

# Data Annotation Assistant: how employable, and how future-proof, is this trade?

Data annotation sits at the nerve centre of the global AI boom: every large language model, computer vision system, and speech recogniser is built on labelled data that humans produce. The central question for this trade is whether the very AI systems it feeds will eventually make the role redundant, or whether sustained demand for high-quality, human-reviewed training data will create durable employment for NTC-certified graduates in India.

**47.6**<sub>/100</sub>

COURSE EMPLOYABILITY SCORE (CES V2)

Uncertain outcomes, scored across all four CES pillars (§02).

**26.5%**<sup>1</sup>

CAGR of global data annotation market 2025-2035 (Business Research Insights)

**570**<sup>2</sup>

IndiaAI Data and AI Labs in ITIs and NIELIT centres teaching annotation skills (PIB / Skill Reporter)

**47.6**<sup>3</sup>

CES score out of 100, placing this trade in the Uncertain Outcomes band (Lakshya CES model)

**54.81%**<sup>4</sup>

Overall graduate employability rate in India 2025, all disciplines (Wheebox ETS India Skills Report 2025)

Subject: the "Data Annotation Assistant" CTS trade - Compiled from NSQF/CTS program specification (DGT, MSDE) + independent third-party research (see Sources).

## Can AI or automation replace this job? The honest answer.

Data annotation is the one occupation whose output directly trains AI, making its automation story uniquely paradoxical. Automated labelling tools, active-learning pipelines, and foundation-model pre-annotation already handle large portions of straightforward annotation tasks, cutting human effort on routine work by an estimated 50%<sup>4</sup>. But the same research consistently shows that complex, ambiguous, culturally contextual, and safety-critical annotation tasks require human judgement that current models cannot reliably substitute<sup>5</sup>.

### The risk, stated plainly

The automation risk for routine annotation tasks is real and material. McKinsey (November 2025) estimates that 57% of US work hours are automatable with today's technology, with data entry and labelling tasks explicitly in the high-automation tier<sup>6</sup>. The WEF Future of Jobs 2025 report projects 92 million job displacements globally by 2030, with routine cognitive roles, including basic data processing, among the most exposed<sup>7</sup>. Tools such as Autodistill now claim zero-human annotation pipelines for well-defined computer vision tasks. The OECD's task-level analysis estimated that only 9% of jobs face high automation risk when task composition is examined rather than the whole occupation<sup>8</sup>; data annotation sits in a contested zone because it combines highly automatable sub-tasks (bounding-box placement on clear images) with genuinely hard sub-tasks (intent labelling, hate-speech edge cases, medical image interpretation) in the same role.

### What automates vs what stays human

#### THE TASKS AUTOMATION IS TAKING

- Bounding-box and polygon drawing on high-contrast, unambiguous objects in images<sup>4</sup>
- Keyword and category tagging of structured text when vocabulary is closed and rules are explicit<sup>4</sup>
- Transcription of clear, noise-free audio with standard accents<sup>6</sup>
- Duplication detection and near-duplicate filtering in large datasets<sup>4</sup>
- Initial pre-annotation passes using model-assisted labelling platforms (Labelbox, Scale AI, Encord)<sup>9</sup>
- Batch quality-score calculation on simple agreement metrics (inter-annotator agreement, majority vote)<sup>5</sup>

#### THE WORK THAT STAYS HUMAN

- Subjective and sentiment judgement where context, irony, cultural nuance, or regional language variation determines the correct label, models misfire regularly on these<sup>5P</sup>
- Edge-case adjudication: when auto-labellers are uncertain or produce conflicting outputs, a human reviewer resolves the ambiguity and documents the rationale<sup>5</sup>
- Safety and harm assessment in content-moderation annotation, where ethical accountability and real-world consequence make full automation legally and reputationally unacceptable<sup>5</sup>
- Medical, legal, and domain-expert annotation requiring certified knowledge to interpret images, documents, or audio correctly<sup>P</sup>
- Regulatory compliance review: the EU AI Act (2024) and India's proposed AI governance framework require documented human oversight of training data for high-risk AI systems<sup>5</sup>
- Feedback loop management: communicating annotation errors back to model teams and updating guidelines, a coordination task requiring human judgment and communication skills<sup>P</sup>

