

Expert en ingénierie des données

RNCP36921 · RNCP · Publiée

L'essentiel

Niveau de qualification	Niveau 7
Code(s) NSF	326
Formacode(s)	31023
Code(s) ROME	M1802, M1805
Type d'enregistrement	Enregistrement sur demande
Date d'enregistrement	29/09/2022
Date d'échéance	29/09/2025

Certificateur(s)

ASSOCIATION POUR LA FORMATION A L'INFORMATIQUE ET AU NUMERIQUE · SIRET : 39350481600355

ASSOCIATION POUR LA FORMATION A L'INFORMATIQUE ET AU NUMERIQUE (AFINUM) · SIRET : 39350481600231

Partenaires habilités

24 organismes habilités, d'après la liste publiée par France compétences : 24 à former et organiser l'évaluation.

Organisme	SIRET	Habilitation
ASSOCIATION POUR LA FORMATION A L'INFORMATIQUE ET AU NUMERIQUE	39350481600330	Former et organiser l'évaluation
ASSOCIATION POUR LA FORMATION A L'INFORMATIQUE ET AU NUMERIQUE (AFINUM)	39350481600090	Former et organiser l'évaluation
ASSOCIATION POUR LA FORMATION A L'INFORMATIQUE ET AU NUMERIQUE (AFINUM)	39350481600215	Former et organiser l'évaluation
ASSOCIATION POUR LA FORMATION A L'INFORMATIQUE ET AU NUMERIQUE (AFINUM)	39350481600223	Former et organiser l'évaluation
ASSOCIATION POUR LA FORMATION A L'INFORMATIQUE ET AU NUMERIQUE (AFINUM)	39350481600231	Former et organiser l'évaluation
ASSOCIATION POUR LA FORMATION A L'INFORMATIQUE ET AU NUMERIQUE (AFINUM)	39350481600249	Former et organiser l'évaluation
ASSOCIATION POUR LA FORMATION A L'INFORMATIQUE ET AU NUMERIQUE (AFINUM)	39350481600272	Former et organiser l'évaluation
ASSOCIATION POUR LA FORMATION A L'INFORMATIQUE ET AU NUMERIQUE (AFINUM)	39350481600298	Former et organiser l'évaluation
ASSOCIATION POUR LA FORMATION A L'INFORMATIQUE ET AU NUMERIQUE (AFINUM)	39350481600306	Former et organiser l'évaluation
ASSOCIATION POUR LA FORMATION A L'INFORMATIQUE ET AU NUMERIQUE (AFINUM)	39350481600314	Former et organiser l'évaluation

Liste complète des 24 partenaires habilités : <https://data.aktaconseil.fr/certifications/rncp36921-expert-en-ingenierie-des-donnees/partenaires>

Résumé de la certification

OBJECTIFS ET CONTEXTE

Aujourd'hui, le volume croissant des données disponibles, leur variété sans cesse élargie et la rapidité avec laquelle elles sont créées et sont susceptibles de générer elles-mêmes de la donnée, ont fait basculer nombre d'entreprises dans l'ère du big data. Certaines entreprises s'orientent alors de plus en plus vers des organisations dites « data centric », avec la constitution d'équipe et de compétences techniques dédiées.

La clé du succès avec le big data ne réside pas dans la quantité de données collectées par une entreprise, mais dans la capacité de traitement et la manière dont elle utilise réellement ces données collectées pour créer de la valeur. Ceci a

pour conséquence de faire apparaître de nouveaux besoins métiers autour de cette data, comme celui de Data Engineer, également dénommé expert en ingénierie des données, avec des difficultés de recrutement liées notamment au volume insuffisant de candidats.

Avec cette certification, l'EPSI souhaite répondre au besoin croissant en compétences liées à l'ingénierie des données, et proposer aux entreprises des profils de certifiés d'un niveau technique informatique élevé et dotés d'une expertise pointue en data.

