**ESG Survey Automation Design**

**Business Need:**

Existing Enterprise Sustainability Team’s Survey Questionnaire Analysis, responding to the questions is time consuming process and evidence collection for auditing become challenging. Hence team is looking for the automation of this process which yields accurate Sustainability Reports, References & Responses to Stakeholders.

**Objective:**

The goal is to develop a robust, automated information collection and report generation system capable of parsing complex reports related to various categories concerning the ESG and other business documents to extract relevant data based on Survey Questionnaire related to the company. Automated Solution should fetch the Questions from Survey documents (.pdf,.txt and webpages and others) and record references & citations from related sources (e.g., document link, page number, paragraph/data) which address each question in the survey.

Additional Objective includes scrapping information available from internet as part of publicly accessible information about us.

**Scope:**

* Development of AI Search/Chatbot solution reads the data from various sources like .pdf,.txt and webpages and generate reports.
* Cloud system able to parse the multiple source/relevant data files for the Questions.
* Integrate NLP Libraries to the current GCP (Google Cloud Portal) to identify the Key Entities/Relationships and sentiment analysis within the text.
* Establish connections to external APIs or DBs to fetch reporting information from Documentation source files.
* Develop functions to retrieve and organize the data for the analysed entities.
* Implement reports generates best referential data based on the closest match.
* User Interface allows bank analyst to upload excel/pdf questionnaire and support documents as input to generate report along with the generative AI Based report.

**Out of Scope:**

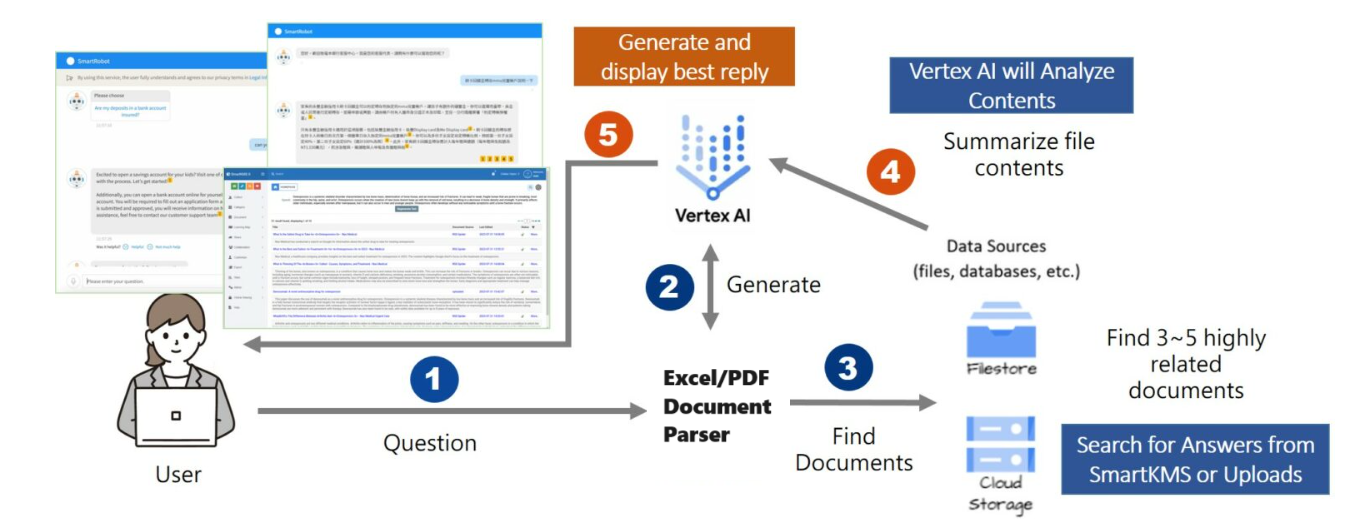
* Developed Automated AI solution exclude other survey sources and reference files which are not included in the integrated Cloud environment.

