

# Intelligent Document Processing

Eliminate manual data entry and access unstructured text data from scanned documents for business-critical insights

## Business Process Automation as a Driver of Cost Optimization

Organizations in industries as diverse as insurance, healthcare, supply chain, and finance rely on manual processing or utilize expensive, legacy OCR solutions to process claims, invoices, shipment orders, contracts, and other unstructured documents. As part of business processes, legacy solutions represent a huge operations bottleneck and introduce unnecessary expenses to be eliminated.

## Automated Information Extraction & Document Processing

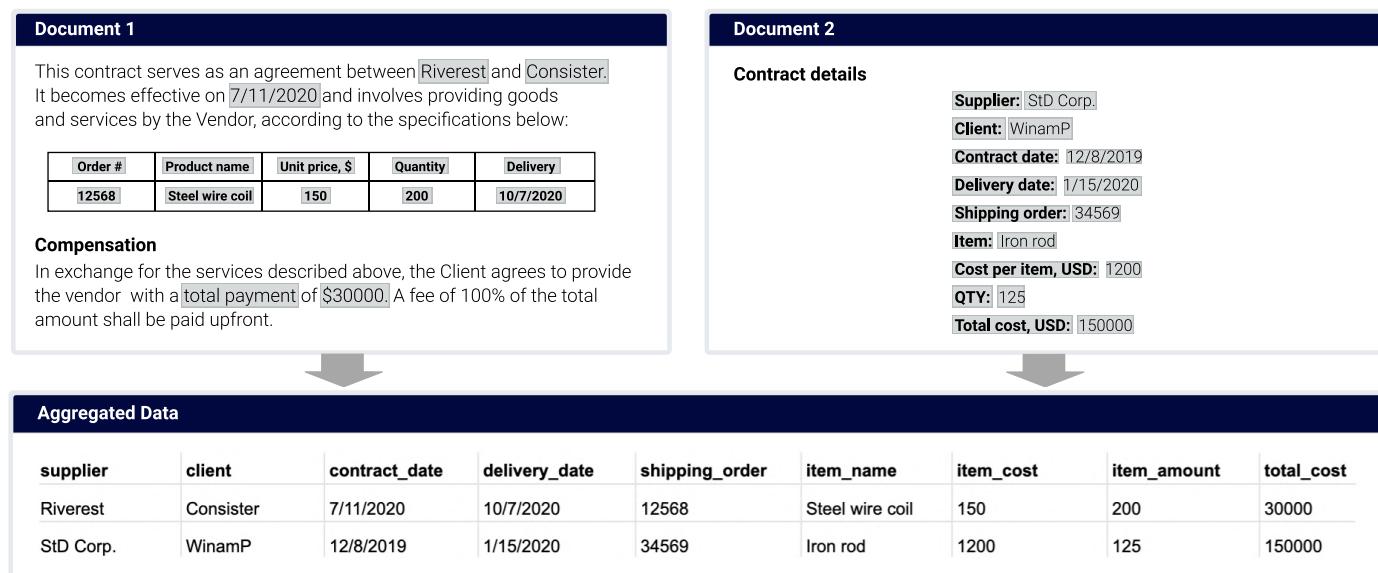
Intelligent Document Processing is an AI-powered solution that automatically extracts and converts unstructured text data from scanned documents into searchable and reusable structured formats. Our proprietary technology built on top of Amazon Textract goes beyond optical character recognition (OCR) – it identifies and matches the contents of various types of documents, aggregating them in tables. Powered by Natural Language Understanding (NLU), it comprehends the text, thus preventing mismatches caused by wording, and flags disputed cases for human review, to increase accuracy and ensure the system self-learns and improves over time while keeping humans in the loop.

## Key Benefits

<b>Accurate Extraction and Conversion</b> Capable of ensuring 80-95% accuracy of text extraction, the solution flags disputed cases for manual review to prevent any errors in data.	<b>Continuously Improved Accuracy</b> Powered by a Natural Language Understanding technology, the solution continuously retrains the ML model and improves its self-sufficiency.
<b>Streamlining Business Processes</b> As human review is needed only if a confidence score is lower than 95%, companies can scale document processing and reduce the costs.	<b>Faster Access to Analytics Initiatives</b> Scanned documents are processed at scale, and businesses get faster access to structured data to drive AI and analytics initiatives.

# Intelligent Document Processing

## How It Works



Intelligent Document Processing Solution combines the capabilities of Amazon Textract and Provectus' proprietary technology powered by NLU. In this solution, Amazon Textract is used to extract characters, forms, and tables while Provectus' custom NLP models extract key metadata from texts. This allows to comprehend the contents of the document, match different variants of text with the same meaning (i.e. shipping order, order #, etc.), and scale across different domains. The solution improves the accuracy of text extraction over time by learning from new data and corrected cases. It features a business dashboard with performance metrics to improve total visibility into document processing.

