Playground v0.0.2

Prerequisites:

We will ask the user to give a work email before he/she can have access to the playground.
Once the user gives a work email, we generate a UserID for him/her

Note: We hope to do Usage tracking and limiting, but not required for this version.

Description

Take user inputs, or allow users to take our sample queries
Allow users to decide the endpoints (GPT4 by default, others TBD)
Allow users to choose one of the model for comparison
If select a sample query, choose an endpoint and a compara model, and hit run

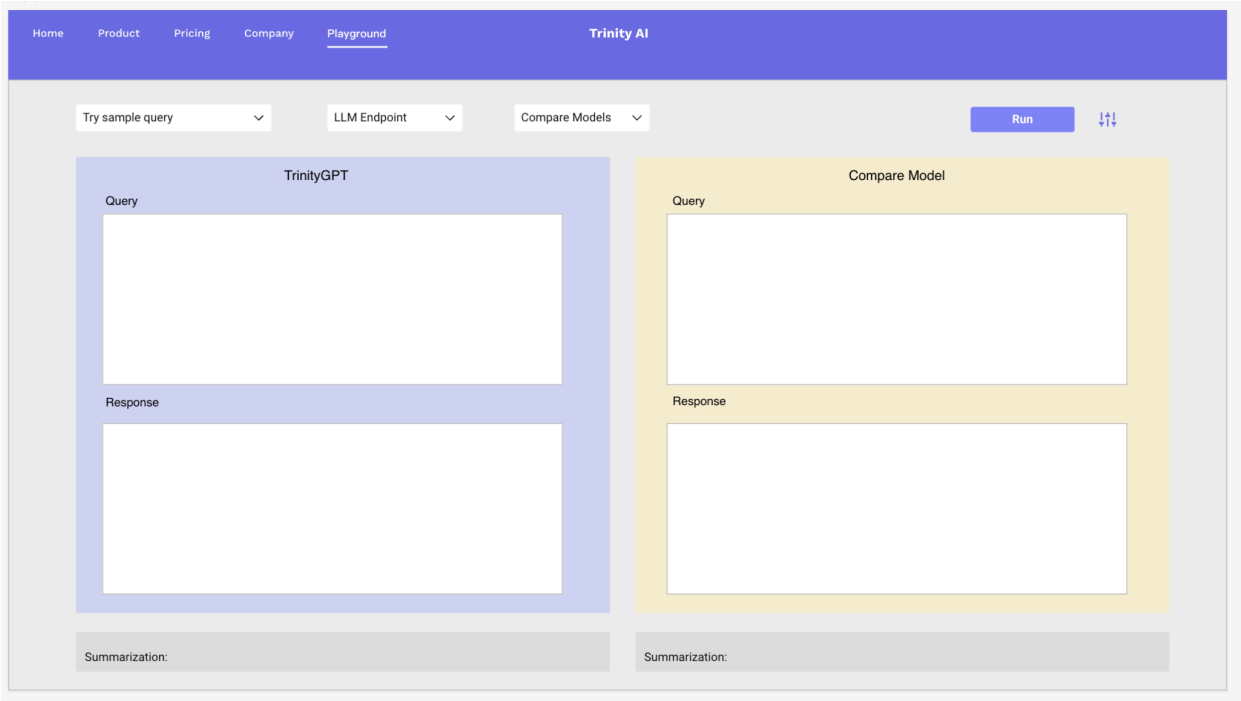
Then
1.Show how we protect the prompt by highlighting the sensitive parts in the text
2. Show responses
3. Show summarization for the results

