**Concevoir et implémenter une solution d'IA - Parcours 2 : Renforcé « Professionnels IT »****Informations générales****Objectif général :**

- Préparer des données nécessaires à un projet d'intelligence artificielle
- Développer et mettre en œuvre une solution d'intelligence artificielle
- Maintenir et améliorer une solution d'intelligence artificielle

Publics : Professionnels de l'IT dont :

- le cœur de métier n'est pas initialement l'exploitation des données,
- l'expertise et les finalités se concentrent sur la gestion et l'optimisation des systèmes et de l'infrastructure informatiques et non sur l'exploitation des données pour prendre des décisions (développeur informatique, spécialiste BDD, consultant AT, spécialiste infrastructure, spécialiste systèmes, réseaux et sécurité, intégrateur logiciel...)

Pré-requis :

- Exercer un emploi comportant des activités orientées vers la création d'infrastructures, de logiciels, vers le traitement et l'analyse de données ou encore vers la maintenance des systèmes ;
- Des connaissances en mathématiques et statistiques, en particulier dans l'exploitation de données (statistiques descriptives, probabilités, statistiques inférentielles, analyse exploratoire des données, modélisation statistique) ;
- Justifier d'une expérience en programmation (1er niveau de maîtrise des langages de programmation, connaissances algorithmes et structures de données, architecture logicielle, systèmes d'exploitation).

Durée : 18 jours (128 heures)

Modalités : Mixte

Régions : Guadeloupe – Martinique – Guyane

Prix Inter Applicable : 15 000,00 €

Nombre de stagiaires : Entre 4 et 12

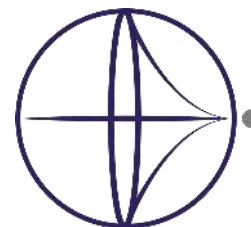
Prix Intra Applicable : 67 500,00 €

Nombre de stagiaires Intra : Entre 4 et 12

Coût de la certification applicable : 400,00 €

Modalités d'évaluation : 1. Cas d'usage unique (1h) (épreuve orale) 2. Questionnaire à visée professionnelle (45 min) (en ligne)

Modalités de certification : Réussite des 2 modalités d'évaluation (cas d'usage unique et questionnaire) - Le coût de la certification est compris dans le coût du stage.



Contenu de la formation :

La documentation

- Documenter les jeux de données (datasheet, documentation technique)
- Documenter le flux de traitement des données (donnée source jusqu'à l'exploitation, chaîne d'approvisionnement des données)
- Documenter le cycle de vie de la donnée

Les techniques de traitement de données

- Prendre en compte les techniques de génération de données (données synthétiques, confidentialité différentielle, etc.)
- Maîtriser les techniques d'augmentation de données

La préparation des données

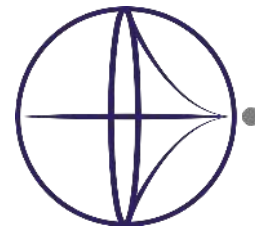
- Evaluer la qualité et la pertinence des données (visualisation, indicateurs (de cohérence), distribution, etc.)
- Identifier les biais les plus courants et leurs atténuations et évaluer les risques résiduels
- Intégrer et maîtriser les méthodes de base de data-cleaning (renseigner données manquantes, identifier les données aberrantes, etc.)

Adapter la solution selon les enjeux sociétaux et les besoins clients

- Comprendre les impacts de la solution par rapport au destinataire direct et indirect
- Connaître, identifier et intégrer les différents risques éthiques et sociétaux associés à l'utilisation de l'IA en fonction du cas d'usage
- Connaître la réglementation en lien avec la confidentialité et utilisation des données et de l'IA pour les usages sensibles
- Connaître les besoins métiers et comprendre les besoins exprimés
- Comprendre le cas d'usage en fonction des besoins métiers

Mesure et suivi de la performance

- Identifier et corriger les dérives du modèle (apprentissage en continu avec des dérives liées aux nouvelles données)
- Identifier, définir et mesurer les indicateurs de performance dès la conception de la solution
- Analyser et ré-évaluer de manière périodique les indicateurs de performance
- Choisir et documenter le modèle de stockage adapté en fonction du cas d'usage et des données sources
- Maîtriser les techniques de versionnage et de mise en production des modèles
- Maîtriser les techniques de Mlops et frameworks associés (MLFlow, ...), conteneurisation (docker) et les articuler avec les techniques de Devops
- Développer des tableaux de bord pour le suivi des performances et les métriques associées
- Intégrer les retours utilisateurs et les limites d'utilisation (retour du contrôle vers un humain)