ACTIVITÉS VISÉES

Analyse et définition de la stratégie Big Data alignée avec la stratégie « business » de l'entreprise, en collaboration avec la DSI et les experts métiers : Veille technologique sur les outils Big data (solutions matérielles et logicielles) / Analyse du besoin Data / Définition d'une stratégie Big Data

Pilotage d'un projet de développement d'une plateforme Big Data (architecture et infrastructure) dans un environnement agile : Mise en œuvre du projet / Suivi des réalisations du projet et management d'équipe

Construction d'une plateforme Big Data permettant la collecte, l'assemblage, le traitement et le stockage des données générées par les systèmes d'une entreprise : Conception de l'architecture d'une plateforme Big Data / Mise en œuvre technique et déploiement d'une plateforme Big Data

Déploiement, administration et supervision d'une plateforme Big Data : Déploiement d'une plateforme Big Data / Administration et supervision d'une plateforme Big Data

Préparation et mise à disposition des données d'une plateforme Big Data aux équipes utilisatrices (data scientist, équipe B.I., décideurs et experts métiers) : Structuration de données (définition de modèles) / Qualité et sécurisation des données / Mise à disposition des données

COMPÉTENCES ATTESTÉES

Mettre en place un dispositif de veille technologique en français et en anglais sur les outils, méthodes et solutions big data afin d'anticiper les technologies émergentes concernant les données (data driven, intelligence artificielle, quantique...) dans l'objectif de restituer les résultats aux décideurs

Collecter les besoins Data des directions métiers de l'entreprise afin de bâtir ou faire évoluer la/les plateformes de données dans le respect des besoins et attentes des directions métiers et de vérifier l'alignement stratégique

Analyser la stratégie de l'entreprise en étudiant son environnement et son fonctionnement afin de pouvoir établir un diagnostic de son système de gouvernance de données

Cartographier le système de gouvernance de données en traçant l'ensemble des sources et composantes de chaque flux de données afin d'analyser le patrimoine de données de l'entreprise et d'avoir une bonne connaissance dans son ensemble des systèmes de collecte, stockage et traitement de données

Identifier les informations sensibles, les risques, les zones critiques et les chemins d'attaque possibles du système de gouvernance de données existant à l'aide de la cartographie afin de contribuer à définir une politique de sécurité S.I. avec le/la RSSI

Elaborer la stratégie Big Data en collaboration avec la DSI en analysant les objectifs et la stratégie générale afin de proposer un projet de création ou des projets d'évolution, d'adaptation ou de migration du système de gouvernance de données en accord avec le développement de l'entreprise

Contribuer à définir une politique de sécurisation des données efficace en termes de protection contre le piratage et la perte de données en adéquation avec la sécurité S.I. mise en place par le/la RSSI et en collaboration avec la DSI de l'entreprise

Proposer des projets d'évolution du système de gouvernance des données au DSI et aux décideurs en les priorisant afin d'accompagner le développement de l'entreprise en cohérence avec la stratégie définie

Définir des indicateurs clés de performance selon une méthode spécifique (SLA, TRS...) afin de concevoir des tableaux de bord d'analyse de performance d'un système de gouvernance de données et déceler des pistes

éventuelles d'amélioration

Définir des indicateurs clés de performance opérationnelle selon une approche Lean afin d'élaborer des propositions d'amélioration continue du système de gouvernance de données en termes de coûts informatiques et de satisfaction utilisateurs

Définir des indicateurs clés Green IT et liés à l'Informatique responsable afin d'évaluer les impacts RSE des ressources informatiques du système de gouvernance de données et d'élaborer des propositions d'amélioration continue du système

Identifier l'ensemble des étapes de réalisation du développement d'une plateforme Big Data pour organiser le projet en tâches et livrables en répartissant les activités en fonction des ressources humaines, techniques et financières à mobiliser

Concevoir les cahiers des charges technique et fonctionnel d'un projet de développement d'une plateforme Big Data à l'aide des besoins utilisateurs collectés afin de cadrer le développement

Gérer un projet agile en utilisant les méthodes et outils adaptés à ce mode de fonctionnement pour tester, modifier et procéder par itération afin de réduire les délais de remise du projet de développement d'une plateforme Big Data

