

Expert en ingénierie de données

RNCP40875 · RNCP · Publiée

L'essentiel

Niveau de qualification	Niveau 7
Code(s) NSF	326
Formacode(s)	31025, 31026, 31028, 11052
Code(s) ROME	M1805, M1806, M1803, M1802
Type d'enregistrement	Enregistrement sur demande
Date d'enregistrement	25/06/2025
Date d'échéance	25/06/2027

Certificateur(s)

EFREI PARIS · SIRET : 39889833800023

Partenaires habilités

14 organismes habilités, d'après la liste publiée par France compétences : 14 à former et organiser l'évaluation.

Organisme	SIRET	Habilitation
CALTEA	93893471800017	Former et organiser l'évaluation
CALTEA	93893471800025	Former et organiser l'évaluation
CALTEA	93893471800033	Former et organiser l'évaluation
CODA	92219486500031	Former et organiser l'évaluation
CODA	92219486500049	Former et organiser l'évaluation
CODA	92219486500056	Former et organiser l'évaluation
DECODE	93951037600022	Former et organiser l'évaluation
DRS SCHOOL	85398915000018	Former et organiser l'évaluation
ECOLE 18.06 ALSACE	94480869000027	Former et organiser l'évaluation
INGETIS	79349617500035	Former et organiser l'évaluation

Liste complète des 14 partenaires habilités : <https://data.aktaconseil.fr/certifications/rncp40875-expert-en-ingenierie-de-donnees/partenaires>

Résumé de la certification

OBJECTIFS ET CONTEXTE

Le métier d'Expert en ingénierie de données est devenu de plus en plus important ces dernières années avec l'explosion des données générées par les entreprises et les utilisateurs, les Bigs Data et l'Intelligence Artificielle (IA) dont l'IA générative. Les experts en ingénierie de données sont des spécialistes de la gestion de la data, qui collectent, stockent, traitent et analysent de très grands volumes de données brutes afin de les transformer en informations utiles et à haute valeur ajoutée pour une organisation tout en garantissant leur sécurité et conformité. Leur rôle est fondamental dans la création d'une infrastructure de données fiable et efficace permettant aux

entreprises de prendre des décisions éclairées et de rester compétitives dans un environnement de plus en plus complexe. L'Expert en ingénierie de données travaille en étroite collaboration avec les Data Scientists et les Data Analysts pour garantir la qualité et la sécurité des données, ainsi que pour développer des outils d'analyse et de visualisation.

Un expert en ingénierie de données dispose d'une triple expertise technique lui permettant d'être le garant de la conception et le développement d'une architecture de données efficiente, de piloter et d'implémenter des solutions d'Intelligence Artificielle (IA), d'implémenter et d'automatiser le déploiement de solution(s) sur le cloud et "on-premise" (c'est à dire au sein de l'environnement de stockage de données d'une entreprise) et de déployer une stratégie de management et de gouvernance de la donnée.

Son expertise en tant que pilote de projet et manager lui permet de garantir la réussite des projets et l'accompagnement des ressources humaines internes ou externes par un management adapté à la taille , culture et attentes de l'entreprise dans une démarche inclusive fondée sur l'intelligence collective.

ACTIVITÉS VISÉES

Construction et développement des architectures de stockage de données

Conception et développement de solutions de traitement de données

Préparation et Visualisation de données

Implémentation d'algorithmes de machine Learning

Mise en œuvre de projets avec l'IA générative

Mise en œuvre de solutions de stockage et de traitement de données dans le cloud

Intégration et déploiement continus des données

Gestion des infrastructures

Définition d'une stratégie de management de données

Élaboration d'une stratégie de sécurité et de sécurisation des données

Piloter un projet Data

Manager des équipes projets Data

COMPÉTENCES ATTESTÉES

Concevoir et développer une base de données relationnelle universellement accessible en réponse aux besoins d'un client/d'une gouvernance, en mobilisant les technologies et les langages de requêtes adaptés, en collaboration avec l'architecte de base de données, afin de fournir une solution fiable et adaptée à la mise à disposition de données structurées.

Concevoir et développer une base de données non?relationnelle universellement accessible permettant la mise à disposition des données semi-structurées et non structurées, en utilisant des technologies NoSQL les plus adaptées, en étroite collaboration avec l'architecte de base de données et les Data Scientists pour répondre aux besoins de traitement analytique et d'intelligence artificielle.

Concevoir et construire un lac de données (Data Lake) universellement accessible tenant compte des contraintes et des règlements de sécurité et de protection des données en vigueur en choisissant les architectures, les indicateurs de performance et les solutions de stockage appropriées, en coordination avec la gouvernance des données, afin d'intégrer les données de sources multivariées et d'optimiser leur exploitation.

Architecturer des infrastructures scalables et résilientes en utilisant des systèmes de haute disponibilité, en collaboration avec l'architecte de données et l'architecte d'infrastructures pour permettre une adaptation à la croissance de l'entreprise et une résilience face aux pannes dans le cadre du traitement de données massives.

Créer une API (Application Programming Interface) en utilisant les technologies permettant de rendre accessibles les données et facilitant la communication entre divers services informatiques en respectant les normes et les standards en vigueur et en collaborant avec les équipes IT pour assurer l'interopérabilité des systèmes.

Implémenter un système distribué en utilisant des technologies de streaming, en travaillant avec les équipes techniques et métiers, afin de traiter les données sur des périodes définies et de répondre aux besoins d'analyse rapide

Transformer les données provenant de différentes sources en prenant en compte la variété de données et en mobilisant des outils d'intégration permettant l'analytique de données à grande échelle, en collaboration avec les parties prenantes, pour une meilleure intégration, formatage et stockage des données multidimensionnelles.

Optimiser la performance des pipelines en utilisant les techniques d'intégration et de mise en scène adéquates afin de garantir la performance et la qualité des analyses en aval exploitables par les collaborateurs internes.

