



---

# FORMATIONS ET COMPETENCES CLOUD COMPUTING et BIG DATA EN France

---

## Rapport Big data

Rapport final – 3 décembre 2015

*Etude réalisée par le cabinet Ambroise Bouteille et associés et l'IDATE*



---

## 0. Rappel des objectifs et de la méthode

---

1. Des innovations technologiques...
2. ... ouvrent de nouvelles fonctionnalités et usages pour les entreprises
3. Un sujet d'avenir, en émergence, autour duquel de nouvelles offres se développent
4. L'impact de ces bouleversements sur les modes de travail entre clients et prestataires
5. Les besoins en compétences liés au Big data
6. Les enjeux en termes de formation liés au Big data
7. Quantification des volumes d'emploi liés au Big data
8. Cartographie et analyse de l'offre de formation
9. Préconisations
10. Annexes



## Les objectifs

NB : ce rapport porte sur le Big data, les commanditaires ayant souhaité que les deux sujets –cloud et big data – fassent l’objet de rapports distincts. Les travaux ont toutefois été menés de front, car faisant l’objet d’une seule et même étude (terrain partiellement commun : certaines entreprises, notamment les entreprises utilisatrices, lorsque possible ont été interrogées sur les deux sujets)

### Repérer les besoins induits par le Cloud computing et le Big data

Evolutions des organisations des entreprises  
Impacts sur l’emploi et les fonctions métiers : besoins en  
recrutement et compétences

### Cartographier l’offre de formation spécialisée sur le Cloud computing et le Big data

Apprécier l’adéquation entre  
cette offre de formation, tant  
initiale que continue, à répondre  
aux besoins actuels et futurs des  
professionnels

En fonction des décalages pouvant être  
repérés, formuler des  
préconisations sur l’offre de  
formation



## La méthode

Les grandes étapes

Les points-clés de la méthode

### Phase 1 : analyse des besoins en recrutement et compétences liés au Big data et Cloud computing

#### *Etape 1.1 : Cadrage des travaux*

- comité de pilotage
- analyse documentaire

#### *Etape 1.2 : Entretiens qualitatifs auprès des professionnels du numérique*

- Groupe de travail d'experts métiers
- 35 entretiens auprès d'entreprises de la branche et d'entreprises de secteurs utilisateurs
- 10 entretiens auprès de salariés
- comité de pilotage

#### *Etape 1.3 : Quantification des volumes d'emploi liés au Big data et au Cloud computing*

- à partir des études déjà réalisées et des éléments remontés lors des entretiens qualitatifs, élaboration de projections à horizon 5 ans sur les volumes d'emploi
- comité de pilotage

### Phase 2 : Analyse de l'offre de formation en matière de Cloud computing et Big data

#### *Etape 2.1 : Cartographie de l'offre de formations françaises relatives aux Cloud computing et Big data*

- recherche documentaire pour le recensement de l'offre de formation initiale et continue
- 5 entretiens auprès d'organismes de formation

#### *Etape 2.2 : Appréciation de la capacité de l'offre identifiée à répondre aux besoins des entreprises*

- 2 entretiens auprès d'interlocuteurs du Ministère de l'Economie, suivant les travaux des plans Big data et Cloud computing
- comité de pilotage

### Phase 3 : Préconisations et communication des travaux

#### *Etape 3.1 : Elaboration des préconisations*

- élaboration des préconisations (à l'aune de l'ensemble des éléments et conclusions des précédentes phases)
- 2 entretiens auprès des experts métiers mandatés par l'Opiiec pour tester ces préconisations
- comité de pilotage

#### *Etape 3.2 : Présentation et synthèse des travaux*

- présentation en commission Opiiec
- présentation en CPNEFP



## Les entreprises et personnes interrogées

Dans le cadre de l'ensemble de l'étude ont été interrogés :

- sur le **big data** :

Entreprises	Personnes interrogées
<b>Sentelis</b>	Shahnaz El Fassi, responsable recrutement
<b>Groupe HLI</b>	Sandra Verrecchia, RRH
<b>Sodifrance</b>	Franck Mazin
<b>SAS</b>	Mouloud Dey, Directeur « Emerging Business solutions Big data, Business Analytics, Data Science »
<b>Affinitech</b>	Vincent Heuschling, directeur
<b>Saegus</b>	Frédéric Brajon, directeur et co-fondateur
<b>Oracle</b>	Sébastien Vige, (presales Western Europe)
<b>Microsoft</b>	Damien Cudel, chef de marché plateforme applicative Big data et cloud
<b>IBM</b>	Jea-Charles Cointot, Cloud & smart infrastructure development executive
<b>Hurence</b>	Thomas Baillet, CTO
<b>Hortonworks</b>	Vincent Bonnot (directeur France), John Kreisa (VP international marketing)
<b>Energiency</b>	Sébastien Duburque
<b>Syncsort</b>	Stéphane Heckel
<b>Open</b>	Laurent Benazera, directeur du recrutement ; Sandrine Chevallier, HR manager

- sur le **cloud computing** :

Entreprises	Personnes interrogées
<b>B2Cloud</b>	Catherine Nohra-China
<b>L'Acoms</b>	Frans Imbert-Vier, directeur associé
<b>Viavoo</b>	Sébastien Louvet, COO
<b>Appian</b>	Laurent Chailley
<b>Antemeta</b>	Bertrand Bombes de Villiers
<b>Qarnot computing</b>	Paul Benoit, président
<b>Numergy</b>	Philippe Tavernier, DG ; Gwenola Mouilleseaux, RRH
<b>Nuagelabs</b>	Rapahël Nicoud, CEO



<b>Multiposting</b>	Romain Herbeau, directeur commercial ; Alexandre Sadones, CTO
<b>Clever Cloud</b>	Quentin Adam, CEO
<b>Gestion-Air</b>	Thomas Proust, CEO
<b>Optimcloud</b>	Eric Lecoquierre
<b>Nuageo</b>	Benoit Degrand, Associé
<b>In Web technologies</b>	Didier Perrot
<b>Sopra-Steria</b>	Thierry Luc
<b>Cedexis</b>	Julien Coulon, fondateur

- parmi des **entreprises utilisatrices** (c'est-à-dire, les entreprises clientes de la branche) :

Entreprises	Personnes interrogées
<b>SNCF Gares et connexion</b>	Marie-Caroline Bénézet Lanfranchi, Directrice du digital
<b>Neopost</b>	Hugues Le Bars, Chief Data Officer
<b>Michelin</b>	Agnès Mauffrey, DSI
<b>Groupe BNP Paribas</b>	Jocelyn Sedes, chef de projet, représentant du pôle urbanisation SI
<b>BPCE</b>	Bernard Gouraud (ex-DSI, qui exerce désormais une activité de conseil auprès de la DG)
<b>Association nationale des DSI</b>	Pierre Delort, Président de l'ANDSI, ancien DSI de l'INSERM, auteur de l'ouvrage <i>Le Big data</i> , aux PUF (collection « Que sais- je ? »)

- des **salariés** exerçant des métiers du cloud et big data :

Prénom des personnes interrogées	Poste occupé
<b>Marion</b>	Consultante data science
<b>Charles</b>	Data scientist
<b>Arnaud</b>	Data scientist
<b>Romain</b>	Data scientist
<b>Morgan</b>	Business analyst / data scientist
<b>Aloïs</b>	Architecte cloud et virtualisation
<b>Yoann</b>	Architecte PHP
<b>Mustafa</b>	Architecte virtualisation et cloud
<b>Samia</b>	Consultante Cloud transformation
<b>Eric</b>	Ingénieur systèmes et réseaux



- des organismes ou réseaux d'organismes de formation :

Nom de l'organisme	Personne interrogée
<b>Data ScienceTech Institute</b>	Léo Souquet
<b>ISEP</b>	Denis Beautier
<b>Pasc@line</b>	Christian Colmant
<b>Ecole Polytechnique</b>	Erwan Le Pennec
<b>EPITA</b>	Christian Dujardin

- ainsi que deux réunions de travail avec les experts-métiers des organisations professionnelles (Syntec Numérique et CINOV) :

Groupe de travail	Experts présents
<b>02/07/2015 – Syntec Numérique</b>	Christian Martin (Witkey), Hugues Le Bars (Neopost), Benjamin Crombé (Consort NT), Catherine Coq (Orange)
<b>10/09/2015 – CINOV IT</b>	Alain Prallong (Président de CINOV-IT), Stéphane Coire (administrateur), Jean-Denis Tsati (administrateur)

- ainsi que des échanges avec des interlocuteurs au sein du Ministère de l'Economie

Nom de l'organisme	Personne inter
<b>DGE</b>	Fabien Terrailot
<b>DGE</b>	Cédric Mora



---

## Le périmètre de cette étude

---

Tant le big data que le cloud computing concernent un **périmètre plus large que celui *stricto sensu* des professionnels du numérique au sein de la branche** des métiers du Numérique, de l'Ingénierie, des Etudes et conseil et de l'Événementiel. En effet, ces technologies, et surtout les usages possibles qu'elles induisent, sont susceptibles de trouver résonance dans n'importe quelle entreprise/association/organisme public : potentiellement, ce sont donc toutes les entreprises de tous les secteurs d'activités possibles qui sont amenées à conduire et mettre en place des projets Cloud et Big data et d'y recourir.

Ainsi, nul doute que ces deux sujets vont également avoir des **échos en termes d'emploi et de besoins en compétences dans tous les secteurs utilisateurs**. Certaines entreprises peuvent en effet internaliser ces projets cloud / big data, mettre en place des équipes dédiées, etc. En outre, dans la mesure où tant le cloud que le big data révolutionnent les usages numériques des entreprises, ce sont l'ensemble des métiers de chaque entreprise qui peuvent être impactés (par exemple, un vendeur disposera de plus d'éléments sur la satisfaction et les attentes clients, etc. ; les services marketing disposeront d'outils nouveaux pour cerner les besoins clients, auront accès à davantage d'informations, etc. ; les services opérationnels au sein des entreprises auront une plus grande capacité d'autonomie et d'initiatives concernant l'IT, etc.). Autrement dit, tant le cloud que le big data sont susceptibles **d'impacter l'ensemble des métiers**. Il s'agit là d'un sujet passionnant, mais d'une ampleur considérable, qui nécessiterait, pour atteindre son ambition l'engagement d'un Contrat d'Etudes Prospectives, a minima, entre plusieurs branches (prestataires et utilisatrices).

Les entreprises du numérique et donc de la branche ont **une valeur ajoutée toute particulière pour accompagner les entreprises (leurs clients) dans leur transition numérique**. Pour cela, il est nécessaire de comprendre ces enjeux du point de vue des clients. En outre, sur les métiers de l'IT, certains métiers ou compétences (au sens d'emploi, de poste) peuvent être exercés tant chez les prestataires que chez les clients.

C'est pourquoi, parmi le panel de professionnels interrogés, ont été **incluses quelques « entreprises utilisatrices »** (c'est-à-dire des entreprises clientes des professionnels de la branche), afin notamment de mieux comprendre quelles attentes nouvelles celles-ci pouvaient avoir vis-à-vis de leurs prestataires dans leurs démarches big data et/ou cloud.

Néanmoins, rappelons que comme convenu ce sont bel et bien les problématiques de compétences des **entreprises de la branche** (ESN, éditeurs de logiciel, hébergeur, intégrateurs, etc.) **qui sont le cœur de sujet de cette étude**.





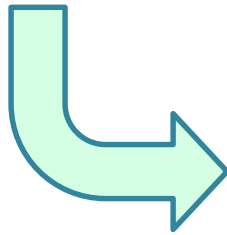
**En synthèse :**

(en vert pâle le périmètre de cette étude)

Professionnels du numérique (entreprises de la branche) :  
ESN, constructeurs, intégrateurs, éditeurs de logiciels, etc.)

Métiers « techno »  
communs, impactés  
par le cloud et le big  
data

Nouvelles compétences  
/ évolution des  
compétences  
nécessaire pour :  
vendre, conseiller,  
implémenter, exploiter  
au quotidien



Entreprises clientes (tous secteurs d'activités :  
transport, secteur public, banque/assurance,  
restauration, industrie, santé, télécoms, commerce,  
etc.)



## 0. Rappel des objectifs et de la méthode

---

## 1. Des innovations technologiques...

---

2. ... ouvrent de nouvelles fonctionnalités et usages pour les entreprises
3. Un sujet d'avenir, en émergence, autour duquel de nouvelles offres se développent
4. L'impact de ces bouleversements sur les modes de travail entre clients et prestataires
5. Les besoins en compétences liés au Big data
6. Les enjeux en termes de formation liés au Big data
7. Quantification des volumes d'emploi liés au Big data
8. Cartographie et analyse de l'offre de formation
9. Préconisations
10. Annexes



## Le Big data est plus qu'une révolution technologique

### Un buzz autour du « Big data », qui conduit à certains abus de langage

Le terme « Big Data » ne possède pas de définition universelle et désigne à la fois les propriétés des données et les méthodes de traitement spécifiques. Le terme « Big » qualifie un **volume de données très important** (sans qu'un seuil minimal ait été déterminé) et généralement des données de formats différents, dites « **non structurées** » (en opposition aux bases de données traditionnelles structurées), en correspondance avec l'accroissement des données en temps réel en provenance notamment des réseaux sociaux ou des capteurs présents dans les téléphones mobiles ou les objets connectés. Les données non structurées ne sont pas gérables par les solutions traditionnelles (datamining, tableurs).

Sans doute lié à un effet de mode, tout type de traitement de la donnée tend à être estampillé « Big data », tout poste traitant des informations chiffrées se prévaut du qualificatif de « data analyst » ou de « data scientist ».

Aussi, par abus de langage, le terme « Big data » est souvent employé de façon très large pour catégoriser la data science (qui n'est pas nécessairement du Big data) :

- d'après nos interlocuteurs, la datascience est plus large que le Big data et n'est pas nécessairement prédictive : elle pourrait être définie par tout type de traitement statistique qui nécessite des traitements plus complexes que les outils classiques (tableur, etc.) et fait appel à quelques compétences en programmation
- le Big data est en fait une branche de la data science, qui nécessite des infrastructures (un socle technologique donc) complexes et spécifiques (puissance de calcul, calcul parallèle, architecture distribuée, etc.), dont l'avancée technologique reste avant tout **l'assimilation de données non-structurées**, et l'accélération de son traitement, qui tend vers le traitement en temps-réel, bien que cela « relève encore du mythe » aujourd'hui.

### Le Big data est susceptible de remodeler en profondeur la façon dont les entreprises font du business

Les **champs d'application possibles sont encore loin d'être défrichés** (cf. infra), toutefois les premiers cas d'usage démontrent que le Big data est fortement créateur de valeur pour l'entreprise.

Loin de prétendre à l'exhaustivité, quelques exemples :

- Les *success stories* entrepreneuriales du 21<sup>ème</sup> siècle sont, pour la plupart, celles des start up qui ont créé leur business model autour de la donnée (*data-driven business model*) : Airbnb, Blablacar, etc.
- Dans le secteur des télécoms, la prédiction du *churn* (score d'attrition) permet aux opérateurs (dans un contexte économique par ailleurs contraint pour eux) de réaliser des économies notoires (il est moins coûteux d'investir pour fidéliser un client existant que pour développer des parts de marché)
- L'analyse fine et prédictive des comportements clients permet au secteur de la banque/assurance de mieux gérer les risques clients



- Etc.

### **Ainsi, le Big data ne saurait se cantonner à la sphère de l'IT**

Le Big data **impacte en profondeur les modèles économiques des entreprises**. C'est donc l'ensemble des métiers de l'entreprise qui sont concernés (NB : par métier, nous entendons ici les fonctions de l'entreprise, par opposition à l'IT)

D'une manière générale, les réflexions autour du Big Data au sein des entreprises et pour leurs prestataires concernent le Data Management au sens large (partage de la donnée entre les silos, constitution d'un data lake, etc.).

### **Pour les entreprises utilisatrices (ou potentiellement utilisatrices), le Big data est un sujet stratégique avant tout**

En effet, les initiateurs (pionniers) des projets Big data au sein des entreprises utilisatrices sont rarement les DSI (même si celles-ci peuvent être associées) :

- les directions commerciales et marketing ont sans doute été les premières à s'intéresser en profondeur au Big data et à chercher à capter, exploiter et valoriser la donnée dans une optique de création de valeur
- au sein des entreprises les plus avancées en la matière, la direction générale est pleinement impliquée dans ces projets. Le cœur du Big data étant justement, non pas la data, mais la prise de décision grâce à la data.



---

## Le Big Data repose toutefois sur quelques innovations technologiques

---

**Avant d’aller sur le Big Data proprement dit, les entreprises semblent avant tout se concentrer sur la maîtrise de leurs données**

Avant d’assimiler des données provenant de sources données tierces (dont l’open data), souvent présentées comme la vraie révolution du BD, la plupart des entreprises ayant des projets en la matière se concentrent déjà sur une **première étape de maîtrise de la donnée interne**.

Le traitement des données repose d’abord souvent sur les « **ancêtres** » du **Big Data**. Les méthodes de traitement du Big Data dérivent en partie des technologies de *data mining* et de *business intelligence*, déjà en place dans de nombreuses entreprises, mais s’appuient aussi sur **quelques nouvelles technologies liées au traitement et à la visualisation** de ces données particulières.

Les principales technologies sont des briques logicielles d’infrastructure (MapReduce, initiée par Google, ou encore Hadoop) assurant la mise en parallèle de tâches de calcul ou la gestion de bases de données non-structurées (NoSQL).

La principale rupture vient donc surtout des données elles-mêmes, en particulier celles qui proviennent des utilisateurs finaux, désormais plus faciles à capter via les développements de l’internet et une infrastructure logicielle adaptée (notamment en cloud).

Les plus grandes ruptures apportées par le Big Data sont donc le **développement des données non structurées** et la **mise en place d’une centralisation des données** (déjà tentée sans succès par certains acteurs autour du BI). Ceci implique notamment:

- Constitution du **data lake** / partage de la donnée entre silos
- Organisation de la **gouvernance** autour du data management.

**Au cœur du Big data, deux aspects concomitants : une révolution dans le traitement et l’exploitation de la donnée, permise par des innovations technologiques et digitales**

Parmi lesquelles :

- la génération d’un **volume** considérable de données (réseaux sociaux, IoT<sup>1</sup>, etc.), d’une grande hétérogénéité (**variété**)
- les avancées permises par le cloud (à ce sujet, cf. rapport cloud computing)
- la capacité à traiter des données non structurées et en temps réel (**vitesse** de traitement)
- l’accélération des puissances de traitement
- les architectures distribuées et le calcul parallèle

---

<sup>1</sup> Internet of Things (objets connectés)



- l'amélioration des algorithmes de machine learning et d'intelligence artificielle
- etc.

La révolution « Big data », si elle n'est pas strictement IT, est cependant la conséquence directe de ces innovations technologiques : c'est la conjugaison des possibles offerts par ces innovations technologiques et de leur exploitation pour une aide à la décision (**valeur** issue de l'analyse des données)

Il existe de nombreuses techniques d'analyse de données, issues notamment des mathématiques ou de la biologie (analyse génétique, réseaux neuronaux). Les technologies big data reprennent, entre autres, des algorithmes utilisés par le data mining ou par l'analyse statistique, mais aussi d'autres procédés tels que :

- la **fusion et l'intégration de données** (ou « croisement ») : il s'agit de différentes techniques qui intègrent et analysent les données provenant de sources multiples pour en extraire un aperçu d'une manière plus efficace et potentiellement plus précise que si l'analyse avait été faite à partir d'une source unique de données. Des techniques de traitement du signal peuvent être utilisées pour mettre en œuvre certains types de fusion de données. Par exemple, les données des capteurs de l'Internet des objets peuvent être combinées pour obtenir une meilleure vision de la performance d'une raffinerie de pétrole. Également, les données de médias sociaux peuvent être combinées avec les données de ventes en temps réel d'un produit, afin de déterminer quel effet a une campagne de marketing sur le comportement d'achat des clients.
- Le **"machine learning"** : cette technique d'intelligence artificielle permet aux ordinateurs de faire évoluer leurs comportements en se basant sur des données empiriques. Un des principaux objectifs de la recherche d'apprentissage machine est d'apprendre à reconnaître automatiquement des schémas complexes et à prendre des décisions intelligentes sur la base des données déjà analysées jusqu'à présent. L'algorithme de traitement du langage est un exemple de "machine learning". Ce programme analyse le langage humain et apprend progressivement son fonctionnement pour ensuite être capable de traduire la langue. Les systèmes de traduction que l'on retrouve sur Internet se basent pour certains sur ce type d'algorithme.
- L'**optimisation** : il s'agit de méthodes mathématiques destinées à améliorer la performance de processus en fonction d'un ou plusieurs paramètres (par exemple, coût, rapidité ou fiabilité). Il peut s'agir de l'amélioration des processus opérationnels comme la planification, le routage, l'aménagement d'installations. Dans le domaine stratégique, l'optimisation aide à la prise de décisions stratégiques telles que la stratégie de gamme de produits, l'analyse des investissements, ou la stratégie de portefeuille de R&D.

### **Le cloud computing : une opportunité pour le développement du Big Data**

Lorsqu'une entreprise souhaite utiliser des technologies Big Data, il est souvent nécessaire qu'elle possède des serveurs suffisamment puissants pour traiter les données par la suite. Les infrastructures déjà présentes dans l'entreprise ne sont généralement



pas suffisantes pour pouvoir supporter la charge de calcul. Big Data n'est donc pas un concept « scalable » ou adaptable à toute entreprise et à toute situation au vue des capacités de calcul nécessaires pour sa mise en œuvre.

Le cloud computing permet toutefois de mettre à disposition des infrastructures de grande capacité, avec une facturation à l'usage. Au lieu d'utiliser un serveur pendant longtemps, le cloud offre la possibilité d'utiliser beaucoup de serveurs pendant peu de temps, notamment pour effectuer des calculs sur de grandes quantités de données. Le cloud offre aux petits acteurs qui ne sont pas en mesure d'acquérir de grosses infrastructures, la possibilité de faire du traitement de données, du big data, en étant facturé à l'usage.



0. Rappel des objectifs et de la méthode
  1. Des innovations technologiques...
- 

## 2. ...ouvrent de nouvelles fonctionnalités et usages pour les entreprises

---

3. Un sujet d'avenir, en émergence, autour duquel de nouvelles offres se développent
4. L'impact de ces bouleversements sur les modes de travail entre clients et prestataires
5. Les besoins en compétences liés au Big data
6. Les enjeux en termes de formation liés au Big data
7. Quantification des volumes d'emploi liés au Big data
8. Cartographie et analyse de l'offre de formation
9. Préconisations
10. Annexes





## La diversité des données ouvre le champ des usages possibles

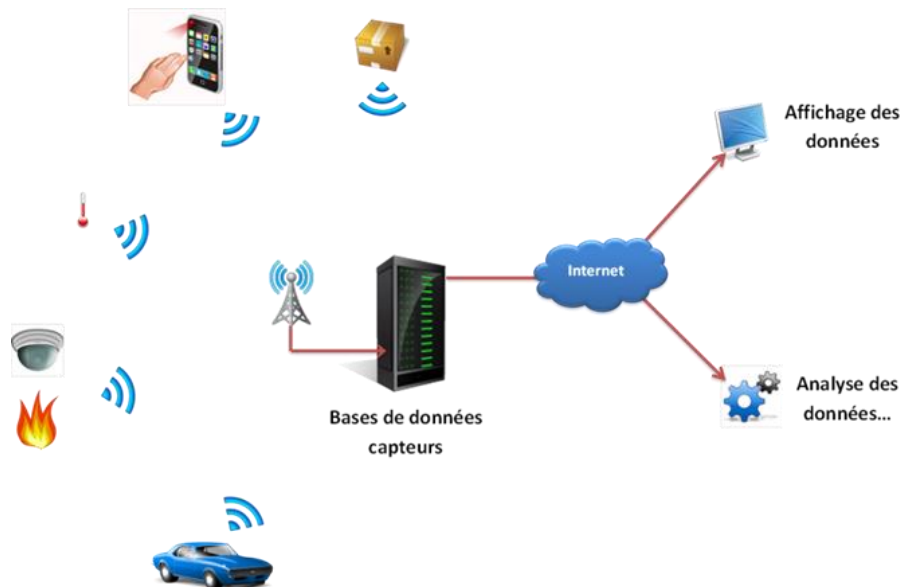
Le Big Data est également né de la diversité très importantes des données désormais utilisables et disponibles, que ce soit au sein même d'une entreprise, ou « publiquement », en général sur Internet. Le développement des réseaux sociaux et des objets connectés ont accentué la création automatique de données, et qui peuvent aujourd'hui présenter un intérêt de croisement et d'analyse.

### De multiples sources de données existent aujourd'hui

La collecte d'informations tend à s'accélérer, avec notamment le développement de nombreuses sources de données. C'est par exemple le cas des capteurs (mouvement, environnement, médical, localisation...) qui se retrouvent aussi bien sur les téléphones eux-mêmes que sur des machines tierces (véhicules, GPS, sondes météo...) connectées à l'Internet directement ou via un smartphone.

Le champ des possibilités de l'Internet des objets en termes de génération de données devient ainsi quasi infini, même si sa mise en place reste progressive et lente. Le schéma ci-dessous détaille les flux de données de l'Internet des objets.

Figure 1: Flux (simplifié) des données de l'Internet des objets



Source : IDATE

En plus des **données déclaratives** (nom, adresse, âge, sexe, centres d'intérêts ...) des usagers qu'ils fournissent lors d'une inscription sur un site ou lors d'un achat sur Internet, les fournisseurs de services peuvent récupérer les informations non pas auprès



des usagers directement mais auprès d'autres fournisseurs de services disposant de bases de données.

C'est par exemple le cas entre Facebook et Deezer (streaming de musique), où Deezer permet à ses utilisateurs de se connecter directement à son service lorsqu'il est également connecté à Facebook. Cette méthode évite d'avoir à s'identifier deux fois. Les acteurs de l'Internet s'échangent ainsi des données pour la mise au point de nouveaux services, mais ils peuvent aussi en collecter librement sans effort auprès d'acteurs, notamment publics, libérant leurs données. Ces initiatives sont regroupées dans le concept d'"**open data**" : certaines entreprises et administrations mettent à disposition leurs bases de données pour que d'autres acteurs en tirent profit, mais aussi par souci de transparence vis-à-vis du public.

Comme dit précédemment au sujet des usages, les **réseaux sociaux** agrègent des quantités de données très importantes sur le profil de leurs utilisateurs, et également sur leurs relations avec des amis, avec des groupes d'intérêts ou avec des marques. L'ensemble de ces relations constituent un "graphe social" très large, permettant par exemple l'étude des comportements des individus.

Dans la même optique, il est possible d'exploiter les **données clients** des entreprises ou les **dossiers médicaux** des patients. Dans tous les cas, des problématiques de confidentialité se posent, notamment au niveau de l'accord que donne l'utilisateur pour utiliser ses données. À l'heure actuelle, l'exploitation des informations personnelles reste un sujet sensible : la plupart des utilisateurs ne semblent pas enclins à autoriser leur diffusion dans la plupart des cas, en particulier commerciaux. Les dossier médicaux sont, au moins en France, personnels et confidentiels, et ne peuvent être exploités que par des médecins ou par le patient lui-même.

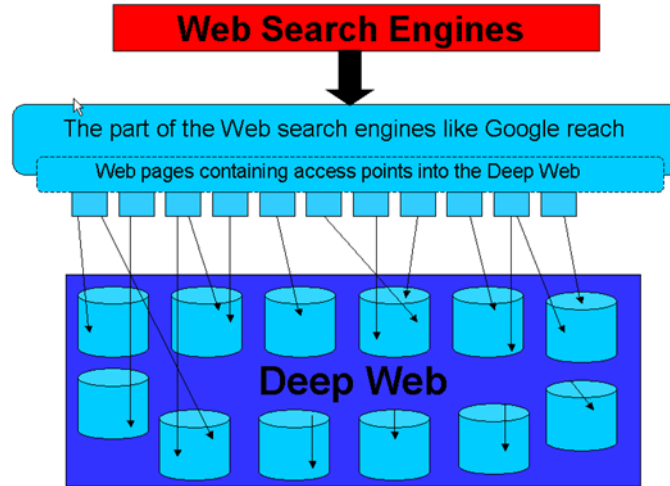
Quelques acteurs constituent des bases de données massives en les récupérant (plus ou moins légalement, parfois à la limite des pratiques autorisées par le copyright) par des robots informatiques ou par des requêtes, en particulier autour du **Web invisible**.

Le Web invisible représente plus de 70% du Web, et se caractérise par le fait qu'il n'est pas accessible directement depuis un moteur de recherche. Les contenus et pages du Web invisible sont néanmoins accessibles et présents (indirectement) sur Internet, et représentent une mine d'informations inexploitées. Ces informations sont souvent accessibles uniquement via des formulaires ou via un accès payant aux données. Les robots qu'utilisent les moteurs de recherche pour parcourir le Web et ajouter de nouvelles pages à leurs bases de données ne sont pas en mesure d'accéder aux pages du Web invisible : les données à répertorier peuvent être trop volumineuses, les pages peuvent être codées de façon à ce que les robots ne les répertorient pas, ou encore des pages peuvent être générées de façon dynamique (création spontanée et à usage unique de la page).

Les 30% restant étant le Web dit "visible" et accessible par un moteur de recherche par exemple. Le schéma ci-dessous représente le Web invisible ("deep Web") et le Web visible, en ayant comme point de départ les moteurs de recherche. Ce schéma indique notamment les points d'entrées au niveau du Web visible qui permettent d'accéder au Web invisible.



Figure 2: Web visible et invisible



Source: IDATE



---

## L'exploitation des Big data offre des possibles insoupçonnés aux entreprises ; les champs d'application (les *usecase*) sont multiples et la plupart sont à défricher

---

Les usages de Big Data peuvent être assez vastes. Les secteurs d'activités générant de grandes quantités de données sont les premiers utilisateurs de ce concept. Tous les secteurs d'activités sont potentiellement concernés, certains utilisant d'ailleurs déjà de manière intensive les solutions de datamining/BI (finance, recherche, logistique, industrie, etc...).

**Indépendamment du secteur d'activité, les technologies Big Data peuvent permettre de créer de nouveaux services, d'améliorer et d'optimiser des processus, produits ou services existants, ou de créer de nouveaux modèles économiques.**

**Les secteurs pionniers sur le Big data sont d'ailleurs des secteurs « BtoC », tournés vers les utilisateurs finaux**

- *Le e-commerce*
- Les télécoms
- La banque/assurance

**Les premiers *usecases* sont d'ailleurs essentiellement des préoccupations marketing**

En effet, l'essentiel de nos interlocuteurs (tant entreprises que salariés) ont souligné que les cas d'usage se focalisaient essentiellement sur l'analyse des scores d'attrition (*churn*) et la connaissance des comportements des consommateurs.

La vente et le marketing génèrent beaucoup de données, notamment sur les clients : types de produits ou services achetés, lieu d'achat, profil démographique, etc. L'analyse de ces données peut apporter des informations clés aux entreprises.

Dans le domaine de **la relation client**, les éditeurs de logiciels permettant de gérer les relations client (contact, facturation, prospection, marketing) intègrent de plus en plus des fonctionnalités d'étude de bases de données client. L'analyse de ces bases de données permet de mieux connaître le comportement et les attentes des clients, en intégrant notamment des données extérieures variées, comme celles issues des réseaux sociaux.

Les acteurs de la **publicité en ligne** font progressivement appel à des technologies Big Data pour traiter les informations qu'ils récupèrent sur les visiteurs. L'analyse individuelle et en temps réel de ces informations permet par exemple de générer des publicités ciblées qui s'afficheront sur une page web visitée par cet utilisateur.

La **vente au détail** peut bénéficier, dans les différents maillons de sa chaîne de valeur, des techniques Big Data. Au niveau marketing, les acteurs du e-commerce peuvent proposer des recommandations de produits en temps réel à leurs clients en se basant sur les produits vus, les achats effectués, et sur leur profil (âge, adresse, sexe, ...). Plus généralement, l'étude du comportement des clients dans le magasin (déplacements,



attitudes, action sur le produit) permet par exemple d'optimiser le placement des offres promotionnelles dans le magasin.

### Mais les champs d'application à imaginer sont infinis

De nombreux champs d'application sont imaginables, dans tous les univers professionnels. Par exemple :

- **Services Internet grand public**

Les acteurs Internet manipulent de fait de très grands volumes de données non structurées. Ils se sont concentrés dans un premier temps sur des données non personnelles, à travers le croisement de données de crawling du web. Ils s'intéressent désormais fortement aux données personnelles à des fins publicitaires ou e-commerce (cf ci-dessus).

Certains **moteurs de recherche** sont d'ores et déjà basés sur ce type de technologie. Ils font appel notamment à des techniques d'extraction très rapides et en temps réel des résultats d'une recherche pour permettre un affichage instantané. En ce qui concerne la recherche d'image, des techniques de reconnaissance d'images ou vidéos couplées à des technologies Big Data pourront émerger dans les années à venir, pour remplacer les systèmes actuels qui se basent, pour la plupart, sur l'analyse du contexte de l'image/vidéo (texte autour de l'image, solutions speech-to-text, transformant les paroles en textes).

De la même façon, la plupart des systèmes de **traduction** quasi instantanée de documents disponibles sur internet s'appuient sur des traductions différentes de mêmes documents et font une analyse statistique des résultats.

- **Recherche scientifique**

La recherche scientifique est un segment majeur d'application du Big Data. Des pans entiers de la science reposent en effet sur l'exploitation de grands volumes de données pour y trouver des corrélations empiriques.

Au niveau **scientifique**, les applications sont importantes, notamment dans la recherche. En **géologie**, les chercheurs doivent traiter les données apportées par plusieurs milliers de capteurs répartis sur différents points du globe. En faisant une analyse de ces informations, les chercheurs peuvent prévoir l'éruption prochaine d'un volcan ou l'arrivée d'un séisme. Les **météorologues** doivent eux aussi gérer des sources et des quantités d'informations très importantes, notamment pour établir des prévisions météo fiables, à plusieurs jours. Les données proviennent de multiples sondes et doivent être couplées à des images satellites, en prenant en compte plusieurs paramètres (vent, température, humidité, pression).

Dans le milieu médical, les généticiens ont besoins de grandes capacités de calcul pour traiter les données ADN et **génétiques** d'organismes vivants.

L'agence spatiale européenne (ESA) a également utilisé des technologies Big Data pour cartographier la galaxie en traitant les données issues d'un satellite.



- **Secteur public**

Dans le secteur public, que ce soit au niveau national ou niveau des collectivités territoriales, de nombreuses données existent, et peuvent être utilisées et valorisées par des technologies Big Data. Il peut s'agir de données « personnelles » liées aux citoyens (impôts, santé, emploi, infractions, ...), ou autres (géographiques, financières, démographiques,...). Les données personnelles devant être d'abord anonymisées pour être utilisables.

Plusieurs usages du Big Data peuvent être envisagés auprès des administrations publiques et notamment l'**optimisation et amélioration de l'efficacité** des tâches effectuées dans les collectivités. A ce stade, les exemples concrets restent rares, les déploiements étant peu nombreux et restant souvent confidentiels.



0. Rappel des objectifs et de la méthode
1. Des innovations technologiques...
2. ...ouvrent de nouvelles fonctionnalités et usages pour les entreprises

---

### 3. Un sujet d'avenir, en émergence, autour duquel de nouvelles offres se développent

- 
4. L'impact de ces bouleversements sur les modes de travail entre clients et prestataires
  5. Les besoins en compétences liés au Big data
  6. Les enjeux en termes de formation liés au Big data
  7. Quantification des volumes d'emploi liés au Big data
  8. Cartographie et analyse de l'offre de formation
  9. Préconisations
  10. Annexes



---

## Quelques barrières structurelles au développement du Big Data

---

### Un changement (une révolution ?) nécessaire autour de la vision de l'usage des données

Le plus gros frein semble plutôt d'origine **culturelle**. Alors que les entreprises sont habituées à des données en silos en vue de garder un certain « pouvoir », le Big Data ne prend tout son sens qu'avec la mise en place de la **centralisation/le décloisonnement des données** (dans les data lakes). Le poids du chef d'entreprise et la mise en place d'une organisation de type CDO sont toutefois de nature à contrer ces freins culturels via une gouvernance à un niveau élevé. Les directions métiers et les DSI freinent souvent les projets, qui remettent en cause l'existant et les règles métiers.