The Setting icon on the left of RUN button is where

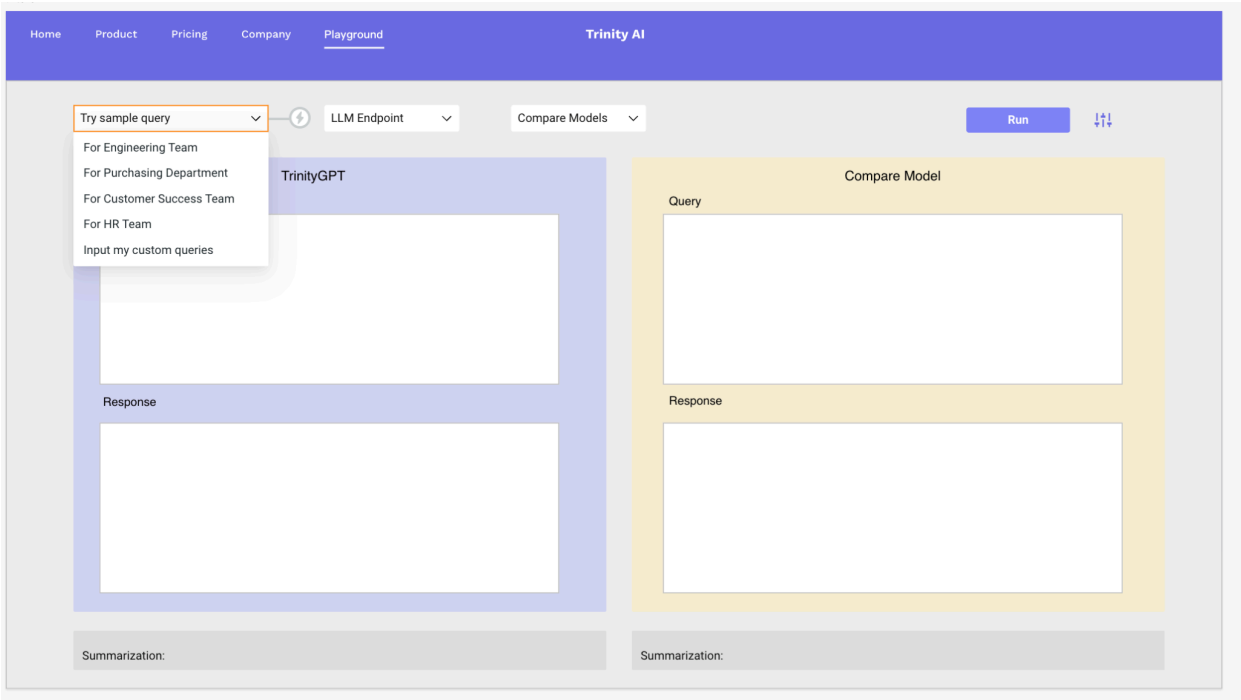
- Users can set up policies
- Decide Rating models