Préparer les données en les transformant et en les nettoyant, en utilisant des outils appropriés, en collaboration avec les gestionnaires et les analystes de données, afin d'assurer une qualité optimale et universellement accessible de

ces dernières pour les différents besoins métiers dont l'analyse et le reporting.

Élaborer une communication infographique visuelle inclusive en construisant des tableaux de bord interactifs en collaboration avec les équipes métiers et les data analystes afin de communiquer les résultats d'analyses, d'assurer l'extraction de connaissances en temps réel et favoriser la prise de décision éclairée par la gouvernance.

Mettre en place des processus d'analyse exploratoire de données en utilisant des techniques statistiques et des outils adaptés, en collaboration avec les équipes métiers, pour générer des insights exploitables pour les décisions stratégiques.

Définir une stratégie d'intégration de l'IA en identifiant les cas d'usage pertinents et en évaluant leur impact sur les processus métiers, en concertation avec les responsables IA et les responsables métiers, afin d'aligner les objectifs de l'IA aux exigences de l'écosystème, aux parties prenantes ainsi qu'aux objectifs de la gouvernance.

Développer des modèles prédictifs dont du "machine Learning" pour identifier de nouveaux comportements et usages en collaboration avec les équipes métiers et les Ops, afin de fournir des insights exploitables pour la prise de décision par la gouvernance.

Évaluer la performance des modèles de prédiction développés en analysant leurs résultats à l'aide de métriques adaptées et de leur degré d'écoresponsabilité, en les comparant avec d'autres modèles et en prenant en compte les besoins et attentes des parties prenantes, pour garantir l'efficacité et la pertinence du modèle retenu.

Identifier les cas d'usage possibles de l'IA générative en évaluant leur pertinence en faveur des processus métiers et en choisissant les modèles appropriés en fonction des objectifs définis par la gouvernance, en collaboration avec le responsable IA et les équipes métiers, pour maximiser le pilotage par la valeur ajoutée des données.

Développer des solutions basées sur des modèles d'IA générative en utilisant les modèles de base "foundation models" et les modèles de langage adaptés, en concertation avec les équipes IT et autres parties prenantes, afin de créer des solutions innovantes, facilement utilisables et adaptées au contexte de l'organisation

Évaluer la qualité des résultats générés par les modèles d'IA générative en mettant en place des métriques d'évaluation adaptées aux cas d'usage et en ajustant les paramètres du modèle pour améliorer les performances des applications et solutions proposées.

Mettre en œuvre des solutions de stockage de données dans le cloud en utilisant des stratégies adaptées en prenant compte les veilles technologiques et légales ainsi que l'impact environnemental de la solution, en coordination avec l'architecte cloud et les équipes IT pour permettre une exploitation optimale des données tout en choisissant des solutions adaptées aux besoins de stockage et d'archivage.

Concevoir et développer des pipelines de traitement de données dans le cloud en chargeant, transformant et mettant les données à disposition des utilisateurs, en coordination avec les analystes de données, pour garantir l'efficacité d'exploitation des données .

Mettre en place une politique de sécurité des données dans le cloud en développant des stratégies de chiffrement et de gestion des accès, en collaboration avec les équipes de sécurité informatique, pour assurer la conformité avec les réglementations en vigueur, les normes de Qualité et de sécurité de l'organisation ainsi que la protection des données sensibles.

Optimiser les solutions de stockage et de traitement dans le cloud en définissant des KPI de performance, de sécurité et de conformité en coordination avec l'architecte et l'ingénieur cloud, pour garantir la disponibilité des services, leur amélioration continue et optimiser les coûts d'infrastructure.

Piloter une stratégie d'intégration continue des pipelines de données en s'appuyant sur une veille technologique et d'innovation pour identifier et adopter les meilleures solutions, en mobilisant des technologies de containerisation et d'ordonnancement, en collaboration avec les équipes DevOps, afin de garantir l'automatisation et l'efficacité des processus de traitement des données.

Développer et automatiser des pipelines d'intégration continue et de développement continu (CI/CD) permettant la livraison rapide, fiable et conforme des applications, en intégrant des mécanismes de tests automatisés et en travaillant avec les ingénieurs DevOps, les équipes de développement et de qualité logicielle, afin de garantir la qualité des livrables, réduire les erreurs humaines et accélérer les mises en production.

Automatiser le processus de traitement et de stockage de données en utilisant des outils d'automatisation de flux de données, en concertation avec la gouvernance, les ingénieurs DevOps et les autres équipes métiers afin d'améliorer la qualité, la sécurité, la fiabilité et la conformité des données.

Assurer l'efficacité des processus CI/CD dans un contexte MLOps en supervisant, évaluant et déployant automatiquement les modèles actualisés, en gérant leurs versions avec des outils adaptés innovants et/ou mis à jour en collaboration avec les Data Scientists, afin d'optimiser le cycle de vie des modèles d'apprentissage machine.

Administrer les ressources cloud et on-premise de manière optimale en surveillant et ajustant l'utilisation des infrastructures à l'aide d'outils de monitoring en collaboration avec l'ingénieur et l'administrateur cloud, afin d'assurer

la disponibilité, la sécurité, la conformité et la résilience des systèmes.

Automatiser les processus de gestion d'infrastructure en utilisant les outils d'"infrastructure as code" (IaC) en coordination avec les équipes DevOps, pour optimiser et rendre efficaces la gestion et le déploiement des infrastructures.

Définir une gouvernance de données en mettant en place les politiques et les standards, en collaboration avec les équipes de gouvernance pour établir les rôles, responsabilités à partir du T.O.M (Target Operating Model) et propriété des données afin d'aligner la gouvernance de données avec les objectifs stratégiques.

Mettre en place les bonnes pratiques de gestion et de sécurisation de données en appliquant les réglementations en vigueur pour la protection de données, en collaboration avec le gestionnaire de données, afin de garantir la transparence et le respect de la vie privée de toutes les parties prenantes.