Le Big Data reste par ailleurs **encore mal connu**, souffrant notamment de l'amalgame évoqué auparavant entre les différentes solutions traitant des données (BI, etc...). Ces possibilités sont aussi encore mal appréhendées par les potentiels utilisateurs.

### De nombreuses questions autour du cadre juridique

Les entreprises ne sont pas toujours pas au clair sur les traitements qui peuvent être mis en œuvre, notamment pour le développement de nouveaux modèles économiques. Au-delà du cadre juridique, les réticences des clients sur l'utilisation de leurs données freinent certains développements. Les entreprises vont devoir revoir leurs processus de collecte d'informations (notamment autorisations d'utiliser les données).

### Des performances technologiques parfois insuffisantes

La **capacité de traitement en temps réel** apparaît comme un plus dans l'aide à la décision que permet le Big Data. Toutefois, les solutions comme Hadoop ne répondent pas toujours à ces contraintes.

L'offre Big Data doit aussi s'étoffer dans la **capacité d'indexation** des données non structurées afin de rendre l'information recherchable pour intégrer dans les processus métiers.

Néanmoins, l'offre évolue vite et de nombreuses solutions sont déjà en place et satisfaisantes, y compris pour aborder les problématiques de sécurité.

### Des modèles économiques mal définis pour des projets essentiellement au stade du POC

Nos interlocuteurs n'ont pas été en mesure de nous citer des projets « big data » déployés en France. Tous travaillent sur des **POC (proof of concept), mais avec encore peu d'industrialisation**. Il n'y a donc pas encore de démonstration de retour sur investissement dans de nombreux cas (mauvais choix de business case). Le marché apparaît nettement plus développé aux Etats-Unis.





Les utilisateurs les plus avancés en termes de BI et data warehouse peuvent aussi rencontrer des difficultés pour **identifier à court terme des modèles économiques nouveaux**, puisqu'ils utilisent déjà de manière intensive les outils d'analyse des données. L'explosion en termes de données (autour notamment de l'IoT) devrait changer la donne, même si celle-ci reste encore modérée.

Les difficultés à identifier de nouveaux modèles économiques sont par ailleurs renforcées par les coûts pour mettre en œuvre le Big Data, notamment au point de vue humain. Les ressources qualifiées (notamment data scientists) sont rares et très chères, faute de formation encore adaptée.

### **Manque de compétences internes**

Les solutions Big Data proposées à l'heure actuelle ne sont pas toujours des offres clés en main (il s'agit dans la plupart des cas d'une intégration d'Hadoop autour des logiciels existants de BI et datamining, ainsi que d'une intégration dans des solutions de cloud). Comme évoqué ci-dessus, les ressources humaines qualifiées capables de gérer de telles solutions sont trop limitées et une pénurie est attendue dans les prochaines années.



---

## Vis-à-vis du Big data, trois « types » d'entreprises peuvent être distingués

---

### **Type 1 : Les entreprises qui créent<sup>2</sup> de la donnée, mais ne l'exploitent pas (encore)**

Ces entreprises constituent un immense vivier de prospects pour les entreprises de la branche.

### **Type 2 : Les entreprises qui créent de la donnée et l'exploitent**

Ces entreprises sont à l'heure actuelle celles qui font ou sont susceptibles de faire appel aux entreprises de la branche pour les accompagner dans leur transformation numérique et dans la valorisation de leurs données

Ces entreprises peuvent :

- soit gérer ces projets elles-mêmes (pilotage et internalisation complète des projets « big data »)
- soit faire appel à des prestataires pour les accompagner dans ce virage et dans la valorisation de leurs données

### **Type 3 : Les entreprises qui ne créent pas de données, mais les exploitent**

Il s'agit des entreprises entrant pleinement dans le périmètre de cette étude (c'est-à-dire les prestataires des entreprises de type 1 et 2)

---

<sup>2</sup> On parle ici de données créées au sens capture de données. Il est à noter que la capture de données peut être d'origine interne ou externe (récupération de données auprès de tiers, via notamment des APIs)



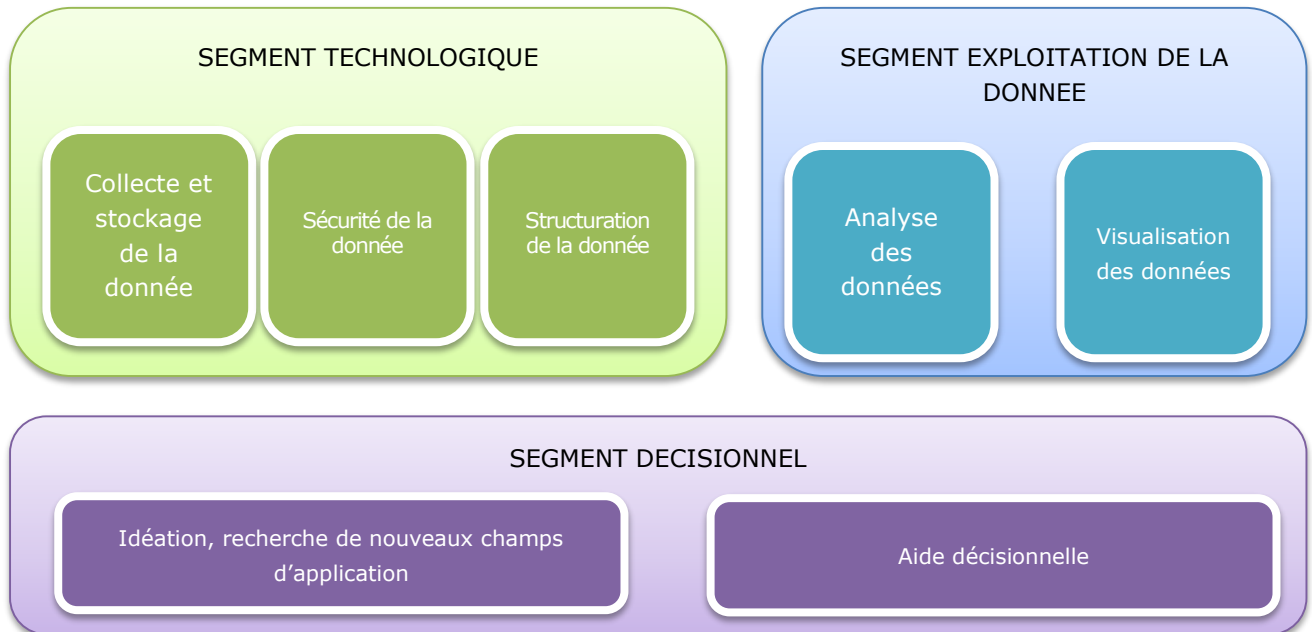
---

## Une offre qui va au-delà de l'informatique pure

---

Les prestataires de Big data (type 3 de la segmentation précédente) se retrouvent sur différents segments :

- Sur le **segment technologique**, on retrouve notamment les fournisseurs d'infrastructures. Dans la mesure où les infrastructures s'appuient essentiellement sur des infrastructures cloud, on retrouve grosso modo les mêmes acteurs. Les technologies clés mises en œuvre incluent notamment Hadoop et la programmation distribuée, NoSQL et le machine learning.
- Sur le segment « **exploitation de la donnée** », on retrouve une multitude d'acteurs, d'univers très différents :
  - ceux qui interviennent sur la partie technologique et qui proposent également des solutions d'exploitation et d'analyse (qui semblent évoluer vers un modèle de Big data as a service)
  - des sociétés d'études et de conseil, qui vont croiser les données pour les faire parler
  - une multitude de start up, se positionnant sur ce créneau
  - les services internes des entreprises utilisatrices (bien souvent des équipes de statisticiens, au sein des directions métiers)
- Sur le **segment « décisionnel »**, des prestataires se positionnent également (essentiellement des cabinets de conseil en management, qui se saisissent des sujets digitaux, mais n'interviennent pas uniquement sur le big data)



Alors que la mutation vers le cloud entraîne une évolution des fournisseurs et des offres (cf. rapport cloud computing), on ne peut pas parler d'évolution structurelle pour le Big data :

- Sur la partie technologique, les infrastructures reposant sur le cloud, le lecteur est invité à se tourner vers le rapport dédié au cloud computing qui explicite les évolutions à l'œuvre
- Sur la partie analyse et exploitation des données, les prestataires se sont en fait dotés d'un domaine d'activité nouveau (il s'agit d'une prestation complémentaire), mais demeurent dans un système transactionnel avec leurs clients (je vends un service)
  - Le changement notoire signalé par nos interlocuteurs est que les projets sont plus petits (quelques mois) et ainsi de facturation moindre que les projets SI classiques



---

## Les principaux acteurs impactés par ces évolutions

---

### Transformation numérique à tous les étages de l'entreprise

Le Big Data est un **élément moteur de la transformation numérique des entreprises**. Comme son nom l'indique, il s'agit toutefois plus d'une **évolution des règles métiers que des métiers eux-mêmes** au sein des entreprises. Les métiers vont de plus en plus intégrer les données disponibles (et les traitements associés) dans leurs prises de décision.

De même, le changement de relation avec les prestataires n'est pas évident (au-delà des aspects déjà évoqués pour le cloud). De nombreux prestataires seront surement nécessaires pour gérer la complexité des solutions.

### Fournisseurs/activités menacés

La DSI pourrait voir son rôle amoindri et recentré sur la fourniture du socle d'infrastructure basée le plus souvent sur le cloud (cf. rapport cloud). Elle pourrait toutefois récupérer de nouvelles fonctions liées à l'infrastructure (cf. ci-dessous).

Les activités les plus menacées sont celles autour de la gestion et du traitement/analyse de la donnée structurée que vient progressivement remplacer le Big Data. C'est notamment le cas des profils BI, gestion de bases de données et des data business analysts.

### Activités en développement

Le Big Data implique le développement d'une multitude de nouveaux métiers et profils, notamment autour du data lake. (cf partie 4)



0. Rappel des objectifs et de la méthode
  1. Des innovations technologiques...
  2. ...ouvrent de nouvelles fonctionnalités et usages pour les entreprises
  3. Un sujet d'avenir, en émergence, autour duquel de nouvelles offres se développent
- 

#### 4. L'impact de ces bouleversements sur les modes de travail entre clients et prestataires

---

5. Les besoins en compétences liés au Big data
6. Les enjeux en termes de formation liés au Big data
7. Quantification des volumes d'emploi liés au Big data
8. Cartographie et analyse de l'offre de formation
9. Préconisations
10. Annexes



## De forts espoirs de la part des entreprises clientes qui, dans le même temps, se questionnent

### Le Big data, un défi pour l'entreprise toute entière

Le Big data **n'est pas un sujet confiné à la sphère de l'IT**, bien au contraire. Certes, comme expliqué supra, l'approche Big data est rendue possible par des innovations technologiques (capacité à exploiter des données non structurées, en temps réel, etc.). De plus, pour être en mesure d'exploiter ces « big data » et d'en tirer de la valeur, il faut d'abord les capter, les stocker, les qualifier, ce qui suppose une infrastructure sous-jacente (puissance de calcul, etc. donc de l'IT).

Néanmoins, à la différence du cloud computing (qui reste essentiellement un sujet « techno »), le Big data est un **sujet d'entreprise, un sujet décisionnel** : l'enjeu est de **révolutionner la manière dont l'entreprise fait du business**, en faisant des données une source de croissance potentielle. Il s'agit d'optimiser les processus, d'optimiser sa performance, définir la stratégie d'entreprise. L'ensemble des directions, y compris (notamment) la **direction générale**, sont impliquées (et non la seule DSI).

### Dans la plupart des entreprises, une gouvernance autour des données qui reste à mettre en place

Ceci étant dit, si les avantages possibles du Big data sont fréquemment évoqués (usecases emblématiques, médias, etc.), des **avancées considérables dans les entreprises sont encore nécessaires**, d'après certains de nos interlocuteurs.

D'une part, comme l'ont évoqué la totalité des entreprises utilisatrices interrogées, **casser les silos**, pour partager les données (création d'un data lake) implique une évolution radicale des cultures d'entreprises. Le partage des données n'est pas, historiquement, dans la culture des entreprises ; ainsi, une certaine frilosité, voire réticence au changement, a parfois été évoquée lors des entretiens.

Un autre défi concerne la **traçabilité** des données, alors que les enjeux autour de la sécurité des données augmentent (cf. ci-dessous). En effet, les données proviennent à la fois des entreprises elles-mêmes (entreprises qui peuvent être des multinationales) et surtout de sources externes ; elles peuvent donc répondre de législations différentes.

De nouveaux rôles doivent ainsi être définis au sein des entreprises clientes : qui pilote la donnée ? qui en assure la traçabilité ? etc.

En la matière, les solutions retenues par les entreprises sont protéiformes : mise en place d'une fonction Chief data officer ; rôle dévolu à la DSI, dont le rôle de data steward (responsable de la qualité des données) est renforcé, etc.

### Une augmentation des enjeux autour de la sécurité des données et de l'éthique dans leur utilisation

Du fait notamment de deux aspects :

- Les **données sont désormais génératrices de valeur**, les business-model peuvent être *data-driven* : aussi les données ont-elles un caractère profondément stratégique. D'où un enjeu renforcé autour de la sécurité de la donnée, pour en éviter les fuites.



- L'utilisation des **données personnelles** pose naturellement des questions de déontologie et d'éthique dans leur exploitation.

Ces deux aspects se trouvent d'autant plus renforcés que la **tentation du shadow IT** n'est jamais loin. En effet, on assiste à une tendance de fond (non spécifique au big data) où les directions-métiers s'autonomisent par rapport à l'IT, du fait notamment de solutions le permettant (cloud computing, cf. rapport dédié à ce sujet). Aussi, chaque direction-métier est-elle en mesure de mener un projet data « par elle-même », en ne considérant que les apports potentiels et faisant fi des risques associés (par exemple, des directions-métiers qui mettraient leurs données dans le cloud). Ce qui pose bien évidemment des enjeux en termes de sécurité de la donnée.





---

## A l'heure actuelle, des clients avant tout en phase d'exploration

---

### L'inconnu est en effet au cœur du Big data

Une expression, qui commence à être fréquemment usitée, lorsque l'on parle du Big data est qu'avec le Big data « *avant on trouvait ce que l'on cherchait, avec le Big data, on trouve ce que l'on ne cherche pas forcément* ». Cette expression illustre bien le fait que l'inconnu est au cœur du Big data : **il est par essence impossible de dire si la question business soulevée par l'entreprise lui apportera un retour sur investissement**. Le Big data est encore aujourd'hui une démarche exploratoire avant tout.

### De plus en plus de POC, mais peu de véritables projets Big data en phase d'industrialisation

Parmi nos interlocuteurs, aucun n'a été en mesure de nous évoquer un projet qui aurait dépassé le stade du POC (proof of concept).

### Une nécessaire évangélisation

Au-delà du buzz autour du Big data, nos interlocuteurs se sont accordés sur le fait qu'un travail d'évangélisation pour convaincre de l'utilité du Big data restait à mener.

Cette évangélisation est d'autant plus nécessaire pour « gommer » les aspects techniques des projets Big data. Avant de mettre en place des infrastructures potentiellement onéreuses (cluster Hadoop, etc.), il s'agit de convaincre de l'intérêt du Big data en avançant dans une logique quick-win (démontrer que l'on peut obtenir des résultats rapidement, sans moyens démesurés).

L'enjeu est fort pour les prestataires : il s'agit de s'inscrire dans cette posture, d'une part, et de mener de projets dans des cycles très courts, d'autre part (point développé ci-après).



## En apparence, aucune transformation radicale de l'articulation entre clients et prestataires, mais des méthodologies qui doivent se réinventer pour répondre à des problématiques nouvelles

### Le Big data, comme une prestation complémentaire...

Les prestataires ont d'autant plus d'intérêt à construire et proposer une offre big data que, le ROI étant encore incertain pour les entreprises clientes, celles-ci peuvent avoir une plus **grande tentation de se faire accompagner par des prestataires externes** dans un premier temps (moindre prise de risque).

Par la suite et supposant une acculturation large au Big data et son adoption massive dans la plupart des entreprises, certaines entreprises pourraient avoir tendance à internaliser ce sujet, compte tenu de son caractère hautement stratégique. Néanmoins, d'après nos interlocuteurs, « les entreprises auront toujours recours à des prestataires », dans la mesure où l'activité ne pouvant être lissée, les entreprises feront toujours appel à des ressources externes. Toujours selon nos interlocuteurs, les mieux positionnés seront ceux qui auront investi ce créneau le plus tôt possible, car ils auront ainsi développé une capacité de benchmark (multipliant les expériences dans divers secteurs), leur permettant d'apporter des insights précieux à leurs clients.

Ainsi, à la différence du cloud qui redessine les lignes entre clients et prestataires (du fait d'une modification profonde des offres et des modèles économiques), le Big data apparaît comme une **prestation complémentaire**, que les prestataires ajoutent à leurs arcs.

### ... mais une prestation qui implique un mindset différent...

Il s'agit d'une **exploration permanente** : tout projet Big data est un pari, il n'est pas possible de savoir à l'avance quels en seront les résultats et le ROI pour l'entreprise (en dehors de quelques use cases type churn).

### ... et donc des méthodologies nouvelles...

- **Des projets Big data qui ne sont pas nécessairement « Big »**
  - En effet, selon nos interlocuteurs, dans la mesure où il est impossible de deviner la pertinence et surtout le résultat de tout nouveau business case, les méthodologies classiques de gestion de projet (expression de besoin, spécifications générales, spécifications détaillées, etc.) sont inadaptées et défaillantes.
  - Au contraire, il s'agit de « **démarrer petit pour, si possible, aller loin** ». Plutôt que de mettre en place des infrastructures lourdes pour collecter, stocker puis exploiter les données (des clusters Hadoop), il peut être plus pertinent de s'orienter vers des POC légers (développement d'applications légères, évitant de mobiliser trop de ressources) dans un premier temps, puis d'enrichir ultérieurement cette première version de l'appliquatif ainsi développé.
- **Le raccourcissement du cycle de vie des projets** et la **méthode agile**

De plus en plus, les clients souhaitent des résultats rapidement, et ce d'autant plus lorsque les besoins émanent des directions métiers ou de la direction générale



(comme c'est le plus souvent le cas). D'autre part, il est déterminant que les projets Big data s'inscrivent dans une logique de quick-win. Enfin, le big data étant par nature une démarche exploratoire, la mise en place des projets ne peut se faire que de façon itérative. Les équipes doivent donc travailler en mode agile y compris avec avec les équipes du client.

**Les équipes projets deviennent donc plus complexes et transversales.** Conséquence de la méthode agile, les équipes doivent être embarquées entre le métier (équipe du client), l'IT (architecte base de données, architecte infrastructure, chief protection officer chez le client) et les profils de data science (développeur, développeurs de modèles, data-designer, etc.).

**En outre, les interlocuteurs chez le client se multiplient :**

- les directions métiers sont impliquées
- mais surtout les services juridiques, avec un enjeu crucial autour de l'utilisation qui est faite des données.

Les prestataires doivent donc être en mesure de composer avec une multitude d'interlocuteurs non-techniques.

- Une logique **Devops**, dans la mesure où le Big data s'appuie souvent sur le Cloud  
En effet, les solutions cloud apparaissent comme les plus optimales pour répondre à la fois aux enjeux de stockage et d'analyse des données :
  - le stockage de données massives
  - le cloud permet l'ajustement des ressources et de la puissance de calcul, indispensable pour des traitements massifs de données volumineuses en temps réel (le traitement en temps réel est en effet générateur de forts pics d'activités).

La notion de stabilité importe énormément dans le cloud (pour des raisons évidentes de continuité et qualité de services) ; de fait les prestataires doivent intégrer une logique Devops, inhérente au cloud : il convient de penser ses développements avec la contrainte de la mise en production, les développeurs doivent donc travailler avec les équipes d'infrastructures.



0. Rappel des objectifs et de la méthode
  1. Des innovations technologiques...
  2. ...ouvrent de nouvelles fonctionnalités et usages pour les entreprises
  3. Un sujet d'avenir, en émergence, autour duquel de nouvelles offres se développent
  4. L'impact de ces bouleversements sur les modes de travail entre clients et prestataires
- 

## 5. Les besoins en compétences liés au Big data au sein de la branche

---

6. Les enjeux en termes de formation liés au Big data
7. Quantification des volumes d'emploi liés au Big data
8. Cartographie et analyse de l'offre de formation
9. Préconisations
10. Annexes



## Les compétences recherchées

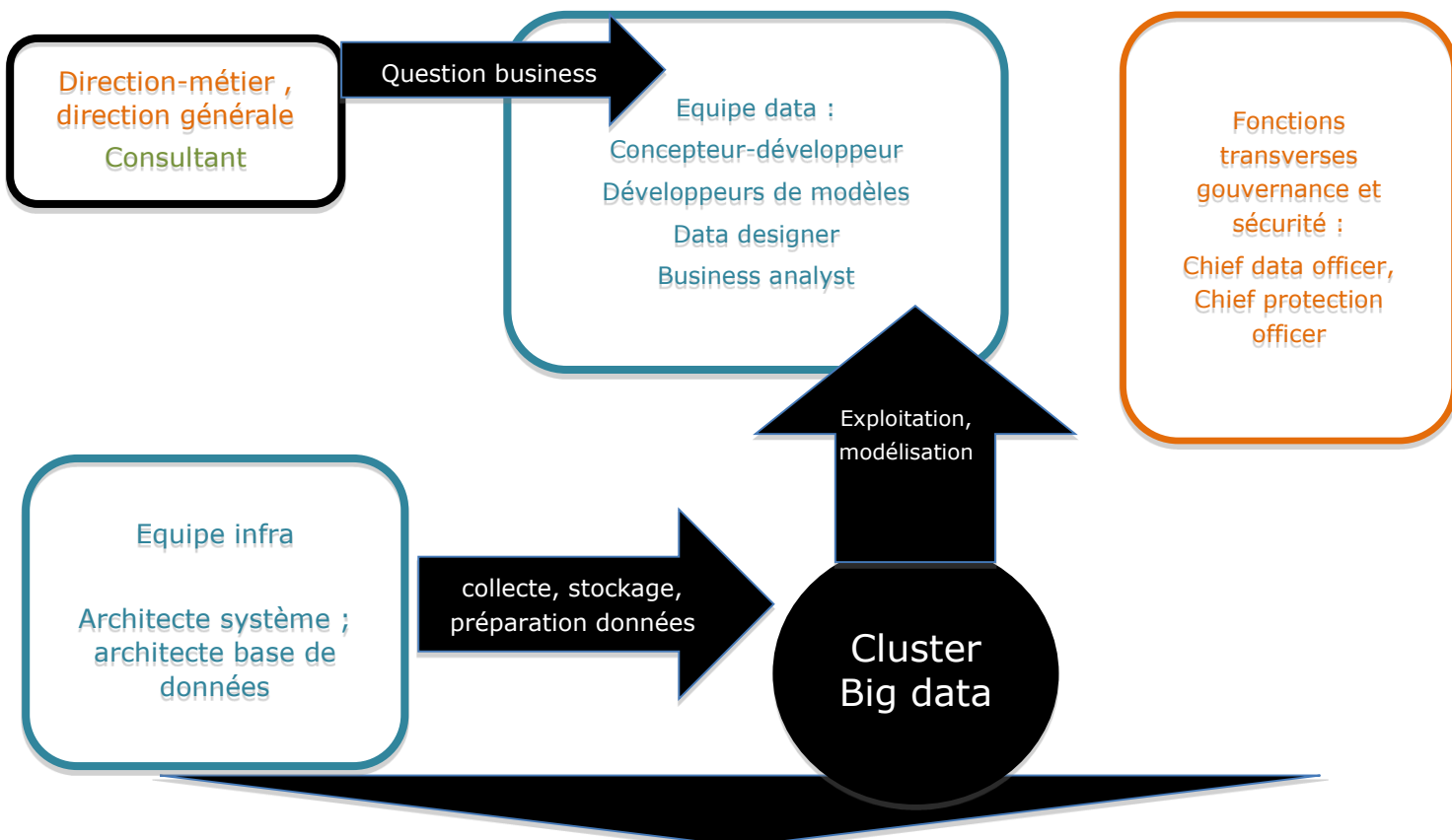
De nouvelles fonctions se créent, tant au sein des entreprises clientes que chez les prestataires

Certaines de ces fonctions sont communes aux clients et aux prestataires du numérique

En **bleu** sont indiquées les fonctions que l'on peut retrouver indifféremment chez les prestataires et leurs clients

En **vert** sont indiquées les fonctions que l'on retrouve majoritairement chez les prestataires

En **orange** sont indiquées les fonctions qui sont (ou ont vocation) à être au sein des entreprises clientes



Conséquences opérationnelles sur les différentes direction-métiers de l'entreprise (exemples)

Direction marketing : meilleure connaissance des comportements consommateurs, développement de nouveaux services

Centre de maintenance : nouvelle organisation, mise en place d'une maintenance préventive, résultante d'une meilleure prédiction des pannes

Vente : réorganisation des points de vente car compréhension fine des attentes client

Etc.



## Les métiers de l'infra : des évolutions « logiques » des équipes d'urbanisation et d'architectes cloud

**Bien souvent effacés derrière la popularité du data scientist, les métiers de l'infra sont fondamentaux pour que les données deviennent des « big data »**

Même s'ils outrepassent la sphère IT, les projets Big data nécessitent des équipes d'infrastructure pour définir les architectures les mieux à même de collecter, stocker, protéger et préparer les données. Ce sont ces équipes qui mènent la réflexion sur l'infrastructure à mettre en place pour porter les projets Big data, avec en ligne de mire les questions d'optimisation et de performance.

L'implication des équipes d'infra dans les projets Big data se renforcera lorsque les projets rentreront en phase d'industrialisation (mise en production des algorithmes de data science). Des compétences en virtualisation sont particulièrement recherchées.

### Essentiellement une évolution des compétences des profils IT

Il ne s'agit pas de nouveaux métiers, mais de métiers dont le rôle se renforce au sein des entreprises et qui doivent s'enrichir de **compétences supplémentaires, spécifiques au Big data.**

Les équipes IT (architecte cloud, ingénieur système et réseau, administrateur de base de données, développeur sur les couches basses, etc.) doivent intégrer des compétences complémentaires notamment liées à :

- la **maîtrise des environnements** Hadoop et MapReduce
- la **sécurité** des données (meilleur cryptage, etc.)
- les bases de **données non relationnelles** (NoSQL, technologies Cassandra, MongoDB)
- nouvelles **méthodes d'indexation** (ElasticSearch)
- les méthodes pour le **nettoyage et la qualification** des données (Pig)

En outre, comme expliqué supra, ces équipes d'infra doivent dorénavant être en capacité de **travailler de façon intégrée** (méthode agile) avec de nouvelles équipes (les équipes de data, cf. ci-dessous). Tout particulièrement, les équipes d'infra se repositionnent en **posture de conseil quant à la stabilité future des environnements** et modèles envisagés par les équipes de data (elles anticipent les phases d'industrialisation, que les équipes data n'ont pas nécessairement en ligne de mire au stade du POC).

### Ces métiers se retrouvent indifféremment chez le client et chez le prestataire, avec des différences toutefois selon la taille des structures

Au sein des **fournisseurs et éditeurs** de matériels qui conçoivent et proposent des solutions et environnement, ces profils se retrouvent tout naturellement (à noter que dans le cadre d'extension de leur offre vers du BDaaS, ils sont amenés à constituer des équipes data en interne).



Certaines **grandes ESN** ont mis en place une offre Big data, à côté de leurs prestations traditionnelles. Aussi retrouve-t-on ces équipes en leur sein, qui interviennent dans une posture de conseil IT. Néanmoins, selon les entreprises utilisatrices entendues, si ces compétences des équipes infra contribuent à un projet big data, certaines ESN seraient restées dans des schémas classiques de gestion de projet (expression de besoin, spécifications générales et détaillées, test et recette, etc.), peu adaptés aux projets Big data. Les projets Big data seraient, selon nos interlocuteurs, gérés uniquement comme des projets de mise en place d'infrastructures (et les grandes structures prestataires ou clientes auraient des difficultés à s'inscrire dans des schémas de pensée de test and learn).

Des **start up ou structures plus légères** ont un mode de fonctionnement différent :

- certaines impliquent fortement les équipes IT du client
- d'autres internalisent ces compétences, mais ne les spécifient pas au sein d'une équipe particulière : ce sont les mêmes personnes détiennent les compétences infra et les compétences data (cf. ci-dessous)



---

## Les métiers de la data : des expertises pointues

---

### Faire parler les données

D'après les entreprises, ce sont les métiers de la data qui concentrent le plus d'enjeux, car au cœur de la démarche et artisans de la manipulation des données.

### « Data scientist », une appellation souvent galvaudée qui regroupe des compétences complémentaires

D'une part, lors de certains entretiens a pu être observée une certaine tendance à catégoriser comme « data scientist » toute fonction dès lors qu'elle implique de travailler avec des données.

D'autre part, une tendance à ne considérer que les compétences en algorithmie et développement de modèle (et en occultant ce faisant d'autres compétences pourtant cruciales en infrastructure et datavisualisation) a aussi été constatée (les interlocuteurs n'en parlant pas sur le fond, alors que ces compétences nous également primordiales sur les profils de la data).

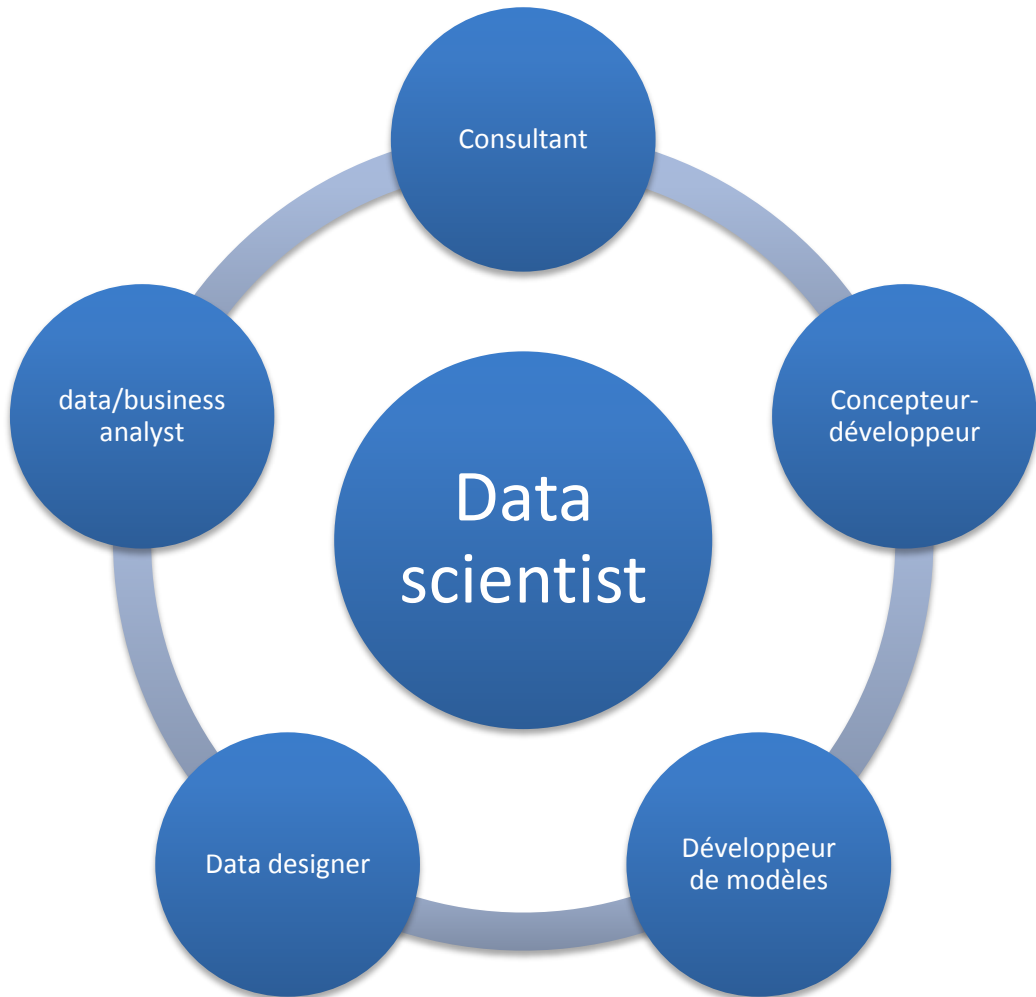
Si les métiers « de la data » ont en commun d'être des profils hybrides mixant des compétences mathématiques/statistiques, informatiques et, éventuellement business, il s'avère qu'au sein du « concept » de data scientist, des fonctions/rôles différents sont à distinguer. En ce sens, **le data scientist n'est pas nécessairement une personne, mais une équipe de compétences complémentaires.**

Selon la taille des structures notamment, l'ensemble de compétences/profils peut être réunis au sein d'une même personne. Par exemple, dans les start up, les profils peuvent être polyvalents, avec éventuellement des colorations plus ou moins prononcées sur certains domaines. Dans les structures plus lourdes, on peut voir des équipes avec des rôles bien définis (comme dans le schéma ci-dessous)

A noter que certains « métiers/fonctions » décrits ci-dessous existaient avant l'apparition du Big data (consultant et business analyst ; il peut s'agir de profils de la BI qui doivent développer des compétences complémentaires)

Les structures plus « lourdes » (et non forcément « matures » sur le Big data) ont parfois eu tendance à ne considérer que le côté « développeur de modèles » pour envisager les contours des compétences du data scientist.







