un audit de qualité de données
Préparer, nettoyer et programmer le recodage des données
Avez-vous besoin de monter en compétence sur "Préparer, nettoyer et programmer le recodage des données" ?
Précisez le nombre de collaborateurs concernés.
Niveau d'expertise - Besoin dans le temps - Interne/externe
Savoir programmer le recodage des données
Savoir traiter les valeurs manquantes
Savoir traiter les valeurs aberrantes
Maîtriser le fonctionnement des bases de données SQL
Pouvoir créer des jeux de données de test
Augmenter les données
Avez-vous besoin de monter en compétence sur "Augmenter les données" ?
Précisez le nombre de collaborateurs concernés.
Niveau d'expertise - Besoin dans le temps - Interne/externe
Connaître les sources de données open-source
Savoir générer de nouvelles variables (feature engineering)
Savoir utiliser un simulateur / jumeau numérique
Savoir utiliser un modèle génératif (GAN)
MODÉLISATION
Réaliser une analyse exploratoire des données
Avez-vous besoin de monter en compétence sur "Réaliser une analyse exploratoire des données" ?
Précisez le nombre de collaborateurs concernés.
Niveau d'expertise - Besoin dans le temps - Interne/externe
Maîtriser les principales méthodes de réduction de dimension non supervisées
Maîtriser les principales méthodes de clustering et de typologies non supervisées
Maîtriser l'analyse statistique des corrélations
Savoir créer un jeu de données pertinent pour la cross-validation
Savoir éviter un biais dans les données d'apprentissage
Maîtriser les différentes méthodes de test d'hypothèses
Concevoir des modèles analytiques d'IA et de machine-learning
Programmer les indicateurs nécessaires au fonctionnement des algorithmes
Avez-vous besoin de monter en compétence sur "Programmer les indicateurs nécessaires au fonctionnement des algorithmes" ?
Précisez le nombre de collaborateurs concernés.
Niveau d'expertise - Besoin dans le temps - Interne/externe
Savoir utiliser la méthode MLOps pour mener à bien un projet
Savoir mettre en ≈ìuvre une stratégie d'inférence
Etre capable de mettre en ≈ìuvre du machine learning non supervisé
Etre capable de mettre en ≈ìuvre du machine learning supervisé

Savoir programmer les indicateurs nécessaires	
Identifier et sélectionner des algorithmes de base à utiliser pour la modélisation	
Avez-vous besoin de monter en compétence sur "Identifier et sélectionner des algorithmes de base à utiliser pour la modélisation" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Connaitre les différentes formes d'apprentissage machine	
Savoir entraîner les algorithmes d'apprentissage	
Connaître les principaux algorithmes d'IA du domaine public	
Savoir utiliser les algorithmes du domaine public pour créer un modèle	
Savoir utiliser les modèles pré-entraînés du domaine public pour créer un modèle (IA générative, vision, texte, parole)	
Savoir développer un modèle d'IA symbolique	
Savoir développer un modèle multiagent	
Prototyper les modèles d'IA	
Avez-vous besoin de monter en compétence sur "Prototyper les modèles d'IA" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Savoir utiliser les principales librairies d'IA en Python	
Etre capable d'assembler la chaîne de traitement des différents modèles	
Savoir optimiser le compromis explicabilité & traçabilité / précision des algorithmes	
Savoir documenter les modèles avec Jupyter Notebook	
Savoir implémenter des pistes d'auditabilité et de traçabilité	
Tester et implémenter des modèles analytiques d'IA	
Tester et décider de rejeter/adopter les modèles en mesurant & évaluant leur pertinence sur la base	
Avez-vous besoin de monter en compétence sur "Tester et décider de rejeter/adopter les modèles en mesurant & évaluant leur pertinence sur la base de leur précision, leur robustesse, leur absence de biais" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Savoir mesurer la performance des modèles	
Savoir créer les indicateurs pour valider ou rejeter les modèles testés	
Savoir calculer et mesurer la variance des résidus d'un modèle quantitatif	
Savoir produire la matrice de confusion d'un modèle de classification	
Maîtriser les méthodes d'optimisation de complexité et de cross-validation des modèles	
Savoir exploiter les résultats d'une stratégie inférentielle de cross-validation	
Savoir s'assurer du respect des hypothèses de fonctionnement des modèles	
Savoir s'assurer de l'absence de biais dans les résultats du modèle	
Savoir mesurer la robustesse statistique des modèles	
Automatiser les modèles d'IA	
Avez-vous besoin de monter en compétence sur "Automatiser les modèles d'IA" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Savoir choisir les hyper-paramètres pour faciliter l'automatisation	
Savoir automatiser l'optimisation des hyper paramètres	

INDUSTRIALISATION	
Optimiser les modèles de datascience	
Avez-vous besoin de monter en compétence sur "Optimiser les modèles de datascience" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Savoir documenter les modèles (par exemple avec Jupyter Notebook)	
Savoir construire le workflow (graphique) d'un modèle	
Etre capable de produire du code industrialisable (par exemple en utilisant des notebooks) et gérer les versions de modèles (par exemple avec Git/GitHub)	
Industrialiser, sécuriser, rendre résilients et utilisables les modèles de datascience	
Avez-vous besoin de la monter en compétence sur "Industrialiser, sécuriser, rendre résilients et utilisables les modèles de datascience" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Savoir spécifier les conditions d'utilisation des modèles d'IA	
Maîtriser la méthode devops en lien avec la méthode MLOps (dataops)	
Savoir orchestrer un ensemble d'algorithmes	
Savoir intégrer les résultats dans les applications digitales de l'entreprise	
Maintenir les modèles et algorithmes en état de fonctionnement	
Avez-vous besoin de monter en compétence sur "Maintenir les modèles et algorithmes en état de fonctionnement" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Maîtriser les méthodes statistiques de monitoring des modèles de Machine Learning	
Savoir identifier les dérives des modèles	
Savoir configurer les seuils d'alertes en cas de dysfonctionnement des modèles	
TRANSFERT AU MÉTIER	
Restituer aux interlocuteurs métiers les résultats des modèles	
Avez-vous besoin de monter en compétence sur "Restituer aux interlocuteurs métiers les résultats des modèles" ?	
Précisez le nombre de collaborateurs concernés.	
Niveau d'expertise - Besoin dans le temps - Interne/externe	
Savoir présenter des visualisations des données compréhensibles par les métiers	
Savoir illustrer les résultats du modèle par des graphiques compréhensibles par les métiers	
Savoir sélectionner les observations illustrant le comportement typique des modèles	
Savoir interpréter les résultats des différents modèles et en expliquer les résultats	
Savoir utiliser des modèles explicatifs pour pouvoir rendre traçables / explicables les modèles	
Savoir expliquer les risques des modèles	
Savoir présenter les biais des modèles	
Savoir mesurer les écarts/non conformités et les communiquer	
Savoir mesurer les impacts des résultats	

QUESTIONS FINALES
Si vous souhaitez être tenu informé de l'avancement des travaux et des résultats du projets, merci de préciser votre adresse e-mail.
Souhaitez-vous nous partager d'autres compétences ou sous-compétences qui n'auraient pas été
Pour conclure, avez-vous des commentaires ou des remarques ?

Annexe 4 – Référentiel v1

REFERENTIEL DE COMPÉTENCES INTELLIGENCE ARTIFICIELLE V1	
1	**COMPÉTENCES TRANSVERSES**
1.1	**ÉVALUER LA FAISABILITÉ D'UN BESOIN MÉTIER**
1.1.1	*Savoir expliquer les bénéfices d'un projet IA aux interlocuteurs métier*
1.1.1.a	Savoir identifier les bénéfices et les impacts de l'IA
1.1.1.b	Etre capable d'identifier les sujets relevant de l'IA
1.1.1.c	Savoir identifier les cas d'usages classiques liés à un métier
1.1.1.d	Savoir scénariser des implémentations possibles de cas relevant de l'IA
1.1.2	*Savoir identifier l'opportunité d'un projet IA*
1.1.2.a	Savoir expliquer les problématiques typiques en IA
1.1.2.b	Savoir évaluer la complexité d'une problématique IA
1.1.2.c	Etre capable d'évaluer la faisabilité d'un projet
1.1.2.d	Savoir identifier les algorithmes utilisables
1.1.3	*Connaître les problématiques éthiques de l'IA*
1.1.3.a	Savoir mettre en place une veille sur l'éthique de l'IA
1.1.3.b	Savoir intégrer une démarche éthique dans les projets IA
1.1.3.c	Savoir mettre en oeuvre des procédures de mesure et contrôle des impacts de la solution IA
1.1.4	*Connaître les enjeux de conformité et réglementaires de l'IA*
1.1.4.a	Savoir mettre en place une veille sur la réglementation de l'IA
1.1.4.b	Savoir évaluer les impacts d'un projet IA au vu de la réglementation
1.1.4.c	Savoir mettre en place le RGPD
1.1.5	*Appliquer les bonnes pratiques pour gérer les risques IA*
1.1.5.a	Savoir mettre en place une méthodologie de mesure et de maîtrise des risques
1.1.5.b	Savoir mettre en place une gouvernance de contrôle des risques
1.1.5.c	Savoir mettre en place une méthodologie privacy-by-design
1.1.5.d	Savoir maîtriser les risques associés aux places de marché pour exposer les données
1.1.6	*Savoir mettre en place une approche sobre de l'IA*
1.1.6.a	Savoir mesurer les dépenses énergétiques de l'IA
1.1.6.b	Savoir proposer des algorithmes IA plus frugaux
1.1.6.c	Savoir évaluer et mettre en place des best practices d'utilisation frugale de l'IA
1.2	**PILOTAGE DE PROJET IA**
1.2.a	Savoir définir une solution projet (angles fonctionnel et métier)
1.2.b	Savoir calculer le ROI et les montants à investir
1.2.c	Savoir réaliser un plan projet
1.2.d	Savoir utiliser les méthodes agiles
1.2.e	Connaître les méthodes de gestion DataOps et DevOps
1.2.f	Savoir encadrer une équipe projet IA
1.2.g	Savoir garantir le respect du calendrier
1.2.h	Savoir accompagner les équipes métiers
1.3	**SAVOIR METTRE EN PLACE LA CONDUITE DU CHANGEMENT**
1.3.1	*Savoir identifier les enjeux stratégiques d'évolution*
1.3.1.a	Savoir identifier le contexte, les enjeux, les objectifs et les contraintes de l'entreprise
1.3.1.b	Savoir identifier les différents processus métier potentiellement impactés par l' IA
1.3.1.c	Savoir produire une étude de cadrage pour accompagner le changement par des projets IA
1.3.1.d	Savoir mettre en oeuvre une approche éthique pour mesurer l'impact des changements
1.3.2	*Savoir concevoir un plan de changement*
1.3.2.a	Savoir identifier les parties prenantes (MOA , MOE) et les canaux de communication
1.3.2.b	Savoir identifier les actions à mener pour optimiser les processus internes, suite au projet IA
1.3.2.c	Savoir produire le plan de changement
1.3.3	*Savoir accompagner le changement induit par un projet IA*
1.3.3.a	Savoir élaborer une stratégie de communication auprès des acteurs du projet IA
1.3.3.b	Savoir impliquer les directions métiers
1.3.3.c	Savoir traduire les besoins en termes de transfert de compétences
1.3.3.d	Savoir élaborer des plans de formation pour accompagner le changement

2	**COMPÉTENCES TECHNIQUES**
2.1	**ARCHITECTURE**
2.1.1	*Concevoir et mettre en œuvre les infrastructures big data*
2.1.1.a	Savoir planifier les moyens à mettre en œuvre de façon opérationnelle
2.1.1.b	Savoir contribuer aux choix d'architecture (on prem, cloud, edge, embarqué)
2.1.1.c	Savoir maîtriser le hardware (GPU, CPU, RAM, Compute) et être capable d'effectuer le dimensionnement
2.1.1.d	Savoir maîtriser les solutions de stockage, notamment NoSQL
2.1.1.e	Savoir mettre en place la containerisation, l'orchestration des différents outils et le monitoring
2.1.1.f	Savoir maîtriser le fonctionnement d'une infrastructure cloud majeure
2.1.1.g	Etre capable d'utiliser des machines sur le cloud (ajouter, supprimer, sauvegarder les machines
2.1.1.g	Savoir maîtriser les sauvegardes et restaurations des données
2.1.1.h	Savoir gérer les services managés cloud
2.1.2	*Concevoir l'architecture data*
2.1.2.a	Savoir maîtriser les architectures data-centric
2.1.2.b	Savoir définir une vision long terme et des principes directeurs
2.1.2.c	Connaitre les architectures et outils cloud, on prem et edge
2.1.2.d	Comprendre les principes de conception du cloud
2.1.2.e	Savoir définir l'architecture générale
2.1.1.f	Maîtriser les architectures fédérées de type data mesh
2.1.2.g	Maîtriser les solutions de federated learning
2.1.2.h	Maîtriser les bonnes pratiques de migration vers un cloud public
2.1.2.i	Connaître les limites des outils
2.1.2.j	Savoir où trouver les données dans les applicatifs de l'entreprise
2.1.2.k	Savoir s'interfacer avec l'ensemble des outils et applications
2.1.3	*Maitriser l'optimisation et le run*
2.1.3.a	Savoir mettre en oeuvre la mesure de la performance des traitements sur l'infrastructure cible
2.1.3.b	Savoir paralléliser les traitements
2.1.3.c	Savoir identifier les goulots d'étranglement et les traitements critiques
2.1.3.d	Savoir optimiser les traitements sur le cloud
2.1.4.e	Savoir optimiser les usages de l'architecture cloud
2.1.3.f	Savoir assurer le maintien en conditions opérationnelles et traiter les incidents
2.1.4	*Maitriser les coûts et la sécurité des architectures à base de cloud*
2.1.4.a	Savoir mesurer le coût du cloud
2.1.4.b	Comprendre le modèle de responsabilité partagée avec le fournisseur de cloud
2.1.4.c	Maitriser les bonnes pratiques pour assurer la sécurité des ressources dans le cloud
2.1.4.d	Savoir mettre en place des solutions de vérification de maîtrise du risque, de cryptage vers le cloud
2.1.5	*Programmer les traitements de données*
2.1.5.a	Maitriser le langage Python
2.1.5.b	Maitriser le langage C++
2.1.5.c	Maitriser le langage Java
2.1.5.d	Savoir utiliser du hardware spécialisé pour l'apprentissage (optimisation de TensorFlow/pytorch sur
2.2	**DONNÉES**
2.2.1	*Auditer les sources de données disponibles*
2.2.1.a	Etre capable de reconnaître le caractère éthique ou non éthique des données disponibles
2.2.1.b	Savoir identifier les données atypiques et/ou aberrantes
2.2.1.c	Savoir identifier et gérer les données manquantes
2.2.1.d	Savoir appliquer les règles du RGPD
2.2.2	*Collecter les données*
2.2.2.a	Savoir où collecter les données
2.2.2.b	Savoir collecter les données (ETL)
2.2.2.c	Savoir constituer un échantillon de données pour le data scientist
2.2.2.d	Savoir établir un dossier de sécurité (limites, habilitations, éthique, règlementaire)
2.2.2.e	Savoir gérer l'anonymisation et le cryptage
2.2.2.f	Savoir gérer l'ordonnancement de la collecte des données
2.2.3	*Auditer la qualité des données*
2.2.3.a	Maîtriser la statistique descriptive
2.2.3.b	Savoir utiliser la librairie Numpy en langage Python
2.2.3.c	Savoir utiliser la librairie Pandas en langage Python
2.2.3.d	Savoir réaliser un audit de qualité et de sécurité des données

