

Expert en science des données pour la connaissance client (MS)

RNCP40855 · RNCP · Publiée

L'essentiel

Niveau de qualification	Niveau 7
Code(s) NSF	114b, 326
Formacode(s)	31026
Code(s) ROME	M1403
Type d'enregistrement	Enregistrement sur demande
Date d'enregistrement	25/06/2025
Date d'échéance	25/06/2028

Certificateur(s)

GROUPE DES ECOLES NATIONALES D ECONOMIE ET STATISTIQUE · SIRET : 13001422800055

Partenaires habilités

1 organisme habilité, d'après la liste publiée par France compétences : 1 à former et organiser l'évaluation.

Organisme	SIRET	Habilitation
CFAI DIAFOR	39048242000028	Former et organiser l'évaluation

Résumé de la certification

OBJECTIFS ET CONTEXTE

Dès qu'une entreprise collecte un grand nombre de données (data) via les outils numériques afin de développer ses relations avec ses clients, utilisateurs et usagers, qu'elle soit de type TPE-PME ou plus grande, et dans tous les secteurs d'activité, elle a besoin des compétences d'experts en science des données pour la connaissance client (data scientists). Ces professionnels sont essentiels pour collecter, nettoyer, analyser et interpréter ces données de manière à en tirer des connaissances exploitables selon la stratégie de l'entreprise.

Le professionnel est capable de transformer les données numériques en opportunités de croissance pour les entreprises. Grâce à ses analyses, les dirigeants prennent des décisions plus éclairées, et en tirent un avantage concurrentiel. L'expert en science des données a un impact direct sur le business de l'entreprise. A cet effet, il occupe une place centrale dans l'économie numérique actuelle et est devenu indispensable dans le développement marketing et la gestion des relations clients.

ACTIVITÉS VISÉES

Activité 1 : Mise en œuvre de processus d'exploration des données et de collecte d'information :

- Analyse du patrimoine informationnel de données de l'entreprise et du système existant
- Extraction et stockage de données dans le respect de la réglementation
- Automatisation de l'extraction en mobilisant les outils de web scraping (extraction des données du web)
- Qualification des données client et pré-traitement

Activité 2 : Proposition de modèle mathématique, statistique et/ou économique permettant de représenter au mieux les relations et les structures dans les données, et approfondir la connaissance client

- Compréhension du comportement d'achat, de réceptibilité à la communication, dans le cadre plus général de la relation client

- Développement des modèles d'analyse prédictive pour la segmentation des clients, la prévision des ventes ou l'optimisation des campagnes marketing (scoring)
- Mettre en œuvre les outils et méthodes statistiques permettant le développement des services de marketing digital (système de recommandation, qualification des leads, ciblage, chatbot, personnalisation des messages)

Activité 3 : Visualisation de la donnée et interprétation des données pertinentes et utile à la connaissance client

- Production d'informations à l'intention des équipes métiers (y compris publics en situation de handicap) permettant l'évaluation, l'organisation, la prise de décision
- Identification des indicateurs utilisés en marketing, ou production d'indicateur nouveau à partir d'une recherche
- Représentations graphiques

Activité 4 : Manager les projets de l'entreprise en valorisant et optimisant l'apport des sciences de la donnée dans le respect des cadres déontologiques, éthiques et du développement durable

- Gestion et planification du projet de campagne marketing ou d'analyse de données client
- Pilotage des ressources du projet

Activité 5 : Partage de la culture de la donnée au sein de l'entreprise

:

- Valorisation de l'utilisation de la donnée
- Prévention des risques réglementaires
- Veille technologique (outil IA, data science....)

COMPÉTENCES ATTESTÉES

Explorer le besoin d'un commanditaire interne en produisant une note ou un cahier des charges précisant les questions statistiques qui seront étudiées, les sources d'informations mobilisées et l'analyse des informations produite afin de valider la bonne compréhension de la demande.

Auditer le système d'information en ayant réalisé une cartographie au préalable afin d'identifier les informations utiles à la réponse aux besoins et les possibilités d'amélioration des systèmes d'information.

Requêter le système d'information à partir des outils de requêtage des bases de données (SQL) et des API de l'entreprise pour collecter les données qui seront utilisés pour répondre à la demande du commanditaire.

Collecter des données sur le web en mobilisant les outils de web scraping afin d'automatiser les extractions permettant de répondre à la demande d'une part et d'autre part d'enrichir le système d'information de façon pérenne.

