



Intelligent Document Processing (IDP) Playbook 2023

June 2023



Contents

Preface	4
1. Introduction to IDP	5
• Importance of data and digitalization in this age	6
• Current challenges	7
• Emergence of IDP and its relevance in digital transformation journey	8
• IDP market overview and buyer trends	10
2. Understanding IDP	11
• Construct of an enterprise-grade IDP solution	12
• Core technologies powering IDP solutions	14
• Intricacies in document and data types handled using IDP	17
• Measuring and improving effectiveness of IDP solutions	19
• Benefits of IDP solution	20
• IDP deployment use cases	21
3. IDP market adoption and buyer trends	24
• IDP adoption trends by geography	25
• IDP adoption trends by industry	26
• IDP adoption trends by process area	27
• IDP adoption trends by buyer size	28
• Drivers for IDP adoption	29
4. Enterprise IDP journey	30
• Understand current state and market capabilities	31
• Building a business case – what is the right outcome target state?	34

For more information on this and other research published by Everest Group, please contact us:

Vaibhav Bansal, Vice President

Samikshya Meher, Practice Director

Ashwin Ambala, Senior Analyst

Ruchin Dwivedi, Senior Analyst

Contents

• Determining the capabilities required to achieve identified outcomes	42
• Identify all determinants and map path	43
• Execute against mapped path	48
5. Selecting the best-fit enterprise-grade solution	50
• Core product capabilities	53
• Solution approach	57
• Product training and support	58
• Technology partner ecosystem	59
• Service partner ecosystem	60
• Hosting options	61
• Commercial model	62
6. Challenges and best practices for IDP adoption	63
• Key barriers and best practices to accelerate value realization from IDP adoption	64
7. Future outlook	66
8. Appendix	69
• Enterprise IDP Capability Maturity Model (CMM)	70
• Environmental determinants	85
• Variance in execution path steps based on environmental determinants	86
• Glossary of terms	90

Preface

In today's fast-paced digital era, businesses face the formidable challenge of managing vast amounts of information and extracting valuable insights from it. Traditional manual document processing methods no longer suffice, often resulting in inefficiencies, errors, and missed opportunities. The advent of Intelligent Document Processing (IDP) has revolutionized the way enterprises handle and process documents, providing them with the tools to transform their operations, drive innovation, and unlock new opportunities for growth.

Since its advent, IDP has gained significant traction in various industries, including finance, healthcare, insurance, legal, and many others. The adoption of IDP solutions has helped enterprises streamline document-centric processes, improve data accuracy, reduce manual labor, and enhance overall productivity. As the IDP continues to evolve, technologies such as generative AI are being integrated into IDP solutions, further expanding their capabilities and potential impact on businesses.

This playbook is specifically crafted to address the needs of enterprise readers seeking to understand and adopt IDP technology as well as practitioners looking to expand and scale their existing IDP implementations. It is aimed to help enterprises, automation experts, and business leaders understand the latest innovation in the processing of

structured, semi-structured, and unstructured documents with the application of heuristics, NLP, NLG, computer vision, and LLMs. It is designed to provide a holistic view of the latest advances in IDP, help understand intricacies related to expected outcomes, offer framework that supports design decisions for IDP application, and equip readers with potential challenges and best practices for successful implementation. From envisioning the digital transformation roadmap to selecting the right IDP solution for enterprises, this playbook will provide actionable guidance and valuable tips at every step.

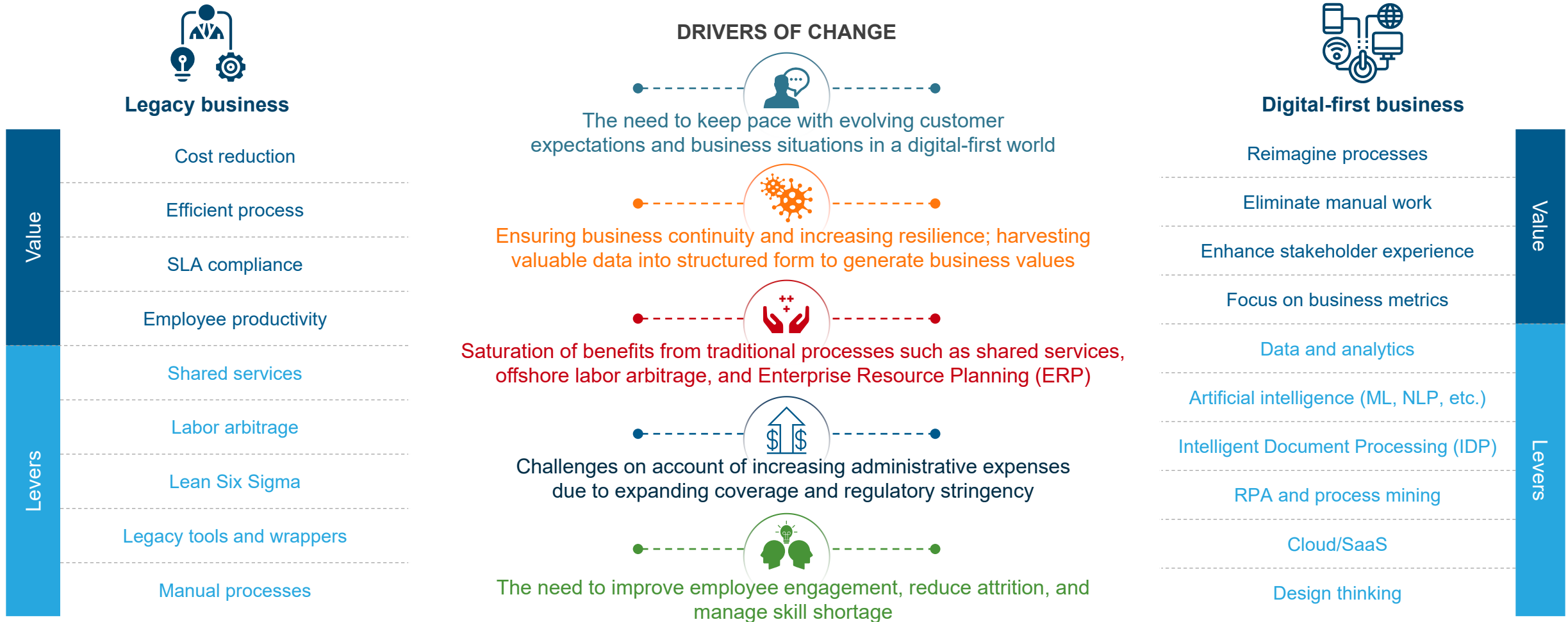


01

Introduction to IDP

- Importance of data and digitalization in this age
- Current challenges
- Emergence of IDP and its relevance in digital transformation journey
- IDP market overview and buyer trends

Evolving into a digital-first business is becoming increasingly important for organizations to remain resilient and competitive



As enterprises move along the digital transformation and optimization journey, they face numerous data-centric challenges

Data-centric barriers in an enterprise digital transformation journey

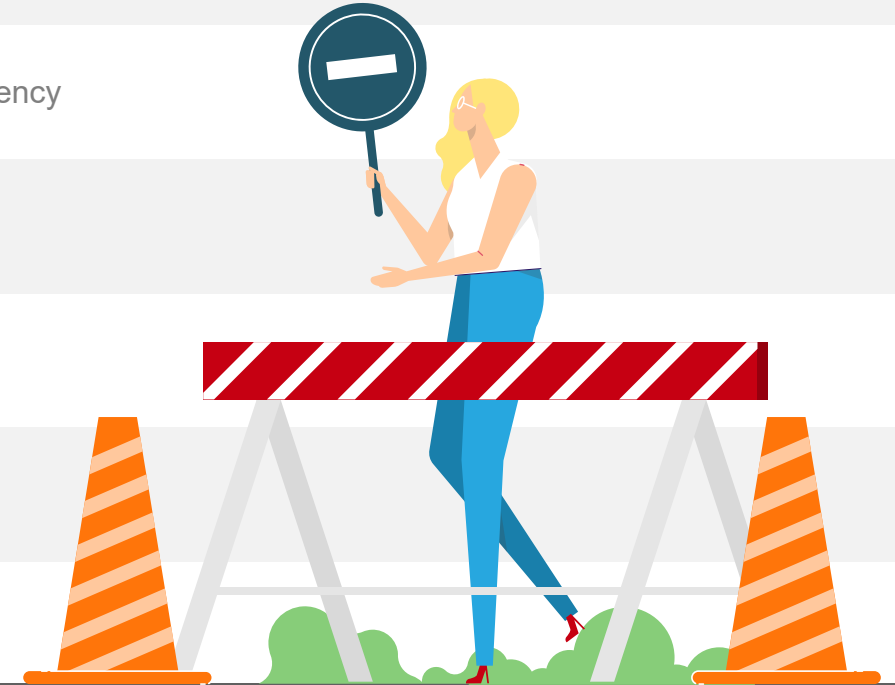
Managing and deriving meaningful insights from a large amount of unstructured data arising from cloud computing

Difficulties in integrating multiple software systems and platforms, leading to data silos and reduced efficiency

Collection and storage of sensitive personal and business data leading to cybersecurity risks

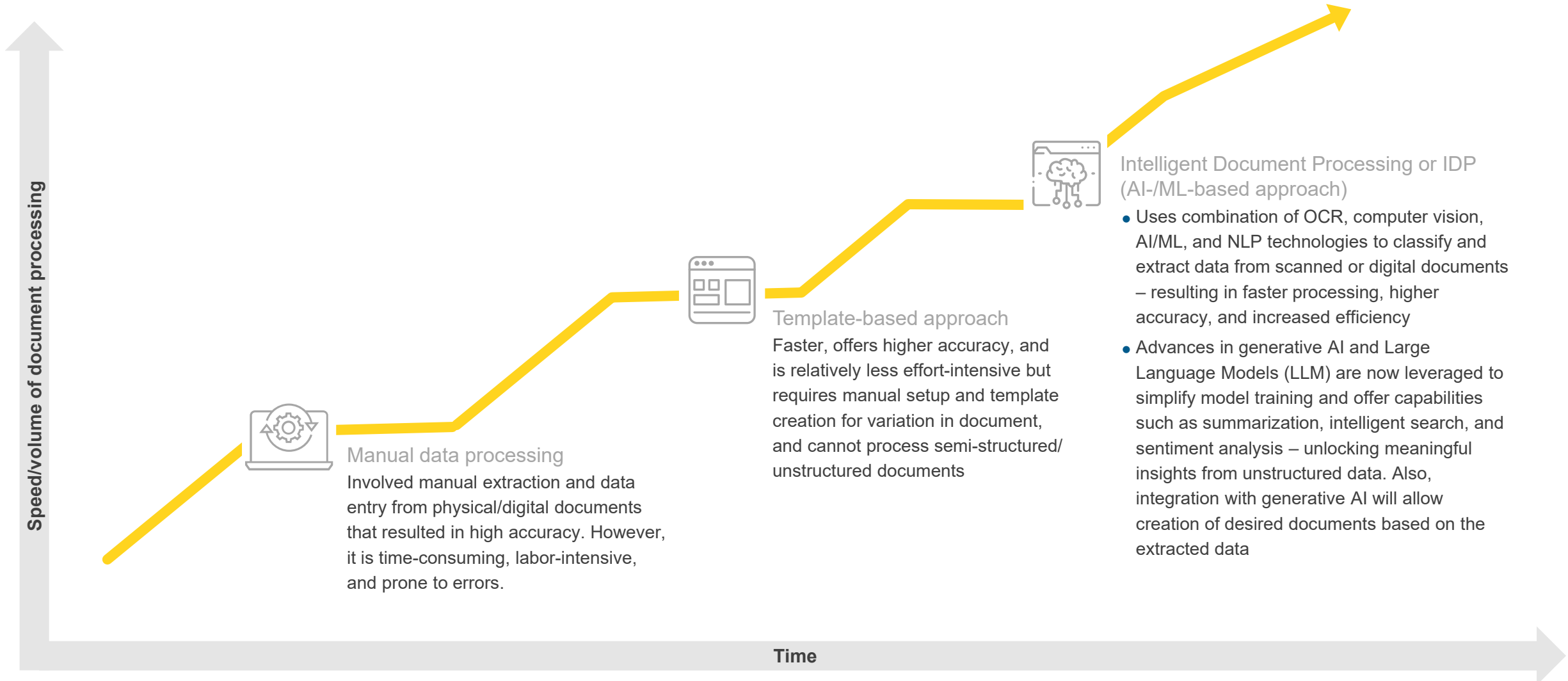
Increase in compliance requirement and the ever-changing regulatory landscape

Changing customer expectations in terms of fast and seamless experience that cannot be achieved with manual operations

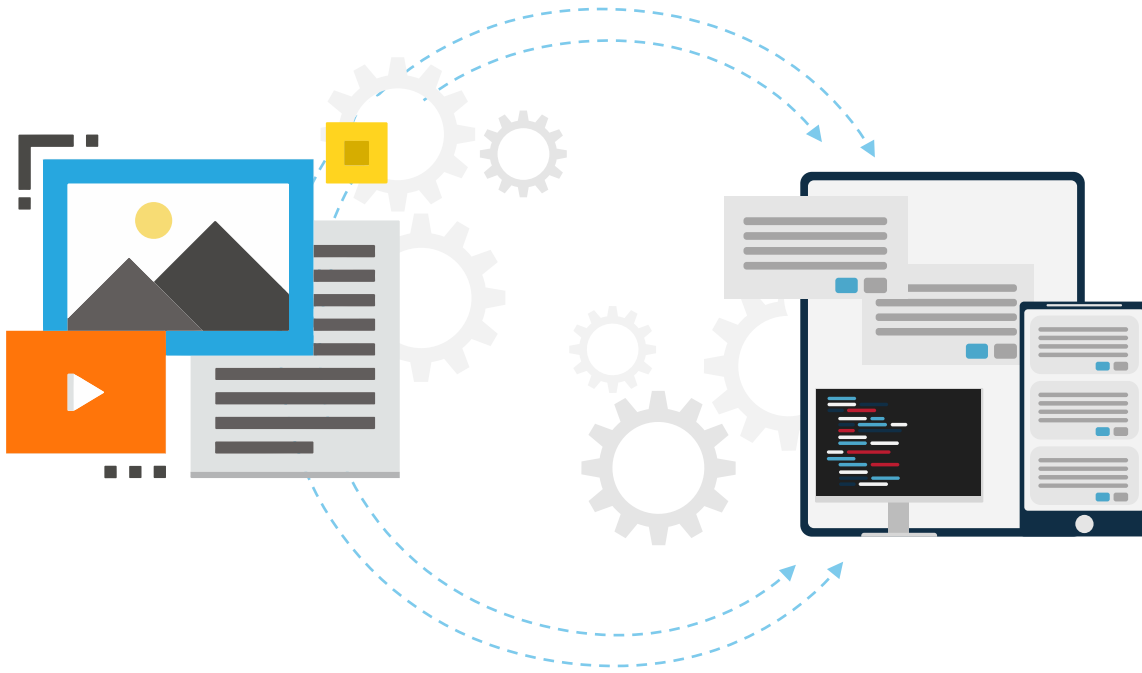


Automating data processing from various unstructured sources can help organizations overcome data-related challenges, achieve their digital transformation goals, and remain competitive in an increasingly data-driven world.

Approaches to document processing have evolved significantly, with AI/ML technologies helping in more efficient document handling



IDP has emerged to help organizations overcome data processing challenges, leveraging AI to address document processing and automate conversion of unstructured data into digital formats



Automate document processing

IDP solutions can ingest, classify, and extract data from semi-structured and unstructured documents using ML techniques, which can be further fed into downstream applications.

Address limitations of traditional RPA/OCR

It adds a layer of probabilistic decision-making over traditional technologies, such as RPA and OCR, thereby achieving a higher degree of automation.

Clean data upfront

It provides high quality and timely data, which can be further fed into other applications for use in downstream processes and helps create data lakes.

Harvesting data locked in historical documents

IDP can help organizations to unlock valuable insights from historical records, long documents (such as contracts), and sources such as websites and social media platforms that would otherwise be difficult or time-consuming to access.

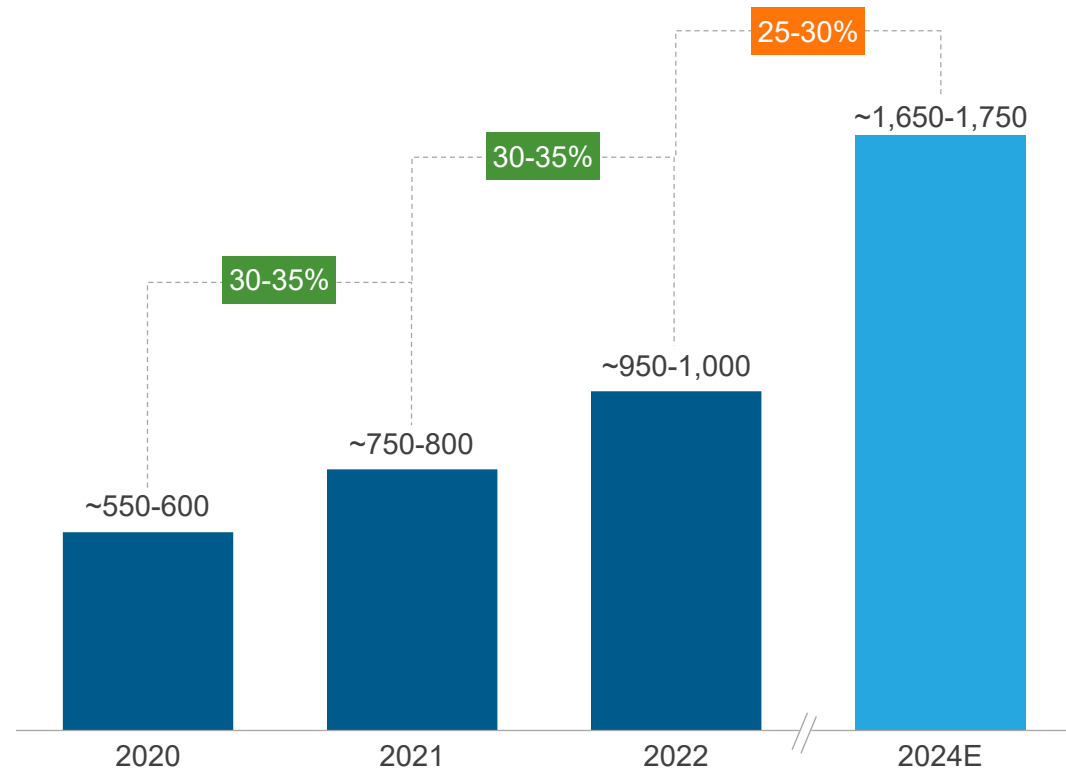
As the volume of unstructured data continues to grow, IDP becomes an increasingly important tool for enterprises that need to manage and analyze large amounts of data.

IDP market overview and buyer trends

The IDP market is experiencing steady growth and is expected to reach US\$1.7 billion by 2024

IDP independent technology providers market size¹
Revenue in US\$ million

XX Growth rate XX Expected CAGR



- The IDP software market stood at approximately US\$975 million in 2022, experiencing a YoY growth of ~30%. The growth is primarily driven by:
 - Growing organizational focus on digital transformation and business process automation to improve efficiency, reduce costs, and increase resilience
 - Increased availability of cloud-based and SaaS solutions offering higher flexibility, scalability, and cost-effectiveness
 - Robust security and compliance features that help address data security and privacy concerns
 - Expansion of Out-of-the-Box (OOTB) support and pre-trained solutions for various process areas and industry verticals
- The market is expected to grow at a CAGR of 25-30% in the coming years, as automation continues to be a key priority for enterprises despite recessionary mindset
- Moreover, advances in IDP technologies to reduce training efforts, overcome challenges associated with unavailability of training datasets, improve extraction accuracy, and offer self-serve capabilities to enterprise users will drive further adoption

¹ Based on revenue estimates of 36 IDP technology providers projected out to estimate the total IDP technology provider market. It does not include revenue generated by service providers, consultancy firms, or system integrators
Source: Everest Group (2023)

02

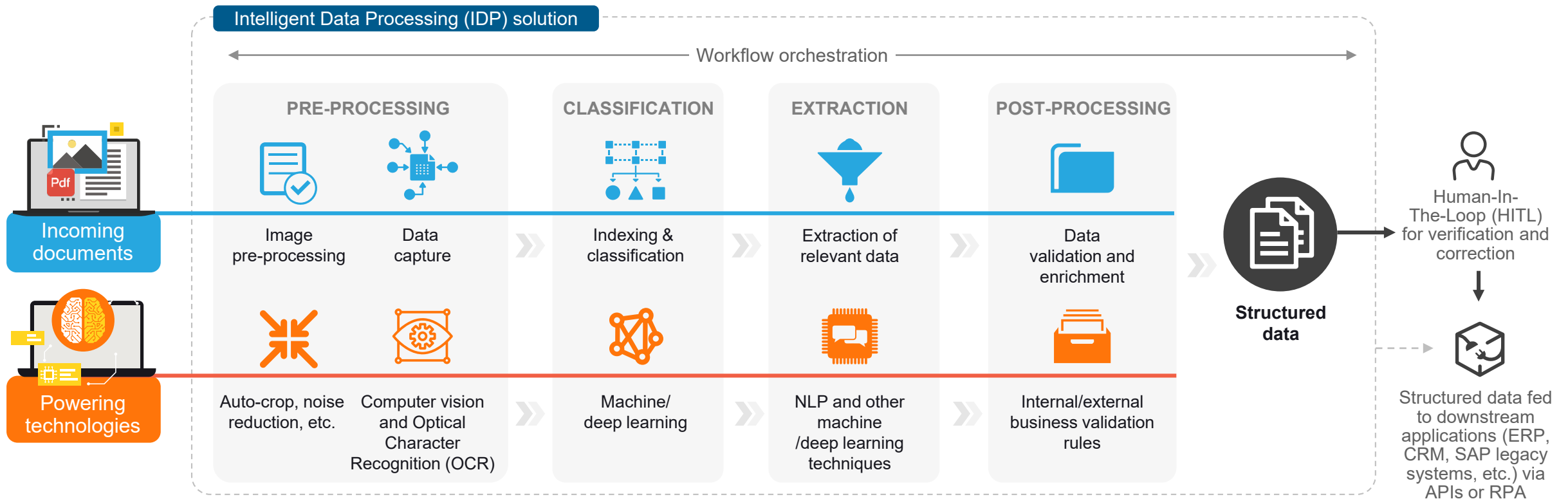
Understanding IDP

- Construct of an enterprise-grade IDP solution
- Core technologies powering IDP solutions
- Intricacies in document and data types handled using IDP
- Measuring and improving effectiveness of IDP solutions
- Benefits of IDP solutions
- IDP deployment use cases

Understanding enterprise-grade IDP solutions (page 1 of 2)

IDP software solutions blend the power of AI technologies to efficiently process all types of documents and feed the output into downstream applications

Enterprise-grade IDP solution



Understanding enterprise-grade IDP solutions (page 2 of 2)

Different steps involved in document processing for an enterprise include pre-processing, classification, extraction, post-processing, and HITL capabilities



Image pre-processing

The document quality is improved – using techniques such as auto-crop, brightness equalizer, motion blur and color marks removal, and reorientation – to increase the accuracy of OCR results.



