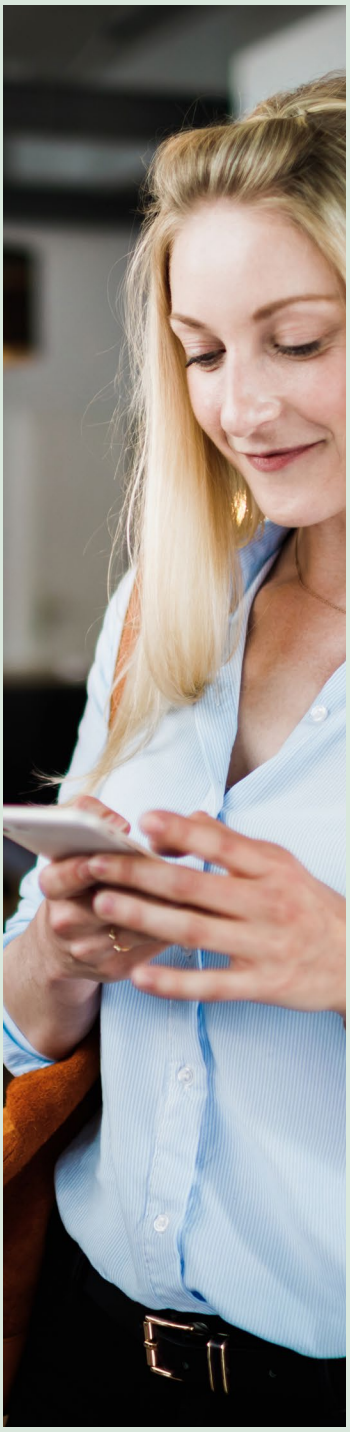
WHITE PAPER



REVOLUTIONIZING THE CONTENT CAPTURE SERVICES WITH NATIVE CLOUD SERVICES

Infosys®
Navigate your next

1. Introduction:

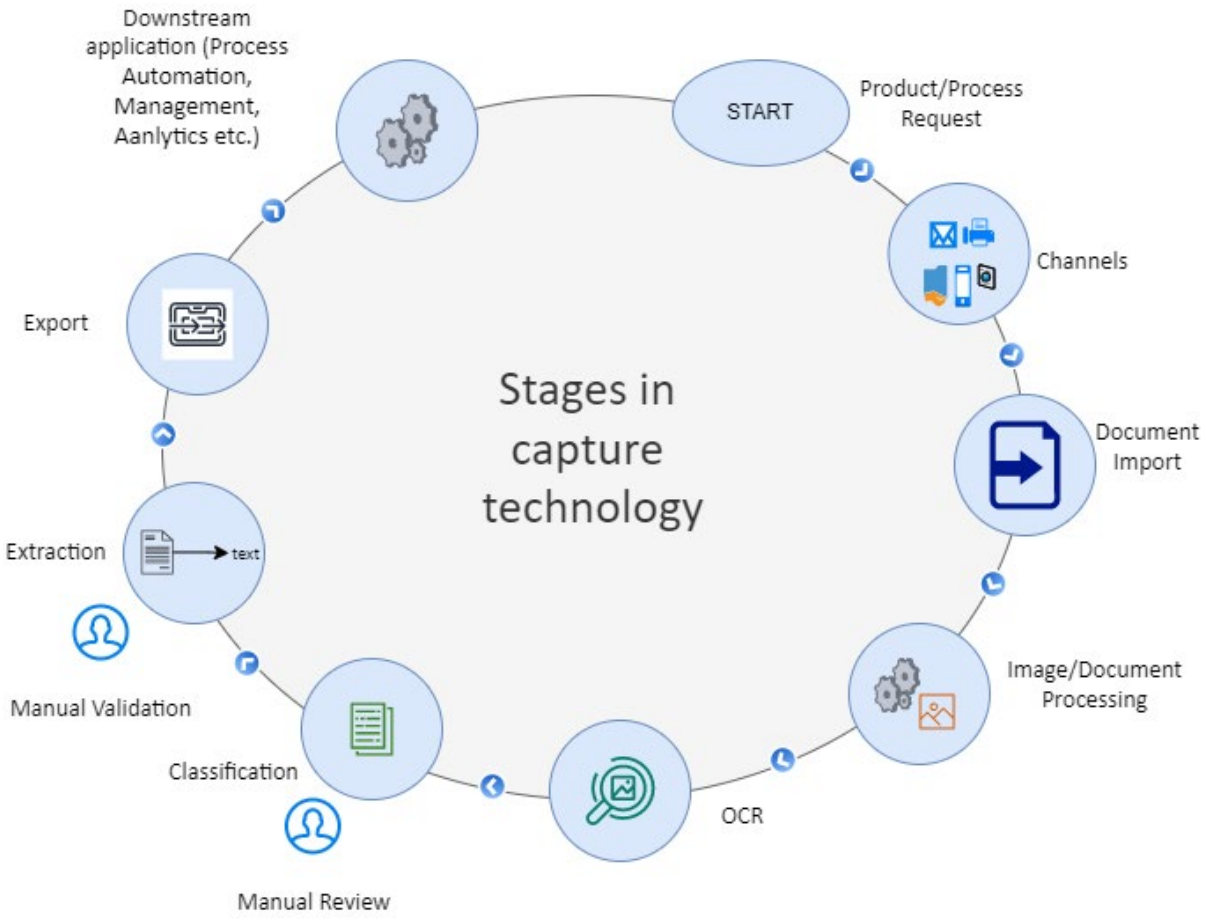
Cloud service providers (CSP) are growing in leaps and bound. Most cloud service providers provide various capture-related service offerings. Most cloud service providers have concentrated heavily on the array of cognitive services, which are helpful for most capture use cases. These cognitive services are also termed as "computer vision" services. Over the period, plethora of data has been exposed to search engines of CSPs. The learning models embedded in cognitive services from these CSPs are built with this huge learning data, which results into accurate results. The cloud infrastructure has also enabled the fast processing of these capture use cases. To summarise, with range of cognitive services, CSP's allow accuracy, quick processing, and high-volume support for capture use cases. This whitepaper details various capture services offered in the cloud, their benefits towards digitization, and content services.

