

2026

Global Capture & IDP Software

Vendor Matrix Report

(Featuring KnowledgeLake)

By Ralph Gammon
Senior Analyst
Infosource Software

July 2026

Contents

Introduction.....	3
Market Definition	3
Market Dynamics.....	4
Agentic Automation: The Next Frontier.....	5
The Ideal Solution	6
Methodology	6
Criteria Used for Evaluation	6
Strategy and Capabilities (y-axis)	6
Execution in the Market (x-axis)	7
2026 Infosource Capture & IDP Vendor Matrix	8
Matrix Review	9
Stars.....	9
Disruptors.....	9
Competitors	9
Explorers	10
Featured Vendor Profile.....	11
KnowledgeLake 	11
Selected Competitor Summaries	12
ABBYY ★	12
Iron Mountain 	13
Kodak Alaris 	13
Tungsten Automation ★	13
About The Analyst	14
Copyright	14

Tables and Figures

Figure 1: Global Capture & IDP Market by Use Case.....	3
Figure 2: Infosource Market Maturity Model	5

INTRODUCTION

Welcome to the 2026 edition of the Infosource Capture & IDP Vendor Matrix Report. It features an evaluation and ranking of more than 25 leading ISVs in the market. Each vendor was evaluated on the strength of its products, as well as its vision and execution in the market.

Market Definition

Infosource defines Capture & IDP software as addressing the ingestion and processing of business inputs, which includes documents and can extend to other media types (voice, video, SMS, etc.). The software does the following:

- acquires
- classifies
- converts and enriches
- validates

business inputs into data for use in business transactions, records management, and case management applications.

For 2025 (based on our 2024 vendor revenue numbers), we estimated the Global Capture & IDP market to be worth approximately \$9B, with EMEA accounting for just over 40% of that and North America approximately 35%. Accounting and Case Management are the two major use cases, each making up about a third of the global market.

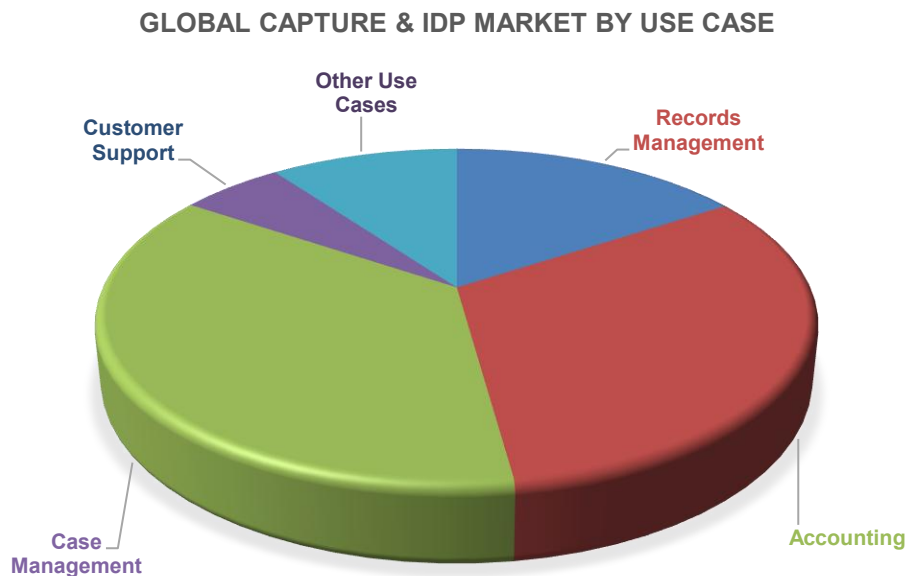


Figure 1: Global Capture & IDP Market by Use Case

Market Dynamics

AI has been the driving force behind the Capture & IDP market since the mid-to-late 2010s, when ISVs started utilizing open source machine learning models and training them on document processing—typically classification and extraction. Large Language Models (LLMs) were introduced into the equation in the early 2020s, and they offer capabilities like the ability to process documents without any training and to understand, on some level, completely unstructured documents. LLMs can also be used for functionality like summarization and ad hoc interrogation of documents via Q&A. On the downside, their token-based pricing models can be unpredictable and expensive, LLMs make mistakes (commonly referred to as hallucinations) and confidence levels are hard to gauge. Also, because of the enormous amount of processing power they utilize, LLMs are typically run on public clouds, which raises security and privacy concerns from some users.