Etablir des tableaux de bord de suivi de performance (qualitative et quantitative) de l'ensemble des ressources allouées à chaque étape-projet pour anticiper, visualiser et corriger les écarts en temps réel afin de limiter les contraintes de ressources et les retards dans la réalisation du projet

Piloter les prestataires extérieurs éventuels gérant les ressources informatiques d'une plateforme de données existante listées dans la cartographie établie afin de sécuriser la mise en œuvre technique

Conduire une équipe projet en diffusant les fondamentaux de l'agilité : adaptation, flexibilité et amélioration continue au sein de l'équipe afin d'être en mesure d'absorber les changements de priorité qui peuvent intervenir dans un contexte de forte contrainte de temps et d'incertitudes

Adopter une stratégie d'accueil aux handicaps afin de favoriser l'inclusion des profils en situation de handicap au sein de l'équipe et permettre leur pleine intégration, en collaboration avec le référent handicap de l'entreprise

Concevoir un processus de communication inclusif régulier au sein de l'équipe afin de synchroniser les activités quotidiennes et mettre en place un fil de discussion à l'aide d'outils numériques

Animer des réunions à distance afin de maintenir une dynamique de groupe et renforcer l'esprit d'équipe des membres en télétravail et/ou à distance

Concevoir un processus de partage d'information afin de faciliter l'inclusion et la collaboration entre les membres en télétravail et/ou à distance en utilisant des outils numériques

Concevoir une architecture de collecte et de restitution de données robuste, évolutive, sécurisée et utilisant l'intelligence artificielle (machine learning) afin d'améliorer en continu sa capacité à prédire les besoins Data des experts métiers utilisateurs

Concevoir une architecture de stockage de données (data warehouse, data lake...) permettant de répondre aux besoins Data des experts métiers et respectant la politique de sécurité des données définie par le/la RSSI

Définir les processus de collecte et d'intégration de données par lot ou en streaming afin de favoriser la vitesse d'intégration et la volumétrie de données dans le respect de l'architecture définie

Mettre en place un système d'ingestion de données structurées et non structurées afin de permettre la manipulation et l'accès aux données ainsi que l'authentification des utilisateurs

Développer une solution de migration inter-systèmes et multi-environnements à l'aide d'un outil de son choix afin de permettre l'intégration de données diverses et l'interopérabilité des différentes sources de données

Développer un pipeline de données et/ou un pipeline ETL prenant en compte l'environnement technologique déployé (infrastructure, services, applications...) dans le respect du cahier des charges de la solution proposée

Créer un lac de données (data lake) afin de collecter des données brutes dans le respect de l'architecture de collecte des données définie dans la solution proposée

Créer un entrepôt unique à partir du référentiel de données établi pour centraliser les informations stratégiques de l'entreprise et répondre rapidement aux besoins métiers

Déployer un processus de collecte, stockage et traitement de données selon une approche ETL (Extract-Transform-Load) ou une approche ELT afin de permettre l'extraction, le stockage et le traitement des données de manière optimale et adaptée aux besoins utilisateurs métiers

Déployer une plateforme de stockage de données, dans un environnement distribué à travers des clusters de nœuds (serveurs machine) afin d'y stocker les données de façon distribuée et de les traiter selon le principe du calcul parallèle

Configurer des clusters de nœuds avec une solution de stockage distribuée afin d'assurer une tolérance de zéro panne

Automatiser des opérations de déploiement, de tests et de maintenance curative et préventive afin de s'assurer de la fiabilité de la solution tout au long de son cycle de vie

Concevoir un ensemble de tests de surveillance des pipelines afin d'anticiper et d'organiser l'évaluation et le dépannage des représentations de données

Administrer la plateforme de stockage de données déployée dans un environnement technologique optimal afin de répondre aux trois critères sine qua non tels que la performance, la capacité et la bande passante de la plateforme

Dimensionner en temps réel les besoins en consommation de ressources en mettant en place l'autoscaling des ressources afin d'anticiper les montées en charge et assurer la fiabilité de la solution