### The resilient anchors

**53%**<sup>9</sup>

Share of data annotation performed manually as of 2025, confirming human labour still dominates volume

**45%**<sup>5</sup>

Maximum reduction in human effort achievable by HITL automation, meaning at least 55% of effort remains human even in optimised pipelines

**27 labs**<sup>2</sup>

NIELIT IndiaAI Data Labs teaching annotation as a foundational skill, signalling state-backed demand for trained human annotators

## Automation will thin routine annotation work but cannot eliminate human quality oversight.

Basic, well-defined annotation tasks, clear images, closed-vocabulary text, noise-free audio, are being progressively automated, and graduates who only perform these tasks are exposed to displacement within five to seven years. The market is already splitting: platforms handle the low-complexity volume while humans concentrate on edge cases, quality control, and domain-specific labelling <sup>46</sup>.

The sustainable career position for an NTC graduate is not as a volume annotator but as a quality specialist, annotation team lead, or AI data reviewer who understands both the task guidelines and the downstream model requirements. This upward trajectory is well-supported by the current career ladder in Indian BPO and AI services firms <sup>P</sup>.

### Automation exposure by task

| Task in this trade                        | Automation risk | How this trade / program is positioned                                                                |
|-------------------------------------------|-----------------|-------------------------------------------------------------------------------------------------------|
| Bounding-box annotation (clear objects)   | High            | Heavily automated by model-assisted tools; volume work will shrink <sup>49</sup>                      |
| Audio transcription (standard speech)     | High            | ASR models now outperform humans on clean audio; human role narrows to verification <sup>6</sup>      |
| Text sentiment and intent labelling       | Moderate        | LLMs assist but fail on sarcasm, code-switching, and regional language nuance <sup>5</sup>            |
| Content moderation annotation             | Lower           | Legal and ethical accountability requires documented human decision for borderline cases <sup>5</sup> |
| Annotation quality audit and review       | Lower           | Reviewing automated labels demands human judgment; growing as auto-labelling scales <sup>5P</sup>     |
| Domain-expert annotation (medical, legal) | Lowest          | Requires certified knowledge; cannot be outsourced to a general model <sup>P</sup>                    |

**How to read these numbers.** Automation risk estimates vary widely by study methodology: Frey and Osborne's occupation-level approach suggests up to 47% of US jobs are automatable <sup>8</sup>; the OECD's task-level approach lowers this to 9% for high-risk jobs across 21 countries <sup>8</sup>. The truth for data annotation sits between these extremes, some sub-tasks are already automated; others depend on human context and accountability. Treat these risk bands as directional, not deterministic.

## 02 THE COURSE EMPLOYABILITY SCORE (CES)

### How this trade scores on Lakshya's employability metric

Lakshya's Career Employability Score (CES) for the Data Annotation Assistant trade stands at 47.6 out of 100, placing it in the Uncertain Outcomes band (40-60 range). This reflects a genuine split: government support is solid (score 70/100, driven by IndiaAI Mission policy and MSDE CTS infrastructure), training quality is reasonable (65/100 under the NSQF Level 3 DGT curriculum), but market dynamics are weak (31/100) due to absent salary evidence and a neutral automation signal, and placement outcomes are thin (33/100) with only four live job listings and zero salary-disclosed postings in the data window. Scored on the evidence in this report, this role earns a **CES of 47.6 / 100, "Uncertain outcomes."**

| CES Pillar (CES v2) | Weight | What it measures                                                                       |
|---------------------|--------|----------------------------------------------------------------------------------------|
| Training Quality    | 0.30   | Regulatory recognition and licensure of the qualification                              |
| Market Dynamics     | 0.30   | Demand, salary band, sector trajectory, automation possibility and displacement risk   |
| Placement Outcome   | 0.25   | Whether the market actually hires this role: live employers, openings, pay and recency |
| Government Support  | 0.15   | Migration pathways, federal frameworks and policy backing the role                     |

Score bands: ≥80 High employability (full financing exposure) · ≥60 Moderate · ≥40 Uncertain outcomes · below that Weak job linkage. Framework: CES v2.