**Constraints:**

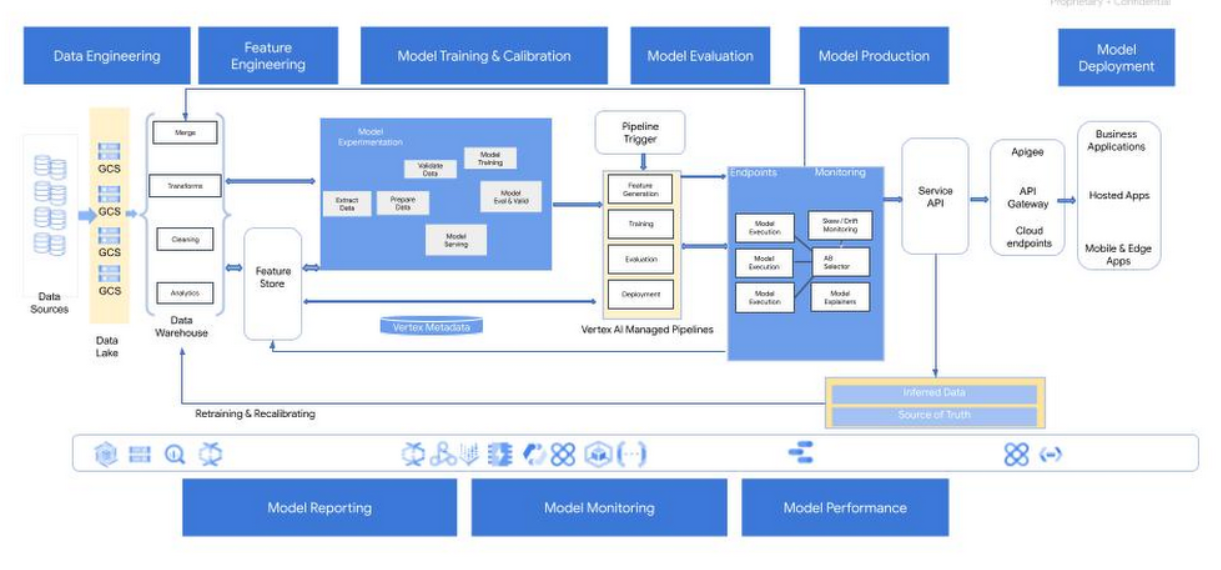
* Developed AI solution will work on the Sources and Data inputs only. Whereas if non-existent data related questions are posted, Chat Bot/Search Solution doesn’t provide any responses.
* Current solution provides the details based on the sample inputs provided from Environment Social Governance categories of Banking. Variances will be handled.
* Due to timeout challenges in the external environment for the Authentication, currently disabled and later can be enabled based on further detail requirement.