Contrôler la bonne application de la politique des données en analysant, traitant et reportant les coûts et les performances de stockage selon les différents critères (licences, supports, évolutivité, performance...) afin de s'assurer de la pérennité de la solution

Rédiger un protocole de maintenance de la documentation technique afin de consigner et assurer la préservation de la solution choisie

Définir les données de référence de l'entreprise à partir des données utilisées pour créer un référentiel de données afin d'assurer la mise à disposition de données cohérentes aux directions métiers

Développer la phase de préparation des données afin de permettre leur chargement en prenant en compte les étapes de découverte, structuration, nettoyage, validation et intégration décrite dans le cahier des charges afin de les rendre consommables par les utilisateurs métiers

Assurer la qualité des données en utilisant les outils de gestion de la qualité de données pour garantir l'exactitude, la cohérence, la synchronisation et la traçabilité des données afin de satisfaire les besoins d'accessibilité des utilisateurs métiers

Appliquer les procédures de sécurité établies par le / la RSSI de l'entreprise afin d'assurer la confidentialité et la sécurité des données et garantir une mise en conformité avec les obligations légales du RGPD

Manipuler les différents services et fonctionnalités des architectures de type Data Lake afin de gérer le cycle de vie des données DLM (Data Life cycle Management)

Proposer des modèles statistiques et de data science (machine learning) à mettre en pratique aux directions métiers afin de détecter des nouveaux services, anticiper des besoins et résoudre des problématiques métiers de l'entreprise

Organiser les sources de données sous forme de résultats exploitables (data visualisation) pour alimenter les outils décisionnels et visualiser les résultats de façon compréhensible permettant d'aider les directions métiers à la prise de décision

Développer divers services de stockage, de gestion de systèmes de bases de données, de production d'ensemble de données nettoyées et améliorées pour l'analyse grâce aux langages adaptés et répondant aux besoins afin de mettre en place l'exploitation de données par les différents métiers

Blocs de compétences

RNCP36921BC01 — Analyse et définition de la stratégie Big Data alignée avec la stratégie « business » de l'entreprise en collaboration avec la DSI et les experts métiers

LISTE DE COMPÉTENCES

Mettre en place un dispositif de veille technologique en français et en anglais sur les outils, méthodes et solutions big data afin d'anticiper les technologies émergentes concernant les données (data driven, intelligence artificielle, quantique...) dans l'objectif de restituer les résultats aux décideurs

Collecter les besoins Data des directions métiers de l'entreprise afin de bâtir ou faire évoluer la/les plateformes de données dans le respect des besoins et attentes des directions métiers et de vérifier l'alignement stratégique

Analyser la stratégie de l'entreprise en étudiant son environnement et son fonctionnement afin de pouvoir établir un diagnostic de son système de gouvernance de données

Cartographier le système de gouvernance de données en traçant l'ensemble des sources et composantes de chaque flux de données afin d'analyser le patrimoine de données de l'entreprise et d'avoir une bonne connaissance dans son ensemble des systèmes de collecte, stockage et traitement de données

Identifier les informations sensibles, les risques, les zones critiques et les chemins d'attaque possibles du système de gouvernance de données existant à l'aide de la cartographie afin de contribuer à définir une politique de sécurité S.I. avec le/la RSSI

Elaborer la stratégie Big Data en collaboration avec la DSI en analysant les objectifs et la stratégie générale afin de proposer un projet de création ou des projets d'évolution, d'adaptation ou de migration du système de gouvernance

de données en accord avec le développement de l'entreprise

Contribuer à définir une politique de sécurisation des données efficace en termes de protection contre le piratage et la perte de données en adéquation avec la sécurité S.I. mise en place par le/la RSSI et en collaboration avec la DSI de l'entreprise

Proposer des projets d'évolution du système de gouvernance des données au DSI et aux décideurs en les priorisant afin d'accompagner le développement de l'entreprise en cohérence avec la stratégie définie

