



FEDERAL GOVERNMENT | WHITEPAPER

THE FUTURE OF DIGITAL GOVERNMENT IS AI-POWERED INTELLIGENT DOCUMENT PROCESSING (IDP)

Hyland®



Documents are everywhere in our personal and business lives — for example, forms for employee onboarding, identification, taxes, FOIA requests and benefits eligibility verification. The federal government spends vast sums of money getting staff to manually input, edit and process data and documents. However, this is mundane, time-consuming data entry and prone to human error.

Thankfully, advances in digital technology have enabled agencies to more efficiently process information and to derive new insight and value from the data contained within their trove of documents.

Intelligent document processing (IDP) is an emerging solution that can transform and digitize government documents quickly and cost-effectively. This increasingly sophisticated intelligent automation technology improves the precision and throughput of document processing beyond traditional optical character recognition (OCR) and robotic process automation (RPA) technologies.

IDP uses artificial intelligence (AI) information extraction to read, recognize and understand the text and structures within the federal government's multitude of differently formatted templates, forms and documents. Federal agencies can then use machine learning (ML) to "teach" the IDP platform how to make sense of documents and classify the data within.



THE OPPORTUNITY TO IMPROVE EFFICIENCIES AND PRODUCTIVITY WITH IDP

Most government data is in a semi-structured or unstructured format — for example, emails, audio and video recordings, images and photographs, reports and studies, forms and faxes. According to FedTech, some analysts have estimated that up to 90% of all data generated by an organization is unstructured.

Wouldn't it be helpful if technology could automatically capture, extract, interpret and process data within those documents?

Documents come in a wide range of formats and layouts: Some contain typed text, some handwritten; others contain checkboxes, tables or handwritten signatures. Some are easier to read than others.

A passport application form, for example, can involve the applicant writing in their name, address, Social Security number and more information as well as completing check boxes and submitting a physical photograph. After submission, a clerk will manually input this information into a computer system. The tedious process can take months and is prone to human error. IDP offers a solution to extract and interpret content and context from government documents using AI and ML models that keep learning over time to become even more accurate and precise.

Agencies could make better use of their staffs' time by instead training an IDP system using supervised learning. This training approach involves giving it correct examples to follow, consisting of a set of documents and their expected output. Once trained, IDP offers a tremendous opportunity to gain new data-driven insight and a competitive advantage from the tremendous volume of data contained within documents.

Once data is extracted in a structured format, IDP can then be combined with RPA to provide agencies with a simple yet effective way to automate processes end to end. The combination of IDP with RPA bridges the gap between unstructured data, human verification processes and structured data that is required to automate processes. For example, data extracted by IDP can be input programmatically into back-end government software systems using RPA to achieve a greater level of automation or can be stored in a repository to augment strategic decision-making.

Effectively extracting data and structuring information is a gateway toward automating many federal government processes that, today, rely on manual inputs and human intervention.

IDP USE CASES FOR FEDERAL AGENCIES

In any government agency, you will find a multitude of different document types and vast swaths of people struggling to meticulously process them with precision and speed. This means that there are many potential use cases to employ IDP such as:



Department of Defense

- Background checks and clearances
- Criminal investigations
- Facilities and equipment inspections
- Complaints and harassment
- Supply chain automation



Department of State

- Passports
- Visas
- Property management
- Personnel relocation
- Logistics



Department of Homeland Security

- Immigration and naturalization records
- Customs documents
- Background checks
- Emergency response
- Criminal investigations



Department of Treasury

- Tax forms
- Audits and appeals
- Personnel relocation and logistics
- Investigations



Department of Labor

- Complaints and appeals
- EEOC investigations
- Workers' compensation
- Policies and procedures



Department of Energy

- Inspections
- Licenses and permits
- Supply chain
- Emergency response



Health and human services

- Benefits applications
- Fraud investigation
- FDA approvals
- Complaints and appeals
- Policies and procedures



Records management and information governance

- Record ingestion and classification
- Compliance with ISO 15489, ISO 16175, FOIA, U.S. DoD 5015.02 CH2 and CH3, EgovG and MoReq



Procurement

- Purchase order processing
- Customer onboarding
- Vendor onboarding
- Contract administration
- Customer contracts
- Tenders



