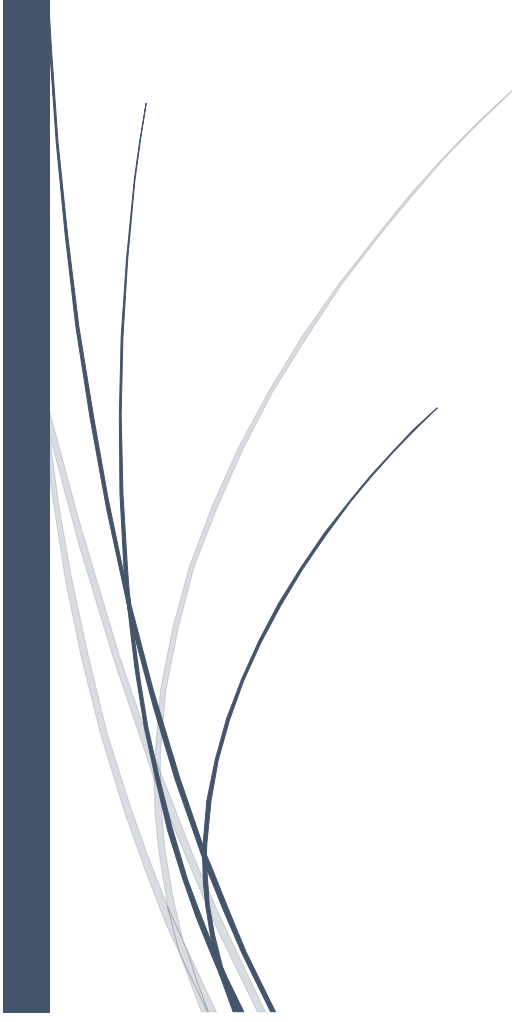
2023

# Intelligent Document Processing



# Contents

INTELLIGENT DOCUMENT PROCESSING ..... 2
PROCESS ..... 2
FS IDP KEY FEATURES FOR COMPANIES ..... 2
NOTABLE USE CASE ..... 2
DRIVE MEASURABLE RESULTS ..... 3
PROCESS THOUSANDS OF DOCUMENTS PER DAY ..... 3
EXPERIENCE ..... 4



© 2023 FINDABILITY SCIENCES | CONFIDENTIAL

INTELLIGENT DOCUMENT PROCESSING

**Extract and organize information from complex content in any document format**

Intelligent Document Processing (IDP) solutions transform unstructured and semi-structured information into usable data. Business data is at the heart of digital transformation; unfortunately, 80% of all business data is embedded in unstructured formats like business documents, emails, images and PDF documents.

Intelligent document processing is the next generation of automation, able to capture, extract and process data from a variety of document formats. It uses AI technologies such as Natural Language Processing (NLP), Computer Vision, Deep Learning and Machine Learning (ML) to classify, categorize, and extract relevant information, and validate the extracted data.

## PROCESS

## FS IDP KEY FEATURES FOR COMPANIES

Enhance customer experience & reduce underwriting time by getting better data from customer documents. Improve accuracy and straight through processing for your claims process.

Findability Sciences makes it possible for you to capture & analyze data from thousands of application forms, address proofs, invoices and certificates of insurance easily. Process documents much faster with improved accuracy and onboard customers in real time.

### NOTABLE USE CASE

**IDENTITY VERIFICATION**

Smooth & efficient verification of multiple KYC Documents as a proof of identity before extending your services. Onboard customers from the safety of their homes in real time.

- Verify driver's licenses
- Match identity from multiple documents
- Match selfie with ID
- Verify with external databases

**INSURANCE APPLICATION FORM**

Extract data from application forms without having to write custom rules. Validate extracted data from other documents received from your customers in real time.



© 2023 FINDABILITY SCIENCES | CONFIDENTIAL

- Extract key value pairs
- Verify from multiple documents
- Optical mark recognition
- Identify signature blocks

**CLAIMS AUTOMATION**

Read proof of loss documents such as invoices with high accuracy. Categorize line items with product catalog to achieve high straight through processing.

- Invoice data capture
- Categorize line items
- Validate values such as tax rate & amounts

**SUMMARISE DOCUMENT**

FP-Summary™, a text summarization engine can be successfully extract key takeaways and summarized it into a compact paragraph. FP-Summary™ runs on a proprietary, extractive algorithm that isolates the most significant excerpts from descriptive content. Applications include processing legal agreements, books, news data, and research papers.