Définir des indicateurs clés de performance selon une méthode spécifique (SLA, TRS...) afin de concevoir des tableaux de bord d'analyse de performance d'un système de gouvernance de données et déceler des pistes éventuelles d'amélioration

Définir des indicateurs clés de performance opérationnelle selon une approche Lean afin d'élaborer des propositions d'amélioration continue du système de gouvernance de données en termes de coûts informatiques et de satisfaction utilisateurs

Définir des indicateurs clés Green IT et liés à l'Informatique responsable afin d'évaluer les impacts RSE des ressources informatiques du système de gouvernance de données et d'élaborer des propositions d'amélioration continue du système

MODALITÉS D'ÉVALUATION

Mise en situation professionnelle reconstituée (MSPR) / Elaboration d'une stratégie Big Data à partir d'une situation réelle ou reconstituée proposée par le certificateur

Phase I- Préparation tutorée de la MSPR par équipe de 3 max. / Durée de préparation 28 heures

Phase II- Production écrite individuelle à réaliser

Phase III- Soutenance orale devant un jury de professionnel / Durée : 45 mn - 15 min de présentation de la réalisation en groupe + 10 min de démonstration technique et entretien individuel (questions en lien avec les compétences)

RNCP36921BC02 — Pilotage d'un projet de développement d'une plateforme Big Data (architecture et infrastructure) dans un environnement agile

LISTE DE COMPÉTENCES

Identifier l'ensemble des étapes de réalisation du développement d'une plateforme Big Data pour organiser le projet en tâches et livrables en répartissant les activités en fonction des ressources humaines, techniques et financières à mobiliser

Concevoir les cahiers des charges technique et fonctionnel d'un projet de développement d'une plateforme Big Data à l'aide des besoins utilisateurs collectés afin de cadrer le développement

Gérer un projet agile en utilisant les méthodes et outils adaptés à ce mode de fonctionnement pour tester, modifier et procéder par itération afin de réduire les délais de remise du projet de développement d'une plateforme Big Data

Etablir des tableaux de bord de suivi de performance (qualitative et quantitative) de l'ensemble des ressources allouées à chaque étape-projet pour anticiper, visualiser et corriger les écarts en temps réel afin de limiter les contraintes de ressources et les retards dans la réalisation du projet

Piloter les prestataires extérieurs éventuels gérant les ressources informatiques d'une plateforme de données existante listées dans la cartographie établie afin de sécuriser la mise en œuvre technique

Conduire une équipe projet en diffusant les fondamentaux de l'agilité : adaptation, flexibilité et amélioration continue au sein de l'équipe afin d'être en mesure d'absorber les changements de priorité qui peuvent intervenir dans un contexte de forte contrainte de temps et d'incertitudes

Adopter une stratégie d'accueil aux handicaps afin de favoriser l'inclusion des profils en situation de handicap au sein de l'équipe et permettre leur pleine intégration, en collaboration avec le référent handicap de l'entreprise

Concevoir un processus de communication inclusif régulier au sein de l'équipe afin de synchroniser les activités quotidiennes et mettre en place un fil de discussion à l'aide d'outils numériques

Animer des réunions à distance afin de maintenir une dynamique de groupe et renforcer l'esprit d'équipe des membres en télétravail et/ou à distance

Concevoir un processus de partage d'information afin de faciliter l'inclusion et la collaboration entre les membres en télétravail et/ou à distance en utilisant des outils numériques

MODALITÉS D'ÉVALUATION

Mise en situation professionnelle reconstituée (MSPR) - Gestion d'un projet de développement de l'architecture et l'infrastructure d'une plateforme Big Data à partir d'une situation réelle ou reconstituée proposée par le certificateur

Phase I- Préparation tutorée de la MSPR par équipe de 3 max. / Durée de préparation 21 heures

Phase II- Production écrite individuelle à réaliser

Phase III- Soutenance orale devant un jury de professionnel Durée : 45 mn / 15 min de présentation de la réalisation

en groupe + 10 min de démonstration technique et entretien individuel (questions en lien avec les compétences)