Les menaces

- Connaître les menaces qui pèsent sur l'élaboration d'une solution d'IA et comprendre l'état de l'art des mécanismes d'atténuation (adversarial example) et évaluer les risques résiduels

Les connaissances générales liées aux modèles d'IA

- Comprendre les modèles d'IA et leurs champs d'application
- Maîtriser les outils (Tensorflow, Keras, ...) pour faire de L'IA
- Compréhension, prise de recul par rapport des cas d'usage
- Intégrer les contraintes opérationnelles (accès aux données, SI, etc...)
- Sensibilisation Écoconception (ex-gestion de flux de données vertueuse, code optimisé, ...)
- Apprendre les bonnes pratiques pour gérer un projet d'Intelligence artificielle
- Maîtriser les bases et les bibliothèques de Python, et leur application à l'intelligence artificielle

La modélisation IA

- Connaissance des différents modèles d'IA et de leurs spécificités (ex. contraintes liées aux différentes formes d'apprentissage)
- Apprendre à développer en Python des modèles de IA (prédiction, classification, réseaux de neurones)
- Maîtriser les environnements de développement de l'IA (Notebook Jupyter, ...)

Les méthodes d'apprentissage

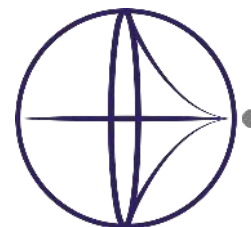
- Apprendre à Optimiser le code & automatiser des tâches
- Optimiser les méthodes d'apprentissage au regard du jeu de données
- Elargir les connaissances sur les bibliothèques du marché.

Industrialisation et architecture

- Bonnes pratiques Mlops, versionning, dév. Continu, CI/CD... Savoir « exposer » l'IA (API, web service, events, etc)
- Développement d'IHM basiques pour première retitution MVP (Tkinter, Flask, Django, ...)
- Avoir une bonne connaissance des principaux éléments d'architectures impliquées pour l'IA
- Être capable de challenger les propositions techniques sous les contraintes performance, économique, écologique, ...)
- Apprendre à utiliser des outils de monitoring d'entraînement de type Tensorboard
- Apprendre à utiliser les bibliothèques d'optimisation de modèles de type Optuna
- Avoir des notions d'architecture d'un système d'information intégrant de L'IA

Oidaneos

Immeuble Le Sommet 42 rue Ferdinand Forest
97122 Baie-mahault
Email : start@oidaneos.com
Tel : 0690649318



Inscription :

- Demande d'informations : <https://www.oidaneos.com/formation/concevoir-implementer-solution-ia-professionnels-it/>
- Mail : start@oidaneos.com
- Téléphone : +590690649318
- Sur le site de l'Opco Atlas : <https://campus.opco-atlas.fr/>



Atlas

certifAtlas

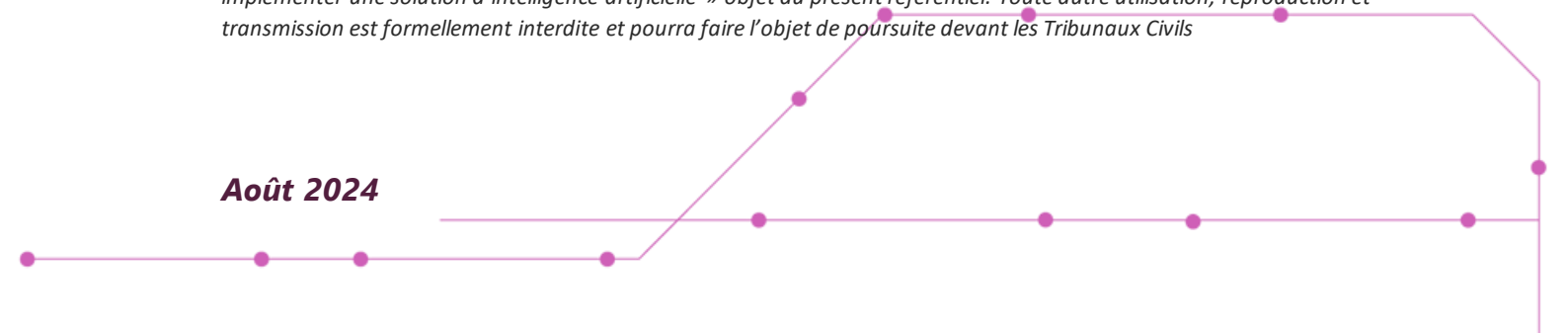
Validez les compétences
stratégiques de votre secteur

