



INTER IIT  
TECH MEET 14.0

NO PREP



 Genuity

Algorithmic Optimisation

# Problem Statement Description

The PS reflects real problems encountered while building scalable AI systems where synthetic data generation, data structure optimization, autonomous knowledge systems, and LLM-driven decision pipelines converge to enable next-generation data intelligence.

The questions are structured as a series of compact coding tasks each intended to be solved at a low level, yet each representing a miniature version of a real component found in modern big data and synthetic data ecosystems. These tasks span multiple interconnected domains, including:

- **Data Structure Transformations & Optimization:**  
Efficiently converting and optimizing complex data representations (graphs, hierarchical JSON, sparse matrices, distributed logs), mirroring the challenges of high-volume data engineering.
- **Loss Engineering & Structural Optimization:**  
Designing and refining custom loss functions with topological, geometric, or structural constraints; focusing on computational efficiency and gradient stability.
- **Autonomous Knowledge Graph Construction:**  
Building small components for entity extraction, relation formation, graph augmentation, and incremental knowledge updatescore primitives for autonomous retrieval and reasoning systems.
- **LLM-Based Decision and Control Modules:**  
Implementing scoring, ranking, filtering, and lightweight state management utilities that enable LLMs to behave as structured decision-making agents.

- **Synthetic Data & Privacy-Aware Functions:**

Crafting controlled noise injections, distribution-matching routines, and privacy-preserving transformations used in synthetic data pipelines.

- **Micro-Pipelines & Utility Components:**

Writing helper utilities for hashing, batching, caching, validation, and structural checks mirroring the modularity required in real AI and big-data workflows.

## Format

The competition consists of 4 rounds. Each round will last 3-4 hours.

**Round 1:** All registered teams will participate.

**Round 2:** All registered teams will participate.

**Round 3:** The top 12 teams will participate.

**Round 4:** The top 4 teams will participate.

**Note:**

- Use of LLMs is allowed, provided the output reflects genuine reasoning and does not replace analysis or understanding.
- Difficulty increases progressively across rounds.

## Prerequisite Knowledge

Participants are expected to have basic familiarity with Topological Data Analysis (TDA), as one of the problems may involve concepts such as:

1. Simplicial complexes
2. Persistent homology
3. Extracting structure from point clouds
4. Geometric/topological insights into high-dimensional data

A conceptual understanding is sufficient; advanced theory is not required