RNCP36921BC03 — Construction d'une plateforme Big Data permettant la collecte, l'assemblage, le traitement et le stockage des données générées par les systèmes d'une entreprise

LISTE DE COMPÉTENCES

Concevoir une architecture de collecte et de restitution de données robuste, évolutive, sécurisée et utilisant l'intelligence artificielle (machine learning) afin d'améliorer en continu sa capacité à prédire les besoins Data des experts métiers utilisateurs

Concevoir une architecture de stockage de données (data warehouse, data lake...) permettant de répondre aux besoins Data des experts métiers et respectant la politique de sécurité des données définie par le/la RSSI

Définir les processus de collecte et d'intégration de données par lot ou en streaming afin de favoriser la vitesse d'intégration et la volumétrie de données dans le respect de l'architecture définie

Mettre en place un système d'ingestion de données structurées et non structurées afin de permettre la manipulation et l'accès aux données ainsi que l'authentification des utilisateurs

Développer une solution de migration inter-systèmes et multi-environnements à l'aide d'un outil de son choix afin de permettre l'intégration de données diverses et l'interopérabilité des différentes sources de données

Développer un pipeline de données et/ou un pipeline ETL prenant en compte l'environnement technologique déployé (infrastructure, services, applications...) dans le respect du cahier des charges de la solution proposée

Créer un lac de données (data lake) afin de collecter des données brutes dans le respect de l'architecture de collecte des données définie dans la solution proposée

Créer un entrepôt unique à partir du référentiel de données établi pour centraliser les informations stratégiques de l'entreprise et répondre rapidement aux besoins métiers

Déployer un processus de collecte, stockage et traitement de données selon une approche ETL (Extract-Transform-Load) ou une approche ELT afin de permettre l'extraction, le stockage et le traitement des données de manière optimale et adaptée aux besoins utilisateurs métiers

MODALITÉS D'ÉVALUATION

Mise en situation professionnelle reconstituée (MSPR) - Conception de l'architecture et construction de l'infrastructure d'une plateforme Big Data sécurisée à partir d'une situation réelle ou reconstituée proposée par le certificateur

Phase I- Préparation tutorée de la MSPR par équipe de 3 max. / Durée de préparation 35 heures

Phase II- Production écrite individuelle à réaliser

Phase III- Soutenance orale devant un jury de professionnel

/ Durée : 45 mn

: 15 min de présentation de la réalisation en groupe + 10 min de démonstration technique et entretien individuel (questions en lien avec les compétences)

RNCP36921BC04 — Administration et supervision d'une plateforme Big Data

LISTE DE COMPÉTENCES

Déployer une plateforme de stockage de données, dans un environnement distribué à travers des clusters de nœuds (serveurs machine) afin d'y stocker les données de façon distribuée et de les traiter selon le principe du calcul parallèle

Configurer des clusters de nœuds avec une solution de stockage distribuée afin d'assurer une tolérance de zéro panne

Automatiser des opérations de déploiement, de tests et de maintenance curative et préventive afin de s'assurer de la fiabilité de la solution tout au long de son cycle de vie

Concevoir un ensemble de tests de surveillance des pipelines afin d'anticiper et d'organiser l'évaluation et le dépannage des représentations de données

Administrer la plateforme de stockage de données déployée dans un environnement technologique optimal afin de répondre aux trois critères sine qua non tels que la performance, la capacité et la bande passante de la plateforme

Dimensionner en temps réel les besoins en consommation de ressources en mettant en place l'autoscaling des ressources afin d'anticiper les montées en charge et assurer la fiabilité de la solution

Contrôler la bonne application de la politique des données en analysant, traitant et reportant les coûts et les performances de stockage selon les différents critères (licences, supports, évolutivité, performance) afin de s'assurer de la pérennité de la solution

Rédiger un protocole de maintenance de la documentation technique afin de consigner et assurer la préservation de la solution choisie

MODALITÉS D'ÉVALUATION

Mise en situation professionnelle reconstituée (MSPR) / Administration et supervision d'une plateforme Big Data sécurisée à partir d'une situation réelle ou reconstituée proposée par le certificateur