Prétraiter les données collectées en procédant à des vérifications statistiques (qualité des données, valeurs manquantes, aberrantes, complétude des données) et en réalisant le cas échéant des recoupements afin de qualifier les caractéristiques des données et aider à la documentation des données.

Documenter les données extraites en produisant l'information nécessaire à son utilisation (sources, traitements réalisés, vigilance sur les biais d'observation, de sélection, les valeurs manquantes, aberrantes....) dans un objectif de conservation et de réutilisation et dans le respect du cadre déontologique

Réaliser des analyses de données en mettant en œuvre les procédures statistiques descriptives et exploratoires afin de décrire et qualifier les différentes données caractérisant les clients et leurs comportements, détecter les incohérences dans les données

et comprendre les corrélations entre les variables.

Modéliser un comportement donné d'un client ciblé mettant en œuvre les modèles de régression y compris dans les cas complexes (grand nombre de variables, multi colinéarité) afin d'identifier les relations entre les variables et mieux comprendre les actes de consommation du client.

Prédire un comportement donné d'un client cible en utilisant les modèles supervisés non paramétriques et en s'appuyant sur les méthodes du machine learning (arbres de décision, forêts aléatoires, méthodes à noyaux, machines à vecteur de support avec noyaux non linéaires) afin d'optimiser la connaissance et la relation client

Traiter les données textuelles recueillies auprès des clients avec les outils du NLP (traitement automatique du langage naturel) afin de réaliser la classification de texte, l'extraction d'informations, et permettre la mise en œuvre des méthodes dite d'analyse de sentiments permettant la compréhension du client.

Modéliser les évolutions d'un comportement client, en fonction du temps et à partir de l'analyse des séries chronologiques (en mettant en œuvre des modèles ARIMA, SARIMA, Holt-Winters, Réseaux de neurones récurrents) afin de faire des prévisions de valeurs futures en fonction de certains déterminants, en particulier ceux sur lesquels une action est possible

Produire les modèles adaptés aux données déséquilibrées et aux valeurs manquantes en paramétrant les procédures statistiques utilisés afin d'obtenir une meilleure fiabilité des informations produites.

Produire des indicateurs répondant aux problématiques métiers en précisant leurs propriétés métrologiques et leurs

limites d'interprétation afin de répondre à la demande des directions métiers et d'accompagner la prise de décision. Réaliser les représentations graphiques les plus pertinentes possibles en vérifiant les règles de la sémiologie graphique et en déterminant les formes graphiques les plus adaptées afin de rendre accessible la compréhension des données au client de façon claire et pédagogique, et ce en intégrant la prise en compte des handicaps visuels et en utilisant les outils de datavisualisation.

Concevoir des dashboard (tableaux de bord) en interaction avec les équipes en ayant documenté les données en amont et en s'assurant de leur automatisation possible afin de compiler les informations essentielles sur les moyens, le suivi de l'activité et la performance et ce, en tenant compte des ressources de l'entreprise. (data notamment)

Construire un discours narratif permettant de susciter l'intérêt de son auditoire afin de mettre en avant les principaux messages de son analyse tout en favorisant sa mémorisation et ce, en s'appuyant sur les formes produites.

Définir le cadre de gestion de projet le plus adapté en tenant compte des contraintes de délais, de ressources et d'objectifs afin de cadrer son projet et gagner en efficience pendant son déploiement

Piloter un projet dans le respect du cadre donné par une feuille de route, incluant la bonne gestion des ressources humaines du projet dans le respect des besoins individuels (notamment en cas de situation de handicap),

le respect des différents jalons afin de s'assurer de la bonne qualité du livrable et du respect de la demande du client et le cas échéant, du cahier des charges.

Diffuser l'usage des nouvelles méthodes d'exploitation des données notamment IA et leur cadre éthique auprès des directions commerciale et marketing en s'appuyant sur une veille active, et en apportant des conseils sur les processus et procédures à mettre en place à tous les stades de l'utilisation des données afin de développer des bonnes pratiques et garantir le respect de la vie privée, la confidentialité des données et le respect des règles de sécurité de l'information.

Partager ses besoins et ses contraintes, ses objectifs, sa culture métier dans le cadre de dispositifs collectifs existant ou le cas échéant, créés à son initiative afin de mieux comprendre les problématiques rencontrées par les différents professionnels des métiers connexes dans l'entreprise autour du traitement de l'information et construire un écosystème de la data plus efficace et rendre plus fluide la collaboration entre les services et métiers

Concevoir un processus structuré de veille technologique et réglementaire en réalisant une sélection d'outils performants par rapport à l'objet des recherches à effectuer afin de s'assurer du maintien de son efficacité opérationnelle et de ces compétences technologiques (notamment en matière d'IA, outils datascience) et assurer le respect du cadre déontologique et éthique.

