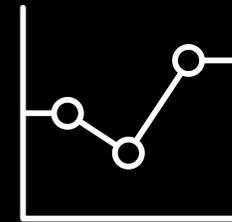
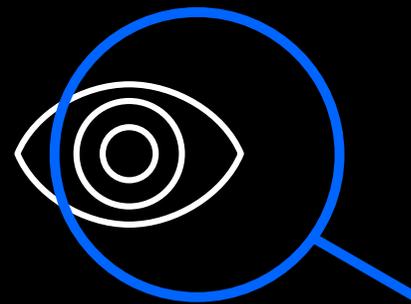
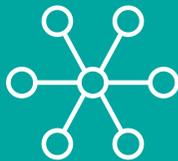


# The essential buyer's guide to data capture and automation

*What to look for when considering capture and automation solutions*





Businesses thrive on data—but only when it is available, relevant and actionable. This eBook introduces the power of automated data capture and serves as a guide for buyers currently assessing data capture and automation solutions—whether it’s part of an overall digital transformation initiative or just to accelerate a document-based transaction.

It also provides important considerations about the ways in which the solution you choose will be used and how it might be integrated with the other parts of your information infrastructure, as well as essential solution capabilities and maturity.

Introduced 30 years ago, data capture has evolved from simple scanning and manual data entry of incoming documents into a key element of full-fledged digital business transformation. The technology has progressed significantly.

Today, you can input a stack of documents and it will tell you what they are, what’s important on them, and how to put the data to work for the business—in seconds.

Information has always been key to business decision making, but those decisions can only be as good as the information available to guide them.

For today’s enterprises, the rich possibilities for customer interaction and internal communication include a huge spectrum of documents—from email and instant chats to correspondence and telephone transcripts—of which only a fraction is readily available today.

Too much of the data that could be used to inform decisions is in the form of unstructured data, or is siloed in one part of the organization rather than widely available. The information is out there—but it’s not easy for business professionals to locate, or for data architects to make it accessible across the enterprise.

# 72%

*Companies that have invested in technology to drive and improve enterprise content and business process management.<sup>1</sup>*

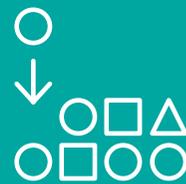
As data capture software solutions have grown more effective and accurate, they have become a critical tool for achieving business and operational objectives. Not only do customer experience and cost reduction, but they also can reveal information hidden in documents that organizations previously couldn't access or leverage.

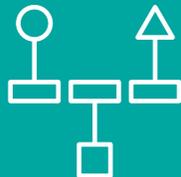
Customer sentiment (good or bad), product quality issues, and important sales trends are often buried inside transactional documents and correspondence. Think of the benefit of being able to flag personal data that requires heightened protection—*before* it becomes a problem.

Intelligent capture systems today are often referred to as document automation, because they make information ingestion faster and more accurate, and can be more flexibly integrated compared to manual data entry that requires constant human participation.

*Modern businesses are leveraging machine learning, artificial intelligence, bots and advanced analytics to completely improve everything from the customer experience to enterprise-resource and supply-chain management, workforce management, sales, marketing and more.<sup>1</sup>*

The latest systems are incorporating artificial intelligence (AI) to “read” documents like a human, to identify and classify the type of document and extract key data. Such systems can efficiently, accurately convert the varied content those sources contain into standard data types, scanning for relevant information and feeding that information into common data stores.





## The importance of cognitive capture

As capture enters its fourth decade, the most significant development in its history is the recent introduction of AI. The greatest challenge for a capture system is locating the data needed in a document, and the second greatest challenge is identifying the type of document.

The new generation of “cognitive capture” does both of these things far more effectively than previous systems. Here’s why: Today’s intelligent capture solutions were built on a strategy of creating templates.

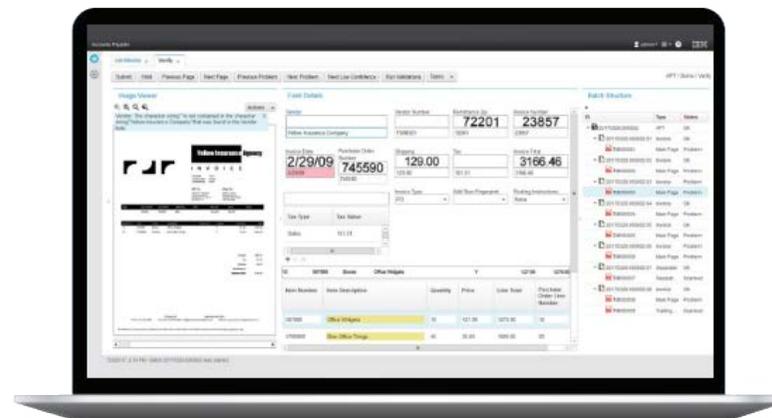
Over the years, capture systems added a wide variety of capabilities such as multiple optical character recognition (OCR) engines, location rules, database lookups and cross-field validations to raise the capture accuracy.