## DRIVE MEASURABLE RESULTS

1. **Save Time**: Process documents in under a minutes
   Free your operations team from repetitive and redundant tasks, so you can save time and shift the focus of your team towards more strategic initiatives. That means you can react more quickly, and customers will love you for it.
2. **Save Costs**: Reduce your back-office costs
   Save up to 50% on document processing and **back-office** costs. **Findability Sciences** APIs for document classification, identification of key attributes, and data verification allow you to free your team from routine paperwork and reduce the number of paid hours they previously spent on these tasks. That means you can invest this money on scaling your business.
3. **Improve Accuracy**: Achieve 99%+ accuracy with human-in-the-loop
   With our human-in-the-loop review tool, you can now automate processes that need next to 100% accuracy. Verify data rather than manual keying and increase the accuracy of your downstream processes.

## PROCESS THOUSANDS OF DOCUMENTS PER DAY

- **EASY TO USE:** Send data through API or upload it in web app and within seconds your processed output is ready to be downloaded or send it directly to other software.
- **DATA SAFETY**: Your data is encrypted with the highest standards during transfer and storage.
- **ANALYTICS & REPORTING**: Create work queues for your operations team and get detailed performance reports. Get weekly and monthly analytics on for your automated workflow.



© 2023 FINDABILITY SCIENCES | CONFIDENTIAL

# EXPERIENCE

The following organizations have been part of the FS' Experience:

| ORGANIZATION | SCOPE OF ENGAGEMENT |
|---|---|
| **A US based financial services company** | Analyzed financial transactions report for a leading, Wall-street based, credit rating agency. This involved classifying, extracting, documenting, and reverse-tagging documents related to Collateralized Loan Obligations (CLOs) and Collateralized Mortgage Obligations (CMOs)<br><br>Findability Sciences used unstructured data processing techniques, powered by NLP (Natural Language Processing) to develop a cognitive AI system that could accept documents in multiple formats as input (pdf, scanned pdfs, spreadsheets, etc.) and extract relevant information such as tables, keywords, and crucial insights. The AI system used proprietary search tree algorithms to ensure the accuracy of extracted information. |
| **A private deemed university located in Mumbai.** | The client to develop and implement an Artificial Intelligence based data extraction software system that would extract specified data from images of documents that accompany student's application for college admission to automate a hitherto manual process of data entry.<br><br>Findability Sciences developed and implemented an Artificial Intelligent enabled data extraction software system that will retrieve information from<br>i. ID documents and marksheets such as: Aadhaar, PAN Card, Driving License, Passport, 10th Marksheet, 12th Marksheet, Degree Certificate, Degree Marksheets<br>ii. Document type: pdf or image formats (jpeg, png, etc)<br><br>The system extracted 5 fields per document. For example, from the marksheets - Name, Date of Birth, Subjects, Marks Scored, Year of Graduation, etc. were extracted. The details were stored as a key-value pair format for the defined fields<br><br>The solution was integrated with the customer's CRM (Salesforce) in the back end. Role-based access was provided for viewing and downloading reports for review. |
| **An American- Japanese multinational pharmaceutical company.** | The client tested multiple batches of pharmaceuticals in their laboratories every day. The lab personnel manually compared test results with the standardized 'certificates of analyses. This approach to quality assurance was time consuming and resulted in loss of productivity.<br><br>Findability Sciences used unstructured data processing techniques, powered by NLP (Natural Language Processing) to benchmark test results against the specifications in certificates of analysis. The AI-system accepted inputs in pdf/ scanned formats and published a tabulated comparison of product specifications as output. The user-interface applied a color code to highlight any mismatches between the various |



© 2023 FINDABILITY SCIENCES | CONFIDENTIAL

| | |
|---|---|
| | specifications of test batch and the corresponding certificate of analysis. The solution was deployed on premise.<br><br>AI-system exhibited 100% accuracy in identifying mismatches between test and standardized results, prompting the client to extend the scope of project to more complex product lines. |
| **Book: The Tech Whisperer** | Findability Sciences Leveraged an AI-powered, cognitive analytics system to compose a chapter on Artificial Intelligence in the bestselling book, "The Tech Whisperer" authored **by Jaspreet Bindra**. The solution deployed used AI algorithm for extraction and summarization of text to analyze large volume of AI related content and summarized it into a compact chapter on AI in the book.<br><br>The AI system was connected to a repository of open-source text content related to AI, ML, Deep Learning, etc. The extractive algorithm built a graph associated with the text, where the graph vertices represented sentences. The edges were established based on the 'similarity' relation between the connecting sentences. The 'weight' associated with each edge, represented the extent of content overlap. The algorithm scored each of the vertices based on the paired 'weight'. Sentences were sorted based on the score and top-ranked sentences were included in the chapter. |



© 2023 FINDABILITY SCIENCES | CONFIDENTIAL