

SIH 2026 — Shortlisted Problem Statements

Prepared for Om | PCCOE, Pune | Matched to: GenAI / RAG, Computer Vision, Flutter & Android, Go

1. [26101] AI-Enabled Skill Intelligence & Learning Platform for India's Official Statistical System

Organization: MoSPI (Ministry of Statistics & Programme Implementation) | **Category:** Software | **Theme:** Smart Education

Background

India's statistical system is adopting AI, ML, big data, and cloud tools faster than its officials can be trained on them. The government's iGOT Karmayogi platform already hosts a large library of courses, but there is no intelligent way to match an official's actual skill gaps to the right courses — recommendations aren't personalized, and there's no fast way to turn existing training material into assessments.

What has to be built

- Auto-build a competency profile per official from designation, department, experience, and past training.
- Map competencies across statistical, technical (Python/R/SQL/AI-ML/Cloud), digital-governance, and managerial domains.
- Use NLP / LLMs / semantic search to recommend personalized learning paths from the iGOT Karmayogi catalogue.
- Integrate with iGOT Karmayogi APIs to pull courses and track completion.
- AI-powered assessment engine that generates MCQs and quizzes automatically from uploaded learning material (docs, slides, videos), with instant grading and explanations.
- Learner and admin analytics dashboards — progress, skill-gap heatmaps, predictive future skill needs.

Dataset provided: nssta.gov.in, mospi.gov.in

Why this fits you: This is a RAG + LLM pipeline end to end — ingest documents, embed them, retrieve, and generate quizzes/recommendations — exactly the LangChain/LlamaIndex stack you already work in. MoSPI is a smaller, less-contested ministry track compared to MoD/ISRO/DRDO, so competition for a selection slot is lower, and the scope is clean enough to fully demo in the hackathon window.

2. [26088] Multilingual Cooperative Governance & Legal Assistance Chatbot

Organization: Ministry of Cooperation | **Category:** Software + Hardware | **Theme:** Agriculture, FoodTech & Rural Development

Problem

Cooperative members, farmers, and rural stakeholders don't know their rights and options under cooperative law — PACS services, crop insurance (PMFBY), financial literacy schemes, grievance redressal — mainly because of language barriers and no easy way to ask questions in their own dialect.

What has to be built

- Multilingual conversational chatbot (text + voice) covering cooperative laws, by-laws, and Ministry schemes.
- PMFBY / crop insurance guidance and financial literacy assistance.
- Grievance redressal support built into the conversation flow.
- Speech-to-Text and Text-to-Speech for rural, low-literacy users.

- Mobile and web deployment.

Why this fits you: A RAG chatbot over scheme/legal documents, wrapped in a Flutter app — you've already shipped a mobile app, so the front end is a repeat of skills you have, not new ground. It's a rural-impact story judges respond well to, and the build is realistic in the time you'll actually have.

3. [26165] AI/NLP Engine to Detect Serious Injury & Fatality (SIF) Precursors in Safety Reports

Organization: Oil India Limited (OIL) | **Category:** Software | **Theme:** Smart Automation

Background

OIL collects large volumes of free-text unsafe-act / unsafe-condition and near-miss reports through its HSSE platform, but they're triaged manually on a monthly/quarterly cycle. Industry data shows minor incidents and fatalities don't share the same root causes — so the ~20-25% of reports carrying real fatal potential need to be flagged separately and fast, not buried in a quarterly review.

What has to be built

- Classify each incoming report as SIF-potential vs non-SIF-potential.
- Tag each report to the relevant IOGP Life-Saving Rule (Energy Isolation, Hot Work, Confined Space, Line of Fire, etc.).
- Surface recurring precursor patterns — activity, location, barrier failure — on an interactive dashboard.
- Rank sites/activities by SIF-precursor density so HSE teams know where to intervene first.

Dataset provided: OIL's actual UA/UC, near-miss, and incident reports.

Why this fits you: A well-scoped text-classification + pattern-mining problem with real data handed to you — no scraping or synthetic data needed. PSU statements like this historically draw far fewer strong teams than the flashy AI-for-defence ones, and the deliverable (classifier + dashboard) is realistic to finish and demo well.

4. [26177] Autonomous AI Drone for Search & Rescue — People and Hazard Detection

Organization: Qualcomm Inc | **Category:** Hardware | **Theme:** Robotics and Drones

Background

India is highly disaster-prone (floods, cyclones, earthquakes, landslides), and the first few hours after a disaster are critical for finding survivors — but ground assessment is slow and dangerous in damaged terrain. An autonomous drone with on-device AI can give responders real-time aerial intelligence, even with no network connectivity.

What has to be built

- Autonomous navigation (GPS and GPS-denied) with obstacle avoidance for damaged environments.
- On-device AI inference — detect people/survivors and hazards without relying on the cloud.
- Multi-sensor fusion: RGB + thermal cameras, IMU, GPS for accurate victim localization.
- Hazard classification — fire, floodwater, debris, unstable structures, landslide zones, exposed electrical lines.
- Geo-tagged live disaster maps showing survivor locations, hazards, and safe access routes.
- Command-center dashboard for responders with automatic alerting and rescue prioritization.

Why this fits you: This is a direct extension of your ArUco/OpenCV/Raspberry-Pi navigation and CV work on the hotel delivery robot — same core skills (real-time detection + navigation on constrained hardware), just a new sensor set and domain. Industry-sponsored statements like Qualcomm's also tend to have clearer evaluation criteria and better mentor access. If a real drone build isn't feasible in the timeframe, the detection/mapping stack can be demoed on a ground rover instead — that's the part judges are actually evaluating.

5. [26097] AI Voice Assistant for Livelihood Mapping & Skilling Recommendations (PM-AJAY, SC Communities)

Organization: Ministry of Social Justice & Empowerment (MoSJE) | **Category:** Software | **Theme:** Agriculture, FoodTech & Rural Development

Background

Under PM-AJAY's Grant-in-Aid component, Scheduled Caste beneficiaries are supposed to be matched to skill-training programs that fit their aspirations and local job market — but low digital literacy, language barriers, and text-heavy forms cause mismatched enrollments and high training dropout.

What has to be built

- Multilingual, voice-based virtual assistant that conducts a conversational interview instead of a form.
- Collects education, family occupation, current livelihood, skills/interests, mobility constraints, employment preference.
- AI/ML profiling engine that recommends NSQF-aligned training programs and region-specific livelihood opportunities.
- Must work in low-connectivity, low-tech environments — IVR phone calls, WhatsApp voice notes, lightweight mobile/kiosk apps.

Why this fits you: Conversational AI + recommendation logic in regional languages — a dialogue system on top of an LLM, similar architecture to the MoSPI and Cooperation picks above but centered on voice-first, low-connectivity delivery. MoSJE is an underserved ministry track, and the social-impact framing (SC community upliftment) tends to score well with judging panels.

Note: Details above are condensed from the official SIH 2026 Problem Statement Directory (229 statements total). Verify exact wording, evaluation parameters, and datasets against the official portal before final submission, as ministries occasionally revise statement text after publication.