They became quite effective at capturing data from structured layouts such as tax returns, medical claims and even invoices. But input a letter, a contract or even a

stack of 200 different documents for a mortgage application, and results drop off quickly.

Unlike the template-based systems, cognitive capture “reads” the entire document, much like a human does, and it makes note of headers and footers, data labels, tables, text formatting (such as bold, italics and larger fonts) and other features of the document that provide context.

Cognitive capture also uses natural language understanding to recognize phrases and determine the “emotion” of the document. Then it compares all that information to a library of documents it has already processed, and it makes a determination.



Cognitive capture can easily identify a letter’s sender address at the top, block paragraphs in the middle and signature at the bottom. Then it will factor in key phrases

such as “my wife and I are unsatisfied,” “not what was agreed to,” or “our attorney recommends.” The technology can conclude that the document is a dispute letter and that the customer is not happy. It can even recommend that the document be forwarded to customer service immediately. No template-based system is capable of doing that.

When business information is trapped in unstructured documents, it remains essentially invisible. Until now, most companies simply scanned these documents, indexed them with a date and document number and stored them in a repository.

If they needed data on these documents, they would devote valuable employee time for labor-intensive manual information entry. More importantly, opportunities for increased response time and operational efficiencies were lost.

Only now, with these new cognitive systems, has data capture made it possible to capture all documents—not just structured layouts—and make that data actionable with a minimum of human intervention.

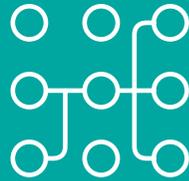
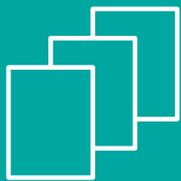
And today’s organizations can use that information to drive immediate business decisions and longer-term strategic decisions. With the information in these unstructured data sources transformed into consistent, accessible information, business processes can be accelerated and customer experiences improved.

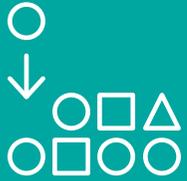
The most valuable data capture solution is one that can help spot, classify and extract valuable information based on its source, its context, its form, and its content, then place it into data repositories optimized for easy searching and retrieval.

## Avoiding the pitfalls of a naïve capture process

A company deploying a new capture system goes into the project with high hopes and a need to achieve real business results. Yet all too often, capture solutions become hampered by increased processing time or a greater need for manual user intervention than previously expected.

These can lead to lower productivity, but even worse, provide greater opportunity for errors to sneak into the captured data. Longer response time, too, can harm the customer experience; a longer turnaround for data capture means that any downstream process that depends on that data is also delayed.





The need for manual intervention means that knowledge workers' skills are steered away from high-priority tasks to help overcome technology gaps, setting up an unfortunate situation where the automation solution doesn't deliver on its initial promise, or, even worse, actually reduces business and operational efficiency.

Now is the time, when you are embarking on a data capture journey, to find the right solution for the capture challenge you face, and put the right plan in place to deploy it.

## How IBM approaches data transformation with capture

Every industry has its own unique document types, every organization handles documents in its own unique way according to its policies and procedures, and every document can vary, based on document quality, contrast, font type, embedded images and layout.

Rather than cobbling together disparate tools and services to address these multiple challenges, the IBM approach to data capture is based on tightly unified components and capabilities that form a comprehensive whole.

For IBM, data capture is an integral part of the [IBM® Automation Platform for Digital Business](#), which unites in a single platform five important elements, the first being capture. The others include:

- **Task automation**—Using software to automate vital but repetitive human tasks, to minimize error-prone manual work
- **Content management**—Providing secure management and instant access to a wide range of content, to ease business tasks and to help assure governance and compliance

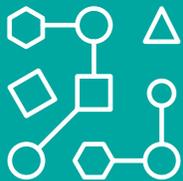
*Today's business automation has evolved to combine automation, cognitive and analytics to aid digital transformation. It brings together elements of human intelligence and software automation to help enterprises augment their digital workforces, redefine the customer experience and streamline their engagement with the business ecosystem.*

- **Workflow management**—Offering tools to help enterprises model, automate, change, monitor and optimize core business processes
- **Business decision automation**—Encapsulating business rules for greater consistency, accountability and speed in detecting problems

## End-to-end: Client experience, ease of use and systems integration

Digital transformation is all about improving customer experience. While it began with the introduction of new mobile and social capabilities, now the emphasis has shifted to automation of front- and back-office processes.