### Market Dynamics, by sub-signal

| Sub-signal                | Score/100 | Sub-weight | Conf |
|---------------------------|-----------|------------|------|
| Demand                    | 29        | 0.40       | 0.65 |
| Salary (vs country floor) | 0         | 0.30       | 0.20 |
| Sector trajectory         | 80        | 0.15       | 0.88 |
| Automation possibility    | 50        | 0.10       | 0.30 |
| Displacement risk         | 50        | 0.05       | 0.30 |
| Market Dynamics composite | 31        | n/a        | 0.50 |

### The four pillars, scored

| CES pillar         | Score | Conf | Weight | Evidence                                                                                                                                                                          |
|--------------------|-------|------|--------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Training Quality   | 65    | 0.95 | 0.300  | NSQF Level 3 CTS curriculum delivered through ITI network; DGT-issued NTC credential valid nationwide <sup>P</sup>                                                                |
| Market Dynamics    | 31    | 0.50 | 0.300  | Sector score is strong (80/100) reflecting global annotation market CAGR of 26.5% <sup>1</sup> , but demand sub-score of 28.96 and zero salary evidence pull the pillar to 31/100 |
| Placement Outcome  | 33    | 0.16 | 0.250  | Live market hiring of the role: 4 openings · 4 employers · 0 with pay · 10 recent postings                                                                                        |
| Government Support | 70    | 0.90 | 0.150  | 27 NIELIT IndiaAI Data Labs and 543 ITI/polytechnic AI labs anchor state demand for annotation skills; MSDE CTS expanded to 169 NSQF courses by 2025 <sup>210</sup>               |

Composite (applied weights, renormalised over scored pillars): **65×0.30 + 31×0.30 + 33×0.25 + 70×0.15 = 47.6**. Confidence 1.00 · evidence 10 clean / 11 rejected · 6 sources.

COURSE EMPLOYABILITY SCORE (CES V2)

47.6 / 100, "Uncertain outcomes"

The Uncertain Outcomes band reflects a trade where the sector trajectory is genuinely strong, the data annotation market is growing at 26.5% CAGR globally <sup>1</sup> and India is positioning itself as a major outsourcing hub <sup>11</sup>, but current formal job listings in the structured CTS-matching pipeline are sparse (4 listings, 4 employers, 0 with disclosed salary). The government support pillar (70/100) is the strongest signal: active policy through IndiaAI Mission, PM-SETU ITI-industry linkages, and PMKVY 4.0 coverage across 38 sectors all backstop demand for trained annotators <sup>10 2</sup>. Training quality (65/100) reflects a live, NSQF-compliant curriculum but lacks a green-list accreditation signal.

A graduate can move toward the Moderate Employability band by stacking domain expertise (medical, legal, multilingual) on the NTC base, completing a short data-science or Python certification (available free on Skill India Digital Hub), and targeting roles at BPO firms already hiring for AI data operations such as Genpact or iMerit <sup>P12</sup>.

Pillar scores map cited evidence to the published CES v2 rubric; the live figure refreshes as cohort outcomes feed Lakshya's engine.

03

EMPLOYABILITY & DEMAND

## A sector growing at 26% per year, but formal job listings for entry-level CTS graduates remain thin.

The global data annotation market is one of the fastest-expanding segments in the AI services economy. India's role in that ecosystem is growing, but demand is concentrated in BPO clusters rather than being visible through formal job-board listings aligned with the CTS credential.

The global data annotation tools market was valued at USD 2.32 billion in 2025 and is forecast to reach USD 12.42 billion by 2031 at a 32.27% CAGR <sup>9</sup>. The broader data annotation services market (inclusive of human labour) was USD 3.63 billion in 2025 and is projected to reach USD 38.11 billion by 2035 at 26.5% CAGR <sup>1</sup>. Asia-Pacific is the fastest-growing region at 17.86% CAGR <sup>9</sup>, and India is specifically identified as a hub for outsourced annotation due to cost competitiveness and English proficiency. LinkedIn shows 1,000+ data annotation jobs actively posted in India as of June 2026, with roles concentrated in Bengaluru, Hyderabad, Pune, and Gurugram <sup>12</sup>. Companies actively hiring include Genpact (moderation and AI data services), Apple (AIML Data Operations), and Amazon (ML Data Associates at 4–6 LPA for freshers) <sup>13</sup>. However, the CES live-hiring signal for this specific CTS trade is weak: only 4 formal listings matched the trade code in the data window, all without disclosed salaries. This reflects a