2.2.4	*Préparer, nettoyer et programmer le recodage des données*
2.2.4.a	Savoir programmer le recodage des données
2.2.4.b	Savoir traiter les valeurs manquantes
2.2.4.c	Savoir traiter les valeurs aberrantes
2.2.4.d	Maîtriser le fonctionnement des bases de données SQL
2.2.4.e	Savoir créer des jeux de données de test
2.2.4.f	Savoir annoter/labelliser les données
2.2.5	*Augmenter les données*
2.2.5.a	Savoir identifier les sources de données open-source utiles
2.2.5.b	Savoir générer de nouvelles variables (feature engineering)
2.2.5.c	Savoir utiliser un simulateur / jumeau numérique
2.2.5.d	Savoir générer des jeux de données par simulateur
2.2.5.e	Savoir utiliser un modèle génératif (GAN)
2.3	**MODÉLISATION**
2.3.1	*Réaliser une analyse exploratoire des données*
2.3.1.a	Savoir utiliser les principales méthodes de réduction de dimension non supervisées
2.3.1.b	Savoir utiliser les principales méthodes de clustering et de typologies non supervisées
2.3.1.c	Savoir utiliser l'analyse statistique des corrélations
2.3.1.d	Savoir mettre en œuvre une visualisation des données
2.3.1.e	Savoir créer un jeu de données pertinent pour la cross-validation
2.3.1.f	Savoir analyser les biais existants dans les données d'apprentissage
2.3.1.g	Savoir éviter un biais dans les données d'apprentissage
2.3.1.h	Maîtriser les différentes méthodes de test d'hypothèses
2.3.2	***Concevoir des modèles analytiques d'IA et de machine-learning***
2.3.2.1	*Programmer les indicateurs nécessaires au fonctionnement des algorithmes*
2.3.2.1.a	Savoir utiliser la méthode MLOps pour mener à bien un projet
2.3.2.1.b	Savoir mettre en oeuvre une stratégie d'inférence
2.3.2.1.c	Savoir mettre en oeuvre du machine learning non supervisé
2.3.2.1.d	Savoir mettre en oeuvre du machine learning supervisé
2.3.2.1.e	Savoir programmer les indicateurs nécessaires
2.3.2.2	*Identifier et sélectionner des algorithmes de base à utiliser pour la modélisation*
2.3.2.2.a	Savoir identifier les différentes formes d'apprentissage machine
2.3.2.2.b	Savoir entraîner les algorithmes d'apprentissage
2.3.2.2.c	Savoir identifier les principaux algorithmes d'IA du domaine public
2.3.2.2.d	Savoir utiliser les algorithmes du domaine public pour créer un modèle
2.3.2.2.e	Savoir utiliser les modèles pré-entraînés du domaine public pour créer un modèle (IA générative, vision,
2.3.2.2.f	Savoir développer un modèle d'IA symbolique
2.3.2.2.g	Savoir développer un modèle multiagent
2.3.2.3	*Connaissance de techniques d'IA spécialisées*
2.3.2.2.a	Maîtriser les techniques de vision par ordinateur
2.3.2.2.b	Maîtriser les techniques de NLP
2.3.2.2.c	Maîtriser les techniques de reconnaissance vocale
2.3.2.4	*Maitriser les techniques d'IA avancées*
2.3.2.4.a	Savoir mettre en œuvre de l'apprentissage sur graphes/graphes de connaissances (GNNs)
2.3.2.4.b	Maîtriser les méthodes de reinforcement learning
2.3.2.4.c	Maîtriser les méthodes de federated learning
2.3.2.4.d	Savoir utiliser les knowledge informed neural networks (physical / geometric)
2.3.2.4.e	Savoir utiliser l'IA neuro-symbolique
2.3.2.4.f	Savoir concevoir et exploiter des ontologies
2.3.2.4.f2	Savoir concevoir et automatiser des modèles de raisonnement
2.3.2.4.f3	Savoir concevoir et mettre en œuvre des architectures d'IA distribuée
2.3.2.4.g	Connaître les méthodes d'optimisation
2.3.2.4.h	Savoir mettre en œuvre de la recherche opérationnelle
2.3.2.4.i	Maîtriser les technologies de edge AI

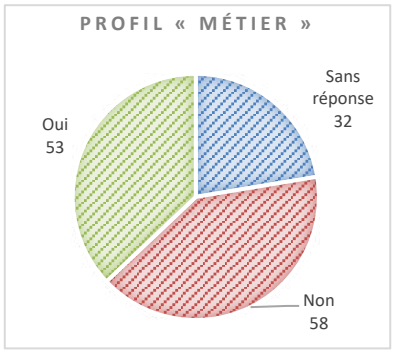
2.3.2.5	*Maitriser les techniques d'IA génératives*
2.3.2.5.a	Savoir générer des prompts et les mettre à disposition des métiers
2.3.2.5.b	Savoir utiliser les prompts et vérifier les informations retournées
2.3.2.5.c	Maîtriser les outils et techniques d'IA générative
2.3.2.5.d	Savoir identifier les limites des outils d'IA générative
2.3.2.5.e	Savoir adapter les modèles LLMs aux données de l'entreprise
2.3.2.6	*Prototyper les modèles d'IA*
2.3.2.6.a	Savoir utiliser les principales librairies d'IA en Python, notamment TensorFlow et PyTorch
2.3.2.6.b	Savoir assembler la chaîne de traitement des différents modèles
2.3.2.6.c	Savoir optimiser le compromis explicabilité & traçabilité / précision des algorithmes
2.3.2.6.d	Savoir documenter les modèles avec Jupyter Notebook
2.3.2.6.e	Savoir implémenter des pistes d'auditabilité et de traçabilité
2.3.2.6.f	Savoir s'assurer de la réversibilité du code
2.3.2.6.g	Maîtriser les outils low-code/no-code
2.3.2.6.h	Maîtriser les services cloud d'IA
2.3.2.6.i	Connaître les bases du Quantum Machine learning
2.3.3	***Tester et implémenter des modèles analytiques d'IA***
2.3.3.1	*Tester et décider de rejeter/adopter les modèles en mesurant & évaluant leur pertinence sur la base de leu*
2.3.3.1.a	Savoir mesurer la performance des modèles
2.3.3.1.b	Savoir créer les indicateurs pour valider ou rejeter les modèles testés
2.3.3.1.c	Savoir calculer et mesurer la variance des résidus d'un modèle quantitatif
2.3.3.1.d	Savoir produire la matrice de confusion d'un modèle de classification
2.3.3.1.e	Maîtriser les méthodes d'optimisation de complexité et de cross-validation des modèles
2.3.3.1.f	Savoir exploiter les résultats d'une stratégie inférentielle de cross-validation
2.3.3.1.g	Savoir s'assurer du respect des hypothèses de fonctionnement des modèles
2.3.3.1.h	Savoir s'assurer de l'absence de biais dans les résultats du modèle
2.3.3.1.i	Savoir mesurer la robustesse statistique des modèles
2.3.3.2	*Automatiser les modèles d'IA*
2.3.3.2.a	Savoir choisir les hyper-paramètres pour faciliter l'automatisation
2.3.3.2.b	Savoir automatiser l'optimisation des hyper paramètres
2.4	**INDUSTRIALISATION**
2.4.1	*Gérer le cycle de vie des modèles*
2.4.1.a	Savoir documenter les modèles
2.4.1.b	Savoir construire un pipeline de modèles, basé sur une architecture complexe
2.4.1.c	Savoir gérer le versionning des données
2.4.1.d	Savoir gérer le versionning des modèles
2.4.1.e	Savoir mettre en place un feature store
2.4.1.f	Savoir produire des templates réutilisables
2.4.1.g	Savoir gérer le suivi des expérimentations
2.4.1.h	Savoir mettre en œuvre le réentrainement des modèles
2.4.1.i	Savoir gérer l'explicabilité et la prise en compte des biais
2.4.2	*Industrialiser, sécuriser, rendre résilients et utilisables les modèles de datascience*
2.4.2.a	Savoir spécifier les conditions d'utilisation des modèles d'IA
2.4.2.b	Maîtriser la méthode devops en lien avec la méthode MLOps (dataops / modelOps)
2.4.2.c	Savoir orchestrer un ensemble d'algorithmes
2.4.2.d	Savoir intégrer les résultats dans les applications digitales de l'entreprise
2.4.2.e	Savoir documenter une application
2.4.2.f	Savoir intégrer la sécurité dans la chaîne data et dans les environnements d'entraînement
2.4.3	*Gérer la supervision et le monitoring*
2.4.3.a	Maîtriser les méthodes statistiques de monitoring des modèles d'IA
2.4.3.b	Savoir identifier et analyser les dérives des modèles
2.4.3.c	Savoir configurer les seuils d'alertes en cas de dysfonctionnement des modèles
2.4.3.d	Savoir mesurer les performances des modèles
2.4.3.e	Savoir mettre en place le monitoring fonctionnel du feedback utilisateur (boucle de rétroaction)
2.4.3.f	Savoir mettre en place une maintenance corrective et/ou évolutive (continuous training)
2.4.3.g	Savoir monitorer les processus de réapprentissage automatique
2.4.3.h	Savoir réaliser la supervision technique/MCO

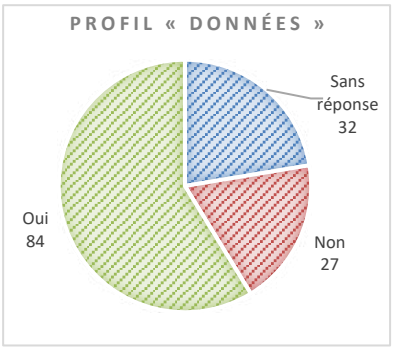
2.4.4	*Tester et déployer*
2.4.4.a	Savoir réaliser les tests d'intégration et tests système
2.4.4.b	Savoir mettre en place les tests d'acceptation des utilisateurs
2.4.4.c	Savoir mettre en place les tests de performance et de charge
2.4.4.d	Savoir maîtriser la chaine CI/CD (continuous integration / continuous development)
2.4.4.e	Savoir mettre en œuvre le pipeline de collecte des données
2.4.4.f	Savoir définir et mettre en œuvre la stratégie de déploiement des nouveaux modèles (amélioration continu
2.5	**TRANSFERT AU MÉTIER**
2.5.1	*Restituer aux interlocuteurs métiers les résultats des modèles*
2.5.1.a	Savoir effectuer du storytelling autour de la donnée
2.5.1.b	Être capable de simplifier la communication à des fins de vulgarisation
2.5.1.c	Savoir présenter des visualisations des données compréhensibles par les métiers
2.5.1.d	Savoir illustrer les résultats du modèle par des graphiques compréhensibles par les métiers
2.5.1.e	Savoir sélectionner les observations illustrant le comportement typique des modèles
2.5.1.f	Savoir interpréter les résultats des différents modèles et en expliquer les résultats
2.5.1.g	Savoir utiliser des modèles explicatifs pour pouvoir rendre traçables / explicables les modèles
2.5.1.h	Savoir expliquer les risques des modèles
2.5.1.i	Savoir présenter et afficher les biais des modèles
2.5.1.j	Savoir mesurer les écarts/non conformités et les communiquer
2.5.1.k	Savoir mesurer les impacts des résultats

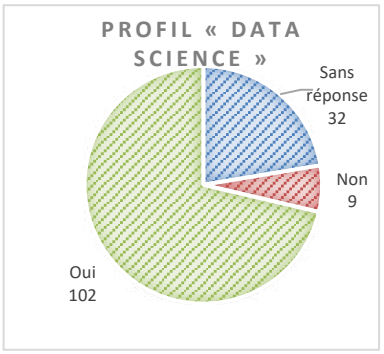
Annexe 5 – Résultats de l'enquête : les ressources

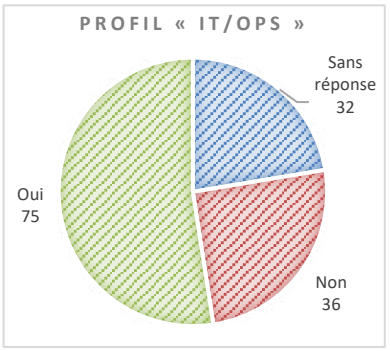
Les ressources IA par type de profil, internes et externes confondues

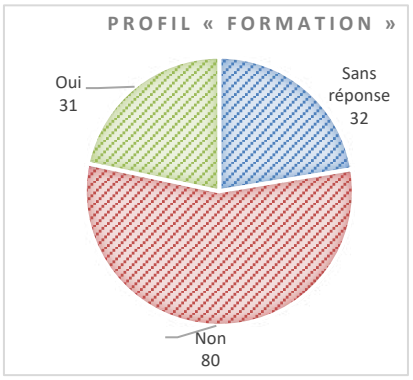
Les entreprises déclarent avoir des ressources IA en interne, avec les profils suivants : métier, données, data science, IT/Ops, formation, chef de projet, conduite du changement, AMOA (assistance à maîtrise d'ouvrages), autre. Les profils data science, données et IT /Ops sont presque toujours présents.

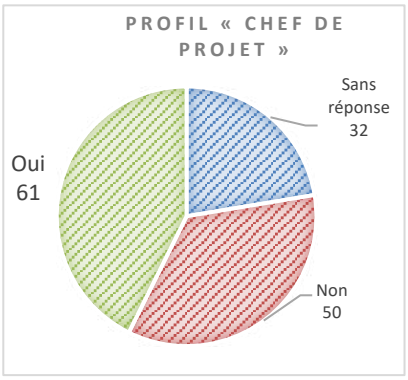


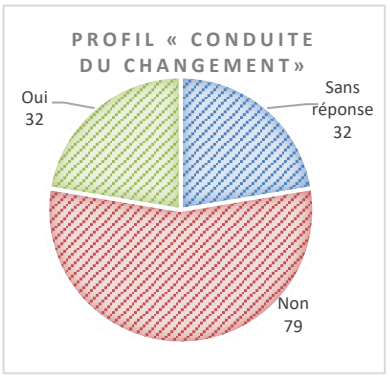


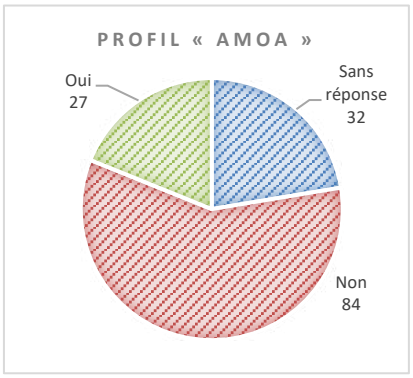






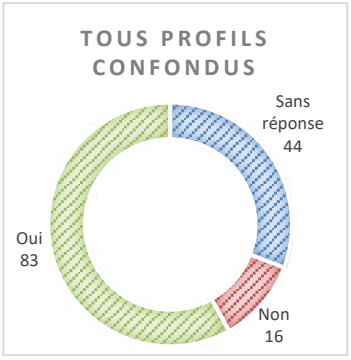


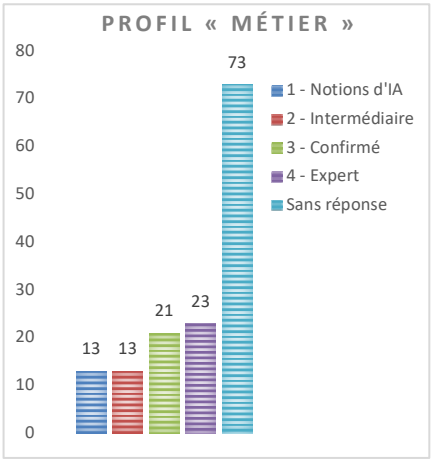


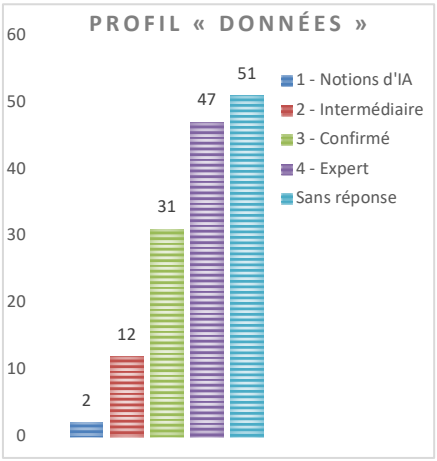


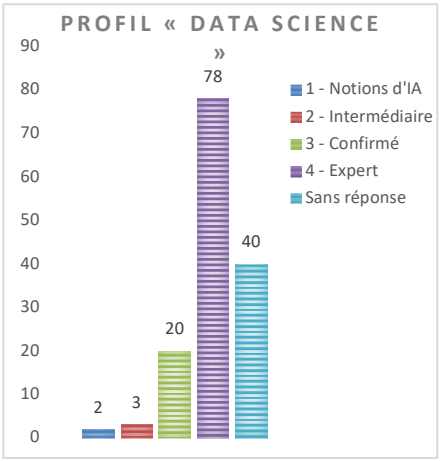
Les ressources IA : formation en IA

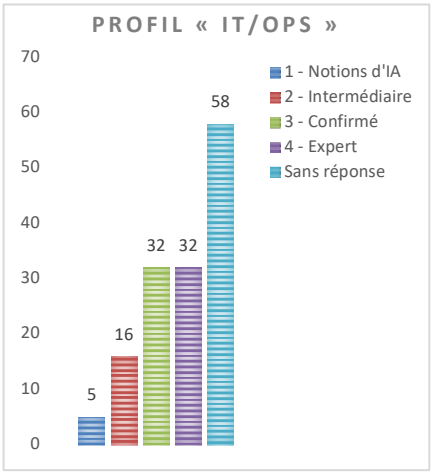
Près de 60% des répondants déclarent que leurs collaborateurs ont suivi des formations, mais beaucoup ne savent pas indiquer leur niveau de formation. Ils disposent en général de beaucoup profils formés en data science, données et IP/OPS, moins en chef de projet et très peu en conduite du changement ou AMOA.