Présentation et référentiel de Certification de la certification « Concevoir et implémenter une solution d'intelligence artificielle »

Branche des bureaux d'études techniques, des cabinets
d'ingénieurs-conseils et des sociétés de conseils (IDCC
1486)

La Commission Paritaire Nationale de l'Emploi et de la Formation Professionnelle (CPNEFP) de la Branche des bureaux d'études techniques, des cabinets d'ingénieurs-conseils et des sociétés de conseils demeure titulaire de l'intégralité des droits de propriété Intellectuelle et/ou matérielle sur l'ensemble du document ci-après dénommé référentiel de certification qui est sa propriété et qui ne peut être utilisé que dans le cadre exclusif de la certification « Concevoir et implémenter une solution d'intelligence artificielle » objet du présent référentiel. Toute autre utilisation, reproduction et transmission est formellement interdite et pourra faire l'objet de poursuite devant les Tribunaux Civils

Août 2024



Sommaire

1. Descriptif de la certification	3
1.1 OBJECTIF ET ARCHITECTURE DE LA CERTIFICATION	3
1.2 ACCES A LA CERTIFICATION	3
1.3 PRINCIPES D'ATTRIBUTION DE LA CERTIFICATION	4
2. Descriptif de l'activité professionnelle visée par la certification.....	4
3. Référentiel de certification	5

1. Descriptif de la certification

1.1 Objectif et architecture de la certification

La certification « Concevoir et implémenter une solution d'intelligence artificielle » permet de reconnaître les compétences liées au développement d'une solution d'intelligence artificielle (IA)

1.2 Accès à la certification

- **Public**

La certification s'adresse à un public de professionnels exerçant des activités orientées vers la création d'infrastructures, de logiciels, vers le traitement et l'analyse de données ou encore vers la maintenance des systèmes et qui souhaitent acquérir des compétences leur permettant de contribuer au développement d'une solution d'intelligence artificielle (IA).

La Branche a identifié trois catégories de publics cibles :

- **Professionnels data** : professionnels de l'informatique et du traitement des données (data engineer, data analyst, business data analyst et architecte data). Ces professionnels sont le cœur de cible de cette certification.
- **Professionnels IT** : professionnels de l'IT, dont le cœur de métier n'est pas initialement l'exploitation des données. Leur expertise se concentre sur la gestion et l'optimisation des systèmes et de l'infrastructure informatiques et non sur l'exploitation des données. Ils peuvent être par exemple : développeur informatique, spécialiste base de données, consultant architecte technique, spécialiste infrastructure, spécialiste systèmes, réseaux et sécurité, intégrateur logiciel métier / spécialiste middleware).
- **Data scientists** : dont les missions en termes de traitement de données sont plus avancées (notamment l'exploration de données avancée, la modélisation statistique et l'utilisation de techniques d'apprentissage automatique et d'intelligence artificielle pour résoudre des problèmes complexes) et qui disposent de compétences déjà solides en mathématiques et statistiques.

- **Pré-requis**

Les pré-requis à l'entrée en formation sont les suivants :

- Exercer un emploi comportant des activités orientées vers la création d'infrastructures, de logiciels, vers le traitement et l'analyse de données ou encore vers la maintenance des systèmes ;
- Des connaissances en mathématiques et statistiques, en particulier dans l'exploitation de données (statistiques descriptives, probabilités, statistiques inférentielles, analyse exploratoire des données, modélisation statistique) ;
- Justifier d'une expérience en programmation (1^{er} niveau de maîtrise des langages de programmation, connaissances algorithmes et structures de données, architecture logicielle, systèmes d'exploitation).

- Voies d'accès

L'accès à la certification « Concevoir et implémenter une solution d'intelligence artificielle » s'effectue par la voie de la formation dans le cadre d'un parcours en continu.

1.3 Principes d'attribution de la certification

Chaque compétence est évaluée selon les critères et modalités d'évaluation définis au sein du référentiel de certification.

