



TEXTTRACT: ANALYZE SCANNED DOCUMENTS

WHAT IS TEXTTRACT?

Textract is a machine learning service that automatically extracts text, handwriting, and data from scanned documents. It can quickly automate document processing.

**“Why trouble
your experts
when Textract
can do it faster?”**



Variety of Documents

PDFs, Images, Tables, Forms
& OCR software



Languages Detected

English, German, French,
Spanish, Italian & Portuguese



Accurately Extracts

Text, Handwriting, Tables &
other data forms

USE CASES



Financial Sector

Interest rates, Application names and Invoice Totals



Public Sector

Government Applications and Federal Tax Forms



Business Operations

NDAs, Lease Agreements, & Terms and Conditions

CURRENT FLOW

PDF or Image



Expert manually inspects files



Sorts files and highlights the mismatches



TEXTTRACT FLOW

PDF or Image



Texttract automatically extracts files



Sorts files and highlights the mismatches



IMPLEMENT TEXTTRACT

Input PDF or Image Files



6. Create PDF



1. Upload

S3 Bucket



5. Generate CSV



2. Trigger

Lambda Layer



4. Analyze Text



3. Creates Job

Texttract