Data capture

OCR and computer vision help to convert the data from the incoming documents into a machine-readable format.



Indexing and classification

As document batches are ingested into the system, the IDP solution identifies and classifies documents according to their types, such as invoices, personal IDs, and application forms, using machine/deep learning algorithms such as text mining, which helps in better understanding about the relevant data to extract.



Extraction

The solution leverages Machine Learning (ML) and rules-based extraction to obtain relevant information based on the business user requirement.



Post-processing for data enrichment and validation

It includes data normalization, data enrichment, and validations against pre-defined taxonomies, data dictionary, business rules, and third-party database to ensure accuracy of information and transformation of data into semantics of the organization.



HITL validations

Based on the confidence level threshold, documents are routed to human operators for validation and correction of the extracted output. The corrections are fed into a continuous feedback loop, ensuring continuous improvement of the ML model.

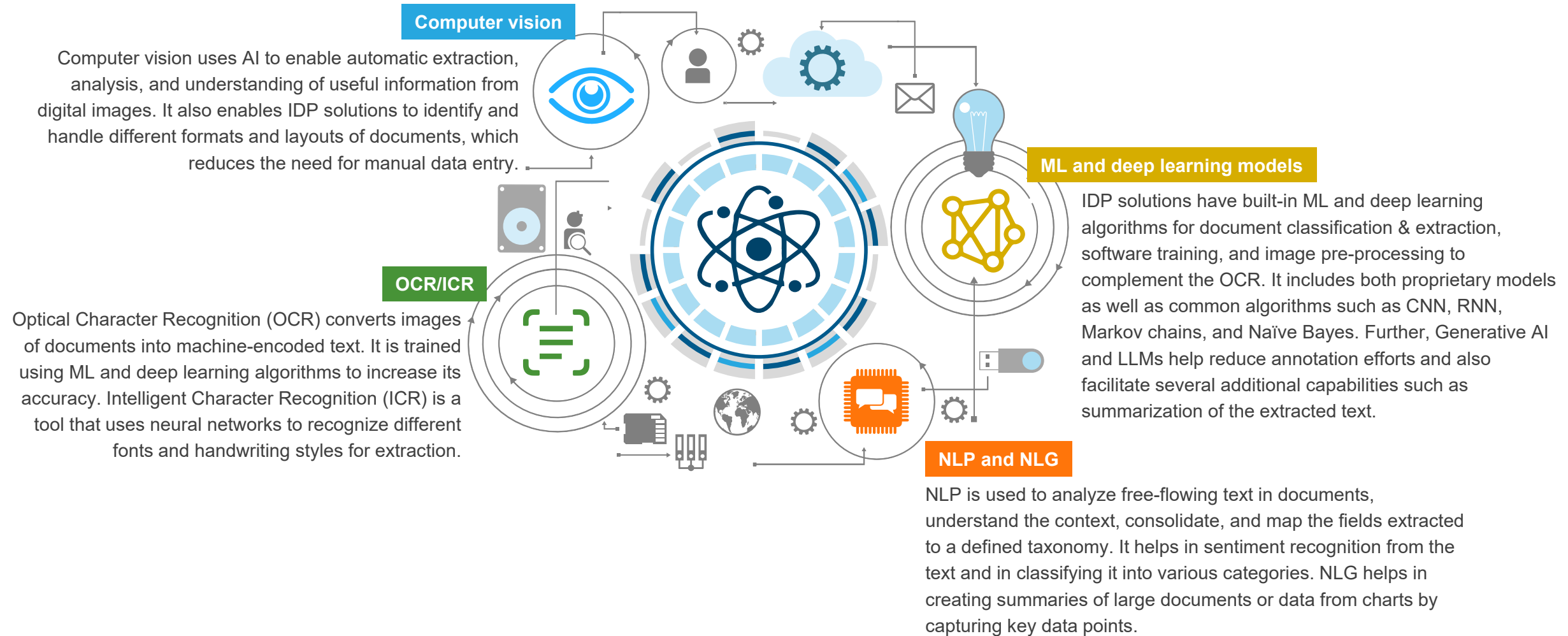


Data entry into downstream applications

The extracted and validated data is then fed into downstream applications through APIs or RPA robots.

Core technologies for IDP solution

IDP solutions are powered by underlying core technologies such as computer vision, OCR, machine and deep learning models, and NLP

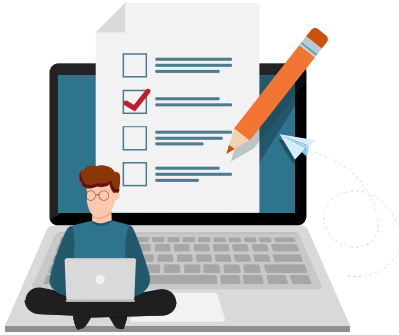


Key document types processed using IDP

IDP solutions can process varied document types including structured, semi-structured, and unstructured documents

01

Structured documents



These documents can be mapped into a pre-defined template with a fixed layout and tags to separate semantic elements.

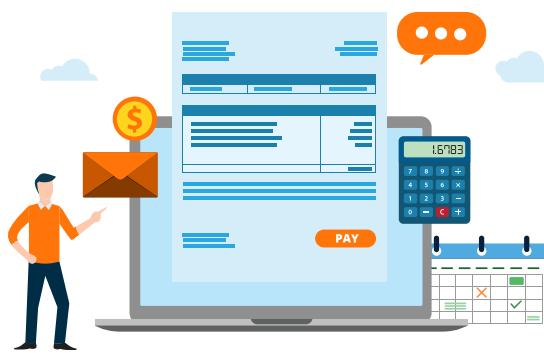
Processing capabilities

- OCR recognizes printed characters and converts images into machine-readable text
- AI/ML helps in handling variance in quality of documents during extraction
- Post-training, high accuracy and STP levels are possible with less human involvement

E.g., standard forms such as taxation, vehicle registration, medical, and education-related

02

Semi-structured documents



Documents which either have variability of layout or of semantic expression but contain some general keys, similar to the documents' organizational structure.

Processing capabilities

- AI/ML is used to train the system to identify, classify, and extract relevant information
- Tags are used, which can be linked to a position or visual elements or a key phrase
- Reasonable accuracy and STP levels can be expected, with some human intervention

E.g., invoices, purchase orders, bill of lading, paystubs, and insurance claims

03

Unstructured documents



Documents that do not conform to a pre-defined data structure and lack keys to separate semantic elements such as text-heavy documents or videos.

Processing capabilities

- NLP is used to interpret and extract data from free-flowing text in natural language
- NLG is used to process information and documents for summarization
- Lower STP levels, as higher degree of HITL is required for the desired output

E.g., annual reports, contracts/agreements, emails, prescription, and research papers






Key underlying data types processed using IDP

Data types within documents significantly contribute to complexities in document processing using IDP






Intricacies in document processing (page 1 of 2)

Overall, the complexity of document processing can be influenced by a variety of factors, including the quality of the incoming documents, document layout, and the data type processed

Factors	Description	Low complexity	Medium complexity	High complexity
 Quality of incoming documents	Quality is defined by legibility of the document	High quality document with clear and well-defined text, with no smudges, blurs, or distortions	Medium quality documents that have some minor inconsistencies and require pre-processing capabilities	Poor quality documents including blurred texts, overlapping data points, and improperly scanned documents
 Document layout/structure	Documents that have a defined template or have some variability in layout	Structured document with relatively simple and standardized layout and well-defined fields for extraction, such as forms	Semi-structured documents with some variability in layout or of semantic expression, but contain some general keys, such as invoices and purchase orders	Unstructured documents with highly variable structure and lack keys to separate semantic elements, such as scientific articles and annual reports
 Scope of processing capabilities	Types of document processing tasks performed using IDP solutions	Simple document processing tasks such as classification and extraction	More complex document processing tasks, such as data extraction, entity recognition, or sentiment analysis	Advanced tasks such as document comparison, summarization, and intelligent search
 Data type for extraction	Data types to be extracted using IDP solution are based on the customization and level of training required for the system	Simple bordered tables and printed text	Check boxes, barcodes, borderless tables, and handwritten block text	Cursive handwriting, signature, logos, stamps, multi-page tables, charts/graphs, audio, and video
 Document variability by source	The number of document origination sources (e.g., number of carriers sending bills of lading) will drive variability and complexity	Small datasets from limited sources	Medium-sized datasets with several hundred of document sources	Large datasets with thousands of documents

Intricacies in document processing (page 2 of 2)

Factors such as document variability by source, applicability of validation rules, and availability of training datasets also impact the complexity of document processing

Factors	Description	Low complexity	Medium complexity	High complexity
 Variability in data fields	Data field level variability requires continuous configuration for higher accuracy of extraction	Simple variability in metadata identifiers for the same information in the document (e.g., date and Dt.)	Data fields at multiple places without a key value pair to identify, such as name, date, and address	Data appearing in multiple languages, duplication of data fields within the document, etc.
 Business rules validation	Business rules for validation and the logic that needs to be applied for extracting relevant information	Simple mathematical rules for verification of extracted data	Configuring and managing business rules from within the platform for removing duplications during reconciliation	Configuring and managing highly conditional business rules for validation from external database through API integration
 Availability of training datasets	Amount of input data or sample documents available to train the AI/ML model	Availability of large volume of sample documents for training	Availability of moderate number of sample documents or datasets for training	Availability of limited sample documents, necessitating synthetic data generation

Measuring and improving effectiveness of IDP solutions

Common KPIs to measure and monitor IDP effectiveness



Straight Through Processing (STP) rates



Accuracy rates



Turnaround Time (TAT)



Speed of processing/throughput

Fallacies in measuring IDP effectiveness

- The performance of an IDP solution is commonly tracked using KPIs such as STP or touchless automation and accuracy rates. Enterprises often expect IDP to deliver high accuracy and touchless automation rates
- However, achieving exceptionally high field-level extraction accuracy in the production phase is a challenge due to high upfront ML model training costs and subsequent model drift due to real-world data
- For use cases requiring higher STP rates, accuracy gets impacted due to the lower confidence thresholds. In a few other use cases, certain metrics such as compliance and error reduction have a greater relevance and a direct impact on the turnaround time
- Enterprises face challenges in achieving exceptional results across multiple metrics due to the underlying trade-offs, which subsequently results in mismatch of enterprise expectations and outcomes delivered by IDP solution

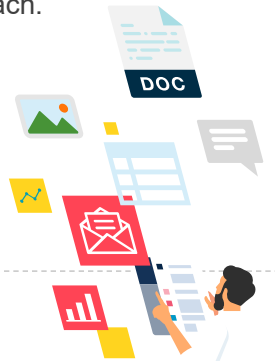
Role of HITL augmentation for superior outcomes

- HITL augmentation plays a critical role in helping enterprises achieve high quality output in terms of accuracy, faster turnaround time, and greater compliance
- During processing, the system alerts human reviewers to specific documents or fields based on the system's confidence level about its accuracy. If a document does not meet the confidence threshold or business validation rule, it is routed to a reviewer for verification and validation
- The human agent can review and rectify false positives, false negatives, or misinterpretations made by the IDP solution, and thereby help improve output
- Further, the verification information can be fed back into the model to retrain the algorithm and improve the overall model performance
- Depending on the criticality of the fields to be extracted, different confidence thresholds can be set up to enable HITL augmentation to ensure higher accuracy

Combining human expertise and judgment with AI algorithms is the key to enhance output of the IDP solution and achieve higher efficiency, accuracy, and overall productivity gains.

IDP solutions offer a range of operational, cost, and strategic benefits to enterprises






Employee productivity	Operational efficiency	Improved governance and compliance	Cost savings	Ensured business continuity	Enhanced customer experience
<ul style="list-style-type: none"> Employee productivity is a key driver of IDP adoption as it aids in offloading repetitive data entry tasks It also helps in improving accuracy of the documents processed with minimum human intervention and enables them to focus on high-value work 	<p>IDP speeds up processing time due to increased Straight Through Processing (STP), higher accuracy, and reduced errors.</p>	<ul style="list-style-type: none"> A robust audit trail can be gathered for all processing activities, enabling traceability and improving process governance Ability to add field-level business validation rules from multiple sources enhances efficiency and compliance 	<p>Overall processing cost for the huge volume of documents of different types (structured, semi-structured, and unstructured) is reduced due to lower number of manual workforce required.</p>	<p>It enables business continuity by building enterprise resilience through a technology-backed approach.</p> 	<p>IDP improves customer satisfaction when used in a customer-facing process such as customer onboarding by reducing wait times, speeding the completion of query resolution, and ensuring a consistent experience.</p>



IDP deployment use cases (page 1 of 3)






ILLUSTRATIVE

Business function	Objective	Business challenge	Solution
F&A 	Processing invoices	<ul style="list-style-type: none"> While common fields such as vendor name, date, and amount are present on all invoices, there is no uniform structure or accepted format for them Also, it is difficult to extract using conventional methods since different vendors use different key identifiers and places for their data 	<ul style="list-style-type: none"> IDP helps in classification, recognition, field extraction, and validation of data from invoices, ensuring better accuracy and STP It facilitates performing three-way matching by collating data from multiple sources and systems, thereby providing structured data to populate accounts, reports, and statements
HR 	Processing CVs and resumes	<ul style="list-style-type: none"> The challenge is to process resumes and CVs, which are in different formats, to extract information such as contact details of the candidate, their educational background, professional experience, and skill set This would help in automating the screening of resumes and the selection of candidates, accelerating the overall hiring process 	<ul style="list-style-type: none"> IDP comes with pre-trained ML models, which are further trained using different samples of CVs and resumes for relevant data extraction The HITL capability helps in model improvement, addressing errors, and modifying extraction guidelines It also facilitates integration of the enterprise's own pre-trained ML model with the solution through bring-your-own-model feature, ensuring greater flexibility and better outcomes
Procurement 	Processing bills of lading	<ul style="list-style-type: none"> These documents vary depending on the various forms of transportation, as well as the various consignee and payment types Also, low document scan quality makes it extremely difficult to recover details such as the cargo number, consigner and consignee information, and pricing and payment information 	<ul style="list-style-type: none"> IDP allows extracting relevant data irrespective of the supplier and document format, including borderless and nested tables, stamps, signatures, and logos It also comes with image pre-processing capabilities, thereby helping to improve the document image quality and the accuracy of extraction

IDP deployment use cases (page 2 of 3)





ILLUSTRATIVE

Industry	Objective	Business challenge	Solution
Banking 	Processing mortgage documents	<ul style="list-style-type: none"> • Mortgage-related documents are of different types including personal ID proofs, address proofs, and bank account details. The challenge is to extract data from a variety of documents with higher degree of accuracy • Also, the documents include sensitive data of personal identification and bank details, which can lead to privacy and data security issues 	<ul style="list-style-type: none"> • IDP offers computer vision and NLP-based extraction models to extract data from different types of documents • Its integration with third-party database helps matching and validation of the extracted information • It also offers redaction capabilities for masking/blurring of personal sensitive information and facilitates storing of the redacted data
Capital markets 	Annual reports and financial statements	<ul style="list-style-type: none"> • The challenge is not only to process documents and extract relevant data including revenue, expenses, assets, liabilities, and equity. It also requires context understanding for post-processing analysis of the free-flowing text documents • Also, generating reports and visualizations based on the extracted and validated data from the documents in the form of charts and graphs is needed by users 	<ul style="list-style-type: none"> • IDP comes with CV-based ML models for extracting relevant data. It also has context understanding capabilities such as sentiment analysis, summarization, document comparison, and intelligent search using NLP and NLG technologies • Its analytics component can generate insights based on the extracted data, such as identifying trends, comparing financial metrics, and benchmarking against industry peers through OOTB pre-built dashboards and custom reporting capability
Insurance 	Processing policy and claims documents	<ul style="list-style-type: none"> • The insurance industry sees huge volumes of policy contracts, claims documents, etc. One of the key challenges that enterprises face is to ensure compliance with the policy terms and to compare policies of different insurance providers • Also, the claims and verification for settlement includes huge volume of documents spanning multiple pages 	<ul style="list-style-type: none"> • IDP solution comes with NLP capabilities that help in understanding the context of the data to be extracted in a free-flowing text document with higher accuracy • It provides advanced capabilities for summarization and intelligent search, ensuring ease of compliance • It is also possible to compare/match certain sections and paragraphs of different documents using the IDP solution

IDP deployment use cases (page 3 of 3)



ILLUSTRATIVE

Industry	Objective	Business challenge	Solution
Healthcare 	Processing medical prescriptions and reports	<ul style="list-style-type: none"> • Handwritten/printed medical reports and prescriptions in a non-standardized structure present challenges such as readability of the text and document quality • Also, processing patient charts in medical reports for different tests undertaken and low tolerance of errors even at the character level, makes data extraction more complex 	<ul style="list-style-type: none"> • IDP solution comes with pre-processing techniques and helps extraction of information from a handwritten document using its underlying OCR/ICR capabilities • It facilitates businesses with semantic understanding through NLP and taxonomy-based models for interpretation of medical data • It helps process various data types such as graphs/charts, images, and QR codes into machine-readable format for downstream applications
Manufacturing 	Certificate of analysis and safety data sheets	<ul style="list-style-type: none"> • These documents can be in email attachments, file systems, or external portals. Also, they require pre-processing to remove noise • Validation checks from external databases are critical to ensure that the extracted data, which is prone to errors, is flagged 	<ul style="list-style-type: none"> • IDP performs pre-processing activities such as skew, orientation, and inverted text correction. It comes with NLP capability to interpret product description or hazard statements on safety data sheets • It uses ML models to extract data such as product name, batch number, and expiration date from different sources of origination and performs rules-based validation to ensure that the extracted data is accurate and complete

03

IDP market adoption and buyer trends

- IDP adoption trends by geography

- IDP adoption trends by industry

- IDP adoption trends by process area

- IDP adoption trends by buyer size

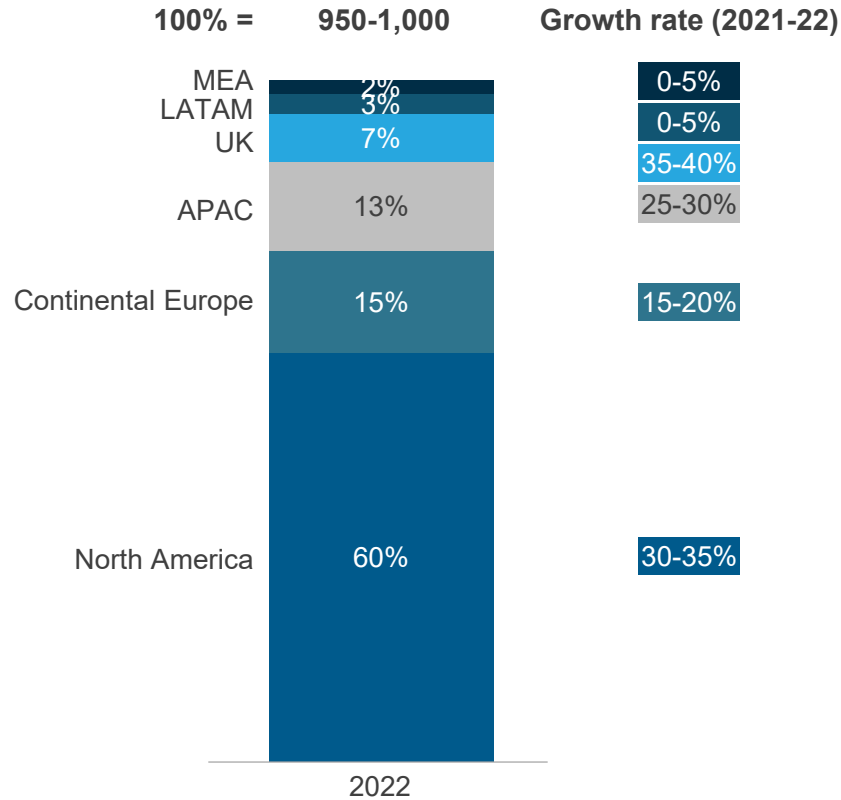
- Drivers for IDP adoption

IDP market overview and buyer trends (page 1 of 5)

North America continues to dominate IDP adoption with more than 50% market share

IDP software market size by buyer geography

License revenue in percentage; US\$ million



- North America continues to be the largest market for IDP software adoption, maintaining a healthy growth rate of around 30-35%. The growth is primarily driven by the mature banking and capital markets, insurance, and healthcare and pharma industries
- The UK region reported the highest uptick in demand over the last year, with a growth rate of over 35%. The rise is mostly propelled by quicker ROI and faster deployments achieved using pre-trained models and OOTB solutions offered by IDP providers
- APAC has also shown strong growth of around 30%. This can be attributed to the emergence of regional players and emergence of new technology partnerships. Also, the increasing support for regional languages by the IDP solution is playing an important role
- Factors such as lack of awareness, regulatory compliance, upfront costs, and lack of strong regional presence are contributing to the untapped potential in LATAM and MEA
- The growth in Continental Europe has been slower than the market average, predominantly due to cost cutting measures undertaken by enterprises due to recessionary tendencies