Superviser le cycle de vie des données dans les projets data-driven en établissant des processus de gestion (création, stockage, utilisation, archivage, suppression) en collaboration avec les équipes IT et de gouvernance des données, afin de garantir la conformité aux réglementations et d'optimiser l'utilisation des ressources tout au long du cycle de vie des données.

Assurer et diffuser une veille réglementaire, légale et normative au niveau national, européen et international en matière de gestion des données et d'intelligence artificielle, afin de garantir auprès de l'ensemble des collaborateurs la conformité des pratiques et des solutions avec les lois, normes ISO et standards en vigueur.

Auditer la qualité des données en utilisant des outils de monitoring et en interaction avec les équipes de gouvernance et le gestionnaire de données, afin de garantir l'adéquation des données aux besoins stratégiques ainsi que les objectifs définis au préalable.

Mettre en place une stratégie d'amélioration continue en prenant en considération les résultats de l'audit, les exigences de l'ensemble des parties prenantes, les meilleures pratiques et labels afin de garantir le maintien de la qualité optimale des données et ancrer la culture de la donnée dans l'ensemble des processus de l'organisation.

Élaborer une stratégie de protection et de sécurisation des données en mettant en place des politiques d'accès aux données, et en développant une stratégie de chiffrement, de gestion des identités et des accès ainsi que de la sécurisation du stockage dans le respect des réglementations et bonnes pratiques ainsi que la conception universelle en vigueur en collaboration avec les équipes de sécurité pour garantir la protection de données.

Mettre en place des stratégies de sécurité pour les infrastructures en définissant des protocoles de protection contre les cybermenaces et de résilience en cas d'attaque, en utilisant des outils de sécurité et un protocole de gestion de crise, en collaboration avec les équipes de sécurité informatique afin d'assurer la protection des infrastructures et la continuité des services.

Définir le périmètre du projet Data avec la gouvernance en dimensionnant l'ensemble des contraintes et risques internes et externes du projet, les délais, les KPI, les ressources à mobiliser afin d'assurer une gestion optimale du projet data.

Proposer la méthodologie projet la plus adaptée au Data et à la ressource humaine disponible en se basant sur les forces et faiblesses des méthodologies usuelles (Agile, Cycle en V, etc.) et sur l'élaboration d'une feuille de route méthodologique afin de garantir l'efficacité du projet Data.

Planifier les étapes clés de la réalisation du projet en organisant l'attribution des missions et responsabilités afférentes aux ressources humaines internes et externes nécessaires afin de garantir le respect et la tenue des différentes étapes du projet.

Définir et suivre les budgets avec la gouvernance en prenant en compte le cahier des charges, les contraintes budgétaires, le ROI, la maîtrise des risques à anticiper afin de garantir l'optimisation financière du projet Data.

Coordonner le niveau d'évolution du projet Data auprès des services en mettant en valeur les points positifs d'avancée, en sensibilisant sur les retards, freins ou difficultés rencontrés pour s'assurer de la diffusion uniforme et universellement accessible de l'information au sein de l'entreprise et le maintien des délais, de la qualité des livrables ainsi que la motivation du collectif.

Diagnostiquer les ressources en compétences existantes et manquantes en interne à partir de la stratégie RH actuelle, et de la politique gouvernance Data afin de sélectionner les profils internes à mobiliser, les recrutements à prévoir ainsi que les prestataires extérieurs à solliciter pour un besoin donné en veillant à bien respecter le principe d'inclusivité.

Identifier les besoins en formation en interne et en externe en collaboration avec la direction des ressources humaines dans une démarche inclusive et égalitaire afin de garantir la montée en compétences et la disponibilité des compétences en interne pour la réalisation d'une transformation durable.

Organiser les activités, les rôles et missions de chacun des membres de l'équipe projet en fonction de leurs compétences et de leurs profils en appliquant une ligne managériale inclusive tenant compte de la diversité, des

situations de handicap et de la multiculturalité afin de favoriser la performance, l'éthique à l'égard de la Data, la réussite du projet ainsi que la Qualité de Vie au Travail (QVT).

Evaluer les compétences ainsi que le niveau d'implication des équipes et des individus sur la base de critères préalablement définis avec la DRH dans le cadre de points d'étape, réunions, ou d'entretien individuel annuel afin de garantir la cohésion d'équipe et l'implication générale dans le déploiement du projet.

Appliquer des méthodes managériales adaptées et inclusives auprès d'équipes inter et multidisciplinaires et multiculturelles, aussi bien techniques que fonctionnelles en adaptant ses actions et décisions aux parties prenantes pour s'assurer de l'adhésion du collectif dans le cadre d'un projet Data avec l'appui de la gouvernance de l'organisation.

Blocs de compétences

RNCP40875BC01 — Construire/superviser et développer une architecture de stockage et de traitement de données

LISTE DE COMPÉTENCES

Concevoir et développer une base de données relationnelle universellement accessible en réponse aux besoins d'un client/d'une gouvernance, en mobilisant les technologies et les langages de requêtes adaptés, en collaboration avec l'architecte de base de données, afin de fournir une solution fiable et adaptée à la mise à disposition de données structurées.

Concevoir et développer une base de données non relationnelle universellement accessible permettant la mise à disposition des données semi-structurées et non structurées, en utilisant des technologies NoSQL*

les plus adaptées, en étroite collaboration avec l'architecte de base de données et les Data Scientists pour répondre aux besoins de traitement analytique et d'intelligence artificielle.

Concevoir et construire un lac de données (Data Lake) universellement accessible tenant compte des contraintes et des règlements de sécurité et de protection des données en vigueur en choisissant les architectures, les indicateurs de performance et les solutions de stockage appropriées, en coordination avec la gouvernance des données, afin d'intégrer les données de sources multivariées et d'optimiser leur exploitation.

Architecturer des infrastructures scalables et résilientes en utilisant des systèmes de haute disponibilité, en collaboration avec l'architecte de données et l'architecte d'infrastructures pour permettre une adaptation à la croissance de l'entreprise et une résilience face aux pannes dans le cadre du traitement de données massives.