In the upcoming sections, first, the background of capture technology and cloud-native capture technology is briefly introduced. Later, various service offerings available in the cloud and how they can be used to benefit organisations in several areas are discussed. Toward the end, the conclusion of the discussion is presented.

2 Background:

2.1 Overview of capture technology

Document processing, management, and archival are a few document-centric critical business functions used in many organizations. The organization receives the documents via multiple channels and in various formats. Capture is an essential pillar in organizations, which creates digital representation of these documents. Which then can be leveraged by enabling digitization, accurate filing, and automation, enabling the transformation of several manual business processes into automated and effective business processes.

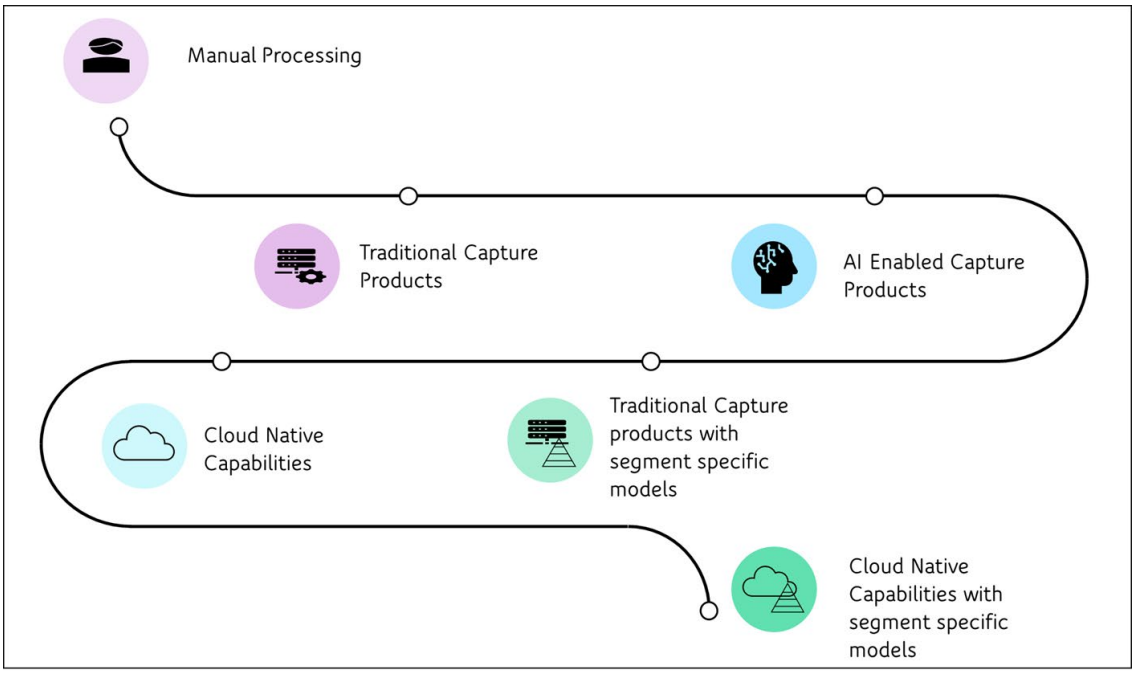


The document capture process has below essential stages:

- **Document Ingestion:** Ingesting the documents from many channels such as emails, fax, scanner, or digital upload.
- **Document enhancement:** This is a stage in which received documents are enhanced to improve the accuracy of OCR results during the subsequent phases of capture.
- **Optical character recognition (OCR):** Text recognition from image-based documents is performed in this stage. The output of this stage is machine interpretable text for further analysis.
- **Document Classification:** In this phase, documents are classified/ categorized into business-meaningful predefined categories.
- **Document Content Extraction:** In this phase, the relevant information is extracted from documents. This extracted information is used in many downstream processing or business processes for business process automation.
- **Document Review/Validation:** This is an optional phase in which manual expertise reviews the document classification results and validates the extracted information
- **Document Export:** This phase exports the classified documents and relevant extraction information to downstream applications such as Document Management System.

2.2 Where industry is heading with capture technology:

Below infographic image explains the maturity journey of capture technology throughout the industry.



2.2.1 Manual Processing

Before revolution of digital world, the information extraction from document needed manual intervention. A human used to read the physical (hard copy) document and then used the required information to process the required further activities.