LLMs in IDP	
Positives	Negative
Process documents without training	Hallucinations
Can be applied to unstructured and complex documents	Hard to gauge confidence levels
Enable document interrogation (Q&A)	Unpredictable and/or potentially expensive pricing
Able to produce summarizations	Typically run on public clouds

LLMs can also be used for features like expediting IDP workflow setup and assisting with validation. In addition to LLMs, we are seeing vendors implement vision-language models (VLMs) and multi-modal models, which enable them to process documents and other types of input forms more holistically, by being able to understand and synthesize various elements such as text, images, and audio and video within a single model. We are also seeing the introduction of Small Language Models (SLMs), which can run on-prem and be trained with domain-specific data.

There are still a lot of elements that go into Capture & IDP solutions in addition to AI models. The foundation for the market was laid more than 30 years ago with the introduction of document scanning, OCR, and template-based document extraction. These elements are still in place in many applications. On top of that have been built techniques like the application of data rules and regular expressions, which are also still in use today, many times in conjunction with AI technologies.



Agentic Automation: The Next Frontier

Here is a look at the current Infosource Capture & IDP Market Maturity Model:

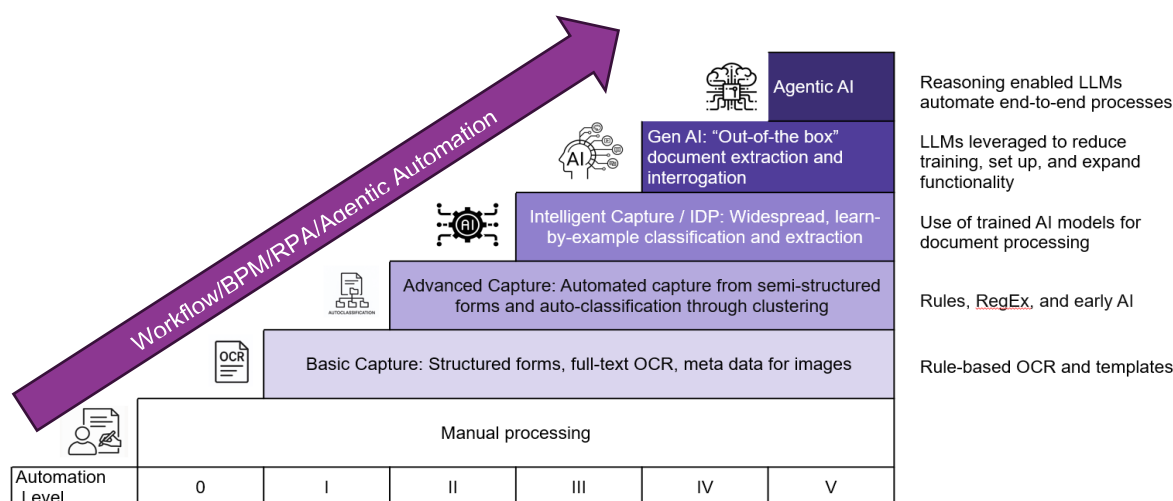


Figure 2: Infosource Market Maturity Model

At the top is Agentic AI, which we have not discussed yet. Agents are the hottest topic related to business process automation, a broader market, which typically incorporates IDP. Technically, an agent refers to a system or program that can autonomously perform a task or an action, either as part of a business process, or it could potentially be automating the entire business process. Capturing and posting data from an invoice, for example, could be considered an agentic task. Automating an entire purchase-to-pay (P2P) process, from placing an order to reconciling the shipping and receiving data, to processing and paying the invoice, could be considered an agentic process. Since we already have a specific term for an IDP transaction, when we speak about agentic AI, we are typically talking about automating processes, as in the P2P example.