Créer une API (Application Programming Interface) en utilisant les technologies permettant de rendre accessibles les données et facilitant la communication entre divers services informatiques en respectant les normes et les standards en vigueur et en collaborant avec les équipes IT pour assurer l'interopérabilité des systèmes.

Implémenter un système distribué en utilisant des technologies de streaming, en travaillant avec les équipes techniques et métiers, afin de traiter les données sur des périodes définies et de répondre aux besoins d'analyse rapide

Transformer les données provenant de différentes sources en prenant en compte la variété de données et en mobilisant des outils d'intégration permettant l'analytique de données à grande échelle, en collaboration avec les parties prenantes, pour une meilleure intégration, formatage et stockage des données multidimensionnelles.

Optimiser la performance des pipelines en utilisant les techniques d'intégration et de mise en scène adéquates afin de garantir la performance et la qualité des analyses en aval exploitables par les collaborateurs internes.

MODALITÉS D'ÉVALUATION

Le bloc est évalué par une mise en situation professionnelle (MSP) réelle ou reconstituée réalisée sous le format d'un projet global de construction-supervision, développement d'une architecture de données. Les évaluations sont constituées d'évaluations formatives et d'évaluations certificatives. Ces dernières sont formalisées par passage devant jury.

Un jury d'évaluation intervient obligatoirement à l'issue de l'ensemble des modalités d'évaluations constitutives de la MSP. Ce jury est habilité

à procéder à la vérification de l'acquisition du bloc de compétences par le candidat.

Le jury d'évaluation est composé de 3 personnes minimum habilitées par le certificateur, dont au minimum 2 sont obligatoirement extérieures au certificateur.

RNCP40875BC02 — Piloter et implémenter des solutions d'IA en s'aidant notamment de l'IA générative

LISTE DE COMPÉTENCES

Préparer les données en les transformant et en les nettoyant, en utilisant des outils appropriés, en collaboration avec les gestionnaires et les analystes de données, afin d'assurer une qualité optimale et universellement accessible de ces dernières pour les différents besoins métiers dont l'analyse et le reporting.

Élaborer une communication infographique visuelle inclusive en construisant des tableaux de bord interactifs en collaboration avec les équipes métiers et les data analysts afin de communiquer les résultats d'analyses, d'assurer l'extraction de connaissances en temps réel et favoriser la prise de décision éclairée par la gouvernance.

Mettre en place des processus d'analyse exploratoire de données en utilisant des techniques statistiques et des outils adaptés, en collaboration avec les équipes métiers, pour générer des insights exploitables pour les décisions stratégiques.

Définir une stratégie d'intégration de l'IA en identifiant les cas d'usage pertinents et en évaluant leur impact sur les processus métiers, en concertation avec les responsables IA et les responsables métiers, afin d'aligner les objectifs de l'IA aux exigences de l'écosystème, aux parties prenantes ainsi qu'aux objectifs de la gouvernance.

Développer des modèles prédictifs dont du "machine Learning" pour identifier de nouveaux comportements et usages en collaboration avec les équipes métiers et les Ops, afin de fournir des insights exploitables pour la prise de décision par la gouvernance.

Évaluer la performance des modèles de prédiction développés en analysant leurs résultats à l'aide de métriques adaptées et de leur degré d'écoresponsabilité, en les comparant avec d'autres modèles et en prenant en compte les besoins et attentes des parties prenantes, pour garantir l'efficacité et la pertinence du modèle retenu.

Identifier les cas d'usage possibles de l'IA générative en évaluant leur pertinence en faveur des processus métiers et en choisissant les modèles appropriés en fonction des objectifs définis par la gouvernance, en collaboration avec le responsable IA et les équipes métiers, pour maximiser le pilotage par la valeur ajoutée des données.

Développer des solutions basées sur des modèles d'IA générative en utilisant les modèles de base "foundation models" et les modèles de langage adaptés, en concertation avec les équipes IT et autres parties prenantes, afin de créer des solutions innovantes, facilement utilisables et adaptées au contexte de l'organisation

Évaluer la qualité des résultats générés par les modèles d'IA générative en mettant en place des métriques d'évaluation adaptées aux cas d'usage et en ajustant les paramètres du modèle pour améliorer les performances des applications et solutions proposées.

MODALITÉS D'ÉVALUATION

Le bloc est évalué par une mise en situation professionnelle (MSP) réelle ou reconstituée réalisée sous le format d'un projet global d'implémentation de solutions d'IA en s'aidant de l'IA générative. Les évaluations sont constituées d'évaluations formatives et d'évaluations certificatives. Ces dernières sont formalisées par passage devant jury.

Un jury d'évaluation intervient obligatoirement à l'issue de l'ensemble des modalités d'évaluations constitutives de la MSP. Ce jury est habilité

à procéder à la vérification de l'acquisition du bloc de compétences par le candidat.

Le jury d'évaluation est composé de 3 personnes minimum habilitées par le certificateur, dont au minimum 2 sont obligatoirement extérieures au certificateur.

RNCP40875BC03 — Implémenter et automatiser le déploiement des solutions de données sur le cloud et on-premise

LISTE DE COMPÉTENCES

Mettre en œuvre des solutions de stockage de données dans le cloud en utilisant des stratégies adaptées en prenant compte les veilles technologiques et légales ainsi que l'impact environnemental de la solution, en coordination avec l'architecte cloud et les équipes IT pour permettre une exploitation optimale des données tout en choisissant des solutions adaptées aux besoins de stockage et d'archivage.

Concevoir et développer des pipelines de traitement de données dans le cloud en chargeant, transformant et mettant les données à disposition des utilisateurs, en coordination avec les analystes de données, pour garantir l'efficacité d'exploitation des données .

Mettre en place une politique de sécurité des données dans le cloud en développant des stratégies de chiffrement et de gestion des accès, en collaboration avec les équipes de sécurité informatique, pour assurer la conformité avec les réglementations en vigueur, les normes de Qualité et de sécurité de l'organisation ainsi que la protection des données sensibles.