2.2.2 Traditional Capture Products

The digital revolution brought some excellent products, technologies, and software that helped to automate the information extraction from the documents. Classification of digital documents based on image or layout evolved to classify documents into specific types or categories such as passport, salary slips, etc. Further on, the fixed form method of extraction, also known as positional-based extraction, helped extraction of information from the documents. Classification and Extraction based on barcodes and patch codes helped further in this digital capture process. With the help of regular expressions or scripting automation of information validation was started.

Further, the use of tooling for learning started, shifting to web-based clients rather than thick clients. Capture solutions providers also added various import connections such as web, hot folder, email, MFP, and bulk scanners to enable ingestion via multiple channels. This paved the way for mailroom automation, and bulk scanning document flows, and several other mechanisms of automation of document ingestions. Even export mechanisms provided various options such as database, CSV, XML, and JSON. Both import and export mechanisms are also enhanced with options like Content Management Interoperability Services (CMIS) which offers vendor independent access to document management repositories.

2.2.3 Traditional Capture products with specific document type models

Capture products also realised that providing extraction for specific document type as feature was the need of hour. For e.g., pre-configured models to extract information from invoices which reduced the learning time. Using the vast learning samples, and contributions from partners and customers, preconfigured models was able to reduce time to market and was the major factor of success

2.2.4 AI Enabled Capture Products

Revolution did not stop at specific document type models; capture solution providers started providing supervised machine learning capabilities. With this implementation, a human support loop is enabled which can enhance the prebuilt learning models dynamically. This was the step towards solving biggest challenge of minimizing the time required for document/template-based learning. Along with this, capture solution providers, started exposing the services to get classification, extraction, and OCR results. Increasing the performance of the classification and extraction services was also the need of the hour. Further multiple OCR engines became the order of the day. Also, identifying and increasing the accuracy of handwritten content got the focus which helped ICR (Image character recognition) to become key part of capture products.

2.2.5 Cloud Native Capabilities

Cloud-native capabilities are being used for capture requirements such as OCR, classification of document etc. They are enabled with AI techniques such as machine learning (ML) and natural language processing (NLP) which analyses the text with the help of pre trained AI models to provide expected results. Cloud native cognitive capabilities provide all services as single entity due to which the capture of information was not limited to documents. Video and audio processing use cases are also getting recognized with the help of audio-to-text conversion and then processed through extraction services. For e. g. customer support activities such as complaint/enquiry can be modernised with use of audio to text recognition.

2.2.6 Cloud Native Capabilities with specific document type models

Various CSPs are now targeting to providing service towards specific document type models, for e.g., Lending DocAI offering provided by Google cloud platform. CSPs are now enabling direct results of classification and extraction with their services for this specific type of document.

2.3 Establishing the contribution of cloud native capture towards digital transformation

Cloud providers have many services with cognitive capabilities which can enable the transformation of the capture process. All major cloud service providers have made significant strides in cloud computer vision, cognitive OCR, and cloud AI capabilities. This enables organizations to develop below capture capabilities with ease:

- *Image processing*
- *OCR (Fetching content from digital documents)*
- *Classification of documents, images*
- *Extraction of relevant information from documents*

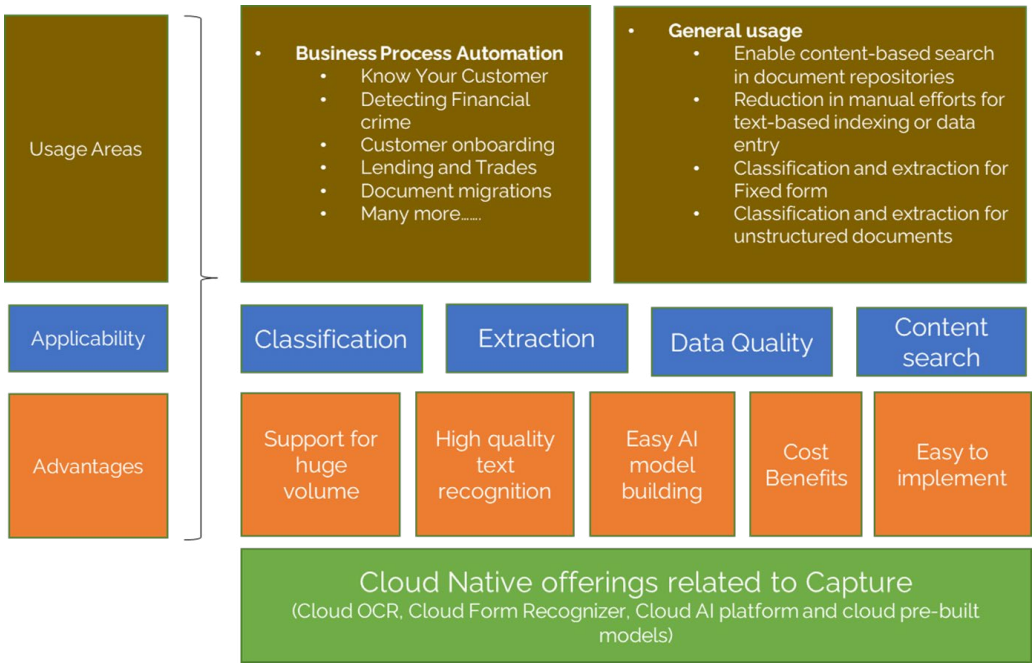
With their cognitive services and computer vision offerings, many cloud providers can enable digitization and automation for the capture processes.