## Consultant

- **Son rôle** : Il intervient vis-à-vis du client comme un « maïeuticien », en faisant émerger les business cases (champs applicatifs) possibles. Son rôle est d'accompagner le client dans sa réflexion pour « qu'il se pose les bonnes questions »
- **Les compétences-clés mises en avant par les entreprises** :
  - avoir une très bonne culture générale, capacité à s'immerger dans des univers professionnels variés, de saisir rapidement leurs problématiques et enjeux business ; être orienté usages et métiers
  - disposer de réelles capacités d'animation de réunion
  - maîtriser le design thinking et processus d'idéation
  - être un grand pédagogue
- **Les autres compétences nécessaires** :
  - disposer de bases techniques en infrastructure, avoir une approche fonctionnelle
  - disposer de connaissances en algorithmique (pour être en mesure d'explicitier l'utilité de tel ou tel type d'algorithme, pour identifier quel type d'algorithme pourrait correspondre à tel ou tel enjeu business)
  - être curieux : s'intéresser constamment aux nouveaux outils possibles, aux nouveaux usages
  - disposer d'une culture juridique de la donnée
  - savoir « lâcher du lest », c'est-à-dire accepter de renoncer à une part de rigueur mathématique/statistique lorsque cela permet de gagner en pédagogie et de clarté des messages

## Concepteur-développeur

- **Son rôle** : il fait le lien entre les équipes d'infrastructures et les développeurs de modèles ; il doit imaginer les solutions pour faire tourner les algorithmes dans un socle distribué.
- **Les compétences-clés mises en avant par les entreprises** :
  - disposer d'un très solide bagage scientifique (mathématiques, économétrie, statistiques) et en algorithmique
  - faire preuve de curiosité et « d'intuition » : être capable d'identifier les technologies disruptives, celles dont on parlera demain
  - maîtriser les approches NoSQL (Hadoop, MapReduce, etc.) ainsi que les techniques d'optimisation, préparation et nettoyage des données (par exemple, Pig)
  - maîtriser les différents types d'algorithmes
  - disposer d'un socle technique en architecture informatique et cloud, connaître les différentes architectures possibles, mais être capable d'imaginer leurs apports/faiblesses pour supporter et prendre en charge les développements envisagés
  - maîtriser un ou plusieurs langages de programmation (notamment sur les couches basses, par exemple C), connaître et savoir utiliser les bibliothèques
- **Les autres compétences nécessaires** :



- être « ouvert d'esprit », c'est-à-dire ne pas s'enfermer dans la technique (quand bien même celle-ci est prédominante à ce profil), pour être en mesure de comprendre ce que font les autres personnes de l'équipe

### Développeur de modèles

- **Son rôle :** il met en place et/ou optimise les algorithmes de traitement des données, pour les faire parler, de façon prédictive.
- **Les compétences-clés mises en avant par les entreprises :**
  - disposer de solides connaissances, voire d'une expertise, en mathématiques appliquées, statistiques
  - être expert en algorithmique et machine learning (apprentissage automatique)
  - être stratège, c'est-à-dire être capable de s'approprier les enjeux business afin de déterminer si telle ou telle approche est pertinente / envisageable / à imaginer (être capable de déterminer quand et comment appliquer tel ou tel algorithme), dans un objectif de création de valeur autour de la donnée et pour le client
  - savoir utiliser et mobiliser les outils classiques d'implémentation des algorithmes (Python, R)
  - savoir utiliser les technologies d'exploitation de données, de calculs distribués (par exemple, Map Reduce, Hadoop, Spark, etc.)
  - être en mesure de travailler avec les bibliothèques en ligne (par exemple, xgboost, etc.)
  - savoir coder, maîtriser un ou plusieurs langages de programmation (plutôt sur les couches hautes, par exemple, R ou Python)
  - comprendre les fonctionnements SQL et NoSQL
- **Les autres compétences nécessaires :**
  - être « astucieux » (pour par exemple, optimiser les développements de modèles : reprendre des briques existantes déjà codées pour les optimiser, plutôt que de tout créer de novo)
  - posséder une vue globale d'une infrastructure informatique, pour « comprendre comment et surtout si son algorithme est capable de tourner »
  - être curieux, s'intéresser aux technologies émergentes

### Data designer

- **Son rôle :** c'est un technico-communicant, qui est capable de représenter les données de façon pédagogique et parlante, de manière à ce que des décisions puissent être prises. Il n'est pas
- **Les compétences-clés mises en avant par les entreprises :**
  - avoir le sens de l'esthétisme
  - des connaissances en ergonomie visuelle et mise en scène des données (théorie des couleurs, storytelling, etc.)
  - faire preuve de communication et de pédagogie
  - maîtriser les outils de dataviz (par exemple, Tableau, etc.)



- **Les autres compétences nécessaires :**

- connaître les différents types d'algorithmes, savoir les optimiser afin d'améliorer les visualisations (par exemple : tableau, Spotfire, etc.)
- savoir utiliser les différents outils et langages utiles dans l'exploitation des données
- savoir manipuler des bases de données SQL et NoSQL
- être curieux, s'intéresser aux outils nouveaux (Vowpal Wabbit, xgboost, etc.), notamment ceux de l'open source

Business analyst (peut être « englobé » avec le consultant)

- **Son rôle :** il décrypte finement les enjeux business et contribue à définir la meilleure approche pour traiter la problématique client
- **Les compétences-clés mises en avant par les entreprises :**
  - être stratégique, décrypter très rapidement les enjeux d'un secteur, comprendre ses problématiques particulières
  - être pédagogique et communicant
  - sans être expert, avoir des connaissances en statistique, machine learning, algorithmique, outils IT de traitement des données pour être en mesure de préconiser une approche technique/scientifique adaptée à la problématique rencontrée
  - de fortes capacités analytiques, pour interpréter les données, en tirer des conclusions opérationnelles
- **Les autres compétences nécessaires :**
  - faire preuve de discernement dans l'utilisation et l'exploitation des données (en lien avec les questions d'éthique)
  - être curieux, notamment sur les sujets d'open data



## Une tendance de fond : des soft skills très attendues

### Apprendre à apprendre, curiosité et « débrouillardise »

Sans nul doute la compétence sur laquelle tant les entreprises que les salariés ont le plus insisté : le Big data étant un **champ nouveau, exploratoire**, les professionnels doivent être curieux, **constamment en éveil** et avoir « l'envie de découvrir de façon permanente ».

### Créativité

S'agissant d'un sujet par essence exploratoire, il importe aux entreprises de dénicher des **profils « astucieux », « qui n'ont pas peur d'essayer »**, dans une **logique d'essai-erreur**. La capacité à fréquemment se questionner et remettre en question ses réalisations a également été évoquée par les entreprises.

### Polycompétence

La recherche de polycompétences est une lame de fond quant aux professionnels du numérique du fait de plusieurs évolutions :

- la généralisation des principes **Devops** chez les clients induit une plus grande **capacité à travailler avec des équipes d'univers différents**.
- spécifiquement pour le Big data, les profils des équipes data notamment sont des **profils hybrides** (cf. supra), qui possèdent à la fois des compétences mathématiques/statistiques, IT et business.
  - La polycompétence sera d'autant plus grande chez les **prestataires de conseil** que leurs équipes data devront, par essence, être à même de travailler **sur différents secteurs** et donc de s'approprier rapidement les enjeux business de multiples secteurs donnés.

### Collaboration

- du fait du développement et du **recours aux « lib »** (open source), de la nécessité de participer aux communautés.
- mais aussi du fait que les expertises se renforçant, quand bien même les professionnels du Big data ont des profils hybrides (IT, statistiques, business), il sera **difficile de concentrer plusieurs degrés d'expertise au sein d'une seule et même personne**. Ainsi, les modes de travail seront **nécessairement collaboratifs** au sein d'équipes menant des projets Big data.

### Communication et pédagogie

Indispensable à triple titre :

- aujourd'hui, les projets Big data connaissent une phase **d'évangélisation** ; il est nécessaire d'avoir des profils avec des compétences en communication pour démontrer, convaincre de l'intérêt des Big data.



- des demandes qui émanent de **l'ensemble des directions-métiers** : aussi une capacité renforcée à expliquer, vulgariser, « rendre simple » des analyses et process auprès de « **non techniciens** »
- les résultats des analyses de données n'ont pas vocation à rester l'apanage de quelques initiés. En effet, le Big data, compte tenu de son rôle décisionnel, va impacter l'ensemble de l'entreprise. Il faut ainsi des compétences renforcées en pédagogie pour **diffuser les messages-clés** tirés, les enseignements à retenir de l'exploitation des données, ainsi que les actions à mettre en place.

### D'après certains de nos interlocuteurs, ces soft skills seraient en phase avec les valeurs des générations Y<sup>3</sup> et Z<sup>4</sup>.

- les **digital natives** seraient plus enclins à collaborer spontanément (les réseaux sociaux, la capacité à communiquer de façon permanente avec l'ensemble de la planète permise par l'internet, etc. expliquent sans doute cette « tendance naturelle »)
- une génération « indécise », ce qui n'est pas forcément négatif : parce que moins sûrs de leur choix, le **droit à l'erreur** (test & learn) est beaucoup **plus communément admis et accepté** par ces générations ; elle correspond en outre à une certaine culture anglo-saxonne, nord-américaine particulièrement, d'où provient aussi ce type de technologies

### Par ailleurs, on notera que ces soft skills ne sont pas spécifiques au Big data, mais sont plutôt la résultante d'une évolution globale du secteur numérique

- tant une **évolution technologique** : des technologies désormais open source
- qu'une évolution **organisationnelle** : évolution des méthodes de conduite de projet (méthode agile, devops)
- et une **évolution de l'IT au sein des entreprises clientes** (l'IT n'est plus un vase clos au sein d'une organisation, mais un fournisseur de services apportant un conseil à des clients internes : aussi davantage de collaboration avec les directions métiers)

<sup>3</sup> Déjà entrée sur le marché du travail (jeunes gens nés entre 1980 et 2000)

<sup>4</sup> Qui arrivera prochainement sur le marché du travail (jeunes gens nés depuis l'an 2000)



## Les enjeux et pratiques GRH

---

### Pour les entreprises de la branche, des enjeux GRH communs à l'ensemble des métiers

---

De façon générale, une pénurie de profils dans l'informatique, entraînant des difficultés de recrutement et une inflation salariale

La quasi-totalité des entreprises interrogées ont fait part de leurs **difficultés de recrutement sur les métiers de l'IT**, soulignant de façon générale, le manque de profils formés au **développement**.

Ces tensions sur le marché contribuent à la formation de bulles salariales, avec des salaires élevés, et ce, dès la sortie d'école.

#### D'où un fort enjeu autour de la fidélisation des profils

*NB : notons que les éléments ci-dessous ont été exprimés par nos interlocuteurs tant sur le cloud que sur le big data, et de façon plus générale, sur l'ensemble du secteur numérique.*

Fidélisation qui n'est pas nécessairement évidente pour les entreprises de la branche pour plusieurs raisons :

- une **course aux talents qui s'opère à l'échelle internationale**, aussi des professionnels ou jeunes très fréquemment et légitimement « happés » par l'attrait d'une opportunité à l'étranger
- des **profils qui « s'ennuient vite »** : selon nos interlocuteurs, il importe que les collaborateurs puissent changer rapidement de sujet, travailler sur des projets différents.
- à tort ou à raison, **l'image « ternie » des conditions de travail** au sein de la branche : certaines entreprises ont déclaré qu'à « un moment ou à un autre », les collaborateurs aspiraient à des conditions de travail « plus soft », des « horaires moins lourds », « moins de pics de charge ». D'ailleurs, certains candidats expriment ce souhait dès les premiers entretiens de recrutement (un recruteur : « on ne nous considère pas comme des entreprises au sein desquelles on souhaite faire carrière »). Selon certains de nos interlocuteurs, le peu d'avantages concédés par la CCN contribue à cette image défavorable des conditions de travail au sein de la branche. Parmi les salariés interrogés, certains ont également déclaré que « le conseil et les SSII, c'est parfait pour se faire les dents, mais c'est surtout une opportunité pour aller ailleurs ».



## Des enjeux et pratiques GRH tous particuliers sur les compétences data

### Au sein des entreprises, une possible immaturité de la fonction RH sur les « data scientists »

**D'une façon générale**, les directions ou responsables RH ont des **difficultés à recruter des profils dits techniques**, car ils ne sont pas toujours en mesure d'évaluer les capacités du candidat sur des sujets très opérationnels. Notons que ces difficultés ne sont pas spécifiques au Big data : elles concernent beaucoup de cas de recrutement dans l'informatique. Pour y faire face, des binômes RH/opérationnels sont mis en place.

Sur le Big data en particulier, cette difficulté semble se doubler d'une autre : **la « définition » du data scientist est parfois très floue au sein des entreprises**. Le data scientist est souvent considéré comme un poste à part entière (alors qu'il s'agit plutôt d'une réunion de compétences complémentaires, cf. ci-dessus). Hormis l'aspect « développement de modèles », « utilisation des algorithmes de machine learning » et « connaissance d'Hadoop », les autres « rôles » (équipe infra, data design, concepteur-développeur, etc.) au sein des équipes de data science sont occultés.

Plusieurs éléments sont sous-jacents à ce phénomène :

- il s'agit de métiers (et non d'un métier) nouveaux, et que les entreprises doivent donc découvrir, s'approprier et définir les contours
- Le big data semble être considéré par tous comme un « sujet d'avenir, en émergence », sur lequel de nombreuses entreprises souhaitent se développer, sans avoir nécessairement mûri leur offre au préalable

### Des relations-écoles très développées

La plupart des entreprises ont fait part de leur volonté (si cela n'est déjà fait) de se rapprocher et nouer des partenariats très étroits avec les cursus d'enseignements supérieurs qui ont mis en place des formations (sous forme de mastères spécialisés par exemple) dédié au Big data.

Ces partenariats prennent plusieurs formes :

- Sponsorship des formations (c'est-à-dire contribution à leur financement, sous forme indirecte : financement d'une chaire, etc.)
- Intervention pendant les formations, participation au programme pédagogique
- Proposition de stages aux étudiants suivants ces cursus

L'objectif pour les entreprises est « de recueillir les talents à la source ».

### Des modes de recrutement inédits

Au-delà des relations très fortes nouées avec les établissements de formation (cf. ci-dessus), les **plateformes de défis en ligne** (kaggle et datascience.net) constituent pour les entreprises un important vivier de recrutement. Ils présentent en plus l'avantage de tester directement les capacités des candidats.





Notons cependant que les défis sont le plus souvent proposés directement par les entreprises utilisatrices et ainsi les « génies ainsi débusqués » s'envolent le plus souvent vers ces entreprises (plutôt que vers leurs prestataires).

Cela étant, ces plateformes constituent malgré tout pour les prestataires une source pour faire de l'approche directe de candidats. En outre, leurs propres collaborateurs y participant souvent, cela permet de développer leur réseau et de favoriser la cooptation.

### **Des interrogations de la part des entreprises sur leur capacité à fidéliser ces profils**

Les entreprises ont exprimés, de façon générale, leur peine à fidéliser les profils. Il semblerait que cette crainte soit encore amplifiée pour les profils « data », dans la mesure où :

- Il s'agit de profils encore rares (car métier en émergence)
- « très scientifiques », ayant nécessairement une appétence prononcée pour « les développements, les trucs de petit génie », ces profils, selon certains de nos interlocuteurs, seraient susceptibles de « s'ennuyer encore plus vite que les autres ».

D'où un enjeu en termes de GRH, à savoir celui de l'arbitrage entre la cherté de ces profils (proposition de niveaux de salaires élevés pour fidéliser des profils encore rares sur la marché) et la variété des missions à leur confier (qui peut être un levier plus décisif que les arguments salariaux pour fidéliser ces profils).



0. Rappel des objectifs et de la méthode
  1. Des innovations technologiques...
  2. ...ouvrent de nouvelles fonctionnalités et usages pour les entreprises
  3. Un sujet d'avenir, en émergence, autour duquel de nouvelles offres se développent
  4. L'impact de ces bouleversements sur les modes de travail entre clients et prestataires
  5. Les besoins en compétences liés au Big data
- 

## 6. Les enjeux en formation liés au Big data

---

7. Quantification des volumes d'emploi liés au Big data
8. Cartographie et analyse de l'offre de formation
9. Préconisations
10. Annexes

A noter : certains enjeux relatifs à la formation (tant concernant les besoins que les pratiques) ont été exprimés par nos interlocuteurs, tant sur le Big data que sur le cloud, car communs aux deux sujets, et au-delà, concernent le secteur du numérique dans son ensemble.

Dans cette partie, nous signalons donc par un trait **bleu** dans la marge les éléments spécifiques au big data



## *Les besoins en formation*

---

### **L'évolution permanente : le dénominateur commun des métiers du numérique**

---

#### **Un univers d'évolutions permanentes et rapides**

C'est un poncif que de souligner que le secteur numérique est sans conteste un secteur qui connaît les mutations et avancées les plus significatives sur ces dernières décennies. En effet, l'innovation est son ADN, sa raison d'être : avancées technologiques, développement continu de nouvelles applications, recherche de nouveaux usages.

Par ailleurs, depuis son émergence dans les années 1970, le secteur a inversé la donne :

- parti d'une logique *pull* (les progrès informatiques étaient la conséquence de besoins croissants de la part des clients)...
- ...le numérique est aujourd'hui générateur de nouveaux usages (c'est-à-dire que ce sont ses avancées qui créent de nouveaux usages, dans une logique *push*)

**Ainsi, des professionnels par essence « malléables » et « adaptables »**

Les métiers du numérique ne cessent donc de se réinventer.

Comme l'a souligné la plupart de nos interlocuteurs, les collaborateurs du secteur sont constamment en phase d'apprentissage : les savoirs techniques deviennent rapidement obsolètes (nouveaux langages, nouvelles technologies qui doivent être intégrées, etc.).

En conséquence, les professionnels de ces secteurs voient leurs métiers constamment évoluer et se réinventer. Par exemple, les développeurs entrés en poste il y a ne serait-ce que 5 ou 6 ans exercent aujourd'hui un métier autre (nouveaux langages, nouveaux outils, etc.).



## Un socle technique complémentaire à développer pour les métiers de l'infra

### Selon nos interlocuteurs, une montée en compétences des profils architectes IT

D'après les entreprises interrogées, sur les équipes d'infrastructure, il s'agit surtout d'un **enrichissement des métiers d'informaticien que d'un nouveau métier** en tant que tel qui serait spécifique au Big data.

A noter cependant que l'enjeu compétences et, par répercussion formation, est de taille car **les profils « infra » doivent nécessairement muter du fait de l'adoption du cloud** (cf. rapport dédié à ce sujet), architecture sur laquelle reposent de nombreux clusters Big data. En revanche, ce besoin n'est pas induit par le Big data en tant que tel.

### De nouvelles technologies à comprendre et maîtriser

Sur la partie « Big data » de ces métiers, nos interlocuteurs ont notamment mentionné des besoins sur :

- la **maîtrise des environnements** Hadoop et MapReduce
- la compréhension des différences derrière les différentes distributions d'Hadoop (Hortonworks, MapR, Cloudera)
- les bases de **données non relationnelles** (NoSQL, technologies Cassandra, MongoDB) et les langages associés
- les nouvelles **méthodes d'indexation** (ElasticSearch)

### La nécessité de sensibiliser ces profils aux enjeux de l'éthique dans l'utilisation des données

S'agissant in fine d'exploiter et valoriser les données, il importe que les équipes d'infra, qui vont particulièrement être responsable de la collecte, du nettoyage et du stockage des données, soient sensibilisées aux enjeux éthiques dans l'usage de la donnée :

- Quelles données puis-je techniquement récupérer, traiter et stocker ?
- Quelles données ai-je le droit de récupérer, traiter et stocker ?
- Quelles données puis-je « moralement » récupérer, traiter et stocker ?



## Pour les métiers de la data, un socle d'expertise pointue en mathématiques appliquées et statistiques, qui doit s'enrichir au quotidien

### Des bases très solides en mathématiques appliquées, statistiques, économétrie

- Un **socle indispensable en mathématiques et statistiques**
  - Selon nos interlocuteurs, les métiers de l'analyse et de l'exploitation de la data reposent « sur un socle théorique en mathématiques, statistiques ou économétrie » (« il faut être matheux »). Tant selon les entreprises que les salariés interrogés, ces fondamentaux sont « stables », en cela qu'ils reposent sur des bases « académiques » (double intégrale, algorithmie, arbre de décisions, réseaux de neurones, etc.)
  - Ainsi, la formation initiale, d'après nos interlocuteurs, est pleinement légitime et a tout son rôle à jouer dans l'enseignement de ces fondamentaux.
- Des **formations initiales de qualité**, selon les entreprises
  - Nos interlocuteurs se sont en effet accordés sur la qualité des cursus scientifiques d'enseignement supérieur en France. Preuve en est de la fuite des talents vers l'étranger, notamment Etats-Unis, où les profils scientifiques formés au sein des écoles françaises sont très appréciés et recherchés.
- Une **appétence pour les cursus spécialisés**, mis en place récemment
  - Les entreprises interrogées suivent de très près la mise en place de cursus « Big data » « Data Science » « Machine Learning » qui se mettent progressivement en place au sein des établissements de formation initiale (cf. Cartographie et analyse de l'offre de formation). D'après nos interlocuteurs, ces cursus sont intéressants dans la mesure où ils permettent aux profils « matheux » de prendre en main les aspects Big data (quand bien même ils devront par la suite continuellement se former) et de nombreux enseignements sont appliqués (cas pratiques, développement d'un projet, etc.)

### Des outils qui évoluent constamment et que les professionnels doivent maîtriser

Le Big data repose en grande partie sur des environnements, langages, technologies **open source** (les « lib »). Aussi a-t-il un caractère **intrinsèquement communautaire et collaboratif** (cf. ci-dessous). Conséquence de cela, les compétences des collaborateurs sur les outils d'exploitation des data doivent constamment se renouveler, les professionnels des équipes data doivent être constamment à l'affût et s'intéresser aux technologies émergentes. C'est donc sur les compétences IT que les besoins en formation se renouvellent à grande vitesse. Ont par exemple été cités par nos interlocuteurs :

- La connaissance des langages de développement haut niveau (aujourd'hui, les plus utilisés sont Python ou R, mais demain, ces langages peuvent évoluer ou un autre peut s'imposer)



- Les technologies IT d'analyse de données (pour les calculs parallèles, Hadoop ou MapReduce font aujourd'hui référence, mais d'autres technologies émergent de type Spark)
- Se tenir au courant des librairies accessibles en ligne (xgboost, libfm, etc.) et de celles qui émergent

**D'autant plus que la data science et le big data sont des sujets nouveaux et donc amenés à évoluer énormément dans un futur proche**

Le Big data a fait couler beaucoup d'encre ces dernières années. Cet engouement médiatique a tendance à faire oublier que le Big data n'en est qu'à ses prémices et que l'essentiel de **son développement reste à venir**.

Aussi, ce nouveau champ est-il susceptible de connaître, dans un futur proche, des développements majeurs faisant évoluer d'autant le besoin d'expertise en la matière.



# Les pratiques en termes de formation

## L'aspect communautaire omniprésent, un challenge pour la formation

### L'avènement de l'Open source

Les développements et analyses Big data reposent quasi-intégralement sur les « **lib** » (technologies open source, librairies en ligne, etc.) dont les codes sont **co-construits** et **co-partagés** par des communautés très actives (cf. ci-dessous)

Les salariés interrogés ont tous insisté très fortement sur cet aspect fondamental.

### Des communautés très actives

L'**aspect communautaire est ainsi intrinsèque au Big data** : les langages sont complétés chaque jour par de nouvelles contributions des communautés en ligne ; les technologies open source ainsi que les librairies en ligne s'enrichissent de façon quasi-continue par les contributions de membres à leur code.

### La maîtrise de l'anglais

A noter ici que cela en appelle à un **renforcement des compétences linguistiques** de l'ensemble des collaborateurs (l'anglais étant la langue du développement et la langue pour interagir dans ces communautés). Or, de nombreuses entreprises ont souligné que les étudiants français, y compris dans l'enseignement supérieur, étaient « *encore à la traîne avec l'anglais* ».

### Ainsi, des besoins en formation permanents et évolutifs...

Les technologies utilisées changent tous les jours et de façon très rapide. Aussi, au-delà d'une prise en main de ces outils, les professionnels doivent constamment mettre leurs compétences à jour qui **évoluent quotidiennement**.

De plus, de nouveaux outils voient constamment le jour et, selon les professionnels entendus, il n'y a pas nécessairement de standard en phase de s'imposer comme une référence de marché. Par exemple, au cours des entretiens, si les technologies Python, R ou MapReduce ont souvent été citées, des outils « en émergence », « pas encore très utilisés » ne l'ont pas moins été : Vowpal Wabbit, xgboost, libfm, blaze.

Les professionnels doivent ainsi être en **veille continue sur les nouveaux outils**, être à l'affût des nouvelles librairies, pour être en mesure d'en saisir les nouvelles possibilités offertes par celles-ci.

### ... qui ne « rentrent » que très rarement dans la définition réglementaire de l'action de formation

Tant les entreprises que les salariés interrogés n'ont pas tant parlé de « formation » que de « **curiosité** » et « **d'investissement personnel** ».



En effet, compte tenu des évolutions permanentes évoquées supra, toute formation (entendue ici au sens classique de l'action de formation) se heurte à deux écueils :

- un **univers constamment mouvant**, aussi des formations obsolètes dès leur élaboration, avant même qu'elles soient dispensées.
- des **connaissances « collaboratives » et « communautaires »** : ainsi, le modèle d'une transmission de savoirs d'un expert vers des apprenants est suranné. La prise en main d'un outil, d'un langage, d'une technologie passe par l'intégration à une communauté et les échanges avec ses membres, échanges qui contribuent eux-mêmes à l'évolution de ces technologies.

Aussi, comme l'ont souligné les entreprises, il apparaît difficile de s'inscrire dans le cadre de programmes pédagogiques prédéfinis. C'est pourquoi, s'agissant d'acquisition des compétences et de formation, beaucoup de nos interlocuteurs (tant les entreprises que les salariés) ont parlé de « bricolage », de « bidouillage », de « curiosité », « d'auto-formation ».





## Ainsi, en matière de formation, tant les entreprises que leurs collaborateurs, s'aventurent hors des sentiers battus de la traditionnelle formation continue

### Bien entendu, tant les entreprises que les salariés ont exprimé leur appétence pour les MOOC

parmi les modalités évoquées par les entreprises et les salariés, il s'agit sans doute de la modalité la plus classique (alors même que les MOOC sont considérés comme des modalités avant-gardistes dans la plupart des autres secteurs). Les salariés peuvent ainsi accéder, de leur propre initiative et à leur gré, à un ensemble de contenus (proposés par des entreprises elles-mêmes, par des passionnés, etc. ex : Coursera)

### Mais, dans le même temps, d'autres modalités innovantes de « formation » se mettent en place

En effet, tant les entreprises que les salariés privilégient des modalités pédagogiques reposant sur l'esprit d'initiative des individus :

- les **forums** des communautés en ligne : selon les salariés interrogés, il s'agit du « meilleur moyen » de se former. Les professionnels interagissent de façon permanente, partagent les « astuces » et leurs dernières découvertes et co-innovent ensemble.
- la participation aux **meet up** : l'extension physique des forums communautaires.
- des **hackathons**, des défis, des compétitions
  - il peut s'agir de compétitions internes à l'entreprise (pour les entreprises, il s'agit de créer une émulation au sein de leurs équipes et d'inviter chacun à se dépasser)
  - tout comme il peut s'agir de défis ouverts à tous, via des plateformes comme kaggle.com ou datascience.net (qui sont également des viviers de recrutement exploités par les entreprises, comme détaillé ci-dessus). Les salariés interrogés ont unanimement déclarés être très friands de ces défis ouverts à tous, en cela qu'ils permettent de se confronter à d'autres spécialistes de la data science et de découvrir des astuces
- la **contribution aux codes** des communautés : certaines entreprises incitent leurs collaborateurs à contribuer aux codes (par exemple Openstack)

### Ces modalités innovantes s'affranchissent du cadre réglementaire des politiques de formation

De la même façon qu'un projet Big data ne s'inscrit pas dans la lignée d'un projet classique (car on ne connaît pas le résultat auquel on souhaite aboutir), ces modalités de formation **ne peuvent s'inscrire dans la définition de l'action de formation** (à l'exception des MOOC), telle que définie par la loi, pour qu'elle soit imputable (objectifs de formation, programme, etc.). En effet, comme l'ont exprimé les entreprises et les salariés interrogés, les apprentissages ne peuvent être figés dans un programme ou un quelconque référentiel.



## Toutefois, de façon plus classique, les entreprises mettent aussi en place des parcours de formation, qui ne peuvent cependant être des parcours-catalogues

Ces parcours sont généralement des **initiations** qui permettent aux collaborateurs de prendre en main les différents outils disponibles. Pour ce faire, les entreprises s'orientent vers des **formations sur-mesure**. En effet, compte tenu de l'évolution à vitesse grand V des outils et langages, ces formations ne peuvent être des formations « catalogues », en cela qu'elles **doivent constamment être réactualisées**.

En outre, nos interlocuteurs ont insisté sur le fait qu'il s'agissait surtout de formation de « **prise en main** » et que ces actions avaient davantage pour objectif « d'initier une démarche », c'est-à-dire de susciter la curiosité des salariés, pour qu'ils s'investissent ensuite dans les logiques évoquées supra (participation aux communautés, forums, etc.).

Néanmoins, sur les aspects autres que « techniques », des modules de formation plus classiques peuvent faire sens (selon nos interlocuteurs), par exemple, sur les enjeux de l'**éthique** dans l'usage et l'exploitation des données ; idem sur les **compétences linguistiques**.

## Un cadre de la formation professionnelle qui n'est donc pas toujours en phase avec les besoins et pratiques des entreprises de la branche et de leurs clients, notamment sur les formations d'adaptation au poste

- Notamment avec l'orientation sur les formations certifiantes. La nécessité de faire évoluer rapidement les référentiels de formations est contrainte en la matière par le dépôt de ces certifications au RNCP ou à l'inventaire CNCP :
  - Les certifications enregistrées voient leur référentiel de compétences figé pour plusieurs années (souvent 6 ans)
  - Une réactualisation est éventuellement envisageable (tous les 3 ans), mais cette période reste déjà bien trop lente pour les besoins des entreprises et de leurs salariés en la matière.
- Avant même la dernière réforme de la formation, les entreprises semblaient rencontrer des **difficultés à faire financer leurs formations qualifiantes/d'adaptation des compétences aux évolutions du poste de travail<sup>[1]</sup> par le FAFIEC**. En effet, les modalités innovantes et mieux adaptées aux besoins des entreprises n'entraient déjà pas dans le cadre des critères d'éligibilité de la formation professionnelle (faute d'objectifs pédagogiques clairement définis, d'un programme pédagogique, etc.) et donc n'étaient **pas ou trop difficilement imputables sur les financements du plan de formation des entreprises**.

En ce sens, les infléchissements apportés par la réforme de la formation vont dans le sens des pratiques déjà mises à l'œuvre par les entreprises :

<sup>[1]</sup> Par oppositions aux formations plus longues comme des passerelles ou des formations diplômantes pouvaient/peuvent toujours parfaitement entrer dans le cadre du financement de la formation professionnelle



- Nos interlocuteurs ont certes exprimé leurs difficultés vis-à-vis de la réforme de la formation professionnelle, notamment le « moindre accès aux financements » (car les financements mutualisés ont été réorientés vers le certifiant).
- Il n'empêche que, d'une certaine manière, l'autre grande évolution de la réforme (suppression du 0,9%) s'inscrit parfaitement dans les pratiques des entreprises évoquées ci-dessus. C'est-à-dire que les entreprises ont désormais une obligation de former les salariés (et non de financer la formation) et ont donc désormais toute latitude dans le « choix des armes » pour ce faire.



## Un phénomène semblable s'esquisse concernant la formation initiale

### Certes, des cursus de formation spécifiques se créent au sein des écoles et universités

En effet, compte tenu de l'expertise pointue attendue en mathématiques appliquées et statistiques, des établissements d'enseignement supérieur (notamment écoles d'ingénieur et d'économie) ont mis en place des **cursus dédiés Big data**, bien souvent accessibles en 2<sup>ème</sup> cycle (cf. partie suivante, Cartographie et analyse de l'offre de formation). Ces cursus reposent sur des **enseignements théoriques et appliqués en mathématiques, économétrie, statistiques** (en cela ils ne diffèrent que peu des cursus scientifiques « classiques ») auxquels s'ajoutent des initiations au développement et aux outils Big data (R, Python, MapReduce, etc.)

A noter que les entreprises reconnaissent la **grande qualité de ces cursus**, et entretiennent d'ailleurs des **relations très rapprochées** avec ceux-ci. Ces relations étroites sont de plusieurs ordres :

- **financier** : les entreprises contribuent au financement des chaires pédagogiques et au financement des cursus
- **pédagogique** : des professionnels s'impliquent dans la définition des programmes
- **recrutement** : des entreprises « sponsorisent » ces cursus, afin notamment de pouvoir recruter leurs étudiants (« *en moyenne, on est sur des promotions de 20-25 personnes ; chaque étudiant, au début même de son cursus a a minima 4 propositions d'entreprises qui l'attendent à sa sortie* »)

### Mais, dans le même temps, de plus en plus de personnes « s'auto-forment » grâce aux cursus pédagogiques disruptifs et collaboratifs qui voient le jour (et qui rencontrent un certain succès)

A l'instar des initiatives pour démocratiser l'enseignement et révéler des talents cachés (telle l'école 42 créée par Xavier Niel), des « cursus » ouverts au plus grand nombre possible sont créés. Par exemple :

- des **initiations partagées** le plus largement possible par les entreprises : dans une logique collaborative. Aussi, de nombreuses entreprises, start up notamment, mettent régulièrement à disposition de tous des vidéos, tutoriels, explications pédagogiques sur tel ou tel outil. La plupart de celles-ci sont accessibles gratuitement en ligne. Celles-ci s'inscrivent pleinement dans une logique d'open-innovation (je partage ce que je peux et, ce faisant, j'ai la possibilité d'améliorer ce que je sais, car la communauté va enrichir et faire évoluer mon produit).
- des sessions de formations **accessibles en ligne**, à des coûts modiques (par exemple : le Wagon).

De telles initiatives vont très certainement contribuer à former toute une population d'autodidactes.



0. Rappel des objectifs et de la méthode
  1. Des innovations technologiques...
  2. ...ouvrent de nouvelles fonctionnalités et usages pour les entreprises
  3. Un sujet d'avenir, en émergence, autour duquel de nouvelles offres se développent
  4. L'impact de ces bouleversements sur les modes de travail entre clients et prestataires
  5. Les besoins en compétences liés au Big data
  6. Les enjeux en termes de formation liés au Big data
- 

## 7. Quantification des volumes d'emploi liés au Big data

---

8. Cartographie et analyse de l'offre de formation
9. Préconisations
10. Annexes



---

## Les sources mobilisées pour quantifier les emplois liés au Big data en France

---

### L'enquête quantitative

Certaines informations nécessaires à la réalisation d'une quantification des emplois liés au Big data ne semblaient pas être disponibles dans les sources à disposition pour la réalisation de cette étude. Par conséquent, l'option de mener directement une enquête quantitative de terrain pour aller chercher ces informations directement auprès des entreprises a été retenue.

Un questionnaire a été élaboré et validé suite à sa première présentation lors du comité de pilotage du 30-09-2015, avec une volonté marquée de limiter le nombre de questions pour ne pas en rendre le remplissage fastidieux pour les entreprises.

Le questionnaire définitif apparaît en annexe du présent rapport.

- Le **Panel de diffusion** du questionnaire

Le questionnaire a été diffusé [dans une première vague](#) auprès des adhérents au FAFIEC sur le périmètre des entreprises du numérique, uniquement auprès des entreprises pour lesquelles un email de contact était disponible. Ces contacts sont généralement issus du service RH ou comptable des entreprises.

7800 emails ont ainsi été adressés aux entreprises, dont 16% n'ont pu être délivrés en raison d'adresses non-valides.

[Une seconde vague de diffusion, plus ciblée, a été réalisée auprès des adhérents du Syntec Numérique d'une part, et du CINOV-IT d'autre part.](#)

- **Résultats de l'enquête**

[La première vague de l'enquête](#) a été ouverte aux répondants pendant 10 jours<sup>5</sup>, dont 6 jours ouvrés, avec 2 relances intermédiaires.

[82 réponses qualifiées](#) ont ainsi été collectées, un résultat limité en termes de représentativité statistique.

[La décision a donc été prise de réaliser une seconde vague d'enquête auprès des adhérents des organisations professionnelles de la Branche, pour une durée de 10 jours supplémentaires, dont 8 jours ouvrés.](#)

[51 réponses qualifiées supplémentaires](#) ont été enregistrées suite à cette seconde vague.

[Au total, 133 réponses d'entreprises](#) constituent donc le panel de l'enquête.

**[Ces 133 entreprises<sup>6</sup> comptant plus de 52 300 salariés au total, le panel de l'enquête représente donc plus de 15% des salariés de la Branche.](#)**

Les caractéristiques des répondants à l'enquête quantitative étaient les suivantes :

---

<sup>5</sup> Sur ce type d'enquêtes quantitatives, les délais habituels laissés aux entreprises pour répondre sont généralement d'un mois minimum

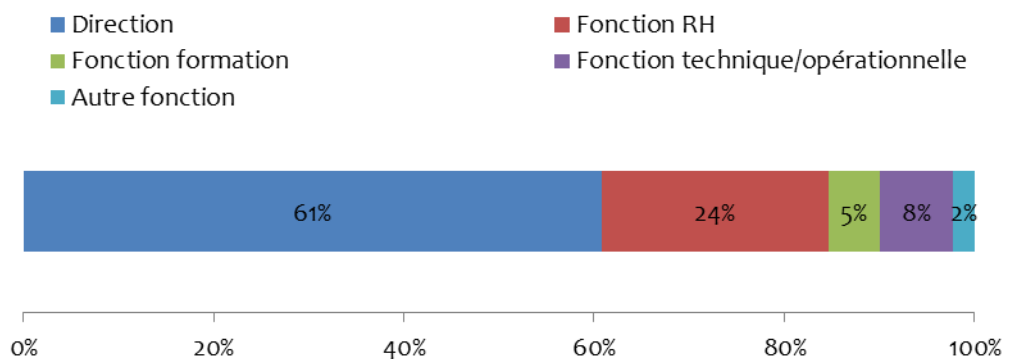
<sup>6</sup> [2 entreprises n'ont pas souhaité répondre à la question portant sur le nombre de salariés de leur structure.](#)



- 61% des répondants étaient directement les **dirigeants des entreprises**, notamment dans les plus petites structures, et 29% étaient des interlocuteurs issus des services RH des entreprises.