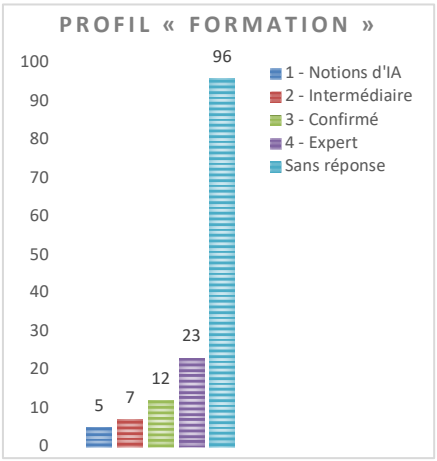


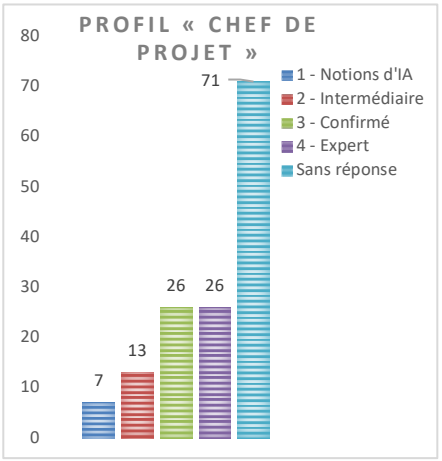

PROFIL « MÉTIER »
80
70
60
50
40
30
20
10
0
13
13
21
23
73
1 - Notions d'IA
2 - Intermédiaire
3 - Confirmé
4 - Expert
Sans réponse

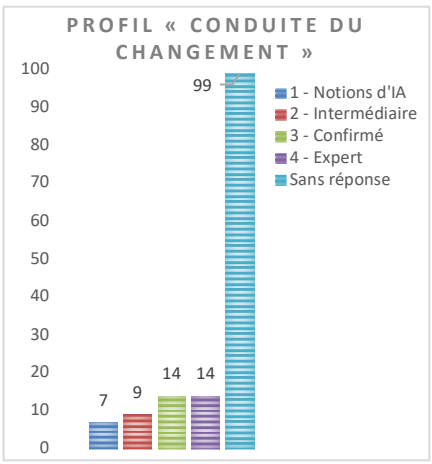

PROFIL « DONNÉES »
60
50
40
30
20
10
0
2
12
31
47
51
1 - Notions d'IA
2 - Intermédiaire
3 - Confirmé
4 - Expert
Sans réponse

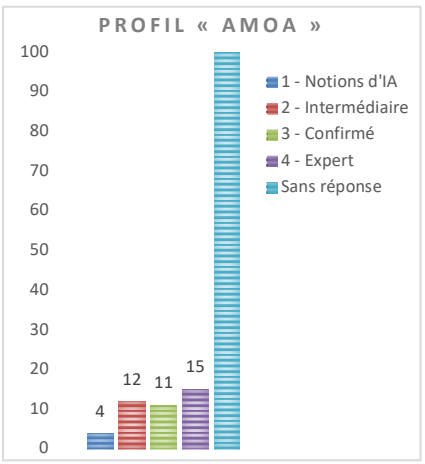

PROFIL « DATA SCIENCE »
90
80
70
60
50
40
30
20
10
0
2
3
20
78
40
1 - Notions d'IA
2 - Intermédiaire
3 - Confirmé
4 - Expert
Sans réponse


PROFIL « IT/OPS »
70
60
50
40
30
20
10
0
5
16
32
32
58
1 - Notions d'IA
2 - Intermédiaire
3 - Confirmé
4 - Expert
Sans réponse


PROFIL « FORMATION »
100
90
80
70
60
50
40
30
20
10
0
5
7
12
23
96
1 - Notions d'IA
2 - Intermédiaire
3 - Confirmé
4 - Expert
Sans réponse


PROFIL « CHEF DE PROJET »
80
70
60
50
40
30
20
10
0
7
13
26
26
71
1 - Notions d'IA
2 - Intermédiaire
3 - Confirmé
4 - Expert
Sans réponse


PROFIL « CONDUITE DU CHANGEMENT »
100
90
80
70
60
50
40
30
20
10
0
7
9
14
14
99
1 - Notions d'IA
2 - Intermédiaire
3 - Confirmé
4 - Expert
Sans réponse


PROFIL « AMOA »
100
90
80
70
60
50
40
30
20
10
0
4
12
11
15
1 - Notions d'IA
2 - Intermédiaire
3 - Confirmé
4 - Expert
Sans réponse

Annexe 6 – Résultats de l'enquête : besoins en compétences

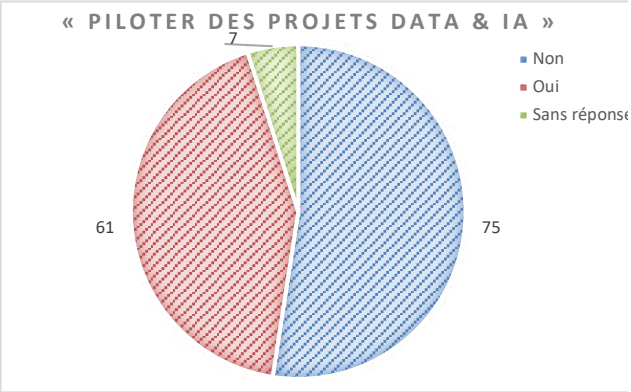
Pour chaque macro-compétence, il est demandé si l'entreprise a besoin de monter en compétence, et dans ce cas de préciser le nombre de collaborateurs concernés, le niveau d'expertise souhaité, le délai (<1 an, > 1 an) et la source de la compétence (interne/externe).

Compétences transverses

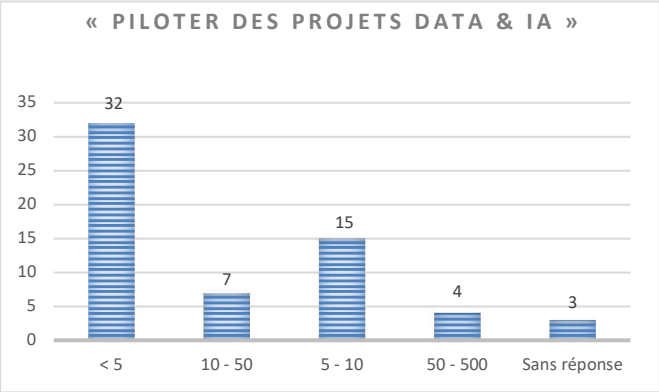
Les entreprises indiquent leurs priorités :

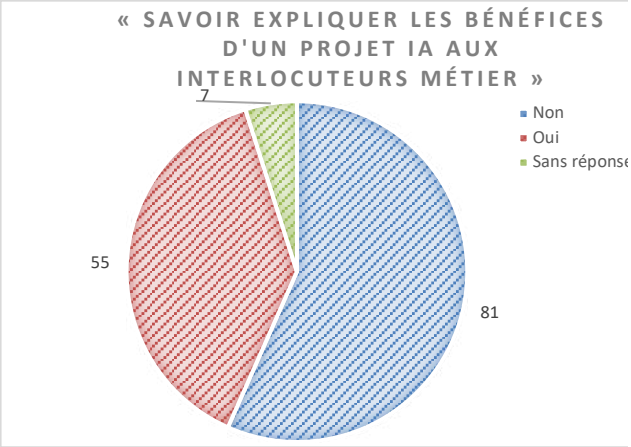
- Connaître les enjeux de conformité et réglementaire de l'IA pour 48,25% (sans doute en préparation de l'IA Act) ;
- Piloter des projets data et IA pour 42,66% ;
- Connaître les problématiques éthiques de l'IA pour 41,96% ;
- Savoir expliquer les bénéfices d'un projet IA aux interlocuteurs métier pour 38,46% ;
- Savoir mettre en place une approche sobre de l'IA pour 37,76%.

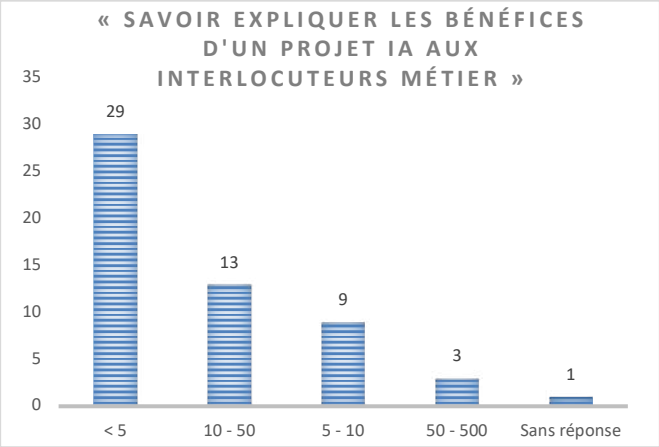
Nombre d'entreprises ayant besoin de monter en compétence sur la compétence



Volumétrie du besoin en compétence

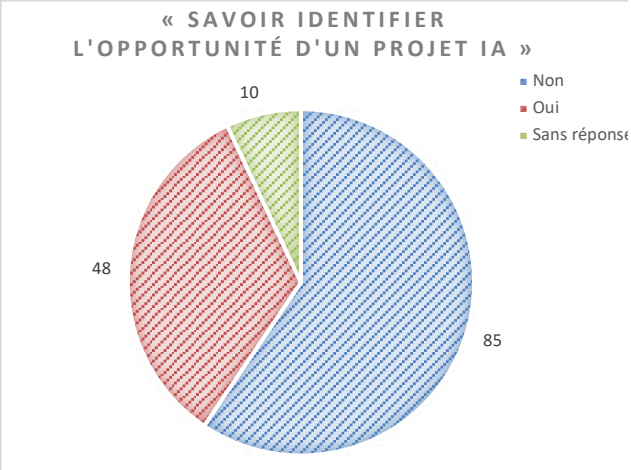


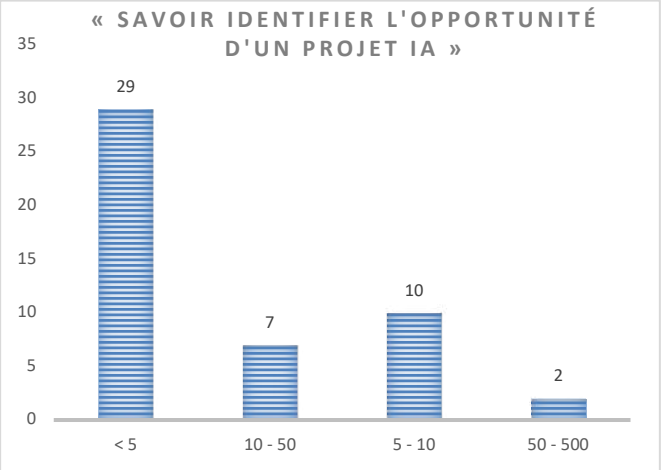


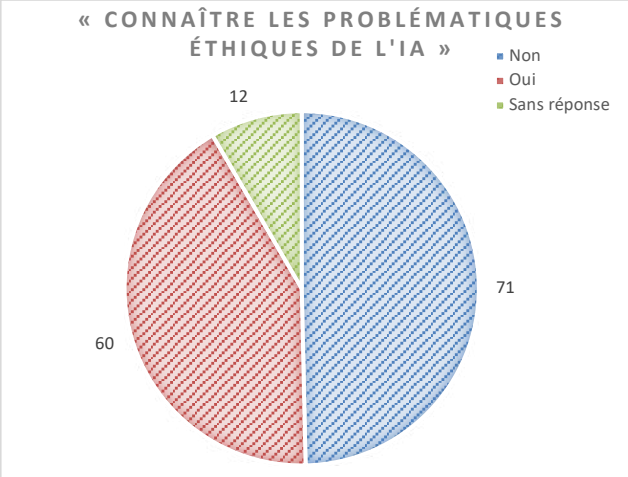


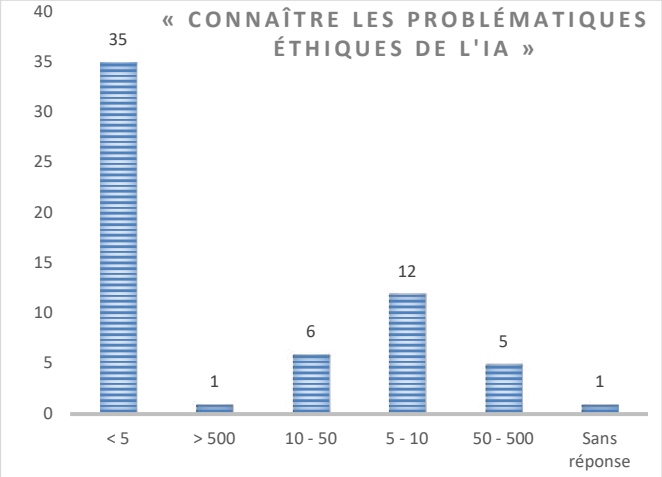
Nombre d'entreprises ayant besoin de monter en compétence sur la compétence