Automating document-driven processes is nothing new. Since the dawn of scanning, we've had workflow software to complement capture. In document-centric applications, workflow is designed to make sure documents are routed to the correct people. Process-centric Agentic AI can be considered the next generation of document workflow where IDP is involved. It offers functionality like recommending the next step that should be taken or even automating the taking of that step based on previous actions and ancillary information, such as can be found in Retrieval Augmented Generation (RAG) databases or other types of content stores. The evolution of LLMs to incorporate more reasoning capabilities is supposed to further empower Agentic AI applications.

While I would say Levels I-IV on the maturity model are already in widespread use, process-automation utilizing Agentic AI is still in its early stages in the mission-critical applications that IDP typically addresses. There is still plenty of human supervision, and we are at least a few years out from the vision we've heard

of agentic-first businesses, where processes are designed with automation as the standard and people just fill in where needed.

The Ideal Solution

Infosource's description of an ideal Capture & IDP solution says that it should have the ability to automate processing of inputs required for business transactions, including unstructured and semi-structured input from any source, in any format. In addition, an IDP solution should be able to classify this input and extract relevant data, then validate and augment that data and apply it to the appropriate systems and workflows, which can involve initiating or advancing a process.

Scalability and accuracy are important. We also factor in the ability to utilize AI and machine learning to reduce implementation time and effort and to broaden automation. A containerized cloud services architecture can be applied to increase flexibility in deployment, as well as improve integration with other platforms. Ideal Capture & IDP solutions should enable end-to-end automation of transactional processes and can also address compliance and analytics.

It's important to note that while AI and agentic automation are definitely the hottest topics related to the Capture & IDP market right now, in the big picture, AI and even LLMs are enabling technologies that assist with automation. And while agentic automation may be an end goal that IDP can contribute to, an IDP solution can stand on its own as an agentic task as well. Just as not every Capture vendor offered workflow, not every IDP vendor is going to provide end-to-end process automation facilitated by LLMs.

METHODOLOGY

Vendors included in this year's matrix have been ranked in a number of individual categories, with their overall rankings a product of the sum of their scores in each category.

Criteria Used for Evaluation

Following is a look at the categories we used to evaluate the vendors for this year's Capture & IDP Vendor Matrix.

Strategy and Capabilities (y-axis)

- **Vision for Capture & IDP:** Even though well less than half the vendors surveyed are today offering what we would consider true agentic process automation, in five years' time, a number envision themselves offering a solution in which document classification, extraction, validation, self-correction, and action are all part of continuous flow. In other words, the application of Agentic AI seems to be the predominant goal of vendors in the market.
- **Recognition Technologies:** This is the first time we have included AI as part of the "engine" mix, along with OCR/ICR and other enabling technologies. AI is pretty versatile and also incorporated in several other categories.
- **Complex Documents:** An area where a lot of AI is used, this category accounted for processing both long and complex documents and the ability to reconcile information across document sets. A surprising number of vendors have good tools and strategies to address this.
- **UI for Document & Agentic Automation:** We've had a category for designing document workflows for several years, which we've extended to include design for agentic automation, since that seems to be the way many vendors are going.
- **End-to-End Automation:** This includes integration with scanning, workflow, BPM, ECM, RPA, ERP and other line of business apps. Some vendors have their own technology in these areas, but strong connectors, APIs, and partnerships are another way of addressing this.

- **Agentic AI:** I threw this out there for the first time, kind of to see how vendors are defining the term themselves. A lot of what came back was discussion on leveraging AI to automate multi-step processes. One interesting note is that these processes can be kind of elastic. For example, we have some vendors with a very narrow focus, like automating document processing around shipping and receiving or A/P in construction, which is a valid Agentic AI strategy. Then we have larger vendors with platforms in place to automate multiple types of complex processes, also a valid strategy.
- **RAG Capabilities:** Because it involves converting documents into “chunks,” meta data and other information that can be stored in a vector database, I’ve always considered the ability to create RAGs (and similar sets of data to be used for grounding LLMs activities) to be a sister to the traditional Capture & IDP market, even though it requires some different tools. We received a range of responses, from ISVs doing full-blown RAG creation, to those leveraging RAGs to ground their agentic processes, to those doing nothing. This category didn’t have the highest weighting but it’s interesting to follow. Perhaps it will be considered as part of End-to-End Automation and/or Agentic AI going forward.
- **Architecture:** This category combines deployment options, scalability and security. Most vendors in the market have done a pretty good job addressing scalability and security with deployment options a mixed bag.
- **Capture:** This incorporates the ability to ingest input from a variety of sources in a variety of formats, although since the IDP market is still primarily about documents, they take precedence. The ability to connect to, and control, scanners was also considered as there is still a lot of paper input, especially in higher volume, production environments.