Dans l'ensemble, cette enquête reflète donc plutôt la vision des interlocuteurs RH des entreprises, car le dirigeant des TPE-PME est directement en charge de la GRH de son entreprise.

### Identité des répondants au sein des entreprises

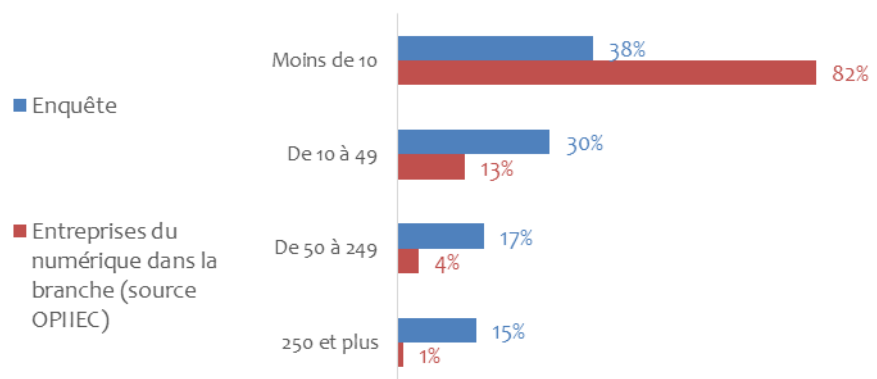


Source : Enquête auprès des entreprises  
Traitements : Ambroise Bouteille et Associés

#### Assiette : ensemble des répondants à l'enquête

Une répartition du panel de répondants assez équilibrée entre les différentes **tailles d'entreprises**, mais démontre une **sous-représentation des entreprises de moins de 10 salariés** dans l'enquête :

### Répartition des entreprises répondantes, en comparaison de la réalité de la Branche



Données : Enquête auprès des entreprises, OPIIEC  
Traitements : Ambroise Bouteille et Associés

#### Assiette : ensemble des répondants à l'enquête

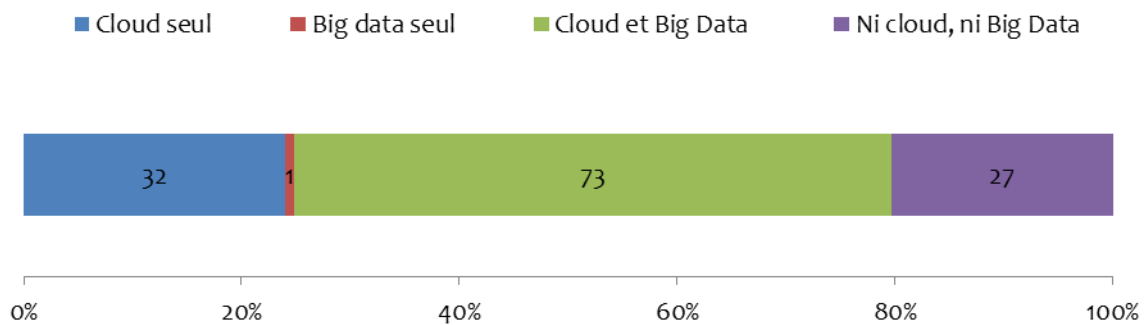


- Pour assurer la représentativité de l'enquête par rapport à la réalité de la branche, un **redressement des résultats** a été effectué sur la base de la réalité de la répartition des entreprises **par taille** d'après les données de l'OPIIEC.

L'ensemble des traitements issus de l'enquête présentés ci-dessous sont donc effectués sur les **réponses pondérées en fonction de la taille de l'entreprise** répondante pour assurer la représentativité du panel de l'enquête.

- Parmi ces 133 réponses, 74 entreprises **se sentent concernées** par la **problématique du Big Data** (déjà positionnées, ou envisageant de le faire) :

### Répartition des entreprises répondantes selon le positionnement actuel ou envisagé

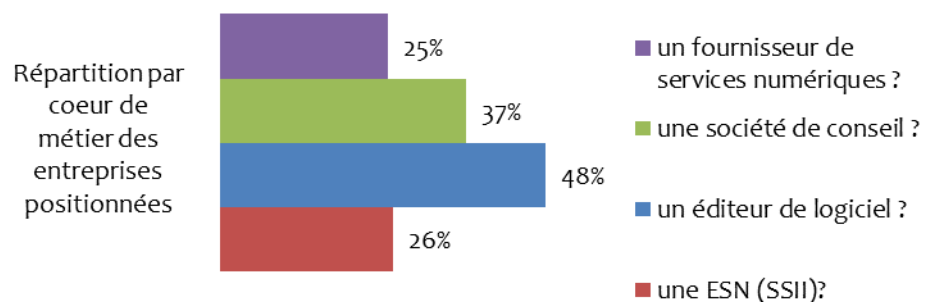


Source : Enquête auprès des entreprises  
Traitements : Ambroise Bouteille et Associés

Assiette : ensemble des répondants à l'enquête

- Les entreprises positionnées sur le Big Data semblent être majoritairement :

### Coeur de métier des entreprises positionnées sur le Big Data ou envisageant de le faire



Source : Enquête auprès des entreprises  
Traitements : Ambroise Bouteille et Associés

Assiette : répondants positionnés ou envisageant de se positionner sur des activités Big Data

- Les **indicateurs clés** issus de l'enquête :





- Répartition des emplois actuels de l'entreprise par famille de métiers ;
- Part des emplois de l'entreprise impactés par le Big data ;
- Prévion de cette part des emplois impactés à horizon 3 ans ;
- Impact de cette évolution de l'introduction du Big data dans l'entreprise par famille métier au cours des 3 prochaines années

### Les chiffres de l'étude sociologique 2014 de l'OPIIEC

Pour quantifier les emplois liés au Big data dans les entreprises de la Branche en France, il était nécessaire de disposer d'une estimation des emplois totaux dans les entreprises sur le périmètre de cette étude.

Pour cela, il a été retenu de se baser sur les données de l'étude sociodémographique de la Branche du numérique, de l'ingénierie, du conseil, des études et des métiers de l'événement – Rapport de synthèse secteur du numérique, du 18 septembre 2014, qui donne les chiffres les plus pertinents en termes de périmètre par rapport à celui considéré dans la présente étude.

- **Indicateur clé** issu de l'étude OPIIEC 2014 : le nombre de salariés des entreprises de la Branche en 2013

### Les hypothèses de croissance de l'emploi dans le secteur du numérique du CEP numérique

Pour compléter les données emploi de l'étude sociologique 2014 de l'OPIIEC, qui ne fournit les chiffres de l'emploi sur le périmètre du numérique que pour l'année 2013, il a été nécessaire d'estimer l'évolution de l'emploi total dans la Branche ; pour cela, les hypothèses d'évolution des emplois par activité présentées dans le CEP du secteur professionnel du numérique ont été retenue.

- **Indicateur clé** issu du CEP Numérique : projection de l'évolution des emplois par secteur du numérique entre 2012 et 2018



## Les tendances d'évolution de l'emploi lié au Big data dans les entreprises de la Branche en France

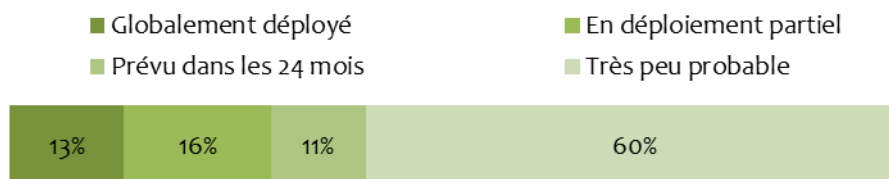
### 1 – Positionnement des entreprises du panel sur le Big Data

- A – Détermination du **taux de positionnement des entreprises répondantes sur une offre liée au Big data**

La première question a donc été de déterminer quelle part des entreprises de la branche sont positionnées aujourd'hui sur des activités de Big data en France, ou comptent se positionner dans les prochains mois et prochaines années.

Il apparaît qu'à la question « Le positionnement de votre entreprise sur le Big data ? », **29% des entreprises répondantes ont déjà déployé une offre en lien avec le Big Data<sup>7</sup>**. A noter que **11% supplémentaires des entreprises répondantes l'envisagent au cours des 24 prochains mois**. En revanche, une majorité des entreprises répondantes (**60%**) ne semble pas envisager –pour le moment- de déployer une offre en lien avec le Big data.

#### Positionnement des entreprises répondantes sur le Big Data



Source : Enquête auprès des entreprises  
Traitements : Ambroise Boutelle et Associés

Assiette : Ensemble des répondants

#### Plusieurs enseignements sont à tirer de ces chiffres :

- Il semble que les entreprises positionnées sur le Big Data soient surreprésentées dans l'échantillon. Nous sommes ainsi convaincus que les entreprises déjà positionnées sur le big data se sont particulièrement mobilisés pour répondre à ce questionnaire (car son annonce précisait cette thématique, elle-même très mobilisatrice).
- A l'inverse, il semble que les entreprises qui ne démontrent pas un intérêt particulier sur le sujet ont massivement renoncé à répondre. Ainsi le 60 % d'entreprises non positionnées nous semblent sous-estimer la réalité sur le périmètre de la Branche

<sup>7</sup> Une analyse extrêmement fine des réponses a été nécessaire pour assurer la plus grande réalité possible de ces réponses, en raison d'incohérences dans les réponses ( ) qui ont dû être corrigées. Les résultats présentés ci-dessus sont ceux issus de la base de réponses corrigée par le cabinet.



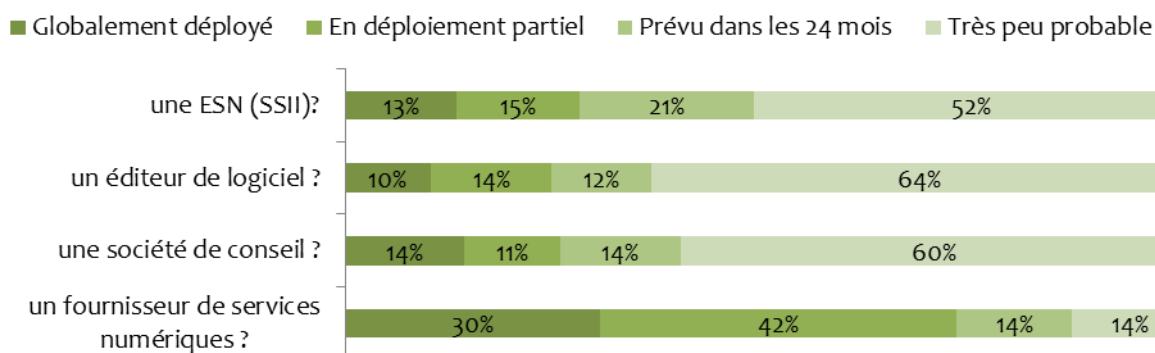
prise dans son ensemble ; il s'agit davantage du pourcentage des entreprises intéressées par le cloud qui ne feront pas de big data.

- Une telle proportion ne pourrait être déterminée avec précision qu'au sein d'une étude plus globale sur un sujet affectant l'ensemble des entreprises de la Branche, c'est-à-dire une étude où les entreprises positionnées et non-positionnées Big Data auraient la même propension à répondre à une enquête, ce qui ne pouvait pas être le cas ici par nature.

Nous avons donc fait le choix de donner les résultats de l'enquête sur l'échantillon des répondants, sans tenter d'extrapolation. Ces entreprises répondantes représentent pour rappel environ 15% des effectifs de la Branche.

Au sein de l'échantillon, les fournisseurs de services numériques semblent être les acteurs qui se sont le plus positionnés sur ce type d'offres dédiées au Big Data (50% aujourd'hui, 86% pourraient avoir déployés une offre d'ici 2 ans). Assez logiquement, les éditeurs semblent moins concernés, puis seuls ceux qui développent des logiciels de collecte, d'analyse, de gestion, etc. des données peuvent se sentir impactés :

### Positionnement des entreprises sur le Big Data selon leur coeur d'activité



Source : Enquête auprès des entreprises  
Traitements : Ambroise Bouteille et Associés

Assiette : Ensemble des répondants

- B – Pour les entreprises répondantes **positionnées** sur une offre Big Data, détermination de la **proportion de leurs emplois touchés**

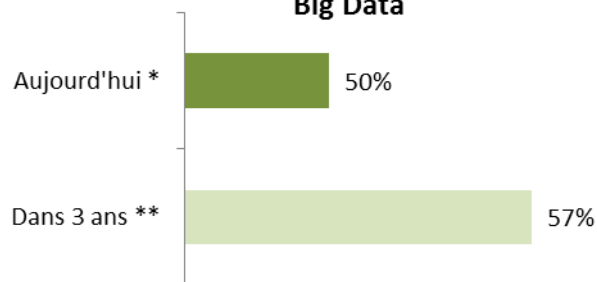
Dans un second temps, il a paru nécessaire de déterminer au sein de l'échantillon parmi les entreprises positionnées ou celles envisageant de le faire quelle proportion de leurs emplois étaient concernés par cette offre liée au Big data<sup>8</sup>, c'est-à-dire **quelle proportion de leurs salariés devaient maîtriser des compétences liées au Big data.**

<sup>8</sup> Question posée : « Quelle proportion de vos effectifs a intégré ou aura à intégrer significativement du Big Data dans leurs activités ? » aujourd'hui et à 3 ans.



Pour les entreprises de l'échantillon disposant dès aujourd'hui d'une offre Big Data, environ un emploi sur deux est touché par cette activité. Cette proportion devrait croître pour atteindre environ 57% des emplois des entreprises positionnées d'ici 3 ans.

### Part des effectifs impactés par le positionnement de leur entreprise sur le Big Data



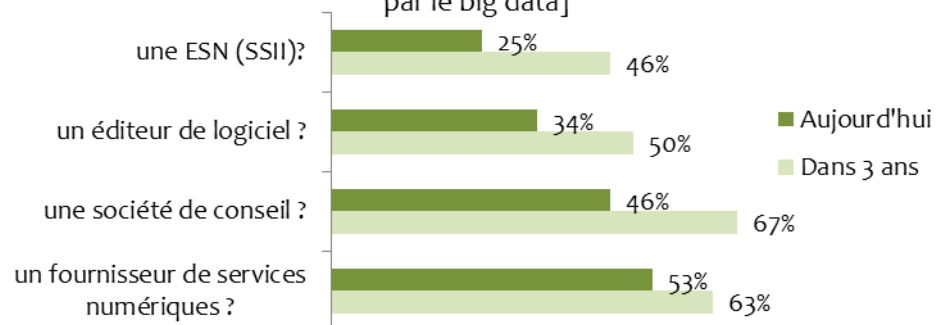
Source : Enquête auprès des entreprises  
Traitements : Ambroise Bouteille et Associés

\* Assiette : répondants déjà positionnés (globalement ou partiellement) sur des activités Cloud

\*\* Assiette : répondants positionnés ou envisageant de se positionner sur des activités Cloud

A assiette constante, sur l'ensemble des entreprises positionnées ou envisageant de le faire, on voit nettement cette progression des effectifs touchés quel que soit le secteur d'activité :

### Part des effectifs des entreprises répondantes touchés par le Big Data selon le coeur de métier de l'entreprise [assiette constante de l'ensemble des entreprises touchés par le big data]



Source : Enquête auprès des entreprises  
Traitements : Ambroise Bouteille et Associés

Assiette : répondants positionnés ou envisageant de se positionner sur des activités Cloud



- C - Détermination du **TCAM de la progression des emplois concernés par le Big data** dans les entreprises de l'échantillon

La progression globale de ces emplois touchés par le Big data au sein de l'échantillon serait de l'ordre de 24% par an selon les résultats de l'enquête.

Nombre d'emplois (Source OPIIEC)	Source	TCAM	2015	2016	2017	2018
<b>Echantillon total des répondants</b>	<b>Source Enquête</b>	1,1%	<b>52 400</b>	<b>53 000</b>	<b>53 600</b>	<b>54 200</b>
Part des entreprises positionnées sur le Big Data (répondant au questionnaire)	Source enquête	17%	<b>29%</b>	34%	<b>40%</b>	47%
Nombre total de salariés des entreprises répondantes positionnées Big Data	-	19%	15 200	18 100	21 400	25 500
Part des emplois touchés par le Big Data au sein de ces entreprises répondantes positionnées	Source enquête	-	<b>50%</b>	52%	55%	<b>57%</b>
Soit : nombre d'emplois touchés par le Big Data dans les entreprises du panel	-	24%	7 600	9 500	11 700	14 500

En comparant ces chiffres aux estimations de progression du chiffre d'affaires sur le marché du Big data<sup>9</sup> en France en arrive à une croissance annuelle assez similaire (Taux de Croissance Annuel Moyen de 24% pour les effectifs et de 23% pour le chiffre d'affaire), qui semble donc plausible :

Indicateur	Source	Unité	2015	2016	2017	2018
Valeur estimée du marché du Big Data en France	Estimation IDATE	M€	257	335	410	491

<sup>9</sup> Pris au sens strict du Big Data par opposition à la BI, c'est-à-dire sur l'exploitation de données non structurées, ce qui n'était pas le cas dans l'esprit de tous les répondants à l'enquête



## Les familles métiers impactées par l'émergence du Big data

### Détermination de la répartition des effectifs actuels des entreprises du numérique

Attention il s'agit des effectifs **totaux** des entreprises répondants, et **pas seulement ceux** touchés par le Big Data.

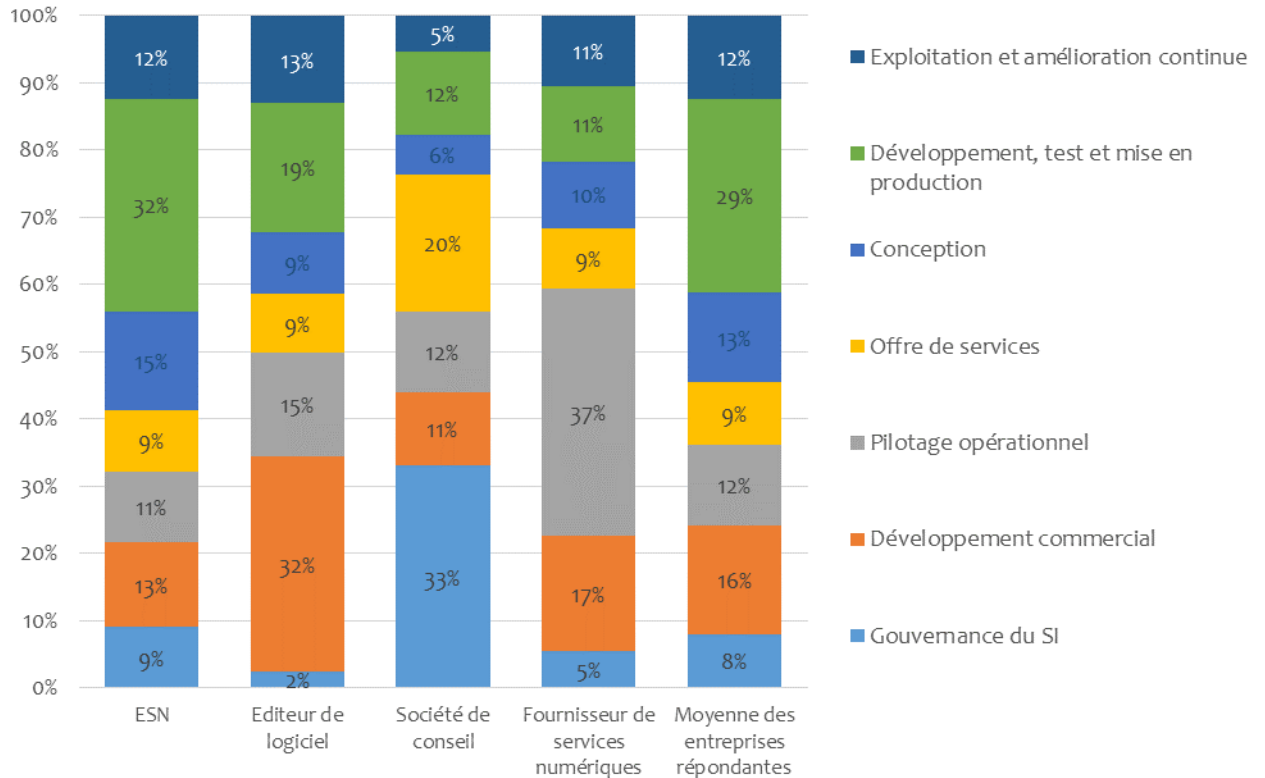
L'enquête auprès des entreprises a permis à celles-ci de nous faire remonter la répartition de leurs effectifs entre les familles suivantes :

- **Gouvernance du SI**, qui inclue des métiers comme Urbaniste, architecte technique, responsable de la sécurité de l'information, responsable juridique, etc.  
En moyenne pour les répondants, cette famille représentait **8% des emplois** de l'entreprise.
- **Développement commercial**, avec des métiers comme Commercial, consultant avant-vente, directeur commercial, etc.  
En moyenne pour les répondants, cette famille représentait **16% des emplois** de l'entreprise.
- **Pilotage opérationnel**, avec des métiers comme chef de projet, PMO, Service delivery manager, etc.  
En moyenne pour les répondants, cette famille représentait **12% des emplois** de l'entreprise.
- **Offre de services**, avec des métiers comme chef de produit/service, consultant métier, responsable des usages numériques, marketing, etc. En moyenne pour les répondants, cette famille représentait **9% des emplois** de l'entreprise.
- **Conception**, avec des métiers comme Consultant technique, designer, concepteur, architecte logiciel, etc. En moyenne pour les répondants, cette famille représentait **13% des emplois** de l'entreprise.
- **Développement, test et mise en production**, avec des métiers comme Développeur, intégrateur, chargé de référencement, administrateur d'infrastructure, consultant système et réseaux, administrateur de bases de données, data scientist, etc.  
En moyenne pour les répondants, cette famille représentait **29% des emplois** de l'entreprise.
- **Exploitation et amélioration continue**, avec des métiers comme responsable support, analyste d'exploitation, expert technique, conseiller support technique, business analyst, etc.  
En moyenne pour les répondants, cette famille représentait **12% des emplois** de l'entreprise.

Ces résultats cachent cependant (et logiquement) des disparités selon l'activité des entreprises :



## Répartition des effectifs des entreprises répondantes par famille de métiers



Assiette : Ensemble des répondants

### Estimation de l'impact du Big data sur les différentes familles métiers<sup>10</sup>

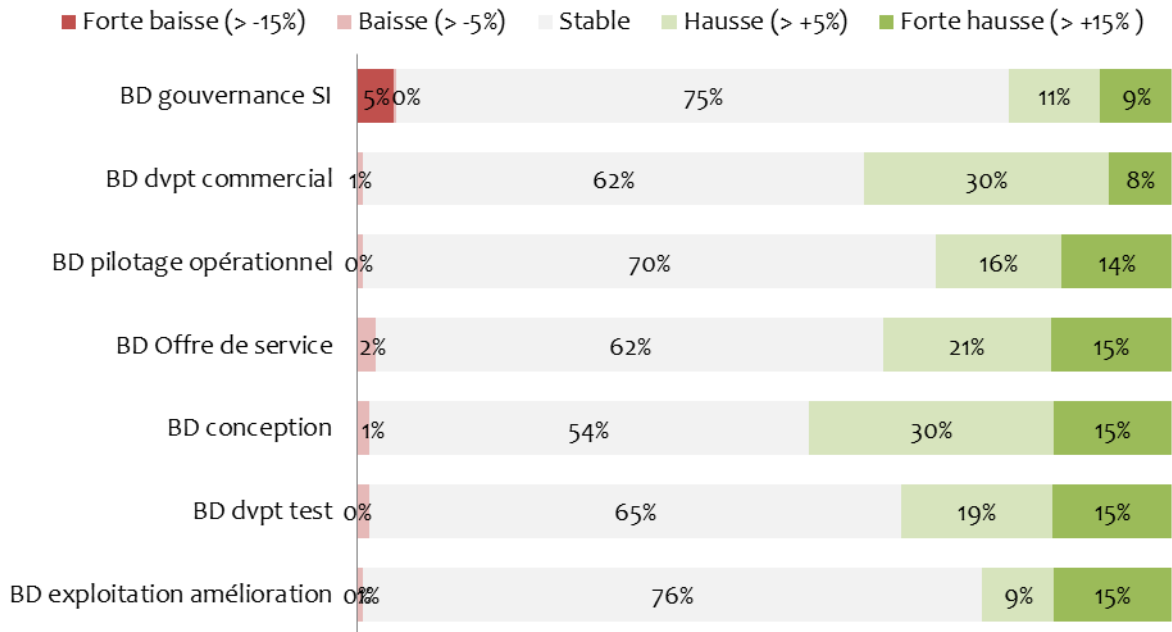
Les entreprises ont également exprimé dans le questionnaire leur perception de l'impact quantitatif du Big data sur leurs effectifs dans chacune de ces familles métiers.

Si au global les avis sont assez partagés, il apparaît tout de même que la plupart des entreprises **perçoivent globalement plutôt une hausse de leurs effectifs** en lien avec le Big data sur chacune des familles métiers, voyant le Big Data plutôt comme une opportunité de développement que comme une menace.

<sup>10</sup> A la différence du rapport sur le Cloud Computing, ces données n'ont pas été pondérées par les effectifs. 2 entreprises répondantes représentant près de 75% du poids des réponses, cette pondération n'aurait pas été pertinente.



## Perception des entreprises répondante sur l'impact de leurs activités Big Data sur l'évolution de leurs effectifs par famille de métiers



Données : enquête auprès des entreprises  
Traitements : Ambroise Bouteille et Associés

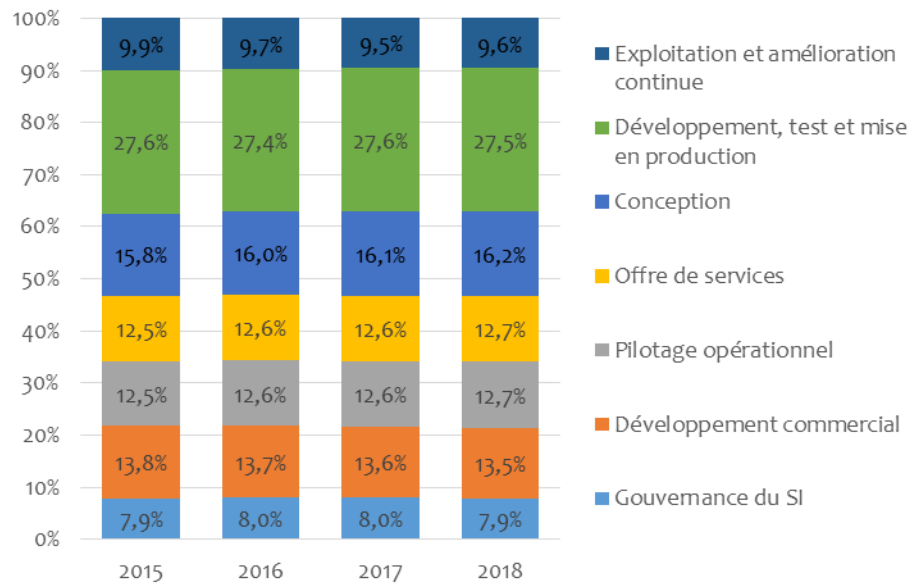
Ainsi, en pondérant ces résultats et en les appliquant aux effectifs par familles de métiers sur [les effectifs des entreprises du panel](#) (par ventilation des effectifs totaux en fonction des familles selon les résultats de l'enquête), on obtiendrait au global [une faible évolution de la répartition des effectifs](#).

Le développement des activités Big Data dans les entreprises du panel n'a donc pas semblé de nature à bouleverser les équilibres entre les familles métiers :





### Evolution projetée du poids des familles de métiers dans les entreprises du panel en lien avec l'émergence du Big Data



Données : OPIIEC, Enquête auprès des entreprises  
Traitements : Ambroise Bouteille et Associés



## Les besoins en compétences liés au Big data mis en avant par les entreprises

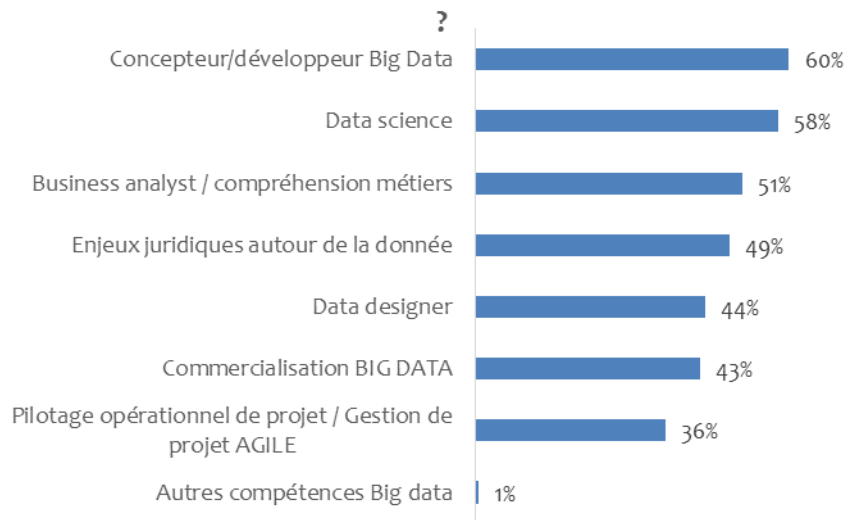
Hormis les tendances quantitatives prévues, le questionnaire a permis de sonder les entreprises sur leurs intentions en matière d'accompagnement de ces évolutions et de leurs besoins en compétences associés à l'émergence du Big data.

### Les compétences clés recherchées par les entreprises

Assez logiquement, les besoins en compétences de développement Big Data (type maîtrise univers Hadoop, etc.) ressortent comme le premier besoin des entreprises pour préparer le développement d'une offre Big Data, de même que la data science, qui avait déjà été citée de manière récurrente par les entreprises lors des entretiens qualitatifs.

L'ensemble des besoins en la matière semble être concentrés sur les 7 compétences citées ci-dessous :

#### Compétences clés BD - Quelles sont les compétences BIG DATA que vos équipes doivent acquérir en priorité ?



Données : Enquête auprès des entreprises  
Traitements : Ambroise Bouteille et Associés

*Assiette : répondants positionnés ou envisageant de se positionner sur des activités Big Data*

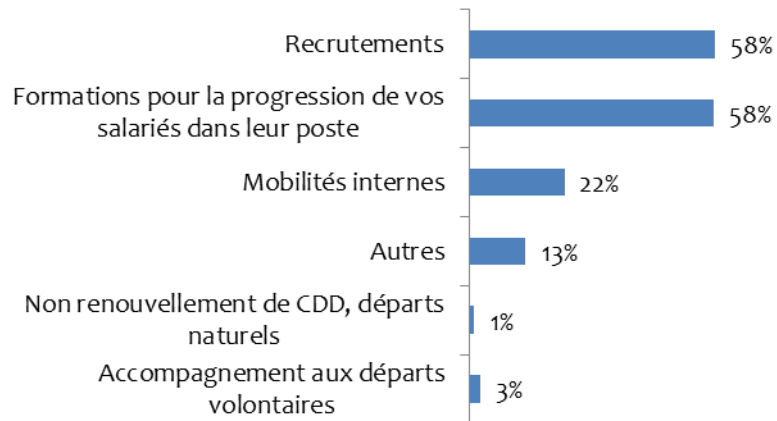
Pour accompagner ces évolutions de besoins en compétences, plusieurs solutions sont envisagées dans les entreprises.

En effet, pour 58% des répondants positionnés sur le Big Data ou envisageant de le faire, la formation et le recrutement seront les principaux leviers d'acquisition de ces compétences clés.

En revanche, peu de départs semblent envisagés par les entreprises, ce qui semble indiquer que les entreprises pensent pouvoir adapter leurs salariés en poste aux nouveaux besoins en compétences générés par le Big data : des passerelles doivent donc être envisageables pour aller vers les métiers/compétences Big data.



### Solutions RH - Quelles solutions RH envisagez-vous pour accompagner ces évolutions quantitatives (tant Cloud que Big data) ?

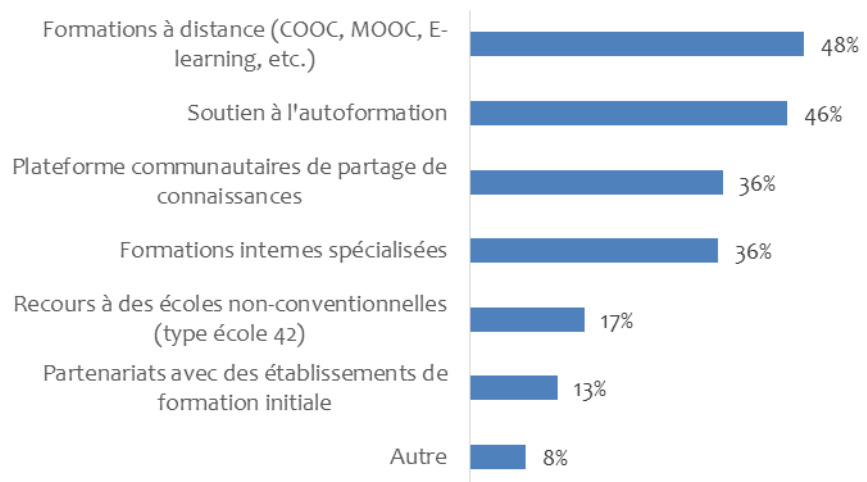


Source : Enquête auprès des entreprises  
Traitements : Ambroise Bouteille et Associés

*Assiette : répondants positionnés ou envisageant de se positionner sur des activités Big Data*

Les entreprises semblent envisager de tester plusieurs options pour accompagner leurs salariés dans l'acquisition de ces compétences :

### Projets innovants - Projetez-vous des solutions innovantes pour relever ces défis de développement des compétences ?



Données : Enquête auprès des entreprises  
Traitements : Ambroise Bouteille et Associés

*Assiette : répondants positionnés ou envisageant de se positionner sur des activités Big Data*



## Des difficultés d'accès à ces compétences perçues comme particulièrement sensibles pour le Big data

Il apparaît en matière de recrutement que les entreprises partagent le constat réalisé qualitativement de difficultés sensibles à recruter des salariés de profils « Big data ».

69% des entreprises interrogées jugent ainsi que les difficultés de recrutement seront supérieures sur les profils Big data que sur des recrutements traditionnels, déjà complexes pour les métiers non-spécialisés Big data.