User stories

0 User Story: As a user, I want to see the Playground UI once given the access.

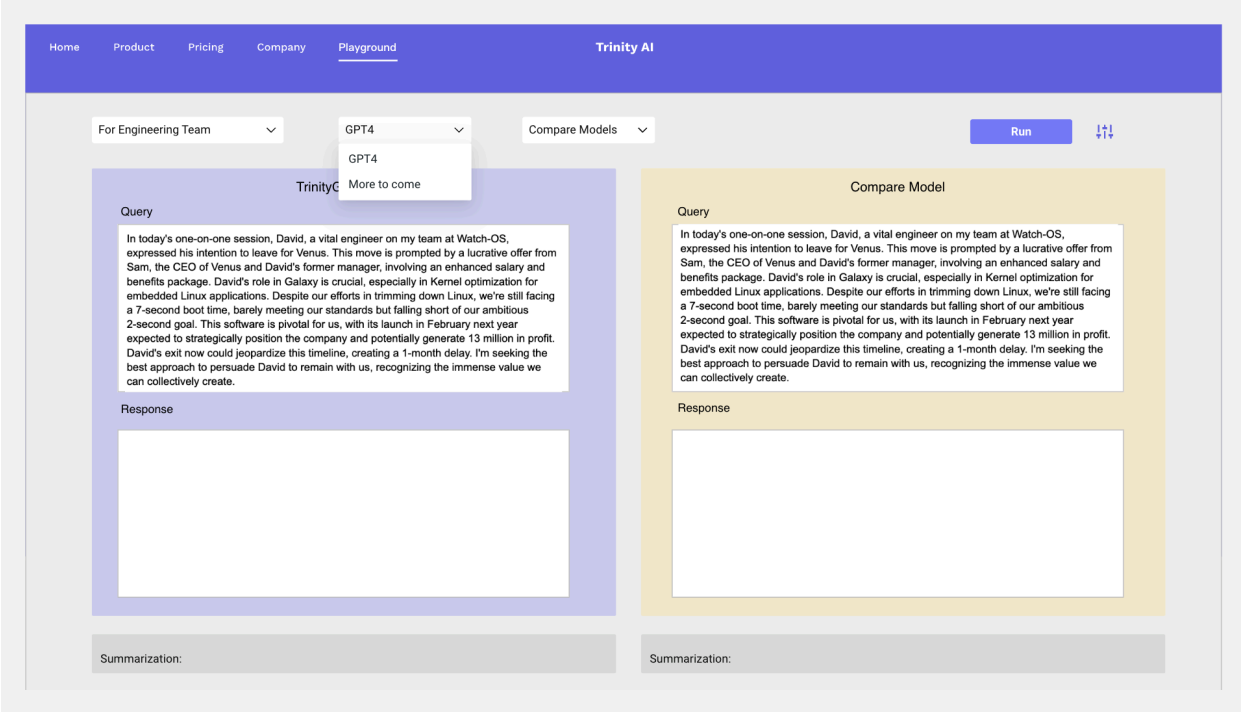


1 User Story: As a user, I want to select from samples and see the query in the Query box, so I can interact conveniently.
Requirement: Implement a model selection dropdown populated with available sample queries.

