RFP Response: Scope of Work (SoW)

Development of a U.S. Robotics and Artificial Intelligence Training and Education Programs System and Database

Table of Contents

Introduction

Draup is pleased to submit this proposal in response to the ARM Institute's Request for Proposal (RFP #25-01) for the development of a U.S. Robotics and Artificial Intelligence Training and Education Programs System and Database. As a global leader in talent intelligence and data curation, Draup specializes in acquiring, validating, and enriching large-scale datasets aligned with specific industry competency frameworks.

In this engagement, Draup proposes to serve as the **data provider**—responsible for identifying, collecting, validating, classifying, and delivering a comprehensive, export-ready database of U.S.-based robotics and AI training providers and programs mapped to the ARM Institute's Robotics and Artificial Intelligence Competency Frameworks. Our proven methodologies in structured data extraction, normalization, and quality assurance ensure that the ARM Institute will receive high-quality, current, and framework-aligned data that can be seamlessly integrated into the existing RoboticsCareer.org platform or any future system the Institute develops.

By focusing exclusively on **data acquisition and delivery**, Draup enables the ARM Institute to leverage our expertise and infrastructure for sourcing and classifying educational program data while retaining the flexibility to design, build, and manage the technology platform in-house or through another vendor.

Objectives

1. **Deliver a Comprehensive Dataset**: Provide a complete, validated list of U.S. training providers and programs relevant to robotics and AI, including all attributes outlined in Section 3.2 of the RFP, ready for direct integration into the RoboticsCareer.org data structure.
2. **Ensure Alignment with Competency Frameworks**: Map each program to the relevant competencies and essential soft skills within the ARM Institute's Robotics and AI Competency Frameworks, ensuring accurate tagging and classification for both current and future framework expansions.
3. **Implement Rigorous Data Quality Standards**: Apply Draup's proprietary validation and normalization processes to ensure data accuracy, consistency in field values, and elimination of duplicates, meeting the ARM Institute's quarterly update requirements.
4. **Provide Flexible Data Delivery Formats**: Deliver the data in ARM Institute–compatible formats (PostgreSQL/MySQL, CSV, or REST API output) to enable seamless system ingestion without requiring Draup to manage or host the system.
5. **Support Scalability and Future Growth**: Structure the dataset to support competency frameworks over the next two years, allowing the ARM Institute to expand scope without reworking the underlying data foundation.

1. Company Background

Company overview

Draup is a leading AI-powered Talent Intelligence and Academic Intelligence platform, built to help enterprise leaders make smarter, faster decisions through deep data insights and intelligent workflow automation.

Founded by Vijay Swaminathan (CEO) and Vamsee Tirukkala (CCO)—visionaries with over 30 years of experience in global consulting and talent analytics—Draup is redefining how organizations approach workforce strategy, academic partnerships, and digital transformation. The founders previously built **Zinnov**, a global consulting firm, and **TalentNeuron**, the industry's first talent intelligence platform (acquired by CEB, a Gartner company).

Talent Intelligence Platform

Draup specializes in delivering multidimensional insights on the external labor market through a secure, cloud-based SaaS platform, custom consulting reports, and API integrations. Key areas of expertise include:

- Talent Supply & Demand
- Compensation Benchmarking
- Skills Architecture & Future Skills
- Diversity & Inclusion
- Reskilling & Workforce Planning
- Library of courses relevant to Knowledge work

Powered by proprietary ML, AI, and GenAI models, Draup analyzes over **30 million+ data points** daily from **75,000+** global sources, enabling near real-time intelligence for HR and business leaders.

Academic Intelligence Platform

Draup also delivers deep academic intelligence that goes far beyond traditional alumni directories. The platform continuously harvests and reconciles publicly available data from university websites, Google Scholar, ORCiD, and other repositories to build a living graph of:

- **50,000+** higher education institutions across 65 countries
- **185,000+** professor profiles
- Millions of academic program records

This enables enterprises to:

- Explore full catalogs of degrees, micro-credentials, and certificates mapped to job roles.
- Monitor graduate output to forecast future talent supply.
- Benchmark institutions using Draup's proprietary **Academic Research Index (0–10)**.

- Track domain expertise in emerging areas like AI, blockchain, and sustainability.
- Identify thought leaders for collaboration, advisory roles, or speaking engagements.

All insights are refreshed on a rolling schedule and accessible via dashboards and APIs, supporting use cases across **recruiting, workforce planning, R&D, and academic partnerships**.

Relevant Experience

In this engagement, Draup proposes to serve as the data provider—responsible for identifying, collecting, validating, classifying, and delivering a comprehensive, export-ready database of U.S.-based robotics and AI training providers and programs. These will be systematically mapped to the ARM Institute's Robotics and Artificial Intelligence Competency Frameworks, ensuring accuracy, relevance, and industry alignment.

Draup's Academic Intelligence capability underpins this effort. Over years of working across Engineering, Telecom, and Knowledge Worker domains, Draup has assembled granular, role-aligned course libraries that connect real-world job tasks to specific skills, and those skills to targeted learning modules. This experience ensures the ARM Institute will receive a database that is:

- Skills-Framework Aligned – Each training program will be classified according to ARM's robotics and AI competencies, with clear mapping to the knowledge, skills, and abilities they develop.
- Domain-Calibrated – Draup's prior work in Engineering, Telecom, and Knowledge Worker pathways demonstrates expertise in assessing technical depth, applied tool usage, and market relevance.
- Validated & Comprehensive – All providers and programs will be vetted for credibility, recency, and completeness, ensuring decision-makers have high-confidence data.
- Export-Ready – The database will be structured for easy integration into ARM Institute systems, enabling instant use in workforce planning, training partner selection, and capability analysis.

