QUALIFICATION FILE – Standalone NOS

Fundamentals of Data Annotation 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: - 91-11-25308300
e-mail: - contact@nielit.gov.in

Table of Contents

Section 1: Basic Details

<table>
<tr><td>1.</td><td>NOS-Qualification Name</td><td colspan="2">Fundamentals of Data Annotation using Python</td></tr>
<tr><td>2.</td><td>Sector/s</td><td colspan="2">IT-ITeS</td></tr>
<tr><td>3.</td><td>Type of Qualification ☒ New ☐ Revised</td><td>NQR Code & version of the existing /previous qualification: NA</td><td>Qualification Name of the existing/previous version: NA</td></tr>
<tr><td>4.</td><td>National Qualification Register (NQR) Code & Version (Will be issued after NSQC approval.)</td><td>NG-3.5-IT-03722-2025-V1-NIELIT</td><td>5. NCrF/NSQF Level: 3.5</td></tr>
<tr><td>5.</td><td>Brief Description of the Standalone NOS</td><td colspan="2">This course provides a comprehensive foundation in Python programming, AI principles, and Data Annotation enabling learners to apply these skills in real-world scenarios. Students will master text, audio, image, and video annotation, utilizing suitable methods and tools to create high-quality datasets for machine learning models. Emphasis is placed on ethical data management, responsible use of annotation tools, and ensuring data quality through cleaning and transformation. The course prepares learners to handle diverse annotation tasks efficiently, understand AI applications across industries, and address challenges like ethical compliance and risk mitigation, equipping them for success in the field of data annotation.</td></tr>
<tr><td>6.</td><td>Eligibility Criteria for Entry for a Student/Trainee/Learner/Employee</td><td colspan="2">a. Entry Qualification & Relevant Experience:
<table>
<tr><th>Academic/Skill Qualification (with Specialization - if applicable)</th><th>Required Experience (with Specialization - if applicable)</th></tr>
<tr><td>Grade 10 pass and pursuing continuous schooling*</td><td>No Experience required</td></tr>
<tr><td>Grade 8 with 2 years of NTC in the field of IT Sector*</td><td>No Experience required</td></tr>
<tr><td>Grade 8th pass*</td><td>3 Year Relevant Experience in IT</td></tr>
<tr><td>Previous NSQF Level 3 Required</td><td>1.5-year relevant experience in IT</td></tr>
</table>
Prerequisite: Fundamentals of Python
*Should have a basic understanding of coding.</td></tr>
</table>

<table>
<tr><td>7.</td><td>Credits Assigned to this NOS-Qualification, Subject to Assessment (as per National Credit Framework (NCrF))</td><td>4 Credits</td><td>8. Common Cost Norm Category (I/II/III): Category-II</td></tr>
<tr><td>9.</td><td>Any Licensing Requirements for Undertaking Training on This Qualification (wherever applicable)</td><td colspan="2">NA</td></tr>
<tr><td>10.</td><td>Training Duration by Modes of Training Delivery (Specify Total Duration as per selected training delivery modes and as per requirement of the qualification)</td><td colspan="2">☒Offline ☐Online ☐Blended
<table>
<tr><th>Training Delivery Modes</th><th>Theory (Hours)</th><th>Practical (Hours)</th><th>Total (Hours)</th></tr>
<tr><td>Classroom (Offline)</td><td>38</td><td>82</td><td>120</td></tr>
</table>
The mode of delivery shall be based on the regional demand and can be offered in any of the above-mentioned modes.
(Refer Blended Learning Annexure-V for details)</td></tr>
<tr><td>11.</td><td>Assessment Criteria</td><td colspan="2">
<table>
<tr><th>Theory (Marks)</th><th>Practical (Marks)</th><th>Project (Marks)</th><th>Viva (Marks)</th><th>Total (Marks)</th><th>Passing %age</th></tr>
<tr><td>100</td><td>60</td><td>20</td><td>20</td><td>200</td><td>50%</td></tr>
</table>
The centralized online assessment is conducted by the Examination Wing, NIELIT Headquarters.
*Assessment strategy shall be as per NIELIT Norms prevailing at times.</td></tr>
<tr><td>12.</td><td>Is the NOS Amenable to Persons with Disability</td><td colspan="2">☒ Yes ☐ No:
a. Locomotor Disability: Leprosy Cured Person, Dwarfism, Muscular Dystrophy and Acid Attack Victims
b. Visual Impairment: Low Vision</td></tr>
<tr><td>13.</td><td>Progression Path After Attaining the Qualification, wherever applicable (Please show Professional and Academic progression)</td><td colspan="2">Professional:
AI Development Associate
Academic:
Data Annotation Assistant → Senior Data Annotator → Data Annotation Specialist → Machine Learning Data Associate → AI/ML Engineer</td></tr>
<tr><td>14.</td><td>How participation of women will be encouraged?</td><td colspan="2">Participation of Women Candidates will be ensured as per Existing government norms. Exclusive batches for women would be run for the proposed course. Funding of women batches from IT for masses scheme /other schemes launched by government from time to time. Funding is</td></tr>
</table>

