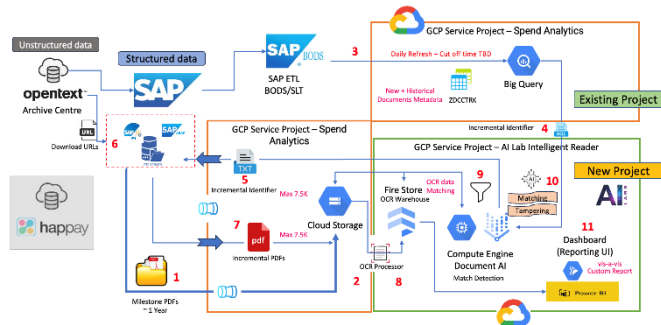# Document AI – Multi Modal Large Language Model

The OCR (Optical Character Recognition) and Multimodal LLM (Large Language Model) based intelligent solution represents a cutting-edge approach in data extraction and classification from vast repositories of scanned documents. Its primary objective is to discern financial documents such as invoices, credit notes, and vouchers with unparalleled accuracy and efficiency, thereby mitigating posting errors and ensuring timely detection of potential discrepancies or red flags.

At its core, this solution harnesses the power of OCR technology to convert scanned images of documents into machine-readable text, facilitating automated data extraction. Additionally, it leverages Multimodal LLM, an advanced form of language model, which integrates various data modalities such as text, images, and possibly even audio, enabling more comprehensive analysis and classification.

One of the key strengths of this solution lies in its ability to handle a massive volume of documents. With a daily volume exceeding 100,000 pages and a cumulative repository of 15 million pages over the past year, amounting to more than 10 terabytes of data, the solution demonstrates scalability and robustness. Despite the vast amount of data, it boasts an impressive expected runtime of less than 4 hours, ensuring timely processing and analysis.

Beyond mere data extraction, the solution plays a pivotal role in regulatory and statutory compliance. By identifying and classifying financial documents accurately, it facilitates the analysis of potential duplicates and enables the linkage of related claims processing. This functionality is crucial for compliance with various regulations and statutes governing financial transactions.

In summary, the OCR and Multimodal LLM based intelligent solution represents a groundbreaking innovation in document processing and analysis. Its ability to extract and classify data with unprecedented accuracy and efficiency, coupled with its scalability and compliance capabilities, makes it an indispensable tool for organizations grappling with vast volumes of financial documents.

Detail Design

Prepared by – Gour Nath [gourprosad.nath@adani.com], AGM, AI Labs, Adani Group
Head of Department - Sreenivas Ramarao [Sreenivas.Ramarao@adani.com], SVP, AI Labs, Adani Group