Optimiser les solutions de stockage et de traitement dans le cloud en définissant des KPI de performance, de sécurité

et de conformité en coordination avec l'architecte et l'ingénieur cloud, pour garantir la disponibilité des services, leur amélioration continue et optimiser les coûts d'infrastructure.

Piloter une stratégie d'intégration continue des pipelines de données en s'appuyant sur une veille technologique et d'innovation pour identifier et adopter les meilleures solutions, en mobilisant des technologies de containerisation et d'ordonnancement, en collaboration avec les équipes DevOps, afin de garantir l'automatisation et l'efficacité des processus de traitement des données.

Développer et automatiser des pipelines d'intégration continue et de développement continu (CI/CD) permettant la livraison rapide, fiable et conforme des applications, en intégrant des mécanismes de tests automatisés et en travaillant avec les ingénieurs DevOps, les équipes de développement et de qualité logicielle, afin de garantir la qualité des livrables, réduire les erreurs humaines et accélérer les mises en production.

Automatiser le processus de traitement et de stockage de données en utilisant des outils d'automatisation de flux de données, en concertation avec la gouvernance, les ingénieurs DevOps et les autres équipes métiers afin d'améliorer la qualité, la sécurité, la fiabilité et la conformité des données.

Assurer l'efficacité des processus CI/CD dans un contexte MLOps en supervisant, évaluant et déployant automatiquement les modèles actualisés, en gérant leurs versions avec des outils adaptés innovants et/ou mis à jour en collaboration avec les Data Scientists, afin d'optimiser le cycle de vie des modèles d'apprentissage machine.

Administrer les ressources cloud et on-premise de manière optimale en surveillant et ajustant l'utilisation des infrastructures à l'aide d'outils de monitoring en collaboration avec l'ingénieur et l'administrateur cloud, afin d'assurer la disponibilité, la sécurité, la conformité et la résilience des systèmes.

Automatiser les processus de gestion d'infrastructure en utilisant les outils d'"infrastructure as code" (IaC) en coordination avec les équipes DevOps, pour optimiser et rendre efficaces la gestion et le déploiement des infrastructures.

MODALITÉS D'ÉVALUATION

Le bloc est évalué par une mise en situation professionnelle (MSP) réelle ou reconstituée réalisée sous le format d'un projet d'implémentation et d'automatisation

du déploiement des solutions de données sur le Cloud et on-premise. Les évaluations sont constituées d'évaluations formatives et d'évaluations certificatives. Ces dernières sont formalisées par passage devant jury.

Un jury d'évaluation intervient obligatoirement à l'issue de l'ensemble des modalités d'évaluations constitutives de la MSP. Ce jury est habilité

à procéder à la vérification de l'acquisition du bloc de compétences par le candidat.

Le jury d'évaluation est composé de 3 personnes minimum habilitées par le certificateur, dont au minimum 2 sont obligatoirement extérieures au certificateur.

RNCP40875BC04 — Déployer une stratégie de management et de gouvernance de la donnée

LISTE DE COMPÉTENCES

Définir une gouvernance de données en mettant en place les politiques et les standards, en collaboration avec les équipes de gouvernance pour établir les rôles, responsabilités à partir du T.O.M (Target Operating Model) et propriété des données afin d'aligner la gouvernance de données avec les objectifs stratégiques.

Mettre en place les bonnes pratiques de gestion et de sécurisation de données en appliquant les réglementations en vigueur pour la protection de données, en collaboration avec le gestionnaire de données, afin de garantir la transparence et le respect de la vie privée de toutes les parties prenantes.

Superviser le cycle de vie des données dans les projets data-driven en établissant des processus de gestion (création, stockage, utilisation, archivage, suppression) en collaboration avec les équipes IT et de gouvernance des données, afin de garantir la conformité aux réglementations et d'optimiser l'utilisation des ressources tout au long du cycle de vie des données.

Assurer et diffuser une veille réglementaire, légale et normative au niveau national, européen et international en matière de gestion des données et d'intelligence artificielle, afin de garantir auprès de l'ensemble des collaborateurs la conformité des pratiques et des solutions avec les lois, normes ISO et standards en vigueur.

Auditer la qualité des données en utilisant des outils de monitoring et en interaction avec les équipes de gouvernance et le gestionnaire de données, afin de garantir l'adéquation des données aux besoins stratégiques ainsi que les objectifs définis au préalable.

Mettre en place une stratégie d'amélioration continue en prenant en considération les résultats de l'audit, les exigences de l'ensemble des parties prenantes, les meilleures pratiques et labels afin de garantir le maintien de la qualité optimale des données et ancrer la culture de la donnée dans l'ensemble des processus de l'organisation.

Élaborer une stratégie de protection et de sécurisation des données en mettant en place des politiques d'accès aux

données, et en développant une stratégie de chiffrement, de gestion des identités et des accès ainsi que de la sécurisation du stockage dans le respect des réglementations et bonnes pratiques ainsi que la conception universelle en vigueur en collaboration avec les équipes de sécurité pour garantir la protection de données.

Mettre en place des stratégies de sécurité pour les infrastructures en définissant des protocoles de protection contre les cybermenaces et de résilience en cas d'attaque, en utilisant des outils de sécurité et un protocole de gestion de crise, en collaboration avec les équipes de sécurité informatique afin d'assurer la protection des infrastructures et la continuité des services.

MODALITÉS D'ÉVALUATION

Le bloc est évalué par une mise en situation professionnelle (MSP) réelle ou reconstituée réalisée sous le format d'un plan d'actions de déploiement d'une stratégie de management et de gouvernance de la donnée. Les évaluations sont constituées d'évaluations formatives et d'évaluations certificatives. Ces dernières sont formalisées par passage devant jury.

Un jury d'évaluation intervient obligatoirement à l'issue de l'ensemble des modalités d'évaluations constitutives de la MSP. Ce jury est habilité

à procéder à la vérification de l'acquisition du bloc de compétences par le candidat.

