Université de technologie de Troyes

Diplôme d'Université
Big Data Engineer

Rédacteur	AD Lab / B.Birregah / L.Amice
Date	07/06/2018
Version du document	1.1
Date de recueil avis CE	7/06/2018
Date de validation CA	14/06/2018

1. Fiche d'identité de la formation

Adresse de la spécialité :
Université de technologie de Troyes
12, Rue Marie Curie
CS42060
10004 TROYES CEDEX
Tél : 03.25.71.76.00

Responsable de la spécialité : Babiga Birregah
Tél : 03.25.71.58.69
Mail : babiga.birregah@utt.fr

Intitulé	Diplôme d'Université Big Data Engineer
Domaine	Big Data
Niveau du diplôme	Spécialisation
Niveau requis	Bac+3 Informatique Bac+2 Informatique avec expérience professionnelle
Durée de la formation	4 mois
Année d'ouverture	2018

2. Introduction

Présentation du projet

Cette formation a pour objet de permettre à des personnes souhaitant développer leurs compétences dans le Big Data de maitriser les aspects suivants :

- La conception et le pilotage d'architectures et de projets Big Data
- Le développement et la mise en œuvre d'un projet Big Data sur « Hadoop »
- Les applications Big Data en batch et en temps réel utilisant le Machine Learning

Dans cette perspective, l'objectif est de proposer une formation complète pour devenir Big Data Engineer dans le cadre d'un Diplôme d'Université de 180 heures en s'appuyant sur l'expertise et la notoriété de l'Université de technologie de Troyes.
Le choix opéré par Africa Data Lab de l'Université de technologie de Troyes est un choix stratégique pour accélérer son développement sur le Maroc. Un partenaire académique reconnu pour Africa Data Lab au travers du corps d'enseignants de l'UTT expérimentés dans le monde de la Data.

La collaboration UTT – Africa Data Lab permettra de co-construire un programme de formation répondant aux besoins des entreprises, tout en bénéficiant de l'expertise des équipes pédagogiques et d'ingénierie des programmes de formations de l'UTT.

Les thématiques de cette formation seront abordées dans le cadre de cours magistraux et de projets de mise en œuvre (cas d'application réel, environnement, cas d'usage métiers, …) dispensés par des enseignants de l'Université de technologie de Troyes (UTT) et d'Africa Data Lab (ADL), spécialisés dans les domaines du Big Data, ainsi que par des professionnels experts dans ce domaine.

L'ouverture est prévue pour fin Septembre 2018, avec une fin de formation prévue en fin Janvier 2019.

Dans le cadre de ce projet, la collaboration entre l'UTT et ADL porte également sur l'organisation de passerelles entre le parcours certifiant « Big Data Engineer » déployé par ADL et le Mastère Spécialisé® Expert Big Analytics et Métriques délivré par l'UTT, en favorisant ainsi l'accès des participants diplômés du parcours certifiant au dit Mastère Spécialisé®.

Le contexte

Aujourd'hui, l'enjeu du big data est économique, il s'agit d'un nouveau segment de développement et de création de valeur pour les entreprises et les administrations.

Les compétences sont un levier de compétitivité pour le Maroc dans le cadre de ses ambitieuses stratégies sectorielles. Il y a une véritable pénurie sur le marché marocain d'experts en Data : d'ici 2020, un écart de 50 % entre l'offre et la demande.

Afin de répondre à cet enjeu, la professionnalisation des acteurs, au travers d'actions de sensibilisation et de formation adaptées aux activités et aux usages métiers s'impose.

Actuellement, seuls quelques acteurs proposent de véritables formations centrées sur les sujets Big Data. En revanche, cette thématique fait en général l'objet soit :

- de formations courtes de 1 ou 5 jours, délivrées par des organismes de formation,
- de formations diplômantes ou certifiantes longues (entre 12 et 18 mois) dans le cadre de formations ingénieurs ou master.