Execution in the Market (x-axis)

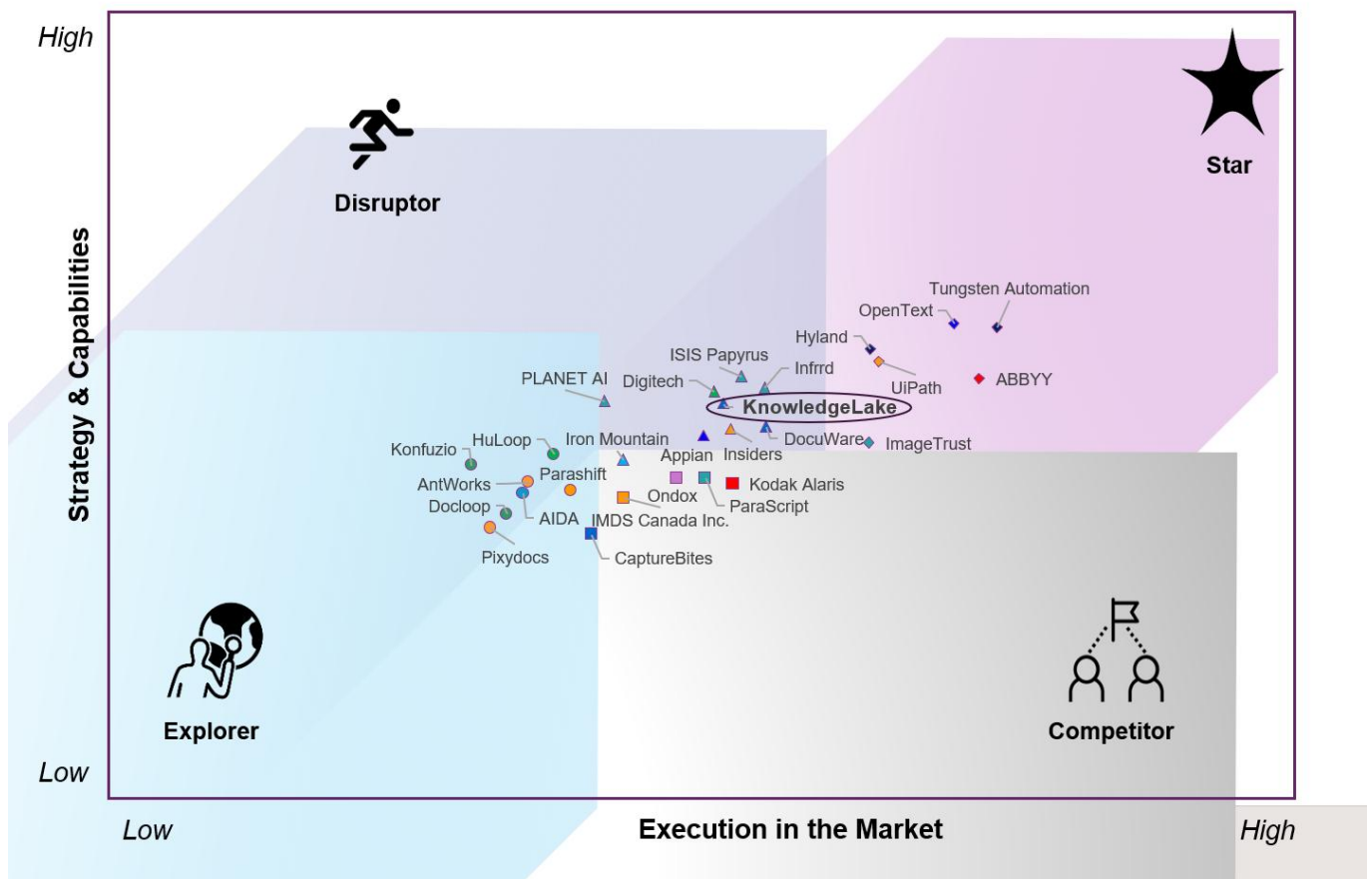
- **Competitive Differentiation:** Basically, we wanted to know what makes a vendor stand out from the competition in the Capture & IDP market, which currently includes several hundred vendors. As you might expect, we received a range of answers here. We also asked for an example of why each vendor won in a competitive bid against another vendor, but, somewhat surprisingly, the majority of respondents declined to answer this.
- **Size of Customer Base:** When considering Execution, size is important, but we weighted it less this year than in the past. Part of that has to do with the fact that even the biggest vendors in the IDP space are now at some level competing against LLM giants and being an enterprise software vendor (as opposed to a hyperscaler) is maybe not the advantage that it once was.
- **Geographical Strength:** This was considered in a couple different ways. If it was a larger vendor with a global footprint, then, how far and wide is their Capture & IDP software installed? Conversely, a strong focus leading to success in a particular region is also considered an advantage.
- **Vertical Strength:** As agentic automation extends IDP deeper into processes, a focus on going deeper instead of broader into targeted vertical markets is increasingly advantageous.
- **Demand Generation:** There are a lot of different techniques utilized to increase visibility, create qualified leads and drive sales in an increasingly crowded market where vendors are competing not just against other IDP vendors but also against do-it-yourself LLM initiatives. The bottom line is how effective these initiatives are in driving new sales.
- **Partner Strategy:** This is a combination of reseller/go-to-market partners and technology partnerships to enable end-to-end automation. Partnerships are becoming increasingly important in the Capture & IDP market as it expands into agentic automation, and vendors have to take their applications deeper into their customers’ processes. Partners provide a combination of vertical expertise and customer familiarity, as well as complementary technologies.



