Today's Capture & IDP Industry

The AI revolution in document capture is here

by: Ralph Gammon, Infosource Software

Over the past decade we have seen artificial intelligence (AI) go from something people were scared of (partially thanks to movies like "2001: A Space Odyssey" and "The Terminator") to something they now look to leverage in everyday life (but are still kind of scared of). During this time, the document processing industry has seen a steady increase in the amount of AI used to automate classification and data extraction.

Early AI implementations were simply billed as "advanced" or "intelligent" capture. They typically were extensions of traditional OCR-based applications, with vendors utilizing learn-by-example algorithms to supplement their offerings. As cloud computing reached a tipping point in the mid-2010s, it opened up accessibility to a number of open-source AI engines. Training these engines for document classification and data extraction was an early killer app for applying AI to business needs. This launched the intelligent document processing (IDP) industry.

Over the past couple of years, we have seen the rise of generative AI, which utilizes large language models (LLMs) to perform a variety of tasks. As inferred by "large," LLMs can be trained on hundreds of billions of parameters, which are essentially variables in neural networks that can be modified through training. The more parameters there are, the more complex the model, which means it requires more processing power — but it also means that the LLM can be trained to do some pretty cool things.

Generative AI is one of those cool things. We have all played with generative AI to answer questions, create images, help with coding, etc. LLMs can also be utilized to do some rudimentary document capture and IDP, such as automatically finding and returning matching pairs on a form and summarizing content.

This has led to some upheaval in the capture and IDP software market, as vendors have been scrambling to implement generative AI technology into their product sets. Here are some initial iterations we have seen:

(1) Integrating publicly available LLMs for basic functions like entity extraction and summarization — This

was the first application and what we primarily saw in 2023.

(2) Security and privacy factors considered — When people started using generative AI and LLMs for business applications, security and privacy concerns were immediately raised due to them being hosted on public clouds. We have seen these concerns addressed through a number of techniques, including guardrails that can be put into place so private data will not be used for training public LLMs.

(3) Co-pilots — This is a generic (and sometimes branded) term that independent software vendors (ISVs) are using to describe generative AI assistants. One of their functions is to enable interrogation of documents with natural language prompts. Co-pilots are also increasingly being used to set up document workflows.

(4) Setup, training and validation — Initially, we saw some AI technology used to create dummy documents that IDP users could train their software with. Generative AI offers to take this to the next level. First, it can automatically detect fields and pairs on real documents and help set up the labeling. It is not perfect, but this technique can significantly reduce initial setup times. We are also seeing an increasing use of LLMs in the validation steps of IDP to reduce the amount of human intervention required.

(5) Document-capture-specific and IDP-specific LLMs — These seem like extended versions of the foundational models that most of the IDP industry is based on. For example, one vendor recently announced an LLM created specifically for transactional documents.

(6) Vertical-specific LLMs — These are versions of LLMs trained on specific sets of vertical data so they better understand context and jargon specific to a market. These are just starting to emerge.

(7) LLMs trained on your own documents — Microsoft is moving forward in this area with its Copilot for Office 365, which can apply generative AI to a user's Microsoft ecosystem.

(8) Managing document workflows — As AI lowers the barrier to entry for utilizing basic capture and IDP functionality, capture's tight incorporation with multifaceted processes becomes more important. We are now starting to see capture and IDP vendors leverage their expertise with AI and their ability to understand content to try and automatically drive workflows.

As AI lowers the barrier to entry for utilizing basic capture and IDP functionality, capture's tight incorporation with multifaceted processes become more important.

Pumping the Breaks (a Bit at Least)

Talk about generative AI is everywhere. LLMs are very powerful and they are going to have wide-ranging influence on the capture and IDP market going forward. But today, it is important to keep a couple of things in mind:

(1) While generative AI may be able to do a certain level of capture and IDP out of the box, it by no means offers a complete application. Generative AI still struggles with elements like confidence levels and hallucinations, which can lead to damaging "false positives" in data capture.

(2) There are still security and cost efficiency concerns. In the long run, if you are running the same process on your documents over and over, using an LLM could be overkill. This is why capture and IDP vendors are still important in the market. They have been working with tools like OCR and other forms of AI for several years. They have built out business and enterprise-level software applications leveraging these technologies. While generative AI certainly has a place in the capture and IDP market, determining how to best utilize it today and going forward is a complex task.

Infosource recently published its "2024 Infosource Global Capture & Intelligent Document Processing (IDP) Vendor Matrix Report," which ranks and discusses the capabilities of the market's leading vendors. What we have seen is that with the infusion of AI technology, capture and IDP software is becoming more powerful and easier than ever to deploy. Businesses that are not taking advantage of it on some level are missing out on efficiency gains.

As a participant in the BTA Channel, document processes should be very familiar to you. Utilizing capture and IDP to make these processes move more quickly and with less human effort can save costs and reduce headaches and paperwork, and these are benefits everyone is interested in. Please reach out to me if you would like to learn more about the "2024 Infosource Global Capture & Intelligent Document Processing (IDP) Vendor Matrix Report" or other matters related to capture and IDP. ■

Ralph Gammon is senior analyst for Infosource Software and the primary author of Infosource's latest report, "2024 Infosource Global Capture & Intelligent Document Processing (IDP) Vendor Matrix Report." Infosource has been covering market sizing, trends, opportunities and leaders in the capture and IDP software market for more than 20 years. From 1998 to 2020, Gammon served as editor and publisher of the Document Imaging Report, the premier insider newsletter covering the ECM/content services space. He can be reached at rg@info-source.com. Visit www.info-source.com.