		available for women participation under other schemes launched by Government from time to time.	
15.	**Other Indian languages in which the Qualification & Model Curriculum are being submitted**	Qualification file is available in English and Hindi language.	
16.	**Is similar NOS available on NQR-if yes, justification for this qualification**	**☐ Yes ☒ No URLs of similar Qualifications:**	
17.	**Name and Contact Details Submitting / Awarding Body SPOC** *(In case of CS or MS, provide details of both Lead AB & Supporting ABs)*	**Name: Dr. Vimala Mathew, Scientist E** **Email: vimala@nielit.gov.in** **Contact No.: 0495-228 7266** **Website: https://nielit.gov.in/calicut/index.php**	
18.	**Final Approval Date by NSQC: 18.02.2025**	**19. Validity Duration:3 Years**	**20. 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/ Electronics and Communication/ allied branches Or MCA Or BCA Or A- Level (IT) with 1 year of Experience in training.
2.	**Master Trainer's Qualification and experience in the relevant sector (in years)** *(as per NCVET guidelines)*	B. Tech in Computer Science/ IT/ Electronics and Communication/ allied branches Or MCA Or BCA Or A- Level (IT) with 2 year of Experience in training.
3.	**Tools and Equipment Required for the Training**	☒Yes ☐No (*If "Yes", details to be provided 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/ Electronics and Communication/ allied branches Or MCA Or BCA Or A- Level (IT) with 2 year of Experience in training.
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 (Subject matter experts) are deployed including NIELIT scientific officers who are subject experts for evaluation of Practical examination/ internal assessment / Project/ Presentation/ assignment and Major Project (if applicable). Qualification is generally B.Tech.
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 (D*etails to be provided 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 candidates will be trained per year

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:** Industry Validation	Available at Annexure-III: Industry Validation
4.	**Annexure: Training Details**	Available at Annexure-IV: Training Details
5.	**Annexure:** Blended Learning *(Mandatory, in case selected Mode of delivery is Blended Learning)*	Available at Annexure-V: Blended Learning
6.	**Annexure/Supporting Document:** Standalone NOS-Performance Criteria Details Annexure/Document with PC-wise detailing as per NOS format (Mandatory-Public view)	Available at Annexure-VI: Standalone NOS- Performance Criteria details
7.	**Annexure:** Performance and Assessment Criteria*(Mandatory)*	Available at Annexure-VII: Detailed Assessment Criteria
8.	**Annexure:** Assessment Strategy*(Mandatory)*	Available at Annexure-VIII: Assessment Strategy
9.	**Annexure:** Acronym and Glossary *(Optional)*	Available at Annexure-IX: Acronym and Glossary
10.	**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	As a Data Annotator, the process involves routine and predictable tasks such as tagging, labeling, and categorizing data (e.g., text, images, or videos) using annotation tools. The role requires a systematic approach to follow predefined guidelines, ensuring data is correctly prepared for machine learning or AI applications. The tasks are repetitive but critical for maintaining consistency in data preparation	The Data Annotator role involves repetitive and routine tasks such as tagging and labeling data, which are predictable in nature. These tasks require adherence to specific guidelines and processes, aligning with the NSQF Level 3 descriptor of routine and structured work	**3.5**
Professional and Technical Skills/ Expertise/ Professional Knowledge	The role requires foundational knowledge of data types (e.g., structured and unstructured), annotation tools (e.g., Labeling, CVAT), and the purpose of annotations in training AI systems. Understanding quality standards, labeling guidelines, and basic troubleshooting of tools is essential. This knowledge ensures accuracy and reliability in annotated datasets	The course provides basic factual knowledge about data annotation, including understanding different types of data (text, images, videos) and the purpose of annotations in machine learning projects. This meets the NSQF Level 3 criterion of basic knowledge relevant to the field	**3.5**
Employment Readiness & Entrepreneurship Skills & Mind-set/Professional Skill	Data Annotators must develop skills in using annotation software to label data efficiently and accurately. They should also be able to detect and correct errors in annotations, ensuring quality output. These practical skills are applied under supervision, with clear instructions and adherence to established standards.	Learners acquire practical skills such as using annotation tools, applying labeling standards, and ensuring accuracy. These limited skills are applied under supervision, consistent with NSQF Level 3 requirements	**3.5**
Broad Learning Outcomes/Core Skill	The role demands basic communication skills to coordinate with supervisors and team members effectively, ensuring alignment on annotation guidelines. Numeracy skills are needed for simple data management tasks, while basic IT skills are essential for operating annotation tools, managing files, and maintaining records in digital formats	The course imparts basic communication skills to interact with team members and supervisors effectively. Learners also gain numeracy skills for managing simple datasets and IT skills for operating annotation software and tools	**3.5**
Responsibility	As a Level 3 Data Annotator, individuals are responsible for ensuring their assigned annotations meet quality and	As a Data Annotator, individuals are responsible for ensuring the quality and accuracy of their	**3.5**