Capture systems are majorly dependent on gigantically learning data. Higher the learning data, the higher the accuracy. Cloud-native offerings in document capture space decrease time andeffort,. CSP has huge set of learning data which has been gathered over long period of time,. Such learning data have enabled CSPs to implement accurate capture service.

CSP also offers advantages such as supporting high-volume processing, accurate results, cost saving, and more straightforward implementation. These cloud-native capture offerings enable transformation and modernization in digital use cases, such as Integrating Content Intelligence for business process automation, auto Classification/Extraction, fixing Data Quality issues, and improving Content Search.

Cloud-native capture offerings with these services enables modernization and automation and thus helps in digital transformation.

The below diagram illustrates the contribution of cloud-native capture offerings toward digital transformation:



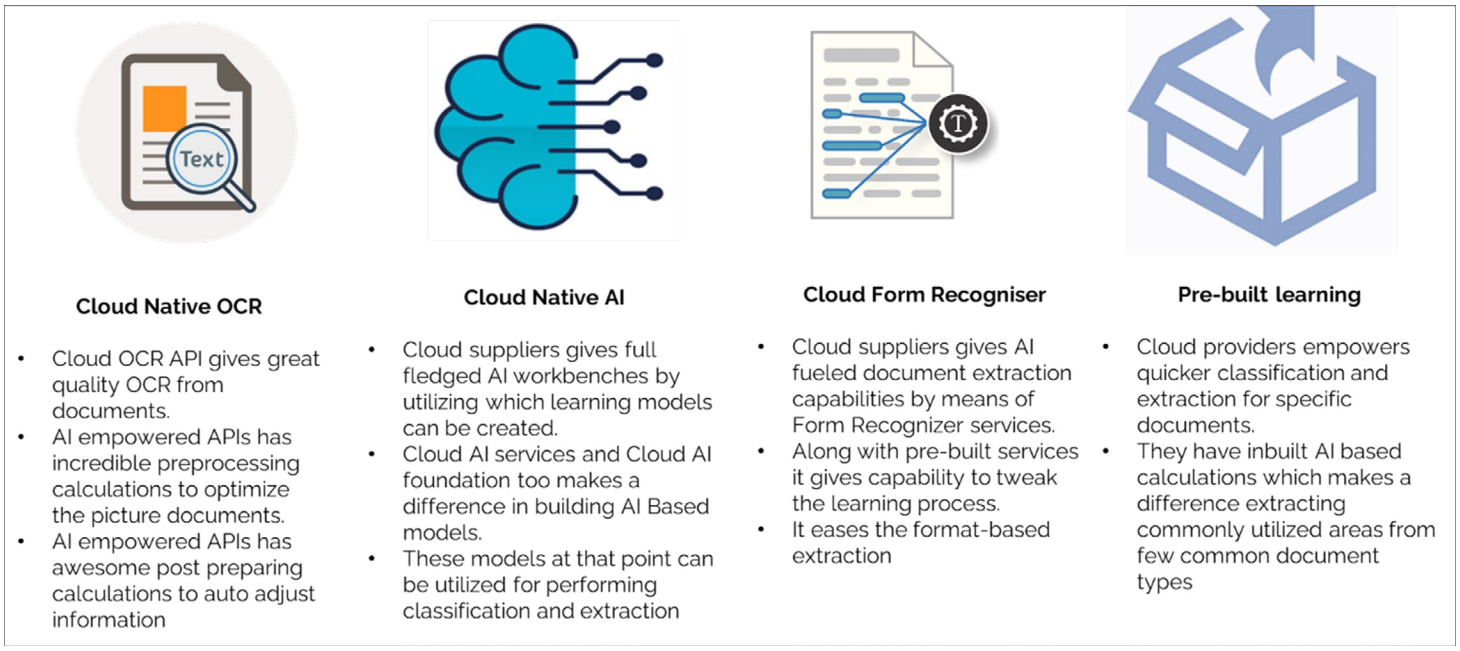
3. Detailing the cloud native capture service:

With their Cognitive Services, CSPs enable cognitive intelligence development into Capture systems that can integrate via REST APIs and Cloud SDK.

All the major CSPs have below dedicated service offerings, which can enhance many capture use cases:

- **Cloud OCR:** CSPs have pre-built services which can receive digitized documents as input and then provide text data from those documents as a response. This OCRed text can then be further used to classify the document or to extract meaningful information from those documents. This OCR text can also be used to build content-based search use cases. Most CSPs have built these OCR services based on huge training data sets which enables the accuracy of OCR. These OCR services can also be used as pay-as-you-go.
- **Cloud AI:** All the major CSPs have many services which can enable artificial intelligence for capture use cases. Techniques such as natural language processing (NLP), machine learning (ML) can easily use or configured as per the need. Organisation can use native cloud AI models or customise these AI models as required for a specific implementation. Organisations do not need any separate platform or infrastructure to use or create these AI models.
- **Cloud Form Recognizer:** It is a cloud service offering that enables training of structured (fixed format) documents to get results such as classification and extraction. Many organizations receive documents in a fixed-format. Classifying or extracting the relevant information from such documents can be done based on the position and format of the document. A specific location or relative location of the text in the document can be used for training the models. Then these models can be used to get the results such as classification and extraction. Cloud form recognizer is easy to set up and use. Organizations do not need a separate platform or infrastructure to set up form recognitions.
- **Pre-Built learning:** Most cloud service providers also have built-in AI models which are exposed directly as API. These prebuilt models can provide general information for a specific document type, such as identity documents or visiting cards. This saves the effort required for the organizations to build these document-specific data models.