## Key Features

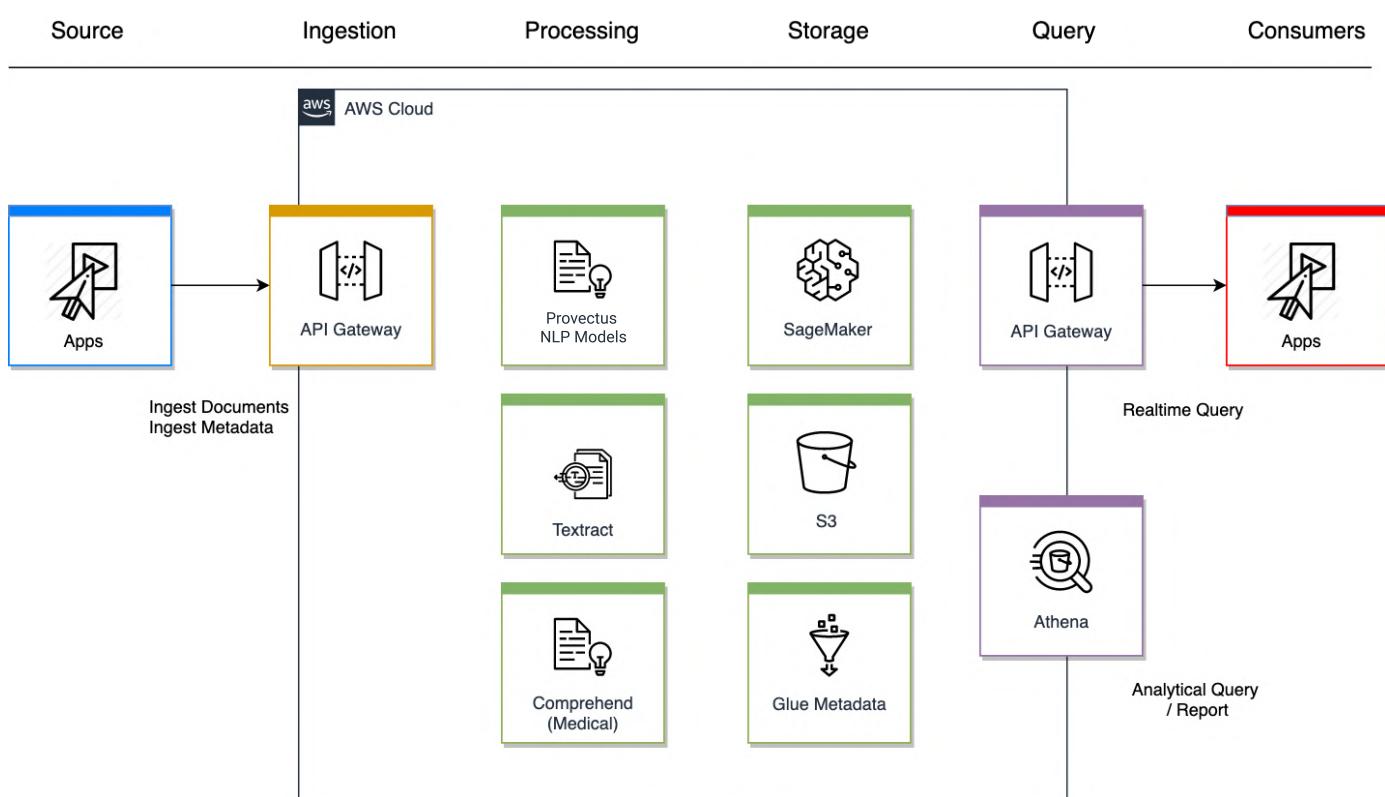
- Pre-built machine learning architecture, including infrastructure for documents data extraction, data ingestion, and data analytics
- Pre-trained machine learning algorithms that learn, adapt, and improve based on data collected over time
- Human-in-the-loop system that allows to review disputed cases and input corrected data into the ML model to increase text extraction accuracy
- Processing of various types of scanned documents: health records, patient registries, tax documents, sales orders, enrollment forms, benefit applications, insurance claims, shipping manifests, invoices, etc.
- Data security methods ensure that the data stored and processed through the system is protected from theft, leakage, and deletion
- Cloud vendor agnostic – The architecture is certified to be used with different cloud providers.

# Intelligent Document Processing

## Real-World Case

The under NDA client, a US-based consumer-directed healthcare solutions provider, sought to automatically extract valuable data from insurance claims. Provectus built an ML solution on AWS that extracts the required information from claim documents automatically and at scale, with accuracy of no less than 80%. The solution is designed so that ML models can be retrained to learn from new data points, to increase text extraction accuracy over time.

## Solution Architecture based on AWS



## About Provectus

Provectus is an Artificial Intelligence consultancy and solutions provider helping businesses achieve their objectives through AI.



## Contact Us

125 University Avenue, Suite 290  
Palo Alto, California, 94301  
+1 (800) 950-9840

[hello@provectus.com](mailto:hello@provectus.com)  
[provectus.com](http://provectus.com)