QUALIFICATION FILE: Advanced Diploma In BigData Analytics (ADBA-Level 7)

NSDA Reference
To be added by NSDA

Revised Application Documentation: Version 1 / 12-Jan-2017

CONTACT DETAILS OF THE AWARDING BODY FOR THE QUALIFICATION

Name and address of submitting body:

NATIONAL INSTITUTE OF ELECTRONICS AND INFORMATION

TECHNOLOGY (NIELIT), CALICUT

NIT CAMPUS POST, KOZHIKODE, KERALA.

PIN – 673601.

Name and contact details of individual dealing with the submission:

Name: Shri. PrasoonKumar KG

Designation: Principal Technical Officer

Mobile: +91 9447305951

Email: prasoon@nielit.gov.in / prasoon@calicut.nielit.in

List of documents submitted in support of the Qualifications File

1. Annexure I – Course Curriculum

2. Annexure II – Candidates Trained

3. Annexure III – Placement Details

4. Annexure IV - Industry validation

5. Annexure V - Evidence of need - Requirement and salary report of 2016 for data Engineers.

6. Annexure VI- Occupational Map as identified by IT SSC mapped to Jobroles.

QUALIFICATION FILE: Advanced Diploma In BigData Analytics (ADBA-Level 7)

SUMMARY

Qualification Title: Advanced Diploma In BigData Analytics Qualification Code: NIELIT/IT/L7/031
Nature and purpose of the Qualification: This Qualification is aligned to Level 7 Engineers to meet the requirements of administration, development, analysis of BigData applications
Body/bodies which will award the qualification: National Institute of Electronics and Information Technology
Body which will accredit providers to offer courses leading to the qualification: National Institute of Electronics and Information Technology
Body/bodies which will be responsible for assessment: Examination Cell, National Institute of Electronics and Information Technology
Occupation(s) to which the qualification gives access: Data Engineer, Administrator, Data Analyst
Licensing requirements: N/A
Proposed level of the qualification in the NSQF: Level 7
Anticipated volume of training/learning required to complete the qualification: 480 Hours
Entry requirements / recommendations: 3year diploma/Graduation in science/commerce/statistics/ engineering or equivalent with good computer programming knowledge.
Progression from the qualification: Professional:

Data Engineer – Senior Data Engineer – Principal Data Engineer

Data Analyst - Senior Data Analyst – Principal Data Analyst

Data Administrator – Senior Data Administrator

Planned arrangements for the Recognition of Prior learning (RPL):

Presently only candidates who undergo training shall be assessed.

It will be incorporated once RPL strategy is finalized.

International comparability where known:

N.A.

Date of planned review of the qualification:

After every 2 years.

Formal structure of the qualification:

Course Structure:

This course consists of seven modules and after the successful completion of all, the candidates will be awarded Advanced Diploma degree.

Module Code	Module Name	Mandatory /Optional	Theory Hours	Practical Hours	Total	Level
ADBA01	Linux tools and scripting	M	15	15	30	6
ADBA02	Java programming	M	15	15	30	7
ADBA11	Python programming	M	40	50	90	7
ADBA21	R programming	M	40	50	90	7
ADBA31	Hadoop	M	60	60	120	7
ADBA32	Hadoop Sub projects	M	15	15	30	7
ADBA41	Project work	M	5	85	90	7
	Total (480 Hours)		190	290	480	7

SECTION 1

ASSESSMENT

Name of Assessment body:

Examination Cell,
National Institute of Electronics and Information Technology
6-CGO Complex, Electronics Niketan
Lodhi Road, New Delhi. 110003.

Name of body checking or verifying Assessments:

Examination Cell,
National Institute of Electronics and Information Technology
6-CGO Complex, Electronics Niketan
Lodhi Road, New Delhi. 110003.

Will the assessment body be responsible for RPL assessment?

RPL Policy will be described as and when available

Describe the overall assessment strategy and specific arrangements which have been put in place to ensure that assessment is always valid, consistent and fair and show that these are in line with the requirements of the NSQF:

The Examination Cell develops assessment strategy unique to each module which bifurcate the theory and practical with higher emphasis on Practical Assessment.

Each Module Assessment consists of marks for internal evaluation, practical examination and viva voce. The internal marks are awarded based on the solution progress of lab exercises and mini projects/problem solving skill and also the percentage of attendance.

ASSESSMENT EVIDENCE

Job Role: Data Engineer/Data Analyst/Data Administrator

Title of Unit/Component: (Detailed Curriculum attached as Annexure-I)