By combining Draup's Academic Intelligence methodology with a rigorous data engineering process, this engagement will deliver more than a directory—it will provide a strategic, skills-mapped training intelligence asset for advancing robotics and AI workforce readiness.

2. Project Methodology

1. *Draup Data Sourcing*

The first step towards the development of a comprehensive database of Robotics Curriculum will start with the identification of publicly available data sources by Draup researchers. This will be followed by a comprehensive sourcing of these data assets through Draup's Big Data Harvesters. Draup's harvester models will target publicly available U.S.-based data sources to identify training providers and programs relevant to robotics and AI.

1.1 Existing Draup Education Corpus Leverage

Draup maintains a validated dataset of:

- **200,000+ active courses and certifications** across STEM, manufacturing, robotics, and AI domains.
- **600,000+ university curricula** from U.S. and global institutions, including course titles, descriptions mapped back to programs offered by the campuses.

Many of these records are already mapped to relevant skill domains through Draup's proprietary enrichment pipelines.
For the ARM Institute, Draup will:

- Filter the corpus to U.S.-based providers/programs.
- Re-map relevant programs to ARM's Robotics and AI Competency Frameworks.
- Update metadata to ensure currency, accuracy, and alignment with ARM's required fields.

This immediately gives ARM a baseline dataset covering thousands of relevant programs before new harvesting begins.

1.2 Incremental Data Harvesting

As Draup sets up the baseline data, Draup's harvester models will in parallel identify, extract, and normalize new program data from public sources, including:

- **Academic Institutions** – University, community college, and technical college websites with program & curriculum listings.
- **Vocational & Technical Schools** – Public pages from trade schools offering advanced manufacturing, robotics, and AI programs.
- **Robotics Manufacturers** – Training portals for industrial robotics vendors (e.g., FANUC, ABB, KUKA).
- **Apprenticeship Programs** – U.S. Department of Labor Apprenticeship Finder and state-level apprenticeship directories.

- **K–12 STEM Programs** – Public charter schools, STEM academies, and robotics clubs with formal training curricula.
- **Online Learning Platforms** – MOOCs and open course listings (Coursera, edX) where relevant to ARM frameworks.
- **Government & Nonprofit Directories** – Labor boards, economic development councils, NCES/IPEDS datasets.

1.3 Harvesting Process

Our proprietary web crawlers and scraping frameworks will extract structured program metadata from these sources, including:

- Provider and program names
- Descriptions and curriculum details
- Delivery mode (onsite, online, hybrid)
- Credential type (degree, certificate, other)
- Program length and tuition (if available)
- Location and contact details

The harvester pipeline is designed for repeatable execution to enable quarterly refreshes with minimal reconfiguration.

2. *Data Consolidation & Standardization*

The harvested data will undergo a comprehensive Extract, Transform, Load (ETL) mechanism which will involve translation of non-English text to standardized English language, classification of various attributes such as title, curriculum provider, authors etc. After the ETL process, the harvested data assets will be collated and grouped with Draup's existing data assets for a single comprehensive library of learning opportunities. All harvested and ETL data will be ingested into Draup's centralized staging environment, where the following operations occur:

1. **Schema Alignment** – All incoming records are mapped to a unified schema aligning to ARM's required data fields (Section 3.2 of RFP).
2. **Deduplication** – Our ML-based record-matching engine uses fuzzy matching on provider names, program titles, locations, and credentials to identify and merge duplicates.
3. **Super Record Creation** – Where the same program appears in multiple sources, data fields are merged to create a single "super record" with the most complete set of attributes.
4. **Attribute Normalization** – Standardizing values for:
    - Credential types (e.g., "B.S.", "Bachelor's" → "Bachelor's Degree")
    - Delivery formats ("Virtual", "Online" → "Online")

- Program lengths (mapped to ARM's categories: <3 months, <1 year, etc.

Tracking fields required for ARM

Draup will maintain an **ARM-specific data schema** from the outset of the project. All harvested program records are mapped to this schema, which contains every field specified in Section 3.2 of the RFP.

- **Schema-Level Controls** – The schema is built into our ingestion and ETL pipelines so that each required field (e.g., Training Provider Name, Program Description, Credential Earned, Competency Matches, CIP Code) has a defined data type, controlled vocabulary (where applicable), and completeness flag.
- **Automated Field Validation** – During ingestion, each record is checked for field presence, format accuracy (e.g., valid URLs, location taxonomy alignment), and standard taxonomy or field rules adherence (e.g., Program Format = Online, Onsite, Hybrid).
- **Update Tracking** – A "Program Status" field is maintained in the schema to capture quarterly changes as New, Updated, No Change, or Removed.

This structured approach ensures that every required data field is continuously tracked, validated, and ready for ARM's integration without the need for post-delivery reformatting. A detailed list of the fields required for ARM and how Draup plans to source this data is provided in the Appendix – Exhibit A

3. *Data Enrichment & ARM Framework Mapping*

Using Draup's **competency mapping models**, each program will be tagged to the specific **Robotics** and **AI** competencies defined by ARM, as well as relevant **soft skills**.