**Recrutement1 - De même, présentez-vous des difficultés de recrutement (supérieures aux difficultés habituelles en IT) de profils adaptés aux exigences du BIG DATA au cours des 3 prochaines années ?**

■ Non, pas du tout ■ Plutôt non ■ Plutôt oui ■ Oui, tout à fait



Données : enquête auprès des entreprises  
Traitements : Ambroise Bouteille et Associés

*Assiette : répondants positionnés ou envisageant de se positionner sur des activités Big Data*



0. Rappel des objectifs et de la méthode
  1. Des innovations technologiques...
  2. ...ouvrent de nouvelles fonctionnalités et usages pour les entreprises
  3. Un sujet d'avenir, en émergence, autour duquel de nouvelles offres se développent
  4. L'impact de ces bouleversements sur les modes de travail entre clients et prestataires
  5. Les besoins en compétences liés au Big data
  6. Les enjeux en termes de formation liés au Big data
  7. Quantification des volumes d'emploi liés au Big data
- 

## 8. Cartographie et analyse de l'offre de formation

---

9. Préconisations
10. Annexes



## Panorama de l'offre de formation sur le Big data

### Méthodologie d'élaboration de cet inventaire

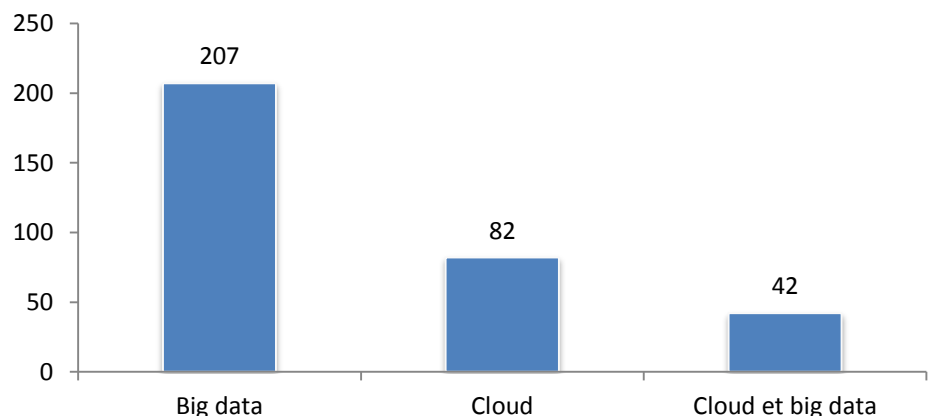
Plusieurs sources ont été investiguées et croisées pour élaborer cet inventaire, parmi lesquelles :

- les formations citées par les entreprises et salariés lors des entretiens
- pour la formation initiale dont l'enseignement supérieur : recherche des formations correspondantes sur :
  - le RNCP
  - la base InterCarif
  - sites spécialisés type Onisep ou l'Etudiant
  - sites web de grandes écoles
  - recherche internet par mots-clés
- pour la formation continue :
  - exploitation de la base des demandes de formation (base brute transmise par le Fafiec en juillet), via une recherche par mots-clés
  - base InterCarif
  - catalogue formation continue des établissements d'enseignement supérieur repérés
  - recherche d'exemples emblématiques de stage de perfectionnement présent sur le marché, à partir d'une recherche par mot-clé sur Internet directement

### Le recensement ainsi élaboré figure dans les annexes

- Au total figurent dans ce fichier près de 330 formations identifiées

**Nombre de formations spécialisées cloud et big data recensées dans l'inventaire excel selon le "segment" (cloud ou big data)**



- L'ambition de ce travail de recensement est multiple :
  - avoir un aperçu de comment est structuré le marché à l'heure actuelle



- se rendre compte de l'offre qui est accessible, ou tout du moins visible, tant pour les entreprises que les salariés

Il ne s'agit pas de prétendre à l'exhaustivité (qui ne pourrait être réalisable, car l'offre évolue en temps réel et est également cachée, non visible car adapté à la demande au cas par cas et non en offre catalogue) mais véritablement d'avoir une vue d'ensemble des tendances (types de formation proposées, la façon dont les sujets Big data et Cloud computing sont abordés par l'offre de formation).

Les traits saillants émergents de ce travail sont soulignés dans la page suivante.

- Le fichier excel en lui-même a été transmis avec ce rapport à l'Opiiec
  - Pour chaque formation, lorsque les informations étaient accessibles sur internet, les champs suivants ont été renseignés :

•

Segment	Intitulé de la formation	Type	Niveau	Public cible	Durée (heures, années)	compétences principales visées / programme de formation	Exemples de formation p
---------	--------------------------	------	--------	--------------	------------------------	---	-------------------------

•

Nom de l'établissement	accessible en Cpro	accessible en contrat d'apprentissage	Ville	Département	Code Postal	Téléphone standard	Mail	Site internet
------------------------	--------------------	---------------------------------------	-------	-------------	-------------	--------------------	------	---------------

•

### A noter que de nombreuses formations abordent à la fois le Big data et le Cloud

Naturellement, il existe aussi des formations spécialisées sur l'un et l'autre sujet.

Cela étant, un nombre non négligeable de formations abordent les deux sujets. Ceci n'est pas surprenant, dans la mesure où tant le Big data que le Cloud sont deux sujets déterminants pour le secteur numérique, et le premier intègre presque nécessairement les compétences du second. Aussi il est « logique » que les formations numériques dans leur ensemble consacrent une part de leur programme à ces deux sujets, sans être nécessairement spécialisées sur l'un ou l'autre.



## Les principaux points à souligner concernant l'offre de formation en Big data

### De prime abord, une offre qui apparaît nombreuse, tant en formation initiale qu'en formation continue

En effet, près de 250 formations spécialisées recensées abordant, d'une manière ou d'une autre, le Big data.

### Les principaux constats concernant la formation initiale

Notons d'abord que les **cursus dédiés/spécialisés « Big data » ou « data science » sont en nombre limité** aujourd'hui...

- Ce sont essentiellement les grandes écoles d'ingénieurs qui semblent avoir pris le pas et initier la mise en place de cursus dédiés Big data, le plus souvent sous la forme de mastère spécialisé (en 3<sup>ème</sup> cycle) ou de master 2 (en 2<sup>ème</sup> cycle)
- Ces cursus sont relativement bien repérés par les entreprises pour plusieurs raisons :
  - la renommée et le « sérieux » des établissements les ayant mis en place
  - au sein des entreprises, des dirigeants peuvent être issus de ces mêmes écoles
  - ces cursus ont aussi été élaborés en lien étroit avec les professionnels, de nombreuses entreprises ayant été associés dans leur conception et les « sponsorisant » (cf. partie 5)
- Ces cursus spécialisés, dont les premières promotions sont encore en formation ou viennent tout juste d'arriver sur le marché du travail, sont par exemple ceux proposés par :
  - l'ENSAE (école de l'Insee) (MS en Data Science)
  - l'ENSAI (master Big data)
  - TelecomParisTech (MS big data : gestion et analyse de données massives)
  - l'EMSI
  - etc.

... mais que l'offre est en **pleine structuration** aussi des autres sont **susceptibles d'émerger**

- En effet, de nombreux établissements sont probablement en cours de réflexion (par exemple, des universités)
- D'autre part, le créneau étant, selon nos interlocuteurs au sein des organismes de formation « porteur », des organismes privés sont susceptibles de se positionner ou de se créer sur ce créneau. Par exemple, le DataScience Institute (start up) vient de lancer ses 2 cursus de formation (accessible tant pour les jeunes en sortie d'études que pour les publics en formation continue)

Probablement conséquence de l'engouement médiatique autour du Big data, **certains cursus de formation initiale ont été renommés**, « rebrandés », Big data

- Il s'agit essentiellement de cursus de management et école de commerce, pour que leur « sémantique » soit en phase avec l'expression du marché





- Toutefois, au-delà de ce changement d'intitulé, les formations n'ont pas forcément évolué grandement dans leur contenu. L'essentiel des programmes est concentré sur les aspects management et un volet sur les enjeux stratégiques du Big data s'y est greffé.
- D'autre part, les débouchés visés par ces formations sont sans doute plus orientés vers l'utilisation des Big data (au sein donc des entreprises utilisatrices plutôt que des entreprises de la branche)

**D'autres formations n'apparaissent pas comme consacrées spécifiquement au Big data**, mais abordent ses enjeux dans un cadre plus général

- Il s'agit généralement des cursus d'écoles d'ingénieurs, dans leurs parcours de formation d'ingénieurs généralistes (par exemple, l'École Polytechnique ou l'ISEP) ou d'écoles comme l'ENS Cachan (formation en machine learning)
- Ces écoles d'ingénieurs délivrant des formations scientifiques pointues (incluant nécessairement le socle scientifique pour la gestion de données massives)
- A noter que, pour les entreprises interrogées, ces cursus, même si non spécialisés « Big data », sont connus et reconnus de qualité. Formant des « têtes bien faites », ils constituent, au même titre que les cursus dédiés Big data qui se mettent en place, des cibles de recrutement privilégiées par les entreprises.

### Les principaux constats concernant l'offre de formation continue

**Des offres multiples que les entreprises n'identifient pas toujours, hormis les formations des éditeurs sur leurs solutions et produits (SAS, etc.)**

- Des formations « catalogues » et « généralistes » qui sont apparaissent comme des « sensibilisation » au Big data (qu'est-ce que le Big data, ses enjeux pour l'entreprise, etc.)
- Sur les aspects scientifiques/techniques, des évolutions fortes et rapides qui font que les offres de formation (selon certaines entreprises) évoluent en permanence et ne peuvent devenir des produits catalogue
- Des éditeurs proposent ainsi des offres « non catalogues », en cela qu'ils donnent accès à un ensemble de ressources (leurs produits, des tutoriels en ligne, etc.) et les individus sont ensuite libres de s'auto-former grâce à ces aides (par exemple, le SAS Analytics U : [http://www.sas.com/fr\\_fr/learn/analytics-u.html](http://www.sas.com/fr_fr/learn/analytics-u.html) )

**Des offres de formation que les entreprises développent en interne**

Certaines grandes entreprises disposent en effet de leur propre école, institut de formation interne, université (terminologie différent selon les entreprises), tant parmi les entreprises utilisatrices que les entreprises de la branche (il s'agit notamment des grandes ESN ou encore des éditeurs de solutions Big data)

Quelques-unes ont déclaré que des formations internes au Big data étaient mises en place.

Toutefois, les entreprises n'ont pas souhaité communiquer plus d'informations sur ces formations internes (contenus, publics auxquels elles s'adressent, modalités, etc.)

Comme évoqué dans la partie précédente, **des « formations » aux modalités innovantes et disruptives sont mises en place en continu**



- Certaines ne sont pas nécessairement des « formations » en tant que telles : hackathons, participation au forum en ligne (il s'agit d'apprentissage « sur le tas », « d'apprendre en faisant »)
- mais une offre dense de formation online est accessible à tous :
  - des MOOC (y compris proposés par les établissements ayant mis en place des cursus spécialisés)
  - de nombreuses vidéos en ligne (par exemple, sous la forme de tutoriels, proposées par des entreprises, des amateurs, des organismes de formation, etc.)Quelques exemples : le Wagon, les post de blog de la société Dataiku ([http://www.dataiku.com/blog/2015/09/02/Free\\_Training\\_06.html](http://www.dataiku.com/blog/2015/09/02/Free_Training_06.html) )



## L'adéquation de l'offre de formation avec les besoins exprimés par les professionnels

### De nombreux cursus de formation initiale spécialisés en big data ont vu le jour récemment, d'autres vont probablement prochainement être créés

Ces cursus présenteraient l'avantage d'offrir des cursus dédiés aux nouveaux métiers de l'exploitation de la data

### L'offre de formation initiale big data est récente et en pleine structuration

- La totalité des formations initiales repérées spécialisées sur le big data est **très récente** : beaucoup n'ont d'ailleurs ouvert qu'à la rentrée 2014, d'autres ont ouvert en cette rentrée 2015
- En conséquence, **la question de leurs référentiels de formation demeure entière**. En effet, parce qu'elles viennent de se mettre en place, l'inscription de ces formations au RNCP n'est pas encore réalisable (il faut attendre que quelques promotions soient sorties). Néanmoins, le caractère plus ou moins rigide et figé d'un référentiel pourra s'avérer problématique, alors même que :
  - - les métiers sont en pleine définition : leurs contours ne sont pas nécessairement encore nettement définis au sein des entreprises
  - - le big data et la data science sont des champs nouveaux, exploratoires : aussi il semble délicat et non forcément désirable d'enfermer les formations dans des référentiels (figés pendant une certaine durée), car ceux-ci seraient frappés d'obsolescence assez rapidement et l'offre ne pourra prétendre à la réactivité attendue pour s'adapter aux besoins des professionnels.

### Ces formations initiales, qui voient progressivement le jour, reposent sur des expertises pré-existantes fortes et pointues en mathématiques, statistiques, informatique

- En effet, beaucoup de ces formations ont été mises en place au sein d'établissements d'enseignement supérieur scientifiques de haut niveau : Ecole Polytechnique, ENSAE, Telecom ParisTech, etc.
- **Ces expertises pointues dans l'exploitation de la donnée correspondent aux besoins des professionnels**. D'ailleurs, lors des entretiens qualitatifs, les entreprises ont estimé que ces « bases théoriques » étaient indispensables et constituaient le « cœur » des compétences attendues
- L'**innovation** consiste dans la **dimension croisée de cette expertise avec les outils IT** de traitement de données massives (NoSQL, langages R et Python, MapR, etc.)



## Les offres de formation initiale se structurent autour de la donnée

ce qui offre ainsi un périmètre plus large que celui seul des statistiques ou de l'IT

- Les programmes de formation sont en effet structurés majoritairement autour du parcours de la donnée (sa collecte, son stockage, son exploitation, etc.)
- Ce qui permet ainsi d'aborder notamment les **aspects juridiques et éthiques** autour du recueil et de l'exploitation des données.

**En revanche, deux aspects semblent aujourd'hui n'être que peu développés, voire manquer, au sein des formations qui voient progressivement le jour :**

### Les enjeux business autour de la donnée

C'est-à-dire les aspects relatifs à la valorisation ultérieure des données exploitées, à la compréhension des secteurs d'application, à l'évolution des business models

- Les formations sont essentiellement tournée sur le couple data science / outils IT
- Or, l'ont expliqué les professionnels interrogés lors des entretiens qualitatifs, le big data n'est pas un diptyque (data science / IT), mais un triptyque (data science / IT / business).
- L'aspect « métier » (comment je tire de la valeur des données et comment j'adapte mon business model à celle-ci) ne semble pas prédominant, voire peu ou pas présent, dans la plupart des formations initiales repérées, tout du moins dans les programmes : la focale n'est pas placée sur les usages business qui peuvent ensuite être fait de la donnée.
- A l'exception près des cursus de BI qui ont été « rebrandés » Big data.
- Toutefois, soulignons que si ces aspects peuvent sembler peu présents dans les programmes, les **modalités pédagogiques des formations initiales semblent se prêter à cette acculturation « métier »** : de nombreux cas d'usage sont traités lors des enseignements et projets collectifs.
- D'autre part, **l'offre est en pleine structuration, elle n'est donc pas figée**. Lors des entretiens auprès des organismes de formation, certains commencent à envisager des « filières complètes » de formation big data (de type « santé et big data », « commerce et big data », etc.) qui pourraient se calquer sur le schéma LMD, et non sur un seul 3ème cycle.

### Les aspects communication et pédagogie

Les aspects pédagogie et communication semblent en revanche, quant à eux, « écartés » de ces formations :

- ils semblent uniquement être abordés sur l'utilisation d'outils de datavisualisation (maîtrise de l'outil Tableau notamment)



- ils semblent uniquement être abordés sur l'utilisation d'outils de datavisualisation (maîtrise de l'outil Tableau notamment)

**A l'instar de l'offre de formation initiale, l'offre de formation continue en big data est également en construction progressive**

**Les établissements de formation initiale, de la même manière qu'ils ont investi la formation initiale, développent également une offre de formation continue**

- Leur offre semble essentiellement être proposée « à la carte », sur demandes des entreprises et en fonction de leurs besoins.
- Cela permet à la fois une adaptabilité à la demande de l'entreprise, mais aussi une réactivité pour intégrer en continu de nouveaux outils (nouvel environnement de calcul parallèle, etc.)

De multiples stages courts de formation continue ont également été repérés. Ceux-ci se focalisent essentiellement soit sur les aspects techniques et mathématiques (formation à un langage, à l'utilisation de tel ou tel type d'algorithme), ou sinon sur une acculturation au big data (qu'est-ce que le big data, quels en sont les enjeux, etc.)

**En parallèle, une offre de formations accessibles au plus grand nombre (sous forme de MOOCs, vidéos en ligne, cours online à prix modique, plateforme de concours de data science) apparaît relativement dense**

- Ces modalités sont plébiscitées par les professionnels et également par les entreprises, qui, comme vu lors des entretiens, misent beaucoup sur l'auto-formation, la curiosité et l'investissement personnel des salariés
- A noter que par son caractère dématérialisé, cette offre n'est pas nécessairement « française » : les entreprises et salariés interrogés ont d'ailleurs bien plus souvent fait référence à Coursera plutôt que FUN<sup>11</sup> ; nombreux sont les MOOCs d'universités outre-atlantique (quand bien même certains établissements d'enseignement supérieur ont également développé les leurs) ; les start up proposant des vidéos en ligne d'usage de leurs technologies peuvent être des pépites françaises (ex : Dataiku), mais aussi US ou allemandes, etc.
- Notons qu'une large partie de ces outils de formations (par exemple : tutoriel sur telle ou telle technologie, tel ou tel langage, etc.) semble déjà nécessiter **un premier niveau d'expertise**, a minima de maîtrise, en la matière.

---

11 France Université Numérique



## **Au global, il semble néanmoins manquer une offre pour les utilisateurs du big data**

- Comme vu précédemment, sur les aspects centrés essentiellement sur les technologies et techniques d'exploitation de la donnée, l'offre de formation semble être en mesure de s'adapter de façon permanente aux besoins des entreprises de la branche (qui se positionneront comme fournisseurs ou conseil auprès de leurs clients).
- Néanmoins, en matière de formation continue, semblent manquer dans le paysage des modules destinés aux utilisateurs possibles du big data (par exemple : le big data pour les juristes, le big data pour le marketing, etc.). Ceux-ci seraient certes susceptibles d'intéresser bien davantage les entreprises clientes. Toutefois, le développement d'une telle offre pourrait bénéficier aux entreprises de la branche de deux façons :
  - d'une part, pour les profils commerciaux (mieux connaître les besoins et spécificités des clients)
  - d'autre part, en cela qu'une telle offre permettrait une acculturation plus large des entreprises clientes au big data (et favoriserait sans doute chez elle l'émergence de projets, que les entreprises de la branche pourraient accompagner). Ils pourraient aussi bénéficier

## **Enfin, tant au niveau de la formation initiale que continue, n'apparaissent pas de façon évidente les enjeux autour de l'agilité et du test and learn**

Comme exprimé par les professionnels interrogés, le big data induit de nouvelles façons de travailler et un mindset différent. Il s'agit en effet d'une exploration permanente : il n'est pas possible de prédire à l'avance le résultat de telle ou telle approche. Aussi, cela impose d'admettre plus facilement le droit à l'erreur, de « déstigmatiser » les échecs en cours de projet, de « franchir la ligne d'arrivée avec des bosses ».

Ce changement de posture est en fait une rupture culturelle. Elle n'est pas vraiment abordée sous l'angle de la formation, mais plutôt sous la forme « d'acculturation » (par exemple, avec les plateformes de concours de type Kaggle.com ou datascience.net). Par ailleurs, un pari est aussi fait sur la capacité des générations Y et Z à insuffler ce nouvel état d'esprit au sein des entreprises (cf. supra)



0. Rappel des objectifs et de la méthode
  1. Des innovations technologiques...
  2. ...ouvrent de nouvelles fonctionnalités et usages pour les entreprises
  3. Un sujet d'avenir, en émergence, autour duquel de nouvelles offres se développent
  4. L'impact de ces bouleversements sur les modes de travail entre clients et prestataires
  5. Les besoins en compétences liés au Big data
  6. Les enjeux en termes de formation liés au Big data
  7. Quantification des volumes d'emploi liés au Big data
  8. Cartographie et analyse de l'offre de formation
- 

## 9. Préconisations

---

## 10. Annexes



## Trois grands volets de recommandations

Volet de recommandations	Les grands objectifs poursuivis	Récapitulatif des actions proposées	Qui pourrait en être le porteur?
1. Convaincre des avancées et des bénéfices offerts par le Big data	<ul style="list-style-type: none"> <li>Dépasser l'engouement médiatique actuel et récent autour du Big data et <b>prouver sa réelle valeur ajoutée</b> (démontrer qu'il ne s'agit pas que d'un « effet de mode »)</li> <li><b>Evangéliser</b> le plus largement possible sur les implications du Big data, afin de <b>développer le marché de clients potentiels</b> pour les entreprises de la branche</li> </ul>	<ul style="list-style-type: none"> <li>Communiquer largement sur les <i>success stories</i> au sein des entreprises utilisatrices, en insistant sur le changement de paradigme sous-jacent à toute approche Big data</li> <li>Former spécifiquement les commerciaux de la branche pour qu'ils puissent eux-mêmes être porteurs de ce message</li> </ul>	Branche, organisations professionnelles, Opiiec Fafiec
2. Accompagner la structuration de l'offre de formation	<ul style="list-style-type: none"> <li><b>Orienter</b>, plutôt que de bâtir <i>de novo</i>, <b>l'offre de formation « big data »</b></li> <li>Faciliter l'accessibilité des offres de formation « big data » au plus grand nombre, quel que soit le statut des apprenants  ... pour, <i>in fine</i>, que les entreprises puissent bénéficier de différents viviers de recrutement et qu'elles soient incitées à <b>diversifier leurs sources de recrutement</b></li> </ul>	<p><b>Au niveau des formations initiales</b></p> <ul style="list-style-type: none"> <li>Ne pas scléroser l'offre de formation qui est en train de se mettre en place...</li> <li>...mais tenter de l'orienter, en termes de contenus, en lui insufflant « l'état d'esprit » Big data</li> </ul> <p><b>Au niveau de la formation continue, spécifiquement pour les salariés de la branche</b></p> <ul style="list-style-type: none"> <li>Pour faire comprendre le Big data et ses implications pour les entreprises clientes, mettre en place des actions collectives consacrées à l'agilité</li> <li>Sur les compétences techniques du Big data, concevoir des actions collectives « avec agilité »</li> <li>Favoriser l'évolution des professionnels du BI vers les métiers du Big data et des professionnels de la gestion des bases de données vers le NoSQL</li> <li>Ce faisant, permettre aux entreprises de compléter leurs pratiques d'auto-formation par une logique de formation plus « structurée »</li> </ul> <p><b>De façon générale</b></p> <ul style="list-style-type: none"> <li>Ne pas se lancer dans une labellisation « Big data » de l'offre de formation</li> </ul>	Organisations représentatives siégeant dans les CPC  FAFIEC  CPNE





3. Développer la notoriété et l'attractivité de la branche auprès des sortants de formation

- **Favoriser l'orientation** des profils Big data **vers** les entreprises de **la branche** (et non pas vers les seules entreprises clientes)...
- ... et ce, quelle que soit la « voie » de formation de ces profils (cursus de formation initiale dédié « Big data », autodidacte, etc.)

- ▶ Inciter les professionnels de la branche qui œuvrent sur des projets Big data à enseigner au sein des formations big data
- ▶ Au niveau des acteurs de la branche, ne pas s'engager dans l'ingénierie de certifications susceptibles d'être inscrites au RNCP (qu'auraient pu être par exemple un CQP ou un titre en partenariat avec un établissement de formation initiale)
- ▶ En revanche, envisager la mise en place de certifications éligibles à l'inventaire CNCP pour les compétences qui s'y prêtent
- ▶ Veiller à l'inscription sur les listes de formations éligibles au CPF des cursus « big data » qui se mettent en place progressivement
- ▶ Pour, in fine, permettre aux entreprises de diversifier leurs sources de recrutement

- ▶ Mener une campagne non pas sur l'attractivité des métiers du Big data, mais sur l'intérêt d'exercer ceux-ci *au sein de la branche*
- ▶ Développer des partenariats avec les établissements de formation initiale
- ▶ Organiser des défis et challenges, portés par les organisations professionnelles, sur les plateformes de concours et défis de Big data

Opiiec

Branche

Syntec  
Numérique,  
CINOV-IT



## Volet 1 : Convaincre des avancées et des bénéfices offerts par le Big data

NB : ce premier volet est quelque peu « hors sujet » dans le cadre de cette étude, car il déborde amplement le sujet de la formation, voire le périmètre des entreprises de la branche. Cela étant, comme cela a été souligné à de nombreuses reprises lors des entretiens auprès des entreprises, le Big data est un sujet nouveau qui reste encore mal connu des entreprises. C'est pourquoi ce volet, qui touche en filigrane à la question du développement du marché du Big data, nous apparaît comme un préalable indispensable au traitement de la question de la formation des professionnels.

### Les grands objectifs poursuivis par ce volet

- Dépasser l'engouement médiatique actuel et récent autour du Big data et **prouver sa réelle valeur ajoutée** (démontrer qu'il ne s'agit pas que d'un « effet de mode »)
- **Évangéliser** le plus largement possible sur les implications du Big data, afin de **développer le marché de clients potentiels** pour les entreprises de la branche

### Pourquoi ? [rappel des constats]

- Le big data est un **sujet encore récent**, et **reste encore mal connu**
- **Un buzz autour du « Big data », qui conduit à certains abus de langage** : tout type de traitement de données tend à être estampillé « Big data ». Ce faisant, les projets voulus Big data restent souvent confinés à une sphère IT ou statistique.
- Or, **le Big data constitue une révolution pour les entreprises clientes de la branche** :
  - Le big data est susceptible de remodeler en profondeur la façon dont les entreprises font du business
  - Le big data constitue un défi pour l'entreprise toute entière : ce n'est pas un sujet confiné à l'IT, il s'agit d'un véritable sujet d'entreprise, un sujet décisionnel, qui impacte de manière transversale la plupart des fonctions de l'entreprise
  - Il implique de nouvelles façons de travailler : changement de la vision autour des données, test & learn, agilité...
- Le **plus gros frein** au Big data est ainsi **d'origine culturelle** : réticence à partager les données, aversion au risque et à l'ambiguïté. Les organisations ne sont pas toutes adaptées à des modes de travail « test & learn »
- Parce qu'**exploration permanente**, le Big data implique un état d'esprit différent : tout projet est un pari, il n'est pas possible de savoir à l'avance quels en seront les résultats et le ROI pour l'entreprise. Les réussites ne sont pas nécessairement immédiates et « franches ».



## Comment ?

- ▶ **Communiquer largement sur les *success stories* au sein des entreprises utilisatrices, en insistant sur le changement de paradigme sous-jacent à toute approche Big data**
  - Ne pas insister sur les seuls succès (sur lesquels, d'ailleurs et peut-être, les entreprises concernées ne voudront pas toujours communiquer, préférant conserver leur avantage concurrentiel), mais sur la façon d'y arriver (« franchir la ligne d'arrivée avec des coups »)
  - Adopter ainsi un message se centrant non pas sur la valeur des données (désormais marronnier des médias), mais sur l'agilité, les nouvelles façons de travailler, etc.
- ▶ **Former spécifiquement les commerciaux de la branche pour qu'ils puissent eux-mêmes être porteurs de ce message**
  - Mettre en place une action collective spécifique « Vendre le big data, une révolution pour les clients »
  - Axer tout particulièrement cette action collective sur les aspects « au-delà de la technique » :
    - « vendre du Big data », ce n'est pas juste vendre une prestation d'analyse de données, c'est révolutionner la façon de faire du client et son business. Comment amener les clients à travailler sur leurs propres business models et leurs nouvelles missions ?
    - le test & learn : comment vendre une prestation dont les résultats sont incertains, comment négocier l'incertitude ?
    - pourquoi l'agilité est-elle indispensable à toute démarche big data et comment le faire comprendre aux prospects ?



## Volet 2 : Accompagner la structuration de l'offre de formation

### Les grands objectifs poursuivis par ce volet

- **Orienter**, plutôt que de bâtir *de novo*, l'offre de formation « big data »
- **Faciliter l'accessibilité** des offres de formation « big data » au plus grand nombre, quel que soit le statut des apprenants  
... pour, *in fine*, que les entreprises puissent bénéficier de différents viviers de recrutement et qu'elles soient incitées à **diversifier leurs sources de recrutement**

### Pourquoi ? [rappel des constats]

- Le Big data est un sujet nouveau et **exploratoire** : ses métiers se redéfinissent au jour le jour
- L'offre de formation en la matière est récente et **en pleine structuration** : des cursus dédiés ont récemment vu le jour et d'autres sont susceptibles d'émerger
- L'aspect **communautaire** omniprésent et quasi-intrinsèque au Big data **constitue un véritable challenge pour la formation** : les technologies, langages, manières de procéder évoluent en permanence et de façon rapide)
- L'offre de formation aujourd'hui existante semble **centrée sur les technologies** et techniques d'exploitation de la donnée (data science / outils IT), mais semble **occulter les aspects « métiers », « agilité », « pédagogie »**
- **Les pratiques actuelles** de formation des professionnels et des entreprises misent avant tout sur la curiosité, **l'investissement personnel** (participation aux forums, communautés, contribution au code, etc.)

### Comment ?

#### Au niveau des formations initiales

- ▶ **Ne pas scléroser l'offre de formation qui est en train de se mettre en place...**
  - Le Big data étant un nouveau métier, il pourrait être tentant, voire « rassurant » de bâtir une filière de formation dédiée, en essayant de définir (voire d'imposer) une « norme », un « référentiel » sur ce que doit être un cursus Big data
  - C'est-à-dire, ne pas essayer d'imposer une « maquette de formation » minimal auxquels les établissements/organismes de formation qui souhaiteraient créer un cursus big data devraient se conformer
- ▶ **...mais tenter de l'orienter, en termes de contenus, en lui insufflant « l'état d'esprit » Big data**
  - En termes de communication (cf. volet 1), **cibler** spécifiquement **les établissements de formation**, notamment sur les aspects non-techniques :



- aspects métiers / business, déclinés éventuellement par industrie
- agilité, faire preuve d'agilité
- pourquoi cela révolutionne la façon de travailler des entreprises, etc. ?
- aspects légaux
- Elaborer une **cartographie des métiers** et/ou compétences Big data et la diffuser tout particulièrement auprès des établissements d'enseignement
  - insister sur les compétences « non techniques » (business, compréhension métier, pédagogie, communication, ergonomie visuelle, agilité, légal, etc.) plutôt que de détailler les compétences techniques (dont la rédaction sera par ailleurs rapidement obsolète si trop détaillée)
  - faire en sorte que cette cartographie puisse être actualisée régulièrement, voire l'alimenter / l'amender de façon collaborative

### Au niveau de la formation continue, spécifiquement pour les salariés de la branche et les entreprises

#### ► Pour faire comprendre le Big data et ses implications pour les entreprises clientes, mettre en place des actions collectives consacrées à l'agilité

- Il ne s'agit pas de mettre en place des actions collectives consacrées à la méthode AGILE (qui figurent d'ores et déjà dans le catalogue du Fafiec)...
- ... mais de formations qui seraient consacrées à l'agilité, non pas comme méthode de gestion de projet, mais comme état d'esprit, façon de penser (pourquoi l'agilité ? comment travailler avec l'incertitude ? savoir manager l'incertitude).  
Pourraient également être abordés les éléments de design-thinking et d'idéation.

#### ► Sur les compétences techniques du Big data, concevoir des actions collectives « avec agilité »

Pourquoi ? : les technologies et techniques d'exploitation des données évoluent en permanence (notamment du fait de l'open source et de l'aspect communautaire ultra-présent)

Ainsi :

- **assurer une veille** sur les technologies émergentes et nouvelles pour être capable de détecter les outils d'avenir
- **déroger au « catalogue »** des actions collectives du Fafiec, en ne définissant pas le contenu de ces actions sur une base annuelle...

... et, en lieu et place, **consacrer une enveloppe budgétaire spécifique** qui permettrait de mettre en place des **actions collectives au fur et à mesure**, sur une base par exemple mensuelle ou trimestrielle (les thématiques et le contenu étant définis et redéfinis plusieurs fois en cours d'année, en fonction des besoins qui émergent au fil de l'eau)



► **Favoriser l'évolution des professionnels du BI vers les métiers du Big data et des professionnels de la gestion des bases de données vers le NoSQL**

- Concernant l'évolution des profils BI vers les métiers du Big data :
  - en effet, ces profils détiennent d'ores et déjà certaines compétences des profils big data, notamment pour la partie business/métier et les aspects communication et pédagogie
  - l'écart sans doute le plus conséquent serait sans doute les aspects IT et algorithmie, aussi :
    - \* veiller à l'inscription progressive sur les listes CPF des cursus Big data qui voient le jour
    - \* analyser les possibilités d'un abondement des heures CPF pour les reconversions vers le Big data
- Concernant le passage au NoSQL pour les
  - envisager une certification éligible à l'inventaire CNCP pour le NoSQL
  - bien que les technologies évoluent très vite (rendant complexe la question de la certification), dans ce cas de figure cela n'apparaît pas problématique, car il ne s'agit pas d'une technologie, mais d'une nouvelle catégorie de gestion de base de données
  - ainsi, naturellement, dans la rédaction du référentiel, ne pas mentionner une ou plusieurs technologie en particulier, mais rester générique (pour éviter qu'il ne se périmé trop vite)

► **Ce faisant, permettre aux entreprises de compléter leurs pratiques d'auto-formation par une logique de formation plus « structurée »**

- Sur les aspects techniques notamment, qui évoluent très vite, car communautaires, ne pas « brider » les initiatives des entreprises et des salariés en la matière (participation aux communautés, hackathon, contribution aux codes, etc.)  
D'autant plus que, comme vu précédemment, la récente réforme de la formation professionnelle (notamment l'élargissement de la définition de la formation éligible et la suppression du 0,9%) s'inscrit dans cette tendance (les entreprises ont davantage le « choix des armes » pour former leur personnel)
- Mais, permettre aux entreprises et leurs salariés d'accéder aux financements mutualisés (professionnalisation, CPF, etc.) via une offre de formation plus « structurée », permise notamment :
  - grâce aux actions collectives évoquées ci-dessus
  - grâce aux certifications type inventaire CNCP évoquées ci-dessus

Cela présente un intérêt certain pour les entreprises :

- pour l'imputabilité et la prise en charge de ces formations
- pour satisfaire à l'obligation de former (et notamment être en mesure de justifier que les salariés tous les 6 ans ont eu accès à une évolution parmi lesquelles une formation)



**Pour l'ensemble des publics (salariés, demandeurs d'emploi, jeunes en formation initiale, autodidactes)**

► **Au niveau des acteurs de la branche, ne pas s'engager à ce stade dans de l'ingénierie de certifications susceptibles d'être inscrites au RNCP (qu'auraient pu être par exemple un CQP ou un titre en partenariat avec un établissement de formation initiale)**

- Car les métiers sont nouveaux et évoluent, se redéfinissent, rapidement
- Aussi, le ou les référentiels qui seraient établis pourraient être frappés d'obsolescence précoce.

► **En revanche, envisager la mise en place de certifications éligibles à l'inventaire CNCP pour les compétences qui s'y prêtent**

- Ne pas considérer les aspects techniques, qui évoluent constamment et pour lesquels toute certification « périmerait » rapidement
- En revanche, créer et mettre en place des certifications de type « inventaire CNCP » (catégorie C) peut être envisageable sur des aspects tels que :
  - l'agilité
  - la méthode AGILE
  - la datavisualisation (ergonomie visuelle, pédagogie)

Apprécier l'intérêt de s'associer à d'autres branches en la matière (clientes ou coprestataires telle la publicité par exemple)

► **Veiller à l'inscription progressive sur les listes de formations éligibles au CPF des cursus « big data » qui se mettent en place progressivement**

- Naturellement, les inscrire sur la liste élaborée par la branche, pour que les salariés de la branche puissent y avoir accès
- Demander à les faire inscrire sur la liste nationale interprofessionnelle, pour qu'elles puissent être accessibles par le plus grand nombre (demandeur d'emploi, salariés d'autres branches)
- Et ainsi, rendre plus aisés des reconversions (par exemple, des profils issus de la BI qui souhaiteraient aller vers les métiers du Big data)

NB : sur la dernière liste des formations éligibles au CPF (publiée le 26/10/2015), seules 3 formations parmi celles repérées dans le cadre du recensement réalisé dans le cadre de cette étude) figurent sur les listes de formation éligibles au CPF :

Organisme éditeur	Niveau	Intitulé de la certification	Certificateur	Code RNC	Code	Code Certi	Code CPF	Public
CPNE de l'ingénierie, des services informatiques et du conseil	Sans niveau spécifique	Certification Microsoft conception de solutions de business intelligence avec Microsoft SQL Server	Microsoft		99957	84472	138328	Salarié
CPNE de l'ingénierie, des services informatiques et du conseil	Niveau II (Bac + 3 et 4)	Licence pro management des organisations spécialité statistique et informatique décisionnelle	Université Lille 2 - droit et santé	4198		42637	146041	Salarié
COPANEF (Comité paritaire interprofessionnel national pour l'emploi et la formation)	Niveau II (Bac + 3 et 4)	Licence pro santé spécialité statistique et informatique décisionnelles pour la santé	Université Paris Descartes	5236		42780	150728	Tout public



► **Pour, in fine, permettre aux entreprises de la branche de diversifier leurs sources de recrutement**

- En complément, **actualiser le référencement des formations « big data »** opéré et initié dans le cadre de cette étude et le mettre à disposition des entreprises, pour que les entreprises puissent disposer d'information sur les formations possibles au sein desquelles elles peuvent recruter
  - l'alimenter par une auto-déclaration des organismes de formation
  - ce travail devrait a priori être porté par les organisations professionnelles, car un tel recensement d'organismes pourrait venir « percuter » les missions du Fafiec relatives à la qualité de l'offre de formation

**De façon générale (tant sur la formation initiale que continue)**

► **Ne pas se lancer dans une labellisation « Big data » de l'offre de formation**

- Car ce serait contraire à « l'esprit » du Big data (champ exploratoire, qui se réinvente en permanence) et le « Big data » n'obéit pas à une quelconque **norme** (dans la mesure où sa raison d'être est l'incertitude).
- Ce pourrait même être contreproductif : les technologies évoluant en permanence, les offres ainsi « labellisées » pourraient être frappées **d'obsolescence** assez rapidement.