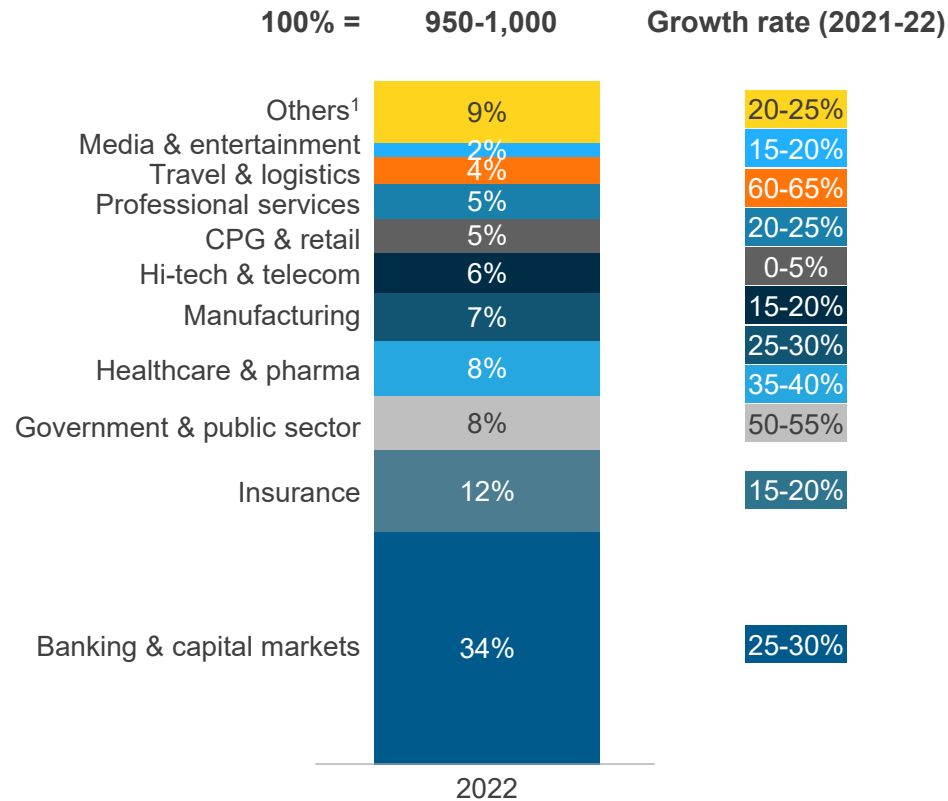
Note: Based on the capability assessment of 36 IDP technology providers
Source: Everest Group (2023)

IDP market overview and buyer trends (page 2 of 5)

The BFSI sector leads IDP adoption, while government & public sector recorded the highest growth

IDP software market size by buyer industry

License revenue in percentage; US\$ million



- Banking and insurance continued to be the largest adopters of IDP solutions and accounted for ~34% and ~11% of the IDP market, respectively. Major growth drivers include maturity of the industry in terms of use cases and additional post-processing analytics capabilities that the solution provides to generate deeper insights, while ensuring regulatory compliance
- Healthcare & pharma and manufacturing continued to experience strong growth, owing to the advanced capabilities of the IDP solution to handle complex use cases such as patient charts, lab reports, scientific papers, certificate of analysis, and safety data sheets. More verticalized pre-built solutions also service the growth in demand
- Government & public sector and travel & logistics have seen fast-paced adoptions post COVID-19, mainly due to the need for digitalization, modernizing operations, reducing costs, and higher transparency
- CPG & retail, hi-tech & telecom, and media & entertainment are largely untapped and have high potential for growth given the growing need to improve efficiencies and reduce reliance on manual processes. With increasing sophistication of the solution, we can expect faster growth and adoption in the future

¹ Others include energy, real estate, shared services, utilities, hospitality, and legal
 Note: Based on the capability assessment of 36 IDP technology providers
 Source: Everest Group (2023)







IDP market overview and buyer trends (page 3 of 5)

BFSI industry-specific processes and F&A account for significant share of the IDP market, while adoption grew rapidly across F&A and procurement process areas

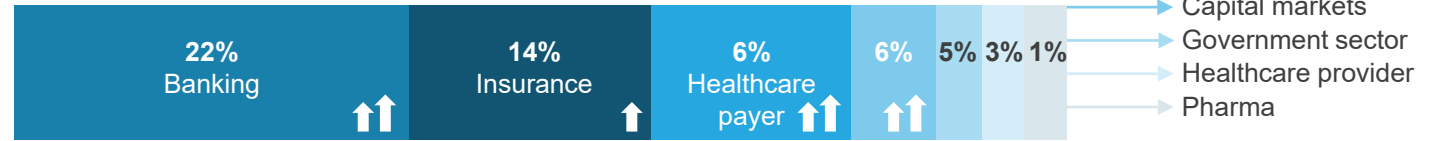
IDP software market size by business process/function

License revenue in percentage; US\$ million

100% = 950-1,000

	Industry-specific	57%	↑↑
	F&A	19%	↑↑
	Procurement	5%	↑↑↑
	HR	3%	↑
	Contact center	2%	↑↑
	Mailroom	1%	↑↑
•••	Others ¹	13%	↑↑↑

Growth rate: ↑ <15% ↑↑ 15-30% ↑↑↑ >30%



- IDP adoption within industry-specific processes continued to grow steadily. The BFSI sector has shown the highest growth based on the need to drive operational efficiency and achieve digital transformation in the financial services industry. Most prominent use cases being loan applications, insurance claims, compliance documents, customer onboarding, and fraud detection
- Apart from industry-specific use cases, IDP solutions are also being leveraged to improve the accuracy and speed of data extraction, classification, and reconciliation in the F&A sector for better decision-making
- IDP solutions are capable of undertaking mailroom operations such as sorting and routing incoming mail, digitizing procurement processes such as purchase orders and vendor management, and streamlining HR applications such as employee onboarding, benefits enrollment, and document management to attain process efficiency, operational costs, and productivity

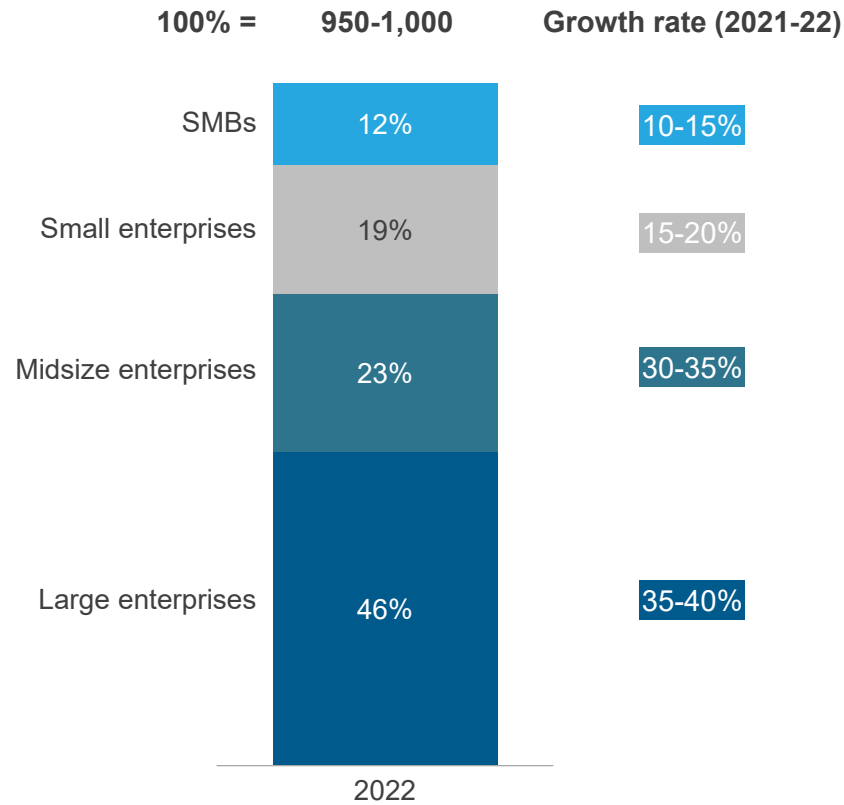
¹ Others include government customs form processing, food and beverages industry-specific processes, etc.
 Note: Based on the capability assessment of 36 IDP technology providers
 Source: Everest Group (2023)

IDP market overview and buyer trends (page 4 of 5)

Large enterprises hold a considerable share of the total market and exhibited the highest growth

IDP software market by buyer size¹

License revenue in percentage; US\$ million



- Large buyers continue to have the highest adoption with a healthy growth rate. This is attributed to the availability of advanced context understanding capabilities of the IDP solution and post-processing activities for gaining drilled-down insights on the data extracted
- Midsize enterprises have also experienced healthy growth in the market. This is because of the constant addition of new use cases where IDP solutions can be deployed for efficiency and productivity gains
- Progressive pricing options, SaaS delivery model, cloud-based deployments, and pre-built OOTB solutions are driving the adoption in small enterprises and SMBs

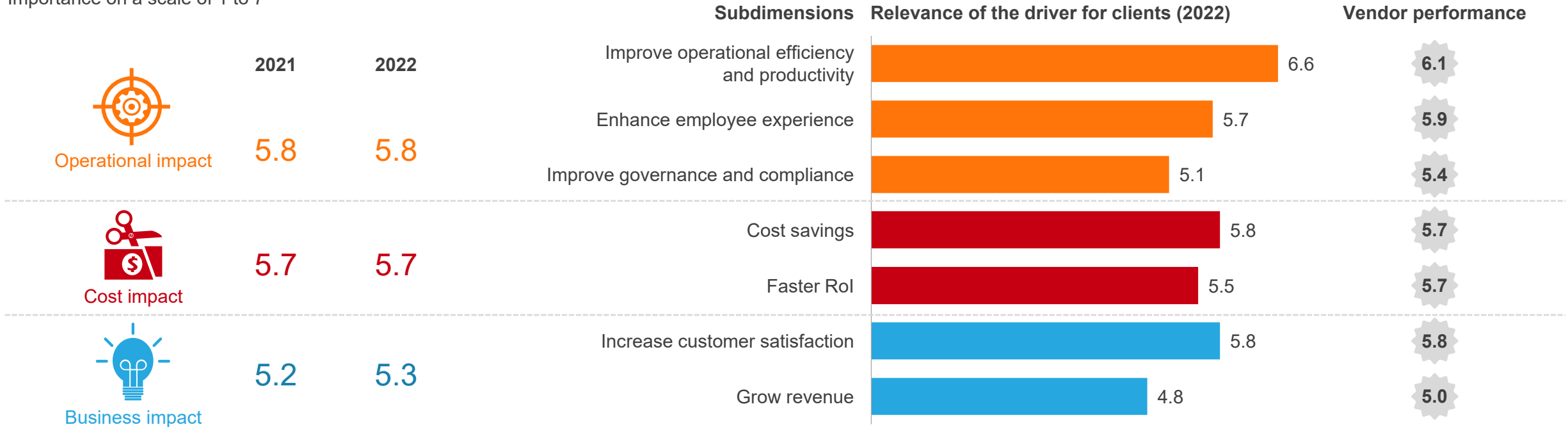
¹ Buyer size is defined as large (>US\$5 billion in revenue), midsize (US\$1-5 billion in revenue), small (US\$50 million-US\$1 billion in revenue), and SMBs (<US\$50 million in revenue)
Note: Based on the capability assessment of 36 IDP technology providers
Source: Everest Group (2023)

IDP market overview and buyer trends (page 5 of 5)

Improving operational efficiency continues to remain the most important driver for IDP adoption, followed by the need to drive cost savings

Key factors driving IDP adoption

Importance on a scale of 1 to 7



- Enhancing operational efficiency and productivity remains a critical factor for the adoption of IDP. In the past years, enterprises have recognized the importance of not only the business impact but also cost impact, particularly with the growing emphasis on faster Rol and enhancing customer satisfaction
- Providers have been successful in assisting enterprises to attain operational efficiencies, which has resulted in high satisfaction among enterprises

Sample: Based on feedback collected from 70+ enterprise buyers in 2023
 Source: Everest Group (2023)

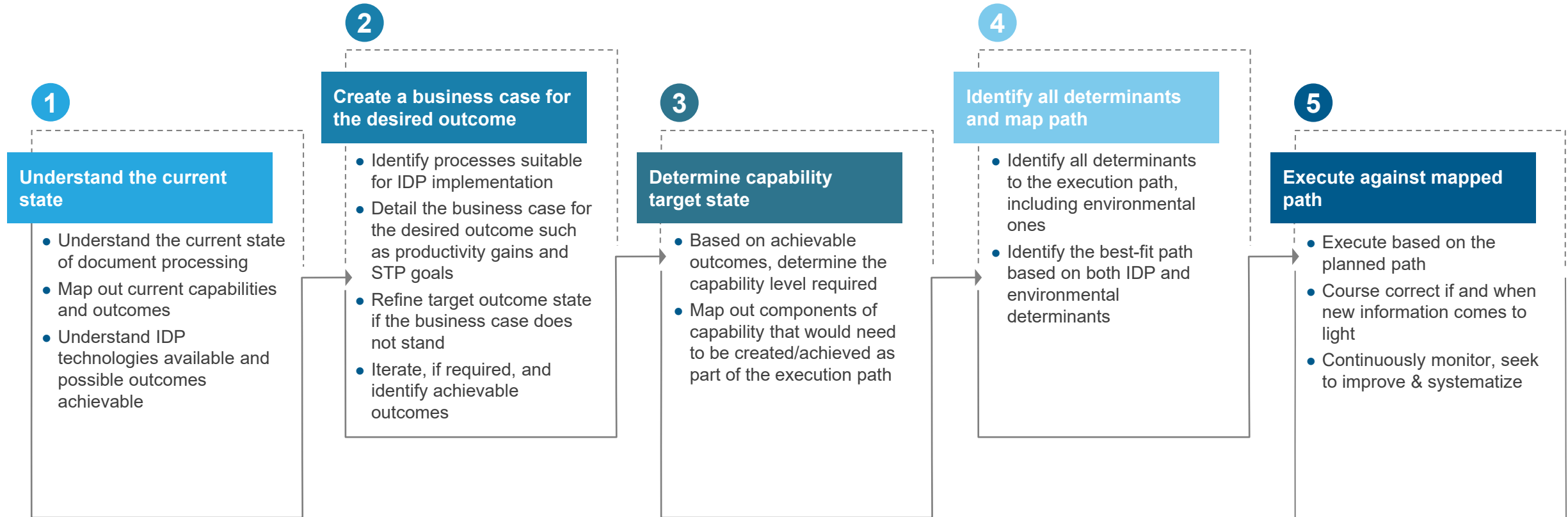
04

Enterprise IDP journey

- Understand the current state and market capabilities
- Building a business case – what is the right outcome target state?
- Determining the capabilities required to achieve identified outcomes
- Identify all determinants and map path
- Execute against the mapped path

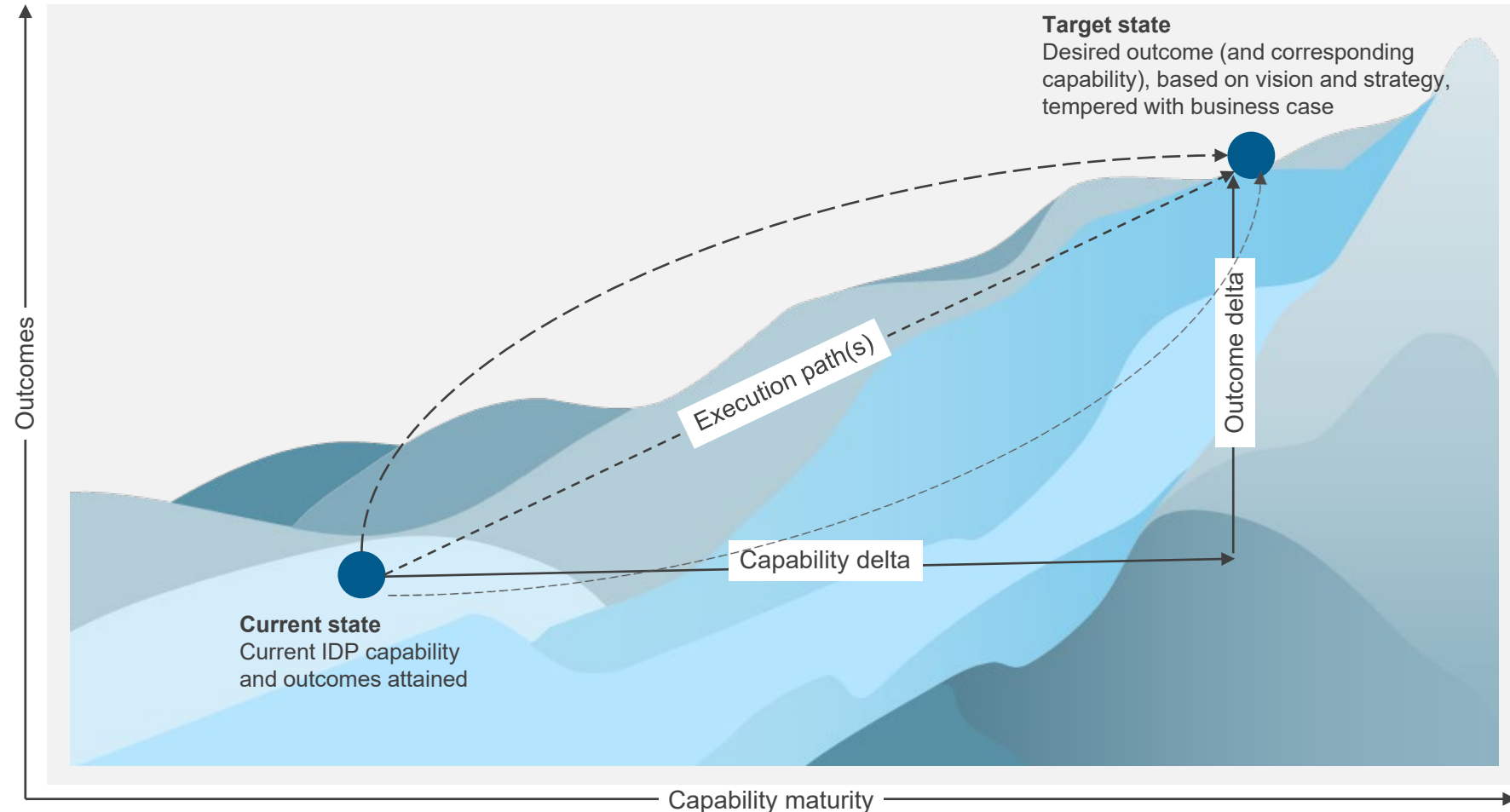
For enterprises that are thinking about IDP adoption programmatically across the organization, the following approach can be taken

Enterprises can break down their IDP journey into five distinct steps



Step 1: Understand the current state (page 1 of 2)

It is important for enterprises to understand their current and desired target outcome states to map a best-fit execution path for their IDP journey



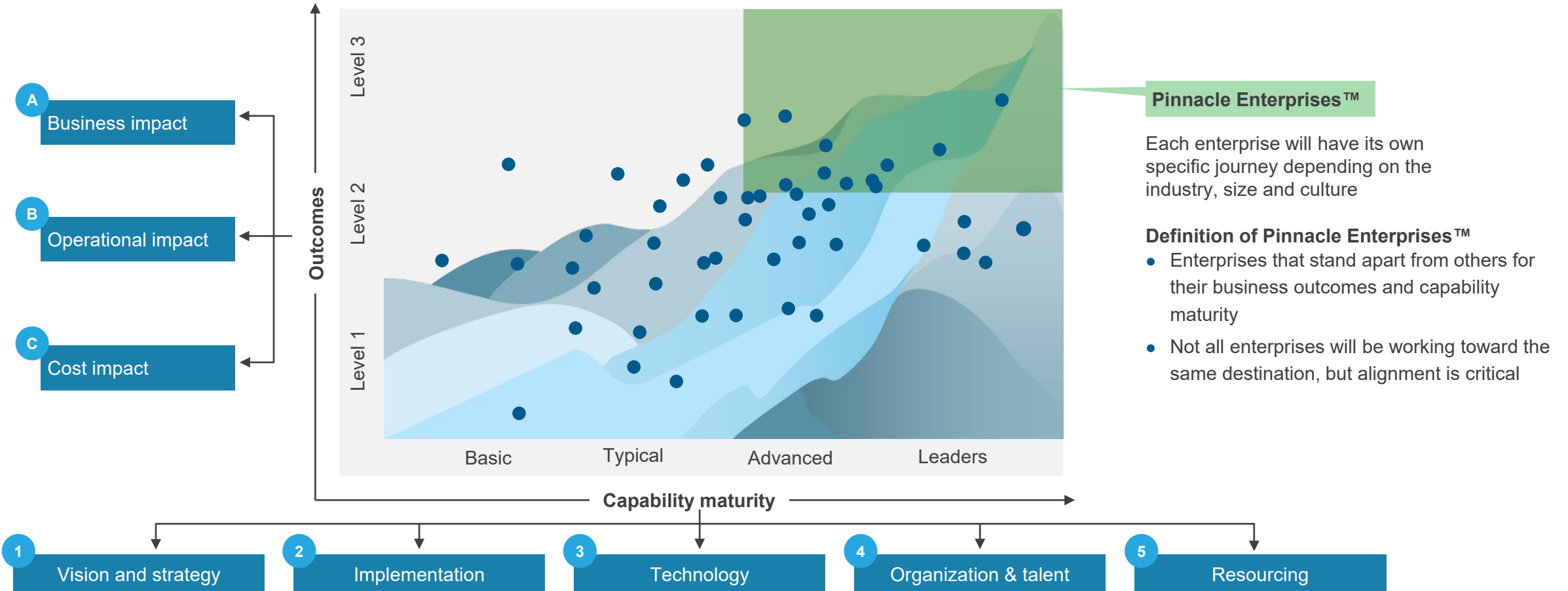
- The IDP journey for every organization begins with an understanding of its current state of maturity and its aspirational target state
- While the current and target states outline the gaps to be bridged, the actual execution path to be followed to bridge those gaps will depend on multiple factors, as described in subsequent pages

Step 1: Understand the current state (page 2 of 2)

The Pinnacle Model[®] provides a framework to help enterprises measure the IDP journey's current and target states, both in terms of outcomes and capabilities

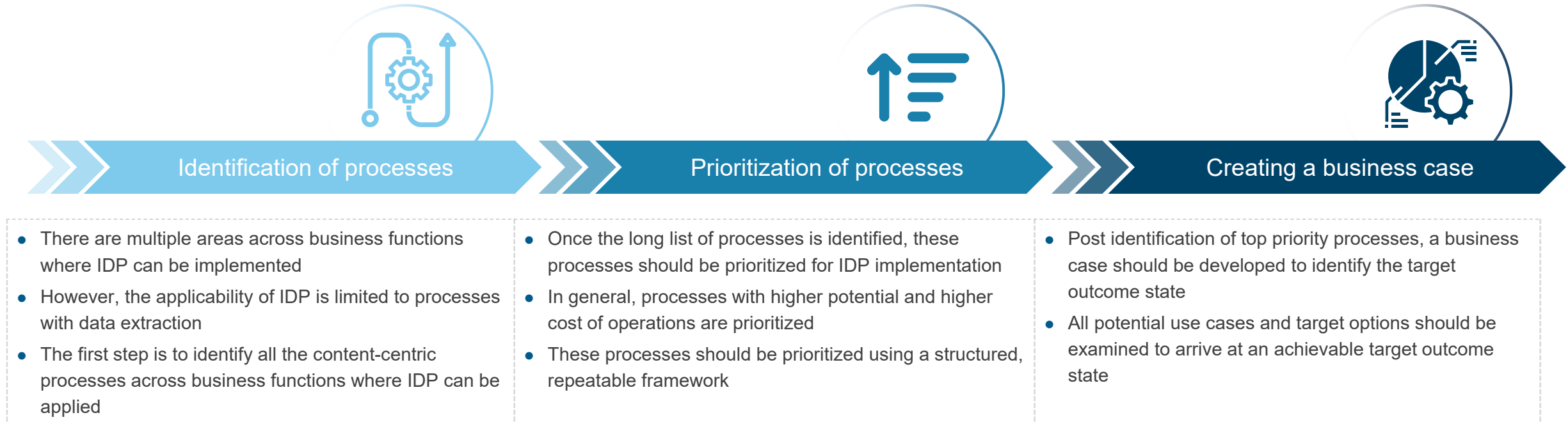
NOT EXHAUSTIVE

Everest Group Pinnacle Model[®] for mapping an enterprise's journey to become a Pinnacle Enterprise[™]



Step 2: Create a business case for the desired outcome (page 1 of 8)

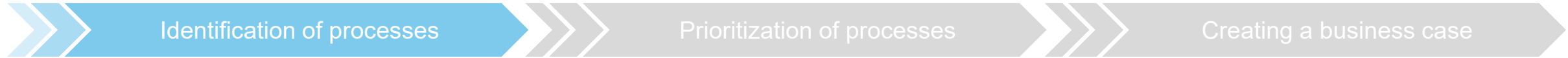
Creating a business case for prioritized processes helps define the target outcome state



Enterprises should identify and prioritize potential processes for IDP adoption. Once potential processes are identified, business case should be created to refine the achievable target outcome state.