Réaliser des notes explicatives à l'attention des parties prenantes et lecteurs non data scientists permettant une prise de décision éclairée à partir d'une bonne compréhension des résultats et de leur limites (interprétabilité, intelligibilité) et tenant compte d'un contexte multiculturel et/ou international ou de la présence d'un éventuel profil en situation de handicap.

Blocs de compétences

RNCP40855BC01 — Collecter des données ou des informations

LISTE DE COMPÉTENCES

Explorer le besoin d'un commanditaire interne en produisant une note ou un cahier des charges précisant les questions statistiques qui seront étudiées, les sources d'informations mobilisées et l'analyse des informations produite afin de valider la bonne compréhension de la demande.

Auditer le système d'information en ayant réalisé une cartographie au préalable afin d'identifier les informations utiles à la réponse aux besoins et les possibilités d'amélioration des systèmes d'information.

Requêter le système d'information à partir des outils de requêtage des bases de données (SQL) et des API de l'entreprise pour collecter les données qui seront utilisées pour répondre à la demande du commanditaire.

Collecter des données sur le web en mobilisant les outils de web scraping afin d'automatiser les extractions permettant de répondre à la demande d'une part et d'autre part d'enrichir le système d'information de façon pérenne.

Prétraiter les données collectées en procédant à des vérifications statistiques (qualité des données, valeurs manquantes, aberrantes, complétude des données) et en réalisant le cas échéant des recoupements afin de qualifier les caractéristiques des données et aider à la documentation des données.

Documenter les données extraites en produisant l'information nécessaire à son utilisation (sources, traitements réalisés, vigilance sur les biais d'observation, de sélection, les valeurs manquantes, aberrantes....) dans un objectif de conservation et de réutilisation et dans le respect du cadre déontologique.

MODALITÉS D'ÉVALUATION

RNCP40855BC02 — Réaliser une modélisation ou une analyse prédictive à partir d'un modèle d'apprentissage avec des données structurées ou non structurées

LISTE DE COMPÉTENCES

Réaliser des analyses de données en mettant en œuvre les procédures statistiques descriptives et exploratoires afin de décrire et qualifier les différentes données caractérisant les clients et leurs comportements, détecter les incohérences dans les données

et comprendre les corrélations entre les variables.

Modéliser un comportement donné d'un client ciblé mettant en œuvre les modèles de régression y compris dans les cas complexes (grand nombre de variables, multi colinéarité) afin d'identifier les relations entre les variables et mieux comprendre les actes de consommation du client.

Prédire un comportement donné d'un client cible en utilisant les modèles supervisés non paramétriques et en s'appuyant sur les méthodes du machine learning (arbres de décision, forêts aléatoires, méthodes à noyaux, machines à vecteur de support avec noyaux non linéaires) afin d'optimiser la connaissance et la relation client.

Traiter les données textuelles recueillies auprès des clients avec les outils du NLP (traitement automatique du langage naturel) afin de réaliser la classification de texte, l'extraction d'informations, et permettre la mise en œuvre des méthodes dites d'analyse de sentiments permettant la compréhension du client.

Modéliser les évolutions d'un comportement client, en fonction du temps et à partir de l'analyse des séries chronologiques (en mettant en œuvre des modèles ARIMA, SARIMA, Holt-Winters, Réseaux de neurones récurrents) afin de faire des prévisions de valeurs futures en fonction de certains déterminants, en particulier ceux sur lesquels une action est possible.

Produire les modèles adaptés aux données déséquilibrées et aux valeurs manquantes en paramétrant les procédures statistiques utilisées afin d'obtenir une meilleure fiabilité des informations produites.

MODALITÉS D'ÉVALUATION

Rapport de méthodologie

RNCP40855BC03 — Formaliser et visualiser les résultats des traitements statistiques

LISTE DE COMPÉTENCES

Produire des indicateurs répondant aux problématiques métiers en précisant leurs propriétés métrologiques et leurs limites d'interprétation afin de répondre à la demande des directions métiers et d'accompagner la prise de décision.