Human resources

- Employee onboarding
- Resume screening
- Identity documents
- Application processing
- Benefits management
- HR records
- References



Finance and accounting

- Invoice processing
- Bank statements
- Collections
- Receipts
- Rebates or returns
- Tax forms
- Expense receipts
- Bank statements



Supply chain

- Customs declarations
- Bills of lading
- Proof of delivery receipts
- Order scheduling and tracking
- Insurance documentation

CRITICAL COMPONENTS OF ANY IDP PLATFORM

To take full advantage of IDP capabilities, federal agencies need to capture/ingest documents and then process, classify and validate them, verifying them either digitally or with a human in the loop. Finally, the data from the documents needs to be intelligently processed and embedded into systems and applications.

Below is a typical flow for an enterprise-grade IDP solution:

- 1. Document capture/ingestion:** Ingest a document in electronic or paper format.
- 2. Image processing:** Process the document image to improve image quality.
- 3. Indexing and classification:** Classify a document into the correct document type using text mining and/or ML algorithms.
- 4. Extraction and validation:** Extract and validate data contained within the document. To do this, an IDP system needs to have been trained on samples of similar documents. Apply predefined validation rules to extract valid information and insights contained within the document and to categorize and organize the document itself.
- 5. Human-in-the-loop verification:** IDP tools provide interfaces that permit human validation of the extracted data as needed to cut down on the use of “people” in close succession. People can check that the correct information and insight have been extracted and further improve the accuracy of the ML model by training it.
- 6. Delivery of data to another system:** Accurate data that has been extracted from a document can then be sent to another computer application or used in a decision-making process (e.g., ERP, CRM Snowflake or a decision model).

HOW TO CHOOSE THE RIGHT IDP PLATFORM

There are a wide variety of IDP platforms to choose from in a growing market.

According to Valuates Reports, the global IDP market is expected to grow from USD 860 million in 2021 to USD 4.15 billion in 2026 at a CAGR of 37%.

But how can an agency decide which IDP platform is best suited to their needs? The selection criteria below can be used to judge the suitability of an IDP platform:

- Accuracy of the extraction, verification, processing and classification engine, including the ability to ingest handwriting, tabular data or bar codes and to understand semantic context.
- Ability to automate the capture, extraction and verification of data from documents and images — of different file types, in different languages and across departments.
- Ability to perform image cleanup functions like line straightening, removing lines and dot shading and character enhancement.
- An intuitive user interface for ease of use, particularly if your agency wants its users to own and train their ingestion models.
- Ability to program and extend an IDP platform using low code. Government agencies are rapidly moving toward highly configurable low-code/no-code platforms with intuitive tools like checkboxes, radio buttons and configurable menus that enable them to make changes at any time, without expense or time-consuming and difficult-to-maintain coding or scripting.
- ML models with the ability to determine where the important information lies — e.g., in images, tables, documents or bar codes.
- Automatic identification of a document or documents, ensuring the correct information is extracted from each page, regardless of the process.
- Out-of-the-box functionality and the ability to expand on that to meet increasingly complex federal requirements.
- Access to key metrics and performance indicators for your processes and metrics that highlight the accuracy and performance of the IDP platform.
- IDP data logging for security risk assessment, compliance and auditing purposes. The ability to log and retain records, enforce structured data retention policies, support information governance and minimize legal risk is ever more important in the data age.
- Built-in intelligence that continuously improves over time, further reducing manual touch points and accelerating document processing. This eliminates the burden of manual processing and data entry as well as accelerates downstream classification and data processing, which improves information accuracy across systems.
- Extensibility to integrate with, and directly deliver data to, other platforms, systems or technology solutions, such as enterprise resource planning (ERP), enterprise content management, data management systems, process management or other technology solutions without disruption. This ensures that teams are working with the most complete and accurate data needed as they make critical process decisions.
- A portfolio of complementary platforms and products offered by the IDP vendor, such as case management, workflow, intelligent automation or RPA technology platforms.