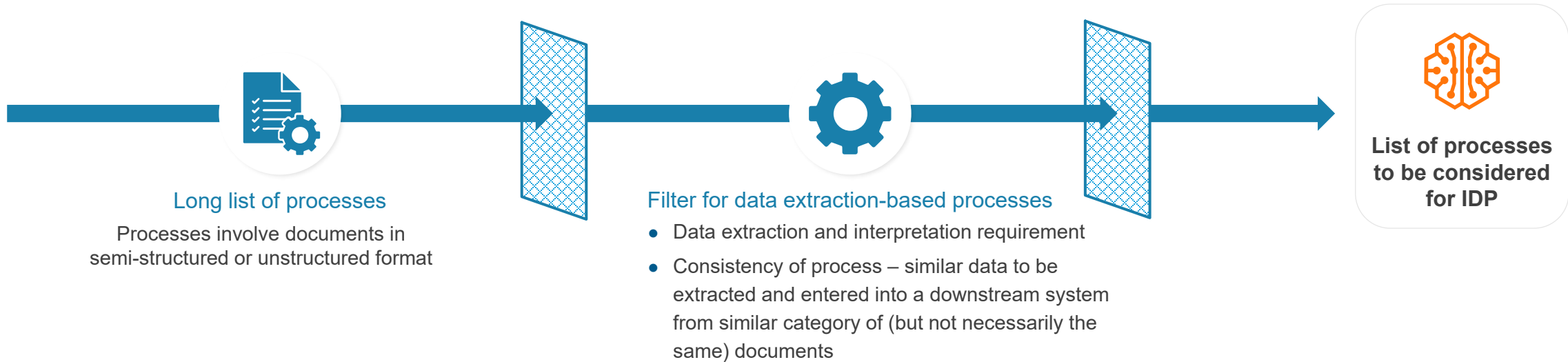
Step 2: Create a business case for the desired outcome (page 2 of 8)

List of data extraction-based processes should be identified through a filtering approach



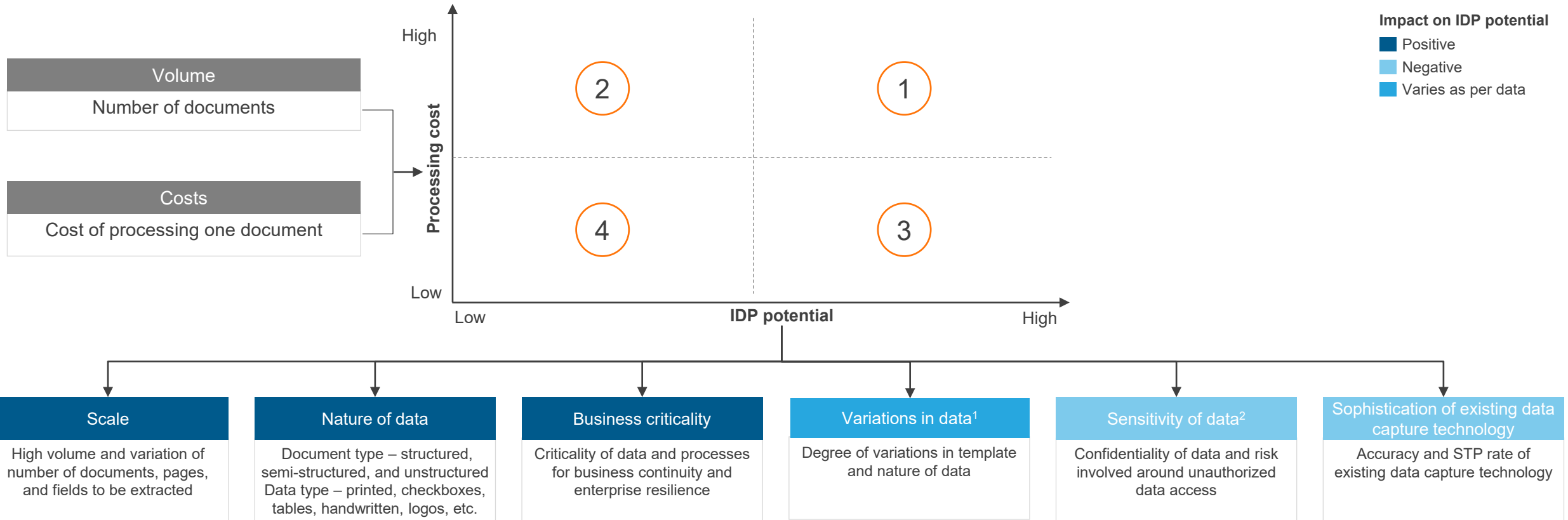
Applicability of IDP

Filtering criteria



Step 2: Create a business case for the desired outcome (page 3 of 8)

The identified data extraction-based processes should be prioritized for IDP implementation, leveraging a structured and repeatable framework



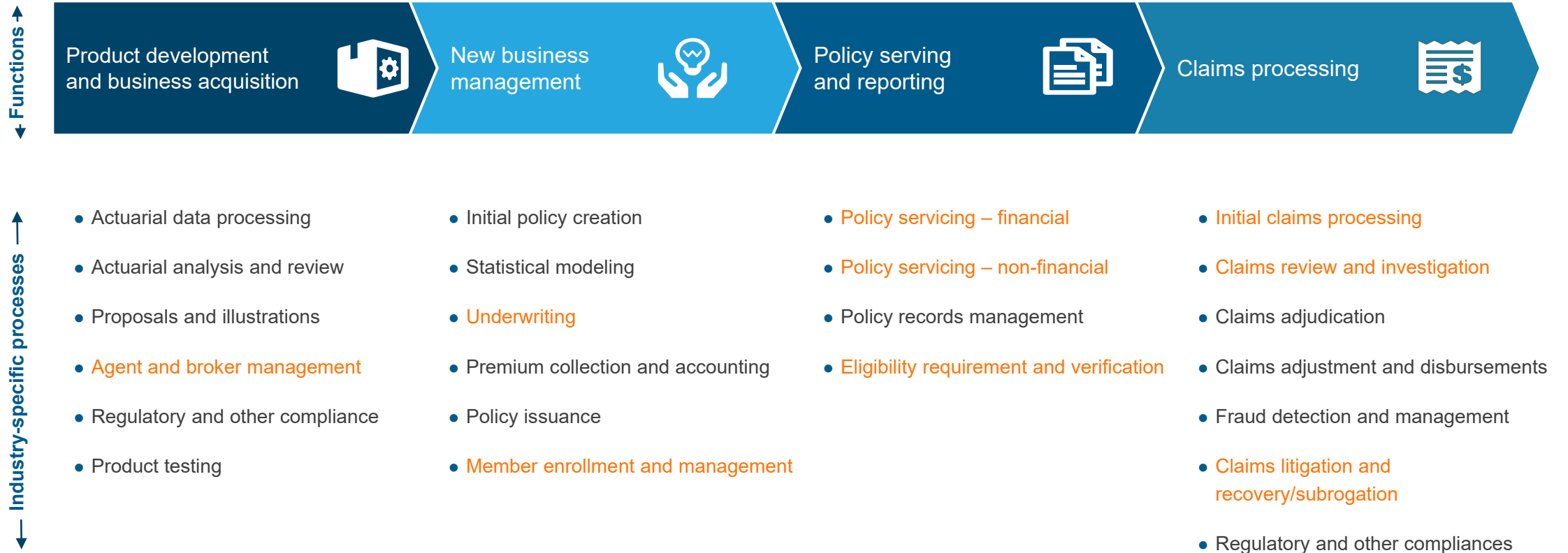
1 Variations in data can have both positive and negative impact on IDP potential. Please refer to the next page for more details
 2 Sensitivity of data is becoming a less critical parameter as there are various ways to circumvent this (such as redaction/masking of data)

Step 2: Create a business case for the desired outcome (page 4 of 8)

Illustration 1: Consider an insurance provider evaluating its P&C insurance business function for IDP implementation

Insurance industry value chain

Processes with high potential for IDP application

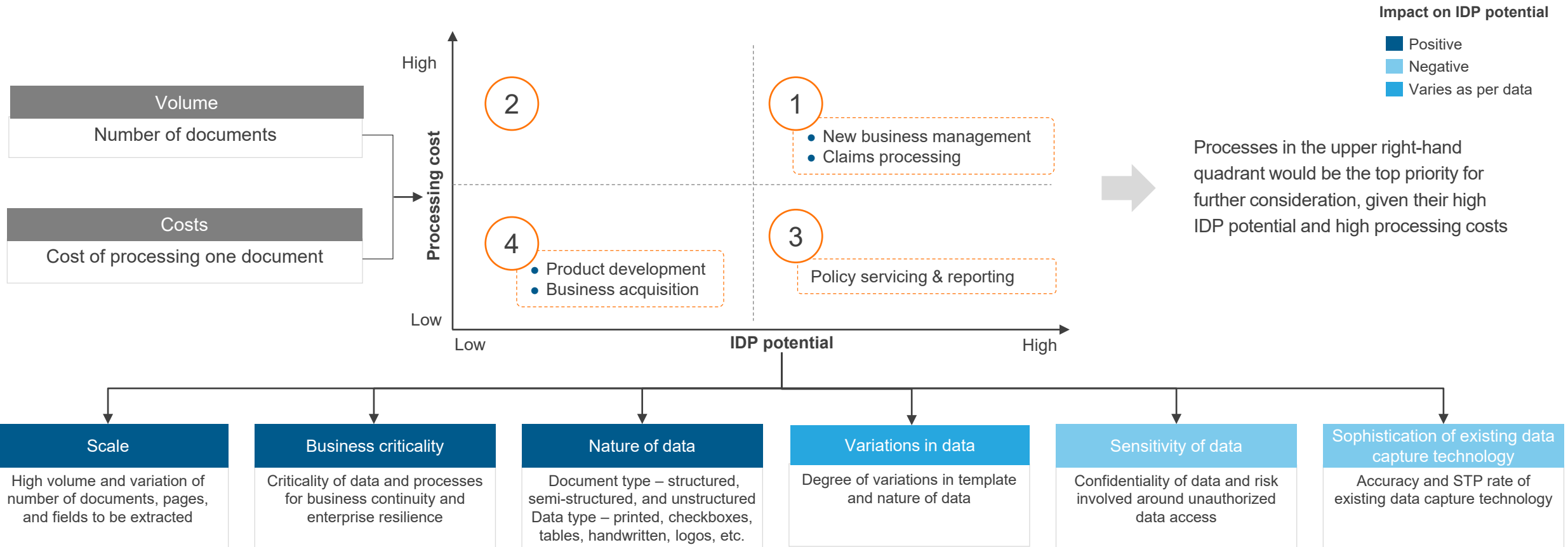


The processes highlighted across the value chain are not exhaustive but captures the most critical areas where the potential of IDP adoption is maximum. Also, the data extracted from these can lead to additional business benefits on other processes of insurance industry such as statistical modelling and claims adjudication.

Step 2: Create a business case for the desired outcome (page 5 of 8)

Once the processes are identified, ones with high processing costs and high IDP potential should be considered for further evaluation

ILLUSTRATIVE



Processes in the upper right-hand quadrant would be the top priority for further consideration, given their high IDP potential and high processing costs

Step 2: Create a business case for the desired outcome (page 6 of 8)

Create business case for the shortlisted process considering cost and benefit components

ILLUSTRATIVE



Business case components



Cost

- The cost of IDP implementation includes:
 - IDP software license, implementation, and maintenance costs
 - HITL augmentation cost, i.e., cost of resources involved in review and validation of the fields extracted with confidence level below the defined threshold
- IDP software license and maintenance costs remain the same across various outcome states. The license cost typically depends on the volume of pages and complexity of the document type processed



Benefits

- For the business case, enterprises should ideally include metrics that are quantifiable and relatively easy to measure, and directly link them to IDP implementation. Examples include potential cost savings due to automation and subsequent productivity gains due to STP or touchless processing
- Information/proxies around other business benefits, such as reduced error rates, easier data accessibility, user experience gains, increase in customer retention rates, and risk reduction and improvement in compliance, may not be available prior to or at the time of business case creation

Step 2: Create a business case for the desired outcome (page 7 of 8)

Considerations when creating a business case also include STP rates and upfront training costs

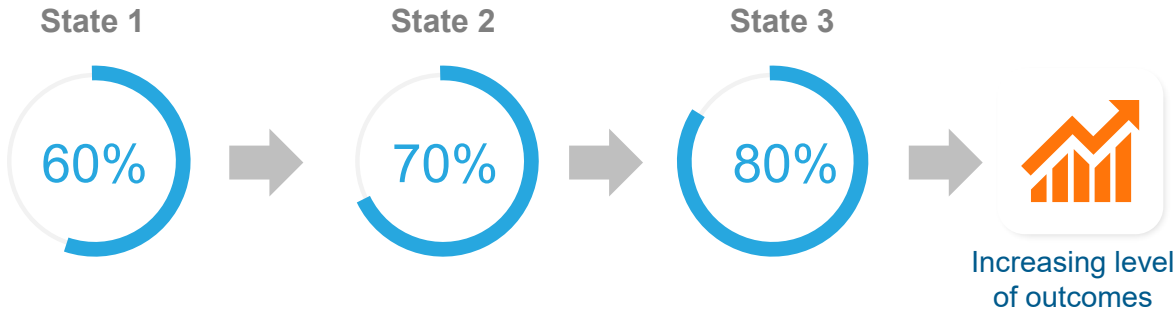
ILLUSTRATIVE



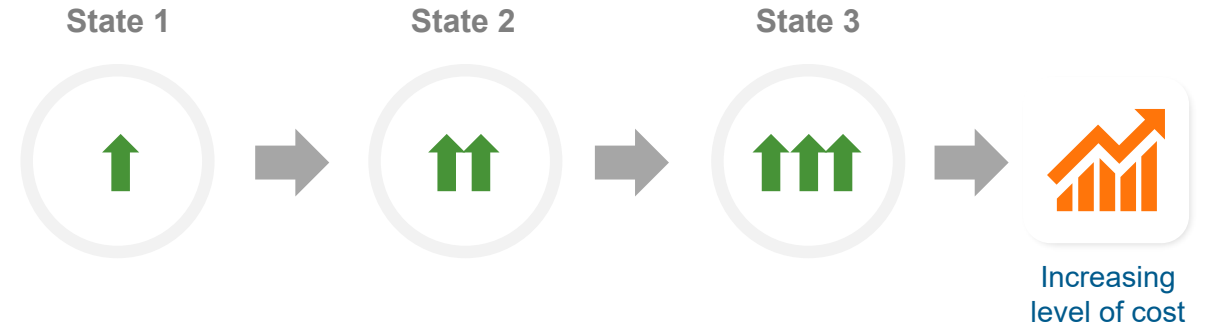
- Typically, business cases are created around STP rates and accuracy level, which reflect the direct cost reduction and productivity gains
- Additionally, with higher STP, the HITL augmentation costs go down as lesser documents are directed for human review
- However, accuracy of the IDP solution and consequently STP rates increase with training, i.e., the system learns and gets better as it is exposed to larger volumes of documents
- As we move from one target state to another (higher STP and accuracy), the system needs to be trained on increasing volumes of documents. This leads to an increase in training cost as we move from one state to another, all other costs remaining

STP rate

STP rates differ based on complexity/nature of document



Training cost



— Sophistication of IDP solutions —

— Static cost (includes license cost, implementation cost, and maintenance cost) —

Note: In some industries such as banking and financial services, compliance requirements and sensitivity of data always demands HITL, even though STP can be achieved

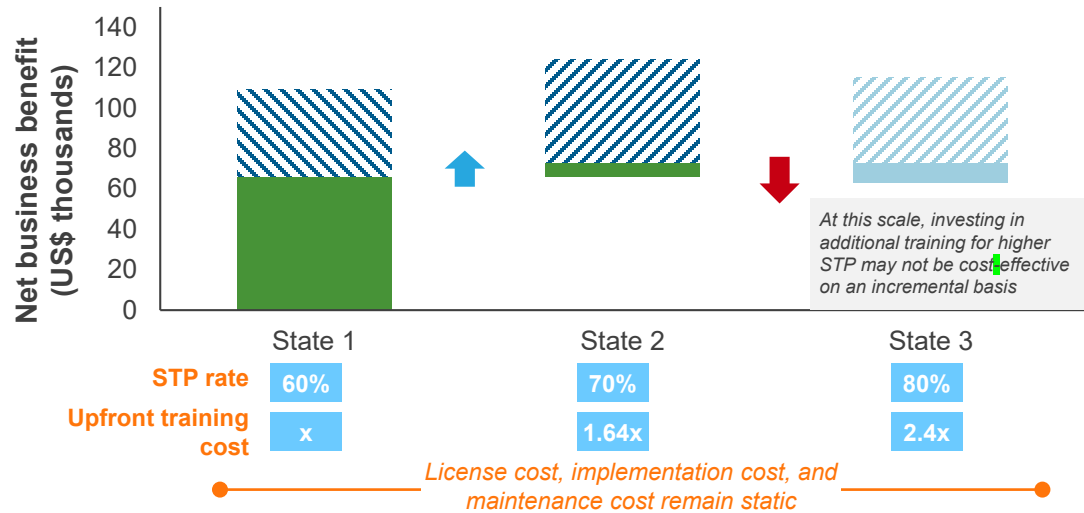
Step 2: Create a business case for the desired outcome (page 8 of 8)

Enterprises need to target appropriate STP given the possibility of diminishing returns

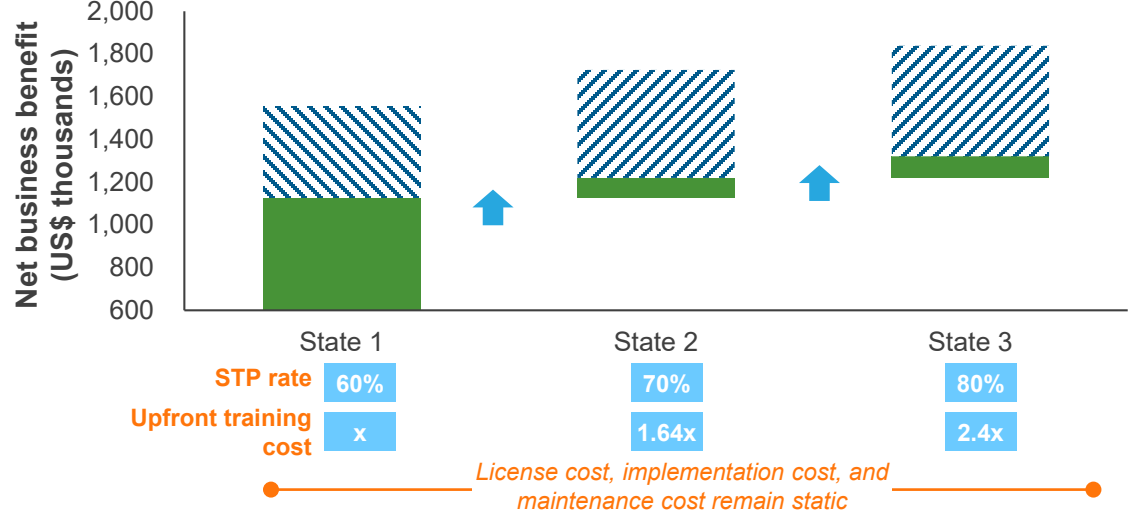
ILLUSTRATIVE



Example 1: Invoice processing; 500,000 invoices



Example 2: Invoice processing; 5,000,000 invoices



- To create business case, enterprises must consider the business benefits (including cost saving due to productivity gains and other benefits such as improvement in customer and employee experience) and underlying costs (including training costs and HITL augmentation costs) for different target states
- For small-scale IDP implementations and particular confidence threshold for extraction, after a certain STP rate, enterprises may not derive incremental returns on the net business benefit due to higher investments in additional training costs. However, as enterprises increase the scale of implementation, the net business benefit can be higher at similar STP rate owing to the higher productivity gains
- Thus, enterprises need to evaluate the target STP rate considering the upfront training cost and determine the net business benefits that can be achieved based on the scale of implementations