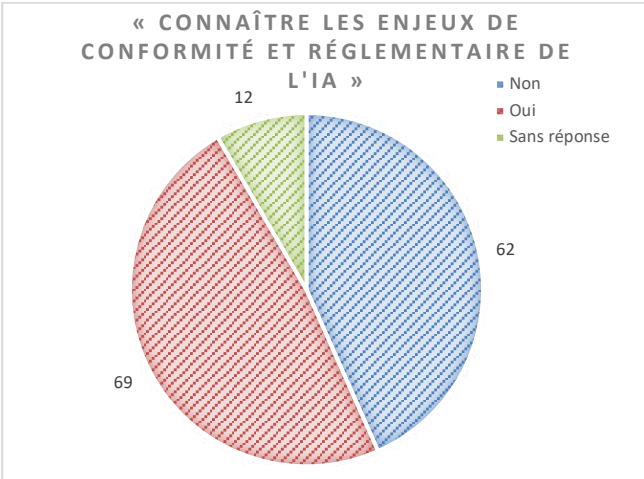
Volumétrie du besoin en compétence

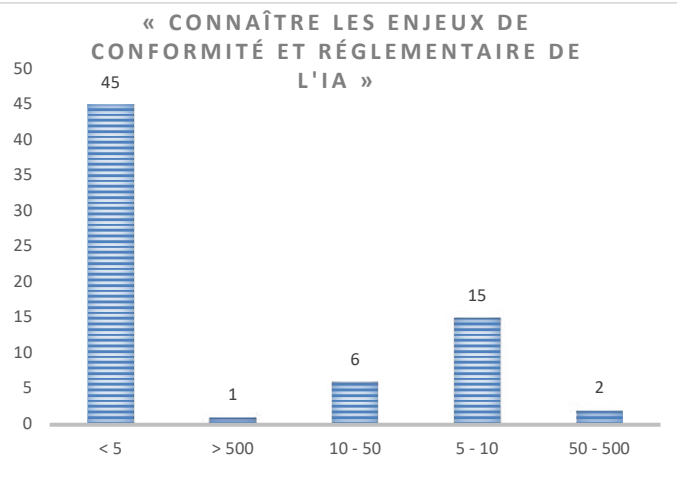






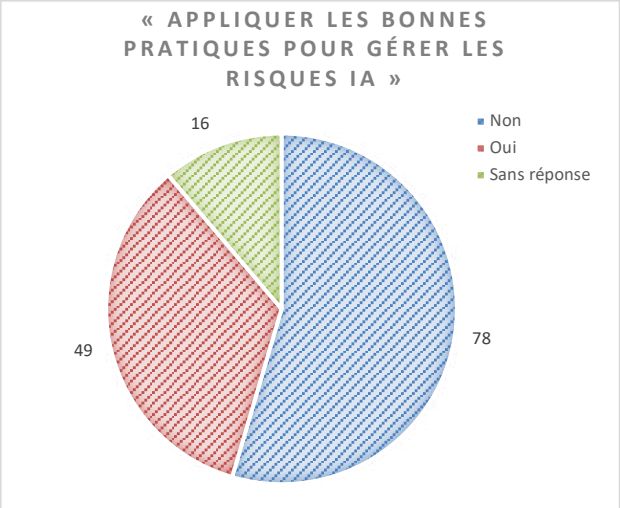


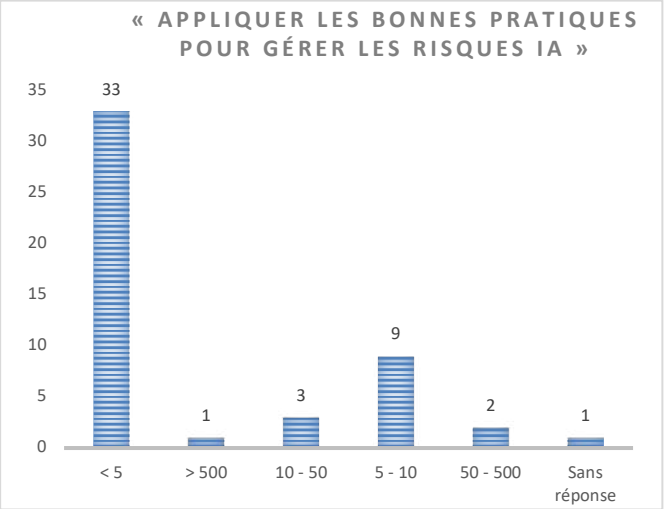


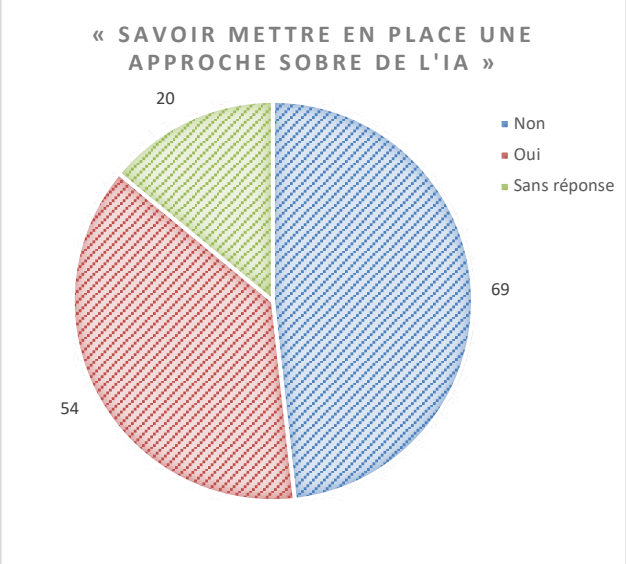


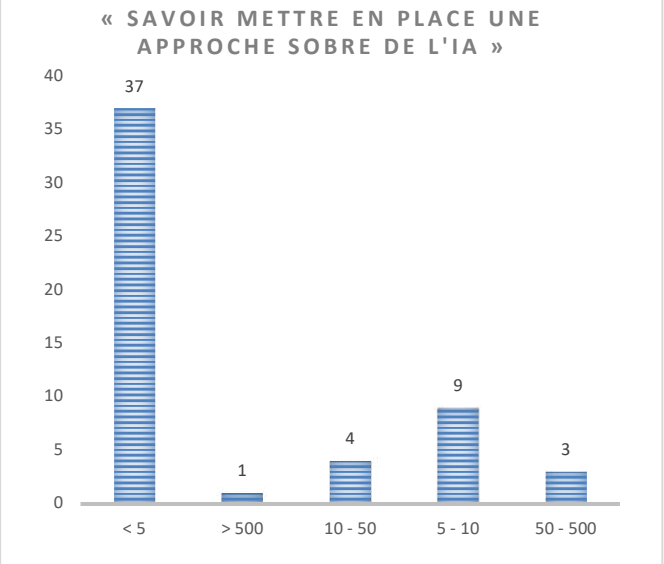
Nombre d'entreprises ayant besoin de monter en compétence sur la compétence

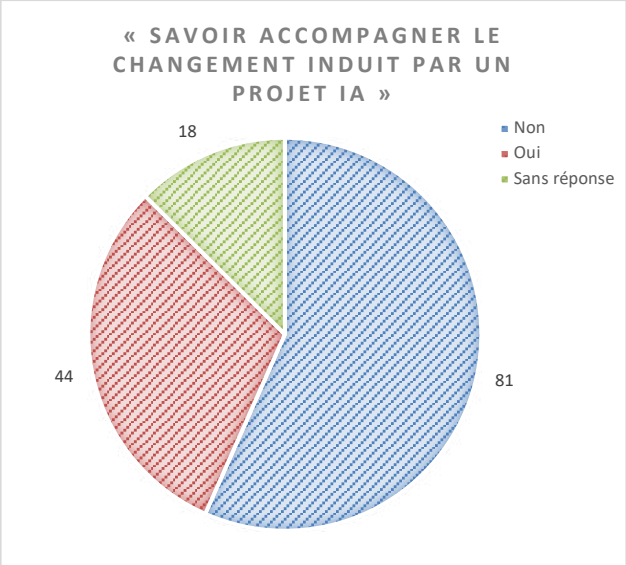
Volumétrie du besoin en compétence

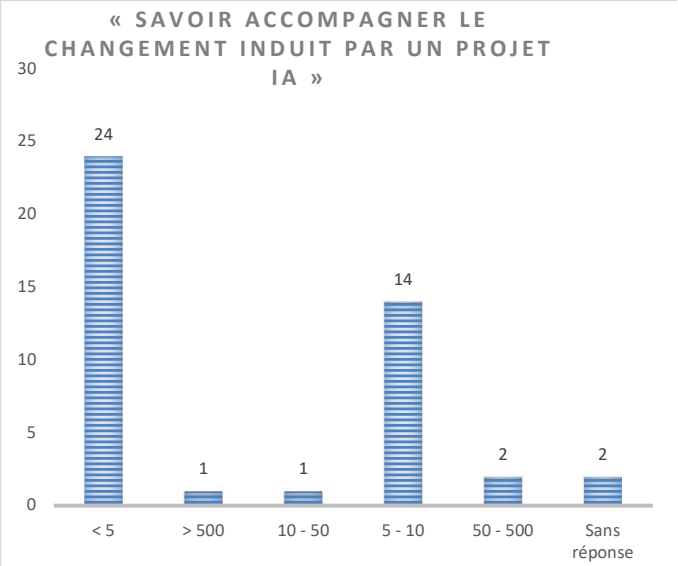




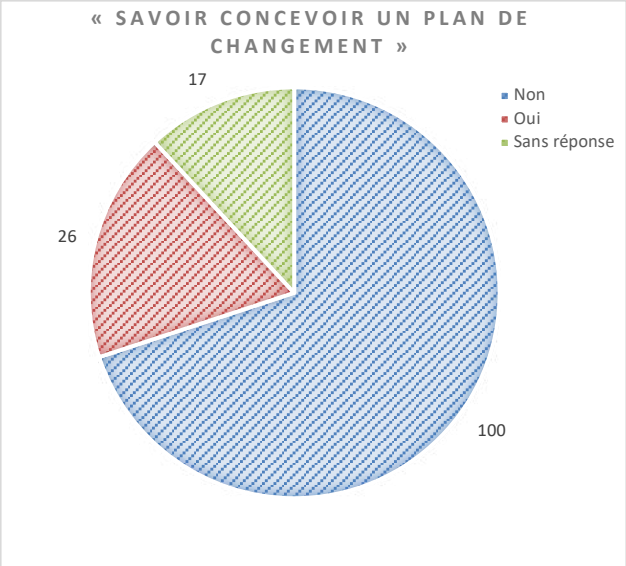




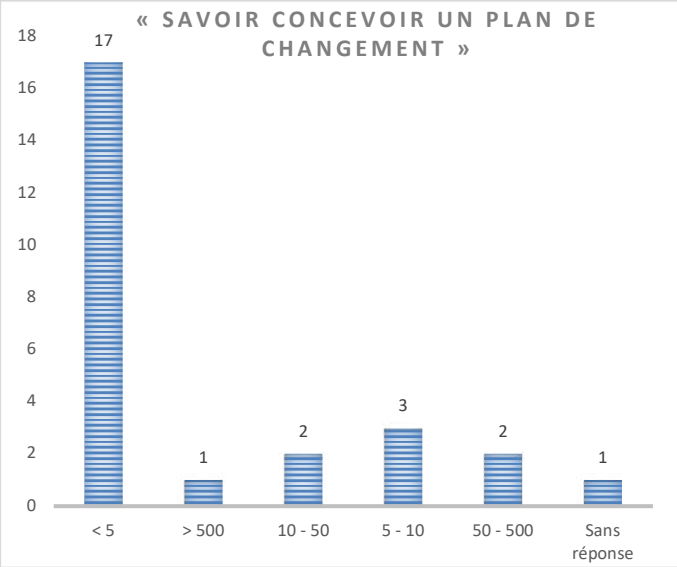


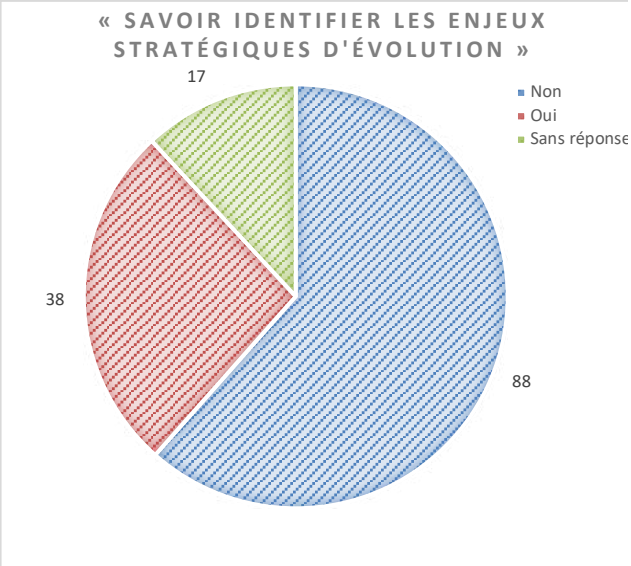


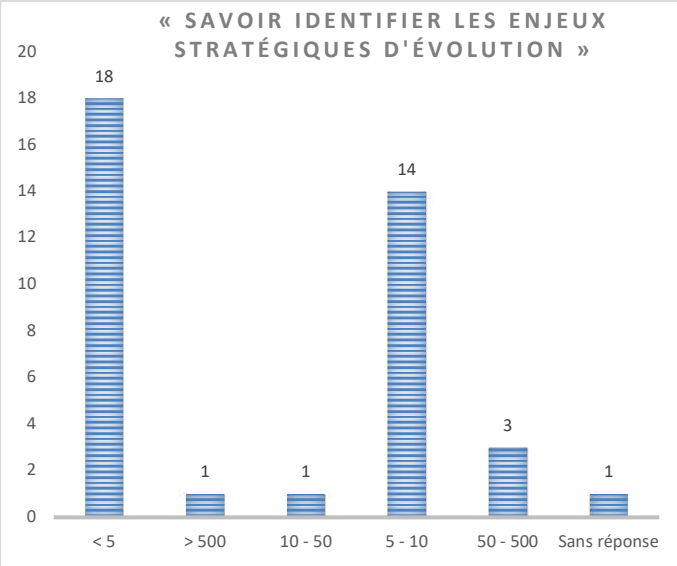
Nombre d'entreprises ayant besoin de monter en compétence sur la compétence



Volumétrie du besoin en compétence







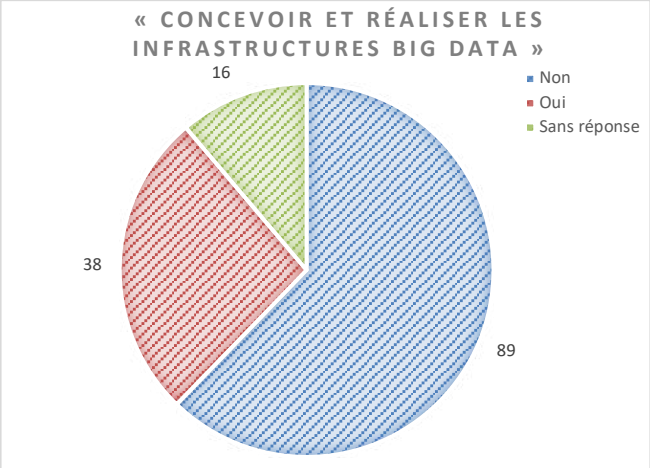
Compétences techniques

Les entreprises indiquent leurs priorités :

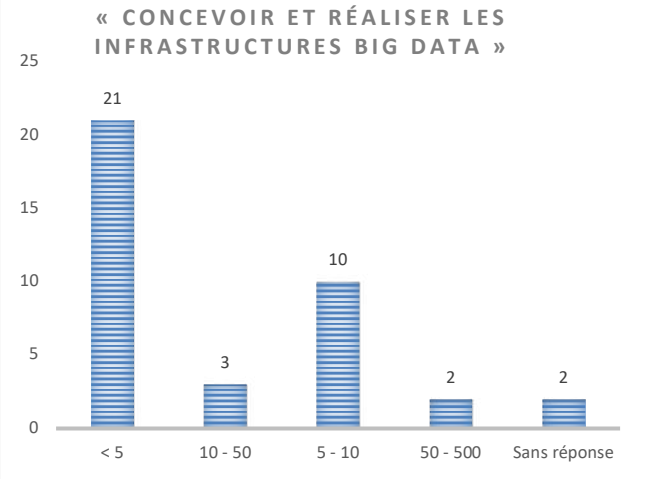
- Restituer aux interlocuteurs métiers les résultats des modèles pour 28,67% est la compétence la plus demandée ;
- Programmer les indicateurs nécessaires au fonctionnement des algorithmes pour 27,27% ;
- Concevoir et réaliser les infrastructures big data pour 26,57%
- Augmenter les données pour 26%
- Identifier et sélectionner des algorithmes de base à utiliser pour la modélisation pour 25%
- Maintenir les modèles et algorithmes en état de fonctionnement pour 25%

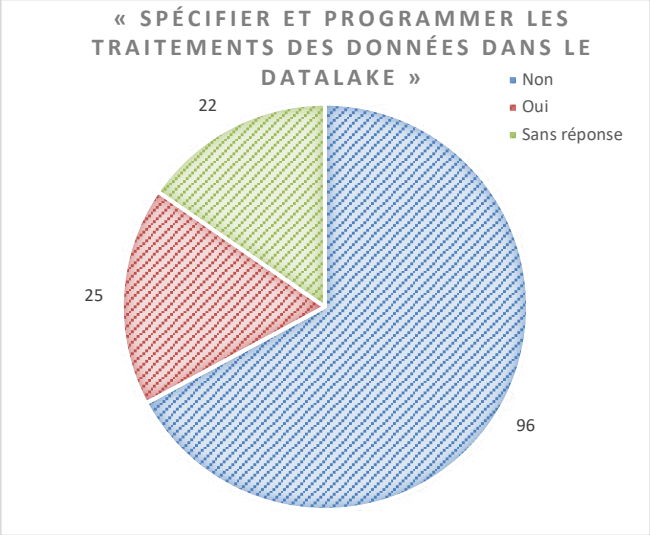
Notons que ces pourcentages sont beaucoup plus faibles que pour les compétences transverses prioritaires (aux alentours de 50%) : ceci reflète certainement la maturité les répondants qui ont déjà mis en place les compétences nécessaires et ne font qu'envisager les formations complémentaires ou les nouveaux recrutements.