3.1 Draup Skills Tagging

Draup's Skills Mapping Methodology employs a structured, data-driven approach to map skills and competencies from diverse curricula, courses, and certifications into a robust taxonomy. The process begins with the **identification and harvesting** of relevant curriculum data from universities, certification libraries, online learning platforms, and specialized sources, ensuring comprehensive coverage across domains. This is followed by **big data-based content harvesting**, where large-scale data pipelines extract both structured and unstructured learning content for analysis.

Draup's proprietary machine learning models then **interpret language and context**, translating domain-specific terminology and regional variations into standardized skill representations. The extracted skills are **mapped to corresponding curriculum elements**, creating a direct linkage

between educational content and industry-demanded competencies. Skills are then **grouped based on similarity, domain, and application**, and further categorized into **core technical skills and soft skills** using Draup's Named Entity Recognition (NER) framework.

The mapped skills undergo a **multi-layered validation process**, starting with Subject Matter Expert (SME) reviews to ensure relevance, followed by Draup expert validation for data-driven accuracy. Finally, the **curriculum database is finalized** into a structured format that provides a detailed mapping of skills and competencies, enabling seamless integration into workforce development, learning & development (L&D) programs, and strategic talent planning initiatives.

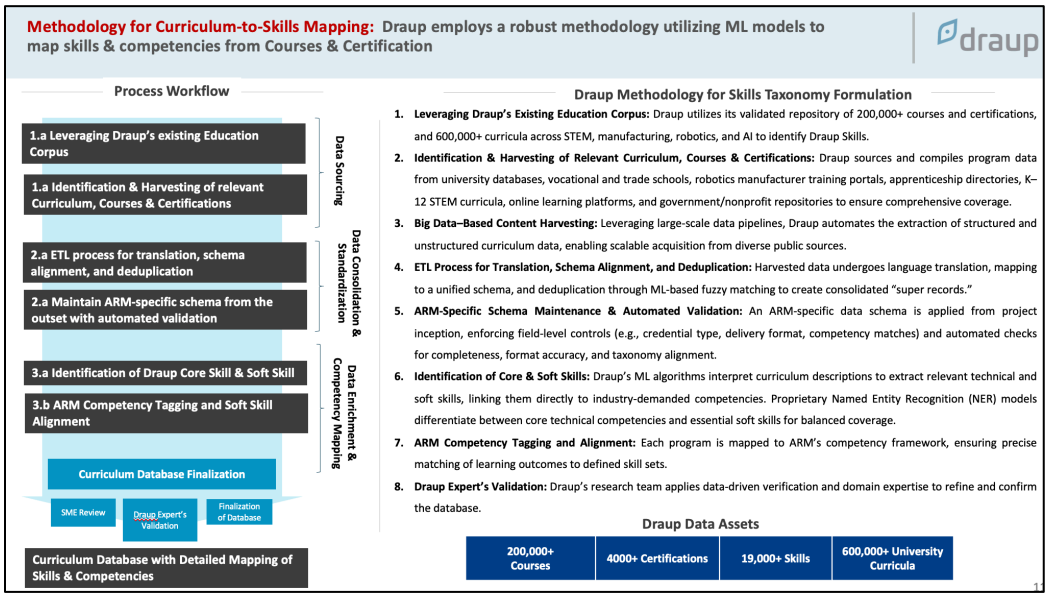

Fig.1 Draup Methodology for Curriculum-to-Skills Mapping

3.2 ARM Competency Tagging

- Upon completion of Draup's comprehensive skills identification process, wherein detailed **Draup Skills** are tagged to the courses & programs, categorized, and validated through its proprietary taxonomy, the next phase will focus on **Competency Tagging**—the alignment of these skills to **ARM's standardized competency framework**. In this stage, skills will be extracted from program descriptions, syllabi, and curriculum outlines, ensuring that each is mapped precisely to Draup's skills taxonomy for consistency and accuracy. This granular data will then be systematically mapped to ARM's established competency lists, which cover key occupational roles including **Robotics Technicians, Specialists, Integrators, AI Data Technicians, AI Implementation Specialists, and AI Developers**.
- This structured mapping process will ensure that Draup's detailed skills intelligence translates directly into ARM's competency architecture, creating a unified view of workforce capabilities across technical and functional domains. By linking program-level learning outcomes to ARM's role-specific competency standards, the Competency Tagging process will support precise gap analysis, targeted workforce development, and

curriculum alignment with industry needs. Ultimately, this integration will strengthen data interoperability between Draup's outputs and ARM's systems, enabling a strategic talent plan.

3.3 ARM Soft Skills Assignment

- Leveraging Draup's proprietary skills tagging approach (as outlined in Section 3.1), soft skills are extracted directly from program descriptions, syllabi, and curriculum outlines. Each identified skill is tagged in alignment with Draup's standardized skills framework to ensure consistency across datasets.
- Following Draup Skills Tagging, the Draup soft skills will be mapped to ARM's established soft skill taxonomy using a **Draup-to-ARM skills referencing framework**. This mapping will create a common language for behavioral competencies, enabling ARM to accurately integrate soft skill intelligence into its broader workforce capability model. The outcome will be a harmonized dataset that links both technical and non-technical competencies to ARM's role profiles, supporting balanced talent assessments, well-rounded workforce development strategies, and training programs that address both domain expertise and critical interpersonal skills.