	consistency standards. However, they work under close supervision, following predefined workflows and receiving regular guidance from supervisors or team leads	assigned annotations. However, they work under close supervision and follow predefined guidelines	

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	Class Room	1 (30 sq m)	1
2	Student Table	30	30
3	Student Chair	30	30
4	Desktop computer with accessories	Minimum i5 PC with 8 GB RAM Installed with: IDEs like PyCharm, Jupyter, AI libraries (TensorFlow, Scikit-learn), Annotation tools like Labelbox, Prodigy, video/image/audio editors, Computer Vision Annotation Tool (CVAT) project management tools, and soft skills training resources.	30

Classroom Aids

The aids required to conduct sessions in the classroom are:

1. LCD Projector

2. Pin-up Board

3. White Board

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 India AI Mission.

Annexure-IV: Training Details

Training Projections:

Year	Estimated Training # of Total Candidates	Estimated training # of Women	Estimated training # of People with Disability
2025-26	500	200	10
2026-27	500	200	10
2027-28	500	200	10

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	Online interaction platforms like JitSi Meet, Bharat VC, Google Meet, MS Teams, etc.	60:40
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 Remote Proctored Software	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:

This course offers a comprehensive 120-hour curriculum designed to equip students with essential programming skills, foundational knowledge in AI and data science, and practical data annotation techniques. Students will learn to set up their Python environment, utilize variables and data types, implement control structures, and handle files while gaining hands-on experience with libraries like NumPy and Pandas. The course also covers key AI concepts, ethical considerations, and data preprocessing techniques, alongside practical applications in data annotation across text, image, video, and audio formats. By engaging in real-world projects sourced from industry partners, students will develop collaborative skills and present their findings, preparing them for careers in technology and data management.

2. Scope:

The scope of this course encompasses a comprehensive exploration of programming fundamentals, AI concepts, and data annotation techniques. Students will gain practical skills in Python programming, including setting up environments, using data types, and implementing control structures, while also learning to apply AI principles in various sectors such as healthcare and finance.

3. Elements and Performance Criteria

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

Foundational in Python Programming:

PC1. Successfully install Python and an IDE, execute basic scripts, and troubleshoot errors.
PC2. Identify and correctly use data types, conditional statements, and loops to build simple programs.

Basics of Artificial Intelligence & Data Science

PC3. Explain fundamental AI concepts, including machine learning and NLP, with relevant industry examples.
PC4. Implement data preprocessing techniques such as data cleaning and normalization on sample datasets.

Introduction to Data Annotation

PC5. Differentiate between supervised and unsupervised learning and explain the role of data annotation.
PC6. Use open-source tools to annotate datasets by tagging and labeling elements accurately.