► **Inciter les professionnels de la branche qui œuvrent sur des projets Big data à enseigner au sein des formations big data**

- Faire davantage connaître le **congé enseignement** (notamment à temps partiel<sup>12</sup>), mis en place par la branche en 2008 et réaffirmé par l'accord du tel que réaffirmé par l'accord du 28 juin 2011 relatif à la formation professionnelle, tant auprès des entreprises qu'auprès des salariés
- Si ce n'est déjà fait, définir en CPNE le Big data comme un domaine pouvant ouvrir **droit au maintien de la rémunération** et du paiement des cotisations sociales, dans le cadre de ce congé enseignement.

Ne pas y assortir de restriction sur les types d'établissements d'enseignement (par exemple, limiter cette possibilité de maintien de la rémunération uniquement si enseignement dans les cursus ingénieur) au sein desquels ces heures seraient réalisées

- Analyser l'opportunité et les **possibilités d'une prise en charge collective**, par la branche, des rémunérations des salariés dont l'employeur ne prendrait pas en charge le maintien de la rémunération
- **Ne pas cibler les seuls « techniciens »** (data scientists), mais aussi les profils davantage « métiers » (consultants, business analysts)

<sup>12</sup> Au maximum 8h par semaine (ou 40h par mois)





## Volet 3 : Développer la notoriété et l'attractivité de la branche auprès des sortants de formation

### Les grands objectifs poursuivis par ce volet

- **Favoriser l'orientation** des profils Big data **vers** les entreprises de **la branche** (et non pas vers les seules entreprises clientes)...
- ... et ce, quelle que soit la voie de « formation » de ces profils (cursus de formation initiale dédié « Big data », autodidacte, etc.)

### Pourquoi ? [rappel des constats]

- Des métiers qui, pour la plupart, peuvent être **exercés indifféremment** aussi bien au sein d'entreprises **clientes** que chez des **prestataires** (entreprises de la branche)
- Les entreprises de la branche sont **moins visibles** auprès des candidats potentiels que les entreprises clientes, car les champs d'application du big data sont chez ces dernières
- une **image parfois « ternie »** des conditions de travail au sein des entreprises de la branche

### Comment ?

- ▶ **Mener une campagne non pas sur l'attractivité des métiers du Big data, mais sur l'intérêt d'exercer ceux-ci au sein de la branche**
- **Profiter de « l'engouement » autour des métiers du Big data**, pour valoriser l'intérêt d'exercer ces métiers au sein d'une entreprise de la branche (et non au sein d'un secteur client)
  - nul besoin de mener une campagne sur l'attractivité des métiers du Big data dans l'absolu, puisque ceux-ci sont déjà largement considérés, par le grand public, comme « les métiers les plus sexy du 21ème siècle »
  - en revanche, il serait pertinent de tirer parti de l'oreille déjà attentive, et sans doute bienveillante et intéressée, du grand public et tout particulièrement des plus jeunes générations envers le Big data pour leur faire connaître les entreprises de la branche et les convaincre d'exercer ces métiers au sein de ces dernières
- Axer la communication tout particulièrement sur les avantages distinctifs des entreprises de la branche par rapport aux entreprises clientes et miser tout particulièrement sur les arguments susceptibles de « séduire » les générations Y et Z :
  - **l'esprit start up** et entrepreneurial
  - **les hiérarchies « flat »**, voire l'absence de hiérarchie, au sein de ces structures
  - le côté **stimulant et l'émulation** de contribuer à une start up



- **l'esprit « pionnier », « visionnaire »,** des entreprises de la branche sur le sujet du Big data
- **l'intérêt et la variété des projets** plus importants chez un prestataire que chez une entreprise cliente (la possibilité de découvrir plusieurs secteurs car on travaille pour plusieurs clients ; les missions qu'externalisent les clients sont celles à la plus haute valeur ajoutée ; etc.)
- En parallèle, **faire connaître la variété des métiers du Big data :**
  - les **profils IT et métier/business**, pourtant capitaux au sein des équipes big data, tendent à être effacés/ignorés derrière la popularité du data scientist
  - aussi, communiquer également sur la **variété des fonctions, des compétences et donc des profils recherchés**

### ► Développer des partenariats avec les établissements de formation initiale

**Plusieurs natures** de partenariats à développer :

- En incitant les professionnels de la branche à **enseigner** au sein des cursus de formation initiale (pour mémoire, cf. volet 1)
- En organisant des **hackathons** inter-établissements, qui pourraient judicieusement associer des cursus « techniques », « matheux » (IT, algorithmique) et « métiers » (écoles de commerce, etc.)
- Le **développement de la bourse à l'emploi du Syntec Numérique :**
  - faire connaître cette bourse auprès des établissements de formation initiale
  - en son sein, y développer un corner « Big data »
  - créer le pendant du volet « offre d'emploi », c'est-à-dire une CVthèque, pour que les établissements de formation puissent inciter leurs apprenants à y déposer leurs CV
- La **mise à disposition de technologies et d'outils** auprès des établissements de formation initiale :
  - **l'idée générale :**
    - \* mettre à disposition des établissements d'enseignement, mais surtout de leurs étudiants une **plateforme de ressources « big data »** qui permettrait **d'accéder à un panel de technologies/ressources « big data »**, celles actuellement « **standard** », mais aussi celles **open source** et celles récentes, **disruptives, en émergence** ou encore **en cours de développement** par les entreprises.
    - \* faire découvrir le **panel d'outils accessibles**, et pas seulement les seuls connus ou privilégiés par le corps enseignant

Les étudiants auraient alors accès à toute une palette de ressources, qu'ils pourraient mobiliser, utiliser, faire appel, lors de leurs projets étudiants, lors de leurs travaux sur cas, etc.

Cela permettrait donc **aux entreprises de la branche :**

- \* **de se faire connaître** par des étudiants au sein des cursus spécialisés, via leurs solutions








- \* de **permettre aux étudiants de se familiariser**, de travailler directement avec ces nouvelles ressources (et non seulement avec les ressources « standard »)
- \* voire, dans une **logique d'open innovation**, de bénéficier des retours des étudiants pour mettre au point, tester et enrichir en continu les ressources ainsi accessibles sur le marché
- **se rapprocher de la DGE et de l'Institut Mines Telecom**, notamment pour voir si cette idée pourrait trouver concrétisation par le développement et l'extension du projet **Teralab** (<https://www.teralab-datascience.fr/fr/>)
- **faire connaître cette plateforme aux entreprises de la branche**, afin que celles-ci acceptent d'être « fournisseuses » de ressources
- Des partenariats pour **élaborer des projets étudiants / fils rouges** au sein des cursus big data
  - dans la plupart des cursus spécialisés « Big data », tout au long de leur formation, les étudiants sont amenés à réaliser un « fil rouge », « projet étudiant » (la terminologie variant selon les établissements), commandité par une entreprise, et qui les amène de ce fait à se pencher sur une problématique métier concrète d'entreprise
  - la branche pourrait de la même façon commanditer des projets au sein de ces cursus. Ces sujets pourraient par exemple avoir trait :
    - \* à des problématiques sociétales (par exemple, en cherchant à tirer parti des données open source)
    - \* à des problématiques propres à la branche elle-même (par exemple, sur les sujets RH : exploitation des données du Fafiec ou de l'Opiiec)
    - \* voire à des problématiques propres aux organisations professionnelles elles-mêmes (comment détecter les sujets de préoccupation de nos adhérents ? etc.)
- **Organiser des défis et challenges, portés par les organisations professionnelles, sur les plateformes de concours et défis de Big data**
- **Pourquoi ?**

Sur ces plateformes de concours (kaggle.com et datascience.net), qui sont véritablement des viviers de recrutement, ce sont logiquement les entreprises clientes, et parmi elles, généralement des grands comptes, qui sont instigatrices des challenges (car elles disposent des données et y trouvent un intérêt direct). Ce sont donc celles-ci qui sont les plus visibles auprès de l'ensemble de la communauté et donc les plus à mêmes de recruter par la suite les talents ainsi dénichés.



A titre d'exemple, les 5 derniers challenges de la plateforme datascience.net au  
26/10/2015 :

### Challenges terminés

	<p><b>Catégorisation de produits pour le e-commerce</b></p> <p>Créer un système automatique de catégorisation des produits à partir de leur description.</p> <p>Challenge terminé</p>	<p><b>15 000 €</b> à gagner</p> <p>Ouvrir</p>
	<p><b>Data Science Game - Part 2</b></p> <p>Welcome to the 1st edition of the Data Science Game!</p> <p>Challenge terminé</p>	<p><b>Projet pédagogique</b></p> <p>Ouvrir</p>
	<p><b>Data Science Game</b></p> <p>Welcome to the 1st edition of the Data Science Game! <a href="http://www.datasciencegame.com">www.datasciencegame.com</a></p> <p>Challenge terminé</p>	<p><b>Projet pédagogique</b></p> <p>Ouvrir</p>
	<p><b>Construction d'un score d'appétence en vente croisée pour un produit d'assurance lors d'une campagne télémarketing</b></p> <p>Dans l'optique d'un meilleur ciblage de ses prospects lors d'une vente croisée, AXA souhaite acquérir un modèle de score d'appétence sur certains contrats d'assurance contre les accidents de la vie, à l'aide de données fournies par diverses filiales du groupe. L'objectif du challenge est ainsi de calculer un score prédisant à partir de ces données la souscription, lors de campagnes de télémarketing, d'un prospect à ce type de produits d'assurance.</p> <p>Challenge terminé</p>	<p><b>13 000 €</b> à gagner</p> <p>Ouvrir</p>
	<p><b>Prédiction de la fréquentation des gares SNCF en Ile-de-France</b></p> <p>Le challenge consiste à construire un modèle permettant d'estimer efficacement le nombre de voyageurs montant par gare pour un jour de semaine, dans les gares SNCF du réseau Ile-de-France, à partir de données open data caractérisant la gare d'un point de vue SNCF (nombre de lignes, nombre de trains par jour, services en gare, etc.), mais également de données externes (caractéristiques socio-démographiques de la ville, présence de transports urbains, etc.)</p> <p>Challenge terminé</p>	<p><b>1 800 € de voyages</b> à gagner</p> <p>Ouvrir</p>



- Comment ?
  - Identifier des « partenaires » (c'est-à-dire des acteurs qui rencontrent une question métier ou qui détiennent des données) qui permettraient de proposer ces challenges.
  - Par exemple :
    - \* des PME (quel que soit leur secteur d'activité) : ce sont essentiellement les grands comptes qui sont présents sur ces plateformes (sans doute du fait de l'investissement financier que cela implique pour les récompenses)
    - \* le monde associatif, les structures de l'Economie Sociale et Solidaire
    - \* se rapprocher d'Etalab, afin de créer des challenges à partir des données publiques, pour adresser des enjeux sociétaux
    - \* les collectivités locales
    - \* les pôles de compétitivité, etc.
  - Poster ces challenges sur les plateformes de concours en faisant apparaître clairement les organisations professionnelles de la branche comme leurs instigatrices.
  - Dans le texte introductif au challenge, en profiter pour « pitcher » sur les entreprises de la branche.
  - Y assortir une récompense, un gain à la clé, financé par les organisations professionnelles



0. Rappel des objectifs et de la méthode
1. Des innovations technologiques...
2. ...ouvrent de nouvelles fonctionnalités et usages pour les entreprises
3. Un sujet d'avenir, en émergence, autour duquel de nouvelles offres se développent
4. L'impact de ces bouleversements sur les modes de travail entre clients et prestataires
5. Les besoins en compétences liés au Big data
6. Les enjeux en termes de formation liés au Big data
7. Cartographie et analyse de l'offre de formation
8. Quantification des volumes d'emploi liés au Big data
9. Préconisations

---

## 10. Annexes

---




---

## Bibliographie

---

DELORT Pierre, 2015, *Le Big data*, 1<sup>ère</sup> édition. Paris : PUF, 127p. « Que sais-je ? »

CIGREF, octobre 2013, *Big data : la vision des grandes entreprises, opportunités et enjeux*, 16p.

CIGREF, octobre 2014, *Enjeux business des données, comment gérer les données de l'entreprise pour créer de la valeur ?*, 58 p.

Quantmetry, 2015, *Dessine-moi un data scientist, 8 leçons pour le recruter* [en ligne : <http://www.quantmetry.com/livre-blanc-dessine-moi-data-scientist/>]

Dataiku, 2015, *The 5 key Challenges to Building a Successful Data Lab – Addressing Challenges in a Data Science Laboratory Environment*, 21 p.

IDATE, mai 2012, *Cloud et Big data, le rôle clé des data centers*, 95p.

OPIIEC, *Référentiels métiers de la branche du Numérique, de l'Ingénierie, des études et du conseil et de l'événement – métiers du Numérique* [en ligne : <http://referentiels-metiers.opiiec.fr/page/1-cartographie-des-metiers-du-numerique> ]

France Stratégie (Commissariat Général à stratégie et à la prospective), Hamel Marie-Pierre et Marguerit David, novembre 2013, *Analyse des Big data, quels usages, quels défis ?*, note d'analyse, 12p.

KATALYSE (étude commanditée par le Fafiec), *Contrat d'études prospectives du secteur professionnel du Numérique*, rapport final, août 2013, 189 p.

Institut Montaigne, avril 2015, *Big data et objets connectés – faire de la France un champion de la révolution numérique*, 228 p.

« *Quelles compétences pour le Big data ?* », le JDN (18/06/2012) [en ligne : <http://www.journaldunet.com/solutions/expert/51837/quelles-competences-pour-le-big-data.shtml> ]

« *Big data, Big compétences* », la Tribune (07/05/2014) [en ligne : <http://www.latribune.fr/opinions/tribunes/20140507trib000828825/big-data-big-competences.html> ]



« *Devenir Data Scientist : les formations et les MOOC Big data* », databusiness.fr [en ligne : <http://www.data-business.fr/guide-meilleures-formations-mooc-big-data-devenir-data-scientist/> ]

« *Les métiers du Big data : comment recruter un data scientist ?* », databusiness.fr [en ligne : <http://www.data-business.fr/metiers-du-big-data-recruter-data-scientist-data-manager/> ]

« *Big data, les filières évoluent vers la double compétence* », educpros.fr (18/03/2013) [en ligne : <http://www.letudiant.fr/educpros/entretiens/big-data-les-filieres-evoluent-vers-la-double-competece.html> ]

« *Big data : quelles compétences pour valoriser les données ?* », séminaire Oresys [en ligne : <http://www.oresys.eu/video/episode-2-big-data-queelles-competences-pour-valoriser-les-donnees/> ]

« *Le Big data, un gros gisement d'emplois* », lavieeco.com [en ligne : <http://www.lavieeco.com/news/la-vie-eco-carrieres/le-big-data-un-gros-gisement-d-emplois-33400.html> ]

Dossier l'Etudiant : « *Big data : quand les données deviennent le nouveau pétrole de l'emploi* » [en ligne : <http://www.letudiant.fr/metiers/big-data-quand-les-donnees-deviennent-le-nouveau-petrole-de-l-emploi.html> ]

Rapport Mettling « *Transformation numérique et vie au travail* », septembre 2015, 69 p.

Markess, 2015, e-book *Big data, analytique & gestion des données, tendances 2015*

SAS, 2014, Survey Summary report : *What makes a Great Data Scientist ?* (en ligne : [http://www.sas.com/content/dam/SAS/en\\_gb/image/other1/events/WMAGDS/DataScientist-survey-report-web%20FINAL.pdf](http://www.sas.com/content/dam/SAS/en_gb/image/other1/events/WMAGDS/DataScientist-survey-report-web%20FINAL.pdf) )





---

## Recensement de l'offre de formation spécialisée Big data

---

Cf. tableaux dans les pages qui suivent



• Les formations « Big data »

Segment	Intitulé de la formation	Type	Niveau	Public cible	Durée (heures, années)	compétences principales visées / programme de formation	Exemples de métiers sur lesquels la formation peut déboucher	Nom de l'établissement	accessible en Cpro	accessible en contrat d'apprentissage	Page formation
Big data	Licence professionnelle Management des organisations Spécialité statistique	initiale/continue	Niveaux I et II			Ces professionnels sont des spécialistes du traitement, de l'analyse, de la synthèse	Ce professionnel peut prétendre aux métiers suivants : Chargé	Université de Bretagne Sud	oui	oui	<a href="http://www.rncp.cncp.gouv.fr/grand-public/visualisationFiche?format=f">http://www.rncp.cncp.gouv.fr/grand-public/visualisationFiche?format=f</a>
Big data	Licence Professionnelle Management des organisations spécialité Statistique et	initiale/continue	Niveaux I et II		3 ans	Ce professionnel conçoit et met en œuvre des bases et des entrepôts de données	Ce professionnel travaille dans les secteurs d'activité où le	Université Lille II - Droit et Santé	oui	oui	<a href="http://www.rncp.cncp.gouv.fr/grand-public/visualisationFiche?format=f">http://www.rncp.cncp.gouv.fr/grand-public/visualisationFiche?format=f</a>
Big data	Licence professionnelle Santé Spécialité statistique et	initiale/continue	Niveaux I et II		3 ans	Comprendre les problématiques et les enjeux dans les grands domaines de	l'industrie pharmaceutique, les organismes sociaux	Université de Bretagne-Sud	oui	oui	<a href="http://www.rncp.cncp.gouv.fr/grand-public/visualisationFiche?format=f">http://www.rncp.cncp.gouv.fr/grand-public/visualisationFiche?format=f</a>
Big data	Licence professionnelle Systèmes informatiques et logiciels	initiale/continue	Niveaux I et II		3 ans	Utiliser les logiciels de traitement de l'information (data mining), les techniques	Le "data miner " travaille en tant que chargé d'études,	Université Paris-Est Marne-la-Vallée	non	non	<a href="http://www.rncp.cncp.gouv.fr/grand-public/visualisationFiche?format=f">http://www.rncp.cncp.gouv.fr/grand-public/visualisationFiche?format=f</a>
Big data	Master Sciences et technologies Mention informatique Spécialité	initiale	Niveaux I et II			11033 : Mathématiques décision 31028 : Intelligence artificielle		Télécom ParisTech	non	non	<a href="http://www.intercariforef.org/formations/master-sciences-et-technologies-mention-">http://www.intercariforef.org/formations/master-sciences-et-technologies-mention-</a>
Big data	Master Sciences et technologies Mention informatique	initiale	Niveaux I et II	Le public de la première année est principalement	2 ans	les agents adaptatifs et les agents décisionnels l'apprentissage et la fouille	Entreprises high-tech Intelligence artificielle et aide à la décision -	Université Pierre et Marie Curie	non	non	<a href="http://www.intercariforef.org/formations/master-sciences-et-technologies-mention-">http://www.intercariforef.org/formations/master-sciences-et-technologies-mention-</a>
Big data	Master Sciences, technologies, santé Mention informatique	initiale	Niveaux I et II	L'inscription de tout étudiant de l'UVSQ ayant acquis l'année	2 ans			Université Versailles Saint-Quentin-en-Yvelines (Université Paris-Saclay)			<a href="http://www.intercariforef.org/formations/master-sciences%2C-technologies%2C-sant-mention-">http://www.intercariforef.org/formations/master-sciences%2C-technologies%2C-sant-mention-</a>
Big data	M2 Modèles et Technologies pour le décisionnel - Decision Support and Business			bac +4	1 an	Les modules couvrent des fondements théoriques comme la fouille de données	les débouchés du parcours DSBI sont les métiers liés aux	Université Paris-Saclay			<a href="http://www.universite-paris-saclay.fr/fr/formation/master/m2-modeles-et-technologies-pour-le-">http://www.universite-paris-saclay.fr/fr/formation/master/m2-modeles-et-technologies-pour-le-</a>
Big data	M2 Data & Knowledge (D&K)		Niveaux I et II	bac +4	1 an	The master program will equip students with the fundamental knowledge,	The Data&Knowledge track will prepare students for careers as	Université Paris-Saclay			<a href="http://www.universite-paris-saclay.fr/fr/formation/master/m2-data-knowledge-d-k-0">http://www.universite-paris-saclay.fr/fr/formation/master/m2-data-knowledge-d-k-0</a>
Big data	M2 Gestion de données dans un monde numérique - Data Management in a Digital World		Niveaux I et II	bac +4	1 an	Le parcours DataScale se focalise sur l'étude et la maîtrise des nouvelles	Les étudiants de cette thématique se verront ouvrir les portes de	Université Paris-Saclay			<a href="http://www.universite-paris-saclay.fr/fr/formation/master/m2-gestion-de-donnees-dans-un-">http://www.universite-paris-saclay.fr/fr/formation/master/m2-gestion-de-donnees-dans-un-</a>
Big data	M2 Data Sciences		Niveaux I et II	bac +4	1 an		Les compétences dans le domaine de l'apprentissage	Université Paris-Saclay			<a href="http://www.lix.polytechnique.fr/bigdata/mathbigdata/">http://www.lix.polytechnique.fr/bigdata/mathbigdata/</a>

Big data	M2 Traitement de l'Information et Exploitation des Données (RIED)		Niveaux I et II	bac +3	2 ans	Le master permet de former des scientifiques capables de concevoir des modélisations		université de Versailles-Saint-Quentin-en-Yvelines (Université Paris-Saclay)			<a href="http://www.tried.uvsq.fr/master-traitement-de-l-information-et-exploitation-des-donnees-tried-">http://www.tried.uvsq.fr/master-traitement-de-l-information-et-exploitation-des-donnees-tried-</a>
Big data	M2 Genomics Informatics and Mathematics for Health and Environment (GENIOMHE)		Niveaux I et II	bac +4	1 an		Designer and developer of databases and biology websites, in	Université Paris-Saclay			<a href="http://www.universite-paris-saclay.fr/fr/formatation/master/m2-genomics-informatics-and-">http://www.universite-paris-saclay.fr/fr/formatation/master/m2-genomics-informatics-and-</a>
Big data	Master Sciences, technologies, santé Mention informatique	initiale	Niveaux I et II			Donner à des étudiants possédant une bonne formation de base en	Chefs de projets en informatique, Concepteurs d'outils	Université Paris-Nord 13	non	non	<a href="http://www.rncp.cncp.gouv.fr/grand-public/visualisationFiche?format=f">http://www.rncp.cncp.gouv.fr/grand-public/visualisationFiche?format=f</a>
Big data	Master Sciences, technologies, santé Mention informatique Spécialité modèles pour le	initiale/continue	Niveaux I et II	bac +3	2 ans	Représentation des données Choisir un modèle de représentation de données		Université Lille 1 - sciences et technologies	non	non	<a href="http://formations.univ-lille1.fr/fr/fiche/di--400099">http://formations.univ-lille1.fr/fr/fiche/di--400099</a>
Big data	Master Sciences, technologies, santé mention statistique et mathématiques appliquées	initiale	Niveaux I et II	bac +3 et +	1 ou 2 ans	Former des analystes et chargés d'étude dominant les techniques et	Les secteurs d'activité concernés vont des études de marché, à la	Conservatoire national des arts et métiers (CNAM)	non	non	<a href="http://portail-formation.cnam.fr/ecole-siti/ingenierie-">http://portail-formation.cnam.fr/ecole-siti/ingenierie-</a>
Big data	Habilitation C_BOBIP_41 - SAP certified application associate - SAP	continue									<a href="https://training.sap.com/shop/app.php/certification/c_bobip_41-sap-certified-application-associate-">https://training.sap.com/shop/app.php/certification/c_bobip_41-sap-certified-application-associate-</a>
Big data	Habilitation C_BOWI_41 - SAP certified application associate - SAP	continue									<a href="https://training.sap.com/shop/certification/c_bowi_41-sap-certified-application-associate-sap-">https://training.sap.com/shop/certification/c_bowi_41-sap-certified-application-associate-sap-</a>
Big data	Habilitation C_TBW60_74 - SAP certified application associate -	continue									<a href="https://training.sap.com/shop/certification/c_tbw60_74-sap-certified-application-associate---">https://training.sap.com/shop/certification/c_tbw60_74-sap-certified-application-associate---</a>
Big data	Habilitation Certification Microsoft conception de solutions de	continue		Ce cours s'adresse aux clients qui sont intéressés à en	5 jours	Décrire les principales caractéristiques et avantages de libre-service BI.					<a href="https://www.microsoft.com/learning/en-us/course.aspx?translate=fr&amp;ID=2">https://www.microsoft.com/learning/en-us/course.aspx?translate=fr&amp;ID=2</a>
Big data	Master pro méthodes informatiques appliquées à la gestion des entreprises	initiale	Niveaux I et II	bac +3	2 ans	maîtrise du développement informatique : Orienté Objet, Distribué, Web,	gestionnaire de systèmes d'information, conseil	FEG - UFR Economie et gestion - antenne de Marseille, Université d'Aix-Marseille		oui	<a href="http://formations.univ-amu.fr/SPBIG5A.html">http://formations.univ-amu.fr/SPBIG5A.html</a>
Big data	Master pro méthodes informatiques appliquées à la gestion des entreprises	initiale	Niveaux I et II	bac +3	2 ans	maîtrise du développement informatique : Orienté Objet, Distribué, Web,	Ingénieur d'étude et de développement ; Gestionnaire	FEG - UFR Economie et gestion-site d'Aix en Provence, Université d'Aix-Marseille		oui	<a href="http://formations.univ-amu.fr/SPBIG5A.html">http://formations.univ-amu.fr/SPBIG5A.html</a>
Big data	MASTER Sciences, Technologies, Santé MENTION Informatique	initiale/continue	Niveaux I et II	bac +3	2 ans	statistiques décisionnelle (entrepôts de données et analyse en ligne, recherche	administrateur(trice) de base de données consultant(e) en	UFR des sciences et techniques, Université de Tours			<a href="http://sciences.univ-tours.fr/formations/master-sciences-technologies-sante-">http://sciences.univ-tours.fr/formations/master-sciences-technologies-sante-</a>
Big data	MASTER Sciences, Technologies, Santé MENTION Informatique	initiale/continue	Niveaux I et II	bac +3	2 ans	statistiques décisionnelle (entrepôts de données et analyse en ligne, recherche	administrateur(trice) de base de données consultant(e) en	INSA CVL - INSA Centre Val de Loire			
Big data	Master pro informatique spécialité informatique décisionnelle et logicielle	initiale/continue	Niveaux I et II	bac +3	2 ans	Maîtriser l'ingénierie d'entreprise (systèmes d'information intégrés, base	consultant(e) en informatique décisionnelle /	UFR de sciences et technologies, Université Blaise Pascal			<a href="http://www.univ-bpclermont.fr/formatation/formation/UBP-PROG19653.html">http://www.univ-bpclermont.fr/formatation/formation/UBP-PROG19653.html</a>
Big data	Master pro informatique spécialité informatique décisionnelle et logicielle	initiale/continue	Niveaux I et II	bac +3	2 ans	Maîtriser l'ingénierie d'entreprise (systèmes d'information intégrés, base	consultant(e) en informatique décisionnelle /	ISIMA - Institut supérieur d'informatique de modélisation et de leurs applications,			

Big data	Master 2 Informatique, Mobilité, Base de Données et Intégration de Systèmes	initiale/continue	Niveaux I et II	bac +3	2 ans	Identifier et formaliser les besoins de leurs clients en vu de l'informatisation	Développeur d'applications et de services notamment	IUP Miage, Université Nice - Sophia Antipolis		oui	<a href="http://unice.fr/formation/formati-on-initiale/sminf12123">http://unice.fr/formation/formati-on-initiale/sminf12123</a>
Big data	Master 2 Informatique, Mobilité, Base de Données et Intégration de Systèmes		Niveaux I et II	bac +3	2 ans	Identifier et formaliser les besoins de leurs clients en vu de l'informatisation	Développeur d'applications et de services notamment	Faculté de sciences, Université Nice - Sophia Antipolis			
Big data	Master professionnel mention Méthodes informatiques appliquées à la gestion,	initiale/continue	Niveaux I et II	bac +3	2 ans	Compétences techniques : Maîtrise des systèmes et	consultant(e) en informatique décisionnelle	Unité de formation d'informatique, Université de Bordeaux	oui	oui	<a href="http://www.u-bordeaux.fr/formation/PRMAMG132/master-professionnel-">http://www.u-bordeaux.fr/formation/PRMAMG132/master-professionnel-</a>
Big data	Master professionnel mention Méthodes informatiques appliquées à la gestion,	initiale/continue	Niveaux I et II	bac +3	2 ans	Compétences techniques : Maîtrise des systèmes et	consultant(e) en informatique décisionnelle	SFC - Service de formation continue collège Sciences et technologies de l'université de	oui		
Big data	Master recherche / professionnel mention Informatique, spécialité	initiale/continue	Niveaux I et II	bac +3	2 ans	Concevoir une modélisation informatique pour tous types de problèmes (III)	Informaticien d'étude ou d'exploitation > Cadre technique	Unité de formation d'informatique, Université de Bordeaux			<a href="http://www.u-bordeaux.fr/formation/PRMAIN111/master-recherche-">http://www.u-bordeaux.fr/formation/PRMAIN111/master-recherche-</a>
Big data	MSc Big Data Analytics for Business	initiale	Niveaux I et II	Bachelor d'une école de commerce ou d'un Bac +4	1 an	Les étudiants apprennent à résoudre des problèmes managériaux en se posant	Digital/Web Analyst Customer Analyst Data Scientist	IESEG - IESEG School of Management, Université catholique de Lille			<a href="http://www.ieseg.fr/programmes/msc-big-data/">http://www.ieseg.fr/programmes/msc-big-data/</a>
Big data	MSc Big Data Analytics for Business	initiale	Niveaux I et II	Bachelor d'une école de commerce ou d'un Bac +4	1 an	Les étudiants apprennent à résoudre des problèmes managériaux en se posant	Digital/Web Analyst Customer Analyst Data Scientist	IESEG School of Management, Université catholique de Lille			
Big data	Diplôme d'ingénieur de l'Ecole nationale de la statistique et de l'analyse de		Niveaux I et II		3 ans			ENSAI - Ecole nationale de la statistique et de l'analyse de l'information			
Big data	Master Big Data (ENSAI)	initiale	Niveaux I et II	M1 mathématiques, informatique ou équivalent ; Dossier,	1 an + stage de 5 mois	Skilled Data Scientist	Digital Marketing - Business Analytics - Risk Management	ENSAI - Ecole nationale de la statistique et de l'analyse de l'information			<a href="http://www.ensai.fr/formation/master-big-data.html">http://www.ensai.fr/formation/master-big-data.html</a>
Big data	Diplôme d'ingénieur de l'Ecole polytechnique universitaire de Lille,					Ingénieur spécialiste du traitement de l'information, informatique et statistique	chargé(e) d'études en marketing data manager	Polytech Lille - Ecole polytechnique universitaire de l'université Lille1, Université de	oui		<a href="http://www.polytech-lille.fr/genie-informatique-et-statistique-p117.html#.VY0a5fntlHw">http://www.polytech-lille.fr/genie-informatique-et-statistique-p117.html#.VY0a5fntlHw</a>
Big data	Master rech. informatique spécialité multimédia and data management	initiale	Niveaux I et II	bac +4	1 an	high level scientific and technological knowledge and working experience, in	work in R&D departments of companies	Polytech Nantes - Ecole polytechnique de l'université de Nantes, Université de Nantes			<a href="http://www.polytech.univ-nantes.fr/mdm/doku.php?id=start">http://www.polytech.univ-nantes.fr/mdm/doku.php?id=start</a>
Big data	Master économie et management des entreprises spécialité systèmes		Niveaux I et II		2 ans		consultant(e) en informatique décisionnelle	FSES - Faculté des Sciences Economiques et Sociales, Université de Lille 1			ses.univ-lille1.fr
Big data	Licence MIASHS – Parcours IDS (Informatique Décisionnelle et Statistique	initiale	Niveaux I et II		3 ans	donner aux étudiants les moyens de comprendre et de modéliser le monde réel		Université Lumière Lyon 2 Institut de la Communication Département d'Informatique et			<a href="http://dis.univ-lyon2.fr/fr/nos-formations/formations-initiales/licence-miashs-parcours-">http://dis.univ-lyon2.fr/fr/nos-formations/formations-initiales/licence-miashs-parcours-</a>
Big data	Première année du Master d'Informatique (M1) Tronc commun + options de pré-	initiale	Niveaux I et II	bac +3	1 an	organiser et analyser l'information pour préparer les décisions dans tous les		Université Lumière Lyon 2 Institut de la Communication Département d'Informatique et			<a href="http://dis.univ-lyon2.fr/fr/nos-formations/formations-initiales/premiere-annee-du-">http://dis.univ-lyon2.fr/fr/nos-formations/formations-initiales/premiere-annee-du-</a>
Big data	Master 2 Informatique Spécialité ECD (Extraction des Connaissances à partir	initiale	Niveaux I et II	bac +4	1 an	A l'issue de la formation, l'étudiant est en mesure de traiter un cas de fouille de	expert en fouille de données / systèmes d'information	Université Lumière Lyon 2 Institut de la Communication Département d'Informatique et			<a href="http://master-ecd.univ-lyon2.fr/">http://master-ecd.univ-lyon2.fr/</a>
Big data	Master 2 Informatique Spécialité IDS, parcours 2 OPSIE (Organisation et	initiale/continue	Niveaux I et II	bac +4	1 an	spécialistes capables d'assurer l'évaluation, l'organisation, la	Conseil et audit informatique • Responsable	Université Lumière Lyon 2 Institut de la Communication Département d'Informatique et			<a href="http://dis.univ-lyon2.fr/fr/nos-formations/formations-initiales/specialite-ids-parcours-2-">http://dis.univ-lyon2.fr/fr/nos-formations/formations-initiales/specialite-ids-parcours-2-</a>
Big data	Master 2 Informatique Spécialité IDS, parcours M2 IIDEE (Ingénierie	initiale/continue	Niveaux I et II	bac +4	1 an	analyser les stratégies des décideurs, trouver les réponses les plus	Fonctions : Consultant en BI, Administrateur des Data warehouses,	Université Lumière Lyon 2 Institut de la Communication Département d'Informatique et			<a href="http://dis.univ-lyon2.fr/fr/nos-formations/formations-initiales/specialite-ids-parcours-">http://dis.univ-lyon2.fr/fr/nos-formations/formations-initiales/specialite-ids-parcours-</a>
Big data	Master 2 Informatique Spécialité IDS, parcours M2 SISE (Statistique et	initiale	Niveaux I et II	bac +4	1 an	savoir comprendre un problème énoncé, le modéliser, y apporter une	Fonctions : statisticiens, chargés puis directeurs	Université Lumière Lyon 2 Institut de la Communication Département d'Informatique et			<a href="http://dis.univ-lyon2.fr/fr/nos-formations/formations-initiales/specialite-ids-parcours-">http://dis.univ-lyon2.fr/fr/nos-formations/formations-initiales/specialite-ids-parcours-</a>
Big data	MATHEMATIQUES ET APPLICATIONS, Parcours STATISTIQUE	initiale/continue	Niveaux I et II	L3 Mathématiques parcours SID	1 an			UFR MIG - UFR de mathématiques informatique gestion, Université Toulouse III			<a href="http://www.univ-tlse3.fr/m1-mathematiques-et-applications-parcours-statistique-informatique-">http://www.univ-tlse3.fr/m1-mathematiques-et-applications-parcours-statistique-informatique-</a>
Big data	Master pro ingénierie mathématique spécialité statistique, informatique,	initiale/continue	Niveaux I et II	Cette formation est ouverte aux étudiants ayant	1 an		biostatisticien(enne) gestionnaire de données cliniques	Faculté des sciences et technologies - Département mathématiques, Université	oui		<a href="http://masterim.univ-lyon1.fr/?q=node/3">http://masterim.univ-lyon1.fr/?q=node/3</a>