Step 3: Define target capability state

Having determined an achievable sets outcome, enterprises should then seek to map out corresponding capability requirements to achieve the desired outcome

ILLUSTRATIVE

 Required capability

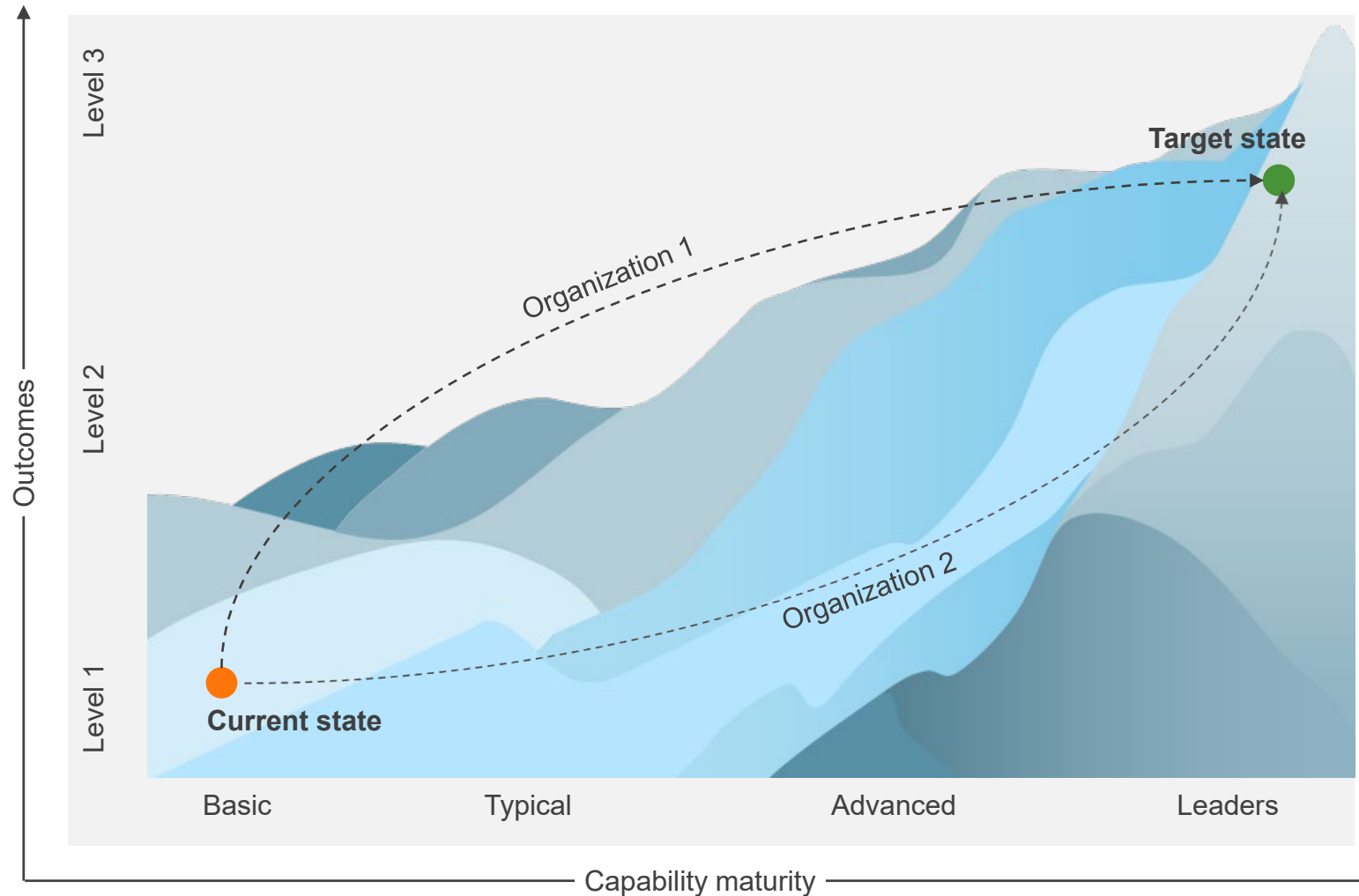
Capability elements	Basic	Typical	Advanced	Pinnacle
Software learning	Manual generation of training data sets during review	Automatic generation of training batches during manual review; automatic feeding of data sets into the system for training	Automatic generation of training batches during manual review along with feature for enterprise users to approve training sets to improve accuracy, approval mechanism at admin level as well	Automatic generation of training batches during manual review, feature for users to approve training sets for better accuracy, approval mechanism at admin level; continuous learning through HITL functionality; using advanced software learning techniques such as transfer learning and n-shot learning
Classification of documents	Do not have the ability to automatically classify documents	Ability to identify discrete documents with low accuracy, leveraging basic statistical approach	Ability to identify discrete documents with medium accuracy, leveraging basic ML-based approach	Ability to identify discrete documents and different pages within a stream of documents with high accuracy, leveraging advanced neural networks
Flexibility with ML algorithms	Different pre-built ML models for different use cases / document types	Different pre-built ML models for different use cases / document types and ability to choose the model by the enterprise user	Different pre-built ML models for different use cases / document types with a feature recommending the best ML model for the user	Feat to ch ability using integ mod

Capability elements	Basic	Typical	Advanced	Pinnacle
Sophistication of document processing	Basic OCR for digitizing content	OCR- and auto ML-based; document classification, data capture, and extraction using ML and validation	Multi-OCR, auto ML, and NLP; document classification and extraction using multi-modal approach and business rules validation; supporting sentiment analysis and entity-based extraction	Multi-OCR, domain ontology, deep learning, auto ML, NLP, NLG, and LLM; document classification and extraction using multi-modal approach and continuous learning; utilizing intelligent search, summarization, intent analysis, and validation
Complexity of the data handled	Block letters (typed) and simple tables	Block letters (typed) and nested and borderless tables	Block letters (typed or handwritten), tables spanning multiple pages, checkboxes, bar codes, QR codes, and logos	Block letters (typed or handwritten), checkboxes, bar codes, QR codes, logos, stamps, charts, signatures, cursive writing, and audio/video
Pre-built use cases	No pre-built use case	Simple use cases involving semi-structured data such as invoice processing, customer onboarding, and claims	Complex use cases involving unstructured data such as contracts and legal documents	Use cases that involve extracting information from free-flowing text as well as NLG
Hosting type	Physical, desktop-based	On-premise, server-based	Private cloud-based, hybrid	Public cloud-based, hybrid
Ancillary technologies	Stand-alone IDP solution	IDP solution integrated with BPM tool and RPA	IDP solution integrated with BPM, RPA, process mining, analytics, and other AI solutions	IDP solution integrated with BPM, RPA, process mining, conversational AI, analytics, and AI solutions including LLM models

Note: Refer to pages 72-84 to understand the Capability Maturity Model (CMM)

Step 4: Identify all determinants and map paths (page 1 of 5)

Given the same current and target states, enterprises' culture, structure, and other environmental determinants influence the routes they take







Two enterprises starting their journeys at the same low level of IDP maturity and wishing to reach the same advanced target state may take significantly different execution paths; the path would largely depend on environmental determinants.

Step 4: Identify all determinants and map paths (page 2 of 5)

IDP execution paths can be broken down into four key phases

Everest Group breaks down enterprises' IDP execution path into four key phases

	Phase 1	Phase 2	Phase 3	Phase 4
	Planning 	Piloting 	Scaling up 	Steady state 
	<ul style="list-style-type: none"> • Business case • Proof of Concept (POC) • Tool selection • Develop IDP strategy 	<ul style="list-style-type: none"> • Select initial use cases • Run pilot projects • Monitor performance • Involve automation CoE 	<ul style="list-style-type: none"> • Refine IDP strategy • Establish / refine the IDP governance model • Scale up across functions / geographies • Embed IDP capabilities in the automation CoE • Identify new opportunities 	<ul style="list-style-type: none"> • Institutionalize the governance model • Continuously improve • Create awareness
Key Focus	<ul style="list-style-type: none"> • Overcome conceptual barriers to IDP implementation and build awareness • Identify opportunities • Develop the business case • Align stakeholders (such as senior management and IT) to get backing for the POC • Select tools/partners • Identify skills sets required • Develop an implementation approach and roadmap 	<ul style="list-style-type: none"> • Get executive backing and funding • Develop/train skill sets for IDP implementation • Select and prioritize processes • Set up infrastructure and design architecture • Implement initial use cases / pilots • Monitor performance • Involve automation CoE • Raise awareness and enable upskilling/reskilling 	<ul style="list-style-type: none"> • Act on lessons learned to refine the IDP strategy and establish/refine the IDP governance operating model • Scale training and teams • Scale up across functions/geographies • Scale up upskilling/reskilling of resources • Set up a dedicated team and continuously identify new opportunities for IDP adoption 	<ul style="list-style-type: none"> • Ensure operations with robust governance and controls • Integrate a culture of innovation and design thinking across the organization

Step 4: Identify all determinants and map paths (page 3 of 5)

Use case: consider two distinct organizations with different characteristics, both seeking to reach similar target states from similar current states in the IDP journey for their expense management process

ILLUSTRATIVE

Scope



Consider two organizations with the same current state and achievable target state

- **Current state:** The organizations run a single T&E system with a workflow for submission of expense claims along with documents of proof by employees, followed by a manual verification and approval process. All the expense details are manually entered by the employees. Each organization also has seven FTEs currently employed in the manual expense proof validation process
- **Achievable target state:** 60% STP rate for all expense proof documents

The ideal execution path for each organization would depend on environmental determinants such as those detailed below:

Environmental determinants

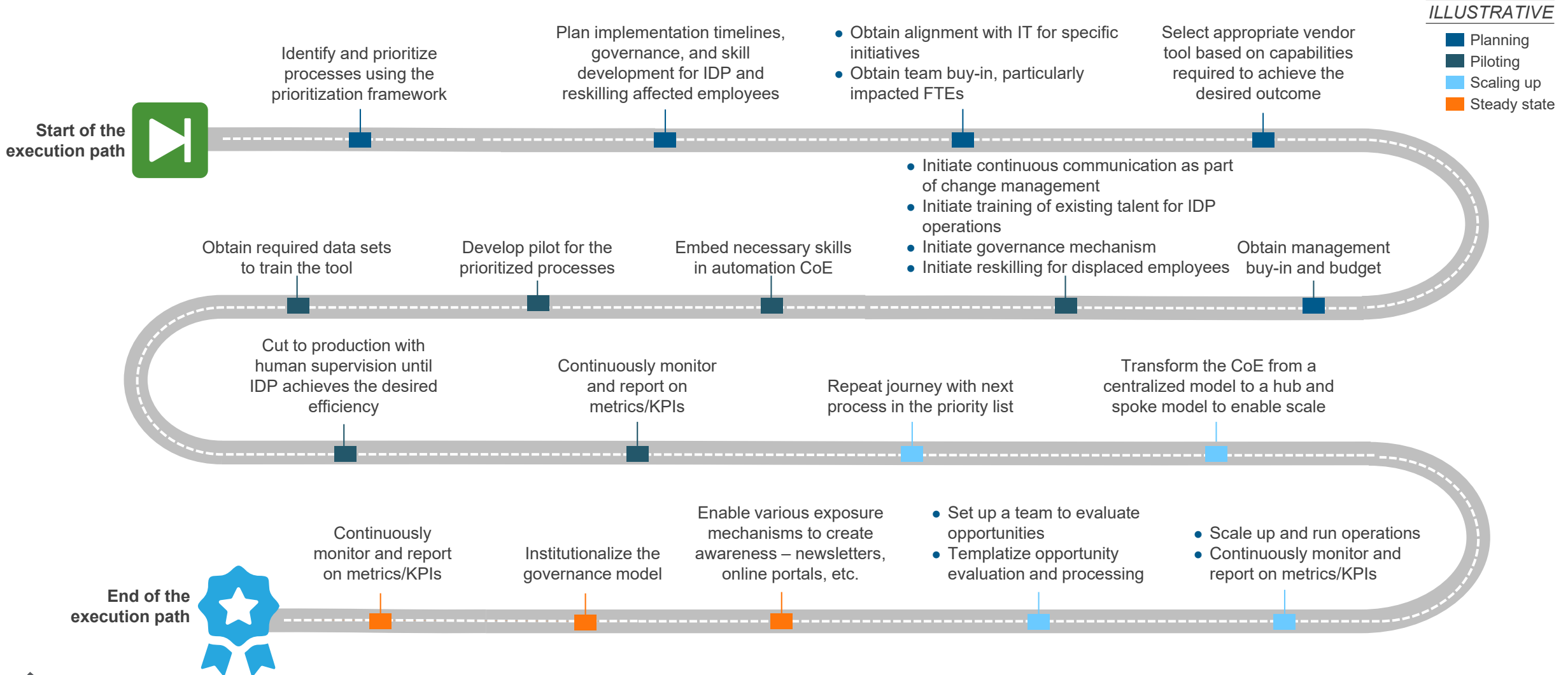
Organization A – conservative energy firm

Organization B – e-commerce firm

Centralized	----- Organization structure -----	Decentralized
High people centricity	----- People centricity -----	Low people centricity
Operations driven	----- Initiating stakeholders -----	IT driven
Low risk appetite; heavy regulation	----- Risk appetite -----	High risk appetite
Centralized workforce	----- Workforce location -----	Distributed workforce
None on the automation front	----- Existing automation partnerships -----	Existing relationship with automation vendor
Few historical documents and extracted values are available	----- Availability of training data -----	All historical documents and extracted values are available

Step 4: Identify all determinants and map paths (page 4 of 5)

While all enterprises will likely follow a set of steps in the execution path ...



Step 4: Identify all determinants and map paths (page 5 of 5)

... the nature of those steps will vary based on environmental determinants¹

ILLUSTRATIVE

Steps	Determinants	Path options
1 Identify and prioritize processes using the prioritization framework	<ul style="list-style-type: none"> Risk appetite Current outcome and capability 	<ul style="list-style-type: none"> Implement one process at a time Implement logical groups of processes sequentially Big bang implementation
2 Plan implementation timelines, development for IDP and reskill		
3a Obtain alignment with IT for IDP	Steps 6a Initiate continuous communication as part of change management	Determinants <ul style="list-style-type: none"> People-centricity Initiating stakeholders
3b Obtain team buy-in, particularly	Path options <ul style="list-style-type: none"> Low to no communication Medium frequency of communication at BU level Frequent communication driven by IT/central team 	
4 Select appropriate provider tool required to achieve the desired	6b Initiate training of existing talent for IDP operations 6c Initiate governance mechanism	NA Risk appetite
5 Obtain management buy-in and	6d Initiate reskilling for displaced employees	People-centricity
	7 Develop pilot for the prioritized process	NA
	8 Obtain required data sets to train the tool	Availability of data
	9 Cut to production with human s achieves the desired efficiency	Steps 12 Embed necessary skills in the automation CoE
	10 Continuously monitor and repo	Determinants <ul style="list-style-type: none"> Organization structure Initiating stakeholders
	11 Repeat journey with the next p	Path options <ul style="list-style-type: none"> Centralized talent pool for IDP managed by automation CoE Decentralized talent pool for IDP with high degree of collaboration with automation CoE Siloed talent pool for IDP collaborating with automation CoE on ad hoc basis
	13a Scale up and run operations	NA
	13b Continuously monitor and report on metrics/KPIs	NA
	14a Set up a team to evaluate opportunities	Organization structure
	14b Templatize opportunity evalu	<ul style="list-style-type: none"> Centrally nominated and controlled Centrally controlled with nominations from business units
	15 Enable various exposure mechanisms to create awareness – newsletters, online web portals, etc.	NA
	16 Institutionalize the governance model	NA
	17 Continuously monitor and report on metrics/KPIs	NA

¹ Refer to pages 86-89 for the overview of the environmental determinants and the various path options for the enterprise at each step

Step 5: Execute against the mapped path (page 1 of 2)

Enterprise IDP journey challenges



Compliance and data security risks

Enterprises are reluctant to provide third-party providers access to user activity data due to growing data security and privacy concerns. Getting approvals from enterprise IT to access data could be a time-consuming process.



Internal resistance

Resistance coming from internal teams in acceptance and adoption of AI and related digital transformation initiatives is a key challenge. And could lead to potential delays and difficulty in achieving the desired RoI.



Improper metrics

Another challenge is the lack of process SMEs' involvement and the inability to identify the right set of metrics/KPIs to track the performance of IDP solutions.



Lack of technology awareness

Organizations could face the lack of business stakeholder buy-in due to limited understanding of document processing technology, its applications, and benefits. They could also be skeptical regarding the credibility of new technology.



Expectation mismatch

Successful implementation of IDP solutions depends on the complexity of use cases. Enterprises, especially business users, sometimes expect unrealistic RoI and time to implementation from IDP solutions due to lack of understanding of ML-based solutions and the hype in the market.

Step 5: Execute against the mapped path (page 2 of 2)

Winning insights for a successful journey



Increase technology awareness

Spread awareness about intelligent document processing and its benefits, address any concerns related to increased transparency, and disseminate success stories for robust change management.



Secure buy-in of business and IT teams

Obtain support from the senior management to ensure smoother implementation and from enterprise IT teams to ensure data security and compliance.



Train/upskill your employees

Collaborate with intelligent document processing providers or their training partners and leverage in-house experts/CoEs to train employees on using the IDP platform.



Collaborate with business units and CoEs

Work closely with other business units, process excellence, and automation CoE to align IDP initiatives with broader transformation initiatives.



Start with a simple project

Start with a process that is limited to a few teams and has a restricted number of applications, defining, and measuring KPIs and monitoring impact. Secure quick wins to establish credibility and accelerate adoption.



Define robust security measures

Conduct regular audits, ensure encryption measures, and maintain access control list. Continuously monitor security threats in real-time and have disaster recovery and business continuity plans.

05

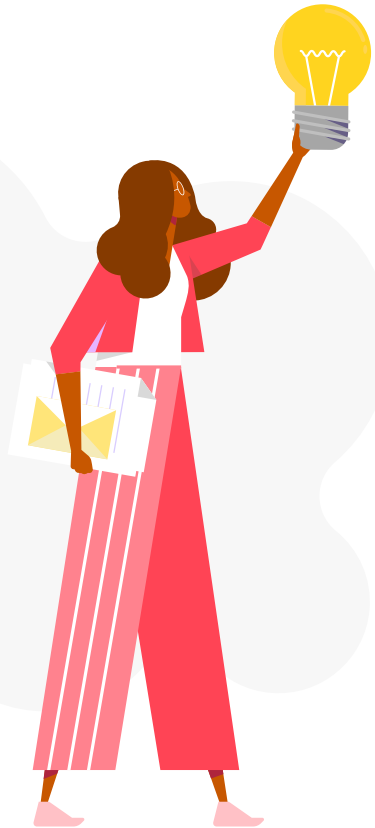
Selecting the best-fit enterprise-grade solution

- Core product capabilities
- Solution approach
- Product training and support
- Technology partner ecosystem
- Service partner ecosystem
- Hosting options
- Commercial model

Selecting the best-fit enterprise-grade IDP solution

To choose the right IDP solution, enterprises need to consider certain key factors including product capabilities, partnership ecosystem, and customer support

Key factors to consider when selecting an enterprise-grade IDP solution



Product capabilities

IDP solution approach

Product training and support

Technology partner ecosystem

Service partner ecosystem

Hosting options

Commercial models



Selecting the best-fit enterprise-grade IDP solution | product capabilities (page 1 of 4)

Overall product capabilities can be broken down into core product capabilities and unstructured document processing capabilities

IDP product capabilities



Core IDP product capabilities
 Capabilities specific to unstructured document processing







Document ingestion	Data extraction	Self-serve capabilities	Summarization
Pre-processing	Post-processing	Security	Intelligent search
Software learning	Interoperability	Monitoring & analytics	Document comparison
Document classification	Workflow capabilities	Entity recognition	Sentiment analysis

Selecting the best-fit enterprise-grade IDP solution | product capabilities (page 2 of 4)

Core product capabilities

Core product capabilities








Capability	Description
 Document ingestion	The common document ingestion modes supported by the platforms include direct upload, bulk upload through File Transfer Protocol (FTP), direct integration or via APIs from document management systems, reading from cloud storage, and emails. Typically, IDP solutions can process documents uploaded in image (JPG and PNG), PDF, DOC, XLS, and TXT formats.
 Pre-processing	This feature enables enterprises to enhance the quality of the ingested documents to increase document processing accuracy. Some of the common pre-processing techniques include auto-cropping, de-skewing, background whitening, and watermark removal.
 Software learning	Software learning refers to the capability of the solution to learn from new data and retrain ML models based on feedback from human verification. Sophisticated technologies such as n-shot learning and transfer learning algorithms help in training the model using fewer sample documents, thereby reducing the liability on enterprises to collect large training datasets.
 Document classification	ML and deep learning algorithms help in document classification. ML algorithms provide the ability to classify unstructured documents. They also help in splitting/classifying pages within a document.
 Data extraction	NLP, ML, and deep learning algorithms are leveraged to extract relevant data from documents. The IDP solution also provides multi-language support for data extraction.
 Post-processing capabilities	This feature enables the IDP solution to integrate with commonly used applications such as SAP, Oracle, Salesforce, and Microsoft using pre-built connectors or API-based integrations. They also allow enterprise users to view/modify extracted data directly from within downstream applications such as SAP, Salesforce, or Excel.