3.3 Target Audience Classification

- Tag based on program-level indicators (e.g., "for high school students" → K-12 Students)

4. *Data Validation & Quality Control*

Our validation process includes **both automated and manual checks**:

4.1 Automated Validation Rules

- Field completeness (no empty required fields)
- URL verification (program links must resolve and return expected HTTP status)
- Location standardization using Draup's Locations Taxonomy Framework
- Credential/format validation against ARM's allowed values

4.2 Manual SME Review

- Human researchers will verify a sample from each source category for accuracy.
- Focus areas for manual review:
    - Correct competency/soft skill mapping
    - Tuition and program length accuracy
    - Proper classification of program type and target audience

5. *Data Delivery & Update Cycle*

5.1 Export Formats

- Delivered in PostgreSQL/MySQL dumps, CSV, or JSON, fully aligned to ARM's schema. Additional supported formats include Parquet, Avro, and TSV for system compatibility.
- All fields structured exactly to ARM's schema.

5.2 Program Status Tracking

- Each program marked as **New**, **Updated**, **No Change**, or **Removed** for each quarterly update cycle.

5.3 Quarterly Refresh Process

- Automated re-harvest from all public sources.
- Apply the same deduplication, enrichment, and validation workflow.
- Provide ARM with both full dataset and change log summary.

5.4 Architecture Overview of Draup Data Feed

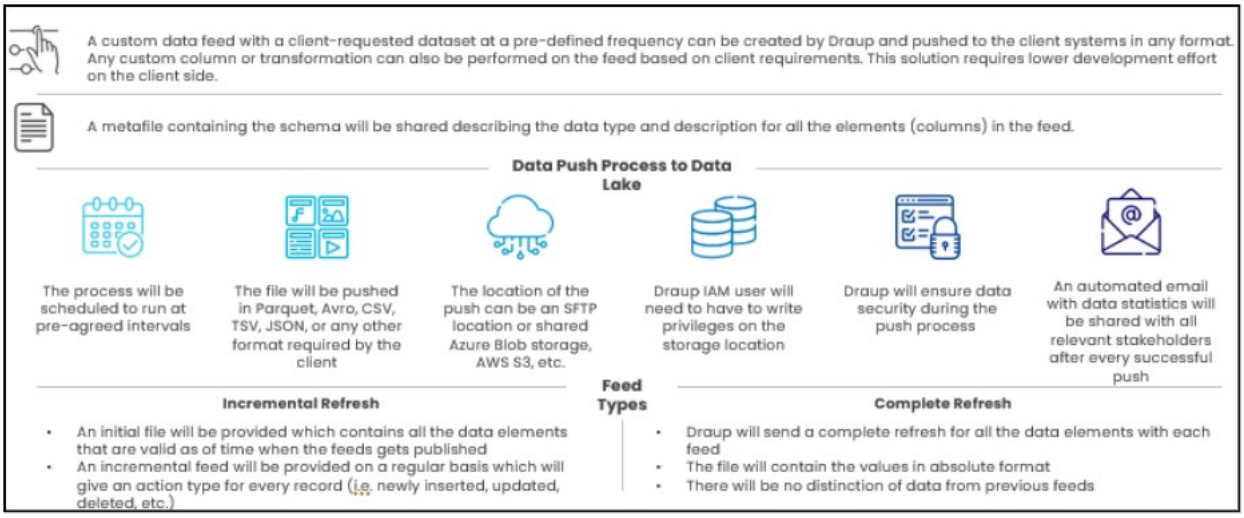

Fig 2: Draup Data Feed Integration

6. *Scalability & Future Framework Support*

- **Framework-Agnostic Design** – Mapping logic is configurable to accommodate 20+ new ARM frameworks without redesigning the pipeline.
- **Program Expansion** – Additional program types or industries (beyond Robotics/AI) can be onboarded by updating source list and competency mapping rules.

3. Technology Stack and tools

Our Harvester Framework delivers a comprehensive, enterprise-grade solution for automated web data collection and processing, capable of handling **billions of data points per month** and processing ~ **50 million requests per day**. The platform combines cutting-edge AI automation with robust engineering architecture to provide end-to-end data harvesting capabilities for various domains of data like Job descriptions, Profiles, Courses, Salary, Products, etc.

The solution is architected across three core components:

1. **Intelligent Configuration & Discovery** - AI-powered automation with human oversight for source discovery and configuration
2. **Robust Data Harvesting & Processing** - Enterprise-grade distributed architecture for scalable data extraction and validation
3. **Comprehensive Observability & Quality Assurance** - Advanced monitoring, analytics, and quality validation framework

Our platform leverages modern cloud-native technologies including LLM integration, distributed processing on AWS infrastructure, and advanced analytics capabilities. The system ensures reliable data delivery with **99.9 % uptime**, fault tolerance, scalability, and operational excellence.

Component 1: Intelligent Configuration & Discovery

Our intelligent automation layer leverages advanced AI and agentic workflows to automatically discover, analyze, and configure data sources with human supervision for critical validations. This component eliminates manual setup overhead by using LLM-powered analysis to understand website structures and automatically generating comprehensive URL sets for career pages, ATS systems, and job aggregators while incorporating manual supervision for critical websites and complex extraction strategies.

