INTELLIGENT DOCUMENT PROGRAM

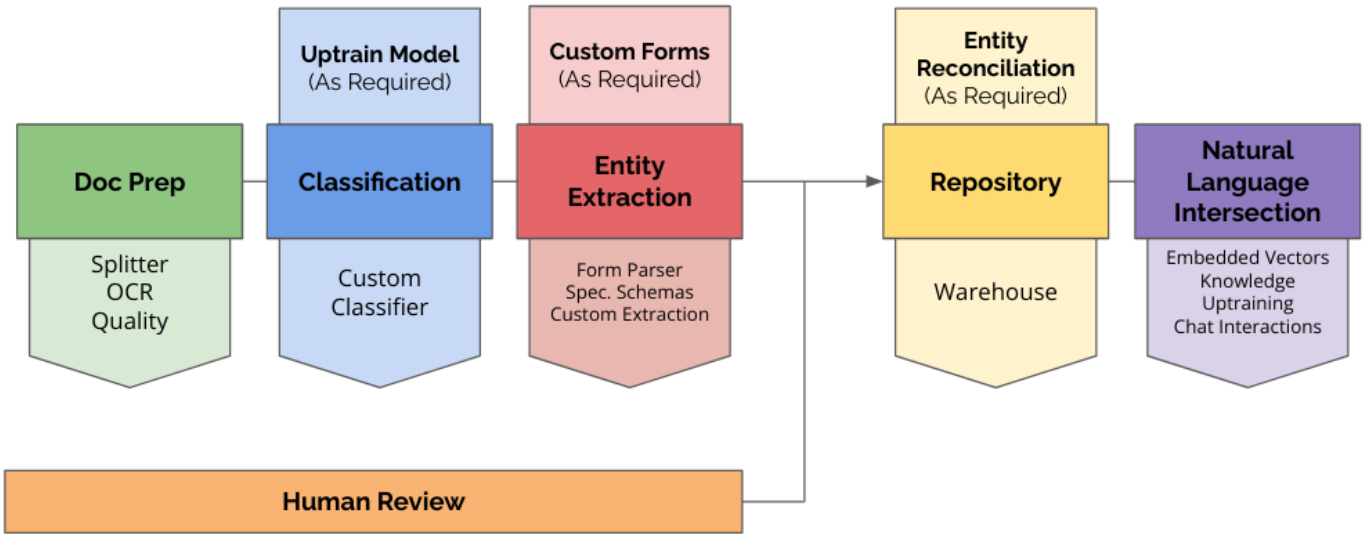
OVERVIEW

Over years of operations, organizations have accumulated a wealth of valuable information within their document libraries. However, this data often remains untapped, and if required, it is typically duplicated manually. This can be an expensive endeavor.

Fortunately, with the emergence of advanced machine learning tools and technologies, the hurdles associated with document processing and analysis can now be automated and streamlined. This breakthrough enables businesses to enhance their decision-making processes and unlock new efficiencies.

Some of the capabilities our Intelligent Document Program enables are:

- Better understanding of documents through the **automated extraction of structured data,** including key entities such as people, organizations, and locations.
- **Content classification** to improve the search and management of document libraries.
- **Document and content comparison** to identify similarities and differences between two or more documents.
- Enhanced **document search**.
- Multilingual document **translation**.



This all results in **improved decision-making processes**.

FINE-TUNING DOCUMENT INGESTION

In the recent past, these problems could only be solved with a tremendous amount of customized modeling and training. Projects were costly and difficult to justify. Thankfully, those times have passed. **Google's Document AI** suite of tools and technologies can now solve a significant portion of the automated document processing and classification challenges.

With Document AI, GDMIL has the flexibility to provide low-cost workflows that can process data from many types of forms and documents "out of the box". Furthermore, these tools allow for "uptraining" or building on top of Google's base models to provide more accuracy or allow for additional form or document types. Lastly, the Document AI Workbench allows for the development of full models by leveraging its training and form element selection tools to minimize custom development costs. Once the information is extracted, Google's Entity Reconciliation API can be used to resolve the semi-structured data appropriately into the business entities within your data warehouse. Google's Enterprise Knowledge Graph allows for Google-like searches through ranked business data.

GENERATIVE AI

GDM Innovation Labs uses these tools within our data practice to augment structured data sources with this untapped document data. Recent advances in generative AI lets us go even further. New tools built on natural language interactions like Google's Bard and OpenAI's ChatGPT also allow for the notion of uptraining. These tools, along with new vector database technologies, allow GDMIL to combine structured data query, document search and generative AI applications with your enterprise's data assets to allow for creative new workplace productivity enhancements.

MVP ENGAGEMENT OFFERING

The Intelligent Document Program will provide insight into key areas including:

- Searchable Document Content at Scale
- Cross-Silo Data Integration and Warehousing
- Analytic Solutions
- Generative AI Knowledge Augmentation
- Natural Language Search

TIMELINE

Week One

- Understanding of Client Goals for Document Automation
- Use Case Finalization
- Content Identification (Forms and Documents)
- Requirements Workshops and One-on-One Meetings with Key Stakeholders
- Agreement on Final Deliverable Details

Week Two

- Extraction "Level" Testing (Validating Document Assumptions)
- Use Case Architecture and Data Modeling
- Document Extraction Training Plan (If required)
- Development and Integration Plan (If required)

Week Three

- Deliverable Review
- Executive Readout Workshop

DELIVERABLES

- Findings Including Program Current State and Desired Future State
- Technical Reference Architecture Recommendations
- 1st MVP Roadmap including Level of Effort and Recommended Timelines
- Platform Pricing Estimates
- Executive Readout Workshop

CONTACT

Should you have any questions, please feel free to contact us at: info@greatdataminds.com.