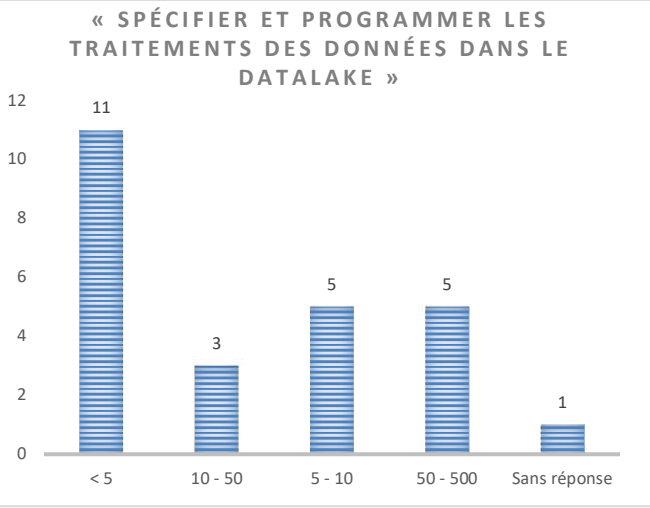
Nombre d'entreprises ayant besoin de monter en compétence sur la compétence



Volumétrie du besoin en compétence

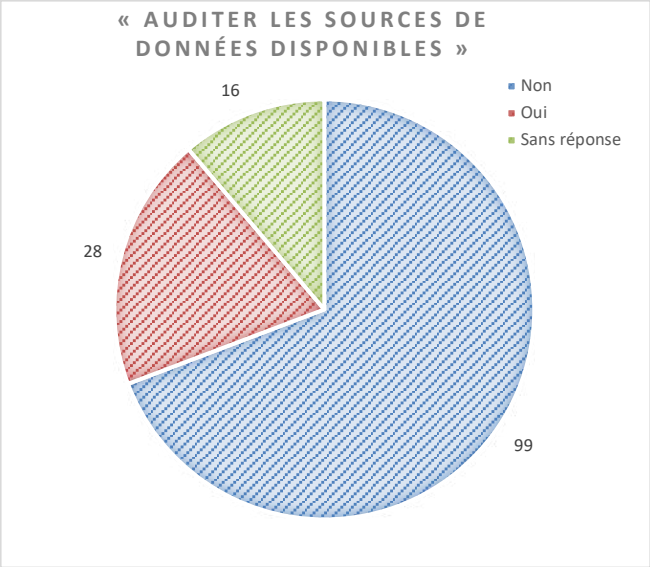


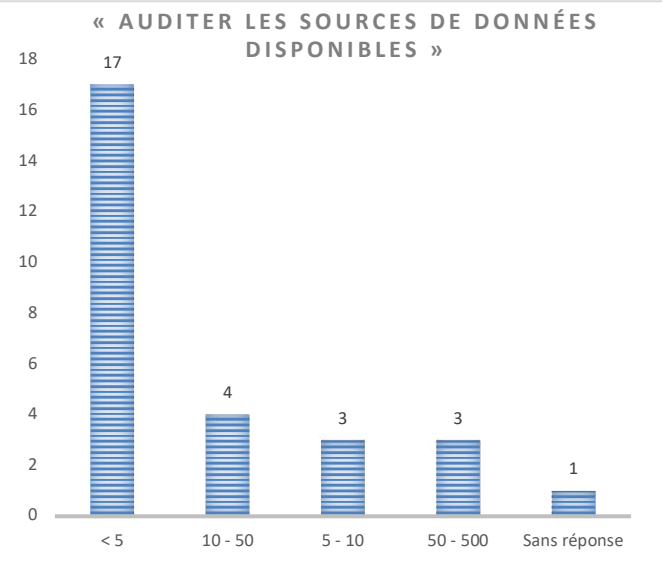


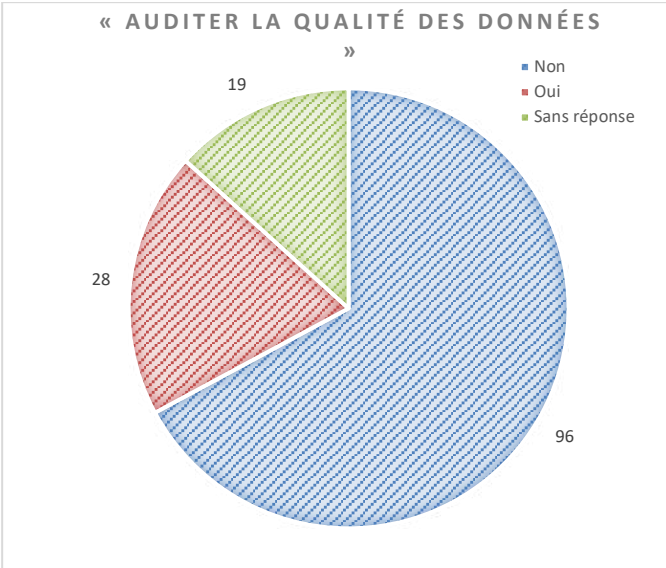


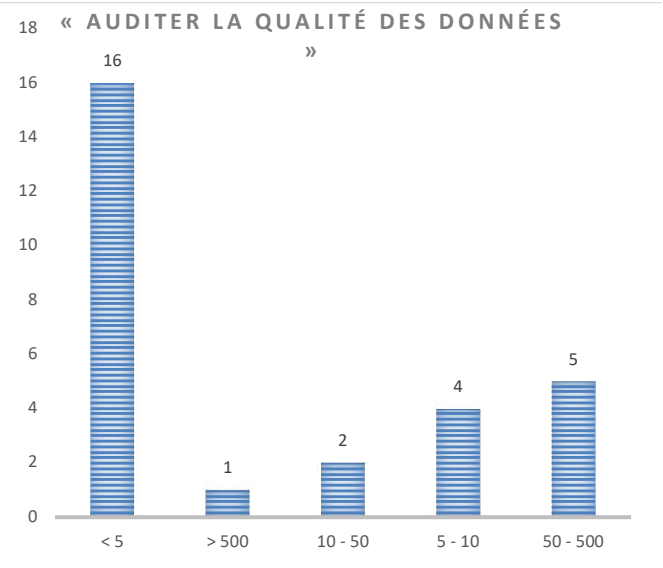
Nombre d'entreprises ayant besoin de monter en compétence sur la compétence

Volumétrie du besoin en compétence



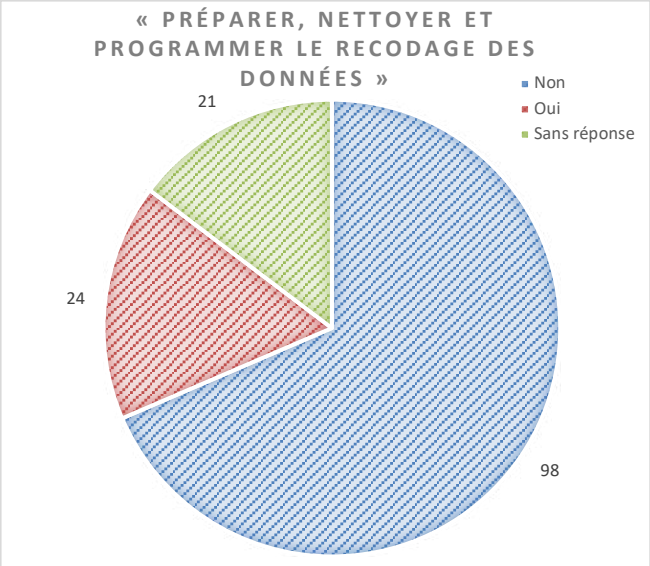


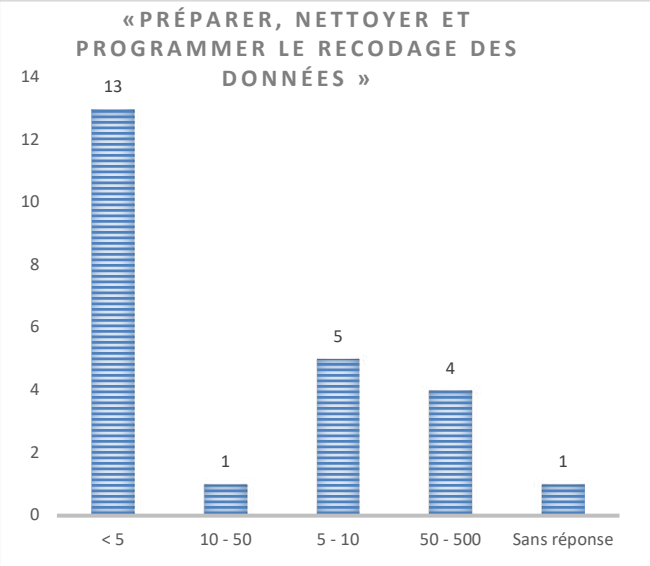


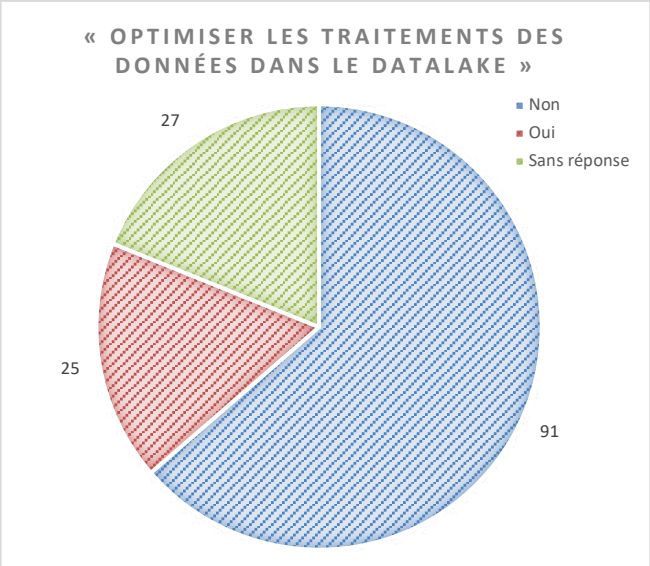


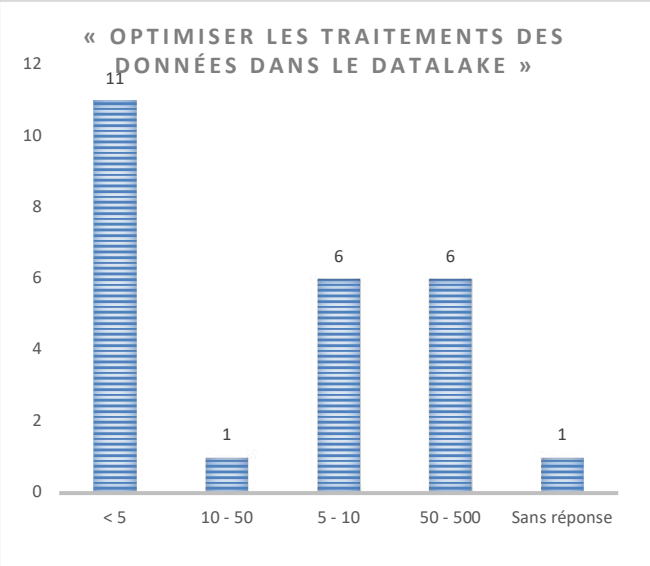
Nombre d'entreprises ayant besoin de monter en compétence sur la compétence

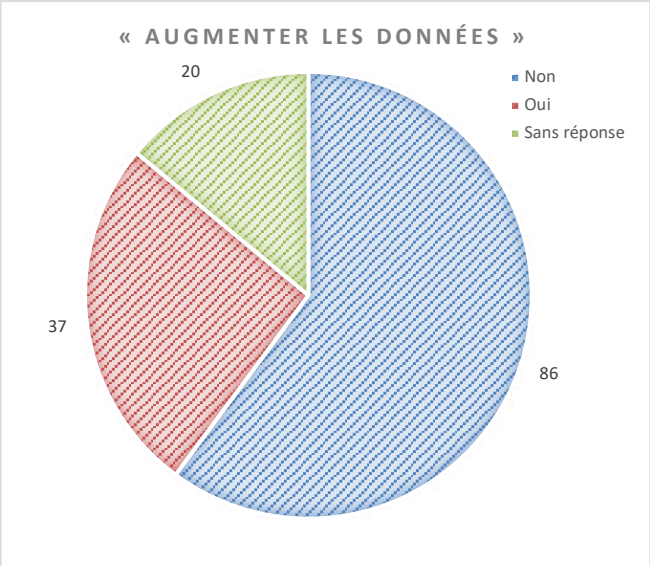
Volumétrie du besoin en compétence

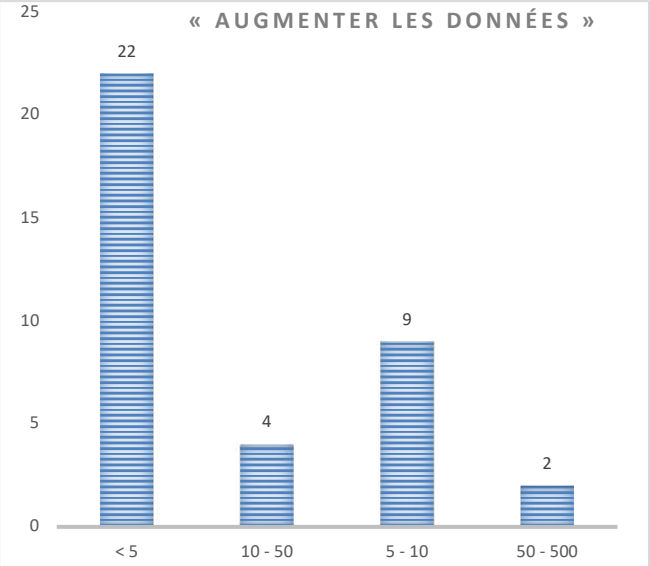






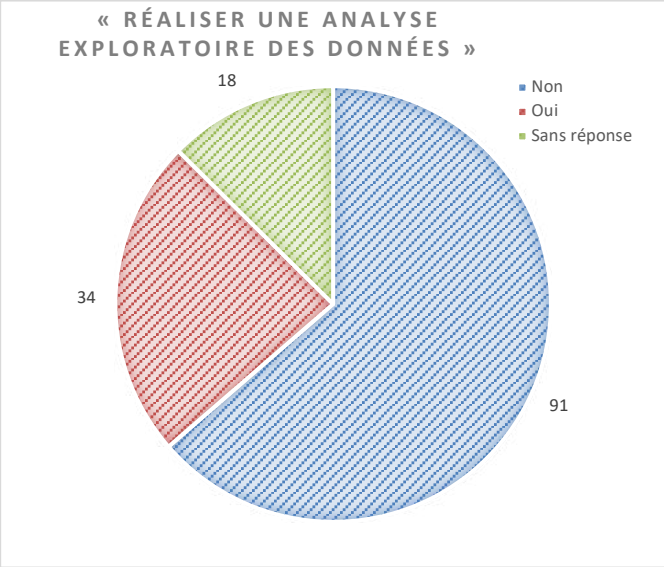


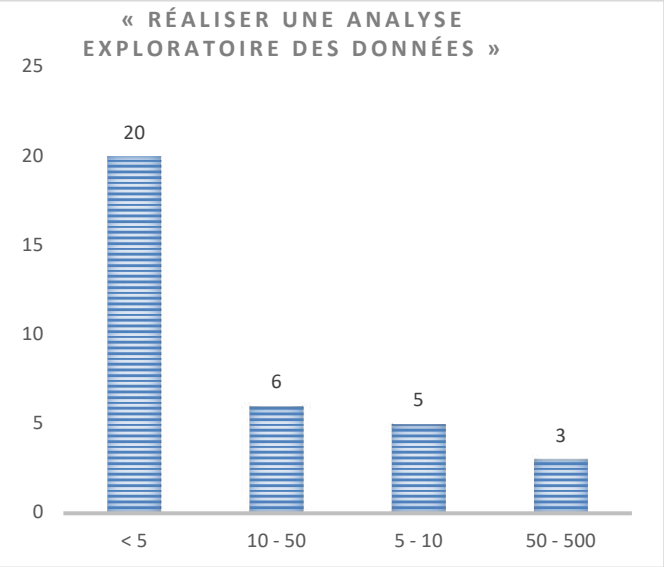


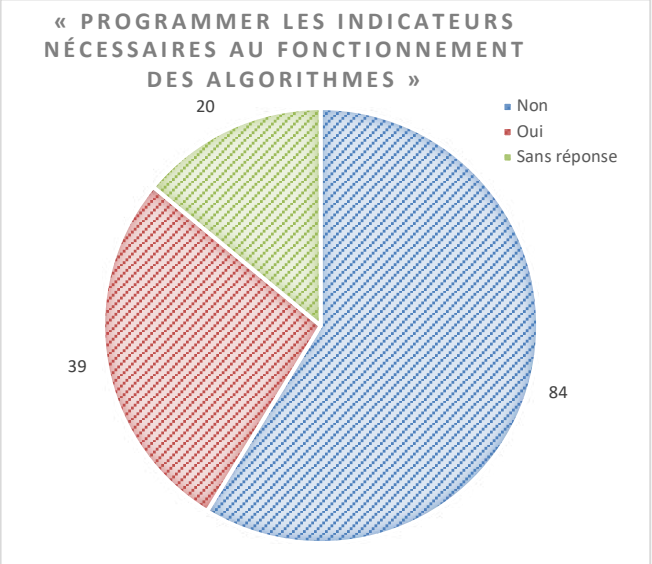


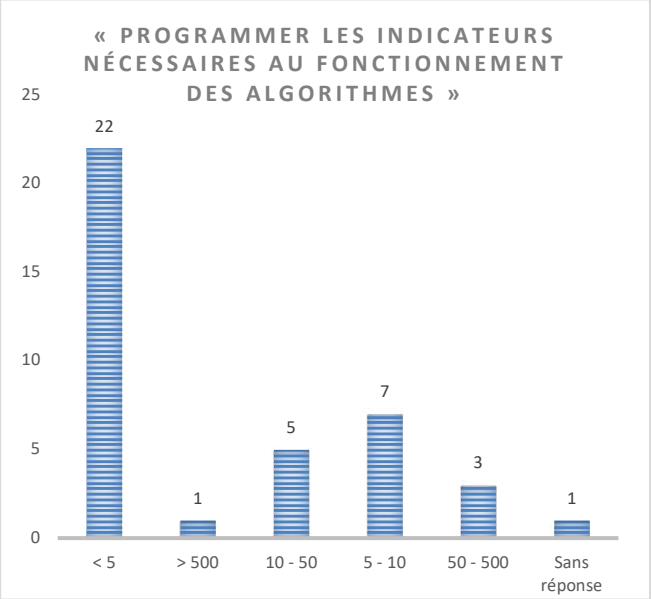
Nombre d'entreprises ayant besoin de monter en compétence sur la compétence

