

Optimize experimentation cost by enabling faster access to historical data for researchers



latentview

Actionable Insights • Accurate Decisions

Client

One of the leading oil-and-natural-gas companies across the globe

The Problem



The R&D scientists used to manually find the necessary documents needed for their experiments, which were cumbersome in the initial phase of the research.



The R&D scientists used to inadvertently end up repeating already conducted experiments due to the inability to access related documents.

The Before State



Material recipes were present in the form of scanned PDFs that contained printed and handwritten steps. These needed to be digitized into a structured format for storage and retrieval.



The LatentView Solution

The Chemical Synthesis Miner solution used the following approach to overcome challenges:



- In-scope and out-of-scope nuances in the digitization process were defined clearly
- A thorough evaluation of the best combination of OCR packages to achieve the best results for handwritten text along with printed text in this use case
- Custom-trained NLP models were used for Named Entity Recognition and Dependency Parsing
- A tailored dictionary was created to identify commonly used chemical terminology
- Constant iterations with the stakeholders to help ratify the progress made



CSM was the pioneer in LV to use a graph database and design an end-to-end pipeline for translating an unstructured PDF document containing printed and handwritten text, identifying entities and their relationships, storing them in a graph database, and retrieving information through a web application and chatbot. This was also one of the first applications to be hosted in an Azure environment. The innovation brought about in this project attracted tremendous traction within LV and helped us gain more projects.

The After State



As indicated by the client, the **expected business impact was to the order of \$3 million/year.**



Using LV's solution, the client can now access the centralized repository, and weeks of effort by the research scientists are cut down to a matter of minutes or even seconds.



Using this solution, the client can now spend the time, material resources, man-hours of effort, money, and even human resources more optimally.

About: LatentView Analytics

LatentView Analytics is a leading global data and analytics service provider helping companies turn data into actionable insights to gain competitive advantage. As a trusted analytics partner to the world's most recognized brands, LatentView solutions provide a 360-degree view of the digital consumer, fuel machine learning capabilities and support artificial intelligence initiatives. LatentView's success is driven by a commitment to deliver unrivalled analytics solutions that enable Fortune 500 companies in the retail, CPG, BFSI, high tech, healthcare and other sectors to predict new revenue streams, anticipate product trends, improve customer retention, optimize investment decisions and turn unstructured data into a valuable business asset. LatentView has offices in Princeton, N.J., San Jose, Calif., London, Singapore and Chennai, India with more than 600 employees globally.

For more information, please visit www.latentview.com or follow us on [LinkedIn](#), Email: sales@latentview.com