QUALIFICATION FILE – Standalone NOS

Fundamentals of Data Curation using Python

☐ **Horizontal/Generic ☐ Vertical/Specialization**

☐ **Upskilling ☐ Dual/Flexi Qualification ☐ For ToT ☐ For ToA**

☐**General ☐ Multi-skill (MS) ☐ Cross Sectoral (CS) ☒ Future Skills ☐ OEM**

NCrF/NSQF Level: 3.5

Submitted By:

National Institute of Electronics and Information Technology (NIELIT)
NIELIT Bhawan,
Plot No. 3, PSP Pocket, Sector-8,
Dwarka, New Delhi-110077,
Phone: - 91-11-2530 8300
e-mail: - contact@nielit.gov.in

Table of Contents

Section 1: Basic Details

1.	**NOS-Qualification Name**	**Fundamentals of Data Curation using Python**	
2.	**Sector/s**	**IT-ITeS**	
3.	**Type of Qualification ☒ New ☐ Revised**	**NQR Code & version of the existing /previous qualification: NA**	**Qualification Name of the existing/previous version: NA**
4.	**National Qualification Register (NQR) Code & Version** *(Will be issued after NSQC approval.)*	NG-3.5-IT-03723-2025-V1-NIELIT	**5. NCrF/NSQF Level: 3.5**
6.	**Brief Description of the Standalone NOS**	The **"**Data Curation using Python" is designed to equip learners with the essential skills and knowledge required to manage, organize, and maintain data throughout its lifecycle. This course covers key aspects of data curation, including data collection, cleaning, annotation, integration, storage, quality control, and visualization. Students will gain hands-on experience and a solid understanding of the tools and techniques used in the data management field.	
7.	**Eligibility Criteria for Entry for a Student/Trainee/Learner/Employee**	a. **Entry Qualification & Relevant Experience:** Table below	

Eligibility table (row 7):

Academic/Skill Qualification (with Specialization - if applicable)	Required Experience (with Specialization - if applicable)
Grade 10 pass and pursuing continuous schooling*	No experience required
Grade 8 with 2 years of NTC in the field of IT Sector*	No experience required
Grade 8th pass*	3 Year relevant experience in IT
Previous NSQF Level 3 Required	1.5-year relevant experience in IT

Prerequisite: Fundamentals of Python

Should have a basic understanding of coding*.

8.	**Credits Assigned to this NOS-Qualification, Subject to Assessment** *(as per National Credit Framework (NCrF))*	4 credits	**9. Common Cost Norm Category (I/II/III)** *(wherever applicable)*: **Category- II**

10.	**Any Licensing Requirements for Undertaking Training on This Qualification** *(wherever applicable)*	NA			
11.	**Training Duration by Modes of Training Delivery** *(Specify Total Duration as per selected training delivery modes and as per requirement of the qualification)*	☒ **Offline Only** ☐ **Online Only** ☐ **Blended**			

Section 11 table:

Training Delivery Mode	Theory (Hours)	Practical (Hours)	Total (Hours)
Classroom (offline)	45	75	120

The mode of delivery shall be based on the regional demand and can be offered in any of the above modes mentioned.
(Refer Blended Learning Annexure-V for details)

12.	**Assessment Criteria**

Theory (Marks)	Practical (Marks)	Project (Marks)	Viva (Marks)	Total (Marks)	Passing %age
100	60	20	20	200	50%

The centralised online assessment is conducted by the Examination Wing, NIELIT Headquarters.

*Assessment strategy shall be as per NIELIT Norms prevailing at times.

13.	**Is the NOS Amenable to Persons with Disability**	☒ **Yes** ☐ **No:** a. Locomotor Disability: Leprosy Cured Person, Dwarfism, Muscular Dystrophy and Acid Attack Victims b. Visual Impairment: Low Vision
14.	**Progression Path After Attaining the Qualification, wherever applicable** *(Please show Professional and Academic progression)*	**Vertical:** Level 4: AI Development Associate **Professional:** Data Curation Assistant → Data Curation Specialist → Data Analyst → Data Engineer → Data Scientist
15.	**How participation of women will be encouraged?**	Participation by women can be ensured through Government Schemes. Occasionally, exclusive batches for women would be run for the proposed courses. Funding is available for women's participation under other schemes launched by the Government from time to time.