Text Annotation:

PC7. Explain text annotation methods like Named Entity Recognition (NER) and sentiment analysis with real-world applications.
PC8. Perform text annotation using open-source tools following project guidelines.

Image & Video Annotation

PC9. Describe image annotation techniques such as segmentation and object detection.
PC10. Demonstrate video annotation skills, including object tracking and frame-by-frame analysis, using appropriate tools.

Audio Annotation:

PC11. Explain audio annotation concepts, including speech-to-text conversion and emotion detection.
PC12. Conduct audio annotation using open-source tools according to defined requirements.

Emerging Trends in AI-Assisted Annotation and Best Practices:

PC13. Identify and explain AI-assisted annotation trends such as 3D annotation and synthetic data generation.
PC14. Apply quality assurance techniques to improve accuracy in annotated datasets.

Quality Control and Best Practices in AI:

PC15. Recognize common data annotation errors and apply strategies for error reduction.
PC16. Explain ethical considerations, including data privacy regulations and responsible AI practices.

Application of Data Annotation:

PC17. Apply learned annotation techniques to solve real-world problems from industry case studies.
PC18. Work collaboratively in teams, present findings effectively, and incorporate feedback for continuous improvement.

4. Knowledge and Understanding (KU):

The individual on the job needs to know and understand:

KU1: Students will understand how to set up the Python environment and install necessary tools, enabling them to execute Python scripts effectively.
KU2: Students will gain knowledge of fundamental AI concepts, including machine learning, deep learning, computer vision, and natural language processing, along with their applications in various industries such as healthcare and finance.
KU3: Students will comprehend strategies for managing large-scale and complex datasets with limited labeled data, preparing them for practical annotation challenges.

5. Generic Skills (GS):

User/individual on the job needs to know how to:

GS1: Students will enhance their technical skills by gaining hands-on experience with programming in Python, utilizing libraries like NumPy and Pandas, and working with various data annotation tools.
GS2: Students will learn to collaborate effectively with peers on group projects, enhancing their teamwork skills.
GS3: Students will enhance their communication skills through presentations of project findings to peers and instructors.

Annexure VII: Assessment Criteria

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 Annotation using Python **NOS Code: NIE/SSC/N1019**	**Foundational in Python Programming:** PC1. Successfully install Python and an IDE, execute basic scripts, and troubleshoot errors. PC2. Identify and correctly use data types, conditional statements, and loops to build simple programs.	11	6	2	2
	Basics of Artificial Intelligence & Data Science: PC3. Explain fundamental AI concepts, including machine learning and NLP, with relevant industry examples. PC4.Implement data preprocessing techniques such as data cleaning and normalization on sample datasets.	11	6	2	2
	Introduction to Data Annotation: PC5. Differentiate between supervised and unsupervised learning and explain the role of data annotation. PC6. Use open-source tools to annotate datasets by tagging and labeling elements accurately.	11	6	2	2
	Text Annotation: PC7.Explain text annotation methods like Named Entity Recognition (NER) and sentiment analysis with real-world applications. PC8. Perform text annotation using open-source tools following project guidelines.	11	7	2	2
	Image & Video Annotation: PC9. Describe image annotation techniques such as segmentation and object detection. PC10.Demonstrate video annotation skills, including object tracking and frame-by-frame analysis, using appropriate tools.	11	7	2	2

	Audio Annotation: PC11. Explain audio annotation concepts, including speech-to-text conversion and emotion detection. PC12. Conduct audio annotation using open-source tools according to defined requirements.	11	7	2	2
	Emerging trends in AI-Assisted Annotation and best practices: PC13.Identify and explain AI-assisted annotation trends such as 3D annotation and synthetic data generation. PC14. Apply quality assurance techniques to improve accuracy in annotated datasets.	22	14	5	5
	Quality Control and Best Practices in AI: PC15. Recognize common data annotation errors and apply strategies for error reduction. PC16. Explain ethical considerations, including data privacy regulations and responsible AI practices.				
	Application of Data Annotation: PC17. Apply learned annotation techniques to solve real-world problems from industry case studies. PC18. Work collaboratively in teams, present findings effectively, and incorporate feedback for continuous improvement.	12	7	3	3
		100	**60**	**20**	**20**
	Total Marks	**200**			

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.