Le jury d'évaluation est composé de 3 personnes minimum habilitées par le certificateur, dont au minimum 2 sont obligatoirement extérieures au certificateur.

RNCP40875BC05 — Piloter et manager des équipes projet data dans une organisation

LISTE DE COMPÉTENCES

Définir le périmètre du projet Data avec la gouvernance en dimensionnant l'ensemble des contraintes et risques internes et externes du projet, les délais, les KPI, les ressources à mobiliser afin d'assurer une gestion optimale du projet data.

Proposer la méthodologie projet la plus adaptée au Data et à la ressource humaine disponible en se basant sur les forces et faiblesses des méthodologies usuelles (Agile, Cycle en V, etc.) et sur l'élaboration d'une feuille de route méthodologique afin de garantir l'efficacité du projet Data.

Planifier les étapes clés de la réalisation du projet en organisant l'attribution des missions et responsabilités afférentes aux ressources humaines internes et externes nécessaires afin de garantir le respect et la tenue des différentes étapes du projet.

Définir et suivre les budgets avec la gouvernance en prenant en compte le cahier des charges, les contraintes budgétaires, le ROI, la maîtrise des risques à anticiper afin de garantir l'optimisation financière du projet Data.

Coordonner le niveau d'évolution du projet Data auprès des services en mettant en valeur les points positifs d'avancée, en sensibilisant sur les retards, freins ou difficultés rencontrés pour s'assurer de la diffusion uniforme et universellement accessible de l'information au sein de l'entreprise et le maintien des délais, de la qualité des livrables ainsi que la motivation du collectif.

Diagnostiquer les ressources en compétences existantes et manquantes en interne à partir de la stratégie RH actuelle, et de la politique gouvernance Data afin de sélectionner les profils internes à mobiliser, les recrutements à prévoir ainsi que les prestataires extérieurs à solliciter pour un besoin donné en veillant à bien respecter le principe d'inclusivité.

Identifier les besoins en formation en interne et en externe en collaboration avec la direction des ressources humaines dans une démarche inclusive et égalitaire afin de garantir la montée en compétences et la disponibilité des compétences en interne pour la réalisation d'une transformation durable.

Organiser les activités, les rôles et missions de chacun des membres de l'équipe projet en fonction de leurs compétences et de leurs profils en appliquant une ligne managériale inclusive tenant compte de la diversité, des situations de handicap et de la multiculturalité afin de favoriser la performance, l'éthique à l'égard de la Data, la réussite du projet ainsi que la Qualité de Vie au Travail (QVT).

Évaluer les compétences ainsi que le niveau d'implication des équipes et des individus sur la base de critères préalablement définis avec la DRH dans le cadre de points d'étape, réunions, ou d'entretien individuel annuel afin de garantir la cohésion d'équipe et l'implication générale dans le déploiement du projet.

Appliquer des méthodes managériales adaptées et inclusives auprès d'équipes inter et multidisciplinaires et multiculturelles, aussi bien techniques que fonctionnelles en adaptant ses actions et décisions aux parties prenantes pour s'assurer de l'adhésion du collectif dans le cadre d'un projet Data avec l'appui de la gouvernance de l'organisation.

MODALITÉS D'ÉVALUATION

Le bloc est évalué par une mise en situation professionnelle (MSP) réelle ou reconstituée réalisée sous le format d'un plan d'actions de pilotage d'équipes dans le contexte d'un projet Data. Les évaluations sont constituées d'évaluations formatives et d'évaluations certificatives. Ces dernières sont formalisées par passage devant jury.

Un jury d'évaluation intervient obligatoirement à l'issue de l'ensemble des modalités d'évaluations constitutives de la MSP. Ce jury est habilité

à procéder à la vérification de l'acquisition du bloc de compétences par le candidat.

Le jury d'évaluation est composé de 3 personnes minimum habilitées par le certificateur, dont au minimum 2 sont obligatoirement extérieures au certificateur.

Secteur d'activité et type d'emploi

SECTEURS D'ACTIVITÉS

L'Expert en ingénierie de données peut intégrer tous types de structures privées ou publiques (entreprise, organisation, administration) de tous secteurs confondus (industrie, énergie, acteurs du e-commerce, constructeurs informatiques, télécommunications, administration, banques, assurances, services, éducation, transport, santé, etc.).

Les activités et compétences professionnelles exercées peuvent varier en fonction de la taille de la structure et du type d'activité adapté au mode d'organisation (groupe, entreprise individuelle, TPE, PME, et le type de secteur d'activité (Commerce et E-Commerce, Banques, Assurances, Construction, Immobilier, Transports, etc.) ou d'institutions publiques (Administrations publiques, Ministères, etc.).

TYPE D'EMPLOIS ACCESSIBLES

Data Engineer - Data Analyst - Data Scientist - Big Data architect - Chief Data Officer (CDO) - Data Steward - Consultant Data et IA - Architecte Data sur le Cloud - Ingénieur Machine Learning - Data Manager - Chef de Projet Data

RÈGLEMENTATIONS D'ACTIVITÉ

L'Expert

en ingénierie de données évolue dans un univers de plus en plus réglementé. Cette évolution réglementaire, légale et normative nécessite la maîtrise et à minima la connaissance du dispositif suivant :

1) L'aspect réglementaire et légale :

Référentiel général d'amélioration de l'accessibilité - RGAA

Règlement Général sur la Protection des Données n°2016/679 (RGPD)

Code de la propriété intellectuelle

, en particulier sur le droit d'auteur, la protection des logiciels et bases de données (Livres I et III) et les brevets (livre 6)

En cybersécurité :

.

Directive dite « NIS 2 »

n°2022/2555 et sa transposition en droit français (actuellement en cours de discussion

)

.

Article 2321-4-1 du code de la défense

et le décret n°2024-421

.

Règlement sur la cyberrésilience

n°2022/0272 (A.2.2)

.

Règlement « identité numérique » dit « eIDAS 2 »

n°2024/1183 (A.2.2)

.