CRITICAL SUCCESS FACTORS FOR MAXIMIZING RETURN FROM AN IDP PLATFORM

Implementing any enterprise technology platform takes careful planning. There can be challenges to deploying an IDP solution like any platform, but done right, intelligent document processing can deliver benefits that maximize your return on investment. To get the best results from an enterprise IDP solution, industry best practices recommend following these five steps:

1. Build a strong business case that articulates the benefits of IDP.

- Run a proof of concept to gauge the suitability of an IDP technology solution within your agency.
- Demonstrate a technology platform solution to key stakeholders and seek their input, direction and feedback.
- Deliver workshops to collect and prioritize a backlog of high-value IDP opportunities.

2. Seek executive approval and sponsorship.

- Get an executive sponsor to champion the IDP program at higher levels.
- Present your business case, including total expected benefits, and your roadmap for implementation to the executive team to get approval for a program budget.
- Set realistic expectations on implementation time and running costs.
- Outline a clear governance structure to track opportunities and benefits over time.

3. Reach out to vendors and select an IDP platform and implementation partner.

- Create a vendor selection matrix and assess each IDP vendor and their platform.
- Engage your chief information security officer (CISO) and your risk, finance and procurement teams to negotiate a win-win deal with your IDP vendor of choice.
- Select a secure, enterprise-wide IDP platform.
- Go to market and select an implementation partner if your vendor does not have an implementation team.

4. Roll out an IDP solution.

- Start small but think big: Scale the IDP solution quickly to deliver significant value back to the agency on a consistent basis.
- Build a center of expertise to own and drive IDP adoption, best practice and use.
- Develop a comprehensive training and change management program to gain and retain buy-in for your IDP program.
- Train the IDP ML models using large volumes of use-case-specific documents and data.
- Focus on data quality and integrity by employing data governance controls that are supported and enforced by data champions.
- Continually monitor IDP metrics and performance indicators for your processes that highlight the accuracy and performance of the IDP platform.

5. Scale your IDP solution across the enterprise.


- Rapidly scale your IDP solution. For example, start with forms or documents in a single team, then rapidly scale across functions, regions and international geographies.
- Utilize complementary technologies, such as intelligent automation or RPA. This will help accelerate the flow of data between applications, systems and people to ensure that your staff members are always working with the most accurate and recent information as they make critical process decisions.
- Continually extol the benefits and impact of your IDP solutions to retain interest and attract further funding and ideas for your IDP delivery program.
- Govern the delivery of IDP to ensure value is delivered over time.

CONCLUSION

In an increasingly digital world, information is quickly becoming a federal agency's most valuable asset. But relying on staff to manually receive, process and input information into core government systems is expensive and leaves room for errors, slower processes and delayed and inaccurate decisions.

Intelligent document processing can accurately extract and speedily classify data held within federal document repositories. IDP platforms are now capable of accurately ingesting high volumes of documents at scale with outstanding precision and speed at a price point that suits most budgets. It is now possible to drive increased efficiencies, improved productivity and enhanced ROI from a successful IDP delivery.

There will always be documents and data fields that fall beyond an IDP system's capabilities. But thanks to AI and ML, IDP platforms are constantly becoming more powerful, accurate and quick to train. IDP allows agencies to unlock hidden information and value from their vast document libraries and frees the intellectual capital of government staff, who can now spend time on more interesting and valuable work to enhance mission delivery. The data unlocked by IDP can enable better decision-making and thereby transform and accelerate today's manual, paper-intensive processes into intelligent digital government.



ABOUT PASCAL BORNET

Pascal Bornet is an expert, author, keynote speaker, influencer and thought leader in artificial intelligence and automation. He is a pioneer in the field of intelligent automation (IA) and the author of the best-seller book, "Intelligent Automation." He has received multiple awards, and he is regularly ranked as one of the top 10 global experts in the fields of AI and automation. Bornet has published articles in Forbes, Bloomberg Businessweek, McKinsey Quarterly and The Times. He is a member of the Forbes Technology Council and a senior advisor for several startups and charities. After 20+ years of experience in consulting and building the Intelligent Automation practices at Ernst & Young and McKinsey, Bornet is currently on the leadership team of Aera Technology, an innovative startup.

Learn more at [Hyland.com/RPA](https://hyland.com/RPA)

Hyland®