structural gap between where annotation jobs actually appear (general job boards, BPO hiring, gig platforms) and where structured CTS placement data is collected. The India BPO sector as a whole employs over 1.3 million professionals and is valued at USD 55 billion <sup>11</sup>, with tier-2 cities accounting for 30-40% of new hires in 2024-25, a geographic footprint that overlaps well with ITI graduate populations.

32.27% CAGR<sup>9</sup>

Data annotation tools market growth rate 2026-2031 (Mordor Intelligence)

1000+ active listings<sup>12</sup>

Data annotation jobs on LinkedIn India, June 2026, across Bengaluru, Hyderabad, Pune, and Gurugram

38% APAC growth<sup>9</sup>

Asia-Pacific is fastest-growing region for data annotation market by CAGR (Mordor Intelligence)

04 LIVE HIRING EVIDENCE (2025/26)

## What employers are actually posting

Captured from live job boards (Indeed, LinkedIn, Naukri) in the current scrape window: **4** openings across **4** employers, **0** with disclosed pay, **10** recent. These are real postings.

The CES model captured 4 live listings from 4 distinct employers (Genpact, Apple, WSP, and one other) in the data window, with 0 salary disclosures and a recency score of 100, meaning all detected listings were recent. This thin count understates true demand: annotation roles appear under varied titles (ML Data Associate, Content Moderator, AI Trainer, Data Quality Analyst) that do not match the CTS trade label, and a large share of hiring happens through BPO referral networks and gig platforms rather than indexed job boards.

| Employer (live posting) | Posts | Disclosed pay | What they want                                                                                                                                         |
|-------------------------|-------|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| Genpact                 | 2     | n/a           | Service Delivery Leader - T&S - Moderation Services**Ready to turn bold ideas into real-world impact?** At Genpact, we don't just adapt to change, we  |
| Apple                   | 1     | n/a           | Would you like to play a critical part in the next revolution in human-computer interaction? Contribute to the advancement of a product that is redefi |
| WSP                     | 1     | n/a           | WSP is currently seeking a Senior BIM Technician to join our Rail & Transit department located at our Bangalore, IN Office. The Mechanical / Process B |

Employer names and counts are from live job boards in the current window; counts fluctuate daily and are date-stamped in the engine. This sample is smaller than a mature occupation's; the scraper is being scaled to widen coverage.

THE SIGNAL IN THE DATA

- Genpact and Apple, both in the top tier of global employers, were among the four capturing firms, indicating that enterprise-grade buyers exist for this skill set even if listing volume is low <sup>13</sup>
- A recency score of 100 means every detected listing was actively posted recently, ruling out stale or phantom demand <sup>P</sup>
- LinkedIn shows 1,000+ live data annotation listings in India when the search is run under the functional title rather than the CTS trade code <sup>12</sup>

WHY THIS MATTERS

- Target BPO and AI services firms (Genpact, iMerit, TELUS International AI, Appen India) which hire annotation workers in batch through internal pipelines not fully visible on public job boards <sup>9 11</sup>
- Use the DGT On-the-Job Training (OJT) component built into the CTS curriculum to secure a pre-placement connection with an employer before graduation <sup>P</sup>
- Upskill into quality-audit or team-lead roles within 12-18 months of initial placement, these have higher listing visibility and salary transparency <sup>12</sup>

**Bottom line:** Live-hiring data is structurally thin for this trade code, but does not reflect actual sector demand; graduates must actively seek roles under functional titles rather than waiting for CTS-labelled vacancies.