Réaliser les représentations graphiques les plus pertinentes possibles en vérifiant les règles de la sémiologie graphique et en déterminant les formes graphiques les plus adaptées afin de rendre accessible la compréhension des données au client de façon claire et pédagogique, et ce en intégrant la prise en compte des handicaps visuels et en utilisant les outils de datavisualisation.

Concevoir des dashboard (tableaux de bord) en interaction avec les équipes en ayant documenté les données en amont et en s'assurant de leur automatisation possible afin de compiler les informations essentielles sur les moyens, le suivi de l'activité et la performance et ce, en tenant compte des ressources de l'entreprise. (data notamment)

Construire un discours narratif permettant de susciter l'intérêt de son auditoire afin de mettre en avant les principaux messages de son analyse tout en favorisant sa mémorisation et ce, en s'appuyant sur les formes produites.

MODALITÉS D'ÉVALUATION

Thèse professionnelle

RNCP40855BC04 — Participer à la mise en œuvre de projets de l'entreprise

LISTE DE COMPÉTENCES

Définir le cadre de gestion de projet le plus adapté en tenant compte des contraintes de délais, de ressources et d'objectifs afin de cadrer son projet et gagner en efficacité pendant son déploiement.

Piloter un projet dans le respect du cadre donné par une feuille de route, incluant la bonne gestion des ressources humaines du projet dans le respect des besoins individuels (notamment en cas de situation de handicap),

le respect des différents jalons afin de s'assurer de la bonne qualité du livrable et du respect de la demande du client et le cas échéant, du cahier des charges.

Diffuser l'usage des nouvelles méthodes d'exploitation des données, notamment l'IA, et leur cadre éthique auprès des directions commerciale et marketing en s'appuyant sur une veille active, et en apportant des conseils sur les processus et procédures à mettre en place à tous les stades de l'utilisation des données afin de développer des bonnes pratiques et garantir le respect de la vie privée, la confidentialité des données et le respect des règles de

sécurité de l'information.

Partager ses besoins et ses contraintes, ses objectifs, sa culture métier dans le cadre de dispositifs collectifs existant ou, le cas échéant, créés à son initiative afin de mieux comprendre les problématiques rencontrées par les différents professionnels des métiers connexes dans l'entreprise autour du traitement de l'information et construire un écosystème de la data plus efficace et rendre plus fluide la collaboration entre les services et métiers.

Concevoir un processus structuré de veille technologique et réglementaire en réalisant une sélection d'outils performants par rapport à l'objet des recherches à effectuer afin de s'assurer du maintien de son efficacité opérationnelle et de ses compétences technologiques (notamment en matière d'IA, outils datascience) et assurer le respect du cadre déontologique et éthique.

Réaliser des notes explicatives à l'attention des parties prenantes et lecteurs non data scientists permettant une prise de décision éclairée à partir d'une bonne compréhension des résultats et de leur limites (interprétabilité, intelligibilité) et tenant compte d'un contexte multiculturel et/ou international ou de la présence d'un éventuel profil en situation de handicap.

MODALITÉS D'ÉVALUATION

Rapport d'expérience professionnelle

Secteur d'activité et type d'emploi

SECTEURS D'ACTIVITÉS

L'expert en science des données pour la connaissance client (MS) exerce ses fonctions dans tous les secteurs économiques (industrie, banque-assurance, santé, énergie, nouvelles technologies...) et dans tous types d'entreprises.

Il peut également exercer en tant qu'indépendant.

TYPE D'EMPLOIS ACCESSIBLES

Datascientist

Chef de projet Data

Chief data officer

Ingénieur data scientist

Expert data science

Data manager officer

RÉGLEMENTATIONS D'ACTIVITÉ

Le métier d'expert en science des données pour la connaissance client (MS) n'est pas encadré par une réglementation spécifique. En revanche, un expert en science des données pour la connaissance client (MS) doit se conformer aux principes du RGPD (Règlement Général sur la Protection des Données) dans l'exercice de son métier afin d'intégrer la notion de protection des données dès la conception de ses projets, mais aussi aux technologies émergentes en intelligence artificielle.

Il doit également prendre en compte la question de la propriété intellectuelle et industrielle lorsqu'il définit les limites d'utilisation des données ou métadonnées.

