

# IBM Certified Data Architect - Big Data

RS3133 · RS · Publiée

## L'essentiel

|                       |                            |
|-----------------------|----------------------------|
| Code(s) NSF           | 326                        |
| Formacode(s)          | 31054                      |
| Type d'enregistrement | Enregistrement sur demande |
| Date d'échéance       | 31/12/2021                 |

## Certificateur(s)

CIE IBM FRANCE · SIRET : 55211846503644

## Partenaires habilités

5 organismes habilités, d'après la liste publiée par France compétences : 4 à former et organiser l'évaluation et 1 à organiser l'évaluation.

| Organisme                       | SIRET          | Habilitation                     |
|---------------------------------|----------------|----------------------------------|
| ARROW ECS                       | 38416992600027 | Former et organiser l'évaluation |
| GLOBAL KNOWLEDGE NETWORK FRANCE | 40394420000142 | Former et organiser l'évaluation |
| LEARNQUEST FRANCE SAS           | 80073704100015 | Former et organiser l'évaluation |
| PEARSON FRANCE                  | 68201927800055 | Organiser l'évaluation           |
| TECH DATA FRANCE                | 72206563800065 | Former et organiser l'évaluation |

## Résumé de la certification

### OBJECTIFS ET CONTEXTE

Tout type d'entreprise ou association ayant plusieurs postes informatisés, et entreprises du secteur informatique à partir d'une personne

Nous évoluons actuellement dans un monde connaissant une croissance exponentielle des données et, parallèlement, des flux de données.

Ces données proviennent de partout : de messages sur les sites de médias sociaux, d'images numériques et de vidéos publiées en ligne, d'enregistrements transactionnels d'achats en ligne, de signaux GPS de téléphones mobiles, de capteurs utilisés pour collecter les informations climatiques, pour ne citer que quelques sources.

Le concept de Big Data représente un ensemble de techniques concernant des données qui correspondent à plusieurs des caractéristiques suivantes : volume massif, grande variété ainsi que grande vélocité.

Les données, structurées (ex.: fichiers d'un progiciel), ou non structurées (ex.: galeries de photos), jouent un rôle prédominant auprès des entreprises qui les capturent, les gèrent et les analysent. Ces connaissances leur permettent de prendre rapidement des décisions métier bien informées.

Le professionnel certifié IBM Certified Data Architect - Big Data travaille en étroite collaboration avec l'Architecte en solutions et le client avec pour objectif de convertir les exigences opérationnelles de ce dernier en une solution-cible Big Data.

L'architecte Big Data est expert en matière de technologies applicables dans ce domaine. Il comprend la façon dont celles-ci s'articulent et sait les intégrer et les associer au mieux pour résoudre efficacement un problème opérationnel Big Data.

### COMPÉTENCES ATTESTÉES

A l'issue de sa certification IBM Certified Data Architect - Big Data, le spécialiste sera en mesure de démontrer les compétences suivantes :

- Modéliser pour l'entreprise des systèmes de traitement des données à grande échelle
- Prendre part aux décisions de l'entreprise portant sur l'architecture de ses matériels et logiciels.
- Concevoir, en tenant compte de la complexité des données, des systèmes et des modèles capables de traiter toutes sortes de données (structurées, semi-structurées, non structurées), variant dans leurs volumes, vitesses (y compris traitement des flux) et vérités.
- Prendre efficacement en charge les contraintes de gouvernance et de sécurité de l'information associées au système.

Modélisation de systèmes de traitement des données

- Concevoir un modèle de données et de flux de données qui répond aux besoins métier
- Définir la structure des données saisies

Prise de décisions sur l'architecture des matériels et des logiciels.

