



# Synthèse

## Projet « REFCO-IA »

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

<b>1</b> – Introduction	3
<b>2</b> – Étapes du projet	3
<b>3</b> – Livrables attendus	4
<b>4</b> – Impacts sur les formations	8
<b>5</b> – Phase 2 du projet	9
<b>5</b> – Conclusions et préconisations	10

# 1. Introduction

Le projet REFCO-IA fait partie de l'appel à manifestation d'intérêt (AMI) « Compétences et métiers d'avenir » qui 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**.

REFCO-IA porte sur le volet « réalisation d'un diagnostic des besoins en compétences et en formations » de l'AMI. Il est à l'initiative d'un consortium de onze acteurs<sup>1</sup>, 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*<sup>2</sup> et *reskilling*<sup>3</sup> de leurs effectifs. Les nouvelles technologies 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 un projet de diagnostic des formations IA ayant fait également l'objet d'un dépôt par un autre consortium. 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.

Le projet a débuté le 17 novembre 2022 et a été découpée en 2 grandes phases. La première phase de 6 mois, avait pour ambition la production d'un référentiel de compétences IA et l'identification des compétences les plus recherchées à court et moyen termes. Cette première phase s'est conduite en deux étapes : une première étape 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.

A la suite de la production du référentiel et des premiers résultats obtenus par compétence, une deuxième phase du projet a été lancée en juin 2023 pour élaborer une nouvelle enquête qui s'est clôturée en décembre 2023 visant à identifier les besoins en formation (volumétrie) par profil métier à horizon 2024 et 2030 et affiner les premières estimations issues de la phase initiale du projet.

## 2. Étapes du projet

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 avons 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
<b>Total</b>	<b>69</b>	<b>58</b>	<b>62</b>	<b>75</b>	<b>264</b>

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.

Lors de la réponse à l'AMI, 4 livrables ont été identifiés : L.1.1 : Référentiel de compétences v0 ; L.1.2 : Questionnaire ; L.2 :

<sup>1</sup> Membres du Consortium : Business & Decision, Datavaloris, ESSEC, Hub France IA, IA Institut, JustAI, La Poste, LePont, Numeum, Société Générale, Société Informatique de France (SIF)

<sup>2</sup> Upskilling : montée en compétence.

<sup>3</sup> Reskilling : acquisition de compétences nouvelles dans le cadre d'un changement de métier ou d'une montée en compétence transversale.

Questionnaires et comptes-rendus (l'enquête) ; L.3 : Référentiel de compétences v1 ; L.4 : Préconisations.

Nous les détaillons dans la section suivante.

## 3. Livrables attendus

### L1.1 - Le Référentiel de compétences v0

#### ● Méthodologie

Le livrable L1.1 a été produit à partir des référentiels existant chez les membres du consortium et sur France Compétences<sup>4</sup>. Ces documents étant des référentiels métiers, nous avons dû fournir un travail d'analyse et de restructuration. Ce référentiel exclut les compétences générales (mathématiques, statistiques, ou *soft skills*) : un métier devra alors être caractérisé par ses compétences générales en plus des compétences IA requises.

Nous n'avons pas adressé la question des **pré-requis** et de la formation initiale des personnes. 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 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**. Les compétences proposées devraient être envisagées dans le long terme et fournir un socle solide pour permettre l'évolution des métiers.

#### ● Structure

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

Nous avons distingué deux types de compétences IA :

- **Compétences transverses** : elles regroupent les compétences nécessaires aux tâches préliminaires au projet, 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.

### L1.2 - Le questionnaire

#### ● Méthodologie

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. **Tous les résultats présentés font donc référence au Référentiel v0**. Le questionnaire est disponible en ligne<sup>5</sup>. 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.

#### ● Structure

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** : informations générales sur l'entreprise répondante à partir de questions fermées ;
2. **Compétences transverses** : issues du référentiel v0, complétées de questions, déclenchées par la réponse « oui » à la question « Avez-vous besoin de monter en compétence sur ... » ;
3. **Compétences techniques** : comme pour les compétences transverses ;
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.