- **Ease of Use and Customization:** We formerly just called this “ease of use” figuring that was an advantage, but the ability to customize software to meet the demands of complex applications is also advantageous. Ideally, a software platform should be flexible and easy to use. We know that is not often the case – although vendors are moving that way with the help of GenAI.
- **Deployment Options:** We used to consider it an advantage if a vendor offered a true SaaS platform, and it still can be, but for larger vendors, or those in niche markets, on-prem and/or hybrid can be the best way to go to market.
- **Case Studies and Customer References:** I was somewhat disappointed in the quality and quantity of case studies available showing truly modern and innovative IDP applications. I also understand a lot of the customers doing really high-level stuff want to keep it quiet for competitive and security reasons, so I credit any vendor who is able to get quality case studies and even make references available for analyst access.

2026 INFOSOURCE CAPTURE & IDP VENDOR MATRIX

2026 Infosource Global Capture & IDP Vendor Matrix



Matrix Review

Stars

This year's Matrix features six Stars, four of which return from last year:

- **ABBYY** continues to develop a strong portfolio of recognition engines, which it complements with LLM technology. Its large global customer base and strong partner channel help it on the x-axis.
- **Hyland** is a document-focused content services market leader that has a strong portfolio of Capture & IDP technology. Its SaaS IDP application is a key piece of its agentic automation platform.
- **ImageTrust** makes its Matrix debut as a Star. It is a flexible Capture & IDP application developed by Image Access, sold primarily through partners, and targeted at production environments.
- **OpenText** continues to be a leader in the market, with a strong global customer base and impressive end-to-end and agentic automation capabilities.
- **Tungsten Automation** returns as a Star after not participating last year. Tungsten has a large global customer base and partner channel, as well as a deep and wide breadth of Capture & IDP technology.
- **UiPath** is an RPA market leader that has grown its IDP business as a way of processing unstructured content. UiPath is heavily invested in agentic automation.