Le cadre réglementaire actuel

de la sécurité numérique ANSSI et le suivi de son évolution ? : S'informer sur la réglementation | ANSSI.

En maîtrisant? :

la sécurité des systèmes d'information ; la confiance numérique ; et

la notification des incidents et vulnérabilités.

Loi numérique : vers une meilleure protection des citoyens et des entreprises en ligne | Ministère de l'Économie, des Finances et de l'Industrie et Ministère chargé du Budget et des Comptes publics

– Loi « Sécuriser et réguler l'espace numérique » dite « SREN »

LOI n° 2021-1485 du 15 novembre 2021 visant à réduire l'empreinte environnementale du numérique en France (1) - Légifrance

LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique - Dossiers législatifs - Légifrance

DMA : le règlement sur les marchés numériques ou Digital Markets Act | vie-publique.fr

Politique publique numérique responsable : quels dispositifs ? - Numérique écoresponsable

La loi Informatique et Libertés | CNIL

LOI n° 2023-566 du 7 juillet 2023 visant à instaurer une majorité numérique et à lutter contre la haine en ligne (1) - Légifrance

Au niveau de l'IA :

.

Règlement Européen sur l'Intelligence Artificielle (RIA)

.

(Proposition de) Directive « produits défectueux » n°2022/0302 (COD) et sa future transposition

.

(Proposition de) Directive sur la responsabilité en matière d'IA n°2022/0303 (COD) et sa future transposition

Au niveau de la Data :

.

Directive « Open Data » n°2019/102

.

Règlement sur la gouvernance des données (DGA) n°2022/868

.

Règlement sur les données n°2023/2854

.

Règlement sur les données non-personnelles n°2018/1807

Au niveau des Plateformes / services en ligne :

Règlement sur les services numériques (DSA) n°2022/2065

En prenant

en considération le statut de manager et posture au sein des organisations, le cadre réglementaire et légal est complété par les veilles sous l'angle du droit du travail

:

La prise en compte de la Mise en œuvre des traitements courants de « gestion du personnel » :

.

Article 9 du code civil (protection de l'intimité de la vie privée)

.

Article 8 de la Convention européenne de sauvegarde des droits de l'Homme (protection de l'intimité de la vie privée)

.

Article L1121-1 du code du travail (droits et libertés dans l'entreprise)

.

Article L1222-3 et L. 1222-4 du code du travail (information et contrôle des salariés)

.

Article R. 2312-14 du Code du travail, (base de données économiques, sociales et environnementales - BDESE)

.

Article L2323-32 du code du travail (information/consultation du CSE)

.
Articles 226-1 et suivants du code pénal (protection de la vie privée)
.
Articles 226-16 et suivants du code pénal (atteintes aux droits des personnes résultant des traitements informatiques)
.
La loi Informatique et Libertés
.
Le règlement général sur la protection des données
II) Au niveau des normes ISO en lien avec les métiers visés,
cet expert peut être sollicité sur la maîtrise/expertise ou connaissance des normes suivantes :
ISO - La famille ISO 9000 — Management de la qualité
dont Norme ISO 9001/2015 portant amendement de février 2024 ;
.
La norme ISO 9001:2015 en lien avec les exigences relatives au système de management de la qualité?;
.
L'amendement ISO A1 :2024, publié le 23 février 2024, en réponse directe à l'urgence climatique clarifiant les attentes
en matière de gestion des enjeux climatiques pour des normes spécifiques ;
ISO - La famille ISO 14000 — Management environnemental
?;
.
Norme ISO 14040 et 10044 :2006 en lien avec le management environnemental — Analyse du cycle de vie —
Principes et cadre ;
.
La norme ISO 14062 :2002 : En lien avec l'intégration des aspects environnementaux dans la conception et le
développement de produit (le terme « produit » englobant à la fois les biens matériels et les services) ;
ISO - La famille ISO/IEC 27000 — Management de la sécurité de l'information
?;
.
La norme ISO 27001 :2022 en lien avec la Sécurité de l'information, cybersécurité et protection de la vie privée —
Systèmes de management de la sécurité de l'information ;
.
La norme ISO 27701 en lien avec la gouvernance et les mesures de sécurité à mettre en place pour les données
personnelles ;
ISO 20121:2024 - Systèmes de management responsable appliqués à l'activité événementielle — Exigences et
recommandations de mise en œuvre
La norme ISO 2600, en lien avec la responsabilité sociétale ;
et enfin la norme ISO 42001, si l'Expert exerce au sein d'un organisme fournisseur ou utilisateur d'IA et notamment
pour le bloc optionnel B05.

Marché du travail

Indicateurs au niveau des familles professionnelles (FAP) des métiers visés, agrégés au niveau national — France Travail / DARES (open data).

Tension de recrutement (DARES 2024)	Très élevée (5/5) Moyenne pondérée par l'emploi ; les familles ci-dessous s'échelonnent de 4 à 5 sur 5.
Facteurs principaux	Manque de main-d'œuvre disponible, Lien formation-emploi
Salaire médian (mensuel)	2 892 €
Diplôme le plus fréquent	Diplôme de niveau bac+5 ou plus (doctorat, master ou équivalent) en communication et information (41 % des actifs en emploi)

Familles professionnelles (FAP)	M1X80 — Techniciens d'étude et de développement en informatique · tension 4/5 (Élevée) M1X81 — Techniciens de production, d'exploitation, d'installation, et de maintenance, support et services aux utilisateurs en informatique · tension 4/5 (Élevée) M2X90 — Ingénieurs et cadres d'étude, recherche et développement en informatique et télécom · tension 5/5 (Très élevée) M2X91 — Chefs de projet et directeurs de service informatique · tension 4/5 (Élevée) M2X93 — Experts et consultants en systèmes d'information · tension 5/5 (Très élevée)
Projets de recrutement (BMO 2026)	42 073
Jugés difficiles à recruter	50 %
Emplois saisonniers	2 %
Familles professionnelles (FAP)	M1X80 — Techniciens d'étude et de développement en informatique M1X81 — Techniciens de production, d'exploitation, d'installation, et de maintenance, support et services aux utilisateurs en informatique M2X90 — Ingénieurs et cadres d'étude, recherche et développement en informatique et télécom M2X91 — Chefs de projet et directeurs de service informatique M2X93 — Experts et consultants en systèmes d'information

Voies d'accès

PRÉREQUIS À L'ENTRÉE EN FORMATION

Dans le cadre de la formation initiale, être titulaire au minimum d'un diplôme ou d'un titre de niveau 6, généraliste ou professionnel, dans les métiers du numérique (le cas échéant justifier de la maîtrise des disciplines de l'informatique ou de développement logiciel), du management ou du pilotage de projets.