Let's understand details of the cloud native capture service offerings with help of below image:



3.1 Summary of offering available at major cloud players for above :

Below table explains the cloud capture specific offerings available with major cloud service providers.

Category	Azure	aws	Google Cloud
OCR	• Azure Computer Vision • Azure OCR • Azure Read Text	Amazon Textract	Vision API (Detect Text)
Fixed form-based extraction	Azure Form Recognizer	Amazon Textract	Document AI - Form Parser
AI (Classification/Extraction)	Azure AI	• Amazon Comprehend • AWS Machine Learning	• Natural Language AI • Document AI
Segmented Learnings	• Receipts • Business Cards • Invoices • Identity Documents	• Regular Invoices / Bills • Financial Documents • Medical Documents • Handwritten Documents • Pay slips or Employee Documents	• Lending DocAI • Procurement DocAI

4 Comparison of different cloud offering

All the cloud service providers have all the possible cloud native capture offerings. It isn't easy to compare the offerings as they are ever-growing and improving but let's see some differences and basic comparisons in certain situations.

4.1 Comparison of OCR offering with different cloud providers and traditional OCR

Below table indicates a comparison assessment on different OCR offerings. Below are the parameters for assessment:

- o No image enhancements
- o Same image must go through all selected OCR engines
- o Only evaluate OCR quality for several image scenarios
- o Try and test methodology for comparison

Image Scenario	Capture Product with Recostar	Capture Product with Tesseract	Azure OCR	Azure Read Text	AWS Textract
Business cards with colors and variation of fonts					
Image with same background color with italic and normal fonts					
Image with variation in background color					
Handwritten forms (Bad Sample)					
ID documents with some marking					
Image with typed and inconsistent font					

Different ML framework availability

AWS	AZURE	GCP
TensorFlow, MXNet, Keras, Gluon, Pytorch, Caffe2, Chainer, Torch	TensorFlow, Scikit learn, Pytorch, MS Cognitive Toolkit, Spark ML	TensorFlow, Scikit learn, Keras, XGBoost

5 Key considerations to select an offering based on the customer needs

Below are some key considerations which can help to select a right offering for the business use cases based on the customer needs.

- If the CSP is already selected. For example, if an organisation has strategically selected a CSP then selecting the cognitive services specific to chosen CSP.
- Selection based on benchmark on the document samples. This technique is useful for organisation which processes certain types of documents and wants specific results for them.
- If CSP provides the already trained segment wise models. For example, if an organisation has requirements for classifying the incoming identity documents, then they can choose the cloud native capture offerings which directly provides pre-built learning models for identity documents.
- Cost comparison-based selection. This technique is used to compare the cost of cloud native capture offerings and then cost-effective solution is selected.
- Compatibility for application integrity. Here organisation evaluates the existing solutions and their integration capability with specific cloud native capture offerings.
- Based on available knowledge pool. For example, if organisation has bigger knowledge pool for google then that organisation can select GCP for cloud native capture offerings.

6 Cloud native Capture based modernization for ECM

Enterprise Content Management (ECM) applications manages the documents in various phases of business processes across the enterprises. In those phases some important capture specific steps are as below:

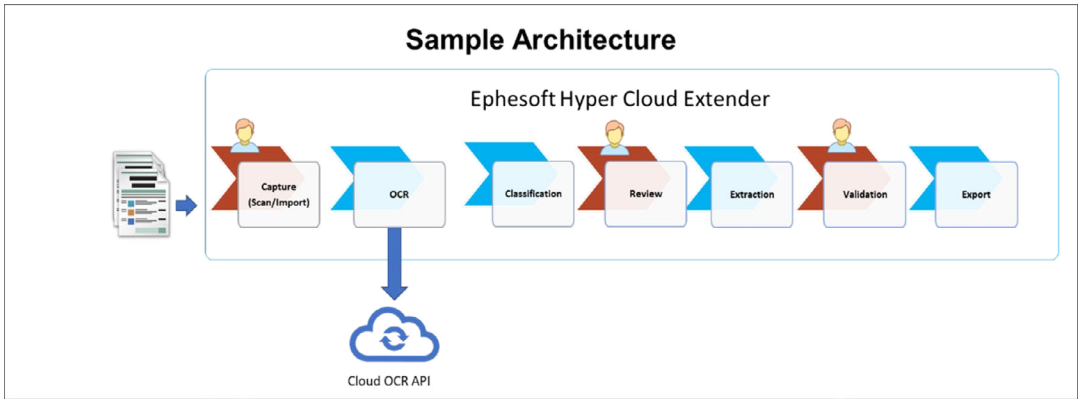
- Digitisation of document
- Classification of documents
- Extracting the relevant information from the documents

Cloud-native capture can reduce the go-to-market time of these steps with quicker integrations, high-volume processing, accuracy, and cost benefits. The vast data acquired by the CSPs as part of their search engines, results in more accurate AI models for cloud cognitive services.