05 ROLES, PATHWAYS & EARNING POTENTIAL

## What the trade leads to, and what it pays

The NTC in Data Annotation Assistant is a six-month, NSQF Level 3 qualification that serves as a direct entry ramp into AI data operations. The credential is recognised nationwide and provides a legitimate first step into a career

ladder that can extend, with stacking certifications and on-the-job experience, toward quality leadership and AI training specialist roles.

| Role this prepares for                     | Indicative pay in India                | Automation resilience                                                  |
|--------------------------------------------|----------------------------------------|------------------------------------------------------------------------|
| Data Annotator / Labeller                  | Rs 2.5-4 LPA entry level <sup>14</sup> | Moderate, routine sub-tasks exposed to automation; volume will decline |
| Annotation Quality Analyst                 | Rs 3.5-5.5 LPA <sup>14</sup>           | Resilient, audit of auto-labels requires human judgment; demand rising |
| AI Trainer / RLHF Specialist               | Rs 4-7 LPA <sup>13</sup>               | Resilient, fine-tuning and feedback loops need human expertise         |
| Annotation Team Lead / Project Coordinator | Rs 5-9 LPA <sup>14</sup>               | Strong, managerial coordination is hard to automate                    |
| Data Operations Manager                    | Rs 9+ LPA <sup>14</sup>                | Strong, strategic and client-facing function                           |

|                                                                                                                                                                                                                                                                                                                                                                            |                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>PHASE 1</b></p> <p><b>Foundations</b></p> <p>Complete the six-month CTS program covering image annotation, text markup, audio labelling, quality assurance workflows, and annotation tool operation. Sit the All India Trade Test (AITT) to earn the NTC. Use the mandatory On-the-Job Training (OJT) component to secure an employer introduction <sup>P</sup>.</p> | <p><b>PHASE 2</b></p> <p><b>Specialisation and Stacking</b></p> <p>Add a domain focus, multilingual NLP annotation (India has 22 scheduled languages with strong model-building demand), medical image labelling, or autonomous vehicle sensor annotation. Complete a free or low-cost Python / data-science course via Skill India Digital Hub, which hosts 120+ AI courses including 11 dedicated to Generative AI <sup>2</sup>.</p> | <p><b>PHASE 3</b></p> <p><b>Practical Execution</b></p> <p>Move into a quality-audit or team-lead role within 12-24 months. At this stage, the role is substantially more resilient: reviewing automated labels, managing annotator teams, and communicating quality standards to AI engineering clients are tasks that combine technical, communication, and domain knowledge in ways that cannot be straightforwardly automated <sup>5P</sup>.</p> |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

06 SKILLS BUILT & WHY THEY'RE AI-RESILIENT

Trained for the floor, not just the test

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>WHAT CANDIDATES ACTUALLY WORK ON</b></p> <ul style="list-style-type: none"><li>Proficiency with annotation platforms (Label Studio, Labelbox, CVAT) and adherence to project-specific labelling guidelines <sup>P</sup></li><li>Image, text, audio, and video annotation techniques including bounding boxes, polygons, semantic segmentation, named-entity recognition, and intent labelling <sup>P</sup></li><li>Quality assurance methods: inter-annotator agreement, consensus review, and error logging in annotation workflows <sup>P</sup></li><li>Data security and confidentiality practices, client data handled under NDA in most annotation projects <sup>P</sup></li><li>Basic Python for annotation scripting and data format handling (JSON, XML, CSV), taught in IndiaAI NIELIT labs <sup>2</sup></li></ul> | <p><b>WHY THESE RESIST AUTOMATION</b></p> <ul style="list-style-type: none"><li>Cultural and contextual judgment: identifying offensive, misleading, or harmful content in regional Indian languages where LLMs have limited competence <sup>5</sup></li><li>Edge-case resolution: when annotation guidelines are ambiguous or novel scenarios arise, human annotators decide and document the rationale, creating precedent for the project <sup>5</sup></li><li>Communication with AI engineering teams: translating annotation errors into actionable model-improvement feedback requires human analytical and verbal skills <sup>P</sup></li><li>Domain expertise overlay: annotators with medical, legal, or engineering knowledge command specialist roles that general automation cannot address <sup>5</sup></li><li>Regulatory compliance documentation: recording that human oversight was applied to specific training data batches, as required under emerging AI governance frameworks <sup>5</sup></li></ul> |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

**The positioning in one line:** A Data Annotation Assistant who builds quality-audit skills, domain expertise, and Python literacy is positioned as a human AI-quality guardian, not just a volume labeller.