Disruptors

This year's Matrix features nine Disruptors who are stronger on Strategy and Capabilities, while pushing forward with Sales and Marketing Execution to become Stars:

- **Appian** is a BPM market leader that is growing its IDP business as part of a comprehensive AI-driven automation strategy.
- **Digitech** is a long-time Capture and ECM player primarily in the mid-market, whose innovative Sys.tm platform, which incorporates multiple LLMs for IDP, could be a game changer.
- **DocuWare** is featured for the first time. Primarily known as a mid-market content services leader (owned by Ricoh), it recently acquired an IDP vendor, whose technology has been integrated into its platform.
- **Infrd** has a strong IDP and agentic automation platform specializing in complex document-driven processes, such as mortgage processing and managing engineering drawings.
- **Insiders** is a long-time leader in the DACH Capture & IDP market that was recently acquired by a German ERP and agentic automation specialist.
- **Iron Mountain** continues to advance its IDP capabilities and is extending into agentic automation through its InSight DXP platform.
- **ISIS Papyrus** has an end-to-end document automation platform that it is increasingly augmenting with more AI technology.
- **KnowledgeLake** is a Capture & IDP SaaS pioneer that has expanded its platform. It was recently acquired and brought in a new CEO to help accelerate growth.
- **PLANET AI** is a recognition technology specialist that recently expanded into agentic AI with some targeted process management applications.

Competitors

Competitors have a strong presence in the market and acceleration of their vision and technical innovation will advance them on the y-axis. This year's Matrix features five Competitors:

- **CaptureBites** is a mid-market Capture & IDP specialist that sells through a network of partners, primarily in Europe.
- **IMDS Canada** is being featured in our matrix for the first time. It is a production Capture & IDP specialist with its own multi-language recognition technology.
- **Ondox** is a digital mailroom and document automation specialist that has its own IDP technology and is also a leading Tungsten Automation partner.
- **Kodak Alaris** is a leading global scanner vendor with complementary Capture & IDP software that is gaining adoption.
- **Parascript** is a recognition technology specialist known for handwriting recognition, but it also offers a full suite of IDP software, as well as market leading fraud detection technology.

Explorers

Explorers have intriguing technology and are just starting to gain momentum in the market. We have seven Explorers in this year's Matrix:

- **AIDA** is an IDP and agentic automation application developed by Italy-based Technology Cognition Lab, S.r.l. The product is growing its footprint through partnerships throughout western Europe.
- **AntWorks** specializes in high-volume complex document sets in the insurance and financial services spaces. In 2025, it was acquired by GTT Data Solutions, an Indian systems integration firm.
- **Docloop**: Docloop is an IDP and agentic automation vendor with a SaaS platform focused on providing trade document intelligence for the shipping and logistics industries.
- **HuLoop** incorporates IDP as part of a broader work optimization platform. It has had some success focusing on the mid-sized banking market.
- **Konfuzio** is an IDP product developed by Helm & Nagel, a Germany-based specialist at using AI software to automate processes. Konfuzio is focused on meeting data sovereignty requirements that are important in regulated industries in Germany.
- **Parashift** is a SaaS vendor that specializes in no-code document classification and extraction. To date, it has gone to market primarily through partners in western Europe.
- **Pixydocs** is a SaaS application developed by Zerapix. It is especially strong in accounts payable (AP) processes and is currently being targeted at the construction industry.

Note: All vendors featured in this year's Matrix report responded to an extensive questionnaire, answered follow-up questions and in many cases set up briefings to further explain their product and go-to-market strategies.

FEATURED VENDOR PROFILE

KnowledgeLake

Corporate Name & HQ: KnowledgeLake, St. Louis, MO

Capture & IDP Product: Intuitive AI

KnowledgeLake offers a full suite of SaaS automation and content management tools, headlined by Capture & IDP. It brings these offerings to market on two fronts: the KnowledgeLake Platform for enterprise and government customers and DocumentOS, a bundle aimed at BPOs that includes professional services and metrics reports.

KnowledgeLake was a pioneer in SaaS Capture & IDP and has subsequently fleshed out its suite with the addition of RPA, content management, and workflow technology. KnowledgeLake has a longstanding partnership with Microsoft, which has recognized the ISV as a leading partner multiple times. Its software is hosted on Azure, and KnowledgeLake leverages AI from both MS and OpenAI.