Below are few ways by which ECM applications can be modernized with help of cloud native capture:

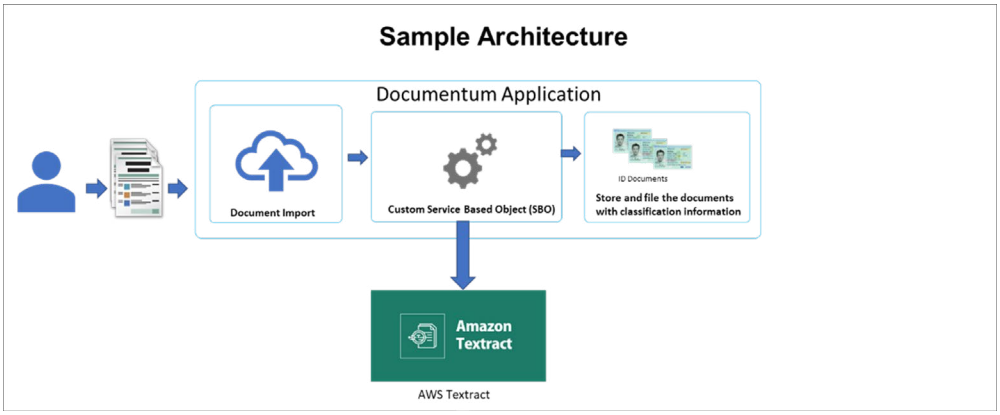
- Traditional capture platforms consuming cloud native OCR

For example, product such as Ephesoft Hyper Cloud Extender. Below is sample Architecture:



ECM applications consuming cloud cognitive services to perform document classification/categorisations

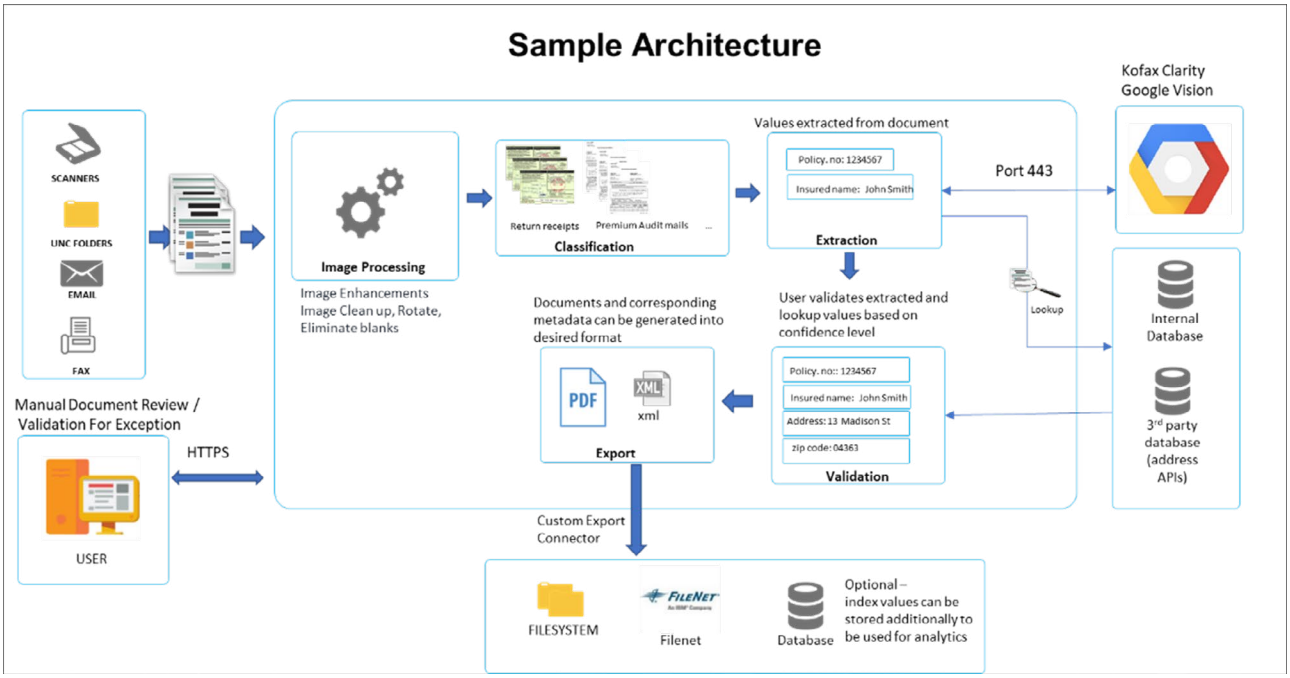
- Uploading a document in product like Documentum or OpenText can be customised to integrate with AWS Textract to fetch document classification of identity documents. Below is sample architecture:



- ECM applications consuming cloud cognitive services to extract relevant information from documents

Product like Documentum or OpenText can be customised to integrate with Azure Cognitive Services API to extract the relevant document information.