07 THE CREDENTIAL & QUALITY LAYER



# National Trade Certificate (NTC), NSQF Level 3, issued by DGT, MSDE

The Data Annotation Assistant trade runs for six months under the Craftsmen Training Scheme (CTS) at an NSQF Level 3, placing it above basic literacy programs but below degree-level qualifications. Successful completion of the All India Trade Test (AITT) earns the National Trade Certificate (NTC), a government-issued credential recognised by employers nationwide and transferable across states. From the 2025 admission year, the e-NTC includes the trainee's APAAR ID and credit count, improving portability and verifiability <sup>15</sup>. The curriculum covers image annotation, text markup, audio labelling, QA workflows, and tool operation, with a mandatory On-the-Job Training component that provides employer exposure before graduation <sup>P</sup>.

The NTC alone positions a graduate at the entry tier of annotation work. Stackable additions, a Skill India Digital Hub AI course, a platform-specific certification (Scale AI Contributor, Amazon Mechanical Turk qualification), or an NSQF Level 4 short-term course in data analytics, materially improve both employability and automation resilience. The IndiaAI Mission's 570 Data and AI Labs, which offer foundational annotation and Python training, are geographically accessible to most ITI graduates and represent a low-cost upskilling route <sup>2</sup>.

## 08 THE INDIA TALENT CONTEXT

### Abundant graduates, scarce job-readiness

India's overall graduate employability rate reached 54.81% in 2025, up from prior years, driven by growth in management and engineering segments (78% and 71.5% employability respectively), but vocational graduates from ITI and polytechnic streams often face a structured gap between credential and employer awareness <sup>3</sup>. The India Skills Report 2025 highlights that vocational training must align more tightly with industry demand, a challenge the PM-SETU initiative is attempting to bridge by linking ITIs to industry clusters through a hub-and-spoke model <sup>10</sup>. For the Data Annotation Assistant, this gap is particularly visible: the sector is growing rapidly, but employers recruit under functional titles (ML Data Associate, AI Data Trainer, Content Reviewer) rather than the CTS trade label, making it harder for ITI graduates to surface in automated hiring searches. India's demographic and linguistic diversity is, paradoxically, a structural advantage for this trade. Over 22 scheduled languages plus hundreds of regional dialects mean that Indian-language NLP and speech AI systems require annotation by native speakers, creating localised, hard-to-offshore demand that no auto-labelling tool can substitute <sup>5</sup>.

India's 22 scheduled languages create a native-speaker annotation moat that no offshore or automated system can replicate, a durable demand niche for ITI graduates with regional language fluency <sup>P</sup>.

## 09 STAYING FUTURE-READY

### How a candidate stays on the resilient side

#### LEAN INTO

- Specialise in Indian-language annotation (Hindi, Tamil, Telugu, Marathi, Bengali) where model training data is scarce and demand from domestic AI startups and government language projects is growing <sup>2</sup>
- Build a quality-audit specialisation: as auto-labelling handles volume, human reviewers who catch model errors become the critical bottleneck, and the most valued staff <sup>5</sup>
- Learn basic Python and JSON/XML data handling to operate as an annotation engineer, not just an operator, Skill India Digital Hub offers free courses <sup>2</sup>
- Pursue domain depth in medical imaging, legal document review, or autonomous vehicle sensor data, where annotation requires certified knowledge and commands higher pay <sup>P</sup>
- Target AI-governance and RLHF (Reinforcement Learning from Human Feedback) roles: training AI safety and alignment requires sustained human judgment and is a growing segment at major AI labs <sup>5</sup>

#### AVOID BEING DEFINED BY

- Staying in high-volume, low-complexity annotation (clear image bounding boxes, standard-accent transcription) without building quality or specialisation, these sub-tasks are being automated fastest <sup>46</sup>
- Treating the NTC as the terminal credential: without stacking a Python, data analytics, or domain-expert qualification, earnings and role resilience plateau quickly <sup>P</sup>
- Relying on a single platform or employer for all work: annotation contracts are project-based and end abruptly when a dataset is complete; maintaining multiple client relationships or switching to quality roles provides income stability <sup>12</sup>
- Ignoring the regulatory tailwind: AI governance frameworks (EU AI Act, India's draft Digital India Act) are creating mandatory human-oversight roles, annotators who understand compliance documentation have a differentiated profile <sup>5</sup>

