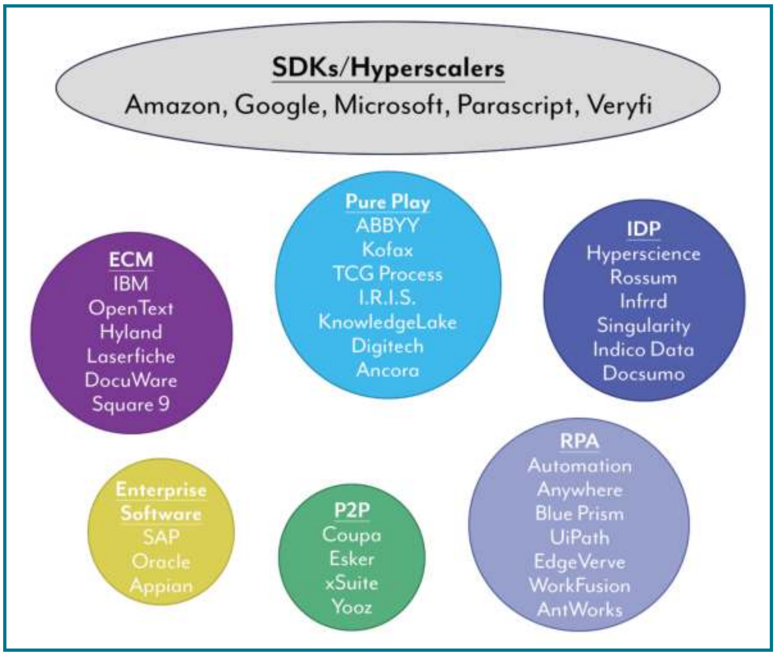
Applying AI to Capture

Carving out a successful IDP strategy

by: Ralph Gammon, Infosource Software

Infosource Software is currently compiling its annual ranking of vendors to be published in its 2023 Capture & IDP Vendor Matrix. Infosource (and, previously, HSA, which Infosource acquired in 2000) has been ranking capture software vendors and analyzing the market for more than 20 years. However, what used to be a limited set of vendors competing with a limited set of applications has exploded in the past five years, primarily due to the introduction of AI into the capture software space.



What exactly is capture software? Infosource defines it as applications that understand and extract meaningful, accurate and usable information primarily from unstructured and semistructured input. This input can include scanned and electronic documents (such as PDFs), as well as fax and other electronic inputs like social media feeds, voice messages, video and still images. Capture acquires, classifies and converts information into enhanced usable data for business transactions, analytics, records management, discovery and compliance applications.

Capture software has been around for a long time. It first emerged as a way to onboard images into electronic document management systems. Optical character recognition (OCR) was applied to help index these images, and then utilizing OCR to capture from imaged forms became a market in and of itself. Eventually, the data- and image-capture markets merged, and more advanced recognition technology was applied to enable capture of data from semistructured forms like invoices, as well as to enable auto-classification of mixed document sets. That is where the capture market sat until about five years ago, when intelligent document processing (IDP) showed up.

At its heart, IDP is the application of AI and machine learning to capture data from documents. The previously mentioned advanced recognition applications also utilized AI and machine learning to some extent, but IDP has taken it to a whole new level. The impetus behind the birth of IDP was a combination of more powerful processors to run AI algorithms (including better and less expensive cloud computing options) and more AI expertise and algorithms coming to the fore.

At first, we heard about a couple of AI specialists who decided to use document processing as a proving ground for their technology. Once they started having some success — oftentimes in applications that had been deemed impossible with legacy capture solutions — we started to hear about more and more AI specialists launching dedicated IDP businesses.

Because it was one of the first practical applications of AI, funding soon followed. From 2019 through 2021, we saw venture capital pour into IDP vendors at a rate we had never seen the likes of in our 20 years of covering the capture market. A few companies even raised more than $100 million. This is not even considering well-funded RPA vendors that quickly recognized the parallels between IDP and their technology, which is designed to move structured data between computer systems.

The bottom line is that a market that once featured less than 50 vendors has now expanded to several hundred. See the graphic on page 16 for an idea of what today's capture and IDP market landscape looks like.

... The implementation of AI is going to have far-reaching effects on businesses going forward ... IDP is a great place for you to start your AI-powered journey.

The vendors in the sky blue (Pure Play) and purple (ECM) bubbles are the two you are probably most familiar with if you have played in the capture and ECM market. Working our way counterclockwise, we are all familiar with the leading enterprise software vendors who have historically participated in capture on some level either through their own applications or through partnerships. The P2P vendors focus on accounts payable and procure-to-pay applications with some specific integrations and workflows related to that.

RPA vendors entering the capture market, as well as IDP startups, are listed in the bubbles on the right-hand side of the chart. It is no coincidence that this then brings us back to the Pure Plays, due to the close relationship between capture and IDP.

What Makes a Top Capture & IDP Vendor?

What are we looking for in a leading capture and IDP vendor and what should you be looking for if you are considering a partner in this space? We consider many factors, including vision, technology stack, market share, partner channel and more. Boiling it all down for you, here are three major factors that I consider most important to becoming a "star" in our rankings:

(1) Size — Most of our leaders are names you might be familiar with and they appear in the ECM and Pure Play bubbles. These companies have large installed bases and support teams, experience in and an understanding of the market, and established partner programs. They also have resources to develop their own IDP technology and most of them have done just that, so far eschewing the opportunity to add more AI and machine learning through the acquisition of IDP startups.

(2) Cloud strategy — Historically the capture market has been slow to transition to cloud services, with perpetual software licenses outselling cloud services more than four-to-one as recently as 2018. But a number of factors — including the growth of IDP, the COVID-19 pandemic and an overall greater trust of the cloud — reduced that ratio to less than two-to-one in 2021. That said, a significant portion of capture and IDP sales still go into security-sensitive markets like financial services and government, so there will always be some demand for on-premise solutions. Software as a service (SaaS) is a great way to efficiently deliver capture and IDP, but it is not for every application. Some sort of hybrid strategy is probably optimal going forward if the vendor has the resources to support it.

(3) Intelligence — There is a reason IDP has become such a hot topic. It really does some cool stuff, like expanding the types of documents automated processing can be applied to and helping reduce notoriously long and cumbersome setup times — both for new implementations and for adding new document types to existing applications. IDP has helped bring the concept of low code/no code to the capture market, which is an important piece of the technology stack going forward.

Stay Calm & Embrace the Future

If the introduction of IDP into the already dynamic capture and ECM/content services space seems a bit overwhelming, do not panic. I was speaking at a conference less than a year ago to a room full of document management technology users; when I asked how many knew what IDP was, no one came up with "intelligent document processing." It is still very early in the IDP adoption curve, so there is plenty of time to get up to speed.

That said, if you are working with a capture vendor, you want to verify that it has an IDP strategy because it is the way of the future. At the very least, it expands and improves what capture can do. At the extreme, it can serve as an assist for your customers in automating the processing of all their incoming communications, saving them time and money, and freeing up resources to do more important things.

As I think we are all beginning to see, the implementation of AI is going to have far-reaching effects on businesses going forward. As an expert in the document world, IDP is a great place for you to start your AI-powered journey. ■

Ralph Gammon is a senior analyst for Infosource Software. In this role, he covers market sizing, trends, opportunities and leaders in the capture and intelligent document processing software space. Gammon joined Infosource in 2017 as regional manager for the Americas. From 1998 to 2020, he served as editor and publisher of the Document Imaging Report, the premier insider newsletter covering the ECM/content services space. Gammon can be reached at rg@info-source.com. Visit www.info-source.com.