Selecting the best-fit enterprise-grade IDP solution | product capabilities (page 3 of 4)

Core product capabilities

Core product capabilities








Capability	Description
 Interoperability	<p>This feature enables the IDP solution to integrate with commonly used applications such as SAP, Oracle, Salesforce, and Microsoft using pre-built connectors or API-based integrations. They also allow enterprise users to view/modify extracted data directly from within downstream applications such as SAP, Salesforce, or Excel.</p>
 Workflow capabilities	<p>This feature provides an intuitive user interface with drag-and-drop features enabling enterprise users to create and edit workflows for approval within the platform.</p>
 Self-serve capabilities	<p>IDP solutions offer provision for users to add new document types for processing. Users can also add/modify fields to be extracted. The fields can be defined and configured directly through the GUI. A few providers also offer the ability to configure threshold confidence levels for data extraction, thereby helping optimize STP rates.</p>
 Security & risk management	<p>The IDP solution should offer security and risk management features such as:</p> <ul style="list-style-type: none"> • Role-based Access Control (RBAC) • Adherence to security standards such as GDPR • PII redaction/removal • Data encryption during transit and at-rest
 Monitoring & analytics	<p>The monitoring and analytics dashboard provides the ability to monitor, measure, benchmark, report, and generate custom reports for business KPIs to better suit the enterprise requirements. Some metrics that are generally tracked are:</p> <ul style="list-style-type: none"> • Speed of processing • Field-level accuracy • Batch-level STP rates • Number of errors, time taken to fix errors, etc. • AI governance metrics for model performance (e.g., F1 score)

Selecting the best-fit enterprise-grade IDP solution | product capabilities (page 4 of 4)

Unstructured document processing capabilities

Unstructured document processing capabilities



Capability	Description
 AI capabilities for entity recognition	This refers to the capability of the IDP solution to identify a group of words from an unstructured document and classify them into pre-defined categories (such as name and location) to improve the accuracy of information extraction.
 Intelligent search	This feature enables enterprise users to search through a free-flowing text document based on the contextual meaning of the query rather than the literal meaning of the key words.
 Document comparison	Document comparison is leveraged to analyze the difference in meaning between two or more documents (such as comparison of clauses in a standard contract vs. individual contract based on its meaning).
 Sentiment analysis	This refers to the capability of the solution to identify the sentiments (positive, negative, and neutral), emotions (angry, happy, and sad), and urgency in textual data and draw insights from customer reviews and surveys.
 Document summarization	This feature refers to the ability of the solution to extract relevant information from one or more text documents and create a summary of the document(s) by leveraging Natural Language Generation (NLG) models.

Selecting the best-fit enterprise-grade solution | solution approach

The IDP software market today primarily comprises two types of solutions: pre-built and platform-based approach

Types of solutions in IDP software market

Pre-built approach



Refers to the IDP software solutions offered as a package, where the providers or the implementation partners will predominantly oversee the customization, configuration, and deployment of the solution

- **Pros:**
 - Dedicated resources and solutioning experts can assist enterprises in deploying the product across complex use cases
 - Consistency in terms of accuracy and performance
- **Cons:** inability to build new use cases or tinker with the existing ones



Platform approach

Denotes the IDP software solutions offered as platforms (can be part of a larger intelligent automation platform), which AI-savvy enterprise users can use to build and deploy specific use cases themselves, or with support from the technology provider or implementation partners

- **Pros:**
 - Enterprises can build use cases on their own
 - Allows for experimentation with choosing the best-fit models and a greater degree of control
- **Cons:** Given the current state of maturity and scarcity of skilled resources, most enterprises end up having to use external support for new use cases, thereby diluting the promise of a platform solution

Currently, both models are viable in the market. Given the scarcity of skilled resources today, most platform solutions end up being made available as pre-built solutions, except in the case of mature enterprises with dedicated data science talent.

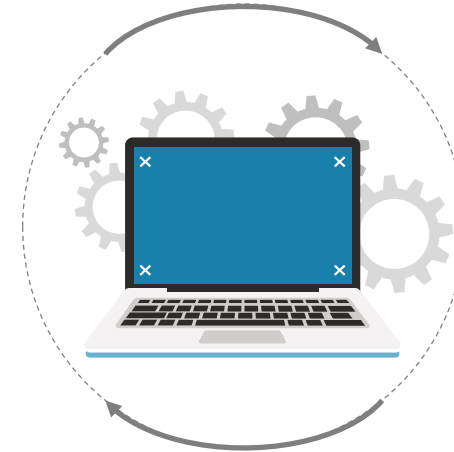
Selecting the best-fit enterprise-grade solution | product training and support

Two main tenets of product training and support



Robust product support and maintenance services

- Product support and maintenance are very important for a smooth journey experience and, hence, enterprises should look for an IDP provider that offers robust and continuous product and maintenance support
- Uniform product update/release cycles and maintenance services help enterprises avoid/minimize challenges with their deployments, especially when IDP is being used for multiple processes. An enterprise's IT and operations department should not be burdened with issues resulting from inadequate product support, but rather be assisted to ensure the most efficient use of limited resources



Easy access to comprehensive product training

- Since the market is in an early stage, most providers provide in-person or classroom training to enterprises, but some providers have started to offer online training courses
- Self-paced online training courses, with training documentation that can be downloaded and viewed offline, allow enterprises to accelerate the learning curve for human agents and accentuate their HITL capability, ensuring efficient handling of complex documents
- Many providers are also developing online community ecosystems where users can interact with each other for brainstorming and troubleshooting. Platforms may also come with embedded help tools to help learn on the job

Selecting the best-fit enterprise-grade solution | technology partner ecosystem

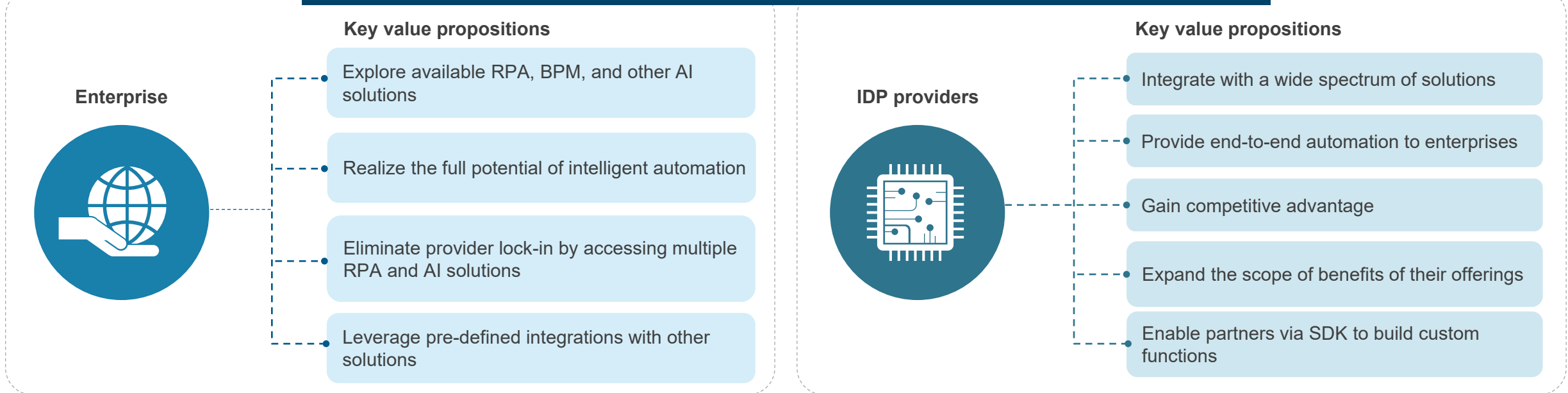
Ecosystem of partners for collaborative technologies



Why technology partner ecosystem is an important consideration for enterprises while choosing an IDP provider

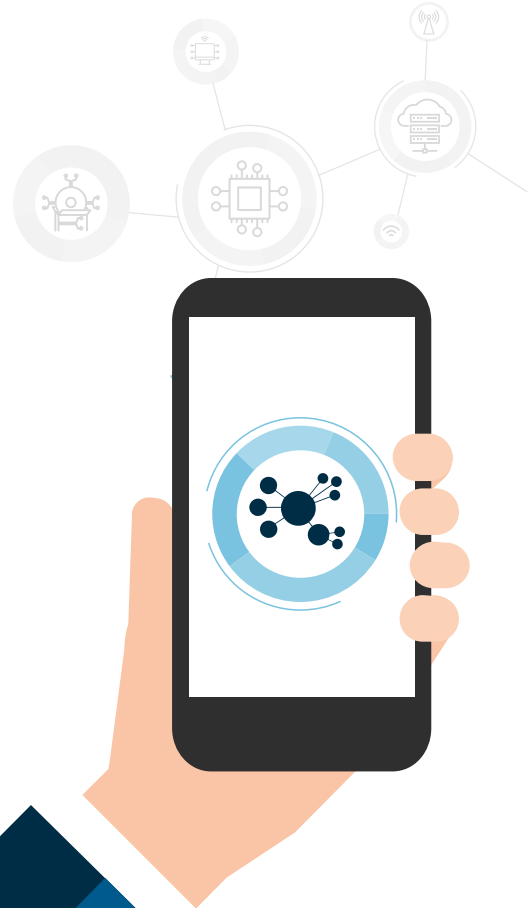
- IDP solutions are capable of extracting data from complex documents, but they can be combined with complementary technologies such as RPA, BPM, and other AI technologies to perform end-to-end automation
- Solutions packaged with IDP and RPA help enterprises improve operational efficiency and increase cost savings
- If an enterprise's chosen IDP provider can provide access to a broad partner ecosystem for collaborative technologies, it would help the enterprise to expand its automation capabilities in an expedient fashion. It would reduce the hassle to reconfigure the deployment and integrate with collaborative technologies
- Some IDP providers partner with providers of best-in-class AI technologies such as NLP and OCR to provide flexibility to enterprises

What a robust partner ecosystem of collaborative technologies means for enterprises and IDP providers



Selecting the best-fit enterprise-grade solution | service partner ecosystem

Ecosystem of services partners



Why training partners are important for enterprises while choosing an IDP provider?

- Training partners provide their implementation and training expertise to help enterprises configure and deploy IDP solutions
- Since IDP solutions are evolving rapidly, training partners are required to train employees on new features/capabilities of the solution
- For enterprises to achieve global scale of deployments, they need training to be available in a variety of languages. This is one of the key value propositions that they bring to the table

Why implementation partners are important for enterprises while choosing an IDP provider?

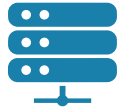
- Implementation partners include system integrators that help enterprises in implementing IDP solutions and overcoming challenges in deployment
- They can also be leveraged to overcome challenges in areas such as governance, business case realization, and scaling up
- Enterprises can also leverage them to set up automation CoEs

Why BPO partners are important for enterprises while choosing an IDP provider?

- Many IDP providers offer the option of a managed-services construct through partnerships with leading BPO providers
- Enterprises can choose to outsource HITL and validation activities to a BPO partner
- This can significantly reduce the time taken for implementation and eliminate the need to train employees

Selecting the best-fit enterprise-grade solution | hosting options

IDP solutions can be hosted either on-premise or cloud, based on enterprise data security requirements



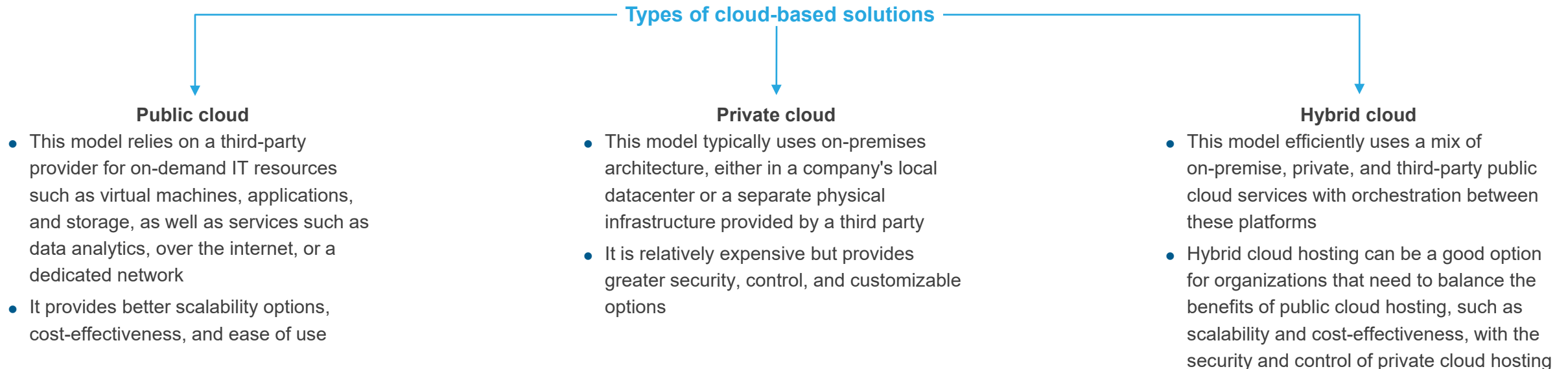
On-premise solution

On-premise hosting, also known as in-house or self-hosting, refers to the practice of hosting applications, data, and infrastructure on local servers that are physically located on a company's premises. It provides control over data and infrastructure, which can be beneficial in terms of security and compliance.



Cloud-based solution




In cloud-based solutions, enterprises are not limited to the capabilities of their existing hardware. Providers offer virtual servers and other hosting resources that can be configured to meet the specific needs of the IDP application. Solutions are scalable, maintenance free, and relatively fast to implement.




Selecting the best-fit enterprise-grade solution | commercial models

One-time fee
The IDP provider or the System Integrator (SI) charges a one-time implementation fee that covers integrating the solution with the client's existing system, training the software, and providing implementation support


Subscription-based pricing models

-  **Volume-based licensing**
Enterprises pay on a monthly or annual basis, depending on the volume and complexity of the documents. This may be on a per-page or per-document basis.
-  **Per-process-based model**
Enterprises pay on a per-process or per-case basis, irrespective of the number of pages processed in the use case. The price of the use case depends on various factors such as complexity.
-  **Fixed-fee model**
Each installation is licensed monthly/annually to operate on a single machine, irrespective of the number of documents/processes handled by the IDP product. This makes it cost-effective for some enterprises.

Progressive pricing model

-  **Outcome-based pricing**
Mutually agreed pricing based on the quality of output or outcomes (e.g., minimum STP rate or accuracy rate), typically observed when IDP solutions are included as part of the broader BPS/IT contracts.

Perpetual pricing model

-  **One-time fee model**
This requires enterprises to pay a one-time upfront fee to purchase a perpetual license and an annual maintenance fee. While it might result in higher savings in the long run, an increasing need for flexibility to scale has driven down the adoption of this model.

While assessing an IDP provider, enterprises should consider its investment and product roadmap to ensure a smooth journey in the future

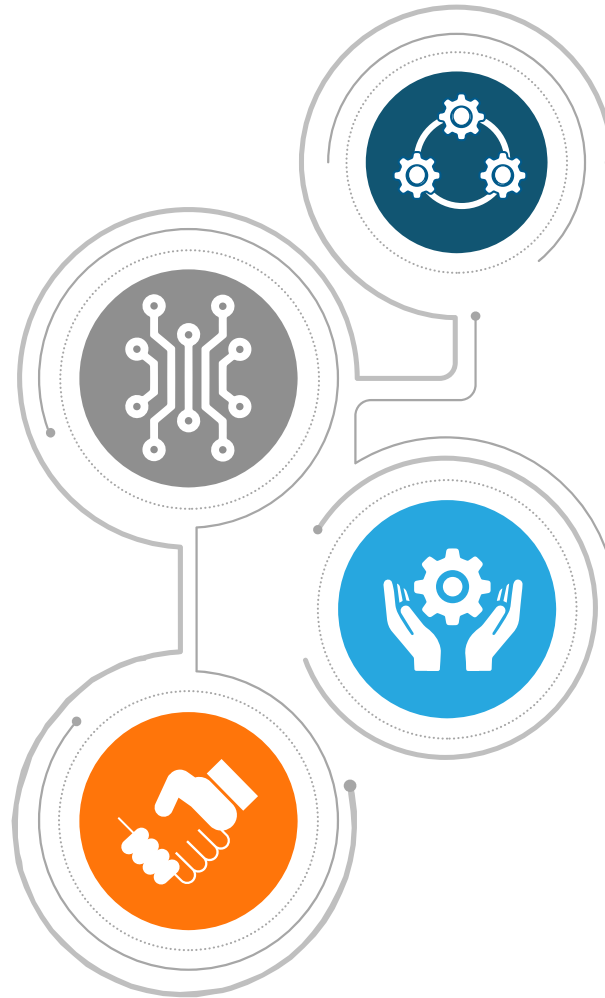
Provider's investments and product roadmap

Technology

Enterprises should consider the IDP provider's future investments in the solution/technology such as NLP, LLM, and generative AI and assess whether it aligns with their objectives

Technology partnerships

IDP providers that have technology partnerships with best-in-class intelligent automation solution providers help enterprises in their automation journey



Processes and use cases

Knowing the processes, document types, and data types for which data can be extracted in the future, is an important factor while selecting an IDP partner. Enterprises should also consider the roadmap for pre-built solutions, additional languages supported, etc.

Training & support

Future roadmap of training & support for new use cases, product upgrades, etc., are essential to continue a smooth journey in the future

06

Challenges and best practices for IDP adoption

- Key barriers and best practices to accelerate value realization from IDP adoption

Key barriers and best practices to accelerate value realization from IDP adoption (page 1 of 2)

Challenges

Challenges such as noise, low-resolution scans, poor image quality, and inconsistent document structures can impact data extraction reliability.

Large amount of high-quality training data can be challenging to obtain with limited resources. Also, data labeling and model selection requires significant resources and expertise with deep understanding of data and the problem being solved.

Processing large volumes of documents efficiently and in a timely manner is a significant technical challenge.



Data quality noise



Availability of training data



Scale and volume of processing

Best practices

Pre-processing techniques such as noise removal, image enhancement, and de-skewing are needed to improve document quality and reduce noise interference.

- To solve the challenge for unavailability of training data, a few IDP providers support transfer learning and n-shot learning that enables model training with fewer datasets. Some providers also leverage generative AI to generate synthetic datasets to create representative samples for training
- Proper labeling and annotation, along with a balanced dataset, help prevent bias. Regular updates and continuous monitoring of performance enables adaptation to evolving patterns
- A few IDP providers have integrations with LLM models to automate the data labeling process and reduce annotation efforts
- Based on volume requirements, enterprises should opt for IDP solutions capable of handling requisite document volumes without compromising performance while optimizing processing speed, resource allocation, and parallel processing techniques
- Prioritizing critical documents, implementing intelligent queuing, and continuously monitoring system performance for optimal efficiency is crucial

Key barriers and best practices to accelerate value realization from IDP adoption (page 2 of 2)

Challenges

There are trade-offs between precision and automation, and it is important to manage false positives and negatives while optimizing the system to meet specific business requirements and minimize errors and manual intervention.



Balancing between accuracy and STP rates

Challenges arise from declining accuracy and outdated results of an IDP solution due to variations in real-world documents.



Model drift

The process of data mapping and synchronization with Enterprise Content Management (ECM) platforms, document repositories, or workflow management systems can be a complex technical challenge.



Integration with existing system

Handling sensitive information in document processing makes security and compliance a critical consideration.



Security and compliance

Best practices

- IDP solutions should ensure high-quality training data, and regular model retraining, along with threshold optimization, error analysis, and feedback loop
- Combining automation with human review for complex cases by implementing HITL mechanisms to validate results helps improve accuracy and the overall output quality

Underlying AI/ML models should continuously learn and adapt to evolving document variations and patterns. This is typically achieved through HITL feedback.

Enterprises must ensure IDP solution compatibility with various APIs to ensure smooth data flow and efficient collaboration between systems. Availability of pre-built connectors further ensures seamless integration.

Ensuring data privacy through auto-redaction, implementing access controls, and compliance with industry regulations such as GDPR or HIPAA will help enterprises overcome data security and compliance concerns.

07

Future outlook

Outlook for 2023-24 (page 1 of 2)

Continuing growth in demand

- Despite recessionary headwinds, the IDP technology provider market is expected to continue its growth momentum owing to enterprises' continued focus on cost savings and productivity gains
- The ability of IDP solutions to handle multiple types of higher complexity use cases such as contracts, financial statements, research documents, blueprints, and engineering diagrams has also paved the way for broader adoption

Go-to-market strategy

- IDP providers are expected to offer more pre-trained ML models and out-of-the-box verticalized solutions to serve industry-specific complexities in document processing, enabling quicker deployment
- Listing of IDP solutions on hyperscaler marketplaces such as AWS and Microsoft Azure is expected to pick up pace for wider reach
- IDP providers are anticipated to form more strategic partnerships with service providers and system integrators for various services, including reselling, training, consulting, and implementation
- Industry-specific accreditations, certifications, and partnerships will continue to gain traction among IDP providers, going forward
- The recessionary tendencies are expected to usher in more progressive commercial models including outcome-based pricing options that are linked to productivity gains
- The market may also see demand for money-back guarantee programs in case the pre-determined business outcomes are not achieved in the stipulated time frame by IDP providers

Cloud adoption

Cloud deployment is expected to grow with availability of various hosting options such as virtual private clouds, hybrid deployment, and multi-cloud deployment. Further, there is an increasing trend toward adopting microservices architecture and fully managed SaaS offerings.