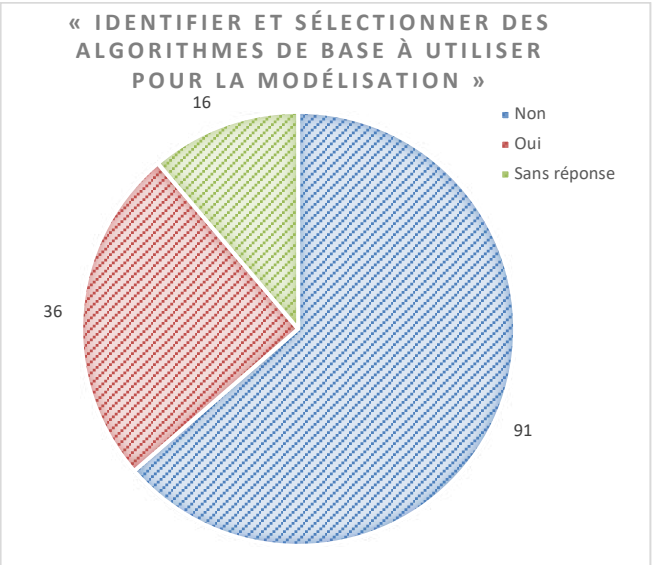
Volumétrie du besoin en compétence

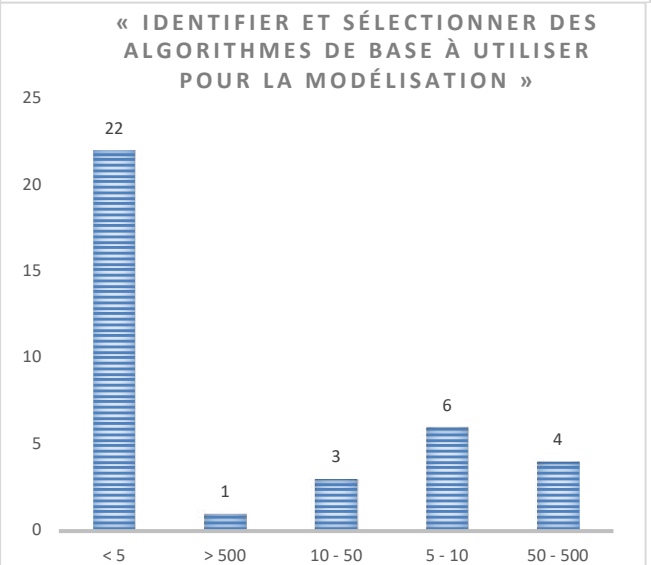




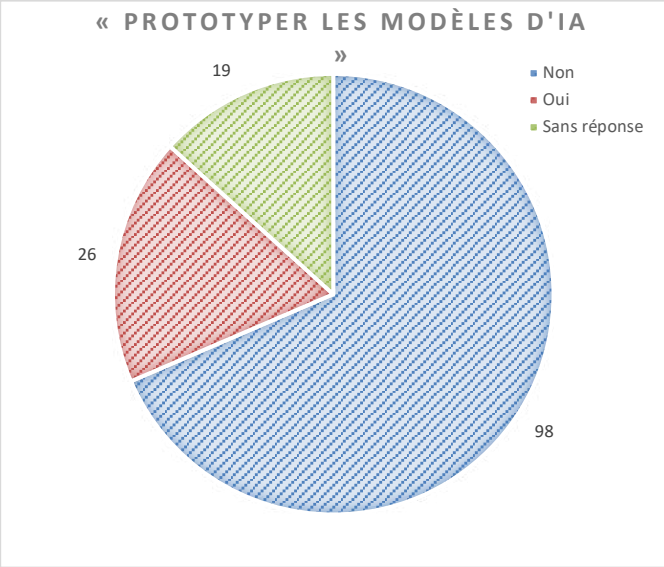




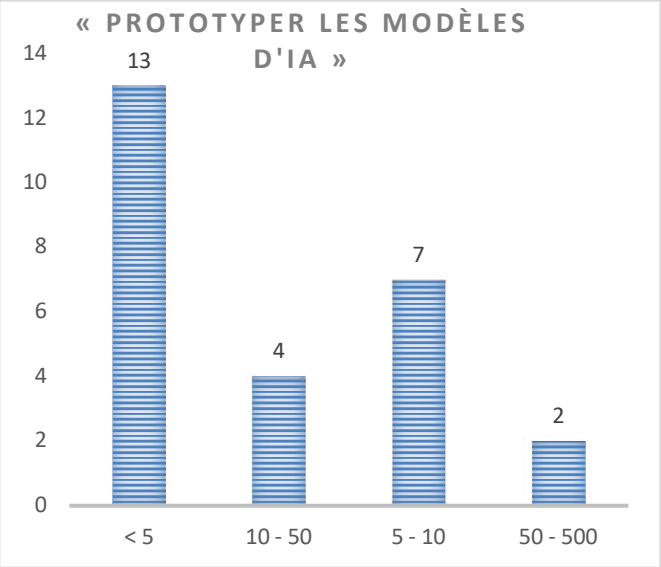


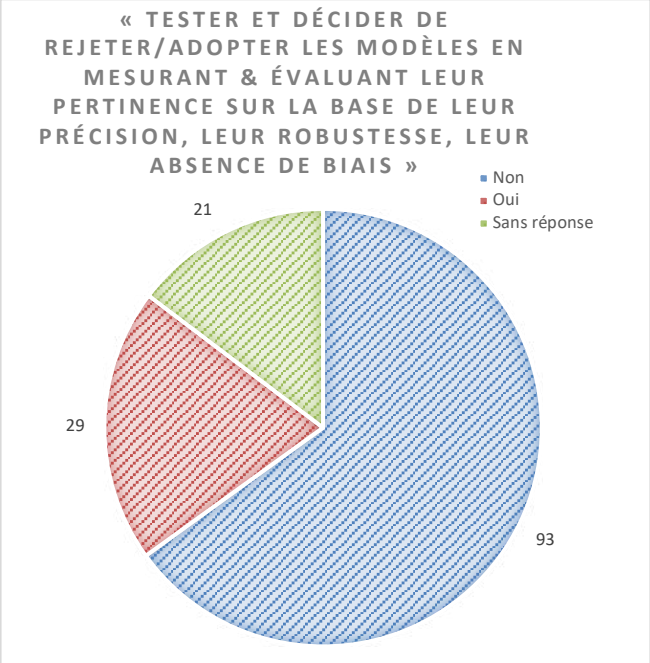


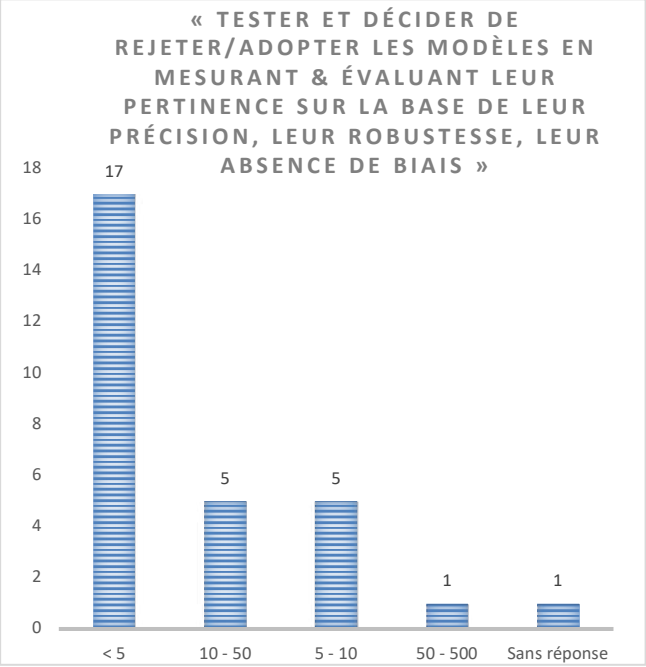
Nombre d'entreprises ayant besoin de monter en compétence sur la compétence



Volumétrie du besoin en compétence

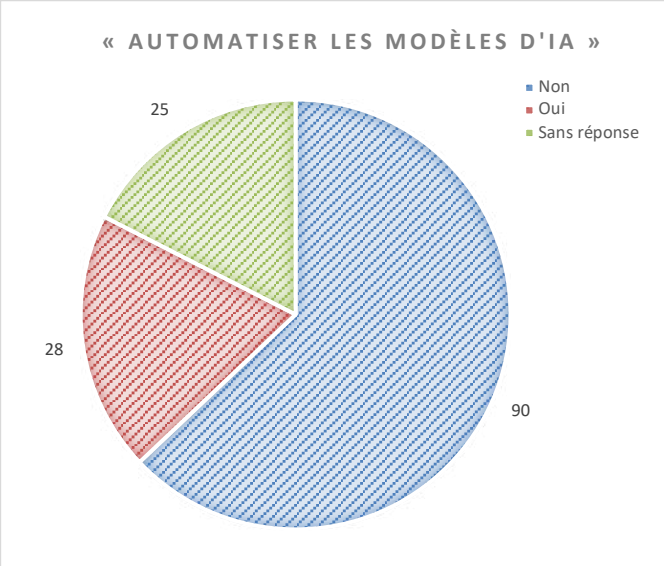


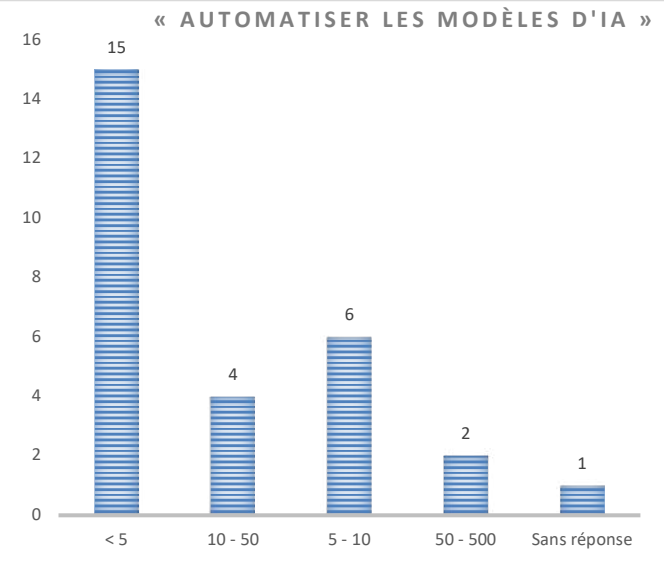


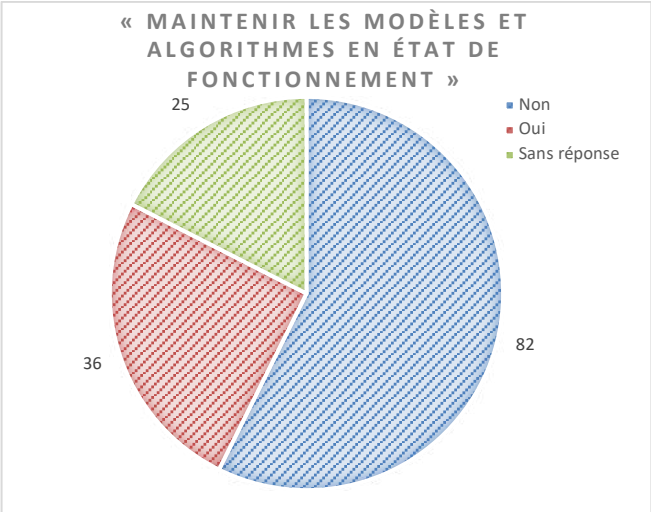


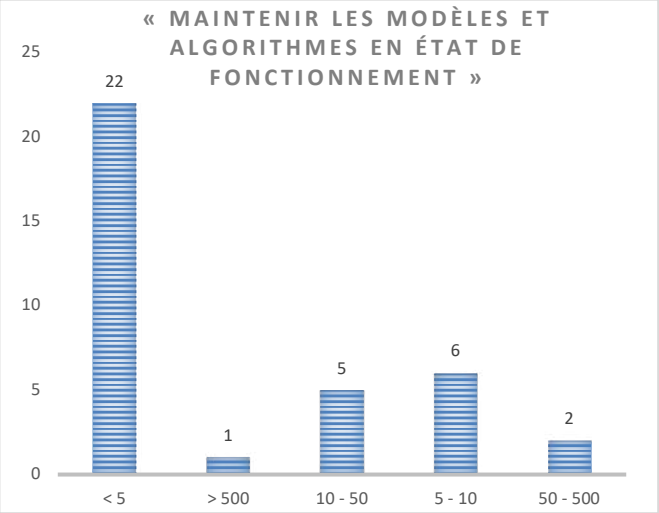
Nombre d'entreprises ayant besoin de monter en compétence sur la compétence

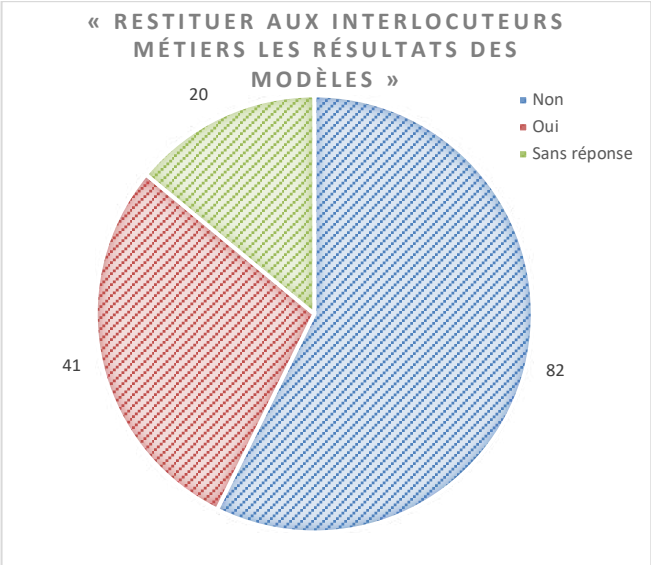
Volumétrie du besoin en compétence

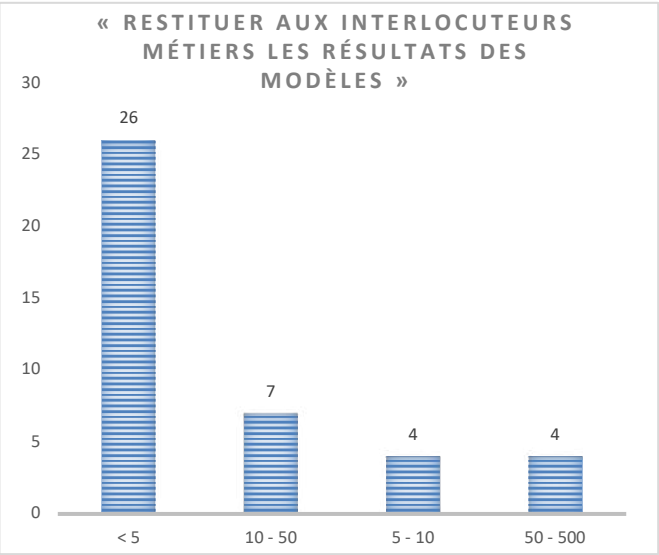












Annexe 7 – Tableau récapitulatif des besoins de montée en compétences

Comme discuté précédemment (Annexe 6), les besoins exprimés les plus nombreux sont en matière de compétences transverses (voir les rangs indiqués). En particulier, les liens avec le métier (faisabilité des besoins métier, conduite du changement) semblent avoir été particulièrement ciblés par les entreprises, ce qui correspond à leur déclarations (annexe 5) indiquant qu'ils ne disposent pas de profil conduite du changement ou AMOA.