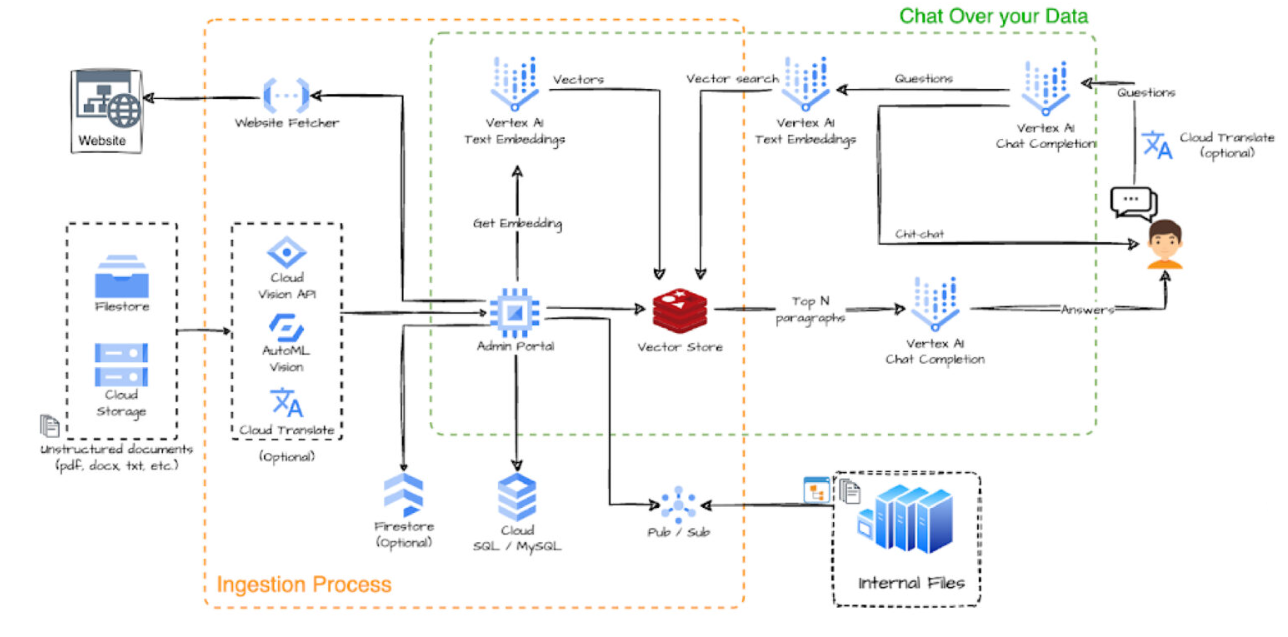
**Google Cloud Environment Vertex AI Search is selected to develop the solution**

**User API/Interface Interactive View**



**Typical Vertex AI Model Processing**

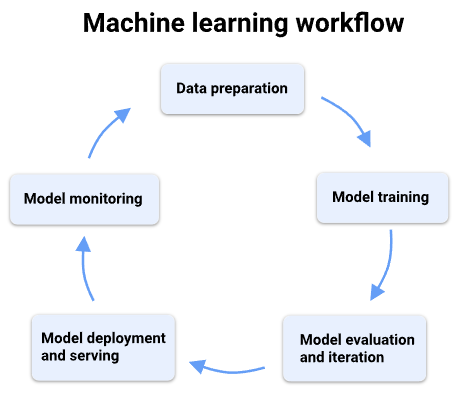
****

****

**Steps Involved in the Vertex AI Processing**

**- Data parsing and Database Creation for analysis**

For this Hackathon we are selecting the Vertex AI Search and Conversation, Big Query and Cloud storage APIs inbuilt in the GCP Cloud environment since the Survey Data is both structured and unstructured.

****

**Data Preparation:**

ESG Survey Questionnaire documentation from various third-party sources will be collected and ingested to Vertex Database.

We will be uploading all the source Survey Questionnaire and Supporting Data Inputs (Reports and Other theoretical insights) to the Data Store which acts as a database for further processing.

Big Query API Provides data retrieval capabilities from Structured Data sources whereas Vertex AI API reads the data from unstructured files.

**Model Processing**

Text Bison LLM model is used in the Model Training which provided the capabilities of single-turn instruction tasks like classification, extraction, summarization and generation.

* Search summarization is enabled to create the summaries from the extracted similar questions/phrases across the Survey PDFs.