Dans le cadre de la formation continue, être titulaire d'une certification de niveau 6 ou d'un diplôme équivalent (diplôme étranger...) dans le domaine du numérique ;

À titre dérogatoire, le candidat qui justifie d'un niveau 5 et d'une expérience professionnelle significative dans le numérique peut être reçu pour un entretien de sélection accompagné d'un examen par une commission spécifique.

PRÉREQUIS À LA VALIDATION

Néant

COMPOSITION DU JURY — FORMATION INITIALE, FORMATION CONTINUE, APPRENTISSAGE, CONTRAT DE PROFESSIONNALISATION, VAE

Le jury de certification est composé de 5 personnes, dont au minimum 3 sont obligatoirement extérieures au certificateur :

3 membres externes à l'Efrei incluant le Président de Jury :

- Salariés et/ou indépendants exerçant les métiers visés au minimum depuis 2 ans et n'ayant pas quitté le champ professionnel d'exercice des métiers visés depuis plus de 5 ans
- À minima titulaire d'un diplôme ou d'un titre/certification à finalité professionnelle de même niveau que la certification visée,
- À titre dérogatoire : la capitalisation d'une expérience professionnelle justifiée d'au moins 5 années dans les secteurs d'activités et les métiers visés de la certification, peut donner lieu à habilitation de membre de jury

2 membres internes à l'Efrei : à minima titulaires d'un diplôme ou d'un titre/certification à finalité professionnelle de même niveau que la certification visée, et ce, quel que soit la voie d'acquisition du diplôme ou du titre. Le cas échéant, leur expérience dans les secteurs d'activités de la certification est justifiée.

Financement en alternance (NPEC)

Niveau de prise en charge d'un contrat d'apprentissage par les OPCO, fixé par France compétences par branche professionnelle (CPNE).

Prise en charge usuelle	9 050 €
Selon la branche	8 423 € – 9 800 €
Prise en charge	Branches concernées
9 050 €	Valeur par défaut · 186 branches
9 800 €	CPNE Casinos, CPNE de l'ingénierie, des services informatiques et du conseil, CPNE de l'édition de livres, de l'édition phonographique et de l'édition de musique, CPNE des Télécoms, CPNE des industries de santé (CPNEIS)
8 696 €	CPNE FP Commerce de Détail de l'Horlogerie Bijouterie, CPNEFP Commerce succursaliste de la Chaussure, CPNEFP Commerce à distance, CPNEFP Grands Magasins/ Magasins Populaires, CPNEFP Import-Export
9 276 €	CPNEF de l'Exploitation cinématographique et de la Distribution de films
9 094 €	CPNEFP de la blanchisserie, laverie, location de linge, nettoyage à sec, pressing et teinturerie
9 005 €	CPNE FP Branche des magasins du Bricolage
8 813 €	CPNEFP Professions de la Photographie
8 775 €	CPNEFP de la branche des entreprises du bureau et du numérique (commerces et services)
8 600 €	CPNEFP de l'habillement, mercerie, chaussure et jouet
8 500 €	CPNEFP de la métallurgie
8 423 €	CPNEFP de la Branche des Caisses régionales de Crédit Agricole et autres organismes

Base légale

Date de décision	25/06/2025
Durée de l'enregistrement	2 ans
Date d'échéance de l'enregistrement	25/06/2027
Date de dernière délivrance possible	25/06/2031

Statistiques

Année	Certifiés	Dont VAE	Insertion 6 mois	Métier 2 ans
2025	16	16	100 %	0 %
2024	27	0	54 %	0 %

Liens avec d'autres certifications

ANCIENNES VERSIONS

RNCP36739

CORRESPONDANCES (PAR BLOCS)

RNCP36739, RNCP36921, RNCP37172, RNCP38584, RNCP39590

Pour plus d'informations

LIEN VERS LE DESCRIPTIF

Mastère data engineering - Formation en alternance Efrei

Fiche officielle (France Compétences)

<https://www.francecompetences.fr/recherche/rncp/40875/>

Où se former

Formations référencées	14
Organismes	3
Régions couvertes	1
Prix indicatif	2 400 € – 15 000 €
Durée	40 – 525 h

Offre de formation — MonCompteFormation (Caisse des Dépôts), licence ouverte.

Se former en apprentissage

Formations en apprentissage	1
Organismes gestionnaires (CFA)	1
Organismes formateurs	1
Régions couvertes	1
Départements couverts	1
Organismes certifiés Qualiopi	1
Sessions	de 10/2025 à 09/2028

Catalogue des formations en apprentissage — réseau des Carif-Oref. Sans frais pour l'alternant : formation financée par l'OPCO de la branche (NPEC).

Entrées et sorties de formation

Entrées en formation	12
Parcours menés à terme	0
Taux d'achèvement	0 %
Période	04/2026 – 09/2026

Entrées et sorties — Caisse des Dépôts (open data), tous organismes confondus.

Demande CPF — formations engagées

Dossiers CPF engagés	15
Montant total engagé	42 050 €
Coût moyen par dossier	2 803 €
Durée moyenne	183 h
Période	04/2026 – 08/2026
France compétences	85 %
OPCO	12 %
Reste à charge titulaire	3 %