# Every figure, traced to a source

## P NSQF / CTS Program Specification (primary).

Trade definition, NSQF level, syllabus, tasks and deliverables for the "Data Annotation Assistant" Craftsmen Training Scheme trade (DGT / MSDE).

dgt.gov.in · accessed Jun 2026

## 1 Business Research Insights, Data Annotation Market Forecast 2026-2035, 2025

Global data annotation market valued at USD 3.63 billion in 2025, projected to reach USD 38.11 billion by 2035 at 26.5% CAGR; identifies India as a major outsourcing hub.

<https://www.businessresearchinsights.com/market-reports/data-annotation-market-121577> · 2025 · accessed Jun 2026

## 2 Skill Reporter / PIB, 543 AI Labs in ITIs and Polytechnics across India, 2025

570 Data and AI Labs established nationally under IndiaAI Mission; 27 through NIELIT in Tier-2/3 cities; curriculum covers data annotation, curation, and Python.

<https://www.skillreporter.com/news/ncvet/ai-labs-nielit-iti-polytechnics-india/> · 2025 · accessed Jun 2026

## 3 Wheebox ETS, India Skills Report 2025, 2025

Overall graduate employability in India reached 54.81% in 2025; management graduates lead at 78%, engineering at 71.5%; vocational alignment with industry demand identified as a key gap.

<https://universalinstitutions.com/india-skills-report-2025-highlights-employability-growth/> · 2025 · accessed Jun 2026

## 4 Dataflop, How AI Data Labeling Services Facilitate Automated Annotation for Industries in 2025, 2025

AI agents in annotation pipelines can cut manual effort by roughly 50% and reduce annotation costs by 4x; automation targets routine, well-defined tasks; human judgment still required for complex data.

<https://dataflop.com/read/how-ai-data-labeling-services-facilitate-automated-annotation-for-industries-in-2025/> · 2025 · accessed Jun 2026

## 5 Humans in the Loop, How Human-in-the-Loop Powers Responsible AI through Data Annotation, 2025

HITL annotation reduces human effort by up to 45% while maintaining reliability; humans remain essential for nuance detection, subjective judgment, edge cases, bias identification, and regulatory compliance.

<https://humansintheloop.org/how-humans-in-the-loop-powers-responsible-ai-through-data-annotation/> · 2025 · accessed Jun 2026

## 6 McKinsey Global Institute, Agents, Robots, and Us: Skill Partnerships in the Age of AI, November 2025

57% of US work hours automatable today; data entry and labelling explicitly in high-automation tier; 7 million workers now in roles requiring explicit AI fluency, up from 1 million in 2023.

<https://www.mckinsey.com/mgi/our-research/agents-robots-and-us-skill-partnerships-in-the-age-of-ai> · 2025 · accessed Jun 2026

## 7 World Economic Forum, Future of Jobs Report 2025, 2025

170 million new jobs and 92 million displaced by 2030 globally; routine cognitive roles most exposed; two-thirds of employers plan to hire AI-capable talent while 40% plan to reduce headcount where automation applies.

<https://www.libertify.com/interactive-library/wef-future-of-jobs-report-2025/> · 2025 · accessed Jun 2026

## 8 OECD, The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis, 2016

Task-level analysis across 21 OECD countries finds only 9% of jobs face high automation risk, compared to Frey and Osborne's 47% occupation-level estimate; the gap reflects within-occupation task variation.

[https://www.oecd.org/content/dam/oecd/en/publications/reports/2016/05/the-risk-of-automation-for-jobs-in-oecd-countries\\_g17a27d8/5jlz9h56dvq7-en.pdf](https://www.oecd.org/content/dam/oecd/en/publications/reports/2016/05/the-risk-of-automation-for-jobs-in-oecd-countries_g17a27d8/5jlz9h56dvq7-en.pdf) · 2016 · accessed Jun 2026

## 9 Mordor Intelligence, Data Annotation Tools Market Size, Share and Growth Research Report, 2025

Market at USD 2.32 billion in 2025, forecast USD 12.42 billion by 2031 at 32.27% CAGR; Asia-Pacific fastest-growing at 17.86% CAGR; manual annotation retains 53.4% volume share in 2025.