<sup>4</sup> RNCP et RS pour Intelligence Artificielle et Data Science

<sup>5</sup> Disponible à l'URL : <https://hfia.limesurvey.net/784422?lang=fr>

## L2 - Questionnaires et comptes-rendus : l'enquête

### • Méthodologie

Le questionnaire a été envoyé par les partenaires à leurs réseaux de contacts, visant un échantillon représentatif des entreprises sollicitées (grands groupes, PME/ETI, startups). 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.

### • Résultats

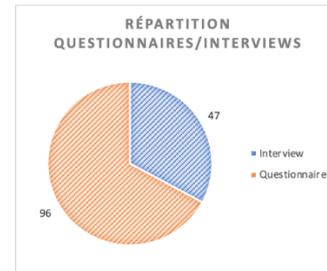
Le tableau ci-dessous présente le nombre d'entreprises ayant répondu à la phase d'enquête par typologie.

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

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**.

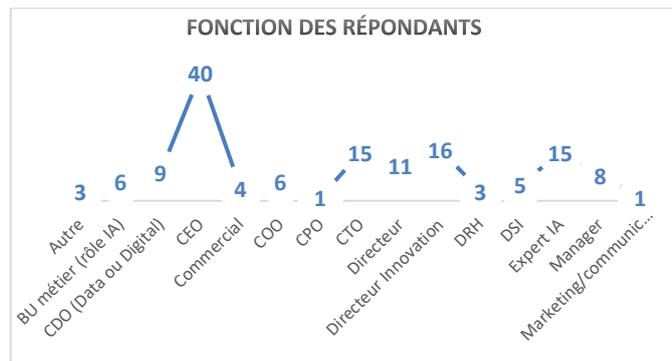
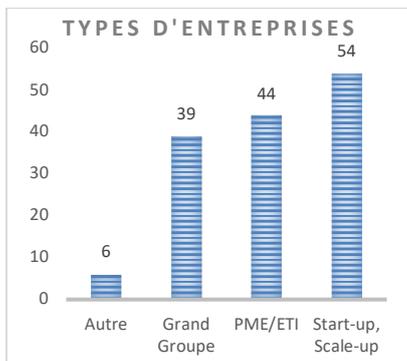
### Répartition des questionnaires et entretiens au sein de l'enquête

La figure montre une répartition équilibrée entre questionnaires et entretiens, avec 67% pour les premiers et 33% pour les seconds. 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.



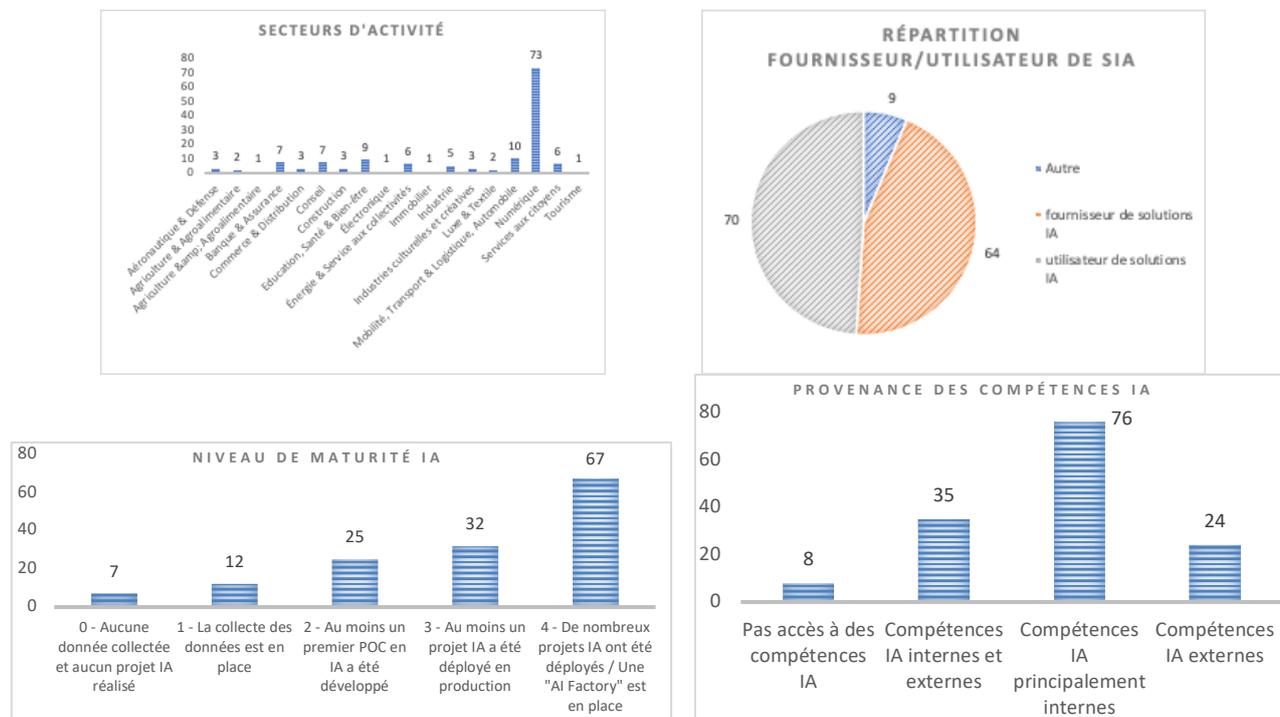
### Typologie des répondants : types d'entreprises et fonctions