Big data	Master pro économétrie et statistique appliquée spécialité compétence		Niveaux I et II	bac +3	2 ans			Collegium droit, d'économie et de gestion, Université d'Orléans			
Big data	Master de mathématiques - parcours SPMA	initiale	Niveaux I et II	bac +3	2 ans	Modélisation aléatoire (modèles et outils de la statistique décisionnelle et	ingénieurs statisticiens ou mathématiciens	Collegium droit, d'économie et de gestion, Université d'Orléans			<a href="http://www.univ-orleans.fr/sciences-techniques/maths/master-de-">http://www.univ-orleans.fr/sciences-techniques/maths/master-de-</a>
Big data	Master informatique spécialité exploration informatique des données et décisionnel	initiale	Niveaux I et II	bac +3	2 ans	être capable d'appréhender et de suivre l'ensemble du cycle de vie du logiciel,	consultant(e) en informatique décisionnelle	Sup Galilée - Institut Galilée, Université Paris 13			
Big data	Master pro informatique spécialité informatiques appliquées à la gestion des	initiale/continue	Niveaux I et II	En MASTER 1 : Licence 3 parcours MIAge, Licence	2 ans	double compétence, à la fois en informatique (systèmes, réseaux, bases	Cadres spécialisés dans l'ingénierie des systèmes d'information	FST - Faculté des sciences et techniques, Université de Haute Alsace Mulhouse		oui	<a href="http://www.fst.uha.fr/?page=formations&amp;idformation=18">http://www.fst.uha.fr/?page=formations&amp;idformation=18</a>
Big data	Master mathématiques appliquées spécialité traitement statistique de			bac +3	2 ans		data manager ingénieur(e) mathématicien(ne)	ENSAE ParisTech - Ecole nationale de la statistique et de l'administration économique			<a href="http://www.onisep.fr/Ressources/Univers-Formation/Formations/Post-">http://www.onisep.fr/Ressources/Univers-Formation/Formations/Post-</a>
Big data	Mastère Spécialisé en Data Science	initiale/continue	Niveaux I et II	Le recrutement standard correspond à des	360 heures d'enseignement + un stage	3 piliers de compétences : • un pilier méthodologique, modèles et méthodes	analyste statisticien et data scientist	ENSAE ParisTech - Ecole nationale de la statistique et de l'administration économique	oui	oui	<a href="http://www.ensae.fr/formations-navhorizontale-172/mastre-specialis-et-cess-navhorizontale-">http://www.ensae.fr/formations-navhorizontale-172/mastre-specialis-et-cess-navhorizontale-</a>
Big data	Master pro informatique des organisations spécialité MIAge	initiale/continue	Niveaux I et II		2 ans	L'objectif de la formation est de former des cadres dans le domaine de	Les métiers de conseil en ingénierie des systèmes d'information	UFR 27 - UFR 27 Mathématiques et informatique, Université Panthéon-Sorbonne		oui	<a href="http://www.univ-paris1.fr/ws/ws.php?_cmd=getFormation&amp;_oid=UP1-">http://www.univ-paris1.fr/ws/ws.php?_cmd=getFormation&amp;_oid=UP1-</a>
Big data	Master pro méthodes informatiques appliquées à la gestion des entreprises	initiale/continue	Niveaux I et II				consultant(e) en informatique décisionnelle	UFR des sciences économiques et gestion (et IAE Gustave Eiffel), Université Paris-Est		oui	<a href="http://www.onisep.fr/Ressources/Univers-Formation/Formations/Post-">http://www.onisep.fr/Ressources/Univers-Formation/Formations/Post-</a>
Big data	Master pro biologie intégrative : physiologie et neurosciences spécialité	initiale		bac +3	2 ans		gestionnaire de données cliniques	Faculté des sciences et technologies - Département biologie, Université Claude			<a href="http://www.onisep.fr/Ressources/Univers-Formation/Formations/Post-">http://www.onisep.fr/Ressources/Univers-Formation/Formations/Post-</a>
Big data	Master pro informatique spécialité méthodes informatique appliquées à la			bac +3	2 ans		architecte des systèmes d'information consultant(e) en	Faculté des sciences et technologies - Département informatique, Université Claude			<a href="http://www.onisep.fr/Ressources/Univers-Formation/Formations/Post-">http://www.onisep.fr/Ressources/Univers-Formation/Formations/Post-</a>
Big data	Master informatique des organisations spécialité systèmes d'information et	initiale/continue		bac +4	1 an	extraire des Systèmes d'Information opérationnels (Bases de Données) les	consultant(e) en informatique décisionnelle	Faculté d'informatique, Université Toulouse I Capitole			<a href="http://www.ut-capitole.fr/formations/nos-diplomes/masters/master-m2-">http://www.ut-capitole.fr/formations/nos-diplomes/masters/master-m2-</a>
Big data	Licence pro santé spécialité statistique et informatique pour la santé	initiale/continue	Niveaux I et II	bac +2	1 an	• Comprendre les problématiques et les enjeux dans les différents	• Assistant-biostatisticien • Programmeur	IUT de Vannes, Université Bretagne-Sud			<a href="http://www.iutvannes.fr/licence-professionnelle-statistique-informatique-sante-morbihan/">http://www.iutvannes.fr/licence-professionnelle-statistique-informatique-sante-morbihan/</a>
Big data	Licence pro santé spécialité statistique et informatique décisionnelles pour la santé	initiale/continue	Niveaux I et II	bac +2	1 an	comprendre les problématiques et enjeux des différents domaines de	biostatisticien(enne) statisticien(ne)	IUT Paris Descartes, Université Paris Descartes			<a href="http://formations.parisdescartes.fr/fr-FR/1/diplome/P5-PROG6506/LICENCE%20PROFESSI">http://formations.parisdescartes.fr/fr-FR/1/diplome/P5-PROG6506/LICENCE%20PROFESSI</a>
Big data	LICENCE PROFESSIONNELLE Systèmes informatiques et	initiale/continue	Niveaux I et II	bac +2	1 an	de rechercher les données externes (via internet ou des enquêtes) susceptibles	chargé d'études data manager gestionnaire de	IUT Paris Descartes, Université Paris Descartes, Paris 5			<a href="http://formations.parisdescartes.fr/fr-FR/5/diplome/P5-PROG6409/LICENCE%20PROFESSI">http://formations.parisdescartes.fr/fr-FR/5/diplome/P5-PROG6409/LICENCE%20PROFESSI</a>
Big data	LICENCE PROFESSIONNELLE Systèmes informatiques et	initiale/continue	Niveaux I et II	bac +2	1 an	capacités à rechercher les données externes susceptibles d'être	Chargé d'études, Data manager, Gestionnaire de données, Veilleur	UFR de mathématiques, Université Paris-Est Marne-La-Vallée	oui		<a href="http://www.u-pem.fr/formations/loffre-de-formations/licences-">http://www.u-pem.fr/formations/loffre-de-formations/licences-</a>
Big data	Master pro informatique spécialité informatique décisionnelle	initiale/continue	Niveaux I et II	bac +3	2 ans	administration de bases de données, modélisation des problèmes (concevoir et	Administrateur de bases de données ou d'entrepôts de	UFR de Sciences site de Marseille Luminy, Université d'Aix-Marseille			<a href="http://formations.univ-amu.fr/SPSIN5E.html">http://formations.univ-amu.fr/SPSIN5E.html</a>
Big data	Master informatiques spécialité data mining and knowledge management	initiale	Niveaux I et II	bac +3	2 ans	knowledgeable about the approaches and the informatics tools necessary		Polytech Nantes - Ecole polytechnique de l'université de Nantes, Université de Nantes			<a href="http://web.polytech.univ-nantes.fr/1382971372533/1/fiche_formation/&amp;RH=1382971139">http://web.polytech.univ-nantes.fr/1382971372533/1/fiche_formation/&amp;RH=1382971139</a>
Big data	Master Business Analytics : Architecture, DATA Exploration and Optimization		Niveaux I et II	This master is available for those holding a BSc in	2 ans			EISTI - Ecole internationale des sciences du traitement de l'information			<a href="http://eisti.fr/fr/node/455">http://eisti.fr/fr/node/455</a>
Big data	Master MIAge	initiale/continue	Niveaux I et II	bac +3	2 ans			UFR Mathématiques et informatique, Université Paris Descartes		oui	<a href="http://www.math-info.univ-paris5.fr/~graff/MiAge/Accueil.html">http://www.math-info.univ-paris5.fr/~graff/MiAge/Accueil.html</a>
Big data	Master of Science Data Analysis and Pattern Classification		Niveaux I et II	To apply you must hold a Bachelor's	2 ans			Télécom SudParis			<a href="http://www.telecom-sudparis.eu/p_fr_formation_post_grade_MSc_1179.html?idm=25">http://www.telecom-sudparis.eu/p_fr_formation_post_grade_MSc_1179.html?idm=25</a>

Big data	Master of Science in Data Science and Business Analytics		Niveaux I et II	Students having obtained (or currently pursuing) a	1 an			ESSEC - Ecole supérieure des sciences économiques et commerciales		<a href="http://www.essec.edu/programs/advanced-masters-msmscs/msc-in-data-science-and-business-">http://www.essec.edu/programs/advanced-masters-msmscs/msc-in-data-science-and-business-</a>
Big data	Mastère spé. Big Data : analyse, management et valorisation responsable		Niveaux I et II	Bac+5 sans expérience ou Bac+4 avec 3 ans	15 mois	capacité à de mener des projets liés au Big Data	Data scientist Data strategist	GEM - Grenoble Ecole de Management		<a href="http://www.grenoble-em.com/mastere-specialise-big-data#Presentation">http://www.grenoble-em.com/mastere-specialise-big-data#Presentation</a>
Big data	MSc Data Management		Niveaux I et II	Professionnels en activité titulaires d'un diplôme universitaire	18 mois		Consultant informatique décisionnelle - big data	PSB - Paris School of Business (ex-ESG)		<a href="http://www.esgms.fr/msc/data-management">http://www.esgms.fr/msc/data-management</a>
Big data	MOOC : Fondamentaux pour le big data	continue						Télécom ParisTech		<a href="http://www.telecom-evolution.fr/fr/moocs/mooc-fondamentaux-pour-le-big-data">http://www.telecom-evolution.fr/fr/moocs/mooc-fondamentaux-pour-le-big-data</a>
Big data	CES Data Scientist			Cette formation s'adresse à des ingénieurs,	25 jour(s)	Maitriser les techniques récentes de gestion et d'analyse de grandes masses		Télécom ParisTech		<a href="http://www.telecom-evolution.fr/fr/formations-certifiantes/ces-data-scientist">http://www.telecom-evolution.fr/fr/formations-certifiantes/ces-data-scientist</a>
Big data	Big data : premiers succès et retours d'expérience	continue		Directeurs de programmes, chefs de projets big data et	1 jour	elle permet aux participants : d'apprécier l'étendue de cas d'usages réels possibles		Télécom ParisTech		<a href="http://www.telecom-evolution.fr/fr/formations-courtes/big-data-premiers-succes-">http://www.telecom-evolution.fr/fr/formations-courtes/big-data-premiers-succes-</a>
Big data	Big data : enjeux stratégiques et défis technologiques	continue		Décideurs et ingénieurs désireux d'appréhender les	2 jours	Explorer la manière dont le big data prend appui sur l'environnement		Télécom ParisTech		<a href="http://www.telecom-evolution.fr/fr/formations-courtes/big-data-enjeux-">http://www.telecom-evolution.fr/fr/formations-courtes/big-data-enjeux-</a>
Big data	Big Data : panorama des infrastructures et architectures distribuées	continue		Décideurs et ingénieurs désireux de comprendre	2 jours	Comprendre les paradigmes de calculs distribués et les solutions de passage à		Télécom ParisTech		<a href="http://www.telecom-evolution.fr/fr/formations-courtes/big-data-panorama-des-">http://www.telecom-evolution.fr/fr/formations-courtes/big-data-panorama-des-</a>
Big data	Data science : introduction au machine learning	continue		Ingénieurs et techniciens ayant besoin de méthodes	2 jours	Acquisition de bases solides dans le domaine du machine learning. Comprendre la		Télécom ParisTech		<a href="http://www.telecom-evolution.fr/fr/formations-courtes/data-science-introduction-">http://www.telecom-evolution.fr/fr/formations-courtes/data-science-introduction-</a>
Big data	Introduction à la sécurité du big data	continue		Cette formation s'adresse plus particulièrement aux	2 jours	Acquisition d'une vision globale des problèmes de sécurité liés au big data.		Télécom ParisTech		<a href="http://www.telecom-evolution.fr/fr/formations-courtes/introduction-la-securite-">http://www.telecom-evolution.fr/fr/formations-courtes/introduction-la-securite-</a>
Big data	Visualisation d'information (InfoVis)	continue		Ingénieurs, chefs de projets, data scientists intéressés	2 jours	Connaissance des principes de base de la visualisation des données et de		Télécom ParisTech		<a href="http://www.telecom-evolution.fr/fr/formations-courtes/visualisation-">http://www.telecom-evolution.fr/fr/formations-courtes/visualisation-</a>
Big data	Web sémantique et extraction de données	continue		Ingénieur, chef de projet, data scientist, intéressé par	2 jours	Comprendre les bases du web sémantique. Connaître les bases de la		Télécom ParisTech		<a href="http://www.telecom-evolution.fr/fr/formations-courtes/web-semantique-et-">http://www.telecom-evolution.fr/fr/formations-courtes/web-semantique-et-</a>
Big data	Introduction à R	continue		De bonnes connaissances en probabilités et des	2 jours	L'objectif de la formation est de présenter R et les tests d'hypothèses pour étudier		Télécom ParisTech		<a href="http://www.telecom-evolution.fr/fr/formations-courtes/introduction-r">http://www.telecom-evolution.fr/fr/formations-courtes/introduction-r</a>
Big data	Mastère spécialisé Big Data : gestion et analyse de données massives	initiale	Niveaux I et II	Bac + 5, bac + 4 et 3 ans d'expérience professionnelle ;	1 an	collecte, représentation, modélisation, analyse et traitement des données,		Télécom ParisTech		<a href="http://www.telecom-paristech.fr/formation-continue/masteres-specialises/big-">http://www.telecom-paristech.fr/formation-continue/masteres-specialises/big-</a>
Big data	Mastère spécialisé Big Data : analyse, management et valorisation responsable	initiale	Niveaux I et II	Jeunes diplômés de Grandes Ecoles (management ou	15 mois	Former des cadres : Rapidement opérationnels	Data scientist Data strategist Chief data officer	Grenoble INP-Ensimag Saint-Martin-d'Hères		<a href="http://ensimag.grenoble-inp.fr/formation/big-data-analyse-management-et-valorisation-">http://ensimag.grenoble-inp.fr/formation/big-data-analyse-management-et-valorisation-</a>
Big data	M2 MSIAM : Master of Science in Industrial and Applied Mathematics	initiale	Niveaux I et II	Anyone holding a first year of master (60 ECTS credits) in	2 ans			Grenoble INP-Ensimag Saint-Martin-d'Hères		<a href="http://msiam.imag.fr/doku.php">http://msiam.imag.fr/doku.php</a>
Big data	Certificat Big Data - UPMC - Faculté de mathématiques Paris		Niveaux I et II	Inscription en M2 mathématiques et applications	1 an + Stage d'application			Faculté de mathématiques, université Pierre-et-Marie-Curie		<a href="http://www.letudiant.fr/etudes/annuaire-enseignement-superieur/formation/formation-">http://www.letudiant.fr/etudes/annuaire-enseignement-superieur/formation/formation-</a>
Big data	MSc Data Management		Niveaux I et II	Professionnels en activité titulaires d'un diplôme universitaire	18 mois		Consultant informatique décisionnelle - big data	Ecole d'ingénieur généraliste en informatique et technologies du numérique, Groupe Efrei		<a href="http://www.efrei.fr/master-science-data-management#.VZKc_PntlHw">http://www.efrei.fr/master-science-data-management#.VZKc_PntlHw</a>
Big data	Majeure Business Intelligence	initiale	Niveaux I et II	Diplôme d'ingénieur, d'école de commerce ou		A l'issue de la formation, l'ingénieur est capable de :	Chef de projet SI décisionnel, Architecte SI décisionnel, Data	Ecole d'ingénieur généraliste en informatique et technologies du numérique, Groupe Efrei		<a href="http://www.efrei.fr/majeure-business-intelligence#.VZKbGfntlHw">http://www.efrei.fr/majeure-business-intelligence#.VZKbGfntlHw</a>
Big data	MSc Executive Big Data Analyst European Master of Science	initiale/continue	Niveaux I et II	APPLICANTS WITH A 3-YEAR BA, BSC OR BENG	16 mois	understand and use Information Systems for Big Data;		DSTI Sophia-Antipolis		<a href="https://www.datasciencetech.institute/master-programme-msc-executive-big-data-analyst/">https://www.datasciencetech.institute/master-programme-msc-executive-big-data-analyst/</a>

Big data	MSc Data Scientist Designer European Master of Science	initiale/continue	Niveaux I et II	APPLICANTS WITH A 3-YEAR BSC OR BENG	2 ans	analyse, design, build and monitor IT & Big Data architectures;		DSTI Paris / DSTI Sophia-Antipolis			<a href="https://www.datasciencetech.institute/master-programme-msc-data-scientist-designer/">https://www.datasciencetech.institute/master-programme-msc-data-scientist-designer/</a>
Big data	Master mathématiques et informatique appliquées aux sciences humaines et sociales	initiale/continue	Niveaux I et II	Admission en M1 : - Licence d'informatique,	2 ans	A la sortie de la formation, l'étudiant sera capable de concevoir, d'optimiser et	directeurs de projets en informatique, concepteurs d'outils et	UFR mathématiques, informatique, technologies, sciences de l'information et de la	oui		<a href="http://www.ufr-mitsic.univ-paris8.fr/presentation-bigdata-minfo.php">http://www.ufr-mitsic.univ-paris8.fr/presentation-bigdata-minfo.php</a>
Big data	DIGITAL MARKETING & DATA ANALYTICS, BI-CURSUS IIM/EMLV	initiale	Niveaux I et II	Titulaire d'un diplôme de niveau bac +3 (Licence) :	2 ans	connaître parfaitement les aspects techniques du web et du mobile / pas	Chef de projet Digital Traffic Manager Media Trader Web	EMLV - Ecole de Management Léonard de Vinci			<a href="http://www.emlv.fr/programmes/digital-marketing-data-analytics-bi-cursus-iimemlv/">http://www.emlv.fr/programmes/digital-marketing-data-analytics-bi-cursus-iimemlv/</a>
Big data	DIGITAL MARKETING & DATA ANALYTICS, BI-CURSUS IIM/EMLV	initiale	Niveaux I et II	Titulaire d'un diplôme de niveau bac +3 (Licence) :	2 ans	connaître parfaitement les aspects techniques du web et du mobile / pas	Chef de projet Digital Traffic Manager Media Trader Web	IIM - Institut de l'Internet et du Multimédia			<a href="http://www.iim.fr/cursus/digital-marketing-data-analytics-bi-cursus-">http://www.iim.fr/cursus/digital-marketing-data-analytics-bi-cursus-</a>
Big data	Diplôme d'Ingénieur ESIEA – Master bac +5	initiale	Niveaux I et II	Bacheliers	5 ans			ESIEA, ÉCOLE D'INGÉNIEURS DU MONDE NUMÉRIQUE			<a href="http://www.letudiant.fr/fiches/etudes/fiche/esiea-grande-ecole-d-ingenieurs.html">http://www.letudiant.fr/fiches/etudes/fiche/esiea-grande-ecole-d-ingenieurs.html</a>
Big data	Programme grande école - CFA Irisup' Clermont-Ferrand	initiale/continue	Niveaux I et II	Bac + 2 ; Concours	Après recrutement par concours à			CFA Irisup' Clermont-Ferrand	oui		<a href="http://www.cfa-irisup.fr/formations/commerce-management/item/37-ecole-">http://www.cfa-irisup.fr/formations/commerce-management/item/37-ecole-</a>
Big data	Ingénieur diplômé de l'ESAIP, spécialité informatique et réseaux	initiale	Niveaux I et II	Admission en cycle préparatoire 1ère année :	3-5 ans	collecter, agréger, stocker, accélérer le traitement des données via des	Data Scientist, Data Miner, Chief Data Officer, Consultant en	ESAIP Ecole d'ingénieurs			<a href="http://www.esaip.org/formations/ingenieur-informatique">http://www.esaip.org/formations/ingenieur-informatique</a>
Big data	BACHELOR INTERNATIONAL DATA MANAGER	initiale	Niveaux I et II	Admission après un baccalauréat toutes séries ou un	3 ans	Développement d'applications web mobiles pour soutenir la stratégie		ESAIP Ecole d'ingénieurs			<a href="http://www.esaip.org/formations/bachelor-data-manager">http://www.esaip.org/formations/bachelor-data-manager</a>
Big data	BACHELOR INTERNATIONAL DATA MANAGER	initiale	Niveaux I et II	Admission après un baccalauréat toutes séries ou un	3 ans	Développement d'applications web mobiles pour soutenir la stratégie		ESSCA - Ecole supérieure des sciences commerciales d'Angers			<a href="http://www.essca.fr/programmes/bachelor-masteres-specialises/bachelor-international-">http://www.essca.fr/programmes/bachelor-masteres-specialises/bachelor-international-</a>
Big data	Mastère spécialisé Mix Digital et Business Développement	initiale	Niveaux I et II	Bac + 5, bac + 4 et 3 ans d'expérience professionnelle ;	13 mois + stage entreprise 4	Développer des business models digitaux performants au plan	Marketing : Responsable E-commerce,	ESSCA - Ecole supérieure des sciences commerciales d'Angers			<a href="http://www.essca.fr/programmes/bachelor-masteres-specialises/mastere-specialise-mix-">http://www.essca.fr/programmes/bachelor-masteres-specialises/mastere-specialise-mix-</a>
Big data	Mastère spécialisé management digital et stratégie de marque	initiale	Niveaux I et II	Bac + 5, bac + 4 et 3 ans d'expérience professionnelle ;	13 mois + stage entreprise 4-6	définir, comprendre et relier les enjeux digitaux des marques avec leurs	Directeur ou manager marketing et digital Brand Digital Manager	ESSCA - Ecole supérieure des sciences commerciales d'Angers			<a href="http://www.essca.fr/programmes/bachelor-masteres-specialises/mastere-specialise-">http://www.essca.fr/programmes/bachelor-masteres-specialises/mastere-specialise-</a>
Big data	BACHELOR DIGITAL MANAGEMENT	initiale	Niveaux I et II	bacheliers de toutes filières, S, ES, L, bac technologiques	3 ans	compétences fondamentales en management ;	Responsable fonctionnel web, Digital Project	ESSCA - Ecole supérieure des sciences commerciales d'Angers			<a href="http://www.essca.fr/programmes/bachelor-masteres-specialises/digitalmanagement/">http://www.essca.fr/programmes/bachelor-masteres-specialises/digitalmanagement/</a>
Big data	Formation Développeur	initiale		bachelier	2 ans	Programmer	Développeur	Devschool			<a href="http://www.devschool.fr/?utm_source=fiche-etudes&amp;utm_medium=web&amp;utm_campaign=formation-developpeur">http://www.devschool.fr/?utm_source=fiche-etudes&amp;utm_medium=web&amp;utm_campaign=formation-developpeur</a>
Big data	Formation Développeur	initiale		bachelier	2 ans	Programmer	Développeur	IIM - Institut de l'Internet et du Multimédia			<a href="http://www.letudiant.fr/fiches/etudes/fiche/devschool-l-ecole-des-developpeurs.html">http://www.letudiant.fr/fiches/etudes/fiche/devschool-l-ecole-des-developpeurs.html</a>
Big data	Programme grande école - Groupe ESC Clermont Clermont-Ferrand	initiale/continue	Niveaux I et II	Admission principale en 1ère année :	3 ans	Maîtrise du système d'information de l'entreprise, notamment	Media et Web Data Manager (initiation) – Responsable CRM	Groupe ESC Clermont	oui	oui	<a href="http://www.esc-clermont.fr/programme/programme-grande-ecole-acces-prepa/">http://www.esc-clermont.fr/programme/programme-grande-ecole-acces-prepa/</a>
Big data	MSc in Digital Marketing and CRM	initiale	Niveaux I et II	Bachelor, bac + 4 ; Dossier, test d'anglais.	2 ans	Theoretical and practical knowledge of digital marketing, ability to analyze	Digital Marketing Manager and Brand Manager	IESEG School of Management, Université catholique de Lille			<a href="http://www.ieseg.fr/en/programs/msc-digital-marketing-crm/">http://www.ieseg.fr/en/programs/msc-digital-marketing-crm/</a>
Big data	MBA management et technologies de l'information et de la communication (TIC).	initiale	Niveaux I et II	Admission principale en 1ère année :	2 ans	Le MBA Management & Architecture des Systèmes d'Information forme des	Chef de projet, ingénieur de développement,	Ionis School of Technology and Management (STM), IONIS Education Group			<a href="http://www.ionis-stm.com/gestion-informatique/management-">http://www.ionis-stm.com/gestion-informatique/management-</a>
Big data	MBA MANAGEMENT & INFORMATIQUE DÉCISIONNELLE	initiale	Niveaux I et II	bac +2-5	2 ans	Le MBA Management & Informatique décisionnelle forme des experts	Ingénieur d'affaires, chef de projet, consultant, auditeur,	Ionis School of Technology and Management (STM), IONIS Education Group			<a href="http://www.ionis-stm.com/formation-informatique/master-">http://www.ionis-stm.com/formation-informatique/master-</a>
Big data	MBA MANAGEMENT, MARKETING & BIG DATA	initiale	Niveaux I et II	bac +2-5	2 ans	Avec le MBA Management, Marketing & Big Data, nous formons des managers et	Métiers de l'analyse marketing et de la gestion des big data :	Ionis School of Technology and Management (STM), IONIS Education Group			<a href="http://www.ionis-stm.com/formation-informatique/master-big-data">http://www.ionis-stm.com/formation-informatique/master-big-data</a>
Big data	MBA MANAGEMENT & WEB INTELLIGENCE	initiale	Niveaux I et II	bac +2-5	2 ans	Le MBA Management & Web Intelligence forme des experts techniques	Chef de projet e-business/e-marketing, ingénieur business	Ionis School of Technology and Management (STM), IONIS Education Group			<a href="http://www.ionis-stm.com/formation-informatique/master-">http://www.ionis-stm.com/formation-informatique/master-</a>

Big data	Ingénieur diplômé de Télécom Lille SPÉCIALISATIONS	initiale	Niveaux I et II	Admission en cycle préparatoire 1ère année :	5 ans			Télécom Lille, université Lille 1 - Sciences et technologies	oui	oui	<a href="http://www.telecom-lille.fr/specialisations-sciences-et-technologies">http://www.telecom-lille.fr/specialisations-sciences-et-technologies</a>
Big data	Licence professionnelle Systèmes informatiques et logiciels, Spécialité Nouvelles	initiale/continue	Niveaux I et II	Toute personne ayant la maîtrise d'un langage de	1 an	A partir d'un cahier des charges ou du collectage des besoins de l'entreprise,	Développement de sites web dynamiques. Développement ou	UFR des sciences, Aix-Marseille université			<a href="http://richard-ostrowski.eu/SILNTI/description.html">http://richard-ostrowski.eu/SILNTI/description.html</a>
Big data	Licence pro. gestion et traitement statistique de bases de données	initiale/continue	Niveaux I et II	Bac + 2 en informatique, mathématiques ;	1 an			UFR des sciences, Aix-Marseille université	oui		<a href="http://www.letudiant.fr/etudes/annuaire-enseignement-superieur/formation/formation-">http://www.letudiant.fr/etudes/annuaire-enseignement-superieur/formation/formation-</a>
Big data	Licence Informatique, Parcours Méthodes informatiques appliquées à la	initiale/continue	Niveaux I et II	Le public étudiant visé est un public intéressé par une	1 an	Maîtriser les concepts et outils de base en informatique de gestion.		UFR des sciences, Aix-Marseille université			<a href="http://lic-info.dil.univ-mrs.fr/parcours-MIAGE.html">http://lic-info.dil.univ-mrs.fr/parcours-MIAGE.html</a>
Big data	Licence mathématiques et informatique appliquées aux sciences humaines et	initiale/continue	Niveaux I et II	Bacheliers	3 ans			UFR des sciences, Aix-Marseille université			<a href="http://www.letudiant.fr/etudes/annuaire-enseignement-superieur/formation/formation-">http://www.letudiant.fr/etudes/annuaire-enseignement-superieur/formation/formation-</a>
Big data	Master pro. informatique décisionnelle	initiale/continue	Niveaux I et II	Admission en M1 : - Licence d'informatique	2 ans			UFR des sciences, Aix-Marseille université			<a href="http://informatique-sciences.univ-amu.fr/master-informatique?code=SPSIN5E">http://informatique-sciences.univ-amu.fr/master-informatique?code=SPSIN5E</a>
Big data	Master Informatique, Spécialité Sciences de l'Information et des Systèmes	initiale/continue	Niveaux I et II	bac +4	1 an	Aptitude à la recherche et à la veille technologique.		UFR des sciences, Aix-Marseille université			<a href="http://informatique-sciences.univ-amu.fr/master-informatique?code=SPSIN5B">http://informatique-sciences.univ-amu.fr/master-informatique?code=SPSIN5B</a>
Big data	<a href="#">Master Informatique Spécialité Recherche d'Informations multimédia (RIM)</a>	initiale/continue	Niveaux I et II	bac +4	1 an	L'objectif de la spécialité RIM est de former des spécialistes du traitement	Cette formation vise à produire des informaticiens	UFR des sciences, Aix-Marseille université			<a href="http://informatique-sciences.univ-amu.fr/master-informatique?code=SPSIN5F">http://informatique-sciences.univ-amu.fr/master-informatique?code=SPSIN5F</a>
Big data	Master 2 CCI (Compétence Complémentaire en Informatique)	initiale/continue	Niveaux I et II	depuis un master 2 ou un diplôme d'ingénieur (toutes	1 an	L'objectif de cette formation est de donner une double compétence en	Ingénieurs ou chefs de projets dans la compétence initiale	UFR des sciences, Aix-Marseille université			<a href="http://informatique-sciences.univ-amu.fr/master2-cci">http://informatique-sciences.univ-amu.fr/master2-cci</a>
Big data	Master Statistiques et traitement des données	initiale/continue	Niveaux I et II	M1: Etudiants titulaires d'une licence de	2 ans	Utiliser les principaux logiciels de statistiques (notamment SAS et R) , de	chercheur en mathématiques fondamentales ou	UFR sciences et technologies, université Blaise-Pascal			<a href="http://www.univ-bpclermont.fr/formation/formation/UBP-PROG19554.html">http://www.univ-bpclermont.fr/formation/formation/UBP-PROG19554.html</a>
Big data	Master Analyse et modélisation des données	initiale/continue	Niveaux I et II	M1: Etudiants titulaires d'une licence de Biologie	2 ans	Utiliser les bases de données et les logiciels de bioinformatique : annotation	Bioinformaticien en études, recherche et développement /	UFR sciences et technologies, université Blaise-Pascal			<a href="http://www.univ-bpclermont.fr/formation/formation/UBP-PROG19549.html">http://www.univ-bpclermont.fr/formation/formation/UBP-PROG19549.html</a>
Big data	Master pro. ingénierie de la santé - Management des Opérations Cliniques et Data	initiale/continue	Niveaux I et II	Admission en M1 : - Licence parcours ingénierie de la	2 ans	Réalisation du Data Management d'une base de données cliniques réelles	Gestionnaire de Bases de Données médicales (Data	Faculté des sciences, université de Montpellier	oui	oui	<a href="http://ingenierie-sante.edu.umontpellier.fr/formation/master/modm-gestion-">http://ingenierie-sante.edu.umontpellier.fr/formation/master/modm-gestion-</a>
Big data	Master bio-informatique, connaissances, données - Master STIC pour la santé,	initiale/continue	Niveaux I et II	Licence d'informatique, santé ou équivalent	2 ans			Faculté des sciences, université de Montpellier			<a href="http://www.lirmm.fr">http://www.lirmm.fr</a>
Big data	Master données, connaissances et langage naturel (DECOL)	initiale/continue	Niveaux I et II	Licence d'informatique ou équivalent	2 ans			Faculté des sciences, université de Montpellier			<a href="http://www.fdsweb.univ-montp2.fr/etudier-a-la-fds/catalogue-de-formation-675">http://www.fdsweb.univ-montp2.fr/etudier-a-la-fds/catalogue-de-formation-675</a>
Big data	Master STIC pour l'écologie et l'environnement, spécialité spatialisation et technologies	initiale/continue	Niveaux I et II	Licence sciences de la vie, sciences de la Terre,	2 ans	Développer la capacité à acquérir des données complexes d'écologie de	Thèse ou Ingénieurs en collectivités territoriales, structures	Faculté des sciences, université de Montpellier			<a href="http://www.masters-biologie-ecologie.com/blog/parcours-stic-ecologie/">http://www.masters-biologie-ecologie.com/blog/parcours-stic-ecologie/</a>
Big data	Licence pro. métiers de l'informatique : systèmes d'information et gestion de	initiale/continue	Niveaux I et II	L2 informatique, mathématiques, chimie, physique,	1 an	Concevoir des architecture de systèmes de Bases de données et de systèmes	Développeur ou Analyste ; Chef de projet ; Administrateur	Institut universitaire de technologie d'Evry, université d'Evry-Val d'Essonne	oui	oui	<a href="http://www.iut-evry.fr/genie-electrique-et-informatique-industrielle/lpageii/33-lp-systemes-">http://www.iut-evry.fr/genie-electrique-et-informatique-industrielle/lpageii/33-lp-systemes-</a> <a href="http://www.cfa-">http://www.cfa-</a>
Big data	Master innovation, marché et science des données	initiale/continue	Niveaux I et II	M1 en économie, mathématiques, informatique ou	1 an	La formation s'adresse à des étudiants désireux d'acquérir une expertise à	Analyste Marketing relationnel, Analyste	CFA EVE Evry		oui	<a href="http://eve.fr/mod_formation/public/detail/id/98">eve.fr/mod_formation/public/detail/id/98</a>
Big data	Licence pro. métiers de l'informatique : systèmes d'information et gestion de	initiale/continue	Niveaux I et II	Bac + 2 informatique, gestion scientifique ou technique, mise à	1 an	Le diplôme a pour objet de former aux comportements et techniques de conception	Analyste d'exploitation informatique, - Analyste en	CFA EVE Evry		oui	<a href="http://www.cfa-eve.fr/mod_formation/public/detail/id/38">http://www.cfa-eve.fr/mod_formation/public/detail/id/38</a>
Big data	Licence pro. bases de données	initiale/continue	Niveaux I et II	DUT informatique, BTS SIO, L2 sciences et technologie avec	1 an	développer et administrer des bases de données, d'extraire des données	développeur d'applications autour des bases de données,	Savoie Mont-Blanc - IUT Annecy-le-Vieux	oui		<a href="http://www.iut-acv.univ-savoie.fr/index.php?id=62">http://www.iut-acv.univ-savoie.fr/index.php?id=62</a>
Big data	Licence pro. intégration des systèmes voix et données (ISVD)	initiale/continue	Niveaux I et II	BTS SIO, DUT réseaux et télécommunications,	1 an			Savoie Mont-Blanc - IUT Annecy-le-Vieux	oui		<a href="http://www.letudiant.fr/etudes/annuaire-enseignement-superieur/formation/formation-">http://www.letudiant.fr/etudes/annuaire-enseignement-superieur/formation/formation-</a>