- Décider de la technologie à appliquer pour garantir une évolutivité horizontale et verticale
- Fixer les besoins en termes de stockage de données en fonction de leur volume
- Désigner la technologie Big Data la plus adaptée aux besoins spécifiques du client (par ex. Hive/HBase ou Cloudant)
- Attribuer le format de stockage et la compression adaptés aux exigences du client
- Définir les besoins en infrastructures matérielles et logicielles
- Privilégier une solution basée sur le cloud par rapport à une solution interne lorsque le contexte y est plus favorable (et établir des plans de migration de l'une vers l'autre)

- Démontrer l'opportunité d'utiliser, dans un contexte approprié, les technologies suivantes :

o Cloudant, SQL ou NoSQL, Open Data Platform, BigInsights, BigSQL, Hadoop, BigR et SPSS, BigSheets, Streams, Netezza, DB2 BLU, GPFS/HPFS, Spark, YARN

Conception de systèmes et de modèles capables de traiter toutes sortes de données

- Concevoir l'intégration des composants matériels et logiciels requis
- Concevoir les connecteurs/interfaces/API's entre la solution Big Data et les systèmes existants
- Définir les sorties
- Cerner les exigences à prendre en compte pour remplacer et/ou fusionner avec les solutions opérationnelles existantes
- Elaborer la solution répondant au SLA du client (accord sur le niveau de services)
- Définir les exigences en matière de réseau en fonction des besoins client

Prise en charge des contraintes de gouvernance et de sécurité de l'information

- Repérer les besoins en haute disponibilité
- Cerner les exigences d'une restauration après sinistre
- Fixer les exigences techniques pour le stockage et la réplication des données
- Fixer les exigences techniques pour se prémunir contre la perte des données
- Fixer les exigences de sécurité

## Voies d'accès

### PRÉREQUIS À L'ENTRÉE EN FORMATION

Préalablement à sa préparation à la certification, le professionnel devra détenir les connaissances et aptitudes suivantes:

- Connaissance de données et ses zones d'enjeux/risques potentiels
- Capacité à transformer les exigences fonctionnelles en spécifications techniques
- Capacité à aboutir à une architecture physique en partant d'une solution globale ou d'une architecture logique
- Gestion des clusters
- Exigences réseau
- Principales interfaces
- Modélisation des données
- Capacité à identifier pour la solution les exigences non fonctionnelles et à apporter un soutien technique
- Latence, flexibilité, haute disponibilité, réplication et synchronisation des données, restauration après sinistre
- Performance globale (performance des requêtes, gestion de la charge de travail, optimisation de la base de données)

- Proposer les meilleures pratiques ou les plus recommandées en matière de déplacement, manipulation et stockage des données dans une solution Big Data
- Options techniques d'ingestion des données
- Options et ramifications du stockage de données, dont exigences supplémentaires et enjeux imposés par les données dans le Cloud
- Techniques de requêtes de données et leur disponibilité en matière de soutien à l'analyse.
- Généalogie des données et gouvernance des données
- Diversité des données et volume des données
- Mettre en place la sécurité des données et fournir des orientations dans ce domaine, conseiller en particulier sur la sécurité LDAP, les fonctions/la sécurité des utilisateurs, le monitoring des données, la sécurité des données à caractère personnel
- Logiciels à connaître en priorité :  
BigInsights, BigSQL, Hadoop, Cloudant (NoSQL)
- Logiciels à connaître accessoirement :  
Information Server - Intégration à partir de BigInsights - Balanced Optimization pour Hadoop -  
Fonction Push Down de JAQL - Gouvernance des données - Spécificités de BigInsights liées à la sécurité -  
Metadata Workbench - Intégrer Optim avec BigInsights - DataClick pour BigInsights -  
BigMatch - Guardium (monitoring) - Outils d'analyse des données (SPSS) - BigSheets -  
Support technique de Hadoop/BigInsights - Disponibilité des données et support dans leur requête -  
Flux de données - Interface/intégration avec BigInsights - Données de streaming -  
Analyse de données intégrée - Netezza - DB2 BLU - Bases de données Graph -  
Machine Learning (système ML)

## Base légale

Date d'échéance de l'enregistrement 31/12/2021

---

## Pour plus d'informations

LIEN VERS LE DESCRIPTIF

Lien de la certification :

### Fiche officielle (France Compétences)

<https://www.francecompetences.fr/recherche/rs/3133/>