16.	Other Indian languages in which the Qualification & Model Curriculum are being submitted	Qualification file is available in English and Hindi languages.		
17.	Is similar NOS available on NQR-if yes, justification for this qualification	☐ Yes ☒ No URLs of similar Qualifications:		
18.	Name and Contact Details Submitting / Awarding Body SPOC *(In case of CS or MS, provide details of both Lead AB & Supporting ABs)*	Name: Vanlalhmangaiha Khiangte Email: vhmangaiha@nielit.gov.in Contact No.: 9774582997 Website: https://nielit.gov.in/index.php		
19.	Final Approval Date by NSQC: 18.02.2025	20. Validity Duration: 3 Years		21. Next Review Date: 18.02.2028

Section 2: Training Related

1.	Trainer's Qualification and experience in the relevant sector (in years) *(as per NCVET guidelines)*	B. Tech in Computer Science/ IT/ allied branches Or MCA Or BCA Or M Sc (IT/CS) Or NIELIT A- Level Qualified with 1 Years of experience in teaching
2.	Master Trainer's Qualification and experience in the relevant sector (in years) *(as per NCVET guidelines)*	B. Tech in Computer Science/ IT/ allied branches 2 Years of Experience in teaching Or M Sc (IT/CS) Or MCA Or BCA with 2 Years of Experience in teaching.
3.	Tools and Equipment Required for the Training	☒Yes ☐No *(If "Yes", details to be provided in Annexure)* Details available in Annexure II
4.	In Case of Revised NOS, details of Any Upskilling Required for Trainer	Nil

Section 3: Assessment Related

1.	**Assessor's Qualification and experience in relevant sector (in years)** *(as per NCVET guidelines)*	B. Tech in Computer Science/ IT/ allied branches 2 Years of Experience in teaching Or M Sc (IT/CS) Or MCA Or BCA with 2 Years of Experience in teaching.
2.	**Proctor's Qualification and experience in relevant sector (in years)** *(as per NCVET guidelines), (wherever applicable)*	The assessor carries out theory online assessments through the remote proctoring methodology. Theory examination would be conducted online, and the paper comprise of MCQ. Conduct of assessment is through trained proctors. Once the test begins, remote proctors have full access to the candidate's video feeds and computer screens. Proctors authenticate the candidate based on registration details, pre-test image captured and I- card in possession of the candidate. Proctors can chat with candidates or give warnings to candidates. Proctors can also take screenshots, terminate a specific user's test session, or re-authenticate candidates based on video feeds.
3.	**Lead Assessor's/Proctor's Qualification and experience in relevant sector (in years)** *(as per NCVET guidelines)*	External Examiners/Observers, including NIELIT scientific officers who are subject matter experts, are assigned to evaluate practical examinations, internal assessments, projects, presentations, assignments, and major projects (if applicable). Qualifications is generally B.Tech degree.
4.	**Assessment Mode** *(Specify the assessment mode)*	Online for Theory Online/ Offline/ Blended for other assessment components depending on the region where the assessment is conducted
5.	**Tools and Equipment Required for Assessment**	☒ Same as for training ☒ Yes ☐ No *(Details in Annexure II)*

Section 4: Evidence of the Need for the Standalone NOS

Provide Annexure/Supporting documents name.

1.	Government /Industry initiatives/ requirement (Yes/No): Yes
2.	Number of Industry validation provided: The NOS are offered through IndiaAi Labs for Empowering the youth by imparting training in Emerging AI Technologies under India AI Future Skills pillar of IndiaAI Mission.
3.	Estimated number of people to be trained: 500 persons per year shall be trained
4.	Evidence of Concurrence/Consultation with Line/State Departments (In case of regulated sectors): (Yes/No): NIELIT is recognized as AB and AA under Government Category. NIELIT is an HRD arm of MeitY, therefore, the Line Ministry Concurrence is not required.