- **LLM-Powered Analysis with Human Oversight**: Automatically analyses website structures and content patterns using advanced NLP techniques, with manual supervision for critical website configurations and complex extraction strategies
- **Intelligent Web Discovery**: Performs automated web search and crawling to discover comprehensive URL sets, validates endpoints, and establishes source-specific access patterns with human review for mission-critical sources
- **Agentic Configuration Workflows**: Orchestrates end-to-end setup processes using autonomous agents for URL building, metadata extraction, schema generation, and configuration optimization with manual intervention capabilities

- **Input Metadata Generation**: Automatically extracts and structures website metadata including page parameters, filters, pagination logic, and content schemas with expert validation for complex architectures

Tech Stack: Temporal Workflow Engine, Apache Airflow, PostgreSQL, SERP, Playwright, LLM and NLP

Component 2: Robust Data Harvesting & Processing Architecture

Our enterprise-grade harvesting engine processes intelligent configurations through a distributed, fault-tolerant architecture designed for high-volume data extraction on AWS cloud infrastructure. The system orchestrates sophisticated data collection workflows across multiple priority queues with intelligent rate limiting, comprehensive proxy management, and advanced duplicate detection, ensuring all extracted data undergoes rigorous schema validation before MongoDB storage.

- **Producer-Consumer Architecture**: Input producers generate extraction tasks based on intelligent configurations while specialized workers process high-priority listing and standard detail extraction operations across distributed queues
- **Intelligent Distribution & Proxy Management**: Custom distribution service manages per-source rate limiting, priority-based task allocation, proxy rotation, and fair resource distribution to maximize throughput while preventing blocking
- **Advanced Extraction with Duplicate URL Detection**: Combines HTTP clients for standard requests with headless browsers for JavaScript-heavy content, includes anti-bot protection through proxy systems, and implements intelligent duplicate URL checker
- **Data Validation & Storage**: Multi-layered data validation against generated schemas, comprehensive duplicate URL detection across sources, and structured persistence in MongoDB with automated quality assurance

Tech Stack: Celery, Redis, Playwright, FlareSolverr, Scrapoxy, MongoDB, OpenSearch, AWS, EKS

Component 3: Comprehensive Observability & Quality Assurance

Our operational excellence layer provides complete visibility into system performance, data quality, and business metrics through advanced monitoring and validation frameworks. This component ensures reliable operations through real-time performance tracking, automated data coverage analysis, comprehensive quality validation, and intuitive dashboards using Superset with Trino for actionable insights into harvesting effectiveness across all configured sources.

- **Real-Time Performance Monitoring**: Tracks system health, processing rates, success ratios, and resource utilization across all harvesting components with custom metrics for queue performance, worker efficiency, and AWS infrastructure
- **Data Coverage & Quality Validation**: Automated analysis of data completeness, schema compliance, freshness metrics, and coverage gaps with intelligent alerts for data quality issues and extraction anomalies
- **Advanced Dashboarding & Visualization**: Interactive dashboards using Superset powered by Trino for high-performance analytics, displaying harvesting metrics, source performance comparisons, and data volume trends with drill-down capabilities
- **Centralized Logging & Alerting**: Comprehensive log aggregation from all system components with intelligent alerting for failures, performance degradation, and data quality issues, plus configurable notification rules

Tech Stack: Prometheus, Grafana, Superset, Trino, Elasticsearch, Alert Management Systems, FastAPI, AWS EKS, Locust (Performance Testing)

4. Project Timeline, including major milestones.

Project Start Date: Project kick-off immediately after issuance of PO.

Total Duration: ~12 weeks for initial delivery, followed by ongoing quarterly refresh cycles.

Steps	Timeline	Milestones	Deliverables
Step 1 – Project Initiation & Planning	Weeks 1–3	1: Project kick-off meeting with ARM Institute 2: Confirm data schema based on ARM's required fields 3: Finalize public source list and validation rules 4: Configure reference layer integration from Draup's 200K+ courses & 40K+ curricula for gap analysis and competency mapping alignment	Finalized project plan, schema, and source list Source-to-field mapping document
Step 2 – Initial Data Sourcing	Weeks 4–6	1: Reference layer processed for ARM-specific filtering & mapping 2: Public data harvesting begins from prioritized source list (academic, vocational, manufacturers, apprenticeship, K–12, government datasets) 3: Weekly ingestion and schema alignment in staging environment 4. Initial deduplication and normalization completed on first large batch of harvested records	Interim sourcing progress report with coverage stats and gap analysis findings
Step 3 – Data Consolidation & Enrichment	Weeks 6–8	1: Merge harvested public data with enriched metadata from reference layer (no direct carry-over without validation) 2: Competency & soft-skill mapping applied to all consolidated records 3: Normalization of credentials, formats, lengths, and location fields completed	Enriched and mapped dataset in staging environment Competency mapping summary document
Step 4 – Data Validation & Quality Control	Weeks 8–10	1: Automated validation checks completed across full dataset 2: Manual SME review of targeted samples from each source category 3: All identified corrections applied and re-validated	Validation report (coverage %, error rates, sample findings) Final clean dataset in staging, ready for export

Step 5 – Delivery & ARM Handover	Weeks 10–12	1: Final dataset export in ARM-compatible formats (PostgreSQL/MySQL, CSV, JSON) 2: Documentation package delivered (data schema, source list, methodology, validation summary) 3: Training session for ARM Institute team on data structure and update process	Final ARM-ready dataset Full documentation package
Step 6 – Ongoing Quarterly/Monthly/Weekly Updates	Post-Go-Live	1: Automated public harvesting refresh run quarterly 2: Deduplication, enrichment, and validation applied each cycle 3: Quarterly change log delivered (New, Updated, No Change, Removed)	Quarterly refreshed dataset in agreed format Change log and coverage report

5. Customer Service / Maintenance Plan

Customer Service

Draup proposes the following multi-layer staffing plan, with all key personnel committed for the duration of the onboard and active partnership period.