L'offre de formations courtes est principalement portée par des cabinets de conseil ayant une activité de formation, à l'image d'ITAB Academy qui propose une formation Big Data d'une durée de 40h ou OCTO Academy.

Concernant l'offre universitaire, il existe des formations académiques longues dans le domaine du Big Data, telles que :

- l'université UIR qui propose un master Big Data sur 2 ans (M1 & M2)
- l'université Mohamed V de Rabat propose un Mastère spécialisé® Big Data
- l'école ENSIAS propose un master de recherche Sciences de Données et Big Data

Il n'existe pas à l'heure actuelle sur le marché de formation certifiante, s'inscrivant à la fois dans le cadre de la formation continue – en terme de durée – positionnée dans une logique de certification.

Afin de répondre à cette double finalité, le projet de Diplôme d'Université « Big Data Engineer» est proposé dans un cadre partenarial entre l'Université de technologie de Troyes (UTT) et Africa Data Lab, visant à conjuguer leur expertise en matière du Big Data et leur volonté de développer ce parcours de formation dans un premier temps au Maroc et éventuellement en Afrique par la suite.

3. Objectifs de la formation

Objectifs pédagogiques

- Comprendre les enjeux et les règles de gouvernance des projets Big Data
- Identifier et dimensionner les outils pour la conception d'architectures Big Data
- Maîtriser les outils de l'écosystème des architectures distribuées
- Développer des applications de collecte, de traitement et de visualisation temps réel

Objectifs professionnels (domaines de compétences visés)

- Développement et administration des plateformes et projets Big Data
- Conception de solutions de stockage de données hétérogènes
- Modélisation, traitement et valorisation de données complexes

4. Programme de la formation (pilotage, contenus, modalités, durée et intervenants)

Pilotage pédagogique de la formation

Un Comité de pilotage est organisé afin de partager avec les différentes parties prenantes le suivi et l'avancement de la formation. Des indicateurs clés de qualité et de performance (KPI) seront mis en place et partagés lors de ce Comité de pilotage, réunissant des représentants de l'UTT et d'Africa Data Lab.

Programme

Intitulé	Contenu	Durée
DE01 **Introduction au Big Data**	1.1 : Présentation de l'écosystème Big Data 1.2 : Méthodologies et process de mise en œuvre 1.3 : Enjeux et gouvernance de la donnée en entreprise	10h
DE02 **Base de données avancées (NoSQL, graphes, etc.)**	2.1 : Comparatif des différents types de Bases de Données (structurées, semi-structurées, et non structurées) 2.2 : Architecture des différentes bases NoSQL 2.3 : Les bases NoSQL clé/valeur, documents, graphes	30h
DE03 **Ecosystème Hadoop**	3.1 : Présentation de l'écosystème Hadoop 3.2 : Architecture de projets Big Data 3.3 : Conception de projets Big Data hautement parallélisés 3.4 : Manipulation des données avec Hive, Flume	40h
DE04 **Intelligence artificielle et Machine Learning**	4.1 : Introduction au Data Mining 4.2 : Mise en forme des données : Feature Engineering 43 : Présentation des différents algorithmes de machine learning 4.4 : Industrialisation des modèles et passage à l'échelle	20h
DE05 **Big Data : applications en batch**	5.1 : Développement et optimisation d'applications avec Apache Spark 5.2 : Création d'applications distribuées avec Java en se basant sur le paradigme MapReduce. 5.3 : Automatisation et orchestration des traitements avec Oozie	30h
DE06 **Big Data: Applications en streaming**	6.1 : Déploiement et gestion d'un cluster pour traiter les données en streaming 6.2 : Gestion des applications en temps réel avec Kafka 6.3 : Développement d'applications avec Apache Spark Streaming	30h
DE07 **Conception et gestion de projets Big Data**	7.1 : Comprendre les spécificités technologiques, organisationnelles et juridiques d'un projet Big Data 7.2 : Mise en place et pilotage d'une équipe projet Big Data 7.3 : Problématiques de protection et de sécurité des projets Big Data	20h

Tout au long de la formation, ADL organisera des séminaires pour traiter les nouvelles tendances du Big Data.