La figure ci-dessous (gauche) montre un équilibre dans la typologie des répondants, 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 CEO, 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

Comme on le voit dans les figures ci-après, 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 :



Nous avons rencontré des difficultés de recrutement qui tiennent 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. 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.**

### Conclusions générales de l'enquête

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. Parmi celles-ci, les 5 macro-compétences les plus mentionnées sont :

- 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. 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.

Nous voyons que la première macro-compétence technique la plus recherchée – qui arrive en 9<sup>ème</sup> position du classement global – est « Restituer aux interlocuteurs métier les résultats des modèles ». 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 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, avec des profils plutôt experts.

Ces éléments démontrent la maturité des entreprises interrogées, 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. **Les entreprises non matures, au fur et à mesure**

de leur montée en maturité, devront monter en compétences sur les compétences techniques IA.

## L3 - Le Référentiel de compétences v1

### ● Méthodologie

Notre référentiel v1 a été produit en améliorant la structure de la version v0. Des travaux d'analyse ont permis d'établir la validité de ce référentiel. Le référentiel v1 a donc 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 traite uniquement de ce qui est indispensable pour les projets d'IA et omet volontairement des sujets spécifiques à la *data* comme la gouvernance des données par exemple.

### ● Structure

Tout comme le référentiel v0, le référentiel v1 est constitué de deux grandes parties :

- **Compétences transverses** : cette partie comprend 10 macro-compétences et 40 sous-compétences. Dans cette partie, quelques sous-compétences ont été ajoutées et/ou réorganisées, mais elle diffère peu entre les référentiels v0 et v1.
- **Compétences techniques** : cette partie 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 et d'IA générative ;
  - Nous avons largement revu la partie Industrialisation pour mieux prendre en compte la gestion du cycle de vie, la supervision et le *monitoring*.

### ● Validation du référentiel : les persona

#### Méthodologie

Le questionnaire porte sur les besoins de compétences et non sur les besoins en nombre de personnes. Il a donc été décidé de construire des « Persona<sup>6</sup> » et de leur attribuer des compétences sur les versions v0 et v1 du référentiel. Nos Persona correspondent à des métiers existants. Ils ne sont pas décomposés en sous-métiers.

Ces Persona nous ont servi 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. Le nombre de compétences attachées à chaque Persona est : 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* jouent un rôle fondamental.

#### Analyse de données des Persona

Nous avons réalisé une **analyse factorielle** qui 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

<sup>6</sup> Profils de personnes

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 représentant chacune 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.