Les évaluations ont pour objet d'apprécier la mise en œuvre des compétences définies au sein du référentiel de certification selon les critères définis.

2. Descriptif de l'activité professionnelle visée par la certification

La certification s'adresse à un public de professionnels exerçant des activités orientées vers la création d'infrastructures, de logiciels, vers le traitement et l'analyse de données ou encore vers la maintenance des systèmes et qui souhaitent acquérir des compétences leur permettant de contribuer au développement d'une solution d'intelligence artificielle (IA). La certification vise des compétences transversales pour les métiers techniques de l'IA/IT permettant de contribuer au développement et à l'implémentation d'une solution d'IA.

Les composantes visées sont les suivantes :

- La préparation des données nécessaires au projet d'Intelligence Artificielle
- Le développement et la mise en œuvre d'une solution d'Intelligence Artificielle
- Le maintien et l'amélioration d'une solution d'Intelligence Artificielle

La certification s'articule autour de 9 compétences, qui devront toutes être maîtrisées à l'issue de la formation :

- C1. Identifier un jeu de données pour répondre aux besoins métiers et aux cas d'usage, en tenant compte des enjeux de pertinence et de cohérence.
- C2. Identifier les risques éthiques et sociétaux à prendre en compte dans le cadre de l'exploitation de la solution d'IA pour prévenir les dérives éventuelles, en tenant compte du cadre réglementaire.
- C3. Préparer les données pour renforcer leur intégrité et leur pertinence en vue du développement de la solution IA, en mobilisant les techniques de traitement adaptées et en tenant compte des attendus (besoins métiers, cas d'usage etc.) identifiés en phase de cadrage du projet
- C4. Choisir un modèle IA pour disposer d'une solution adaptée et performante par rapport aux cas d'usage, en mesurant sa pertinence et en mobilisant une démarche scientifique
- C5. Entraîner le modèle d'IA de façon automatique et supervisée pour valider la pertinence des solutions envisagées, au regard des cas d'usage énoncés par le métier.

- C6. Implémenter le modèle d'IA en intégrant les briques technologiques (moteurs, reporting, suivi des prévisions etc.) au sein de l'environnement technique choisi pour exploiter la solution.
- C7. Contribuer à la conception et à l'évaluation de la proposition d'architecture cible, en identifiant les contraintes avec l'appui des acteurs pertinents, pour garantir les performances attendues.
- C8. Mesurer la performance et les impacts de la solution d'IA pour maintenir son application fonctionnelle, conformément aux cas d'usage et aux enjeux identifiés.
- C9. Adopter une démarche d'amélioration continue de la solution IA, pour garantir son évolution au fil du temps, dans le respect des exigences de la commande initiale et en tenant compte des évolutions des besoins utilisateurs et des données mobilisables.

3. Référentiel de certification

Se référer au tableau suivant.

Description du métier, de l'activité ou de la situation professionnelle à partir duquel le dispositif visant la certification est initié :		
La certification s'adresse à un public de professionnels exerçant des activités orientées vers la création d'infrastructures, de logiciels, vers le traitement et l'analyse de données ou encore vers la maintenance des systèmes et qui souhaitent acquérir des compétences leur permettant de contribuer au développement d'une solution d'intelligence artificielle (IA). La certification vise des compétences transversales pour les métiers techniques de l'IA/IT permettant de contribuer au développement et à l'implémentation d'une solution d'IA. Les composantes visées sont les suivantes : - La préparation des données nécessaires au projet d'Intelligence Artificielle - Le développement et la mise en œuvre d'une solution d'Intelligence Artificielle - Le maintien et l'amélioration d'une solution d'Intelligence Artificielle		
Référentiel de compétences	Référentiel de certification	
	Modalités d'évaluation	Critères d'évaluation
C1. Identifier un jeu de données pour répondre aux besoins métiers et aux cas d'usage, en tenant compte des enjeux de pertinence et de cohérence.	<u>Un cas d'usage unique permettant d'évaluer toutes les compétences du référentiel, qui consiste en :</u> 1. Une production à partir d'un ou plusieurs jeu(x) de données. 2. La production d'un journal de bord permettant de consigner la méthode et les partis pris adoptés au fil du travail réalisé à partir du ou des jeu(x) de données.	Le jeu de données créé est représentatif et adapté aux besoins métiers et aux cas d'usage : • Les besoins métiers sont correctement identifiés. • Les cas d'usages sont correctement décrits. • Les données a priori pertinentes (et celles nécessaires a minima) pour alimenter le modèle d'IA sont identifiées. • L'existence, la disponibilité et l'accès des données sont vérifiés. • Le cas échéant, des solutions alternatives sont envisagées en cas d'indisponibilité, d'absence ou d'inaccessibilité des données.
C2. Identifier les risques éthiques et sociétaux à prendre en compte dans le cadre de l'exploitation de la solution d'IA pour prévenir les dérives éventuelles, en tenant compte du cadre réglementaire.	Ces deux productions sont restituées par le candidat dans un même cahier électronique (notebook) pouvant rassembler du texte, des images, des formules mathématiques et du code informatique exécutable (notebook type Jupyter ou Kaggle). Le choix du notebook est indiqué par le jury d'évaluation en amont de la production.	Les risques éthiques et sociétaux associés à l'exploitation de la solution d'IA sont correctement identifiés et pris en compte : • Les chartes éthiques (européennes et françaises) sont connues et appliquées. • Les impacts éthiques et sociétaux liés au déploiement de la solution d'IA sont connus et leurs conséquences comprises. • Les biais potentiels ou existants sont identifiés lorsque cela est possible. • Les dilemmes éthiques sont identifiés le cas échéant. • Les risques éthiques et sociétaux identifiés sont portés à la connaissance des acteurs concernés (commanditaire, juristes, etc.).