Outcomes to be assessed	Assessment criteria for the outcome	Means of assessment			
		Internal	Practical	Viva voce	Total
Module:ADBA01 Linux tools and scripting: Basic administration of Linux machines using Linux commands and shell scripting	Installation and configuration of Linux Commands for managing files and processes, users etc. Learn Shell scripts	10	40	10	60
Module:ADBA02 Java Programming: basics of Java and Object Oriented programming in Java	Execute Java programs in Linux environment. Class (inner class) based java programs that have inheritance, polymorphism and Abstraction. Create interfaces and packages	10	40	10	60
Module:ADBA11 Python Programming: Developing applications using python	Development and execution of python programs in Linux Environment. Tkinter to develop GUI applications Develop MySQL based	25	125	20	170

	applications, Process n-Dimensional arrays using numpy Plot Graphs using matplotlib pandas for data analysis				
Module:ADBA21 R Programming: Can perform data analysis using R Programs	Development and execution of R programs in Linux environment using RStudio, using different data structures Reading data from different media like HTML, JSON, CSV, MySQL Developing statistical analysis applications that include different visualization techniques.	25	125	20	170
Module:ADBA31 Using Hadoop Administering and developing applications using Hadoop ecosystem	Install and configure Hadoop in pseudo distributed and full distributed mode. Development and execution of machine learning /mapreduce applications using java Integrating Python and Hadoop Integrating R in Hadoop ecosystem	25	175	20	220
Module:ADBA32 Using Hadoop sub projects	Hadoop sub projects like Hive, Pig , OOzie, Hbase, Sqoop and Flume to assist Hadoop	10	40	20	70

Module:ADBA41 Project Work / Thesis	Development and implementation of a data analysis project that uses Hadoop and java or python or R.	25	175	50	250
	Total	130	720	150	1000

Grade Sheet

Grade	% Marks	Grade	% Marks	Grade	% Marks
S	90% and above	A	75% to <90%	B	60% to <75%
C	50% to <60%	D	40% to <50%	F	<40%

SECTION 2

EVIDENCE OF LEVEL

Title/Name of qualification: Advanced Diploma in BigData Analytics		Level: 7	
NSQF Domain	Key requirements of the job role	How the job role relates to the NSQF level descriptors	NSQF Level
Process	Design and Develop BigData Applications using Hadoop, Java Python and R Programming	Data science is a non-linear, iterative process that involves asking questions, getting data, exploring them, communicating analyzing modeling and visualizing data. The data Engineers/analysts will also lead their work through implementation and then test and quantify its impact on their organization.	7
Professional knowledge	Applying statistical methods on data and doing software engineering to process and implement the same.	At least a basic understanding of statistics is vital as a data Engineer. It is important to have a strong software engineering background	7

		and the companies want to see that the candidate is a (data-driven) problem solver too.	
Professional skill	Machine learning techniques that can be implemented using R or Python libraries or java.	It is important to not just be familiar with the tools necessary to visualize data, but also the principles behind visually encoding data and communicating information.	**7**
Core skill	Applying appropriate machine-learning algorithm.	To be a data Engineer, the candidate need to know how and when to apply an appropriate machine-learning algorithm.	**7**
Responsibility	Conducting full lifecycle analysis to include requirements, activities and design. Developing analysis and reporting capabilities. Monitoring performance and quality control plans to identify improvements.	A solution built with great features and the right algorithm that demonstrates tangible value from a real-life business problem is just the beginning of the data Engineer's job. For that value to be realized, the candidate must manage his solution through its journey in the organization.	**7**

SECTION 3

EVIDENCE OF NEED

> ### What evidence is there that the qualification is needed?
>
> In 2012, Harvard Business Review named data scientist the "sexiest job of the 21st century". More recently, Glassdoor named it the "best job of the year" for 2016. In India, too, companies are scrambling to hire data scientists, sending salaries soaring. According to TeamLease, a staffing solutions company, data scientists with around 5 years' experience are earning over 75 lakh per annum as compared to 8-15 lakh for CAs and 5-8 lakh for engineers with the same experience level. So who exactly are data scientists? It's a loose term but these are the guys who know how to look at the data a company generates, and derive the all-important insights it needs to garner more business and enhance customer experience in this age of social media. With companies across industries striving to bring their research and analysis (R&A) departments up to speed, the demand for qualified data scientists is rising."India will face a demand-supply gap of 2,00,000 analytics professionals over the next three years. Even in the US, only 40 out of 100 positions for analytics professionals can be filled," said Rituparna Chakraborty, co-founder & senior VP of TeamLease Services.
>
> (This information is collected from http://timesofindia.indiatimes.com/)
>
> One pdf file attached showing the requirement and salary report of 2016 for data engineers.

> ### What is the estimated uptake of this qualification and what is the basis of this estimate?
> Estimated uptake is 40 students per batch, with 2 batches per year and on the basis of facilities and infrastructure in NIELIT Calicut.

> ### What steps were taken to ensure that the qualification(s) does/do not duplicate already existing or planned qualifications in the NSQF?
> This qualification has comprises both technical and analytic skills and can be linked to any qualification higher than this one, existing or to come.

> ### What arrangements are in place to monitor and review the qualification(s)? What data will be used and at what point will the qualification(s) be revised or updated?
> Based on feedback by participants, employers and based on market survey (about / requirement of latest technology in BigData industry) the qualification will be reviewed in every year.

SECTION 4

EVIDENCE OF RECOGNITION AND PROGRESSION

What steps have been taken in the design of this or other qualifications to ensure that there is a clear path to other qualifications in this sector?

This QF is a specialised training in data science and there is no higher level qualification found available in the NQR at present

SECTION 5

EVIDENCE OF EVIDENCE OF INTERNATIONAL COMPARABILITY

List any Comparisons which have been established:
 NIL.