Nous avons ensuite réalisé un **clustering hiérarchique** en travaillant sur les Persona et les compétences au travers d'une *heatmap*. Les résultats montrent que les compétences sont bien identifiées pour chaque Persona avec des blocs de compétences bien différenciés. Seules quelques compétences communes apparaissent entre Persona voisines.

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 compétences** 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.

## 4. Impacts sur les formations

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 nous permet 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 ou le marketing par exemple, pourraient inclure dans leur cursus des cours de base sur l'IA dans un but d'acculturation.

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<sup>7</sup>, 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.

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.

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<sup>8</sup> et dans notre échantillon était la suivante :

	Population	Echantillon
Nombre d'entreprises	4 114 289	144
Numérique (Information et Communication)	153 727	73
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
<b>TOTAL DES BESOINS EN FORMATION</b>	55 785	64 144	84 520	56 937	99 272	57 200
<b>niveau</b>	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

<sup>7</sup> Massive Open Online Course

<sup>8</sup> <https://www.insee.fr/fr/statistiques/2498421>

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.

## 5. Phase 2 du projet

### • Étapes du projet et indicateurs de performances

A la suite de la présentation des résultats du projet REFCO-IA, les commanditaires ont souhaité un approfondissement du diagnostic en particulier pour répondre aux questions suivantes :

1. Quels sont les profils de métiers de l'IA en cours et à venir ?
2. Quelles est la volumétrie des besoins des entreprises sur ces différents profils à horizon 2024 & 2030 ?
3. Quels sont les niveaux d'études attendues de ces profils ?

Il est attendu également des préconisations sur :

1. La méthodologie à adopter pour tenir à jour la base de données des profils métiers de l'IA
2. Les tendances à venir sur les métiers de l'IA
3. Les besoins en formation (initiale ou continue)

Pour répondre aux attentes de cette deuxième phase du projet, le Hub France IA a piloté seul la mise en place du contenu de l'enquête et a pu s'appuyer sur les membres du consortium ainsi que sur les commanditaires pour dérouler l'enquête entre le mois de septembre et le mois de décembre.

Ainsi le projet a été découpé en trois grandes étapes :

1. Élaboration d'une liste de profils métiers à partir des compétences identifiées en phase 1 et les données de l'OPIIEC ;
2. Définition précise les profils métiers et le niveau d'étude estimé ;
3. Construction d'un questionnaire d'enquête permettre d'établir les volumétries attendues par profils et par niveau de formation ;
4. Établir les préconisations attendues sur les besoins en formation par profil, valider les estimations de niveau d'étude établies en 2, et d'identifier de façon précise si les tendances sont plutôt de l'ordre de la formation continue des collaborateurs déjà en place ou le recrutement de profil spécialisé, ceci ayant un impact sur le type des formations à mettre en place.

### • Livrable

Pour cette deuxième phase du projet, le livrable principal consiste en l'analyse par profil métier des besoins en volume attendu par les entreprises à horizon 2024 et 2030, et précisant les niveaux de formation estimées pour ces profils.

Un livrable connexe correspond donc à une de liste de profils métier avec les compétences utiles à ces profils. Ces compétences font références aux compétences élaborées dans le cadre du référentiel établi lors de la première phase de projet.

Cette seconde enquête a clairement montré que les besoins sont principalement à très court terme (2024) pour plus de 60% sur tous les profils, excepté celui de chercheur en IA pour lequel les besoins sont équilibrés entre 2024 et 2030. Le manque de projection à horizon 2030 s'explique par la très forte évolutivité de l'Intelligence Artificielle actuellement, avec une intégration progressive mais réelle à l'heure actuelle de l'IA générative, et un manque de recul sur son impact à terme sur les emplois. Par ailleurs, nous estimons que la majorité des répondant évaluent, compte tenu de leur typologie, comme une priorité de se renforcer

sur les profils clés plutôt à court terme, tout en accompagnant leurs collaborateurs dans une montée en compétence sur l'IA.

La majorité des entreprises interrogées ont des besoins plutôt sur des volumétries faibles ce qui s'explique par le profil des entreprises répondantes qui sont majoritairement des structures de petite taille ou de taille moyenne.