Organisation pédagogique

La formation « Big Data Engineer » est construite sur une durée de 180h sur 4 mois en temps partiel.

Le rythme de la formation est le suivant :

- Mardi : 19h – 21h
- Vendredi : 19h – 21h
- Samedi : 8h30 – 12h30 & 13h30 – 17h30

Chaque étudiant se verra attribuer un « mentor » qui l'accompagnera tout au long de sa formation.

Méthodes et moyens pédagogiques

La formation est composée de sept (7) modules qui alternent enseignements théoriques et travaux pratiques. Chaque module est conçu pour fournir les concepts et les outils utiles au au projet professionnel de l'étudiant.

Le DU Big Data Engineer est une formation orientée « projet ». Son principal objectif est de répondre aux besoins des entreprises dans le domaine du « Big Data ». Tout au long de la formation, un projet « atelier » est mis en place permettant de reprendre des cas d'usages métiers. Ce projet « atelier » correspond à une mise en situation des apprenants au contexte et environnement « réel » en entreprise.

Notre formation offrira également l'opportunité de préparer les certifications délivrées par les éditeurs (Cloudera, Hortonworks, Spark, AWS, etc.)

Par exemple, le module « Ecosystème Hadoop » permet d'acquérir toutes les compétences nécessaires pour réussir la certification « Hadoop de Hortonworks ».

5. Modalités d'évaluation et de certification

Contrôle de connaissances et modalités d'évaluation

Chaque module donnera lieu à une évaluation des compétences acquises (QCM ou Travaux Pratiques) notée sur 20.

Dans le cadre du projet « atelier » qui se déroulera tout au long de la formation, les participants devront présenter, à la fin de la formation, leur réalisation sous forme d'une soutenance qui sera notée sur 20.

Obtention de la certification

L'obtention de la certification DU « Big Data Engineer » est conditionnée par le suivi et la validation de l'ensemble des modules du programme.

Le participant doit obtenir une note égale ou supérieure à 10 pour chacun des modules et une égale ou supérieure à 10 pour la réalisation et la soutenance du projet « atelier ».

L'évaluation de chaque module :

- Coefficient 1 : Note du module (QCM, travaux pratiques, …)
- Coefficient 2 : Note du projet « atelier » relatif au contenu du module

Jury de diplôme

Le jury de diplôme du DU « Big Data Engineer » est composé de 3 personnes :

- Le responsable pédagogique du programme au titre de l'UTT,
- Un enseignant de l'UTT participant aux enseignements dans le cadre du DU Big Data Engineer
- Le responsable pédagogique du programme au titre d'ADL.

Le jury de diplôme décerne le DU « Big Data Engineer» aux stagiaires ayant répondu aux obligations administratives et pédagogiques exigées.

6. Public cible, prérequis et conditions d'admission

Public cible

La formation « Big Data Engineer », s'adresse à tout public salarié, demandeur d'emploi ou en reconversion professionnelle et souhaitant acquérir des compétences Big Data :

- Chefs de projet ou responsable métier
- Responsable système d'informations
- Développeurs informatiques
- Consultants en business intelligence
- Ingénieurs d'étude, de recherche et développement
- Architecte système et logiciel
- Consultants techniques
- Consultants business
- Statisticiens et Mathématiciens
- Data scientist

Prérequis

Bon niveau de connaissances en Algorithmique et en structures de données

- Maitrise d'un ou plusieurs langages de programmation : Java, Scala, C, R, Python, …
- Maitrise des notions clés de bases de données

Conditions d'admission

L'admission se fait sur dossier et entretien individuel.

Les résultats seront envoyés par courrier et e-mail à l'issue de chaque session d'entretien.
Afin de valider les prérequis, un test sous forme de QCM est prévu lors de l'entretien individuel.