Section 5: Annexure & Supporting Documents Check List

Specify Annexure Name / Supporting document file name

1.	**Annexure:** NCrF/NSQF level justification based on NCrF/NSQF descriptors *(Mandatory)*	Available at Annexure-I: Evidence of Level
2.	**Annexure:** List of tools and equipment relevant for NOS *(Mandatory, except in case of online course)*	Available at Annexure-II: Tools and Equipment
3.	**Annexure:** Training Details	Available at Annexure-IV: Training Details
4.	**Annexure:** Blended Learning *(Mandatory, in case selected Mode of delivery is Blended Learning)*	Available at Annexure-V: Blended Learning
5.	**Annexure/Supporting Document:** Standalone NOS- Performance Criteria Details Annexure/Document with PC-wise detailing as per NOS format (Mandatory- Public view)	Annexure-VI: Standalone NOS- Performance Criteria details
6.	**Annexure:** Performance and Assessment Criteria *(Mandatory)*	Available at Annexure-VII: Detailed Assessment Criteria
7.	**Annexure:** Assessment Strategy *(Mandatory)*	Available at Annexure-VIII: Detailed Assessment Strategy
8.	**Annexure:** Acronym and Glossary *(Optional)*	Available at Annexure-IX: Acronym and Glossary
9.	**Supporting Document:** Model Curriculum *(Mandatory – Public view)*	*Available at* Annexure-A: Model Curriculum

Annexure-I: Evidence of Level

NCrF/NSQF Level Descriptors	Key requirements of the job role/ outcome of the qualification	How the job role/ outcomes relate to the NCrF/NSQF level descriptor	NCrF/NSQF Level
Professional Theoretical Knowledge/Process	Individuals after acquiring the knowledge will have the ability to effectively manage and curate data throughout its lifecycle. This includes ensuring the data is collected, organized, validated, and maintained according to established standards and best practices. By applying their theoretical knowledge, the assistant will contribute to the creation of high-quality, accessible datasets that support informed decision-making. They will also be equipped to handle data governance issues, ensuring compliance with privacy, security, and ethical guidelines.	Work in a familiar, predictable, routine, situation of clear choice	3.5
Professional and Technical Skills/ Expertise/ Professional Knowledge	Individuals will have the skills to apply specialized technical tools and methodologies to effectively curate and manage data. Their technical skills enable them to perform data cleaning, validation, and organization tasks efficiently, while their professional knowledge supports adherence to data governance, privacy, and security protocols. This combination of skills ensures that data is maintained in compliance with industry standards, facilitating seamless data retrieval and usage. The assistant's proficiency in querying languages, data visualization, and processing tools allows them to support data analysis and contribute to informed decision-making within the organization	Factual knowledge of field of knowledge or study	3.5
Employment Readiness & Entrepreneurship Skills & Mind-set/Professional Skill	Individuals gains the practical skills necessary to secure employment, including the ability to work effectively in teams, communicate technical concepts clearly, and adapt to the evolving demands of the data industry. They are trained to take initiative, solve problems independently, and demonstrate attention to detail in data curation tasks. For entrepreneurial pursuits, they are equipped with the knowledge to manage and curate data in ways that add value to business operations, potentially starting their own data management services or consultancy. This combination of professional skills, mindset, and technical expertise ensures the assistant is prepared for both employment and entrepreneurial ventures in the data sector.	Recall and demonstrate practical skill, routine and repetitive in narrow range of application, using appropriate rule and tool, using quality concepts	3.5
Broad Learning Outcomes/Core Skill	The assistant will develop core skills include critical thinking, attention to detail, and problem-solving, which are crucial for maintaining data integrity and ensuring its accuracy. The assistant will develop strong communication skills, enabling them to effectively collaborate with team members and present data insights	Language to communicate written or oral, with required clarity, skill to basic arithmetic and algebraic principles, basic	3.5

	clearly to stakeholders. With these core skills, the assistant is prepared to tackle real-world challenges, ensuring effective data handling while supporting organizational goals.	understanding of social political and natural environment.	
Responsibility	Data Curation using Python can manage and oversee key aspects of the data curation process while adhering to organizational standards and best practices. The assistant is equipped to take full responsibility for the effective and ethical management of data throughout its lifecycle.	Responsibility for own work and learning.	3.5

Annexure-II: Tools and Equipment (lab set-up)

List of Tools and Equipment

Batch Size: 30

S. No.	Tool / Equipment Name	Specification	Quantity for specified Batch size
1	Desktop computer with accessories	12th Generation, 8GB DDR4, 512GB SSD, Graphics Installed with: Installation of Python, including development environments like Jupyter Notebook, Anaconda Distribution, Comprehensive libraries such as pandas, numpy, matplotlib, scikit-learn (for Python)	30
2	Student Chair	-	30
3	Student Table	-	30
5	LCD Projector	1	
6	Deskjet printer	1 No.	A4

Classroom Aids

The aids required to conduct sessions in the classroom are:

1. LCD Projector / Smart Board
2. Pin-Up Board
3. Whiteboard Markers

Annexure-III: Industry Validations Summary

* Course is to be offered through IndiaAi Labs for Empowering the youth by imparting training in Emerging AI Technologies under India AI Future Skills pillar of IndiaAI Mission.