These key personnel will be supported by multiple resources across Draup teams (support, SME, data analysts, engineering). Our team members will be available during ARM's business hours and accessible via video conferencing, emails, phone or in-person meetings as needed.

1. **Customer Success Manager (CSM)**: ARM will be assigned a named Customer Success Manager who acts as the strategic point of contact and the first level for any escalations. The CSM ensures alignment with your goals, needs, helps in driving product adoption, shares best practices, ensures timely issue resolution and facilitates Quarterly Business Reviews (QBRs) to track value for your investments.
2. **Delivery Lead (DL)**: A dedicated, named Delivery Lead will oversee all aspects of execution and deliverables, ensuring optimal platform performance, high data quality, and the timely rollout of features and enhancements. The DL will also deliver custom reports in alignment with the defined scope by ARM stakeholders. The Delivery Lead (DL) will collaborate closely with the Customer Success Manager (CSM), Customer Support Teams and Customer Teams (ARM users and key stakeholders) to proactively resolve delivery blockers, manage ARM's evolving priorities, and ensure seamless project coordination.
3. **Customer Support Team (CST)**: The Customer Support Team will serve as the primary point of contact for addressing ARM's inquiries, troubleshooting issues, tracking support tickets, and providing timely resolutions to ensure seamless user & delivery experience. Operating in close coordination with the Customer Success Manager (CSM), and Delivery Lead (DL), the CST team will manage support tickets, track recurring issues, and escalate critical matters when necessary. Leveraging robust knowledge bases and best practices, the CST team will not only resolve technical and functional challenges but also proactively share insights to enhance platform performance, user satisfaction, and overall customer success.

   Draup provides technical & functional to customers. Functional support addresses specific queries related to Draup data, taxonomy, data formats, data updates, and Data/ML models and analysis. Technical support focuses on data support and delivery queries within the defined project scope, ensuring smooth and reliable access to the required datasets and deliverables.

We provide 24x5 multi-channel support via an integrated ticketing platform, email, and phone. Our support team is distributed across USA and India, enabling continuous coverage with SLA driven response times during business days.

Service Maintenance:

Draup maintains a comprehensive maintenance program to ensure the availability, accuracy, reliability, and timeliness of the data we are committing to deliver to ARM.

Our delivery and support teams will work in close collaboration with ARM team to ensure uninterrupted data service. In the event of planned maintenance, ARM team will be notified in advance, with clear timelines and contingency measures to prevent service or delivery disruption.

Our service maintenance plan includes:

1. **Continuous Data & Quality Monitoring**
   Draup will perform comprehensive data integrity checks against multiple data sources leveraged by our platform. These checks will combine automated and manual validation processes to detect and address data anomalies, inconsistencies, and obsolete records, ensuring the highest standards of accuracy and reliability.
2. **Regular Data Updates**
   Draup maintains scheduled refresh cycles for each data attribute based on relevance, source refresh cycles and priority. High-priority data sets are updated in real-time or near-real-time to ensure maximum accuracy and relevance.
3. **Enhancements**
   Draup deploys new capabilities and data improvements without disrupting ongoing data delivery, ensuring continuity of service while maintaining transparent communication with customers through release notes, impact assessments and value of enhancements.
4. **Customer Feedback**
   Draup will maintain direct feedback loops with ARM team (users and key stakeholders) to identify evolving requirements and rapidly incorporate corrections or enhancements based on ARM team's inputs.

6. Cost Proposal

Sr. No.	Item & Description	Duration	Proposed Pricing (USD)
1	**Initial Data Infrastructure & Process Setup + First Data Feed** Deployment of data harvesting infrastructure, acquisition of public program data, enrichment and classification aligned to ARM's competency frameworks, and secure delivery of the validated dataset in required formats.	Q1 & Q2	$100,000
2	**Quarterly Dataset Refreshes** Quarterly re-run of the processes from new data acquisition, enrichment, classification & delivery	Q3	$40,000
		Q4	$40,000
	Total Annual Cost in USD		**$180,000**

7. Sample Outputs

Draup transforms the “Programs Data” into a structured intelligence framework by mapping each program against eight key dimensions: Provider Reach, Program Focus, Delivery Flexibility, Cost Accessibility, Duration Intensity, Target Demographics, Geographic Spread, and Credential Level. Using a blend of automated natural language processing and manual verification, competencies, delivery modes, and audience types are consistently tagged and aligned with models such as the ARM Institute’s Robotics and AI Competency Framework. The data is then quantified into percentages, ranges, and coverage counts, revealing patterns such as the share of online programs, tuition distribution, and state-level coverage. This approach converts raw listings into a decision-ready view, enabling stakeholders to benchmark providers, identify gaps, and target program development with precision.