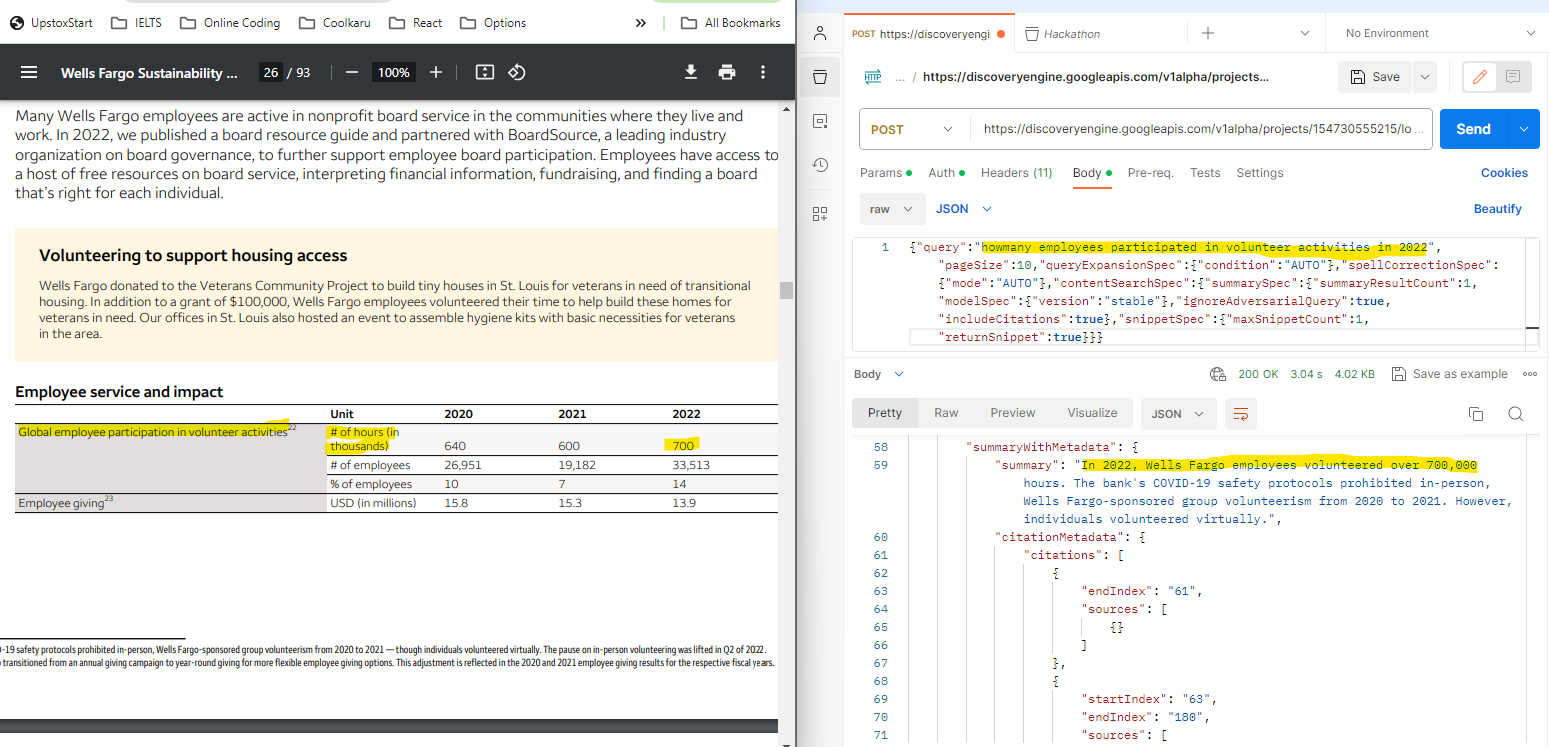
**API Integration with User interfaces**

API endpoints with the Vertex AI inbuilt Search Options and Chatbot Options are integrated within user interface.

Widgets are called to provide the necessary responses to the users.

We have enabled the API services in GCP to provide the API functionality to the end users to integrated with multiple application on need basis.

* We have enabled the OAuth 2.0 token based authorization to consume the API.
* Below is the sample request and response from postman using API.



**User Interface Options**

User will have the following capabilities in the Web interface.

* Search Capability which provides the Survey Responses based on the question posted.
* Chatbot option which provides quick response.
* Summarised Report Metrics for the further reference.