Annexure-IV: Training Details

Training Projections:

Year	Estimated Training # of Total Candidates	Estimated training # of Women	Estimated training # of People with Disability
2025	500	100	10
2026	500	100	15
2027	500	100	20

Data to be provided year-wise for next 3 years.

Annexure-V: Blended Learning

Blended Learning Estimated Ratio & Recommended Tools:

S. No.	Select the Components of the NOS	List Recommended Tools – for all Selected Components	Offline: Online Ratio
1	☒Theory/ Lectures - Imparting theoretical and conceptual knowledge	Online interaction platforms like JitSi Meet, Bharat VC, Google Meet, MS Teams, etc.	60:40
2	☐Imparting Soft Skills, Life Skills and Employability Skills /Mentorship to Learners	NA	NA
3	☒Showing Practical Demonstrations to the learners	Online interaction platforms like JitSi Meet, Bharat VC, Google Meet, MS Teams, etc.	60:40
4	☒Imparting Practical Hands-on Skills/ Lab Work/ workshop/ shop floor training	PCs/Laptops	100:0
5	☒Tutorials/ Assignments/ Drill/ Practice	Online interaction platforms like JitSi Meet, Bharat VC, Google Meet, MS Teams, etc.	50:50

6	☒Proctored Monitoring/ Assessment/ Evaluation/ Examinations	NIELIT Online Examination	Online: 100% Theory; Offline 100% Practical
7	☐On the Job Training (OJT)/ Project Work Internship/ Candidate Training	NA	NA

Annexure-VI: Standalone NOS- Performance Criteria details

1. **Description:**

The " Data Curation using Python" is a standalone National Occupational Standard (NOS) designed for individuals aiming to enhance their skills in data curation. This certification program is tailored to professionals who already have a foundational understanding of data management and wish to specialize further in organizing, validating, and maintaining data for accessibility, usability, and analysis.

2. **Scope:**

The scope covers the following:

- Advanced data curation tools, techniques, and best practices.
- Focus on data collection, validation, organization, and storage with hands-on applications.
- Prepares participants to meet industry demands for managing high-quality, usable datasets.

3. **Elements and Performance Criteria**

To be competent, the user/individual on the job must be able to:

Foundational in Python Programming:

PC1. Set up the Python environment and execute basic scripts.
PC2. Explain AI concepts, including ML, deep learning, computer vision, and NLP, with industry examples.
PC3. Differentiate between supervised and unsupervised learning and understand data annotation's impact.

Basics of Artificial Intelligence & Data Science:

PC4. Explain AI evolution and key components (ML, deep learning, CV, NLP).
PC5. Handle large and complex datasets effectively.
PC6. Identify ethical concerns in AI, such as bias and privacy issues.

Introduction to Data Curation:

PC7. Understand the need and scope of data curation.
PC8. Differentiate between structured, unstructured, and semi-structured data.

Data Collection & Acquisition Methods:

PC9. Understand data collection methods and acquisition techniques.
PC10. Apply data cleaning, transformation, and enrichment techniques.

Data Integration, Storage, and Visualization:

PC11. Apply data integration techniques and understand storage architectures.
PC12. Develop data visualization skills and gain cloud storage knowledge.

Data Quality and Governance:

PC13. Understand key data quality metrics and compliance.
PC14. Learn data governance frameworks.

Advance Data Management Techniques:

PC15. Apply data governance frameworks.
PC16. Use AI-assisted data curation and big data tools.

Application of Data Curation:

PC17. Implement data curation techniques on real-world datasets.
PC18. Collaborate on data curation projects.

4. **Knowledge and Understanding (KU):**

The individual on the job needs to know and understand:

KU1: In-depth knowledge of Python programming concepts and use of its libraries.
KU2: Strong knowledge of the core principles of AI and data science.
KU3: Knowledge of the entire data curation lifecycle which includes data collection, cleaning and integration.
KU4: Techniques for data storage using different tools.
KU5: Tools and methods for data presentation.
KU6: Knowledge about data governance framework.