- Traditional capture product such as Kofax TotalAgility has Cloud AI plugin which consumes cloud offerings from Amazon Web Services, Microsoft Azure, and Google Cloud. Below is sample architecture:

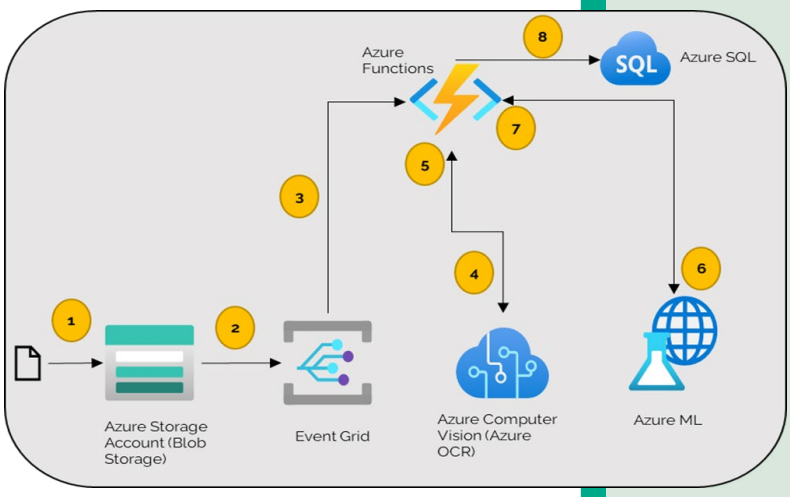


6.1 Reference architecture with explanation

In earlier section we focused on cloud native capture offerings and their direct integration or use with ECM. Let's now focus on standalone end-to-end solutions which can help many use cases. Below are few of the Cloud Native Capture Sample Reference Architectures.

6.1.1 Sample Architecture for batch processing for Classification

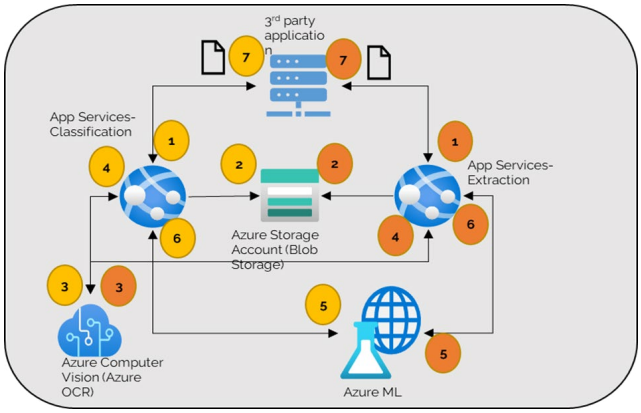
This reference architecture would enable proper categorisation and storage for an organisation who receives huge volume of documents at a time.



- Image will be sent by 3rd party applications to app services
- App services persist documents in blob storage.
- App services triggers Computer Vision API to fetch OCR.
- Once OCR is received then App services triggers Azure ML services to find classification/extraction results
- Once the classification /extraction results has been returned from the Azure ML services, Azure app services responds to 3rd party applications

6.1.2 Sample Architecture for real time processing for Classification/Extraction

This reference architecture would enable use cases such as a robot downloading a document and wants to perform classification and extraction of downloaded document. Or else it is also applicable in document exchange use cases where customer uploads a document where by using this reference architecture real time classification and extraction results can be displayed to the customer..



- Image will be sent by 3rd party applications to app services
- App services persist documents in blob storage.
- App services triggers Computer Vision API to fetch OCR.
- Once OCR is received then App services triggers Azure ML services to find classification/extraction results
- Once the classification /extraction results has been returned from the Azure ML services, Azure app services responds to 3rd party applications

6.2 Future view of the landscape

Below are few examples with details for few futuristic usage of cloud native capture.

6.2.1 Cloud native capture as PaaS

- Cloud Service Providers offerings around cloud native capture can also be wrapped and provided as platform offering where many organisations can use them to train and create models for various ECM use cases.
- For example, Azure form recogniser can be made available to developers, business developers or citizen developers .

6.2.2 Cloud native capture as SaaS

- Cloud Service Providers offerings around cloud native capture can also be provided as API offering where many organisations can use them to fetch direct results such as OCR, Classification and extraction.
- For example, Amazon Textract API endpoint which provides OCR for 3rd party application.

6.2.3 Segmented Service Offering

- Many Cloud Service Providers have focus area to provide segmented services for document classification and extraction.
- For example, GCP have a good focus area on providing segmented offering in Lending space with offerings such as Google Lending DocAI

6.2.4 Capture products bundled with cloud native capture

- All traditional capture products providing in built offerings by integrating with cloud native services
- For example, Ephesoft capture vendor focusing on utilising cloud native offering to gain performance benefits in OCR module.