TR TE*	Famille de compétences	Compétence	Taux **	Rang
TR	Pilotage de projets IA	*Piloter des projets data et IA*	42,66%	2
TR	Évaluer la faisabilité d'un besoin métier	*Savoir expliquer les bénéfices d'un projet IA aux interlocuteurs métier*	38,46%	4
		Savoir identifier l'opportunité d'un projet IA	33,57%	7
		Connaître les problématiques éthiques de l'IA	41,96%	3
		Connaître les enjeux de conformité et réglementaire de l'IA	48,25%	1
		Appliquer les bonnes pratiques pour gérer les risques IA	34,27%	6
		Savoir mettre en place une approche sobre de l'IA	37,76%	5
TR	Mettre en place la conduite du changement	*Savoir identifier les enjeux stratégiques d'évolution*	26,57%	11
		Savoir concevoir un plan de changement	18,18%	23
		Savoir accompagner le changement induit par un projet IA	30,77%	8
TE	Architecture	*Concevoir et réaliser les infrastructures big data*	26,57%	12
		Spécifier et programmer les traitements des données dans le datalake	17,48%	25
		Optimiser les traitements des données dans le datalake	17,48%	26
TE	Données	*Auditer les sources de données disponibles*	19,58%	20
		Auditer la qualité des données	19,58%	21
		Préparer, nettoyer et programmer le recodage des données	16,78%	27
		Augmenter les données	25,87%	14
TE	Modélisation	*Réaliser une analyse exploratoire des données*	23,78%	17
		Programmer les indicateurs nécessaires au fonctionnement des algorithmes	27,27%	10
		Identifier et sélectionner les algorithmes de base à utiliser pour la modélisation	25,17%	15
		Prototyper les modèles d'IA	18,18%	24
		Tester et décider de rejeter/adopter les modèles en mesurant et évaluant leur pertinence sur la base de leur précision, leur robustesse, leur absence de biais	20,28%	19
		Automatiser les modèles d'IA	19,58%	22
TE	Industrialisation	*Optimiser les modèles d'IA*	20,98%	18
		Industrialiser, sécuriser, rendre résilients et utilisables les modèles d'IA	26,57%	13
		Maintenir les modèles et algorithmes en état de fonctionnement	25,17%	16
TE	Transfert au métier	*Restituer aux interlocuteurs métier les résultats des modèles*	28,67%	9

* TR/ TE : Transverse Technique

** Taux d'entreprises ayant exprimé le besoin de monter en compétence

Les volumes des besoins en compétence sont indiqués ci-dessous :

Transverse Technique (TR/TE)	Famille de compétences	Compétence	Rang	<5	5-10	10-50	50-500	>500
TR	Pilotage de projets IA	*Piloter des projets data et IA*	2	32	15	7	4	0
TR	Évaluer la faisabilité d'un besoin métier	*Savoir expliquer les bénéfices d'un projet IA aux interlocuteurs métier*	4	29	9	13	3	0
		Savoir identifier l'opportunité d'un projet IA	7	29	10	7	2	0
		Connaître les problématiques éthiques de l'IA	3	35	12	6	5	1
		Connaître les enjeux de conformité et réglementaire de l'IA	1	45	15	6	2	1
		Appliquer les bonnes pratiques pour gérer les risques IA	6	33	9	3	2	1
		Savoir mettre en place une approche sobre de l'IA	5	37	9	4	3	1
	Mettre en place la co	*Savoir identifier les enjeux stratégiques d'évolution*	11	18	14	1	3	1
TR		*Savoir concevoir un plan de changement*	23	17	3	2	2	1
		Savoir accompagner le changement induit par un projet IA	8	24	14	1	2	1
TE	Architecture	*Concevoir et réaliser les infrastructures big data*	12	21	10	3	2	0
		Spécifier et programmer les traitements des données dans le datalake	25	11	5	3	5	0
		Optimiser les traitements des données dans le datalake	26	11	6	1	6	0
TE	Données	*Auditer les sources de données disponibles*	20	17	3	4	3	0
		Auditer la qualité des données	21	16	4	2	5	1
		Préparer, nettoyer et programmer le recodage des données	27	13	5	1	4	0
		Augmenter les données	14	22	9	4	2	0
TE	Modélisation	*Réaliser une analyse exploratoire des données*	17	20	5	6	3	0
		Programmer les indicateurs nécessaires au fonctionnement des algorithmes	10	22	7	5	3	1
		Identifier et sélectionner les algorithmes de base à utiliser pour la modélisation	15	22	6	3	4	1
		Prototyper les modèles d'IA	24	13	7	4	2	0
		Tester et décider de rejeter/adopter les modèles en mesurant et évaluant leur pe	19	17	5	5	1	0
		Automatiser les modèles d'IA	22	15	6	4	2	0
TE	Industrialisation	*Optimiser les modèles d'IA*	18	18	6	4	2	0
		Industrialiser, sécuriser, rendre résilients et utilisables les modèles d'IA	13	21	9	6	2	
		Maintenir les modèles et algorithmes en état de fonctionnement	16	22	6	5	2	1
TE	Transfert au métier	*Restituer aux interlocuteurs métier les résultats des modèles*	9	26	4	7	4	0

Annexe 8 – Définition des persona

Les sous-compétences nécessaires à chaque persona sont indiquées par « 1 ».

	PERSONA : COMPÉTENCES IA & DATA	Data Scientist	Data Engineer/ Développeur	Chef de projet	IT/Data architect	Métier	ML Ops/ML Engineer
1	**COMPÉTENCES TRANSVERSES**						
1.1	**ÉVALUER LA FAISABILITÉ D'UN BESOIN MÉTIER**						
1.1.1	*Savoir expliquer les bénéfices d'un projet IA aux interlocuteurs métier*						
1.1.1.a	Savoir identifier les bénéfices et les impacts de l'IA			1		1	
1.1.1.b	Etre capable d'identifier les sujets relevant de l'IA			1		1	
1.1.1.c	Savoir identifier les cas d'usages classiques liés à un métier	1		1		1	
1.1.1.d	Savoir scénariser des implémentations possibles de cas relevant de l'IA			1			
1.1.2	*Savoir identifier l'opportunité d'un projet IA*						
1.1.2.a	Savoir expliquer les problématiques typiques en IA	1		1			
1.1.2.b	Savoir évaluer la complexité d'une problématique IA			1			
1.1.2.c	Etre capable d'évaluer la faisabilité d'un projet			1		1	
1.1.2.d	Savoir identifier les algorithmes utilisables	1		1			
1.1.3	*Connaître les problématiques éthiques de l'IA*						
1.1.3.a	Savoir mettre en place une veille sur l'éthique de l'IA					1	
1.1.3.b	Savoir intégrer une démarche éthique dans les projets IA	1		1			
1.1.3.c	Savoir mettre en oeuvre des procédures de mesure et contrôle des impacts de la solution IA			1		1	
1.1.4	*Connaître les enjeux de conformité et réglementaires de l'IA*						
1.1.4.a	Savoir mettre en place une veille sur la réglementation de l'IA						
1.1.4.b	Savoir évaluer les impacts d'un projet IA au vu de la réglementation						
1.1.4.c	Savoir mettre en place le RGPD						
1.1.5	*Appliquer les bonnes pratiques pour gérer les risques IA*						
1.1.5.a	Savoir mettre en place une méthodologie de mesure et de maîtrise des risques			1		1	
1.1.5.b	Savoir mettre en place une gouvernance de contrôle des risques			1		1	
1.1.5.c	Savoir mettre en place une méthodologie privacy-by-design			1		1	
1.1.5.d	Savoir maîtriser les risques associés aux places de marché pour exposer les données			1		1	
1.1.6	*Savoir mettre en place une approche sobre de l'IA*						
1.1.6.a	Savoir mesurer les dépenses énergétiques de l'IA	1		1			
1.1.6.b	Savoir proposer des algorithmes IA plus frugaux	1					
1.1.6.c	Savoir évaluer et mettre en place des best practices d'utilisation frugale de l'IA	1					
1.2	**PILOTAGE DE PROJET IA**						
1.2.a	Savoir définir une solution projet (angles fonctionnel et métier)			1		1	
1.2.b	Savoir calculer le ROI et les montants à investir			1			
1.2.c	Savoir réaliser un plan projet			1			
1.2.d	Savoir utiliser les méthodes agiles		1	1			
1.2.e	Connaître les méthodes de gestion DataOps et DevOps		1	1			1
1.2.f	Savoir encadrer une équipe projet IA			1			
1.2.g	Savoir garantir le respect du calendrier			1			
1.2.h	Savoir accompagner les équipes métiers			1			
1.3	**SAVOIR METTRE EN PLACE LA CONDUITE DU CHANGEMENT**						
1.3.1	*Savoir identifier les enjeux stratégiques d'évolution*						
1.3.1.a	Savoir identifier le contexte, les enjeux, les objectifs et les contraintes de l'entreprise			1		1	
1.3.1.b	Savoir identifier les différents processus métier potentiellement impactés par l' IA			1		1	
1.3.1.c	Savoir produire une étude de cadrage pour accompagner le changement par des projets IA			1			
1.3.1.d	Savoir mettre en oeuvre une approche éthique pour mesurer l'impact des changements			1			
1.3.2	*Savoir concevoir un plan de changement*						
1.3.2.a	Savoir identifier les parties prenantes (MOA , MOE) et les canaux de communication			1			
1.3.2.b	Savoir identifier les actions à mener pour optimiser les processus internes, suite au projet IA			1			
1.3.2.c	Savoir produire le plan de changement			1			
1.3.3	*Savoir accompagner le changement induit par un projet IA*						
1.3.3.a	Savoir élaborer une stratégie de communication auprès des acteurs du projet IA						
1.3.3.b	Savoir impliquer les directions métiers			1			
1.3.3.c	Savoir traduire les besoins en termes de transfert de compétences						
1.3.3.d	Savoir élaborer des plans de formation pour accompagner le changement						

		Data Scientist	Data Engineer/ Développeur	Chef de projet	IT/Data architect	Métier	ML Ops/ML Engineer
2	**COMPÉTENCES TECHNIQUES**						
2.1	**ARCHITECTURE**						
2.1.1	*Concevoir et mettre en œuvre les infrastructures big data*						
2.1.1.a	Savoir planifier les moyens à mettre en œuvre de façon opérationnelle				1		
2.1.1.b	Savoir contribuer aux choix d'architecture (on prem, cloud, edge, embarqué)		1				
2.1.1.c	Savoir maîtriser le hardware (GPU, CPU, RAM, Compute) et être capable d'effectuer le dimensionnement		1		1		
2.1.1.d	Savoir maîtriser les solutions de stockage, notamment NoSQL		1				
2.1.1.e	Savoir mettre en place la containerisation, l'orchestration des différents outils et le monitoring		1				
2.1.1.f	Savoir maîtriser le fonctionnement d'une infrastructure cloud majeure		1		1		
2.1.1.g	Etre capable d'utiliser des machines sur le cloud (ajouter, supprimer, sauvegarder les machines		1				
2.1.1.g	Savoir maîtriser les sauvegardes et restaurations des données		1				
2.1.1.h	Savoir gérer les services managés cloud		1				
2.1.2	*Concevoir l'architecture data*						
2.1.2.a	Savoir maîtriser les architectures data-centric				1		
2.1.2.b	Savoir définir une vision long terme et des principes directeurs				1		
2.1.2.c	Connaitre les architectures et outils cloud, on prem et edge				1		
2.1.2.d	Comprendre les principes de conception du cloud				1		
2.1.2.e	Savoir définir l'architecture générale				1		
2.1.1.f	Maîtriser les architectures fédérées de type data mesh				1		
2.1.2.g	Maîtriser les solutions de federated learning		1		1		
2.1.2.h	Maîtriser les bonnes pratiques de migration vers un cloud public		1		1		
2.1.2.i	Connaître les limites des outils				1		
2.1.2.j	Savoir où trouver les données dans les applicatifs de l'entreprise		1		1		
2.1.2.k	Savoir s'interfacer avec l'ensemble des outils et applications		1		1		
2.1.3	*Maitriser l'optimisation et le run*						
2.1.3.a	Savoir mettre en oeuvre la mesure de la performance des traitements sur l'infrastructure cible		1				1
2.1.3.b	Savoir paralléliser les traitements		1				1
2.1.3.c	Savoir identifier les goulots d'étranglement et les traitements critiques		1		1		
2.1.3.d	Savoir optimiser les traitements sur le cloud		1				
2.1.4.e	Savoir optimiser les usages de l'architecture cloud				1		
2.1.3.f	Savoir assurer le maintien en conditions opérationnelles et traiter les incidents		1				
2.1.4	*Maitriser les coûts et la sécurité des architectures à base de cloud*						
2.1.4.a	Savoir mesurer le coût du cloud				1		
2.1.4.b	Comprendre le modèle de responsabilité partagée avec le fournisseur de cloud				1		
2.1.4.c	Maitriser les bonnes pratiques pour assurer la sécurité des ressources dans le cloud				1		
2.1.4.d	Savoir mettre en place des solutions de vérification de maîtrise du risque, de cryptage vers le cloud				1		
2.1.5	*Programmer les traitements de données*						
2.1.5.a	Maitriser le langage Python	1	1				1
2.1.5.b	Maitriser le langage C++		1				
2.1.5.c	Maitriser le langage Java		1				
2.1.5.d	Savoir utiliser du hardware spécialisé pour l'apprentissage (optimisation de TensorFlow/pytorch sur	1	1				1
2.2	**DONNÉES**						
2.2.1	*Auditer les sources de données disponibles*						
2.2.1.a	Etre capable de reconnaître le caractère éthique ou non éthique des données disponibles		1				
2.2.1.b	Savoir identifier les données atypiques et/ou aberrantes	1	1				
2.2.1.c	Savoir identifier et gérer les données manquantes	1	1				
2.2.1.d	Savoir appliquer les règles du RGPD	1	1				
2.2.2	*Collecter les données*						
2.2.2.a	Savoir où collecter les données		1				
2.2.2.b	Savoir collecter les données (ETL)		1		1		
2.2.2.c	Savoir constituer un échantillon de données pour le data scientist		1				
2.2.2.d	Savoir établir un dossier de sécurité (limites, habilitations, éthique, règlementaire)		1		1		
2.2.2.e	Savoir gérer l'anonymisation et le cryptage		1				
2.2.2.f	Savoir gérer l'ordonnancement de la collecte des données		1		1		
2.2.3	*Auditer la qualité des données*						
2.2.3.a	Maîtriser la statistique descriptive	1					
2.2.3.b	Savoir utiliser la librairie Numpy en langage Python	1					
2.2.3.c	Savoir utiliser la librairie Pandas en langage Python	1					
2.2.3.d	Savoir réaliser un audit de qualité et de sécurité des données	1	1				
2.2.4	*Préparer, nettoyer et programmer le recodage des données*						
2.2.4.a	Savoir programmer le recodage des données	1	1				
2.2.4.b	Savoir traiter les valeurs manquantes	1	1				
2.2.4.c	Savoir traiter les valeurs aberrantes	1	1				
2.2.4.d	Maîtriser le fonctionnement des bases de données SQL	1	1				
2.2.4.e	Savoir créer des jeux de données de test	1	1				
2.2.4.f	Savoir annoter/labelliser les données					1	
2.2.5	*Augmenter les données*						
2.2.5.a	Savoir identifier les sources de données open-source utiles	1	1	1			
2.2.5.b	Savoir générer de nouvelles variables (feature engineering)	1					
2.2.5.c	Savoir utiliser un simulateur / jumeau numérique		1				
2.2.5.d	Savoir générer des jeux de données par simulateur		1				
2.2.5.e	Savoir utiliser un modèle génératif (GAN)	1					