	3. Une soutenance orale Après remise du notebook susmentionné et d'un support de présentation dédié à la soutenance orale (qui peut par exemple être au format power point), le candidat présente oralement son cas d'usage devant le jury d'évaluation. La présentation est suivie d'un temps d'échanges et de questions où le jury peut approfondir un point ou des points particuliers pour s'assurer de la maîtrise des compétences par le candidat. La soutenance peut s'effectuer en présentiel (dans des locaux mis à disposition par l'organisme de formation organisant l'évaluation ou en entreprise par exemple) ou en distanciel. Durée de la soutenance orale : 1 heure (30 minutes de présentation par le candidat + et 30 minutes d'échanges avec le jury. La durée totale de l'évaluation via le cas d'usage est d'environ un mois. Elle est décomposée de la manière suivante : • Production à partir du ou des jeu(x) de données et du journal de bord (notebook) + production du support de la soutenance orale : 3 semaines • Soutenance orale : une semaine après le rendu de la partie écrite. Evaluateurs : Jury d'évaluation composé de deux professionnels externes au dispositif de formation et d'emploi du candidat.	• La vérification par les acteurs concernés des problèmes légaux et éthiques liés au jeu de données est faite.
C3. Préparer les données pour renforcer leur intégrité et leur pertinence en vue du développement de la solution IA, en mobilisant les techniques de traitement adaptées et en tenant compte des attendus (besoins métiers, cas d'usage etc.) identifiés en phase de cadrage du projet.		Les données sont préparées de façon à permettre une exploitation efficace et cohérente par rapport aux attendus du projet : • Les données sont correctement nommées ou renommées. • Le format des données est adapté à l'usage auquel elles sont destinées. • Les données altérées, inexactes ou non pertinentes sont corrigées ou supprimées (blancs effacés, suppression des fichiers vides et des doublons, etc.). • Les traitements effectués sont correctement documentés. • Le choix du modèle de stockage (object storage, base de données de documents, base de données relationnelle, etc.) est adapté aux types de données détenues et aux types d'usages auxquels elles sont destinées. • Le cycle de vie du jeu de données est correctement documenté (accessibilité des données résultant du traitement vérifiée, prise en compte de usages futurs, des caractéristiques des cas d'usage, de la gouvernance, etc.). • Le cycle de vie du jeu de données dûment documenté est soumis aux parties prenantes.
C4. Choisir un modèle IA pour disposer d'une solution adaptée et performante par rapport aux cas d'usage, en mesurant sa pertinence et en mobilisant une démarche scientifique.		Le modèle d'IA choisi offre une solution adaptée et performante par rapport aux cas d'usage : • La pertinence du modèle est évaluée grâce aux bons indicateurs (analyse Receiver Operating Characteristic, ou ROC). • Les contraintes opérationnelles sont effectivement prises en compte. • Les contraintes en matière d'éco-conception sont portées à la connaissance des acteurs pertinents pour choisir la stratégie adaptée. • Les grandes familles d'algorithmes sont connues (outils, contraintes, etc.). • La démarche scientifique (critères, problématiques, etc.) est correctement documentée. • La performance attendue est déterminée (niveau de précision, temps de traitement et d'inférence, prise en compte des déterminants de la performance énergétique, etc.).

