Acclerate Document Processing with **Automated Text Extraction** Powered by Machine Learning

Although we live in an increasingly digital world, the need to process printed documents remains a critical component of every business. But manually collecting data from forms to enter it into a database is an expensive, tedious process prone to errors.
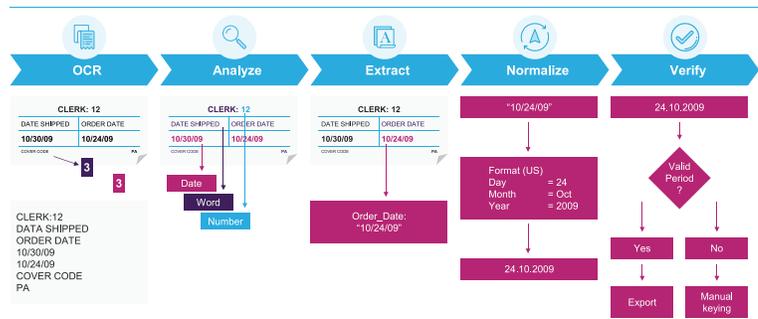
Automating the extraction of text and populating it into the appropriate field in adatabase is the answer, but it's a complex process. Every document, even those of the same type, has a different layout, so an automated text extraction system needs Machine Learning support to teach itself to recognize data so it can be added to the database correctly without human intervention.

BAL Global, one of the world's largest global immigration law firms, receives thousands of documents every day. The firm looked to SpringML to help develop an automated document processing system to extract data from physical documents so users could analyze the data in near real time. Additionally, the system had to support existing workflows so document processing could be accelerated without disrupting established review and approval processes.

SpringML designed a system in which documents were scanned and uploaded to Google Cloud storage as PDF files. The sytem could then be queried using the following data pipleine.

1. Document is scanned, converted to PDF, and then sent to Google Vision API in case the image needs enhancement
2. Several text processing rules are executed to validate text and log any errors
3. Files are stored in BigQuery and are now available for immediate analysis and reporting

B·A·L

- Global law firm specializing in immigration
- Rated "Best Law Firm" by U.S. News & World Report

"SpringML's document processing solution can enable us to augment our staff and processes with Machine Learning. The system delivers the capability to enhance our client service offerings through supplementing quality review processes with smart systems, while at the same time reducing paper processing time. This allows our legal teams to take action faster for our clients. This solution is versatile in that it handles several types of documents. It also provides details on data extraction accuracy, which allows us to be confident of the results."

—Vince DiMascio, CIO, BAL Global

For more information about SpringML's intelligent case routing solution, contact us at info@springml.com.

SpringML