5. **Generic Skills (GS):**

The user/individual on the job needs to know how to:

GS1: Follow instructions, guidelines, and procedures for data handling and management.
GS2: Listen effectively and communicate information accurately, particularly regarding data management processes.
GS3: Apply standard data management protocols and metadata formatting to ensure data quality and accessibility.
GS4: Work collaboratively with teams to manage large datasets and support data-driven decision-making.

<div align="center">

Annexure-VII: Assessment Criteria

</div>

Detailed PC-wise assessment criteria and assessment marks for the NOS are as follows:

NOS/Module Name	Assessment Criteria for Performance Criteria	Theory Marks	Practical Marks	Project Marks	Viva Marks
Fundamentals of Data Curation using Python **NOS Code:** **NIE/ITS/N1016**	**Foundational in Python Programming:** PC1. Set up the Python environment and execute basic scripts. PC2. Explain AI concepts, including ML, deep learning, computer vision, and NLP, with industry examples. PC3. Differentiate between supervised and unsupervised learning and understand data annotation's impact.	11	6	2	2
	Basics of Artificial Intelligence & Data Science: PC4. Explain AI evolution and key components (ML, deep learning, CV, NLP).	11	6	2	2

PC5. Handle large and complex datasets effectively. PC6. Identify ethical concerns in AI, such as bias and privacy issues.				
Introduction to Data Curation: PC7. Understand the need and scope of data curation. PC8. Differentiate between structured, unstructured, and semi-structured data.	15	6	2	2
Data Collection & Acquisition Methods: PC9. Understand data collection methods and acquisition techniques. PC10. Apply data cleaning, transformation, and enrichment techniques.	20	7	2	2
Data Integration, Storage, and Visualization: PC11. Apply data integration techniques and understand storage architectures. PC12. Develop data visualization skills and gain cloud storage knowledge.	15	7	3	3
Data Quality and Governance: PC13. Understand key data quality metrics and compliance. PC14. Learn data governance frameworks.	15	7	3	3
Advance Data Management Techniques: PC15. Apply data governance frameworks. PC16. Use AI-assisted data curation and big data tools.	13	7	3	3
Application of Data Curation: PC17. Implement data curation techniques on real-world datasets. PC18. Collaborate on data curation projects.	-	14	3	3
Total Marks	**100**	**60**	**20**	**20**

Annexure-VIII: Assessment Strategy

This section includes the processes involved in identifying, gathering, and interpreting information to evaluate the Candidate on the required competencies of the program.

Assessment of the qualification evaluates candidates to ascertain that they can integrate knowledge, skills and values for carrying out relevant tasks as

per the defined learning outcomes and assessment criteria.

The underlying principle of assessment is fairness and transparency. The evidence of the outcomes and assessment criteria. competence acquired by the candidate can be obtained by conducting Theory (Online) examination.

About Examination Pattern:

1. The question papers for the theory exams are set by the Examination wing (assessor) of NIELIT HQS.

2. The assessor assigns roll number.

3. The assessor carries out theory online assessments. Theory examination would be conducted online and the paper comprise of MCQ

4. Pass percentage would be 50% marks.

5. The examination will be conducted in English language only.

Quality assurance activities: A pool of questions is created by a subject matter expert and moderated by other SME. Test rules are set beforehand. Random set of questions which are according to syllabus appears which may differ from candidate to candidate. Confidentiality and impartiality are maintained during all the examination and evaluation processes.

Annexure-IX: Acronym and Glossary

Acronym

Acronym	Description
AA	Assessment Agency
AB	Awarding Body
NCrF	National Credit Framework
NOS	National Occupational Standard(s)
NQR	National Qualification Register
NSQF	National Skills Qualifications Framework

Glossary

Term	Description
National Occupational Standards (NOS)	NOS define the measurable performance outcomes required from an individual engaged in a particular task. They list down what an individual performing that task should know and also do.
Qualification	A formal outcome of an assessment and validation process which is obtained when a competent body determines that an individual has achieved learning outcomes to given standards
Qualification File	A Qualification File is a template designed to capture necessary information of a Qualification from the perspective of NSQF compliance. The Qualification File will be normally submitted by the awarding body for the qualification.
Sector	A grouping of professional activities on the basis of their main economic function, product, service or technology.