Metric Category	**Example Metric**	**How It’s Measured**	**Example From Data**
Provider Reach	Number of Programs per Provider Type	Count grouped by *Training Provider Type*	Community College – 4
Program Focus	Top Competency Areas	Frequency of keywords in *Competency Matches*	AI Models, Robotics
Delivery Flexibility	% Programs Offered Online	Programs with *Program Format* = “Online” ÷ Total	38%
Cost Accessibility	Tuition Range	Min & Max of *Total Tuition/Program Cost*	$29 – $12,360
Duration Intensity	Average Program Length	Mean of *Program Length* (converted to months)	~20 months
Target Demographics	Representation of Underrepresented Groups	% of programs listing keywords in *Target Audience*	64%
Geographic Spread	Coverage by State	Count of unique states from *Program Location*	21 states + Online

Metric Category	Example Metric	How It's Measured	Example From Data
Credential Level	Proportion of Degree vs. Certificate vs. Course Programs	Group by *Program Type*	Degree: 54%, Courses: 30%, Cert: 16%

Training Provider Name	Program Name/Title	Program Format	Total Tuition/Program Cost
Baltimore City Community College	Associate of Applied Science in Robotics/Mechatronics Technology	On-Site	$9,797
Garden City Community College	Associate in Robotics and Mechatronics Technology	On-Site	$12,360
Houston Community College	Bachelor of Applied Technology in Artificial Intelligence & Robotics	On-Site	$5,216
AWS Skill Builder	AWS SimuLearn: Robotics Data Visualization	Online	$29
Canvas Network	XRP 101: Introduction to the Experiential Robotics Platform	Online	nan
Coursera	Modern Robotics, Course 2: Robot Kinematics	Online	nan
edX	Robotics: Vision Intelligence and Machine Learning	Online	nan
FutureLearn	Begin Robotics	Online	$79
Oreilly	Hands-On ROS for Robotics Programming	Online	nan
Oreilly	Hands-On Robotics Programming with C++	Online	nan
Swayam	Introduction to robotics	Online	nan
Udacity	Robotics Software Engineer	Online	nan
Udacity	CS 8802, Artificial Intelligence for Robotics: Programming a Robotic Car	Online	nan
Udemy	Introduction to Robotics & Artificial Intelligence	Online	479
Udemy	Artificial Intelligence: Reinforcement Learning in Python	Online	$37

Training Provider Name	Program Name/Title	Program Format	Total Tuition/Program Cost
Bristol Community College	Associate of Applied Science (AAS) in Artificial Intelligence	Online	$9,117
Houston Community College	Bachelor of Applied Technology in Artificial Intelligence & Robotics	On-Site	$5,216
edX	Robotics: Vision Intelligence and Machine Learning	Online	nan
Udacity	CS 8802, Artificial Intelligence for Robotics: Programming a Robotic Car	Online	nan
Udemy	Introduction to Robotics & Artificial Intelligence	Online	479
Udemy	Artificial Intelligence - Easily Explained For Beginners	Online	529
Udemy	Artificial Intelligence: Genetic Machine Learning Algorithms	Online	479
Udemy	Artificial Intelligence: Reinforcement Learning in Python	Online	$37
Udemy	Artificial Intelligence in Bioinformatics	Online	nan
Udemy	Artificial Intelligence - Full course with Deep learning	Online	nan
Carnegie Mellon University	Master of Science in Robotics (MSR)	OnnullSite	$60,400
Carnegie Mellon University	PhD in Robotics	On-Site	$51,250
Arizona State University-Tempe	Master of Science in Artificial Intelligence Engineering (Robotics)	On-Site/Online	$33,720
Arizona State University-Tempe	Master of Science in Robotics and Autonomous Systems (Artificial Intelligence)	On-Site/Online	$37,044
Pennsylvania Western University	Certificate Program in Robotics and Intelligent Systems	On-Site	$11,117

Fig.3 Sample outputs of Courses Identification & Skills Mapping

Sample Record in JSON format


```
{
  "training_status_code": 1,
  "training_provider": {
    "id": "provider_bccc_001",
    "organization_name": "Baltimore City Community College",
    "description": "BCCC is a state-sponsored, comprehensive, degree-granting community colleg
    "type": "Community College",
    "industry": ["Education"]
  },
  "program": {
    "id": "program_bccc_robotics_001",
    "name": "Associate of Applied Science in Robotics/Mechatronics Technology",
    "url": "https://www.bccc.edu/robotics",
    "format": "On-Site",
    "type": "Degree",
    "length": "2 Years",
    "tuition": 9797.0,
    "credential_earned": [
      "Associate of Applied Science in Robotics/Mechatronics Technology"
    ],
    "cipc_code": "-",
    "location": {
      "state": "Maryland",
      "city": "Baltimore"
    },
    "contact": {
      "name": "Michael Kaye",
      "email": "mkaye@bccc.edu",
      "phone": "410-462-8401",
      "title": "Faculty"
    },
    "competencies": [
      "Robot Programming",
      "PLC (Programmable Logic Controller)",
      "Robot / System Troubleshooting",
      "Electronics C Controls",
      "Mechanical Systems"
    ],
    "skills": [
      "Critical Thinking",
      "Problem Solving",
      "Technical Learning Ability",
      "Attention to Detail",
      "Adaptability"
    ],
    "target_audience": [
      "Adult",
      "Minorities",
      "Underrepresented Populations",
      "Women"
    ],
    "target_audience_type": ["Learner"]
  }
}
```


Relevant Draup Customer Case Studies:

Case Study 1 – Major National Technology Industry Association

A leading national trade association in the technology sector engaged Draup to create National Operating Standards (NOS) covering competencies and skills for all IT occupations. Draup developed role-specific NOS frameworks, mapped relevant learning curricula to each, and supported the client in certifying curriculum providers. This ensured the national tech workforce had access to standardized, industry-aligned training backed by a quality-assured provider network.