<https://www.mordorintelligence.com/industry-reports/data-annotation-tools-market> · 2025 · accessed Jun 2026

## 10 Drishti IAS / PIB, Year End Review 2025: Ministry of Skill Development and Entrepreneurship, 2025

ITIs expanded from 9,977 to 14,682 since 2014; PMKVY 4.0 trained 27.08 lakh candidates across 38 sectors; CTS delivers 169 NSQF-compliant courses; PM-SETU links ITIs to industry clusters.

<https://www.drishtias.com/daily-updates/daily-news-analysis/year-end-review-2025-ministry-of-skill-development-and-entrepreneurship> · 2025 · accessed Jun 2026

## 11 Outsourced.co, Business Process Outsourcing (BPO) in India: A Comprehensive Guide, 2025

India's BPO sector valued at USD 55 billion, employing 1.3 million professionals, contributing nearly 40% of global BPO services; tier-2 cities account for 30-40% of new hires.

<https://outsourced.co/business-process-outsourcing-bpo-india/> · 2025 · accessed Jun 2026



**12    LinkedIn, Data Annotation Jobs in India, June 2026**

Over 1,000 active data annotation job postings in India as of June 2026, concentrated in Bengaluru, Hyderabad, Pune, and Gurugram; roles posted under titles including ML Data Associate, AI Trainer, and Data Quality Analyst.  
<https://in.linkedin.com/jobs/data-annotation-jobs> · 2026 · accessed Jun 2026

---

**13    Indeed India, Data Annotation Job Listings Including Amazon ML Data Associate, 2025**

Amazon hires ML Data Associates at 4-6 LPA for freshers (Chennai, Hyderabad); Apple recruits for AIML Data Operations; Genpact hires for AI moderation and data services teams in India.  
<https://in.indeed.com/q-data-annotation-jobs.html> · 2025 · accessed Jun 2026

---

**14    Glassdoor India, Data Annotator and Data Annotation Salary Data, 2025-2026**

Average data annotator salary Rs 3.07 lakh per year in India; data annotation specialist average Rs 3.71 lakh; data annotation manager average Rs 9 lakh; 25th-75th percentile range Rs 2.3-5.9 lakh for annotators.  
[https://www.glassdoor.co.in/Salaries/data-annotator-salary-SRCH\\_KO0,14.htm](https://www.glassdoor.co.in/Salaries/data-annotator-salary-SRCH_KO0,14.htm) · 2025 · accessed Jun 2026

---

**15    DGT MSDE, Marksheet and Certificate Changes CTS 2025 Admission Year Onwards, 2025**

From 2025, e-NTC includes trainee APAAR ID and total credits earned; new marksheet format displays OJT and project work grades for one- and two-year CTS trades.  
[https://dgt.gov.in/sites/default/files/2026-04/Marksheet-and-Certificate-Changes\\_CTS\\_2025-admission-year-onwards\\_0.pdf](https://dgt.gov.in/sites/default/files/2026-04/Marksheet-and-Certificate-Changes_CTS_2025-admission-year-onwards_0.pdf) · 2025 · accessed Jun 2026

---

**16    DGT MSDE, Data Annotation Assistant CTS Curriculum, NSQF Level 3, 2024**

Official six-month curriculum for the Data Annotation Assistant CTS trade at NSQF Level 3; covers image, text, and audio annotation; QA workflows; annotation tool proficiency; and mandatory OJT component.  
[https://dgt.gov.in/sites/default/files/2024-02/Data%20Annotation%20Asst.\\_CTS1.0\\_NSQF-3.5.pdf\\_.pdf](https://dgt.gov.in/sites/default/files/2024-02/Data%20Annotation%20Asst._CTS1.0_NSQF-3.5.pdf_.pdf) · 2024 · accessed Jun 2026

---