TRACK 1 — GENERATIVE AI

Problem Statements

HTH-GA-01: Grounded Multi-Document Compliance Assistant

Category	Details
Industry	Legal / Corporate Governance
Scenario	A mid-size company's legal team manually cross-checks new contracts against internal policy manuals and regulatory guidelines a process that takes days and is error-prone.
Problem Statement	Build a GenAI system that ingests policy documents and a new contract, flags clauses that conflict with or are missing from policy, and cites the exact source passage for every flag.
What Teams Should Build	RAG pipeline over multiple PDFs with per-claim source attribution, a conflict-detection layer, and a confidence indicator distinguishing "explicitly stated," "inferred," and "not found."
24-Hour MVP	Ingest 3–5 sample policy PDFs + 1 sample contract (synthetic), produce a clause-by-clause conflict report with citations.
Data/Input	Publicly available sample contracts + synthetic policy documents (teams write 3–5 policy pages).
Key Constraints	Must not hallucinate citations; must explicitly say "not found" when policy is silent; limited context window forces smart chunking.
Expected Output	Structured conflict report (clause, policy reference, risk level, explanation).
Evaluation Criteria	Citation accuracy, hallucination rate, usefulness of flags, UI clarity.
Bonus Features	Suggested redline language; risk-severity scoring.

HTH-GA-02: Adaptive Corporate Onboarding & Knowledge Coach

Category	Details
Industry	HR / Enterprise Learning
Scenario	New hires take 4–6 weeks to become productive because internal wikis, SOPs, and tool docs are scattered and outdated.
Problem Statement	Build a GenAI onboarding coach that ingests internal docs (simulated), tracks what a new hire has learned via quiz responses, and adapts explanations/next topics based on demonstrated gaps.
What Teams Should Build	Document ingestion + retrieval, adaptive quiz generation, weak-area tracker, personalized daily learning plan generator.
24-Hour MVP	10–15 synthetic SOP docs, quiz generation from docs, adaptive re-ranking of next topics based on quiz score.
Data/Input	Synthetic company SOPs/onboarding docs authored by the team.
Key Constraints	Must not generate quiz questions not grounded in the docs; adaptation logic must be explainable, not a black box.
Expected Output	Personalized learning path dashboard + progress report.
Evaluation Criteria	Grounding accuracy, adaptivity quality, demo clarity.
Bonus Features	Manager-facing readiness report; multi-role onboarding tracks.

HTH-GA-03: AI Meeting-to-Accountability System

Category	Details
Industry	Enterprise Operations
Scenario	Teams lose track of commitments made in meetings; action items are forgotten or misattributed.
Problem Statement	Convert meeting transcripts into structured decisions, action items, owners, deadlines, and unresolved issues — without inventing content not discussed — and track follow-through across multiple meetings.
What Teams Should Build	Transcript parser, structured extraction (decision/owner/deadline), cross-meeting action-item tracker that flags overdue/repeated items.
24-Hour MVP	Process 3–5 synthetic meeting transcripts, extract structured items, show a tracker dashboard flagging items unresolved across meetings.
Data/Input	Synthetic/generated meeting transcripts (teams can script realistic dialogues or use public sample transcripts).
Key Constraints	Zero-hallucination requirement on action items; must handle ambiguous ownership language.
Expected Output	Action-item tracker with status (new/carried-over/overdue) and audit trail to source transcript line.
Evaluation Criteria	Extraction precision/recall, hallucination avoidance, tracker usefulness.
Bonus Features	Slack/email digest generation; sentiment/tension detection in discussions.

HTH-GA-04: Explainable Candidate-Role Fit Engine

Category	Details
Industry	HR Tech / Recruitment
Scenario	Recruiters screen hundreds of resumes per role and struggle to justify shortlist/reject decisions defensibly (bias/audit concerns).
Problem Statement	Build a GenAI system that scores resume-to-JD fit, explains match/gap reasoning per requirement, and generates targeted interview questions — with reasoning that could survive an HR audit.
What Teams Should Build	Structured JD/resume parsing, requirement-level matching with evidence citation from the resume text, gap explanation, interview-question generator tied to gaps.
24-Hour MVP	10 synthetic resumes + 3 JDs, per-requirement match/no-match with quoted evidence, generated interview questions.
Data/Input	Publicly available anonymized resume templates + synthetic JDs.
Key Constraints	Must avoid protected-attribute bias (age, gender proxies); every verdict needs textual evidence, not just a score.
Expected Output	Fit report per candidate with evidence-backed reasoning and prep plan.
Evaluation Criteria	Explainability quality, absence of bias signals, practical usefulness.
Bonus Features	Bias audit report; batch ranking with diversity-aware tie-breaking.

HTH-GA-05: Research-to-Strategy Synthesis Assistant

Category	Details
Industry	R&D / Corporate Strategy
Scenario	Product and R&D teams must digest dozens of research papers/competitor reports before deciding where to invest.
Problem Statement	Build a system where users upload multiple papers/reports and get comparative synthesis — common findings, contradictions, methodology gaps, and a recommended action stance — all citation-grounded.
What Teams Should Build	Multi-document RAG, cross-document comparison engine, contradiction detector, citation-grounded Q&A.
24-Hour MVP	5–8 public research papers on one topic, generate comparison table + synthesized recommendation with citations.
Data/Input	Publicly available arXiv/open-access papers.
Key Constraints	All claims must cite source paper + section; must flag when papers disagree rather than averaging silently.
Expected Output	Synthesis report + comparison matrix + recommended next research direction.
Evaluation Criteria	Grounding accuracy, synthesis insight quality, usability.
Bonus Features	Auto-generated literature-gap map; citation graph visualization.