Case Study 2 – Global Retail Organization

A Fortune 100 retail enterprise selected Draup to curate Massive Open Online Courses (MOOCs) for its corporate learning academy. Draup conducted a detailed skills gap analysis, sourced and validated content from global providers, and ensured alignment with the organization's competency models. The curated learning library enabled targeted, scalable upskilling programs across diverse roles, geographies, and skill levels.

Case Study 3 – Global Professional Services Firm

A top-tier global professional services firm collaborated with Draup to certify learning content for its internal competency framework. The project involved mapping competencies to vet internal and external resources, ensuring quality, relevance, and alignment with strategic priorities. This competency-to-content mapping streamlined global L&D initiatives and delivered consistent, high-quality learning experiences to the workforce.

8. References

1. **Johnson Controls, JCI** – Skills Assessment and Reskilling Assets
   Contact Name: Rajesh Ahuja
   Email: rajesh.ahuja@jci.com

2. **NASSCOM**: Developed national competency standards for IT occupations, mapped aligned learning curricula, and certified training providers.
   Contact Name: Dr. Sandhya Chintala
   Email: sandhya@nasscom.in

3. **Intel**: Upskilling Strategies to transition Technicians to Foundry Talent Pool in United States.
   Contact Name: Vibhu Ganesan
   Email: vibhu.ganesan@intel.com

4. **NASDAQ**: Developed a comprehensive Career Portal for 10k+ Nasdaq Employees to identify viable career paths based on Draup's Reskilling Propensity Index framework
   Contact Name: Karoline Johanna Raets
   Email: karoline.raets@nasdaq.com

9. Litigation

Draup confirms that, since our incorporation in 2021, we have not been involved in any litigation. Furthermore, we are not aware of any pending or threatened legal proceedings that could materially impact on our business operations, financial position, or our ability to fulfill the obligations and deliverables outlined in this proposal.

We maintain robust ethical, legal and compliance frameworks designed to proactively identify, assess, and address potential legal risks before they escalate into material issues. These include regular reviews by our internal legal team, contractual compliance monitoring, and periodic engagement with external legal counsel to ensure our activities remain within applicable laws and regulations (across the globe).

We recognize that our customers rely on Draup not only for the quality of our solutions and services but also for the security and integrity of our operations. We are committed to maintaining transparency and will promptly communicate any future developments, including the emergence of litigation or legal proceedings that could reasonably be expected to affect service delivery or contractual commitments as defined in this proposal. Such communication would include an assessment of potential impacts, and a clear outline of mitigation measures being implemented to ensure uninterrupted service and fulfillment of our commitments to ARM.

Appendix

Exhibit A:

RFP Field	Draup Approach to Source & Process Fields
Training Provider Name	Extracted directly from provider's official program pages; cross-validated with NCES/IPEDS and other authoritative education directories wherever relevant.
Training Provider Description	Parsed from provider's "About" or program overview pages; summarized for consistency.
Training Provider Type	Categorized using ARM's defined taxonomy; assigned via rule-based mapping from provider metadata.
Training Provider Industry	Tagged using Draup's Vertical classification model; validated against provider's stated specialization.
Training Provider Location (Address, City, State, Zip-Code)	Extracted from provider contact pages; standardized to Draup Location Taxonomy
Program Name/Title	Captured exactly as listed by the provider; normalized for capitalization and special characters.
Program Description	Extracted from official program listings; cleaned and standardized.
Program Location (City, State, Online/Onsite)	Derived from program listing; "Online" or "Hybrid" flags applied based on delivery description.
Program Contact (Name, Email, Phone)	Harvested from program info pages; validated with email/phone format checks.
Program Contact Title	Extracted directly from contact listing
Credential Earned	Parsed from program description; normalized to ARM's credential categories (degree, certificate, other). Draup can work with ARM to curate & refine categorizations.
Program URL	Captured from official provider page; verified for validity (HTTP 200 status).
Tuition/Program Cost	Extracted where available; "N/A" if not publicly listed.
Program Format (Onsite, Online, Hybrid)	Assigned from program listing text; standardized to ARM's allowed values.
Program Type (Degree, Certificate/Non-Degree, Other)	Classified according to credential earned and program structure.
Program Length (<3 months, <1 year, etc.)	Derived from program description; normalized to ARM's duration categories.
Competency Matches	Mapped using Draup's skills mapping engine against ARM's competency lists.
Soft Skill Matches	Mapped from course descriptions to Draup's soft skill taxonomy & aligned to ARM taxonomy as per a matching layer.
Target Audience	Identified from program marketing text; assigned tags like K–12 Students, Adult Learners, Veterans.
CIP Code	Assigned from NCES Classification of Instructional Programs database where applicable.
Program Status	Determined via quarterly comparison: New, Updated, No Change, or Removed.