	<u>Un questionnaire à visée professionnelle</u> En complément du cas d'usage, le questionnaire à visée professionnelle permet d'évaluer la maîtrise professionnelle du candidat pour les compétences 1, 2 et 4. Le candidat répond aux différentes questions posées (études de cas, mises en situation fictives, etc.) Les questions sont de différentes natures et peuvent s'appuyer sur des photos et des schémas qui permettent aux candidats de se projeter dans son univers professionnel. Le questionnaire est généré de manière aléatoire. Le questionnaire comporte 15 questions. Le test complet dure 45 minutes. L'évaluation s'effectue en ligne et est automatisée.	- Le type de résultat attendu est identifié (probabiliste, déterministe, etc.). - Le contexte des cas d'usage est pris en compte. - Le modèle d'apprentissage choisi est cohérent par rapport aux résultats attendus. - La pertinence d'utiliser des solutions techniques sur l'étagère est évaluée le cas échéant.
C5. Entraîner le modèle d'IA de façon automatique et supervisée pour valider la pertinence des solutions envisagées, au regard des cas d'usage énoncés par le métier.		La méthode d'entraînement du modèle d'IA est adaptée aux résultats attendus : - Le modèle d'apprentissage est optimisé suivant le contexte du projet. - Le modèle créé a entraîné le cas échéant. - Le modèle choisi a réentraîné le cas échéant. - Les connaissances sont transférées d'un modèle à l'autre le cas échéant. - Les hyperparamètres du modèle sont décrits. - Le travail de feature engineering est effectué lorsque c'est possible (usages classes plutôt que valeurs brutes, réduction de la taille du data set pour optimiser les performances via l'usage de dérivés, etc.).
C6. Implémenter le modèle d'IA en intégrant les briques technologiques (moteurs, reporting, suivi des prévisions etc.) au sein de l'environnement technique choisi pour exploiter la solution.		L'exploitation de la solution est permise par l'intégration des briques technologiques au sein de l'environnement technique choisi : - Le processus de livraison et de déploiement continu est mis en œuvre. - Le versionning est implémenté. - Les besoins d'intégration sont documentés.
C7. Contribuer à la conception et à l'évaluation de la proposition d'architecture cible, en identifiant les contraintes avec l'appui des acteurs pertinents, pour garantir les performances attendues.		La pertinence de l'architecture cible est évaluée avec l'appui des acteurs ressources : - Les principales architectures et leurs contraintes sont connues. - Les contraintes économiques des différents scénarios sont portées à la connaissance des acteurs pertinents pour dimensionner au mieux la proposition d'architecture. - Le cas échéant les acteurs métiers et / ou le commanditaire et / ou les acteurs techniques du projet sont interrogés pour préciser les contraintes techniques liées à la généralisation de la solution.

C8. Mesurer la performance et les impacts de la solution d'IA pour maintenir son application fonctionnelle, conformément aux cas d'usage et aux enjeux identifiés.		La performance de la solution d'IA est correctement mesurée : • Des indicateurs de performance (et seuils associés) adaptés aux attendus de la solution IA sont définis. • La performance de la solution est mesurée grâce au suivi des indicateurs définis. • Les résultats de l'exploitation de la solution IA sont interprétés et présentés aux interlocuteurs concernés. • Les actions adaptées sont le cas échéant déclenchées en fonction des résultats de l'analyse des indicateurs.
C9. Adopter une démarche d'amélioration continue de la solution IA, pour garantir son évolution au fil du temps, dans le respect des exigences de la commande initiale et en tenant compte des évolutions des besoins utilisateurs et des données mobilisables.		La permanence de l'adéquation de la solution d'IA avec les besoins identifiés est assurée grâce à la mise en place d'un cadre d'évaluation pertinent et évolutif : Un système d'évaluation automatisé du modèle d'IA et totalement intégré aux processus CI/CD grâce aux pratiques MLOPS est mis en place • Les métriques (taux de prévision, robustesse, variations de performance, obsolescence, etc.) sont intégrées. • La pertinence des indicateurs de performance est interrogée selon une périodicité définie en phase de cadrage du projet.