Process automation can drive better outcomes, such as faster response and frictionless transactions, and front-line workers can solve problems for customers at the first point of contact. Automated data capture is a crucial element of this strategy.



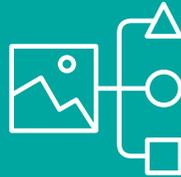
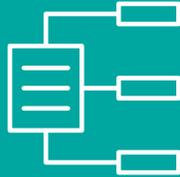
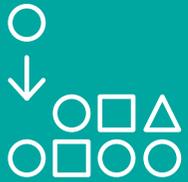
## Leverage data capture from mobile devices

*Mobile devices are increasingly used to capture information, such as from documents that can be shared to accelerate business transactions. For example, a document image captured from a mobile device can be transmitted to complete a mortgage loan application.*

When sourcing a data capture solution, ask three essential questions:

- Does it provide straight-through processing to remove human interaction wherever possible?
- Does it have an AI capability to capture unstructured documents and enable you to expand to include all types of documents?
- Is it tightly integrated with other business automation tools, such as a content repository, workflow and robotic process automation?

Data capture approaches that demand finicky data repository configuration or that lack mature cognitive capabilities can lead to lower employee productivity, an inability to deliver an



end-to-end customer journey, increased workloads for individual departments, and a greater burden related to compliance with data protection regulations.

These inefficiencies stem from inconsistent interfaces, the lack of robust AI-based data-classification capabilities, and the hard-to-avoid complications of multi-vendor integrations.

[IBM Datacap Mobile](#) provides enhanced IBM Datacap image-capture capabilities through a streamlined, mobile user interface for Google Android or Apple iOS devices. The solution enables users to [capture and submit documents on demand](#) and helps increase accuracy with capture and validation at the point of contact, while accelerating content integration into a case file or business process.

## Avoid barriers to business efficiency and operations

For data capture and enterprise content management (ECM) that [can save you the most time and present the best face to customers](#), seek out a solution that:

- Features end-to-end capabilities that will deliver a lower total cost of ownership, rather than one dependent on a third party for a data repository or other key operation
- Presents a unified user interface (UI) and data environment backed by a single experienced provider, rather than a multi-vendor, multi-UI environment, with a range of support services to wade through
- Integrates required capabilities smoothly through the enterprise
- Features AI tools for data capture as a mature core strength, rather than a bolted-on capability

*“Companies should make sure their data capture and automation programs can integrate with all relevant back-office and customer-facing applications. That’s the only way to ensure the newly identified information is being used as part of a contextual picture.”*

*–Frost & Sullivan<sup>1</sup>*

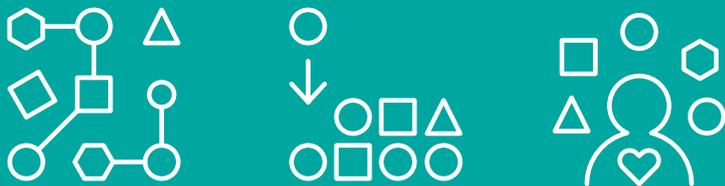
# Enterprise content management

Data capture solutions based on loosely integrated software components can lack cohesion and hinder a smooth user experience. A range of systems may meet the right technical requirements, but if they are equipped with different interfaces, you are establishing what is known as “swivel-chair processes.”

Such processes make users close one window and open another, degrading productivity. Plus, employees may need additional training before the system can help them efficiently generate insights, and higher-skilled, highly compensated employees or consultants may be required to integrate the systems.

For an enterprise that primarily requires data capture today, the business potential of a fully integrated business automation platform may not be obvious.

However, with a growth-oriented integrated platform, additional capabilities beyond data capture can be added



# 24 hours

*Reduction in response to credit requests by reusing information experienced by Banco Galicia with IBM Datacap.<sup>1</sup>*

one at a time, while keeping a familiar interface and enterprise-wide data capture capabilities.

