

## IBM Certified Data Architect - Big Data

CATEGORIE : B

### Vue d'ensemble

Domaine(s) d'activité professionnel dans lequel(s) est utilisé la certification :

Transverse : ■ **Transverse et tout secteur d'activités**

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

Code(s) NAF : **62.02A**

Code(s) NSF : **326**

Code(s) ROME : **M1802**, **M1805**, **M1806**, **M1810**

Formacode : **31054**

Date de création de la certification : **06/08/2015**

Mots clés : **Architecte de données**,  
**Modélisation de données**, **Analyse de données**,  
**Big Data**

### Identification

Identifiant : **3133**

Version du : **06/11/2018**

### Références

Consensus, reconnaissance ou recommandation :

Formalisé :

- [Détails sur les sites :](#)
- [\\*](#)
- [, rubrique certification IBM](#)
- [La recherche s'effectuera ensuite par mot-clé ou n° de test ou de certification sur le site en anglais des certifications IBM.](#)

Non formalisé :

- **N/A**

Norme(s) associée(s) :

—

### Descriptif

#### *Objectifs de l'habilitation/certification*

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 vitesse.

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.

## *Lien avec les certifications professionnelles ou les CQP enregistrés au RNCP*

- 19431
- 12054
- 9874
- 15719
- 11554
- 13082

## *Descriptif général des compétences constituant la certification*

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, vélocités (y compris traitement des flux) et véracités.

Prendre efficacement en charge les contraintes de gouvernance et de sécurité de l'information associées au système.

## *Modalités générales*

Le candidat peut préparer cette certification grâce à plusieurs modules de formations dirigés par un formateur d'une durée de 1 à 3 jours (durée totale des formations recommandées = 7 jours), ainsi qu'à plusieurs lectures web gratuites disponibles sur internet.

## *Liens avec le développement durable*

Aucun

## Public visé par la certification

- multiples : salariés, demandeurs d'emploi

## Valeur ajoutée pour la mobilité professionnelle et l'emploi

### *Pour l'individu*

Le professionnel certifié **IBM Certified Data Architect - Big Data** est garant de l'optimisation des objectifs stratégiques d'une organisation par son expertise quant à l'utilisation des données d'entreprise.

Il définit les standards relatifs à la nature des données, leur collecte et leur mode de stockage et s'assure des modalités de leur sélection, leur intégration, ainsi que leur utilisation dans des systèmes de données.

En définissant la stratégie Big Data des données, le Data Architect endosse une triple responsabilité : une vision prospective sur les cas de figure possible,

### *Pour l'entité utilisatrice*

Notre monde numérique génère une multitude de données au volume et au débit considérables dont l'exploitation constitue un défi technique. Les entreprises sont plus nombreuses à stocker, traiter et exploiter leurs données, quelles qu'en soient la forme et le volume.

- ° Comment repenser le stockage de ces données, quand les infrastructures classiques ne sont plus suffisantes ?
- ° Comment les organiser pour en permettre un accès et une utilisation efficaces ?
- ° Quelles infrastructures mettre en place pour servir les besoins métiers ?

un ajustement en cours de développement et une mise en concordance avec les spécifications d'origine.

Il peut intégrer une équipe complète d'experts des données et apporter sa vision globale de concepteur de l'architecture en s'appuyant sur sa connaissance des enjeux métiers, ou bien intervient simplement en prescripteur auprès d'une Direction des Systèmes d'Information.

Le Big Data et les métiers liés à ce domaine sont particulièrement émergents en France actuellement. Certains géants sont à la pointe et à l'affût de talents. La montée des technologies numériques dans le monde et la quantité massive de données qu'elles génèrent, crée également des emplois liés au Big Data.

De ce fait, le nouveau profil d'Architecte Big Data est très recherché, aussi bien par les grandes entreprises que par les start-up en pleine croissance.

Les Data Architects évoluent au cœur de ces enjeux. À l'intersection de plusieurs métiers, ils sont à la fois des concepteurs d'infrastructures informatiques, des administrateurs et des analystes.

Le professionnel certifié **IBM Certified Data Architect - Big Data** mettra son expertise au service de son entreprise en concevant des solutions fonctionnelles, tout en anticipant les besoins futurs et en respectant les contraintes imposées par les architectures existantes.

## Evaluation / certification

### Pré-requis

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

Couche 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

### Certificateur(s)

- IBM

### Centre(s) de passage/certification

- De nombreux centres Pearson Vue France entière.

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, Cloudera (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)

## Compétences évaluées

### 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 Cloudera)

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 :

Cloudera, 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

*Niveaux délivrés le cas échéant (hors nomenclature des niveaux de formation de 1969)*

N/A

La validité est Permanente

### Possibilité de certification partielle : non

Matérialisation officielle de la certification :

Un Certificat électronique est envoyé par IBM aux candidats ayant réussi, ainsi qu'une carte professionnelle lors du 1er examen réussi. Le candidat accède à son transcript à tout moment sur un portail sécurisé dédié.

## Plus d'informations

### Statistiques

En France, depuis 2015, environ **800 personnes ont obtenu une certification IBM dans le domaine de l'Analyse.**

Le programme de certification professionnelle IBM existe depuis plus de 15 ans et compte entre 1000 et 2000 nouveaux certifiés par an.

### Autres sources d'information

<http://home.pearsonvue.com/IBM>

<http://www-03.ibm.com/certify/>

Lien de la certification :

<https://www.ibm.com/certify/cert?id=50001701>