6.3 Sample use cases

Below are few sample use cases of in ECM modernisation with cloud native capture:

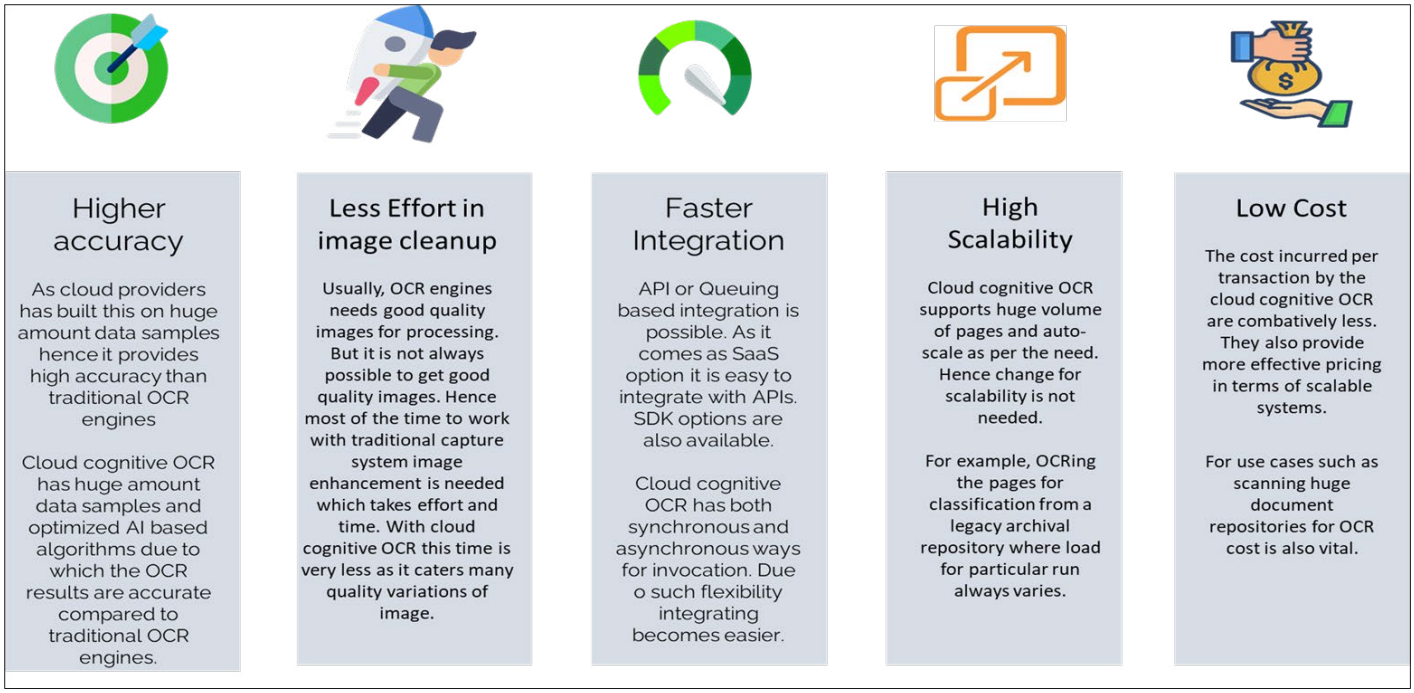
Business Use Cases Model Catalogue	Purpose	Applicability / Sector	Use cases	Business User roles	End user experience of purpose
Contract / Agreement summarization	Get OCR from documents to create summary of contracts and agreement	Finance (BFSI) & Insurance (HIL)	Faster processing and reusing contract & agreements	1. Front office user, Legal officer 2. Legal officer 3. Legal officer, business users 4. Process Automation for most business users	1. Quickly understand important clauses in the contract 2. Quicker evaluation of risk in the contract 3. Faster contract related decision making 4. Automatic alerts on important dates in contract
LIBOR Contract identification	Get OCR from documents to identify LIBOR contracts	Finance (BFSI)& Insurance (HIL)	Repapering LIBOR contracts	1. Process Automation for most business users 2. Process Automation for most business users 3. Legal officer, back-office users, front office users	1. Automatic LIBOR contract identification based on the terms used in contracts 2. Automatic re-papering the contract based on predefined rules 3. Identify similar contracts based on parties in the contract, clauses in the contracts etc.
Document Type Identification	Get OCR from documents to identify documents uploaded by customer	Finance (BFSI) & Insurance (HIL)	Digital Document Exchange	1. Customer uploading documents as a part of a process 2. Customer uploading documents as a part of a process 3. Customer uploading documents as a part of a process	1. Real-time validation of a document uploaded by a user and flagging incorrect documents 2. During document upload by customer, check whether the document is already exists in the backend system and message user that the document already exists and giving option to customer to overwrite previous document with newly uploaded document. 3. Pre-population of data from documents which are being uploaded, so that user do not have to type it again.

7 Benefits of Cloud native Capture offering in ECM modernization

Enterprise content management (ECM) enables many organisations towards digitising and automating below document centric functions such as ingestion, management, storage, archival & retention, and distribution.

7.1 Summary of benefits

Usage of cloud native capture offerings enables several benefits. Below image explains the major benefits of cloud native capture service offerings in modernization of ECM.



8 Conclusion

- Cloud Native Capabilities are proving their worthiness in Capture related requirements.
- Cloud native capabilities overall provides quick, accurate and cost-effective solutions compared to traditional capture products.
- Cloud Native capabilities has cognitive learning inbuilt, and they are also easy to integrate which reduces implementation efforts.
- OCR accuracy out of Cloud Native OCR is good. Which in turn helps accurate classification and extraction. This would help improving automatic case processing with less manual involvement.
- Cost is transaction wise and effective
- As this is on Cloud Native, scalability is easy and faster.

Authors

Nilesh Khot
Senior Technology Architect

Mentor

Girish Pande
Principal Technology Architect

For more information, contact askus@infosys.com

Navigate your next