Artificial Intelligence and Data Capture: A Necessary Marriage for Business Growth

From Siri and Alexa to customer–service chatbots and stock–trading Forex robots, artificial intelligence (AI) has fundamentally changed many aspects of the way we work.

Data capture is no exception.

Automated data capture technology already increases workplace efficiency and decreases business costs. "Intelligent" capture is even more dynamic, leveraging the power of AI and robotic process automation (RPA) to deliver additional benefits to enterprises.

Imagine feeding a batch of invoices into a scanner, stepping away, and letting an attached computer both file and prepare them so you only deal with the exceptions before paying the bills.

Truly intelligent capture software doesn't require templates, keywords, exact definitions, taxonomies, or indexing to get the job done. It can extract the right information and make sense of a wide variety of documents on its own, regardless of size, format, language, or symbols used.

Brainware
by Hyland

Four ways AI is changing data capture

With intelligent capture software, the AI-driven "engine" can be taught—exactly like a new employee – how to perform a data entry task. It can quickly pick up on contextual information and learn to interpret patterns and features in different document types. Moreover, it can validate data against existing systems, providing an additional layer of protection that employees can't duplicate without tedious manual lookups.

Intelligent data capture has changed the game for three major tasks: classification, extraction, and validation.

① CLASSIFICATION

With classification, also known as document sorting, the software learns to recognize different types of documents after the user gives it a few variations and examples. Just as a human can read through some sample documents and intuitively understand similarities and differences, the software doesn't need to see every single version of a contract or check request to recognize it. The machine learning engine cuts down on the rules that need to be applied, resulting in a high level of confidence in document classification with minimal manual effort.

② EXTRACTION

AI has done wonders for data extraction in semi-structured and unstructured documents. Take, for instance, invoice number identification, which usually involves building out complex templates and providing keyword tags and pairings around particular fields and labels.

A new employee can look at a document and immediately locate where invoice numbers are, no matter what the form looks like. Now software can do that, too, without the need for programming. The machine learning engine is trained to understand context, such as what an invoice number is not and what should (or shouldn't) be around the number, so there's a high degree of accuracy in the extraction.

Additionally, extracting data from complex stacked tables, with lines that don't match up (i.e., transcripts), is now a breeze. Mature AI software learns how to understand patterns and formatting, look for different types of information, and identify key data elements without the need for someone to rope and band the information. Only the exceptions would require human intervention.

③ VALIDATION

AI is a boon for tasks that go beyond the "scrape-the-page" approach, such as an advanced search capability that can validate extracted data from a document with existing information in another system. For example, it can match a line item in an invoice with purchase information stored in another system.

AI-driven search also allows for multi-way search, which means it can use multiple pieces of information (i.e., quantity, price, description, and amount) to match an exact item in the back-end system. And even if things aren't precise matches—say, an abbreviation is used in the description of the invoice, but not the back-end system—the software can deduce they are the same item.

Validating precise matches once required the efforts of a patient data-entry clerk. By minimizing this type of work, employees can now work on more interesting tasks, further driving efficiency and minimizing business costs.

> "Validating precise matches once required the efforts of a patient data-entry clerk. By minimizing this type of work, employees can now work on more interesting tasks, further driving efficiency and minimizing business costs."

> "It's important to note that the goal of AI-driven data capture isn't to replace humans, but to drive as much automation as possible with machines that can intelligently carry out tasks. Ultimately, employees are freed from being bogged down by mundane tasks and can take on high-value tasks that require a human mind to do well."

Working in tandem: Intelligent capture and robotic process automation

The enterprise RPA market is booming. So far, it's delivering on its promise of automating complex, rule-based processes, with Forrester projecting an overall market—of which document capture is just a slice—worth $2.9 billion by 2021, up from a mere $250 million in 2016. (That's 10x growth in five years.)

Unfortunately, RPA's processes fail when there's too much variance—and with documents, sometimes you have nothing but variance. For instance, if two documents have the same type of information, but are not in the same exact template, optical character reader (OCR) errors occur.

Enter machine learning. Combined with standard RPA concepts, the AI engine creates a **dynamic variance network**, or a way for software to look at everything in relation to everything else on a document. By calculating vectors between all words on the document and in the target fields, it overcomes OCR errors during the extraction process by using relative location to determine where to find a specific data field instead of a single, memorized location. On top of that, the software monitors the activity of the user, and will make corrections based on that behaviour without human intervention.

In other words, the system gets smarter on its own.

The real value of intelligent data capture

On top of the obvious automation benefits, using intelligent data capture software eliminates guesswork on the setup and programming side as well. But it's important to note that the goal of AI-driven data capture isn't to replace humans, but to drive as much automation as possible with machines that can intelligently carry out tasks. Ultimately, employees are freed from being bogged down by mundane tasks and can take on high-value tasks that require a human mind to do well.

One thing is certain: successful companies aren't built in a templated world. Information and documents continually change, and they must learn and adapt—ideally, with technology that does the same.

Learn more at **Hyland.com/Brainware ››**

Brainware™
by Hyland