In 2025, KnowledgeLake was acquired by the equity investment firm Edison Partners, and in 2026 a new CEO was appointed to drive accelerated growth. To date, KnowledgeLake has sold predominantly direct and in North America, but has plans to extend its reach through an expanded reseller channel.

Vision for Capture and IDP: KnowledgeLake's Capture & IDP currently enables classification, extraction, and validation of document data at scale through a human-AI collaboration model. It combines zero-shot LLM-based classification with QC tools and built-in reporting — without requiring specialized AI or machine learning expertise. Over the next five years, KnowledgeLake's vision is to evolve the software from a document processing tool into a self-improving system of intelligence, where agentic AI catches and corrects issues before they reach human operators, with corrections feeding back into the AI model to continuously improve accuracy. The platform will proactively guide customers with forecasting, anomaly detection, and AI-assisted project onboarding. The end state will be a platform that runs itself — reducing manual labor, eliminating reliance on external tools, and requiring minimal KnowledgeLake involvement to deploy and manage new projects at scale.

Areas of Strength

- **End-to-End Automation:** KnowledgeLake offers a full suite of SaaS content management functionality, with each piece available as a discrete service.
- **Architecture:** KnowledgeLake is delivered 100% through a SaaS platform built on Microsoft Azure. The platform runs in a microservices architecture orchestrated on Kubernetes, allowing individual services to scale independently based on demand. The workflow engine uses a worker-based model, so workloads (including IDP) scale horizontally. The platform inherits Azure's underlying infrastructure, including availability-zone redundancy and managed database replication.
- **Agentic AI:** KnowledgeLake's AI strategy extends its platform with agents that act on documents and processes—under human oversight. Planned capabilities include project-setup agents that identify and resolve onboarding issues; agentic exception reduction; operator-side agents embedded in the QC experience to assist reviewers with common correction patterns and root-cause fixes; and cross-project pattern detection that proactively surfaces quality and throughput anomalies before they hit a human queue. KnowledgeLake's approach prioritizes bounded

autonomy over black-box automation, where customers retain visibility and control through validation interfaces, dashboards, and audit trails.

- **Capture:** KnowledgeLake offers direct scanning integration, supports import from network locations, and has direct email connectors (as well as add-ins for Outlook). File system connectors and import plugins support folder-based ingestion and KnowledgeLake has a REST API to support integrations from external systems.

Areas for Potential Improvement

- **Partner Strategy:** Although KnowledgeLake has architected its product to be channel friendly by enabling licencing and resale of any of its discreet services separately, historically only a small percentage of sales have come through resellers.
- **Vertical Strength:** KnowledgeLake's package for BPOs seems to be gaining traction in the market. Especially as KnowledgeLake moves deeper into agentic automation, it should consider developing focused packages for the other industries it focuses on (including state and local government and education), as well as with partners focused on additional markets.

Overall Outlook

For several years we have ranked KnowledgeLake high on the y-axis (Strategy and Capabilities) as it has been a pioneer in SaaS Capture & IDP and has effectively built out its technology stack while maintaining a viable go-to-market strategy. However, its growth has been limited by factors like its primarily direct sales model and its inability to expand outside North America. After the acquisition last year, the buyer, Edison Partners, has pledged to accelerate KnowledgeLake's growth trajectory.

KnowledgeLake has a strong platform that is designed to scale. It will be up to the new management to implement a sales and marketing strategy to drive corresponding growth. Based on its close work with Microsoft and strong vision for document-centric agentic automation, we expect KnowledgeLake to continue to rank high on our y-axis, and if it can nail down a strategy for reaching a broader market, eventually graduate to a Star level in our matrix.

SELECTED COMPETITOR SUMMARIES

ABBYY ★

Corporate Name & HQ: ABBYY, Austin, TX

Capture & IDP Products: Vantage, FlexiCapture, FineReader Engine

ABBYY is a developer of recognition technology that has long held a leadership position in the Capture & IDP market. ABBYY takes a hybrid approach to document classification and extraction, combining OCR/ICR, machine learning, and LLM models to best fit the needs of an application, considering factors like accuracy, speed, consistency and cost. ABBYY is training its own vision language model (VLM) specifically for the purpose of document processing, which should further improve its pricing and efficiency options.