		Data Scientist	Data Engineer/ Développeur	Chef de projet	IT/Data architect	Métier	ML Ops/ML Engineer
2.3	**MODÉLISATION**						
2.3.1	*Réaliser une analyse exploratoire des données*						
2.3.1.a	Savoir utiliser les principales méthodes de réduction de dimension non supervisées	1					
2.3.1.b	Savoir utiliser les principales méthodes de clustering et de typologies non supervisées	1					
2.3.1.c	Savoir utiliser l'analyse statistique des corrélations	1					
2.3.1.d	Savoir mettre en œuvre une visualisation des données	1					
2.3.1.e	Savoir créer un jeu de données pertinent pour la cross-validation	1					
2.3.1.f	Savoir analyser les biais existants dans les données d'apprentissage	1					
2.3.1.g	Savoir éviter un biais dans les données d'apprentissage	1					
2.3.1.h	Maîtriser les différentes méthodes de test d'hypothèses	1					
2.3.2	***Concevoir des modèles analytiques d'IA et de machine-learning***						
2.3.2.1	*Programmer les indicateurs nécessaires au fonctionnement des algorithmes*						
2.3.2.1.a	Savoir utiliser la méthode MLOps pour mener à bien un projet						1
2.3.2.1.b	Savoir mettre en oeuvre une stratégie d'inférence						1
2.3.2.1.c	Savoir mettre en oeuvre du machine learning non supervisé						1
2.3.2.1.d	Savoir mettre en oeuvre du machine learning supervisé						1
2.3.2.1.e	Savoir programmer les indicateurs nécessaires						1
2.3.2.2	*Identifier et sélectionner des algorithmes de base à utiliser pour la modélisation*						
2.3.2.2.a	Savoir identifier les différentes formes d'apprentissage machine	1		1			1
2.3.2.2.b	Savoir entraîner les algorithmes d'apprentissage	1					
2.3.2.2.c	Savoir identifier les principaux algorithmes d'IA du domaine public	1					
2.3.2.2.d	Savoir utiliser les algorithmes du domaine public pour créer un modèle	1					
2.3.2.2.e	Savoir utiliser les modèles pré-entraînés du domaine public pour créer un modèle (IA générative, vision,	1					
2.3.2.2.f	Savoir développer un modèle d'IA symbolique	1	1				
2.3.2.2.g	Savoir développer un modèle multiagent		1				
2.3.2.3	*Connaissance de techniques d'IA spécialisées*						
2.3.2.2.a	Maîtriser les techniques de vision par ordinateur	1					
2.3.2.2.b	Maîtriser les techniques de NLP	1					
2.3.2.2.c	Maîtriser les techniques de reconnaissance vocale	1					
2.3.2.4	*Maitriser les techniques d'IA avancées*						
2.3.2.4.a	Savoir mettre en œuvre de l'apprentissage sur graphes/graphes de connaissances (GNNs)	1					1
2.3.2.4.b	Maîtriser les méthodes de reinforcement learning	1					1
2.3.2.4.c	Maîtriser les méthodes de federated learning	1	1				1
2.3.2.4.d	Savoir utiliser les knowledge informed neural networks (physical / geometric)	1					
2.3.2.4.e	Savoir utiliser l'IA neuro-symbolique						1
2.3.2.4.f	Savoir concevoir et exploiter des ontologies	1				1	
2.3.2.4.f2	Savoir concevoir et automatiser des modèles de raisonnement	1					
2.3.2.4.f3	Savoir concevoir et mettre en œuvre des architectures d'IA distribuée	1					
2.3.2.4.g	Connaître les méthodes d'optimisation	1					
2.3.2.4.h	Savoir mettre en œuvre de la recherche opérationnelle	1					
2.3.2.4.i	Maîtriser les technologies de edge AI		1				1
2.3.2.5	*Maitriser les techniques d'IA génératives*						
2.3.2.5.a	Savoir générer des prompts et les mettre à disposition des métiers		1				
2.3.2.5.b	Savoir utiliser les prompts et vérifier les informations retournées					1	
2.3.2.5.c	Maîtriser les outils et techniques d'IA générative						1
2.3.2.5.d	Savoir identifier les limites des outils d'IA générative		1				1
2.3.2.5.e	Savoir adapter les modèles LLMs aux données de l'entreprise						1
2.3.2.6	*Prototyper les modèles d'IA*						
2.3.2.6.a	Savoir utiliser les principales librairies d'IA en Python, notamment TensorFlow et PyTorch	1	1				1
2.3.2.6.b	Savoir assembler la chaîne de traitement des différents modèles	1					1
2.3.2.6.c	Savoir optimiser le compromis explicabilité & traçabilité / précision des algorithmes	1					
2.3.2.6.d	Savoir documenter les modèles avec Jupyter Notebook	1					
2.3.2.6.e	Savoir implémenter des pistes d'auditabilité et de traçabilité	1					
2.3.2.6.f	Savoir s'assurer de la réversibilité du code		1			1	1
2.3.2.6.g	Maîtriser les outils low-code/no-code		1				1
2.3.2.6.h	Maîtriser les services cloud d'IA		1				1
2.3.2.6.i	Connaître les bases du Quantum Machine learning		1				1
2.3.3	***Tester et implémenter des modèles analytiques d'IA***						
2.3.3.1	*Tester et décider de rejeter/adopter les modèles en mesurant & évaluant leur pertinence sur la base de leur précision, leur robustesse, leur absence de biais*						
2.3.3.1.a	Savoir mesurer la performance des modèles	1					
2.3.3.1.b	Savoir créer les indicateurs pour valider ou rejeter les modèles testés	1					
2.3.3.1.c	Savoir calculer et mesurer la variance des résidus d'un modèle quantitatif	1					
2.3.3.1.d	Savoir produire la matrice de confusion d'un modèle de classification	1					
2.3.3.1.e	Maîtriser les méthodes d'optimisation de complexité et de cross-validation des modèles	1					
2.3.3.1.f	Savoir exploiter les résultats d'une stratégie inférentielle de cross-validation	1					
2.3.3.1.g	Savoir s'assurer du respect des hypothèses de fonctionnement des modèles	1					
2.3.3.1.h	Savoir s'assurer de l'absence de biais dans les résultats du modèle	1					
2.3.3.1.i	Savoir mesurer la robustesse statistique des modèles	1					
2.3.3.2	*Automatiser les modèles d'IA*						
2.3.3.2.a	Savoir choisir les hyper-paramètres pour faciliter l'automatisation	1					
2.3.3.2.b	Savoir automatiser l'optimisation des hyper paramètres	1					

		Data Scientist	Data Engineer/ Développeur	Chef de projet	IT/Data architect	Métier	ML Ops/ML Engineer
2.4	**INDUSTRIALISATION**						
2.4.1	*Gérer le cycle de vie des modèles*						
2.4.1.a	Savoir documenter les modèles	1					1
2.4.1.b	Savoir construire un pipeline de modèles, basé sur une architecture complexe	1					1
2.4.1.c	Savoir gérer le versionning des données		1				1
2.4.1.d	Savoir gérer le versionning des modèles						1
2.4.1.e	Savoir mettre en place un feature store						1
2.4.1.f	Savoir produire des templates réutilisables		1				1
2.4.1.g	Savoir gérer le suivi des expérimentations						1
2.4.1.h	Savoir mettre en œuvre le réentrainement des modèles						1
2.4.1.i	Savoir gérer l'explicabilité et la prise en compte des biais	1					1
2.4.2	*Industrialiser, sécuriser, rendre résilients et utilisables les modèles de datascience*						
2.4.2.a	Savoir spécifier les conditions d'utilisation des modèles d'IA			1			1
2.4.2.b	Maîtriser la méthode devops en lien avec la méthode MLOps (dataops / modelOps)		1				1
2.4.2.c	Savoir orchestrer un ensemble d'algorithmes				1		1
2.4.2.d	Savoir intégrer les résultats dans les applications digitales de l'entreprise		1		1		
2.4.2.e	Savoir documenter une application		1				
2.4.2.f	Savoir intégrer la sécurité dans la chaîne data et dans les environnements d'entraînement		1		1		
2.4.3	*Gérer la supervision et le monitoring*						
2.4.3.a	Maîtriser les méthodes statistiques de monitoring des modèles d'IA						1
2.4.3.b	Savoir identifier et analyser les dérives des modèles						1
2.4.3.c	Savoir configurer les seuils d'alertes en cas de dysfonctionnement des modèles		1			1	1
2.4.3.d	Savoir mesurer les performances des modèles						1
2.4.3.e	Savoir mettre en place le monitoring fonctionnel du feedback utilisateur (boucle de rétroaction)					1	1
2.4.3.f	Savoir mettre en place une maintenance corrective et/ou évolutive (continuous training)		1			1	1
2.4.3.g	Savoir monitorer les processus de réapprentissage automatique		1				1
2.4.3.h	Savoir réaliser la supervision technique/MCO		1		1		
2.4.4	*Tester et déployer*						
2.4.4.a	Savoir réaliser les tests d'intégration et tests système		1				
2.4.4.b	Savoir mettre en place les tests d'acceptation des utilisateurs					1	
2.4.4.c	Savoir mettre en place les tests de performance et de charge		1		1		
2.4.4.d	Savoir maîtriser la chaine CI/CD (continuous integration / continuous development)		1		1		
2.4.4.e	Savoir mettre en œuvre le pipeline de collecte des données		1		1		
2.4.4.f	Savoir définir et mettre en œuvre la stratégie de déploiement des nouveaux modèles (amélioration contin		1		1		
2.5	**TRANSFERT AU MÉTIER**						
2.5.1	*Restituer aux interlocuteurs métiers les résultats des modèles*						
2.5.1.a	Savoir effectuer du storytelling autour de la donnée			1			1
2.5.1.b	Être capable de simplifier la communication à des fins de vulgarisation			1			
2.5.1.c	Savoir présenter des visualisations des données compréhensibles par les métiers			1			1
2.5.1.d	Savoir illustrer les résultats du modèle par des graphiques compréhensibles par les métiers						1
2.5.1.e	Savoir sélectionner les observations illustrant le comportement typique des modèles						1
2.5.1.f	Savoir interpréter les résultats des différents modèles et en expliquer les résultats						1
2.5.1.g	Savoir utiliser des modèles explicatifs pour pouvoir rendre traçables / explicables les modèles						1
2.5.1.h	Savoir expliquer les risques des modèles			1			1
2.5.1.i	Savoir présenter et afficher les biais des modèles						1
2.5.1.j	Savoir mesurer les écarts/non conformités et les communiquer						1
2.5.1.k	Savoir mesurer les impacts des résultats			1			1
TOTAL DES COMPÉTENCES PAR PERSONAE		**67**	**68**	**39**	**31**	**21**	**54**

Annexe 9 – Analyse factorielle

Cette analyse (voir annexe 7) a pour objectif de comprendre l'interdépendance des diverses compétences en IA au sein des différentes fonctions liées à l'IA dans les entreprises françaises. L'objectif est de de confirmer la possibilité d'établir un référentiel cohérent de compétences.

Pour atteindre cet objectif, une Analyse en Composantes Principales (ACP) avec application d'une rotation de type varimax, une forme spécifique d'analyse factorielle, a été utilisée. L'ACP est une méthode de réduction de dimension qui permet d'extraire les tendances prépondérantes d'un ensemble de données multidimensionnel.

Dans ce contexte, l'ACP a été employée pour analyser la structure de la matrice de données représentant les compétences en IA pour six fonctions clés :

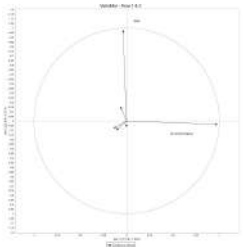
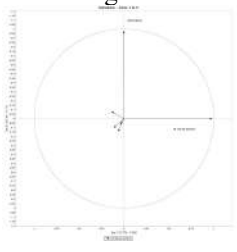
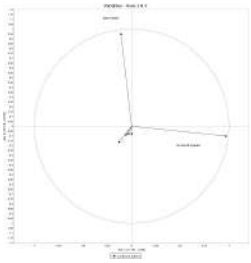
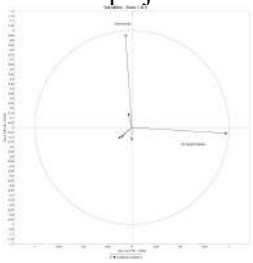
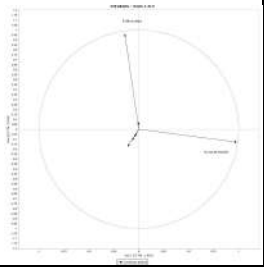
- Data Scientist;
- Data Engineer;
- Chef de projet;
- IT-Data Architect;
- Métier;
- ML Ops-ML Engineer.

Les résultats de l'ACP ont révélé six composantes principales, chaque composante représentant l'une des six fonctions / Persona. Ces composantes ont été interprétées comme des axes.

- Le premier axe a été identifié comme correspondant à la compétence **ML Ops**. Cette compétence englobe la capacité d'opération et de maintenance des systèmes d'apprentissage automatique, ce qui est essentiel pour le déploiement réussi de solutions d'IA.
- Le deuxième axe a été associé à la compétence **Métier**. Cette compétence reflète la compréhension des besoins métier et la capacité de traduire ces besoins en solutions d'IA.
- Le troisième axe a été attribué à la compétence **Data Engineer**. Cette compétence couvre la capacité à concevoir, construire et gérer les systèmes de données nécessaires pour l'IA.
- Le quatrième axe a été associé à la compétence **Data Scientist**. Cette compétence représente la capacité à développer des modèles d'apprentissage automatique et à tirer des informations précieuses des données.
- Le cinquième axe a été lié à la compétence **Chef de projet**. Cette compétence concerne la capacité à gérer les projets d'IA, y compris la planification, la coordination des équipes et le suivi des progrès.
- Enfin, le sixième axe a été attribué à la compétence **IT-Data Architect**, qui comprend la capacité à concevoir l'architecture des systèmes d'information et de données.

L'ordre des axes ne correspond pas à l'importance des fonctions (surtout après varimax) mais peut éventuellement donner une indication sur la diversité des compétences au sein d'une fonction.

Les projections sur les axes principaux montrent l'orthogonalité (indépendance) des fonctions / Persona, ce qui confirme les définitions des Persona :

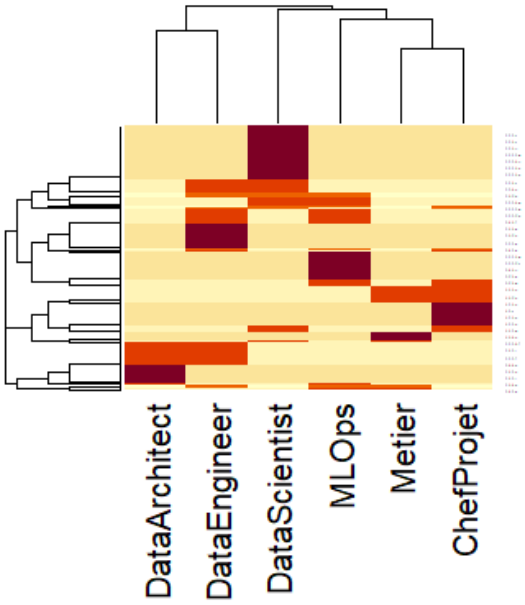
Axes 1 & 2: MLOps /ML engineer & Métier	Axes 1 & 3: MLOps /ML engineer & Data Engineer	Axes 1 & 4: MLOps /ML engineer & Data Scientist	Axes 1 & 5: MLOps /ML engineer & Chef de projet	Axes 1 & 6: MLOps /ML engineer & IT / Data architect
				

Ceci implique que :

- Les compétences sont bien distinctes et indépendantes les unes des autres et forment un référentiel ;
- Une force dans un domaine de compétences n'implique pas nécessairement une force ou une faiblesse dans un autre domaine de compétences ;
- La formation dans une fonction spécifique n'affectera pas nécessairement les compétences dans d'autres domaines.

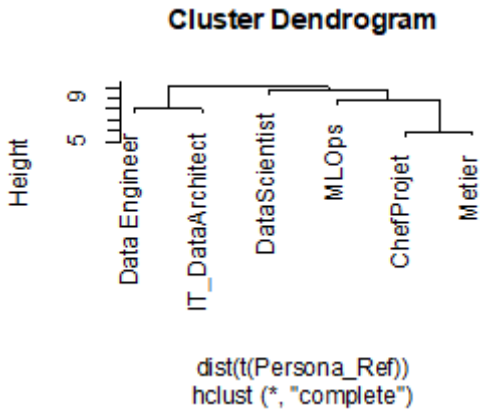
Annexe 10 – Clustering hiérarchique

A partir du référentiel et des compétences rattachées aux Persona, nous avons mis en œuvre une analyse (basée sur le *clustering* hiérarchique) à la fois sur les colonnes (Personas) et sur les lignes (compétences) pour produire la *heatmap* de la figure suivante[12] :



Dans cette figure, plus les blocs de compétences sont foncés, plus ils sont pertinents pour la Persona référencée. Ainsi, la lecture de la figure montre que les compétences sont bien identifiées pour chaque Persona, en raison du peu de blocs de compétences communes et que ces blocs communs sont logiques. Par exemple, des compétences sont communes entre *data engineer* et *data scientist*, entre *data engineer* et *data architect* ou encore entre *data scientist* et *machine learning engineer*. Par ailleurs, une interprétation plus poussée (non visible sur cette figure) fait apparaitre que les blocs de compétences sont bien homogènes par rapport aux macro-compétences du référentiel de compétences v1. Notre référentiel de compétences est donc construit correctement et logiquement par rapport aux profils de personnes (Persona).

Le *clustering* hiérarchique permet de regrouper progressivement les individus en classes, produisant ainsi une hiérarchie sous la forme d'un dendrogramme. Cette méthode permet donc de former ces classes automatiquement et de manière non supervisée. La figure suivante montre le *clustering* hiérarchique réalisé sur les Persona dans le référentiel de compétences :



Nous retrouvons les Persona rattachées aux mêmes classes selon la logique que dicte notre connaissance des métiers : *Data Engineer* et *Data Architect* dans la même classe, Chef de projet et métier dans la même classe, *Data Scientist* à part et proche du *MLOps*. Ceci constitue une validation supplémentaire de notre définition des Persona avec leurs compétences.

[12] *Travail réalisé en collaboration avec Etienne Cuvelier, chercheur au Laboratoire QUARESMI de l'ICHEC Brussels Management School*

Contacts

Françoise Soulié
Conseiller Scientifique
Hub France IA
francoise.soulie@outlook.com

Caroline Chopinaud
Directrice générale
Hub France IA
caroline.chopinaud@hub-franceia.fr

Camille Souillart
Chef de projet
Hub France IA
camille.souillart@hub-franceia.fr