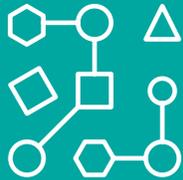
## How IBM addresses enterprise data capture needs

The IBM Automation Platform for Digital Business is an integrated platform of five distinct but tightly coupled automation capabilities that can help businesses drive virtually any business process automation project at speed and scale.

With the capabilities of this platform, IBM provides end-to-end, unified and market-validated tools for ECM, information lifecycle governance and case management.

Because the IBM Automation Platform for Digital Business unites these capabilities and allows them to be used in concert, the value of captured data is magnified.

<sup>1</sup> [“Banco Galicia: Going paperless to support top-notch customer service and unlock massive efficiency gains,” IBM, June 2018.](#)



## From strength to strength

Think of the IBM platform as a complete automation ecosystem, addressing the entire lifecycle of content management, from ingestion, security and workflow to storage, robotics and compliant archiving.

And because IBM is a single vendor, customers can specify, customize and deploy a data capture system in a single process, avoiding the financial and decision-making overhead of a multi-source implementation.

Besides the advantages of dealing with a single vendor for business simplicity, IBM offers a key advantage over other business automation providers: [IBM Content Navigator](#). Content Navigator provides a unified, configurable and customizable user interface across most IBM automation modules—from capture to workflow to content repositories, on-premises and in the cloud.

This standard user interface across the platform increases employee productivity, improves the customer interaction with your organization by delivering a single user experience, and reduces swivel-chair processing. With a catalog of plug-ins available, Content Navigator can support Microsoft Office documents, leverage text analytics, create team rooms for collaboration, and much more.

As an offering within a larger platform, the data capture solution from IBM offers interoperability, a consistent user experience, operational insights, and the ability to easily add additional capabilities, such as OCR engines for capturing handprint or AI services for classification, [minimizing up-front investment](#).

### ***Capitalize on agility and cost effectiveness with a cloud-based solution***

*The importance of deployment and pricing flexibility with data capture solutions cannot be overstated. [IBM Datacap on Cloud](#) is a cloud-based, data-capture-as-a-service solution that delivers all the power of the data capture and imaging capabilities of Datacap, including the ability to apply cognitive capture to complex tasks, in a managed services cloud environment.*

Furthermore, many robotic process automation (RPA) vendors have begun to offer RPA as a solution to simple capture challenges, using their “screen scrape” capabilities. Clients who invest often find they need to add traditional capture to complete the solution.

Knowing that a capture system is integrated with [RPA](#) can be a huge comfort. Instead of knitting together a patchwork of point solutions addressing aspects of data capture piecemeal, the IBM approach creates a data ecosystem that encompasses the entire data lifecycle.

## A reliable roadmap for data capture

Any experienced enterprise architect understands the risks of creeping increases to a project’s scope and cost and the importance of selecting a proven vendor with a development roadmap that keeps pace with the needs of the organization.

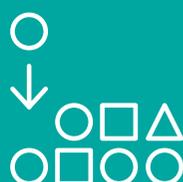
This is especially true when a system requires the integration of software from more than one provider, and the costs of complexity can include incompatible software versions and difficulty in troubleshooting problems. Selecting a single-vendor solution can reduce the headaches of a complex implementation.

Stability is important too. A vendor that has undergone ownership changes or has radically revamped its offerings may suffer in account billing and communication practices relating to its portfolio of services.

Likewise, a capture provider with a limited set of capabilities may break compatibility with ancillary products such as data repositories even while providing technical upgrades to its core features.

Some data capture solutions feature complicated pricing models that require customers to pay per use, or to pay separately for features such as PDF conversion or barcode reading.





IBM provides a long-term roadmap and offers consistent, flexible pricing and deployment options through a single end-to-end source for every component of a data capture solution. By contrast to more narrowly-focused competitors, IBM:

- Provides a single-vendor solution to help ensure that all components work smoothly together, and that their product development roadmaps align
- Offers customers a consistent, straightforward, enterprise-suitable solution pricing and billing experience, allowing customers to buy what they need and scale up to match business needs

## What to look for in a data capture solution

Digital transformation is an ongoing and growing need, and today's market offers numerous data and document capture solutions for enterprises to deploy, including RPA. How can you choose the right solution in this

increasingly crowded field? Here are seven key elements to consider:

1. **Cognitive capabilities:** The more advanced the cognitive capabilities of a solution, the more confidently you can hone in on your most valuable data.
2. **End-to-end simplicity:** Consider whether multiple vendors must be consulted before a data capture solution can be specified, validated and implemented.
3. **Integration:** Data capture's value depends on how smoothly it works with the rest of your IT infrastructure, including delivery to content repositories, launching a workflow or case, enhancing RPA, and providing reliable data for big-data-style content analytics.
4. **Fast time to value:** Are all components ready for rapid installation, or are you buying a toolkit that requires extensive customization?
5. **Consistency:** Look for a system in which the various components feature a common interface, for a smooth user experience and to help minimize expensive user training.

