1. Project Kickoff

Project Goals:

Develop an AI agent that automates and streamlines tax compliance by mapping and analyzing relationships between entities (such as companies, individuals, and
transactions). The agent will extract, categorize, and reconcile data, identify compliance risks, and generate actionable insights to support accurate and efficient tax reporting.

Project Scope:

Focus solely on the development of the AI agent for entity relation mapping and tax compliance.

Exclude cloud, DevOps, and enterprise-scale deployment considerations. Deliver a working prototype that demonstrates core AI functionalities: entity extraction, relationship mapping, compliance rule application, and risk identification.

Expected Deliverables & Milestones:

Week 1: Requirements analysis, dataset preparation, and initial design.

Week 2: Development of entity extraction, relation mapping, and compliance rule modules.

Week 3: Integration, testing, documentation, and final reporting.

Timeline & Dataset:

Total duration: 3 weeks.

Dataset: Use publicly available or synthetic tax data representing entities, transactions, and compliance records.

Team Readiness & Gaps:

The team consists of two members:
Stephen: Background in accounting and tax, responsible for compliance rules, data validation, and domain expertise.

Sujal: Expertise in AI agents, ML, DL, NLP , and LLMs, responsible for AI development, entity extraction, and relation mapping.

The team's combined skills cover both technical and domain requirements; no significant gaps identified.

2. Team Discussions

Core Skills of Team Members:

Stephen: Accounting, tax compliance, regulatory knowledge, data validation. Sujal: AI agents, machine learning, deep learning, natural language processing, large language models.

Responsibility Assignment:
Stephen: Define compliance rules, validate outputs, provide domain insights, curate and annotate datasets.

Sujal: Design and implement AI modules, develop entity extraction and mapping, integrate compliance logic.

Skill Gaps:

No major skill gaps; the team's expertise is well-aligned with project needs.

Programming Languages & Platforms:

Python for all development.

3. Skills and Tools Assessment

Selected Tools & Frameworks:

Python (main language)
NLP libraries (spaCy or NLTK) for entity extraction
pandas for data manipulation

External Resources:

Reference to publicly available tax datasets and regulatory documentation as needed.

Task Assignment & Role Clarity:

Stephen and Sujal have clearly defined roles based on their respective strengths. Regular collaboration to ensure domain and technical integration.

4. Submission for This Iteration

Tasks In-Process:

Currently Stephen is been trying to collect data to try and test upon

Challenges & Solutions:

Challenge: Sourcing real-world tax data.
Solution: Use synthetic data and supplement with publicly available resources or build a fake data which looks like its real and try and test working of our AGENT on it.

Challenge: Ensuring compliance accuracy.
Solution: Leverage Stephen's domain knowledge for rule definition and validation.

Plan Changes:

Timeline condensed to 3 weeks with focused weekly milestones.

Data Link: https://northeastern.sharepoint.com/:f:/r/sites/DS5110entity-relation-mapping-tax-compliance-ai-agent/Shared

Excel Tracker: https://github.com/stephenminan/entity-relation-mapping-tax-compliance-ai-agent/blob/main/Project$_S chedule.xlsx$

Repository Link: https://github.com/stephenminan/entity-relation-mapping-tax-compliance-ai-agent