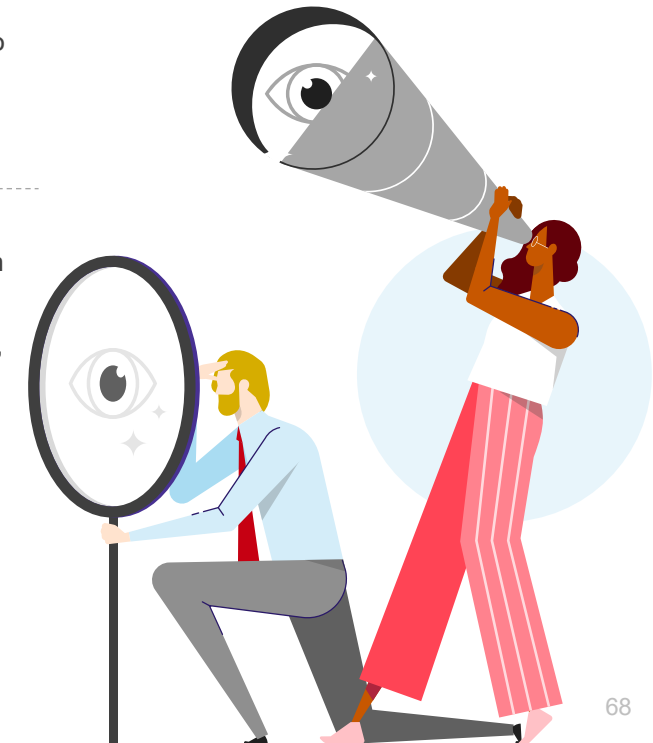
Outlook for 2023-24 (page 2 of 2)

Integrations with LLMs and generative AI

- Integration of IDP with generative AI tools and LLMs such as GPT will continue to gain traction as organizations look for more efficient ways to manage document processing workflows. This is also expected to make the model training process more efficient and user-friendly by leveraging prompt engineering techniques
- LLMs are trained over huge data sets and excel at interpreting natural languages. This would help IDP solutions in processing unstructured documents having free-flowing texts, both handwritten and printed, with higher efficiency and faster turnaround time
- With advanced context understanding capabilities, LLMs enable IDP solutions to perform a higher degree of intelligent search within document repositories and generate summaries of text documents. This would be valuable in helpdesk scenarios and knowledge management systems, which requires conversing with documents
- These integrations are expected to make IDP solutions more productive, not only in terms of classification, interpretation, and extraction of data, but also generating new documents. Potential use cases include creating new contracts based on clauses of pre-existing agreements and questionnaires from documents including an organization's training modules and academic materials. It can have wide applications across verticals including EdTech, journalism, and research

Other technology trends

- Advanced software learning techniques including transfer learning, n-shot learning, and zero-shot learning are expected to see rapid adoption along with synthetic data generation algorithms (e.g., GAN) to train models with smaller number of sample datasets and documents
- Configurability of the IDP platform for improving self-serviceability through features such as creating workflows using drag-and-drop GUI, model training, handling exceptions, and adding/editing business rules for classification, extraction, and validation would further increase
- Ability of the IDP solution to integrate with an enterprise's own ML models through bring-your-own-model feature is expected to gain prominence
- Advanced context understanding capabilities such as generating high-quality summaries, enabling documents comparison, and facilitating intelligent search would either be made available natively as a part of the IDP solution or through partnerships with niche providers
- Demand for the extraction capability of data types such as audio and video and support for processing documents in regional language scripts such as Mandarin, Korean, Japanese, and Arabic is expected to rise

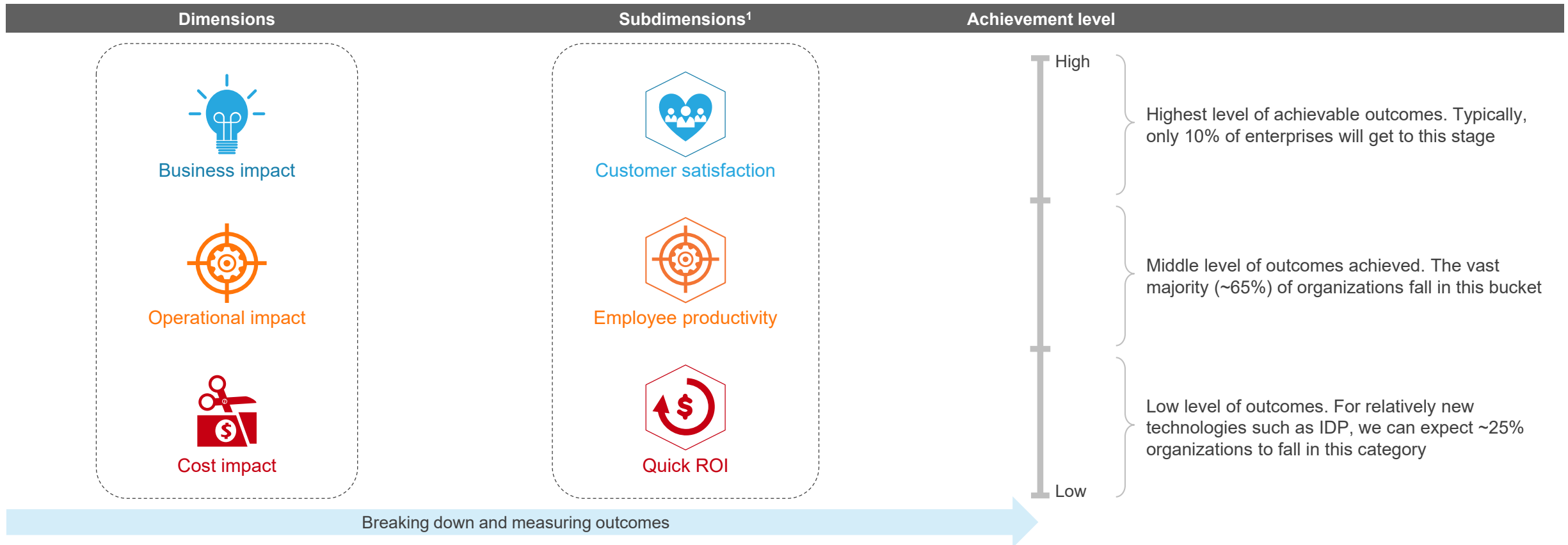


08

Appendix

- Enterprise IDP Capability Maturity Model (CMM)
- Environmental determinants
- Variance in execution path steps based on environmental determinants
- Glossary of terms

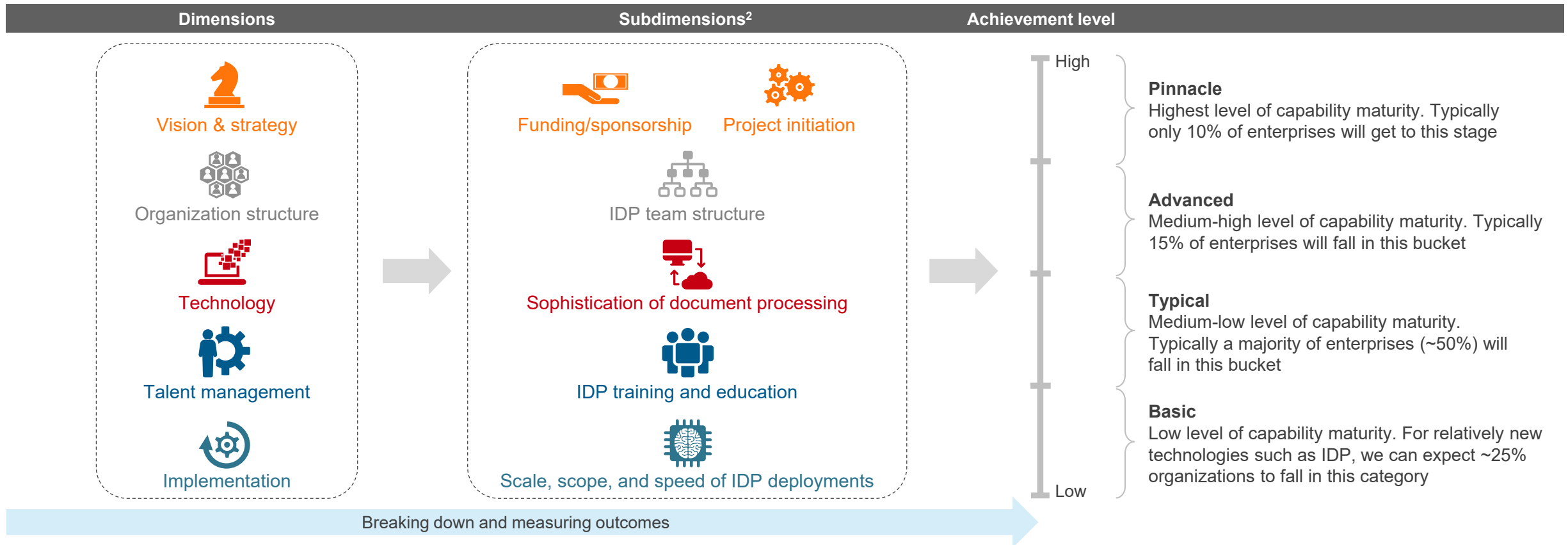
Outcomes: Use the Pinnacle Enterprise® outcomes model to understand your current state and goals for the desired target state



- Overall, we can measure the outcome through a combination of three factors: cost impact, operational impact, and business impact
- We can further break these down into subdimensions that fall into one of the three buckets depending on the level of the outcome achieved. The exact measure of outcomes will vary significantly by the scope of IDP implementation

1 Not exhaustive

Capability: The Pinnacle Enterprises® Capability Maturity Model (CMM)¹ can help enterprises understand their current state of capabilities and subsequently where they want to get to



- Overall, capability is measured through a combination of five factors – vision & strategy, implementation, organization & talent, technology, and resourcing
- Each of these is further broken down into subdimensions, which can be measured as falling in one of the four buckets depending on the maturity level: basic, typical, advanced and pinnacle

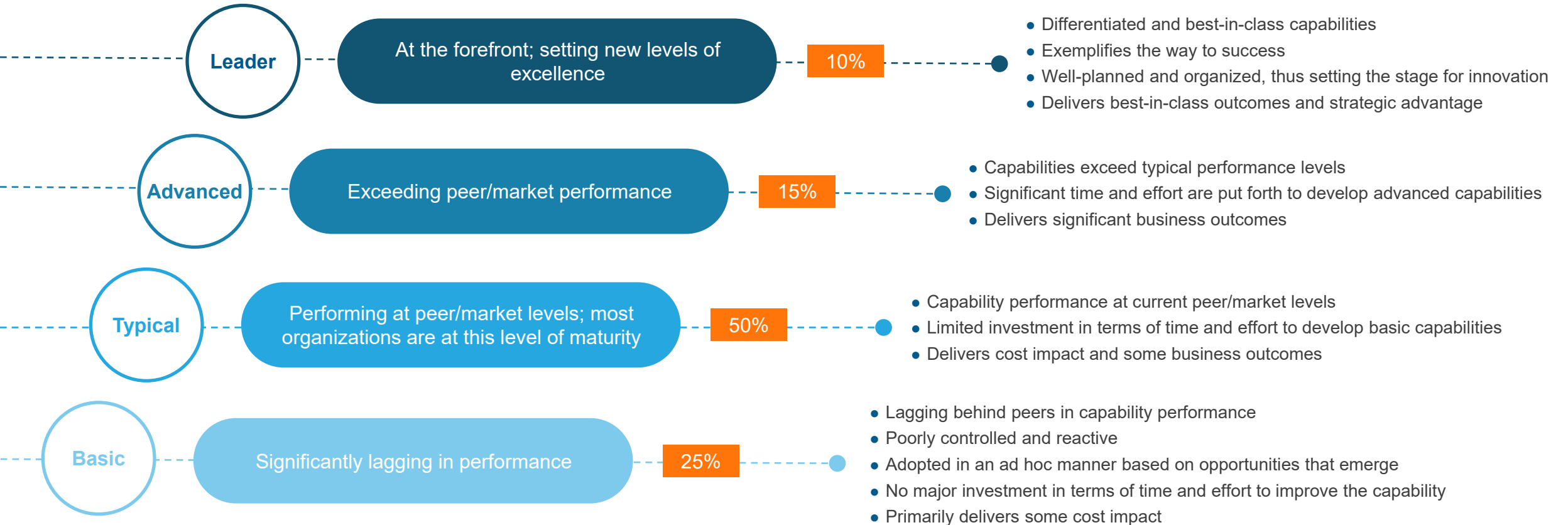
¹ Refer to pages 107-124 for the detailed model, dimensions, and subdimensions

² Not exhaustive

Enterprises' IDP capabilities are assessed across four maturity levels

The four capability maturity levels

XX% Approximate share of enterprises at each maturity level



Everest Group evaluates IDP capabilities according to five key components of enterprises' IDP journey

Journey components

Key focus area



Vision & strategy

- To understand the vision of the organization for IDP and the drivers behind its adoption
- To assess the organization's readiness for IDP adoption from a process and security perspective



Organization structure

- To assess the governance model for IDP initiatives and the extent of collaboration among the implementing groups
- To analyze the talent management strategy for the organizational change caused by IDP adoption



Technology

- To assess the extent to which various components of IDP technologies such as OCR, software learning, computer vision, and analytics are being utilized
- To assess the level of sophistication of various IDP technologies deployed



Talent management






To assess the sourcing strategy, training, and education programs for various IDP skills



Implementation

To assess the scale and scope of IDP adoption along with the pace at which IDP has been adopted – in terms of number and types of documents processed

Enterprise IDP capability is assessed across over 25 capability elements

Journey components	Capability	
 A. Vision & strategy (7 capabilities)	A1. Primary drivers of IDP adoption A2. Funding/sponsorship A3. Project initiation A4. Security & risk preparedness for IDP	A5. Metrics and KPIs for measuring benefits/impact of IDP A6. Metrics and KPIs for measuring effectiveness of IDP initiatives A7. Targeted document types for IDP adoption
 B. Organization structure (8 capabilities)	B1. IDP team structure B2. Scope of automation CoE B3. Roles and responsibilities of CoE B4. Primary use of performance data	B5. Focus on tracking/optimizing the benefits achieved B6. Level of employee engagement B7. Nature of impact on employees B8. Reusability of models
 C. Technology (7 capabilities)	C1. Software learning C2. Classification of documents C3. Flexibility with ML algorithms C4. Sophistication of document processing	C5. Pre-built use cases C6. Hosting options C7. Ancillary capabilities
 D. Talent management (2 capabilities)	D1. Sourcing of IDP talent/skills D2. IDP training and education	
 E. Implementation – scale, scope, and speed (4 capabilities)	E1. Distribution of IDP projects by stage E2. Scale of IDP adoption E3. Scope of IDP deployments across functions E4. Speed of IDP adoption	

Enterprise IDP CMM (page 1 of 10)

Vision & strategy



Capability elements	Basic	Typical	Advanced	Pinnacle
Primary drivers of IDP adoption	Business case focused on generating quick cost savings through better efficiency and quality of output	Business case focused on improving process efficiency, increasing workforce productivity, as well as generating cost savings	Business case focused on improving employee experience, customer experience, governance & compliance along with increasing productivity, efficiency, and business resilience, as well as generating cost savings	Business case focused on bottom- line (cost savings) and also on top-line (new sources of revenue generation), along with improving governance & compliance, employee and customer experience, efficiency, and business resilience
Funding/sponsorship	Primarily sponsored/funded by local/regional business unit budget	Primarily funded by the global shared services budget	Primarily funded by global business function's budget / CoE	Primarily funded by the central enterprise budget; sponsorship from the CXO
Project initiation	Projects are initiated by local/regional business units in a siloed way with the IT team's support	Projects are initiated by local/regional business units with basic/limited support from central automation CoE	Projects are initiated by global business functions OR global shared services; multi-pronged approach with substantial support from the automation CoE at the central level	Projects are initiated by corporate OR global business functions OR global shared services; multi-pronged approach with robust support from both central and business unit level automation CoE

Enterprise IDP CMM (page 2 of 10)

Vision & strategy



Capability elements	Basic	Typical	Advanced	Pinnacle
Security and risk preparedness for IDP	No major changes made to security and risk policies and worked around existing ones to accommodate changes required for IDP	Some changes to security and risk policies made to accommodate IDP environments and scenarios	Proactively evaluated and planned for mitigation of security and compliance risks associated with IDP and associated AI deployments; set up unique risk management protocols and controls for IDP and AI deployments	Included security and risk leaders in IDP projects to proactively evaluate and plan for mitigation of security and compliance risks and unique requirements essential for IDP and associated AI deployments; set up unique risk management protocols and controls for IDP and associated AI deployments
Targeted document types for IDP adoption	Template-based documents (data in pre-defined template) and documents with semi-structured data with no variations	Template-based high-volume documents and documents with semi-structured data with significant variations	Template-based high-volume documents, semi-structured documents with significant variations, and unstructured documents (such as emails)	Template-based high-volume documents, semi-structured data with significant variations, and unstructured content (including engineering diagrams, complex tables, audio, and video)

Enterprise IDP CMM (page 3 of 10)

Vision & strategy



Capability elements	Basic	Typical	Advanced	Pinnacle
Metrics and KPIs for measuring benefits/impact of IDP (such as cost savings, RoI, speed, productivity, accuracy, compliance, and employee experience)	The organization currently does not use any well-defined metrics to measure returns from IDP investments; metrics used are ad hoc, poorly controlled, and reactive/chaotic	The organization has defined some basic cost and efficiency metrics along with existing IT metrics, which are repeatable in projects to measure returns from IDP investments	The organization has defined new metrics (employee experience, productivity, and speed) along with basic cost and efficiency metrics and existing IT metrics, which are repeatable in projects; the metrics are standardized across the organization to track the returns on IDP investments	The organization has defined new metrics (employee experience, speed, productivity, efficiency, cost, etc.) that are standardized across the organization and continuously optimizes the metrics to measure impact of IDP investments in near real-time
Metrics and KPIs for measuring the effectiveness of IDP initiatives (such as accuracy rate, speed of configuration/implementation, STP rate, time taken to process a document, and the number of documents processed)	The organization currently does not use any well-defined metrics to measure effectiveness of IDP initiatives; metrics used are ad hoc, poorly controlled, and reactive/chaotic	The organization uses some basic metrics such as the number of documents processed along with existing IT metrics, which are repeatable in projects, to measure the effectiveness of IDP initiatives	The organization has defined a series of new metrics such as speed of configuration/implementation, STP rate, number of documents processed that are standardized across the organization, and document processing time. These help in tracking and measuring the effectiveness of IDP initiatives as well as define policies, procedures, and practices driven by flexibility to accommodate unique aspects of different business units	The organization has defined new metrics (speed of configuration/ implementation, efficiency of pre-trained algorithms, and AI governance metrics such as F1 scores) that are standardized across the organization; continuously optimizes the defined metrics, policies, procedures, and practices to measure the impact of IDP investments and share best practices across different business units

Enterprise IDP CMM (page 4 of 10)

Organization structure



Capability elements	Basic	Typical	Advanced	Pinnacle
IDP team structure	No dedicated IDP team within the organization; largely handled by existing imaging / data capture	Decentralized structure; each business unit forms a dedicated team for IDP initiatives leveraging both existing imaging / data capture team and additional skill sets	Centralized dedicated IDP team that defines and implements IDP initiatives for the entire organization	IDP team embedded in automation CoE that can either be centralized or decentralized (hub & spoke model) to cross-leverage skill sets to implement IDP initiatives
Scope of automation CoE	Less than 30% of IDP projects are governed by the CoE	Around 30-60% of IDP projects are governed by the CoE	Around 60-80% of IDP projects are governed by the CoE	More than 80% of IDP projects are governed by the CoE
Primary use of performance data	Monitoring performance of IDP applications	Monitors performance of IDP applications; refines the model to improve accuracy	Monitors staff productivity along with performance of IDP applications locally to find gaps in existing processes to optimize and streamline them to increase efficiency	Analyzes performance data centrally to identify gaps in existing algorithms and proactively refine the model across business functions / regions to improve accuracy and STP rates
Focus on tracking/optimizing effectiveness and the benefits achieved	Collection and usage of performance and impact data are ad hoc, sporadic, and uncoordinated	Performance and impact data is collected periodically (quarterly) to produce reports and dashboards to gain new insights that improve operational efficiency	Performance and impact data is collected periodically (monthly) to produce reports and dashboards to gain new insights that improve operational efficiency and enhance efficacy of training algorithms	Performance and impact data is collected/monitored weekly and used in a coordinated fashion to make operational decisions for making the training algorithms more robust, enhancing employee experience, and improving customer satisfaction

Enterprise IDP CMM (page 5 of 10)

Organization structure



Capability elements	Basic	Typical	Advanced	Pinnacle
Roles and responsibilities of the CoE	Drive the roll-out and implementation of IDP projects and ensure coordinated communication with relevant stakeholders; loosely defined roles, responsibilities, and skill sets required	Ensuring quality and compliance through well-defined standards, procedures, and guidelines, owned and developed by the CoE for broader digital initiatives; drive the roll-out and implementation of IDP projects and ensure coordinated communication with relevant stakeholders; some key roles and responsibilities are well-defined	Approves all IDP procedures before they are put into deployment, assesses suitability of IDP vs. other document processing tools for use cases, and ensures quality and compliance through well-defined standards, procedures, and guidelines, owned and developed by the CoE for broader digital initiatives; drives the roll-out and implementation of IDP projects and ensures coordinated communication with relevant stakeholders; well-defined roles, responsibilities, and skill sets required	Cross-leverage of automation/AI training and education program to develop talent for IDP initiatives; approves all IDP procedures before they are put into deployment, assesses suitability of IDP vs. other document processing tools for use cases, and ensures quality and compliance through well-defined standards, procedures, and guidelines owned and developed by the CoE. Drives the roll-out and implementation of IDP projects and ensures coordinated communication with relevant stakeholders; well-defined roles, responsibilities, and skill sets required that are regularly reviewed and optimized
Reusability of models	No reusable models	Reusability of models is limited to business units	Reusability of models across business units and geographies	Reusability of models across business units, geographies, and similar document types (through transfer learning)

Enterprise IDP CMM (page 6 of 10)

Organization structure



Capability elements	Basic	Typical	Advanced	Pinnacle
Level of employee engagement	Few people proactively engaging in some of the IDP initiatives	More believers who engage in IDP initiatives	Organization-wide employee engagement; some internal experts to facilitate engagement; developing a culture of innovation and design thinking	IDP initiatives are recognized as an integral component of the broader digital strategy (automation/AI); rewards system for contribution; Integrated culture for design thinking and innovation
Nature of impact on employees	Some attempt to redeploy/reskill/upskill employees released due to IDP initiatives in other areas (such as minimal investment and management commitment)	Modest attempts made to redeploy employees released due to IDP initiatives by providing alternate career paths (for example, education program set up for reskilling)	Significant attempts made to reskill and redeploy employees released due to IDP initiatives by reskilling them to become SMEs for handling HITL functionality	Significant attempts made to reskill/upskill employees released due to IDP initiatives to do higher value work such as HITL and provide alternate career paths in broader automation initiatives (such as citizen developers)

Enterprise IDP CMM (page 7 of 10)

Technology



Capability elements	Basic	Typical	Advanced	Pinnacle
Software learning	Manual generation of training data sets during review	Automatic generation of training batches during manual review; automatic feeding of data sets into the system for training	Automatic generation of training batches during manual review along with feature for enterprise users to approve training sets to improve accuracy; approval mechanism at admin level as well	Automatic generation of training batches during manual review, feature for users to approve training sets for better accuracy; approval mechanism at admin level; continuous learning through HITL functionality; using advanced software learning techniques such as transfer learning and n-shot learning
Classification of documents	Do not have the ability to automatically classify documents	Ability to identify discrete documents with low accuracy, leveraging basic statistical approach	Ability to identify discrete documents with medium accuracy, leveraging basic ML-based approach	Ability to identify discrete documents and different pages within a stream of documents with high accuracy, leveraging advanced neural networks
Flexibility with ML algorithms	Different pre-built ML models for different use cases / document types	Different pre-built ML models for different use cases / document types and ability to choose the model by the enterprise user	Different pre-built ML models for different use cases / document types with a feature recommending the best ML model for the user	Feature suggesting best ML model to choose for different use cases; ability to create new ML models using low-/no-code capability and integrating enterprise's own ML model with the IDP solution