Big data	Master Mathématiques et applications parcours - probabilités et statistiques des	initiale/continue	Niveaux I et II	Admission en M1 : - Licence de mathématiques de	2 ans	Acquisition de connaissance en probabilités appliquées	ingénieur financier ou informaticien, ou consultant	Paris-Est Créteil - Faculté des sciences et technologie			<a href="http://sciences-tech.u-pec.fr/formations/master-mathematiques-et-applications-">http://sciences-tech.u-pec.fr/formations/master-mathematiques-et-applications-</a>
Big data	Master mathématiques pour la science des masses de données	initiale	Niveaux I et II	Admission en M1 : - Bac + 3 ; Dossier. Admission en M2 :	4 semestres			Ecole polytechnique (EP) Palaiseau			
Big data	Master systèmes d'information pour l'aide à la décision (SIAD)	initiale/continue	Niveaux I et II	Admission en M1 : - Licence d'informatique ou	2 ans	statistiques décisionnelle (entrepôts de données et analyse en ligne, recherche	cadres supérieurs dans le domaine des systèmes d'information	Tours - antenne universitaire de Blois			<a href="http://www.univ-tours.fr/formations/master-sciences-technologies-sante-">http://www.univ-tours.fr/formations/master-sciences-technologies-sante-</a>
Big data	MASTER Sciences, Technologies, Santé MENTION Informatique SPECIALITE	initiale/continue	Niveaux I et II	Vous pouvez postuler en M2 si vous êtes titulaire ne	2 ans	la recherche Opérationnelle la reconnaissance des Formes	services Recherche & Développement	Tours - antenne universitaire de Blois			<a href="http://www.univ-tours.fr/formations/master-sciences-technologies-sante-">http://www.univ-tours.fr/formations/master-sciences-technologies-sante-</a>
Big data	Diplôme ICD	initiale/continue	Niveaux I et II	Admission principale en 1ère année : - Bac ; Dossier, tests,	5 ans			Institut international du commerce et du développement, Groupe IGS (ICD) Blagnac	oui		<a href="http://www.icd-ecoles.com/cycle-master/">http://www.icd-ecoles.com/cycle-master/</a>
Big data	MBA Conseil en Stratégie commerciale et Business Intelligence	initiale/continue	Niveaux I et II	Les candidats titulaires (ou en cours d'obtention)	1 an	Définir une politique de maîtrise de l'information Concevoir et déployer une	Créateur d'entreprise Repreneur d'entreprise Responsable d'un	Institut international du commerce et du développement, Groupe IGS (ICD) Blagnac	oui		<a href="http://www.icd-ecoles.com/conseil-et-strategie-commerciale/">http://www.icd-ecoles.com/conseil-et-strategie-commerciale/</a>
Big data	Licence professionnelle DATA MINING- Décision et traitement de l'information	initiale/continue	Niveaux I et II	bac +2	1 an	Esprit d'analyse - Rigueur intellectuelle - Aimer le travail d'équipe	Data Miner : chargé d'études, gestionnaire de données, veilleur	Centre de formation d'apprentis / CFA Descartes ADEFSA Marne-la-Vallée		oui	<a href="http://www.difcam.com/?pg=enseignement-diplomes&amp;s=niveau&amp;cat=BAC%2B">http://www.difcam.com/?pg=enseignement-diplomes&amp;s=niveau&amp;cat=BAC%2B</a>
Big data	Mastère – Analyste en Data Marketing	initiale/continue	Niveaux I et II	Bac + 4 secteur scientifique, technique ou	1 an	Le quotidien d'un analyste en Data Marketing est: - de veiller à la sauvegarde de		Ecole supérieure de commerce et d'économie numérique (ESCEM)	oui		<a href="http://www.escecm.fr/les-programmes/mastere-specialises/mastere-spe-data-">http://www.escecm.fr/les-programmes/mastere-specialises/mastere-spe-data-</a>
Big data	Master marketing et vente, spécialité marketing et relation client, parcours études et	initiale	Niveaux I et II	Admission en M1 : - Licence de gestion parcours	2 ans			Institut des sciences de l'entreprise et du management, université de Montpellier - ISEM	oui	oui	<a href="http://www.umontpellier.fr/formation/offre-de-formation/">http://www.umontpellier.fr/formation/offre-de-formation/</a>
Big data	Data Science (DSC)	initiale	Niveaux I et II	Admission en M1 : - Dossier. Pas d'admission directe	2 ans	In this programme, students will learn about scalable data collection techniques, data	A professional for a career in a highly innovative area: data	Faculté des sciences, université de Nice Sophia-Antipolis			<a href="http://www.masterschool.eitictlabs.eu/programmes/data-science/">http://www.masterschool.eitictlabs.eu/programmes/data-science/</a>
Big data	Master innovation, marché et science des données	initiale/continue	Niveaux I et II	Admission en M1 : - Dossier. Admission en M2 :	2 ans			UFR des sciences de l'homme et de la société, université d'Evry-Val d'Essonne		oui	
Big data	Master mobilité, bases de données et intégration de systèmes (MBDS)	initiale/continue	Niveaux I et II	Admission en M1 : - Licence d'informatique ;	2 ans			Faculté des sciences, université de Nice Sophia-Antipolis	oui	oui	<a href="http://www.mbds-fr.org">http://www.mbds-fr.org</a>
Big data	MBA Management Stratégique de l'Analyse de données Marketing Massives	initiale	Niveaux I et II	Titulaires d'un Bac + 4 ou Master d'Ecole de Commerce ou Ecole		Développer des compétences avancées en mathématiques et en	Manager de la stratégie commerciale	EPSI			<a href="http://www.epsi.fr/Programmes/MBA-Big-Data">http://www.epsi.fr/Programmes/MBA-Big-Data</a>
Big data	MBA Management Stratégique de l'Analyse de données Marketing Massives	initiale	Niveaux I et II					IDRAC			<a href="http://www.epsi.fr/Programmes/MBA-Big-Data">http://www.epsi.fr/Programmes/MBA-Big-Data</a>
Big data	Master sciences et technologies de l'information et de la communication,	initiale/continue	Niveaux I et II	Admission en M1 : - Licence MIAGE, informatique ou		● Conception et gestion d'une stratégie de communication	● Responsable de communication ● Chef de projet	Institut régional de formation par alternance. Association pour le développement de la formation	oui	oui	<a href="http://www.irfa-apisup.fr/index.php/metiers/info/master-miage">http://www.irfa-apisup.fr/index.php/metiers/info/master-miage</a>
Big data	IBM POWER SYSTEMS II _ MISE EN OEUVRE DE LA VIRTUALISATION _HMC_	continue						ARROW ECS			<a href="https://edu.arrowecs.eu/fr/trainings/open/trainings.html">https://edu.arrowecs.eu/fr/trainings/open/trainings.html</a>
Big data	Science des données	continue				Formulate context-relevant questions and hypotheses to drive data scientific research		Coursera			<a href="https://www.coursera.org/specialization/jhudatascience/1/overview">https://www.coursera.org/specialization/jhudatascience/1/overview</a>
Big data	Formations Big Data	continue		Aux futurs Data Managers : managers et chefs de projet	70h	Stratégie : fournir une compréhension approfondie des enjeux stratégiques et	Data Managers, Data Scientists	Data Business			<a href="http://www.data-business.fr/formations-big-data/">http://www.data-business.fr/formations-big-data/</a>
Big data	FILIERE BI DECISIONNEL	continue						A J C INGENIERIE			
Big data		continue						A6 CONSEIL			

Big data		continue						AFFIL COMP SERV BUSINEE PROCESS SOLUTI			
Big data		continue						ALTI			
Big data	Talend Open Studio for Data Integration : Intégration de données	continue						ALTIC SARL			<a href="http://www.altic.org/formation">http://www.altic.org/formation</a>
Big data	ADMINISTRATEUR POUR HADOOP HADADM1	continue						AMADEUS			<a href="http://www.amadeus.com/web/amadeus/fr_FR-FR/Page-daccueil-Amadeus-Home/Ressources-et-">http://www.amadeus.com/web/a madeus/fr_FR-FR/Page-daccueil- Amadeus-Home/Ressources-et-</a>
Big data	CHEF DE PROJET BI BIG DATA	continue						ANEMA			<a href="http://www.anema.fr/formations-elearning.html">http://www.anema.fr/formations- elearning.html</a>
Big data	Formation Master Data Management	continue						APOLLO FORMATION			<a href="http://www.apollo-formation.com/catalogue-formation/">http://www.apollo- formation.com/catalogue- formations/</a>
Big data	CERTIFICATION MICROSOFT CONCEPTION DE SOLUTIONS DE BUSINESS INTELLIGENCE	continue						APV GROUPE			<a href="http://www.apvgroupe.fr/apvgroupe/formations/formations.php">http://www.apvgroupe.fr/apvgro upe/formations/formations.php</a>
Big data	TRONC COMMUN INSTALLATION BI QLICVIEW	continue						CEGID			<a href="http://www.cegid.fr/erp/erp-business-cegid-logiciel-finance-et-relation-client/p-367.aspx">http://www.cegid.fr/erp/erp- business-cegid-logiciel-finance-et- relation-client/p-367.aspx</a>
Big data		continue						CENS INFO			
Big data	Exploitation et administration de Centreon BI	continue						CENTREON			<a href="https://www.centreon.com/fr/support-technique-expertise/formation/">https://www.centreon.com/fr/su pport-technique- expertise/formation/</a>
Big data	R et Big Data	continue		Une formation à R ou bien connaître le livre 2 j Statistiques avec R				CEPE			<a href="http://www.lecepe.fr/formations/data-science/r-et-big-data-155.html">http://www.lecepe.fr/formations/ data-science/r-et-big- data-155.html</a>
Big data	Le Certificat de Data Scientist	continue		Bien connaître les méthodes statistiques de 15 j				CEPE			<a href="http://www.lecepe.fr/certificats/data-scientist/">http://www.lecepe.fr/certificats/d ata-scientist/</a>
Big data	Bases de données NoSQL : Enjeux et Solutions	continue						CREDIT MUTUEL - ORGANISME DE FORMATION / Le Bischenberg			<a href="http://www.lebischenberg.com/fr/index.html">http://www.lebischenberg.com/fr /index.html</a>
Big data	Formation BIG DATA	continue						DATAIKU			<a href="https://learn.dataiku.com/">https://learn.dataiku.com/</a>
Big data	BIG DATA AVEC HADOOP	continue						DBF - DEVELOPP BUSINESS FORMATIONS			
Big data	SAP BUSINESS INTELLIGENCE	continue						FAYAT			
Big data	CONSULTANT IPE SAP BI BIG DATA NIV1	continue						FORMATION INFORMATIQUE ET TECHNIQUE - FITEC			<a href="http://www.fitec.fr/metiers.html">http://www.fitec.fr/metiers.html</a>
Big data	Formation Big Data	continue						GROUPE CYRES - INGENSI			<a href="http://www.ingensi.com/solutions-big-data/formation-big-data">http://www.ingensi.com/solution s-big-data/formation-big-data</a>
Big data	Formation Big-Data Hadoop & MongoDB	continue						I'FORM			<a href="http://www.iform.fr/menu/3/3/Engagement">http://www.iform.fr/menu/3/3/E ngagement</a>
Big data	PROTEGER LES BASES DE DONNEES ET BIG DATA	continue						INST EUR ENTR PROPRIETE INTELLECTUELLE			

Big data	DUT STID	initiale/continue	Niveau III	Bac ES, S, STI2D, STL, STMG, DAEU B ; Dossier	2 ans	explorer un marché, cibler une clientèle ; optimiser l'expérimentation	chargé d'études	Institut Universitaire de Technologie de Metz (IUT Metz)	oui	oui	<a href="http://www.iut.univ-metz.fr/index.php/les- formations/le-dut/dut-">http://www.iut.univ-metz.fr/index.php/les- formations/le-dut/dut-</a>
Big data	LP Statistique et informatique décisionnelle, sciences de la vie	initiale	Niveaux I et II	Bac +2	1 an	Acquérir une triple compétence en : statistique,		Institut Universitaire de Technologie de Metz (IUT Metz)			<a href="http://www.iut.univ-metz.fr/index.php/les- formations/la-licence-">http://www.iut.univ-metz.fr/index.php/les- formations/la-licence-</a>
Big data	LP Statistique et informatique décisionnelle, marketing	initiale/continue	Niveaux I et II	Bac +2	1 an	Acquérir une triple compétence en : statistique,		Institut Universitaire de Technologie de Metz (IUT Metz)		oui	<a href="http://www.iut.univ-metz.fr/index.php/les- formations/la-licence-">http://www.iut.univ-metz.fr/index.php/les- formations/la-licence-</a>
Big data	Devenez data-scientist	continue		Niveau intermédiaire Business Analyst Consultant	13 cours vidéos & leur quiz 4 heures	Fondamentaux du langage Python Librairies Numpy / Pandas		LE WAGON			<a href="https://ondemand.lewagon.org/tracks/devenez-data-scientist/go">https://ondemand.lewagon.org/tracks/devenez-data-scientist/go</a>
Big data	EXPERT MICROSOFT BI & BASES DE DONNEES	continue						LEE YOUNG MIN - ELDAY INSTITUT			
Big data	sas business intelligence data integration studio architecte	continue						L'ENTREPRISE			
Big data	SI 2.0 (BIG DATA, CLOUD...)	continue						MICROPOLE INSTITUT			<a href="http://www.micropole-institut.com/metiers-de-l-informatique/si-2-0-big-data-">http://www.micropole-institut.com/metiers-de-l-informatique/si-2-0-big-data-</a>
Big data	Utilisation de Odoo et Business Intelligence avec Pentaho	continue						MIND AND GO			
Big data	FORMATION BIG DATA	continue						NATHALIE COLUMELLI - FINANCE TRAINING			<a href="http://www.finance-training.fr/prochaines- formations-cfa-finance/">http://www.finance-training.fr/prochaines- formations-cfa-finance/</a>
Big data	administrer la plateforme Hadoop 2.X	continue						OCTO-TECHNOLOGY - OCTO-ACADEMY			<a href="http://www.octo.academy/fr/stages-pratiques/d/13-big-data">http://www.octo.academy/fr/stages-pratiques/d/13-big-data</a>
Big data	HADOOP AND BIG DATA TRAINING	continue						ORGANISME HORS TERRITOIRE NATIONAL			
Big data	Management du SI et des Projets Informatique	continue						PERFORMAT PLUS			
Big data	Panorama comparatif des solutions BI Open Source	continue						SMILE			<a href="http://www.smile.fr/Services/Formation">http://www.smile.fr/Services/Formation</a>
Big data	CURSUS BI ( POUR LA SG)	continue						TERADATA FRANCE			<a href="http://fr.teradata.com/Layouts/Su bPageLayout.aspx?pageid=12884910953&amp;LangType=1036&amp;LangSel">http://fr.teradata.com/Layouts/Su bPageLayout.aspx?pageid=12884910953&amp;LangType=1036&amp;LangSel</a>
Big data	BIG DATA AVEC HADOOP	continue						VALTECH TRAINING			<a href="http://www.valtech-training.fr/formation/">http://www.valtech-training.fr/formation/</a>
Big data	FORMATION BIG DATA	continue						YSANCE BI SERVICES			<a href="http://www.ysance.com/nos-offres.html">http://www.ysance.com/nos-offres.html</a>
Big data	Master Web intelligence	initiale	Niveaux I et II	Admission en M1 : - Licence d'informatique ou	2 ans		Ingénieur, Analyste-Programmeur, Ingénieur de Recherche	Faculté des sciences et techniques, université Jean-Monnet - Saint-Etienne - FST			
Big data	Erasmus Mundus master course in Data Mining and Knowledge Management	initiale	Niveaux I et II	The Master in DMKM is aimed at students from all over the	2 ans	• perform in depth analysis of information requirements for		Polytech Nantes			<a href="http://www.intercariforef.org/formations/master-sciences%2C- technologies%2C-sant-mention-">http://www.intercariforef.org/formations/master-sciences%2C- technologies%2C-sant-mention-</a>
Big data	Master 2 Spécialité : Extraction des connaissances à partir des données	initiale/continue	Niveaux I et II	Master 1 en informatique, ou en analyse de données	1 an	Génie logiciel et programmation orientée-objet (UML, Java, .Net),	Direction des systèmes d'information Études et	Polytech Nantes			<a href="http://web.polytech.univ-nantes.fr/SI00236/0/fiche_for mation/">http://web.polytech.univ-nantes.fr/SI00236/0/fiche_for mation/</a>
Big data	Big data et data science pour les managers	continue		salariés		mesurer les enjeux comprendre les technologies		DSTI			<a href="https://www.datasciencetech.institute.fr/formation-big-data-et-data-science-pour-les-managers/">https://www.datasciencetech.institute.fr/formation-big-data-et-data-science-pour-les-managers/</a>

Big data	Big data	MOOC			3 semaines	Organize, analyze, and interpret any large, complex dataset, and apply your		Coursera / Université de Californie		<a href="https://fr.coursera.org/specializations/big-data">https://fr.coursera.org/specializations/big-data</a>
Big data	Fondamentaux pour le Big data	MOOC			8 semaines	Programmation Python, Limites des bases de données relationnelles,		FUN / TelecomParisTech		
Big data	Introduction to Big data	MOOC			3 semaines	Welcome to Big Data The Emerging Big Data Stack		Coursera / Université de Californie		<a href="https://fr.coursera.org/learn/intro-to-big-data">https://fr.coursera.org/learn/intro-to-big-data</a>
Big data	Machine Learning with Big data	MOOC				robust set of skills that will allow you to process, analyze, and		Coursera / Université de Californie		<a href="https://fr.coursera.org/learn/machinelearningwithbigdata">https://fr.coursera.org/learn/machinelearningwithbigdata</a>
Big data	Hadoop	MOOC				savoir manipuler Hadoop		Coursera / Université de Californie		<a href="https://fr.coursera.org/learn/hadoop">https://fr.coursera.org/learn/hadoop</a>
Big data	Introduction to graph Analytics	MOOC				model a problem into a graph database and perform analytical tasks		Coursera / Université de Californie		<a href="https://fr.coursera.org/learn/graph-analytics">https://fr.coursera.org/learn/graph-analytics</a>
Big data	Big data et business analytics	initiale / continue	Niveaux I et II	étudiants / salariés / demandeurs d'emploi	entre 12 et 16 mois	Solve business problems by applying analytics methodologies and		HEC	x	<a href="http://www.mba.hec.edu/Learning-Experience/Course-Index/Big-Data-and-Business-Analytics">http://www.mba.hec.edu/Learning-Experience/Course-Index/Big-Data-and-Business-Analytics</a>
Big data	Big data	initiale/continue	Niveaux I et II	étudiants / salariés / demandeurs d'emploi		*Systèmes distribués et architectures de type cloud ;		ENSIMAG et EMSI Grenoble		<a href="http://ensimag.grenoble-inp.fr/lecole/le-mastere-big-data-ensimag-ouvre-en-octobre-2014-">http://ensimag.grenoble-inp.fr/lecole/le-mastere-big-data-ensimag-ouvre-en-octobre-2014-</a>
Big data	Big Data : Gestion et analyse des données massives	initiale/continue	Niveaux I et II	étudiants / salariés / demandeurs d'emploi						
Big data	CES (Certificat d'Etudes spécialisées) Data Scientist			salariés / demandeurs d'emploi	25 jours	Introduction à l'apprentissage statistique □ Objectifs et enjeux de		Télécom ParisTech		<a href="http://www.telecom-evolution.fr/fr/formations-certifiantes/ces-data-scientist">http://www.telecom-evolution.fr/fr/formations-certifiantes/ces-data-scientist</a>
Big data	Big data : premiers sujets et retours d'expérience	continue		Directeurs de programmes, chefs de projets big data	1 jour	Cette formation s'adresse à des responsables décideurs travaillant sur un		Télécom ParisTech		<a href="http://www.telecom-evolution.fr/fr/formations-courtes/big-data-premiers-">http://www.telecom-evolution.fr/fr/formations-courtes/big-data-premiers-</a>
Big data	Big data : enjeux stratégiques et défis technologiques	continue		Décideurs et ingénieurs désireux d'appréhender les	2 jours	Introduction □ Big data : pourquoi s'y intéresser ? Pourquoi		Télécom ParisTech		<a href="http://www.telecom-evolution.fr/fr/formations-courtes/big-data-enjeux-">http://www.telecom-evolution.fr/fr/formations-courtes/big-data-enjeux-</a>
Big data	Big Data : panorama des infrastructures et architectures distribuées	continue		Décideurs et ingénieurs désireux de comprendre	1 jour	Introduction □ Big data - définition et contexte		Télécom ParisTech		<a href="http://www.telecom-evolution.fr/fr/formations-courtes/big-data-panorama-des-">http://www.telecom-evolution.fr/fr/formations-courtes/big-data-panorama-des-</a>
Big data	Data science : introduction au machine learning	continue			1 jour	Introduction au machine learning et à la théorie de l'apprentissage		Télécom ParisTech		<a href="http://www.telecom-evolution.fr/fr/formations-courtes/data-science-">http://www.telecom-evolution.fr/fr/formations-courtes/data-science-</a>
Big data	Ending the R vs Python war							Dataiku (blog)		<a href="http://www.dataiku.com/blog/2015/09/02/Free_Training_06.html">http://www.dataiku.com/blog/2015/09/02/Free_Training_06.html</a>
Big data	How to Develop Indicators Using Healthcare Data?							Dataiku (blog)		<a href="http://www.dataiku.com/blog/2015/04/09/dss-video-series-how-to-develop-indicators-using-">http://www.dataiku.com/blog/2015/04/09/dss-video-series-how-to-develop-indicators-using-</a>
Big data	How to merge different datasets in Data Science							Dataiku (blog)		<a href="http://www.dataiku.com/blog/2015/07/08/Free_Training_02_video.html">http://www.dataiku.com/blog/2015/07/08/Free_Training_02_video.html</a>
Big data	How to build your first predictive model in DSS							Dataiku (blog)		<a href="http://www.dataiku.com/blog/2015/07/24/How_To_Build_Predictive_Model.html">http://www.dataiku.com/blog/2015/07/24/How_To_Build_Predictive_Model.html</a>
Big data	How to make the Best Data Visualizations in Data							Dataiku (blog)		<a href="http://www.dataiku.com/blog/2015/08/06/DataViz_FreeTraining.html">http://www.dataiku.com/blog/2015/08/06/DataViz_FreeTraining.html</a>
Big data	Data Science Starter Program	continue				*Journée 1. Introduction à la science des données. Écosystème du big data,		Polytechnique		<a href="https://www.polytechnique.edu/bigdata/dssp/">https://www.polytechnique.edu/bigdata/dssp/</a>
Big data	SAS Analytics U	éditeur de logiciel qui a créé un		tout public		navigate the SAS windowing environment navigate the SAS		SAS (éditeurs de logiciel)		<a href="https://support.sas.com/edu/sche-dules.html?id=1320&amp;ctry=US">https://support.sas.com/edu/sche-dules.html?id=1320&amp;ctry=US</a> <a href="https://support.sas.com/edu/sche">https://support.sas.com/edu/sche</a>



- Les formations portant à la fois sur le big data et le cloud computing

Cloud/big data	Relever les défis du Big Data	continue		DSI, RSI, Urbaniste, Responsable de Production,	Une journée (7 heures éligibles au	Comprendre le Big Data (Volume, Vitesse, Variété,		ISEP Formation Continue			<a href="http://www.cloud-computing-formation.fr/formations-courtes-cloud-computing/big-data/">http://www.cloud-computing-formation.fr/formations-courtes-cloud-computing/big-data/</a>
Cloud/big data	Chef de projet opérationnel : Options Ingénierie des logiciels- Infrastructure réseaux,	initiale/continue	Niveaux I et II	tout public à partir du niveau bac+3 en développement ou	Jusqu'à 24 mois, selon votre profil	Analyser un besoin, un cahier des charges Concevoir les solutions	Chef de projet applicatif Responsable projet Assistant chef de projet	IPI-INSTITUT DE POLY-INFORMATIQUE - GROUPE IGS	oui		<a href="http://www.letudiant.fr/fiches/etudes/fiche/ipi-institut-de-poly-informatique-groupe-igs.html">http://www.letudiant.fr/fiches/etudes/fiche/ipi-institut-de-poly-informatique-groupe-igs.html</a>
Cloud/big data	Expert Systèmes Réseaux et Bases de Données - Spécialisations réseaux,	initiale/continue	Niveaux I et II	Bac+3/4 ou avec expérience professionnelle	selon votre profil	Il est en mesure d'assurer ainsi une continuité de service optimale.		IPI-INSTITUT DE POLY-INFORMATIQUE - GROUPE IGS	oui		<a href="http://www.inspi.fr/formation-continue/programmes-de-formation/expert-en-systemes-formation/programmes-de-formation-continue/programmes-de-formation/chef-de-projet-formation/">http://www.inspi.fr/formation-continue/programmes-de-formation/expert-en-systemes-formation/programmes-de-formation-continue/programmes-de-formation/chef-de-projet-formation/</a>
Cloud/big data	CHEF DE PROJET INGÉNIERIE LOGICIELLE OU RÉSEAUX ET SÉCURITÉ	continue	Niveaux I et II	Bac+4, ou expérience professionnelle dans	selon votre profil	utiliser une méthode d'analyse et planifier les		IPI-INSTITUT DE POLY-INFORMATIQUE - GROUPE IGS			<a href="http://www.inspi.fr/formation-continue/programmes-de-formation/chef-de-projet-formation/">http://www.inspi.fr/formation-continue/programmes-de-formation/chef-de-projet-formation/</a>
Cloud/big data	Réseaux et télécommunications : présent et avenir	continue		Ingénieurs, ingénieurs d'étude, ingénieurs de	5 jours	Acquisition d'une connaissance globale sur les architectures et services		Télécom ParisTech			<a href="http://www.telecom-evolution.fr/fr/formations-courtes/reseaux-et-formation/">http://www.telecom-evolution.fr/fr/formations-courtes/reseaux-et-formation/</a>
Cloud/big data	Ingénieur diplômé de l'ISEN Lille	initiale/continue	Niveaux I et II	Admission en cycle préparatoire 1ère année :	3-5 ans			Institut supérieur de l'électronique et du numérique de Lille, Université catholique de	oui	oui	<a href="http://www.letudiant.fr/etudes/annuaire-enseignement-superieur/formation/formation-informatique?code=SPSIN5G">http://www.letudiant.fr/etudes/annuaire-enseignement-superieur/formation/formation-informatique?code=SPSIN5G</a>
Cloud/big data	Master Informatique, Spécialité Systèmes d'information répartis (SIR)	initiale/continue	Niveaux I et II	bac +4	1 an	Durant la formation, l'étudiant consolidera d'abord ses connaissances	Architecte en Systèmes d'Information,	UFR des sciences, Aix-Marseille université			<a href="http://informatique-sciences.univ-amu.fr/master-informatique?code=SPSIN5G">http://informatique-sciences.univ-amu.fr/master-informatique?code=SPSIN5G</a>
Cloud/big data	Licence pro. intégration des systèmes voix et données	initiale/continue	Niveaux I et II	Bac + 2 informatique ; Dossier, entretien, tests	1 an	Concevoir l'architecture d'un réseau global d'intégration téléphonie-informatique	Responsable de la sécurité des systèmes d'information (RSSI)	CFA EVE Evry		oui	<a href="http://www.cfa-eve.fr/mod_formation/public/detail/id/50">http://www.cfa-eve.fr/mod_formation/public/detail/id/50</a>
Cloud/big data	Master rec. informatique et télécommunication	initiale/continue	Niveaux I et II	M1 ou équivalent ; Dossier	1 an	appréhender de façon pertinente les multiples facettes des thématiques		ISAE-SUPAERO Toulouse			<a href="http://www.irit.fr/M2RIT/">http://www.irit.fr/M2RIT/</a>
Cloud/big data	Cloud Computing and Services (CCS)	initiale	Niveaux I et II			distributed systems, advanced network technologies, service		Faculté des sciences, université de Nice Sophia-Antipolis			<a href="http://www.masterschool.eictclabs.eu/programmes/ccs/">http://www.masterschool.eictclabs.eu/programmes/ccs/</a>
Cloud/big data	Diplôme d'ingénieur ESIGETEL parcours : Data-center for Cloud Computing	initiale	Niveaux I et II	Bac ou bac +2	5 ou 3 ans			ESIGETEL - ECOLE SUPÉRIEURE D'INGÉNIEURS			<a href="http://www.letudiant.fr/fiches/etudes/fiche/esigetel.html">http://www.letudiant.fr/fiches/etudes/fiche/esigetel.html</a>
Cloud/big data	LOT2_CONCEPTION INTERFACE GRAPH_WEB_SAAS DEV JAVA_EE	continue			2-5 j			PLB CONSULTANT			<a href="http://www.plb.fr/formations.php">http://www.plb.fr/formations.php</a>
Cloud/big data	LOT2_CONCEPTION INTERFACE GRAPH_JAVA WEB_SAAS	continue			63-70 h			DEVELOPPEMENT MULTIMEDIA ET RESEAUX			<a href="http://www.dmr-formation.fr/formations/">http://www.dmr-formation.fr/formations/</a>
Cloud/big data	Administration de Windows server 2012 et Virtualisation	continue						AFIB DECISION			
Cloud/big data	APPROCHE ET REUSSITE PROJET DE VIRTUALISATION	continue			1-17 j			IB - GROUPE CEGOS			
Cloud/big data	Mettre en œuvre et gérer la virtualisation des postes de travail Microsoft - MS10758	continue						SQLI			
Cloud/big data	VIRTUALISATION SERVEUR AVEC WINDOWS SERVER HYPER-V ET SYSTEM CENTER-HYPERV3	continue						GLOBAL KNOWLEDGE NETWORK France			<a href="http://www.globalknowledge.fr/formation-informatique-management/">http://www.globalknowledge.fr/formation-informatique-management/</a>
Cloud/big data	Formation Réseaux Informatiques Data Center Virtualisation	continue						ALLYANS PROFESSIONAL TRAINING & CONSULTING P.T.C SARL			<a href="http://www.allyans.com/formations/">http://www.allyans.com/formations/</a>
Cloud/big data	Formation virtualisation VMware et Microsoft HyperV	continue						Num'Evolution			<a href="http://www.numevolution.fr/nos-domaines-dintervention/">http://www.numevolution.fr/nos-domaines-dintervention/</a>
Cloud/big data	Virtualisation des applications et du poste de travail	continue						M2I FORMATION / M2I TECH			<a href="http://www.m2iformation.fr/formations-informatique/">http://www.m2iformation.fr/formations-informatique/</a>

Cloud/big data	TESTS LOGICIELS METHODES AGILES VIRTUALISATION BUSINESS	continue						RS CONSULTING		<a href="http://rsconsulting.pro/index.php?id_cms=9&amp;controller=cms">http://rsconsulting.pro/index.php?id_cms=9&amp;controller=cms</a>
Cloud/big data	VIRTUALISATION SERVEUR AVEC WINDOWS SERVER HYPER-V ET SYSTEM CENTER-HYPERV3-TEST	continue						EDUGROUPE		<a href="http://www.edugroupe.com/page-recherche.asp?keyword=cloud&amp;id_site=1">http://www.edugroupe.com/page-recherche.asp?keyword=cloud&amp;id_site=1</a>
Cloud/big data	VIRTUALISATION LES FONDAMENTAUX	continue						LEARNING TREE INTERNATIONAL		
Cloud/big data	VIRTUALISATION DES SERVEURS	continue						VAELIA		<a href="http://groupe.nvl.fr/site-vaelia/formations">http://groupe.nvl.fr/site-vaelia/formations</a>
Cloud/big data	VIRTUALISATION DES SERVEURS AVEC WINDOWS SERVER HYPER V ET SYSTEME	continue						ENI SERVICE		<a href="http://www.eni-service.fr/formations/virtualisation-et-infrastructure-du-cloud-">http://www.eni-service.fr/formations/virtualisation-et-infrastructure-du-cloud-</a>
Cloud/big data	WINDOWS 2012, virtualisation avec Hyper-V	continue						ORSYS		<a href="http://www.orsys.fr/?mode=offre">http://www.orsys.fr/?mode=offre</a>
Cloud/big data	Cloud : la synthèse	continue						CAPGEMINI FRANCE		<a href="https://www.fr.capgemini.com/carrieres/developpement-carriere/universite-du-groupe">https://www.fr.capgemini.com/carrieres/developpement-carriere/universite-du-groupe</a>
Cloud/big data	OpenStack: Mettre en place un Cloud OpenSource privé	continue						EGILIA LEARNING		
Cloud/big data	System Center 2012 - Private Cloud Monitoring and Operations	continue			1-5 j			MANPOWER NOUVELLES COMPETENCES / Futurskill training		<a href="http://www.futurskill-training.fr/resultat-formation.html?md=thm&amp;tid=3&amp;li">http://www.futurskill-training.fr/resultat-formation.html?md=thm&amp;tid=3&amp;li</a>
Cloud/big data	Cloud Computing	continue						OBJET DIRECT VISEO		<a href="http://www.viseo.com/fr/offre/business-intelligence">http://www.viseo.com/fr/offre/business-intelligence</a>
Cloud/big data	APACHE SPARK POUR DEVELOPPEUR CLOUDERA	continue						XEBIA IT ARCHITECTS / XEBIA TRAINING		
Cloud/big data	EXECUTIVE CLOUD COMPUTING	initiale/continue						CENTRALE FORMATION		
Cloud/big data	Cloud Campus	continue						ORACLE FRANCE		<a href="http://education.oracle.com/pls/web_prod-plq-dad/db_pages.getpage?page_id=3">http://education.oracle.com/pls/web_prod-plq-dad/db_pages.getpage?page_id=3</a>
Cloud/big data	CLOUDERA POUR DATA ANALYST	continue						SOGETI CORPORATE SERVICES		<a href="http://www.fr.sogeti.com/expertise/institut-de-formation/">http://www.fr.sogeti.com/expertise/institut-de-formation/</a>
Cloud/big data	Data Science Methodology	continue						Big Data University		<a href="http://bigdatauniversity.com/wpcourses/">http://bigdatauniversity.com/wpcourses/</a>
Cloud/big data	Traitement des données non structurées	continue						DICTANOVA		
Cloud/big data	SQL SERVER 2012, BUSINESS INTELLIGENCE	continue						ECOLE SUP INFORM COMMERCE		<a href="http://esic-online.com/new/accueil/informatique/">http://esic-online.com/new/accueil/informatique/</a>
Cloud/big data	FORMATION BIG DATA HADOOP	continue						SOCIETE HURENCE		
Cloud/big data	BUSINESS INTELLIGENCE BI41	continue						SODIFRANCE ISIS		<a href="http://www.sodifrance.fr/formati on">http://www.sodifrance.fr/formati on</a>
Cloud/big data	Administration Système de l'environnement HADOOP	continue						SOFTEAM		<a href="http://formation.softeam.fr/sections/architecture-cloud-big-data">http://formation.softeam.fr/sections/architecture-cloud-big-data</a>
Cloud/big data	IBM COGNOS BI ADMINISTRATION	continue						UMANIS INSTITUT		<a href="http://www.umanis.com/?q=catalogue">http://www.umanis.com/?q=catalogue</a>
Cloud/big data	Big Data, Cloud Computing, & CDN Emerging Technologies	MOOC						Coursera/Yonsei University		<a href="https://fr.coursera.org/learn/big-data-cloud-computing-cdn/">https://fr.coursera.org/learn/big-data-cloud-computing-cdn/</a>

