Intelligent Automation Eliminates Manual Data Entry From Complex Documents

A new breed of recognition technology can automate the processing of complex documents that contain printed and cursive writing.

a2ia
access more

There is an interesting theme that runs throughout the Internet documentary "Humans Need Not Apply": Technological automation doesn't need to be perfect, it just needs to be better than the humans it is replacing. And, it turns out, that's a benchmark that is being met and exceeded in more areas than one could have previously imagined. For instance, some paper-intensive workflows were once thought to be so complex that they required human labor to manage them. That's no longer the case.

Extracting Handwritten Data From Documents

The business world realized great efficiencies and cost savings when OCR (optical character recognition) technology automated the processing of standard forms containing computer-generated text. What about nonstandard forms and cursive handwriting? Similar to the pre-OCR days, most consider expensive human labor to be the only way to effectively manage this type of unstructured data. This isn't true. As the documentary points out, "Just as mechanical muscles made human labor less in demand, so are mechanical minds making human brain labor less in demand." In terms of processing challenging documents and forms, this means automation using OCR and IWR (intelligent word recognition) technology.

Often, however, the industries that could most benefit from this type of automation – healthcare, insurance, finance, and government – still see manual data entry to be the only way to handle complex forms and handwritten information. In addition to data extraction, classification or document sorting can be labor-intensive as well. Oftentimes this needs to be performed in conjunction with, or prior to, recognition. This cumbersome process can be seen during medical coding in healthcare. Documents must first be categorized into various classes (e.g. lab results, administrative forms, healthcare provider notes), and then keywords or phrases are located and extracted – regardless of whether the information is printed or handwritten.

A2iA's classification technique is unique to the market in that it uses artificial intelligence to look at both layout and content to determine a document's type. This allows for workflows across all industries – from medical coding to mailrooms – to speed their access to data by quickly and automatically sorting, extracting, and routing information to the correct location based on a holistic methodology.

Recognition Starts With The Document Type

It's a fruitless effort to try and identify a common set of documents that cut across all industries that could benefit from being automatically processed using OCR and IWR technology. Simply put, documents among companies (let alone entire industries) vary greatly. No matter the industry, however, it's easy to identify areas of potential automation. The obvious opportunities are any processes where a human intervenes to perform any type of manual data entry. In the insurance industry, this might be EOB documentation

or claims processing. For finance, it might be processing lockbox documents or columnar extraction. And, the opportunity to automate the customer onboarding process is shared by most industries.

Industries will also be able to leverage new technology from A2iA that will be performed for both printed and cursive data extraction using the same recognition engine. The technology also eliminates the need for a dictionary, which has traditionally been a requirement for cursive extraction. This offering uses new RNN-based technology developed in-house by A2iA's R&D team. It provides full-text transcription, and it's the first cursive Arabic-recognition engine available. From here, organizations can utilize A2iA's transcription output to enhance various in-house solutions such as existing classification software.

Automation Advantage Can't Be Ignored

Eliminating manual data entry from common paper-based workflows brings some obvious advantages. Chief among them is that recognition software is ultimately less expensive than the human labor it is replacing. Also, recognition software is more accurate, doesn't get tired, take vacation or need an HR department to manage it. But, an automated processing solution also brings a higher level of security by eliminating privacy breeches caused by humans handling sensitive data. Additionally, if third-party processing is being utilized, organizations can be limited as to where the documents can actually be processed (i.e. offshore versus United States).

The paper-intensive markets mentioned previously continue to underutilize data recognition technologies. Opportunities still abound to use OCR technology to automate the process of basic printed information. When it comes to leveraging IWR to process more complex documents and cursive writing, the surface is only now being scratched.

While the opportunities for automation may be vast, it's understandable that users are still skeptical when it comes to cursive recognition and complex data extraction. There is a common question asked by all potential users, "Does it really work?" The short answer is, "Yes." Test the software with your documents to validate its accuracy and efficiency. Then, calculate how many hours of manual labor can be saved and how many documents can be processed with the automation software. Compare those numbers to the output from manual processes. The results can be dramatic (especially in high-volume applications). A2iA has customers that have said automation software has saved them months' worth of time annually by simply using the software instead of manual labor.