Enterprise IDP CMM (page 8 of 10)

Technology



Capability elements	Basic	Typical	Advanced	Pinnacle
Sophistication of document processing	Basic OCR for digitizing content	OCR- and auto ML-based; document classification, data capture, and extraction using ML and validation	Multi-OCR, auto ML, and NLP; document classification and extraction using multi-modal approach and business rules validation; supporting sentiment analysis and entity-based extraction	Multi-OCR, domain ontology, deep learning, auto ML, NLP, NLG, and LLM; document classification and extraction using multi-modal approach and continuous learning; utilizing intelligent search, summarization, intent analysis, and validation
Complexity of the data handled	Block letters (typed) and simple tables	Block letters (typed) and nested and borderless tables	Block letters (typed or handwritten), tables spanning multiple pages, checkboxes, bar codes, QR codes, and logos	Block letters (typed or handwritten), checkboxes, bar codes, QR codes, logos, stamps, charts, signatures, cursive writing, and audio/video
Pre-built use cases	No pre-built use case	Simple use cases involving semi-structured data such as invoice processing, customer onboarding, and claims	Complex use cases involving unstructured data such as contracts and legal documents	Use cases that involve extracting information from free-flowing text as well as NLG
Hosting type	Physical, desktop-based	On-premise, server-based	Private cloud-based, hybrid	Public cloud-based, hybrid
Ancillary technologies	Stand-alone IDP solution	IDP solution integrated with BPM tool and RPA	IDP solution integrated with BPM, RPA, process mining, analytics, and other AI solutions	IDP solution integrated with BPM, RPA, process mining, conversational AI, analytics, and AI solutions including LLM models

Enterprise IDP CMM (page 9 of 10)

Talent management



Capability elements	Basic	Typical	Advanced	Pinnacle
Sourcing of IDP talent	Leverages only technology/service provider resources	Leverages provider resources and existing data capture / imaging resources with relevant training on IDP	Leverages existing resources with proper training and providers for scaling initiatives; cross-skilling IDP and automation resources	Leverages existing well-trained/experienced resources and broader automation resources by cross-skilling IDP and automation resources, enabling resourcing across automation initiatives as per the required bandwidth, and hiring specialized talent externally
IDP training and education	Basic initial IDP training by providers	Well-structured IDP internal training program in addition to training and certification programs offered by IDP providers; focused on implications of IDP	Formal external and internal, well-structured training programs that are continuously reviewed and optimized, including training for technical skills	Well-structured IDP internal and external training programs that are integrated with broader automation training programs, which are continuously reviewed and optimized

Enterprise IDP CMM (page 10 of 10)

Implementation



Capability elements	Basic	Typical	Advanced	Pinnacle
Distribution of IDP projects by stage	Most IDP projects are in planning and pilot stage	Most IDP projects are being scaled up from pilot stage	Most IDP projects are in steady-state implementation stage	Most IDP projects are in continuous improvement stage
Scale of IDP adoption	Less than 20% of the viable documents leveraging IDP solutions	Around 20-40% of documents leveraging IDP solutions	Around 40-70% of documents leveraging IDP solutions	More than 70% of documents leveraging IDP solutions
Scope of IDP deployments across functions	One document category	Two to four document categories	Five to eight document categories	More than eight document categories
Speed of IDP adoption	One IDP license per year on an average	Two to five IDP licenses per year on an average	Five to 10 IDP licenses per year on an average	More than 10 IDP licenses per year on an average

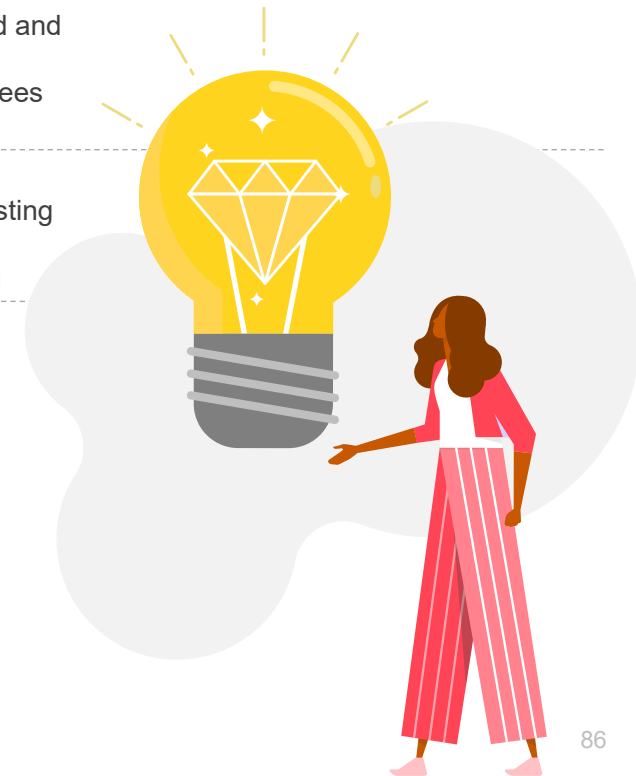
Environmental determinants

 <p>Organization structure</p>	<p>Highly centralized, with some independent decision-making</p>	<p>Partially centralized with portions of decision-making federated to BUs</p>	<p>Largely federated decision-making – BUs have a large degree of freedom to make their own decisions</p>
 <p>People centricity</p>	<p>Highly people-centric organization – the overall culture is people driven rather than efficiency driven</p>	<p>Middle-ground organization with focus on people as assets, with efficiency also playing a role</p>	<p>Efficiency- and/or technology-driven organization</p>
 <p>Initiating stakeholder(s)</p>	<p>Operations-driven initiatives – driven by operations analysts trying to make their jobs easier</p>	<p>IT-driven initiatives with BU support</p>	<p>Centrally-driven, typically by the C-suite or one level below, with all BUs and IT falling in line</p>
 <p>Workforce location</p>	<p>Office-based workforce – FTEs working only from office locations</p>	<p>Hybrid – some FTEs working remotely while some working from office or FTEs working remotely as well as from office as required</p>	<p>Distributed workforce – FTEs working remotely</p>
 <p>Risk appetite</p>	<p>Low risk appetite – need to have multiple layers of checks and balances for any initiative</p>	<p>Medium risk appetite – willing to take risks in select scenarios, especially when dictated by the market</p>	<p>High risk appetite – willing to take risks in the hope of market leadership and payoff</p>
 <p>Existing automation partnerships</p>	<p>No existing automation partnerships</p>	<p>Medium risk appetite – willing to take risks in select scenarios, especially when dictated by the market</p>	<p>Existing partnerships with organizations that also play in the automation space</p>
 <p>Availability of training data</p>	<p>Documents are not maintained properly; few documents and associated values are available and accessible</p>	<p>Documents are managed properly; most of the documents are available, but the associated extracted values are not readily available</p>	<p>Documents are managed properly; most of the documents and associated extracted values are easily available and accessible</p>

Environmental determinants plus current/target outcome and capabilities lead to differences in execution paths (page 1 of 4)

Planning

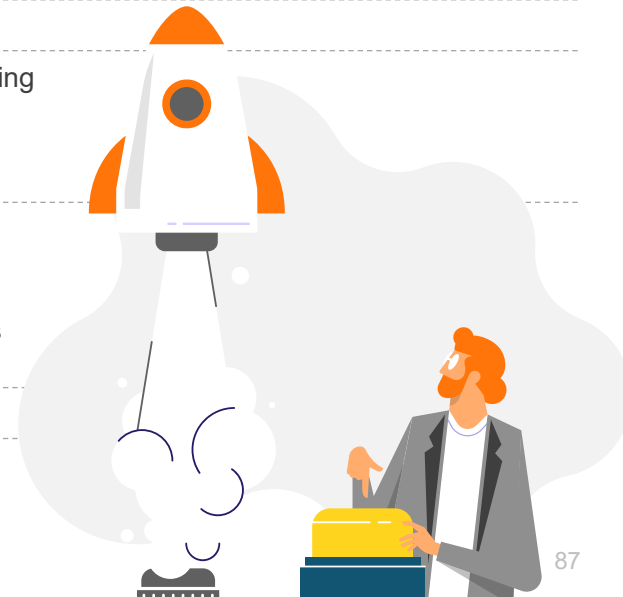
Steps	Determinants	Path options
1 Identify and prioritize processes using the prioritization framework	<ul style="list-style-type: none"> • Risk appetite • Current outcome and capability 	<ul style="list-style-type: none"> • Implement one process at a time • Implement logical groups of processes sequentially • Big bang implementation
2 Plan implementation timelines, governance, and skill development for IDP and reskilling affected employees	N/A	N/A
3a Obtain alignment with IT for IDP implementation	N/A	N/A
3b Obtain team buy-in, particularly impacted FTEs	People centricity	<ul style="list-style-type: none"> • Open communication with the team – affected and unaffected members • Selective communication to impacted employees • Minimal communication
4 Select appropriate provider tool based on capabilities required to achieve the desired outcome	<ul style="list-style-type: none"> • Existing automation partnerships • Risk appetite 	<ul style="list-style-type: none"> • Leverage existing relationships • Evaluate other providers while leveraging existing relationships • Evaluate the entire provider landscape afresh
5 Obtain management buy-in and budget	<ul style="list-style-type: none"> • Organization structure • Initiating stakeholders 	<ul style="list-style-type: none"> • Buy-in and budget at BU level • Buy-in and budget at IT • Buy-in and budget at central level



Environmental determinants plus current/target outcome and capabilities lead to differences in execution paths (page 2 of 4)

Piloting

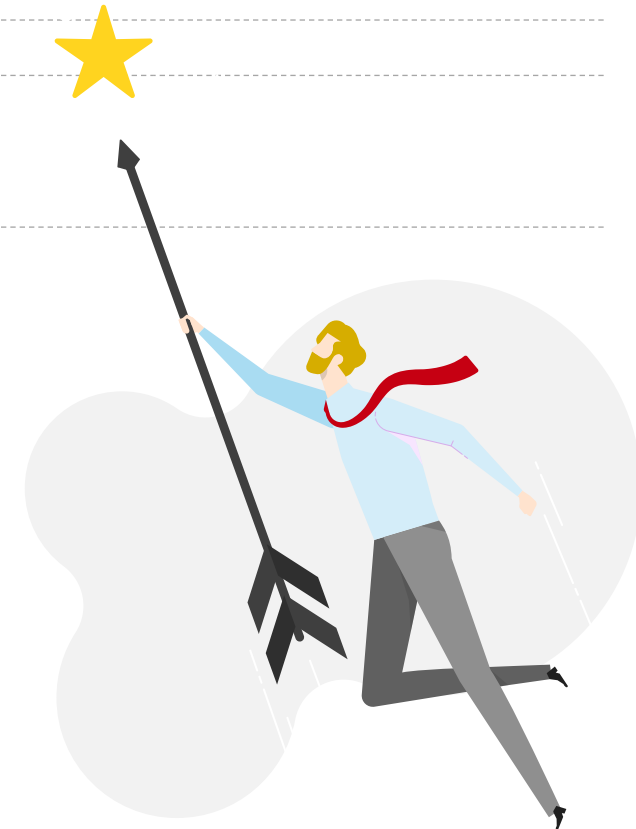
Steps	Determinants	Path options	
6a	Initiate continuous communication as part of change management	<ul style="list-style-type: none"> • People centricity • Initiating stakeholders 	<ul style="list-style-type: none"> • Low to no communication • Medium frequency of communication at BU level • Frequent communication driven by IT/central team
6b	Initiate training of existing talent for IDP operations	N/A	N/A
6c	Initiate governance mechanism	Risk appetite	<ul style="list-style-type: none"> • Minimal, ad hoc governance • Standard set of tracking for metrics • Comprehensive governance, including dashboards for measuring performance, speed, and accuracy
6d	Initiate reskilling for displaced employees	People centricity	<ul style="list-style-type: none"> • No reskilling/upskilling – impacted FTEs may be downsized or reassigned • Upskilling only for high-performing employees, rest reassigned/downsized • Reskilling/upskilling of all employees (all retained)
7	Develop pilot for the prioritized process	N/A	N/A
8	Obtain required data sets to train the tool	Availability of data	<ul style="list-style-type: none"> • Minimal training at production, with model learning during operations • Highly trained model starting with high level of accuracy at production
9	Cut to production with human supervision until IDP achieves the desired efficiency	<ul style="list-style-type: none"> • Availability of data • Risk appetite 	<ul style="list-style-type: none"> • Always employ a HITL • Employ a HITL only for verification of highly sensitive processes • Allow STP where possible, with only exceptions requiring human intervention
10	Continuously monitor and report on metrics/KPIs	N/A	N/A
11	Repeat journey with the next process in the priority list	N/A	N/A



Environmental determinants plus current/target outcome and capabilities lead to differences in execution paths (page 3 of 4)

Scaling up

Steps	Determinants	Path options
12 Embed necessary skills in the automation CoE	<ul style="list-style-type: none"> • Organization structure • Initiating stakeholders 	<ul style="list-style-type: none"> • Centralized talent pool for IDP managed by automation CoE • Decentralized talent pool for IDP with high degree of collaboration with automation CoE • Siloed talent pool for IDP collaborating with automation CoE on ad hoc basis
13a Scale up and run operations	N/A	N/A
13b Continuously monitor and report on metrics/KPIs	N/A	N/A
14a Set up a team to evaluate opportunities	Organization structure	<ul style="list-style-type: none"> • Centrally nominated and controlled • Centrally controlled with nominations from business units • Truly cross-functional, nominally centralized
14a Templatize opportunity evaluation and processing	N/A	N/A



Environmental determinants plus current/target outcome and capabilities lead to differences in execution paths (page 4 of 4)

Steady state

Steps	Determinants	Path options
15 Enable various exposure mechanisms to create awareness – newsletters, online web portals, etc.	N/A	N/A
16 Institutionalize the governance model	N/A	N/A
17 Continuously monitor and report on metrics/KPIs	N/A	N/A



Glossary of key terms used in this report (page 1 of 3)

AI	Artificial Intelligence is the ability of machines to use cognitive computing to mimic human intelligence, such as visual perception, speech recognition, decision-making, and language translation
ABS prospectus	Asset Backed Security prospectus
BI	Business Intelligence refers to technologies, applications, and practices for collection, integration, analysis, and presentation of business information
BPM / process orchestration	Business Process Management (BPM) / Process orchestration solution helps to coordinate tasks and orchestrate the flow of information across disparately designed applications, databases, digital workers, and the human workforce. It includes capabilities of process design, execution (through workflows and orchestration of different BPS technology systems), and monitoring (through analytics)
Buyer	The company/entity that purchases outsourcing services from a provider of such services
Classic process mining	Classic process mining refers to the ability to leverage specialized algorithms to analyze process-related information that is captured in event logs generated by enterprise systems such as ERP, CRM, and SCM, to discover as-is processes, generate process maps, perform conformance check with pre-defined input reference process models, and generate insights for process improvement
Conversational AI	Conversational AI is an ecosystem of advanced technologies such as intelligent virtual agents, agent assist solutions, and analytics & insights that work in an integrated fashion by interacting with each other and sharing information across systems
Task Mining	The ability to capture the user's keyboard, mouse, and potentially other system-level activities performed simultaneously on various desktops to virtually reconstruct the processes and generate a process map capturing the different process variants
Cognitive/smart automation	The ability of a system to learn how to interpret unstructured content such as natural language, and use the analytical capability to derive and present inferences in a pre-defined/structured fashion; for example, a system classifying the mood of a person into one of the pre-defined groups based on their tone and language
Computer vision	A technology that uses AI to enable automatic extraction, analysis, and understanding of useful information from digital images
Deep learning	A subfield of ML concerned with algorithms and inspired by the structure and function of the brain called artificial neural networks
FTE	A way to measure a worker's productivity and/or involvement in a project. An FTE of 1.0 is equivalent to a full-time worker
General AI	A machine that can perform multiple intellectual tasks across a variety of domains; essentially, it mimics all activities performed by a human

Glossary of key terms used in this report (page 2 of 3)

Horizontal business processes	Those processes that are common across the various departments in an organization and are often not directly related to the key revenue-earning business, for example, procurement, finance & accounting, and human resource management
IDP	Intelligent Document Processing is a software product or solution that captures data from documents (e.g., email, text, PDF, and scanned documents), categorizes, and extracts relevant data for further processing using AI technologies such as computer vision, OCR, Natural Language Processing (NLP), and machine/deep learning
ISDA	International Swaps and Derivatives Association
KPI	Key Performance Indicators for processes, services, products, or solutions
ML	It is a type of artificial intelligence that provides computers with learning capabilities without explicit programming
Narrow AI	A machine that performs one narrow task; an expert system
NLP	Natural Language Processing refers to a machine's ability to interpret human languages
NLG	Natural Language Generation The software process to write text in human languages based on structured data
OCR	Optical Character Recognition A technology within computer vision that involves the recognition of printed characters using computer software
PoC	Proof of Concept; a realization of a certain method or idea in order to demonstrate its feasibility or a demonstration in principle with the aim of verifying that some concept or theory has practical potential
RoI	Returns attained from an investment
RPA	Robotic Process Automation refers to a type of rules-based automation technology that helps automate repetitive tasks by mimicking a user's activities. It is non-invasive and typically interacts with a computer-centric task/process through the User Interface (UI) of the underlying software applications
Semi-structured data	Semi-structured content is one that does not conform to the pre-defined structure of content but contains tags / other markers to separate semantic elements and enforced hierarchies. In short, it has a self-describing structure. The placeholders of the content can be in varied sequences
Software-as-a-Service (SaaS)	SaaS is a software licensing and delivery model wherein the software is hosted centrally by a third-party provider and is made available to customers over the internet. It is also referred to as on-demand software

Glossary of key terms used in this report (page 3 of 3)

Transaction-based pricing	Output-based pricing structure; priced per unit transaction with significant price differences between onshore and offshore models
Structured data	Structured content is one that conforms to the pre-defined structure in terms of tags to separate semantic elements and enforced hierarchies of records and fields. Moreover, the placeholders for the content have a pre-defined sequence
Usage-based pricing	Value-based pricing structure; pricing based on per-page or per-document processed
Unstructured data	Unstructured content refers to information that either does not have a pre-defined data model or is not organized in a pre-defined manner. Unstructured information is typically text-heavy but may contain data such as dates, numbers, and facts as well
Vertical-specific business processes	Vertical-specific business processes refer to processes that are specific to a department within an organization and are often directly related to the key revenue-earning business. Examples include lending processes in case of the banking industry and claims processing in case of the insurance industry



Everest Group is a leading research firm helping business leaders make confident decisions. We guide clients through today's market challenges and strengthen their strategies by applying contextualized problem-solving to their unique situations. This drives maximized operational and financial performance and transformative experiences. Our deep expertise and tenacious research focused on technology, business processes, and engineering through the lenses of talent, sustainability, and sourcing delivers precise and action-oriented guidance. Find further details and in-depth content at www.everestgrp.com.

This study was funded, in part, by [AntWorks](#)

Stay connected

Dallas (Headquarters)
info@everestgrp.com
+1-214-451-3000

Bangalore
india@everestgrp.com
+91-80-61463500

Delhi
india@everestgrp.com
+91-124-496-1000

London
unitedkingdom@everestgrp.com
+44-207-129-1318

Toronto
canada@everestgrp.com
+1-647-557-3475

Website
everestgrp.com

Social Media
 @EverestGroup
 @Everest Group
 @Everest Group
 @Everest Group

Blog
everestgrp.com/blog

NOTICE AND DISCLAIMERS

IMPORTANT INFORMATION. PLEASE REVIEW THIS NOTICE CAREFULLY AND IN ITS ENTIRETY. THROUGH YOUR ACCESS, YOU AGREE TO EVEREST GROUP'S TERMS OF USE.

Everest Group's Terms of Use, available at www.everestgrp.com/terms-of-use/, is hereby incorporated by reference as if fully reproduced herein. Parts of these terms are pasted below for convenience; please refer to the link above for the full version of the Terms of Use.

Everest Group is not registered as an investment adviser or research analyst with the U.S. Securities and Exchange Commission, the Financial Industry Regulatory Authority (FINRA), or any state or foreign securities regulatory authority. For the avoidance of doubt, Everest Group is not providing any advice concerning securities as defined by the law or any regulatory entity or an analysis of equity securities as defined by the law or any regulatory entity.

All Everest Group Products and/or Services are for informational purposes only and are provided "as is" without any warranty of any kind. You understand and expressly agree that you assume the entire risk as to your use and any reliance upon any Product or Service. Everest Group is not a legal, tax, financial, or investment advisor, and nothing provided by Everest Group is legal, tax, financial, or investment advice. Nothing Everest Group provides is an offer to sell or a solicitation of an offer to purchase any securities or instruments from any entity. Nothing from Everest Group may be used or relied upon in evaluating the merits of any investment. Do not base any investment decisions, in whole or part, on anything provided by Everest Group.

Products and/or Services represent research opinions or viewpoints, not representations or statements of fact. Accessing, using, or receiving a grant of access to an Everest Group Product and/or Service does not constitute any recommendation by Everest Group that recipient (1) take any action or refrain from taking any action or (2) enter into a particular transaction. Nothing from Everest Group will be relied upon or interpreted as a promise or representation as to past, present, or future performance of a business or a market. The information contained in any Everest Group Product and/or Service is as of the date prepared, and Everest Group has no duty or obligation to update or revise the information or documentation. Everest Group may have obtained information that appears in its Products and/or Services from the parties mentioned therein, public sources, or third-party sources, including information related to financials, estimates, and/or forecasts. Everest Group has not audited such information and assumes no responsibility for independently verifying such information as Everest Group has relied on such information being complete and accurate in all respects. Note, companies mentioned in Products and/or Services may be customers of Everest Group or have interacted with Everest Group in some other way, including, without limitation, participating in Everest Group research activities.

This document is for informational purposes only, and it is being provided "as is" and "as available" without any warranty of any kind, including any warranties of completeness, adequacy, or fitness for a particular purpose. Everest Group is not a legal or investment adviser; the contents of this document should not be construed as legal, tax, or investment advice. This document should not be used as a substitute for consultation with professional advisors, and Everest Group disclaims liability for any actions or decisions not to act that are taken as a result of any material in this publication.