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)	Modérée (3/5) Moyenne pondérée par l'emploi ; les familles ci-dessous s'échelonnent de 3 à 4 sur 5.
Facteurs principaux	Manque de main-d'œuvre disponible
Salaire médian (mensuel)	2 718 €
Familles professionnelles (FAP)	L4X82 — Techniciens et agents de maîtrise chargés d'études socio-économiques · tension 4/5 (Élevée) L5X93 — Chargés d'études socio-économiques · tension 3/5 (Modérée)

Projets de recrutement (BMO 2026)	2 271
Jugés difficiles à recruter	36 %
Emplois saisonniers	2 %
Familles professionnelles (FAP)	L4X82 — Techniciens et agents de maîtrise chargés d'études socio-économiques L5X93 — Chargés d'études socio-économiques

Voies d'accès

MODALITÉS D'ACCÈS

Formation initiale, Formation continue, Contrat de professionnalisation, VAE

PRÉREQUIS À L'ENTRÉE EN FORMATION

Sont recevables les candidats titulaires de :

- ? Titre d'ingénieur diplômé conférant le grade de master (formations évaluées par la Commission des titres d'ingénieur, liste publiée au journal officiel) ;
- ? Diplôme d'une école de management, privée ou consulaire, conférant le grade de master (formations évaluées par la CEFDG, liste publiée au bulletin officiel du MESR) ;
- ? Diplôme de 3ème cycle habilité par les autorités universitaires (Diplôme national de master, DEA, DESS, ...) ou diplôme professionnel de niveau BAC + 5 ;
- ? Diplôme ou attestation de validation d'un niveau équivalent M1 pour des candidats ayant au moins trois ans d'expérience professionnelle en lien avec la formation visée.
- ? Titre inscrit au Répertoire national des certifications professionnelles (RNCP) niveau 7 ;
- ? Diplôme étranger équivalent aux diplômes français exigés ci-dessus.

Conditions d'accès dérogatoires

- Dans la limite de 40 % maximum de l'effectif de la promotion suivant la formation Mastère Spécialisé concernée, sont recevables, après une procédure de Validation des acquis personnels et professionnels (VAPP), les candidatures de personnes, justifiant a minima de 5 années d'expérience professionnelle pour lesquelles les activités exercées ont un lien avéré avec les compétences professionnelles visées par la formation.
 - Par dérogation pour 30 % maximum du nombre d'apprenants suivant la formation Mastère Spécialisé concernée, sont recevables les candidatures de titulaires d'un des diplômes suivants :
 - Diplôme ou attestation de validation d'un niveau équivalent M1 sans expérience professionnelle ou ayant moins de trois ans d'expérience professionnelle en lien avec la formation visée.
 - Diplôme de licence (L3) ou grade de Licence ou titre inscrit au RNCP niveau 6 justifiant d'une expérience professionnelle de 3 ans minimum, en lien avec la formation visée
- Le pourcentage total des dérogations prévues ci-dessus ne doit pas excéder 40%.

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

Le jury est composé sur décision du directeur de l'ENSAI. Le jury est composé de 3 membres :

- un membre permanent (le Directeur de l'ENSAI ou son représentant) ;
- et deux membres non permanents minimum, ayant une activité principale autre que l'enseignement et compétents pour apprécier la nature des acquis, notamment professionnels, dont la validation est sollicitée, c'est-à-dire ayant au moins un titre ou diplôme de même niveau que la certification visée et/ou plusieurs années d'expérience dans le secteur d'activité

COMPOSITION DU JURY — VAE

Le jury est composé de 4 membres :

- deux membres permanents : le Directeur de l'ENSAI ou son représentant, et un enseignant-chercheur
- et deux membres non permanents minimum, ayant une activité principale autre que l'enseignement et compétents pour apprécier la nature des acquis, notamment professionnels, dont la validation est sollicitée, c'est-à-dire ayant au moins un titre ou diplôme de même niveau que la certification visée et/ou plusieurs années d'expérience dans le secteur d'activité.

Base légale

Date de décision	25/06/2025
Durée de l'enregistrement	3 ans
Date d'échéance de l'enregistrement	25/06/2028
Date de dernière délivrance possible	25/06/2032

Statistiques

Année	Certifiés	Dont VAE	Insertion 6 mois	Métier 2 ans
2023	9	0	86 %	—
2022	7	0	80 %	100 %
2021	6	0	83 %	100 %

Liens avec d'autres certifications

CORRESPONDANCES (PAR BLOCS)
RNCP37431

Pour plus d'informations

LIEN VERS LE DESCRIPTIF
<https://ensai.fr/mastere-specialise-data-science-pour-la-connaissance-client/>

Fiche officielle (France Compétences)

<https://www.francecompetences.fr/recherche/rncp/40855/>

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 09/2026 à 08/2027

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