Phase I- Préparation tutorée de la MSPR par équipe de 3 max. / Durée de préparation 28 heures

Phase II- Production écrite individuelle à réaliser

Phase III- Soutenance orale devant un jury de professionnel

/ Durée : 45 mn / 15 min de présentation de la réalisation en groupe + 10 min de démonstration technique et entretien individuel (questions en lien avec les compétences)

RNCP36921BC05 — Préparation et Mise à disposition des données d'une plateforme Big Data aux équipes utilisatrices (data scientist, équipe B.I., décideurs et experts métiers)

LISTE DE COMPÉTENCES

Définir les données de référence de l'entreprise à partir des données utilisées pour créer un référentiel de données afin d'assurer la mise à disposition de données cohérentes aux directions métiers

Développer la phase de préparation des données afin de permettre leur chargement en prenant en compte les étapes de découverte, structuration, nettoyage, validation et intégration décrite dans le cahier des charges afin de les rendre consommables par les utilisateurs métiers

Assurer la qualité des données en utilisant les outils de gestion de la qualité de données pour garantir l'exactitude, la cohérence, la synchronisation et la traçabilité des données afin de satisfaire les besoins d'accessibilité des utilisateurs métiers

Appliquer les procédures de sécurité établies par le / la RSSI de l'entreprise afin d'assurer la confidentialité et la sécurité des données et garantir une mise en conformité avec les obligations légales du RGPD

Manipuler les différents services et fonctionnalités des architectures de type Data Lake afin de gérer le cycle de vie des données DLM (Data Life cycle Management)

Proposer des modèles statistiques et de data science (machine learning) à mettre en pratique aux directions métiers afin de détecter des nouveaux services, anticiper des besoins et résoudre des problématiques métiers de l'entreprise

Organiser les sources de données sous forme de résultats exploitables (data visualisation) pour alimenter les outils décisionnels et visualiser les résultats de façon compréhensible permettant d'aider les directions métiers à la prise de décision

Développer divers services de stockage, de gestion de systèmes de bases de données, de production d'ensemble de données nettoyées et améliorées pour l'analyse grâce aux langages adaptés et répondant aux besoins afin de mettre en place l'exploitation de données par les différents métiers

MODALITÉS D'ÉVALUATION

Mise en situation professionnelle reconstituée (MSPR) / Exploitation d'une plateforme Big Data à partir d'une situation réelle ou reconstituée proposée par le certificateur

Phase I- Préparation tutorée de la MSPR par équipe de 3 max. / Durée de préparation 28 heures

Phase II- Production écrite individuelle à réaliser

Phase III- Soutenance orale devant un jury de professionnel

/ Durée : 45 mn : 15 min de présentation de la réalisation en groupe + 10 min de démonstration technique et entretien individuel (questions en lien avec les compétences)

Secteur d'activité et type d'emploi

SECTEURS D'ACTIVITÉS

Avec l'omniprésence de la data, tous les secteurs d'activités de l'industrie et des services utilisant de la donnée et tous types de structures sont concernés. L'expert en ingénierie des données peut donc exercer ses fonctions :

- Au sein d'un pôle informatique où il intègre le service et l'équipe data, en tant que salarié dans une ETI (Entreprise de Taille Intermédiaire) ou grande entreprise, dans le secteur privé mais également dans le secteur public.
- En tant que consultant sous statut indépendant ou salarié d'une entreprise de service, intervenant auprès d'entreprises de toutes tailles et notamment pour les TPE et PME, quel que soit leur domaine d'activité.

TYPE D'EMPLOIS ACCESSIBLES

Expert en ingénierie des données - Data Engineer / Ingénieur - Ingénieur de données - Ingénieur Big Data

RÉGLEMENTATIONS D'ACTIVITÉ

Le métier n'est pas réglementé, mais l'expert en ingénierie des données exerce ses activités dans le respect des

contraintes réglementaires liées à son domaine d'intervention, le cas échéant en lien avec d'autres professionnels comme le RSSI.