HTH-GA-06: AI Code Review & Release-Risk Assistant

Category	Details
Industry	Software Engineering / DevOps
Scenario	Engineering teams ship faster than they can manually review for security, performance, and maintainability, causing production incidents.
Problem Statement	Build an AI code-review assistant that scans a pull request's diff, flags bugs/security issues/performance risks/code smells with explanations and suggested fixes, and produces a release-risk score.
What Teams Should Build	Static-analysis-augmented LLM review pipeline, issue categorization, fix suggestion, aggregate risk scoring.
24-Hour MVP	Analyze 3–5 sample PRs/diffs (from public GitHub repos), generate categorized review comments + risk score.
Data/Input	Public GitHub repositories/diffs.
Key Constraints	Must distinguish "certain bug" vs. "stylistic suggestion"; false-positive rate must be visible to judges.
Expected Output	Annotated review report + risk score + top-3 must-fix issues.
Evaluation Criteria	Accuracy vs. known issues, explanation clarity, actionable prioritization.
Bonus Features	Auto-link error logs/stack traces to the exact code line; auto-generated test cases for flagged bugs.

HTH-GA-07: Multi-Agent Disaster Response Coordinator

Category	Details
Industry	Public Safety / Emergency Management
Scenario	During a simulated flood affecting multiple regions, responders must coordinate resource allocation, evacuation routing, medical triage, and public alerts under scarce resources.
Problem Statement	Design a multi-agent GenAI system (logistics agent, medical agent, communications agent) that collaborates to produce one coordinated response plan under constrained resources (limited vehicles, shelters, medics).
What Teams Should Build	Multi-agent orchestration, resource-constraint solver, conflict resolution between agents, public communication drafts.
24-Hour MVP	Simulated region map with 3–5 affected zones, fixed resource pool, agents negotiate allocation and output a prioritized action plan.
Data/Input	Synthetic disaster scenario data (population, resources, road network) provided or generated.
Key Constraints	Resources are strictly limited (must show trade-offs, not infinite supply); plan must be explainable to a human commander.
Expected Output	Coordinated response plan with resource allocation table, rationale, and draft public alert.
Evaluation Criteria	Feasibility under constraints, coordination quality, explainability.
Bonus Features	Real-time re-planning when a new zone is added mid-demo.

HTH-GA-08: Evidence-Based Personal/SME Finance Advisor

Category	Details
Industry	FinTech
Scenario	Small business owners and individuals get generic financial advice that ignores their actual transaction patterns.
Problem Statement	Build a GenAI advisor that nalyses real transaction data (synthetic) to identify unusual spending, recurring/hidden costs, and budget violations, then generates a financial plan grounded in the actual numbers — not generic tips.
What Teams Should Build	Transaction analysis pipeline, anomaly/recurring-cost detection, goal-tracking, grounded plan generation referencing actual transactions.
24-Hour MVP	Synthetic transaction dataset (500–2000 rows), detect anomalies/recurring charges, generate a plan citing specific transactions.
Data/Input	Synthetic transaction data generator (teams can build or use provided template).
Key Constraints	Advice must reference actual data points, not boilerplate; must handle irregular income patterns (SME context).
Expected Output	Financial health report with cited evidence and prioritized action list.
Evaluation Criteria	Groundedness, actionability, handling of edge cases (irregular income, subscriptions).
Bonus Features	“What-if” simulator for spending changes.

HTH-GA-09: Schema-Agnostic Natural Language Data Analyst

Category	Details
Industry	Business Intelligence / Analytics
Scenario	Business teams wait on analysts for simple ad-hoc questions about spreadsheets because they can't write queries themselves.
Problem Statement	Build a system where a user uploads any CSV/Excel file and asks natural-language business questions; the system must work without hardcoded column names and explain its calculation, not just show a number.
What Teams Should Build	Schema inference layer, NL-to-query translation (pandas/SQL), chart generation, explanation of computation steps.
24-Hour MVP	Support 3 different unseen CSV schemas at demo time, answer 5+ question types with charts and step explanation.
Data/Input	Public sample business datasets (sales, retail, e-commerce) from open data portals.
Key Constraints	No hardcoded schema assumptions; must gracefully handle ambiguous/unanswerable questions.
Expected Output	Answer + chart + plain-language explanation of how it was computed.
Evaluation Criteria	Generalization across unseen schemas, correctness, explanation quality.
Bonus Features	Follow-up question memory (conversational analytics); anomaly auto-flagging in the dataset.

HTH-GA-10: Controlled Synthetic Data & Edge-Case Generator

Category	Details
Industry	ML Engineering / QA (cross-industry tooling)
Scenario	ML and QA teams need realistic test data with controlled edge cases (fraud patterns, sensor faults, rare events) but can't use real production data due to privacy/compliance.
Problem Statement	Build a GenAI tool that generates realistic synthetic datasets to spec (e.g., "10,000 transactions with 2% fraud, seasonal spikes"), injects controlled anomalies, and validates statistical realism against target distributions.
What Teams Should Build	Spec parser, generative data engine, anomaly injector with configurable rate, distribution-validation report comparing generated vs. target stats.
24-Hour MVP	Generate one dataset type (e.g., transactions or sensor readings) to a natural-language spec, with adjustable anomaly rate and a validation report.
Data/Input	No external data required — purely generative, validated against user-specified statistical targets.
Key Constraints	Generated data must be statistically checkable (not just plausible-looking); must support reproducible seeds.
Expected Output	Downloadable dataset + distribution/validation report + anomaly log.
Evaluation Criteria	Statistical fidelity, controllability, usefulness for downstream ML testing.
Bonus Features	One-click "break my model" mode that generates adversarial edge cases for a provided model.