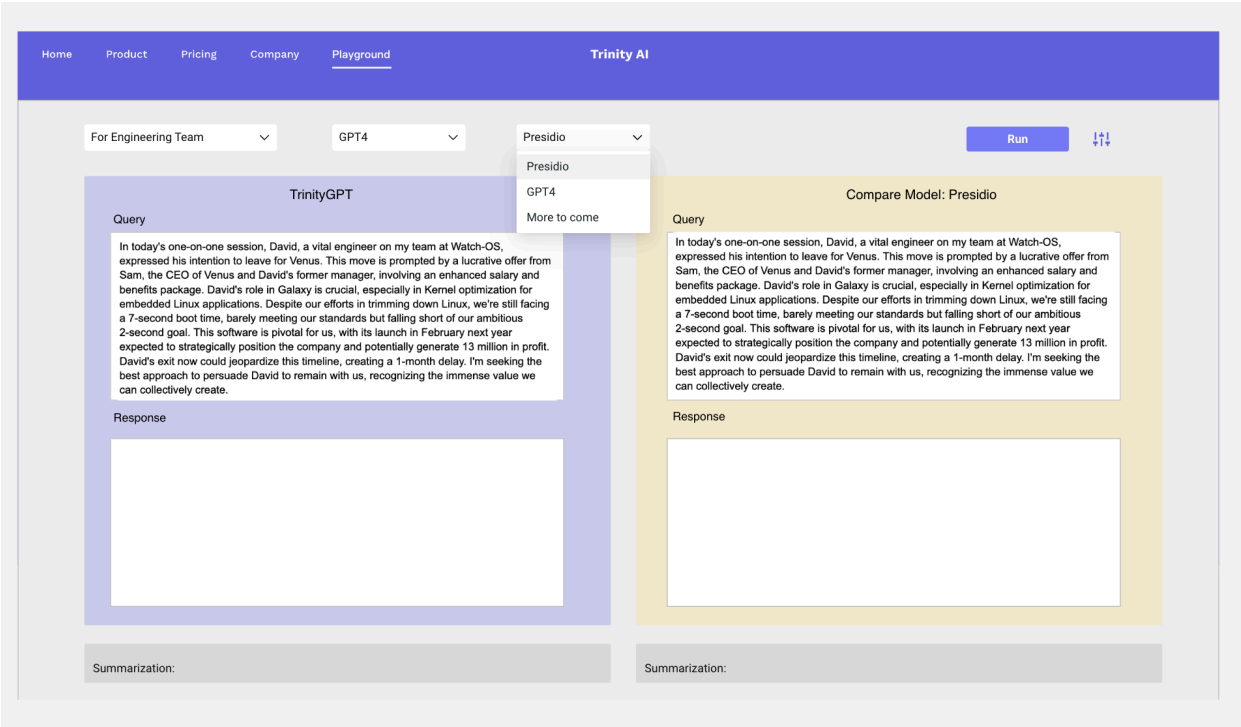


2 User Story: As a user, I want to choose “input my custom queries” and then type in the Query box, so I can interact with my own data.
Requirement: Allow users to give/modify inputs in the Query box.

3 User Story: As a user, I want to choose an available model from the dropdown.
Requirement: Get the available endpoints

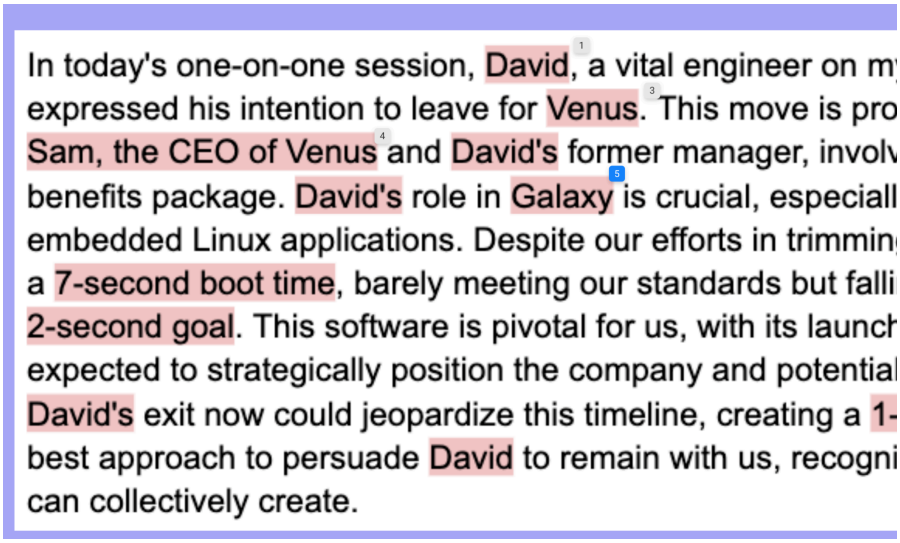


3 User Story: As a user, I want to choose different comparable models, so I can compare results.
Requirement: Implement a model selection dropdown populated with available models.

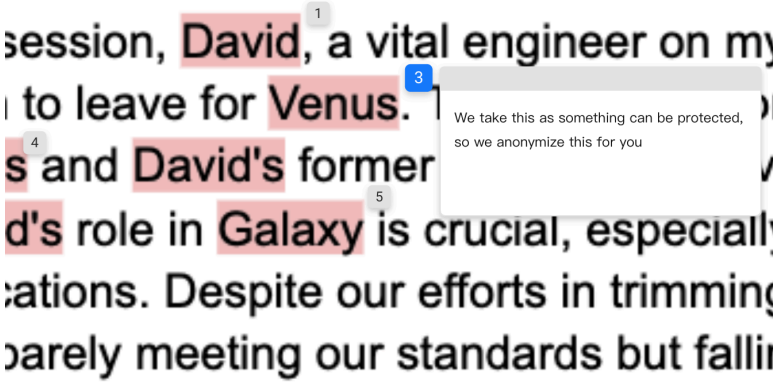


4 As a user, I want to see live results by clicking the RUN button to understand what Trinity can do with the given data/setting.
Requirement: Integrate with the decision engine/existing workflow.