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)	2 942 €
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) M2X90 — Ingénieurs et cadres d'étude, recherche et développement en informatique et télécom · tension 5/5 (Très élevée) M2X93 — Experts et consultants en systèmes d'information · tension 5/5 (Très élevée)
Projets de recrutement (BMO 2026)	32 216
Jugés difficiles à recruter	51 %
Emplois saisonniers	1 %
Familles professionnelles (FAP)	M1X80 — Techniciens d'étude et de développement en informatique M2X90 — Ingénieurs et cadres d'étude, recherche et développement en informatique et télécom M2X93 — Experts et consultants en systèmes d'information

Voies d'accès

PRÉREQUIS À L'ENTRÉE EN FORMATION

Être titulaire d'une certification professionnelle de niveau 6 en Informatique /dans le domaine visé, ou être titulaire d'une certification professionnelle de niveau 5 en Informatique / dans le domaine visé avec une expérience professionnelle d'au moins un an dans les métiers informatiques (infrastructure, développement d'applications, data). Dans le cas où un candidat ne disposerait pas des prérequis définis, il a la possibilité de déposer un dossier de candidature qui sera examiné par une commission réunissant le responsable pédagogique EPSI et le responsable en charge du recrutement des candidats, en vue d'une admission « exceptionnelle ».

Sélection sur dossier de candidature avec CV et lettre de motivation, tests de positionnement et entretien individuel de sélection .

PRÉREQUIS À LA VALIDATION

Chaque bloc de compétences fait l'objet d'une épreuve et peut ainsi être validé indépendamment. Toutefois, l'acquisition de l'ensemble des blocs de compétences ne vaut pas acquisition de la certification pour laquelle une évaluation finale est requise (dossier professionnel) et permet au jury de certification de vérifier la réalité de l'acquisition de l'ensemble du spectre de compétences nécessaires et la compréhension du métier visé.

Dans le cadre de l'accès à la certification par V.A.E., les compétences sont évaluées au travers d'un dossier de validation et d'un entretien devant le jury de certification.

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

La décision d'attribution de la certification est du ressort d'une commission nationale de certification, dit « jury de certification » composé de 2 professionnels extérieurs et du métier visé (employeur ou salarié dans le domaine de l'informatique, développeur applicatif en intelligence artificielle, et non intervenant au sein du parcours de formation

préparant à la certification), et de 2 membres internes à l'organisme.

En amont, intervient un jury de soutenance est composé de trois membres : un membre interne à l'organisme certificateur (de la direction Innovation & Pédagogie, ou le directeur du centre de formation ou son représentant), et deux professionnels experts du métier visé et externes à l'organisme de formation. Le président du jury de soutenance est un professionnel extérieur à l'organisme de formation et est désigné par l'ensemble des membres du jury. Ce jury a pour attribution de proposer au jury de certification la validation ou non d'un bloc.

COMPOSITION DU JURY — VAE

Le jury de certification est composé de 2 professionnels extérieurs et du métier visé (employeur ou salarié dans le domaine de l'informatique, développeur applicatif en intelligence artificielle) et de 2 membres internes à l'organisme

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	8 894 €
Selon la branche	7 697 € – 9 157 €
Prise en charge	Branches concernées
8 894 €	Valeur par défaut · 188 branches
9 157 €	CPNE Casinos, CPNE autoroutes, 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
9 005 €	CPNE FP Branche des magasins du Bricolage
8 974 €	CPNE de la branche de la Distribution Directe
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
7 697 €	CPNEFP de la blanchisserie, laverie, location de linge, nettoyage à sec, pressing et teinturerie

Base légale

Date de décision	29/09/2022
Durée de l'enregistrement	3 ans
Date d'échéance de l'enregistrement	29/09/2025

Liens avec d'autres certifications

NOUVELLE(S) CERTIFICATION(S)
RNCP42157

CORRESPONDANCES (PAR BLOCS)
RNCP35584

Pour plus d'informations

Fiche officielle (France Compétences)

<https://www.francecompetences.fr/recherche/rncp/36921/>