ABBYY is also a leader in mobile capture and makes different aspects of its technology stack available through SaaS, on-prem and private cloud offerings. ABBYY acquired a process mining ISV, but its primary focus is on Capture & IDP integrated with third-party enterprise software applications through a

combination of pre-built connectors and REST APIs. ABBYY goes to market through a combination of direct sales and a variety of partners, which have helped it achieve a broad global footprint.

Iron Mountain

Corporate Name & HQ: Iron Mountain, Portsmouth, NH

Capture & IDP Product: InSight DXP

Iron Mountain is a global leader in information management services, spanning both analog and digital records. It has offered scanning services for several years and in 2024 launched the InSight Digital Experience Platform (DXP), with IDP as a core capability that had evolved from the earlier InSight platform. DXP is a suite of intelligent content offerings that also includes content management and governance, agentic AI, and media management.

InSight DXP has evolved from primarily leveraging hyperscaler document processing platforms to now incorporating Iron Mountain's own trained machine learning, including specialized models for multi-modal capabilities and holistic document analysis. InSight features a low-code interface for designing and deploying document processing workflows. This includes drag-and-drop nodes with pre-built components and natural language customization capabilities. InSight is primarily deployed through a SaaS model.

Iron Mountain has a global base of 250,000 clients to which it provides a range of services related to physical and digital storage, governance, and information management. The initial go-to-market strategy for InSight has been to target these customers.

Kodak Alaris

Corporate Name & HQ: Kodak Alaris, Rochester, NY

Capture & IDP Product: Kodak Info Input Solution

Kodak Alaris is a leader in the production scanning market and offers Capture & IDP Software as a complementary solution. The Kodak Info Input Solution (KIIS) is a browser-based IDP application that provides a number of classification and extraction options. For advanced features, it leverages hyperscaler AI models trained specifically for document processing, as well as LLMs.

KIIS is brought to market through Kodak Alaris' global distribution network, which includes a combination of scanner resellers, ISVs, and consulting firms. The vendor also markets Kodak Capture Pro, a batch scanning software and direct integration with KIIS adds IDP capabilities.

Tungsten Automation

Corporate Name & HQ: Tungsten Automation, Irvine, CA

Capture & IDP Product: TotalAgility

Tungsten is a long-time market leader in the Capture & IDP space that came to prominence under the Kofax brand before changing its name in 2024 to reflect a broader focus on business process and agentic

automation. TotalAgility is its flagship platform that incorporates Capture/IDP, BPM, and RPA and has been infused with AI capabilities.

Tungsten has a global customer base and focuses on automating document-driven processes in targeted verticals like insurance, healthcare, logistics, manufacturing, and the public sector. It has a full spectrum of Capture & IDP technologies, from the ability to drive scanners, to its own market leading OCR/ICR, to machine learning-based AI and LLMs. Tungsten goes to market through a combination of direct sales and a broad partner network.

ABOUT THE ANALYST



Ralph Gammon is a Senior Analyst for Infosource Software. In this role he covers market sizing, trends, opportunities and leaders in the Capture & Intelligent Document Processing software space. Ralph joined Infosource in 2017 as Regional Manager for the Americas. From 1998-2020, he served as editor and publisher of *the Document Imaging Report*, the premier insider newsletter covering the ECM/Content Services space.

COPYRIGHT

© 2026

The entire contents of this Market report are the copyright of Infosource USA, Inc. and Infosource SA. All rights reserved. ***Reproduction or transmission of this publication in any form to any organization or individual outside the noted licensee company without the express permission of the copyright holder is prohibited.***

The information contained herein has been diligently obtained from sources believed to be reliable and has been prepared with care. However, Ralph Gammon and Infosource disclaim any and all warranties as to the accuracy and completeness of this information. Infosource USA Inc., its directors, employees or assistants can accept no liability for any damages or loss occasioned to person, company, or entity due to errors or omissions in the information contained herein or in the interpretation thereof. The opinions expressed herein are the opinions of the author and may change at any time without notice.

Infosource (USA) Inc.

1013 Centre Rd. - Suite 403A

Wilmington, DE 19805

+1 (302) 467-2255