

# GroundTruth Mini AI Hackathon

## Problem Statements

Welcome to the GroundTruth Mini AI Hackathon. Below are the three challenge tracks designed to bridge the gap between digital data and physical world applications.

### H-001 | The Automated Insight Engine

Track: Data Engineering & Analytics

#### 1. Context & Problem

In the AdTech world, we generate terabytes of data daily—foot traffic logs, ad clickstreams, and weather reports. Currently, our Account Managers manually download CSVs and take screenshots to build "Weekly Performance Reports" for clients. This is slow, error-prone, and boring.

**The Vision:** Design a system that ingests raw data from different sources (structured and unstructured), combines them, and automatically generates beautiful, executive-ready PDF/Slide decks with key insights highlighted in natural language.

#### 2. The Challenge

Build an automated pipeline that:

- **Ingests Data:** Accepts data from different sources (CSV, SQL, databases) and transforms it.
- **Reports:** Generates and Exports a formatted PDF or PowerPoint without human intervention.

#### 3. Good to have Technical Requirements

To qualify for judging, your solution can use the following preferred or similar tools:

- **Backend:** Python (Pandas/Polars) can be used for data processing.
- **AI Integration:** You can use an LLM (GPT-4o or Gemini) to generate the textual analysis, not just charts.
- **Output:** The final artifact must be a downloadable file (PDF/PPTX), not just a web dashboard.

### H-002 | Customer Experience Automation

Track: Customer Experience & Conversational AI

#### 1. Context & Problem

Retail customers today expect instant answers. "Is this store open?" "Do you have size 10 in stock?" "Where is my order?" Standard chatbots are dumb—they give generic answers.

## 2. The Challenge

Build a "Hyper-Personalized" Customer Support Agent.

- **Scenario:** A user is standing outside a coffee shop. They text the bot.
- **The AI's Job:** The AI should know the customer's history and context to give a specific answer.
- **Example:** User: "I'm cold." → Bot: "Come inside the Starbucks 50m away. We have a 10% coupon for Hot Cocoa."

## 3. Good to have Technical Requirements

- **Architecture:** Can implement a RAG (Retrieval Augmented Generation) pipeline to fetch specific (PDFs) customer data.
- **Privacy:** No raw user data sent to public LLMs (like chatgpt/gemini). You must mask customer sensitive data like phone numbers before processing.

## H-003 | The AI Creative Studio

Track: Generative AI & Marketing Tech

### 1. Context & Problem

Businesses spend weeks designing variations of the same image content. Design a system to automatically generate variations of given brand creatives/images.

### 2. The Challenge

Leverage Generative AI to build an "Auto-Creative Engine."

- **Input:** A brand's logo, a brand's product image.
- **Process:** The system automatically generates 10+ ad creatives variations.
- **Output:** A zip file containing high-resolution images and matching captions.

### 3. Good to have Technical Requirements

- **Image Generation:** Can use an API like Stable Diffusion, Midjourney, or DALL-E 3.
- **Text Generation:** Ad Content must be generated by an LLM.

---

*GroundTruth Hackathon 2025*