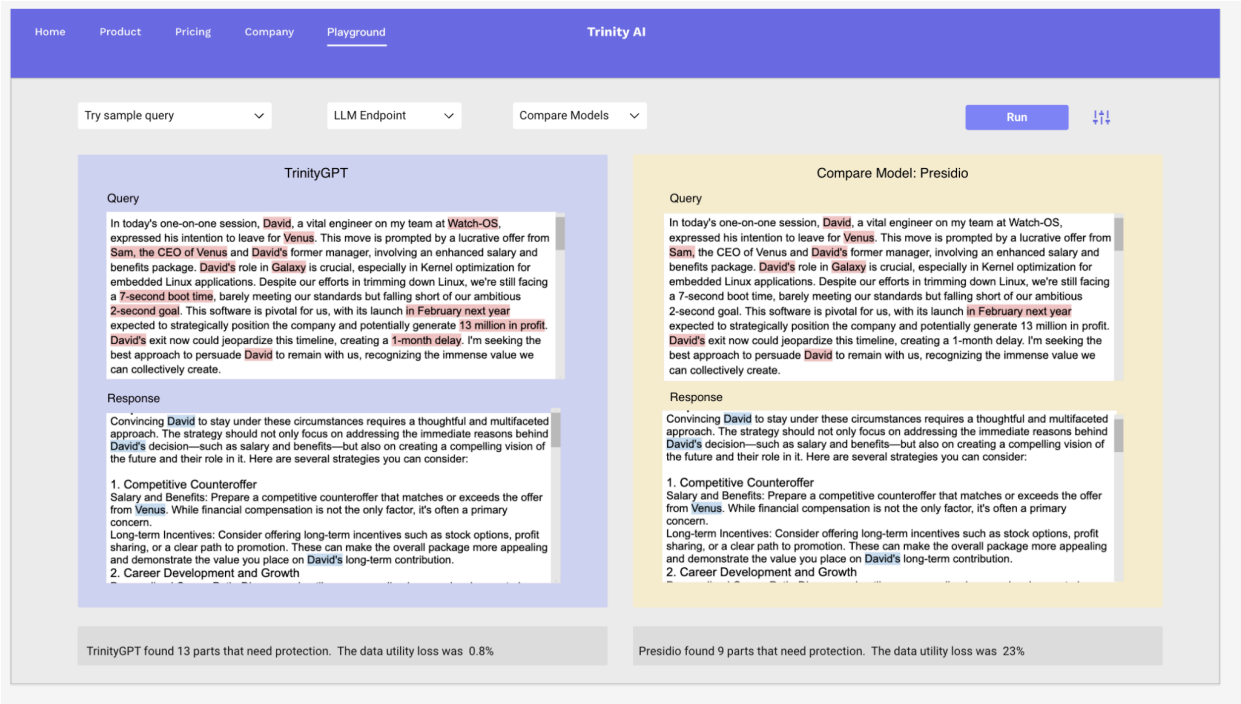
5 User Story: As a user, I want to “Feel” my data is secure, so I can trust the system with sensitive information.
Requirement: Use the mapping from the safety model and highlight anonymized sensitive data.



6 As a user, I want to “understand” why my data is secure by clicking the annotation button, so I can better trust the system
Requirement: Fetch some reasoning from the safety model and display annotations.



7 As a user, I want to see summarized results, so I can quickly understand the key information.
Requirement: Develop a summarization engine that condenses the AI's response without losing critical data points.



Entities and Attributes:

User: UserID (Primary Key),Email
Query: QueryID (Primary Key), UserID (Foreign Key), QueryText, EndpointID (Foreign Key), ModelID (Foreign Key), PolicyID (Foreign Key), RatingModelID (Foreign Key), QueryTimestamp

Relationships

A User can have multiple Queries.
Each Query must have one Endpoint , one compara_Model, …