# 264,000

*Documents scanned, classified and stored per month at Argentina's Banco Galicia with the aid of IBM Datacap and IBM FileNet®.<sup>1</sup>*

1 "Banco Galicia going paperless to support top-notch customer service and unlock massive efficiency gains," IBM Corp, 2016.

6. **Maturity and reliability:** Data capture is a long-term endeavor; for lasting value, seek out a vendor that is reliable, with strong references and a predictable roadmap.
7. **Support:** Select a vendor with its global reach and support infrastructure in mind, and with a history of product maintenance and availability.

## Take the next step

This eBook has presented important attributes and potential drawbacks that enterprise data architects must consider in designing the data capture solution that works best for their organization.

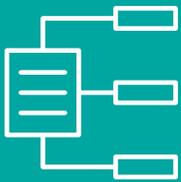
Besides weighing the specific factors that matter most to your own enterprise, consider a solution that exemplifies [today's best practices in business automation](#):

- Focus on how people actually work, and then identify the parts of each process that are repetitive and mundane.

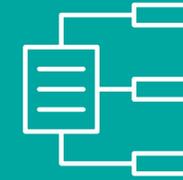
*Data capture is critical to any digital business automation program, but it is only part of the picture. Don't forget the rest: tasks (through robotic process automation, or RPA); content (through a robust content management system); workflow (to provide visibility and consistency to thousands of human processes); and decision making (to increase speed, accuracy and effectiveness).<sup>1</sup>*

–Frost & Sullivan

- Streamline business dealings by extracting and centralizing content from common contract documents, as well as supporting emails, letters, recordings and more.
- Apply automation to internal processes, from employee recruitment to supply-chain management, to drive efficiencies.
- Ensure that data capture and automation programs integrate with all relevant back-office and customer facing applications.



As you consider the ways that data capture could be most useful and most smoothly integrated, consider how each approach you examine fulfills the following business needs:



Does the solution offer ...	IBM	Vendor A	Vendor B
A mature history of usage	✓		
A unified UI across applications	✓		
Full integration with digital business automation tools, including workflow, case management, content management and robotics	✓		
Datastore provisioning	✓		
Fast time to value	✓		
Single-vendor provisioning	✓		
Cognitive capabilities for classification and data location	✓		
Capture capabilities from a managed services cloud environment	✓		
Image capture from smartphones and tablets	✓		
Flexible deployment and licensing options on-premises, in the cloud and on mobile	✓		
References who can vouch for business results	✓		

After you've examined the solutions available to you, learn more about the IBM data capture solution and the IBM Automation Platform for Digital Business by visiting [ibm.com/automation/ibm-automation-platform-digital-business](http://ibm.com/automation/ibm-automation-platform-digital-business), or contact your IBM representative to arrange a demonstration of the integrated, cognitive power of IBM Datacap, and see how it can help unlock insights from the data you already hold.

© Copyright IBM Corporation 2018

IBM Security  
New Orchard Road  
Armonk, NY 10504

Produced in the United States of America  
November 2018

IBM, the IBM logo, ibm.com and FileNet are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml)

Microsoft is a trademark of Microsoft Corporation in the United States, other countries, or both.

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

The client is responsible for ensuring compliance with laws and regulations applicable to it. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the client is in compliance with any law or regulation.

Statement of Good Security Practices: IT system security involves protecting systems and information through prevention, detection and response to improper access from within and outside your enterprise. Improper access can result in information being altered, destroyed, misappropriated or misused or can result in damage to or misuse of your systems, including for use in attacks on others. No IT system or product should be considered completely secure and no single product, service or security measure can be completely effective in preventing improper use or access. IBM systems, products and services are designed to be part of a lawful, comprehensive security approach, which will necessarily involve additional operational procedures, and may require other systems, products or services to be most effective. IBM DOES NOT WARRANT THAT ANY SYSTEMS, PRODUCTS OR SERVICES ARE IMMUNE FROM, OR WILL MAKE YOUR ENTERPRISE IMMUNE FROM, THE MALICIOUS OR ILLEGAL CONDUCT OF ANY PARTY.