Une session de remise à niveau sera, si nécessaire, programmée avant le début des cours.

7. Positionnement de la formation dans l'organisation et modalités de déploiement

Place de la formation dans l'offre des deux établissements

Africa Data Lab est un organisme de formations dédiées aux métiers de la Data qui propose des formations professionnelles et courtes sous plusieurs formats :

- Data Days : Stage de quelques jours pour maitriser les nouvelles tendances et technologies data
- Sur-mesure : Formations sur-mesure afin de répondre aux besoins spécifiques de nos clients
- Certificats : Formation sur 4 mois pour devenir un data expert notamment ce programme « Big Data Engineer » en partenariat avec l'UTT.

Cette formation est au cœur de la stratégie de développement d'Africa Data Lab et fait partie des produits phares.

Au regard de l'Université de technologie de Troyes, le Diplôme d'Université Big Data Engineer est en adéquation avec le cœur de métier des enseignements dispensés au sein de l'établissement en informatique et réseaux.

D'autre part, il viendra renforcer la visibilité de l'établissement dans le domaine du numérique. Ainsi, le DU Big Data Engineer vient densifier l'offre de l'UTT à l'international de développement des compétences déjà mises en œuvre dans le cadre des formations suivantes : Master SSI, DU Délégué à la Protection de Données (DPO) et Mastères Spécialisés® Expert Big Analytics et Métriques (EBAM) et Expert Forensic et Cybersécurité (EFC).

Cette formation s'inscrit par ailleurs dans la stratégie globale de développement de l'UTT en Afrique. C'est donc un partenariat gagnant-gagnant qui offre l'opportunité aux deux parties de se développer au Maroc et plus généralement en Afrique francophone.

Lieu de réalisation de la formation

La formation a lieu dans les locaux d'Africa Data Lab au Technopark de Casablanca.

8. Perspectives

AD Lab a pour ambition de lancer cette première expérience de ce programme au Maroc (Casablanca) et de dupliquer éventuellement cette offre de formation dans le reste du continent africain notamment en Afrique francophone, en partenariat avec l'UTT.

Modèle économique

Partenaires clés

- Université de Technologie de Troyes (UTT) (corps pédagogique, expertise d'ingénierie de formation)
- Formateurs experts Data
- Technopark Casablanca (Ecosystème startup)

Activités clés

- Un format innovant pour répondre aux besoins des entreprises
- Le Big data, un sujet stratégique et un enjeu économique pour toutes les entreprises et administrations

Ressources clés

- Locaux (salles de formation, matériels, …)
- Intervenants formateurs
- Moyens de communication

Offre (proposition de valeur)

- Formation courte et professionnelle encadrée par des experts techniques et métiers
- Data Camp sur 4 mois (180 h)
- Programme complet par module
- Projet d'application tout au long du parcours
- Temps partiel (cours du soir et le samedi)
- Formation diplômante

Relation client

- Mentor et coaching individuel
- Suivi pendant et après formation
- Garant du premier emploi ou reconversion

Canaux de distribution

- Communication commune UTT & ADL
- Marketing digital (réseaux sociaux)
- Médias locaux
- Evénements (meet up, conférences, …)

Segments de clientèle

- Salariés souhaitant se spécialiser / approfondir leurs compétences dans le domaine du Big Data
- Demandeurs d'emploi (reconversion vers les métiers d'avenir)
- Etudiants (Bac+3/4)

Structure des coûts

- Locaux (salles de formation, matériels, …) : Technopark Casablanca
- Interventions des experts Data gérés par ADL
- Interventions des enseignants UTT
- Redevance par inscrit présenté à la diplomation au titre du pilotage pédagogique et de la certification par l'UTT

Sources de revenus

- Coût de la formation : 30 000 - 37 000 HT soit 36 000 - 44 400 TTC
  - 30% à l'inscription
  - 70%/4 par mois (sur 4 mois)
- Mode de paiement : Chèque / Carte Bancaire (virement, …)