AIP CON⑤ / Tyson Foods

CODE MANAGEMENT

AIP + ONTOLOGY SDK FOR SYSTEM MIGRATION

→ PROBLEM

From 2019 to 2021, Tyson developed a data analytics platform on Google Cloud. However, by 2023, the rapid growth of their data led to significant challenges in scaling, performance, and costs.

Issues such as data latency, surging cloud expenses, and poor visibility into data processes began to impact business operations.

To address these problems, Tyson decided to migrate from their legacy data platform (ACE) to the DBT Cloud platform. Nevertheless, the migration process proved more difficult than anticipated, creating a resource strain and requiring specialized expertise.

To navigate these challenges and ensure a smooth transition, Tyson enlisted the help of Palantir.

→ SOLUTION

Using Foundry and AIP, Tyson deployed a workflow to automatically translate code from ACE to DBT.

This process is powered by AIP Logic with GPT-4o, Automate, the Ontology, and the Ontology SDK. It executes automatically, taking human feedback into account.

Human-in-the-Loop Code Translation

As a result, a project initially expected to take 10-15 people three years to complete was done by just one person in 3-4 months.

→ WORKFLOW

→ **Seeding**: The seeding process involves ingesting all legacy ACE models and current DBT translations. This is accomplished by leveraging Palantir's OSDK to load models and their translations into GitLab. A key feature of this step is the embedding and vectorization of model content, which enables effective semantic search.

→ **Insights**: In this phase, the goal is to teach the language model how to translate ACE models to DBT. By analyzing previous translations, we generate actionable insights and rules without the need for additional fine-tuning costs. Human oversight is incorporated to review and approve these learnings, ensuring accuracy.

→ **Translation**: The translation process begins by identifying similar approved translations using semantic search. The language model then applies learned rules and incorporates user feedback before performing the translation. OSDK pings DBT for an external syntax and compilation checks.

→ **Feedback**: User feedback is crucial for enhancing translation accuracy. In this phase, users review and approve translations in the Feedback Cockpit. They can also add feedback to improve future translations, which triggers a re-translation. This iterative process ensures continuous improvement and accuracy.

→ **Harvesting**: Approved translations and configurations are pushed to GitLab using the Ontology SDK. This step prepares the transformations for standard DBT testing and merging into production, ensuring that all changes are seamlessly integrated.

©2024 PALANTIR TECHNOLOGIES, INC.