Enfin, le niveau d'études plébiscité est bac+5, suivi par la formation continue et le niveau bac+8. Les besoins de recrutement à Bac+2/3 sont nettement moindres et quasi-inexistants au niveau Bac.

A noter que la formation continue apparaît plus clairement dans cette deuxième phase de l'enquête comme un souhait des entreprises pour la montée en compétence et ce sur tous les profils. Il est cependant surprenant de voir que les entreprises estiment également qu'un niveau Bac+8 est une option de niveau de formation possible pour l'ensemble des profils.

	2024					2030					Total / Profil
	Bac	Bac +2/3	Bac +5	Bac +8	Formation continue	Bac	Bac +2/3	Bac +5	Bac +8	Formation continue	
<i>Data Scientist</i>	0	2 649	14 129	8 536	4 121	0	752	4 011	2 423	1 170	<b>37 792</b>
<i>Data Engineer</i>	0	12 098	51 416	16 131	21 171	0	1 149	4 882	1 532	2 010	<b>110 388</b>
<i>Développeur IA</i>	0	4 165	12 234	4 946	4 946	0	1 428	4 195	1 696	1 696	<b>35 306</b>
<i>Ingénieur IA spécialisé</i>	169	1 011	8 091	4 888	2 697	73	438	3 502	2 116	1 167	<b>24 152</b>
<i>Chef de projet IA</i>	0	1 165	6 213	2 718	2 718	0	291	1 553	679	679	<b>16 017</b>
<i>IA / Data Architect</i>	119	715	6 076	2 502	2 502	20	120	1 016	418	418	<b>13 907</b>
<i>Métier Augmenté par l'IA</i>	1 310	2 432	7 109	3 180	4 677	211	392	1 145	512	754	<b>21 722</b>
<i>Machine Learning Engineer</i>	161	1 124	7 871	4 498	2 249	38	268	1 875	1 071	536	<b>19 690</b>
<i>Product Owner IA</i>	146	1 611	7 175	2 489	3 075	36	400	1 781	618	763	<b>18 095</b>
<i>Éthicien de l'IA</i>	132	593	2 570	1 714	1 648	48	218	945	630	606	<b>9 105</b>
<i>Data Protection Officer en IA</i>	56	195	1 223	667	611	32	112	706	385	353	<b>4 340</b>
<i>Chercheur en IA</i>	0	0	1 470	3 320	653	0	0	2 358	5 328	1 048	<b>14 178</b>
<i>FinOps / GreenOps Manager</i>	166	717	2 263	993	1 435	7	31	99	43	63	<b>5 818</b>
<b>Total / Niveau</b>	<b>2 257</b>	<b>28 477</b>	<b>127 845</b>	<b>56 586</b>	<b>52 506</b>	<b>466</b>	<b>5 600</b>	<b>28 075</b>	<b>17 456</b>	<b>11 266</b>	<b>330 534</b>

A noter que la formation continue apparaît plus clairement dans cette deuxième phase de l'enquête comme un souhait des entreprises pour la montée en compétence et ce sur tous les profils. Il est cependant surprenant de voir que les entreprises estiment également qu'un niveau Bac+8 est une option de niveau de formation possible pour l'ensemble des profils.

## 6. 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. Nous avons également démontré l'intérêt de **prévoir des cursus de formation bien distincts en fonction des profils à former**. 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<sup>9</sup> (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

<sup>9</sup> <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review>

donc nécessaire de **poursuivre ces travaux pour en tenir compte et pouvoir adapter les offres de formation aux besoins réels du terrain** tout en **mettant régulièrement à jour le référentiel de compétences produit en tant que base commune et ouverte à tous.**

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



### Contacts

Françoise Soulié  
Conseiller Scientifique  
Hub France IA  
[francoise.soulie@outlook.com](mailto:francoise.soulie@outlook.com)

Caroline Chopinaud  
Directrice générale  
Hub France IA  
[caroline.chopinaud@hub-franceia.fr](mailto:caroline.chopinaud@hub-franceia.fr)