Designing Document Automation for a Changing AI Landscape

HOW TO STAY AHEAD OF THE CURVE

AI+IM GLOBAL SUMMIT

Summary

- The Intelligent Document Processing (IDP) landscape is undergoing significant change. Capabilities that were once proprietary—such as document extraction and classification—are increasingly available through hyperscalers, who are investing heavily in AI services that accelerate innovation and reduce barriers to adoption.
- As AI models continue to improve, accuracy scores and confidence thresholds are rising, minimizing the need for manual intervention. At the same time, organizations must support a broad mix of electronic content and remaining physical documents across an expanding range of document types. Strength across this diversity is becoming a key driver of reliable automation outcomes.

About this Session

Hosted by Kodak Alaris

- This session examines **how organizations can respond** to the ongoing race for AI by avoiding early commitment to a single technology or platform. **A future-ready automation strategy requires flexibility**—allowing organizations to adopt the most appropriate AI models as capabilities evolve, while maintaining governance, compliance, and operational control.
- **Attendees will gain practical insight** into building document automation programs where AI supports extraction and classification within a governed framework, rather than making autonomous business decisions. **The session will focus on architectural principles and strategic considerations** that help organizations remain adaptable, compliant, and prepared for continued change.

AI+IM GLOBAL SUMMIT

Key Takeaways

- How the role of proprietary extraction is changing in the IDP market
- Why hyperscaler-driven AI is influencing automation strategy and architecture
- The impact of higher accuracy and scoring confidence on manual processing
- The importance of supporting diverse document types and formats
- How model-agnostic approaches reduce risk and improve long-term flexibility
- Why governance remains essential as AI capabilities expand

Evolution in Intelligent Document Processing

The next AI wave is reshaping IDP.

Intelligent Document Processing

Intelligently extract, validate, and connect information into systems and processes

Enhancing with Generative AI

Leveraging
Large Language Models (LLM)

Specialized Models

Accurate data, better decisions & outcomes

- ✓ **Delegate** work **to trusted AI**
- ✓ Extract more information, accurately
- ✓ **Automate more** & Accelerate DX
- ✓ Deliver **greater insights** & understanding to drive decisions

AI+IM GLOBAL SUMMIT

How has technology evolved?

Phase 1

- Basic Document Management
- Scan, File, and Retrieve
- Light Workflow

Phase 2

- Accounts Payable Automation
- Electronic Forms
- Advanced Workflow

Phase 3

- Intelligent Document Processing
- Artificial Intelligence
- Machine Learning
- Robotic Process Automation
- OCR

Phase 4

- Visual Language Models (VLM)
- Intelligent Processing Automation (IPA)
- Agentic RAG
- MCP
- Chat AI

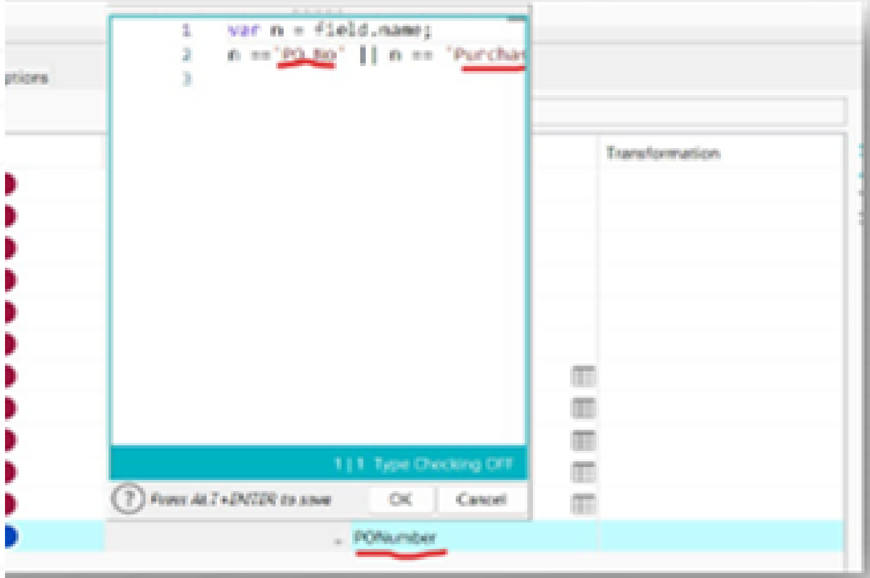
AI+IM GLOBAL SUMMIT

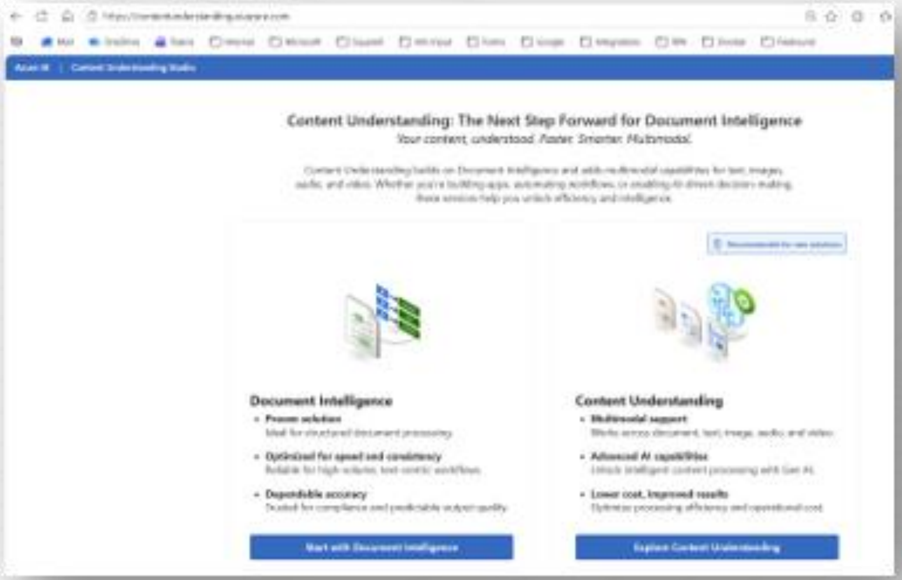
What has changed since the last AI+IM Global Summit?

- **The Old Way**
  - Neural templating with Auto Learning
  - Requires a knowledgebase, few shot
  - Mapping Key-Value Pairs (KVP)
  - Heavily reliant on OCR
- **The New Way**
  - LLM/Generative AI Prompting
  - No knowledgebase, zero shot
  - Uses NLP, visual queues and context to extract data more accurately NOT OCR
  - Works better for unstructured content with no KVP

What you learned before has deprecated

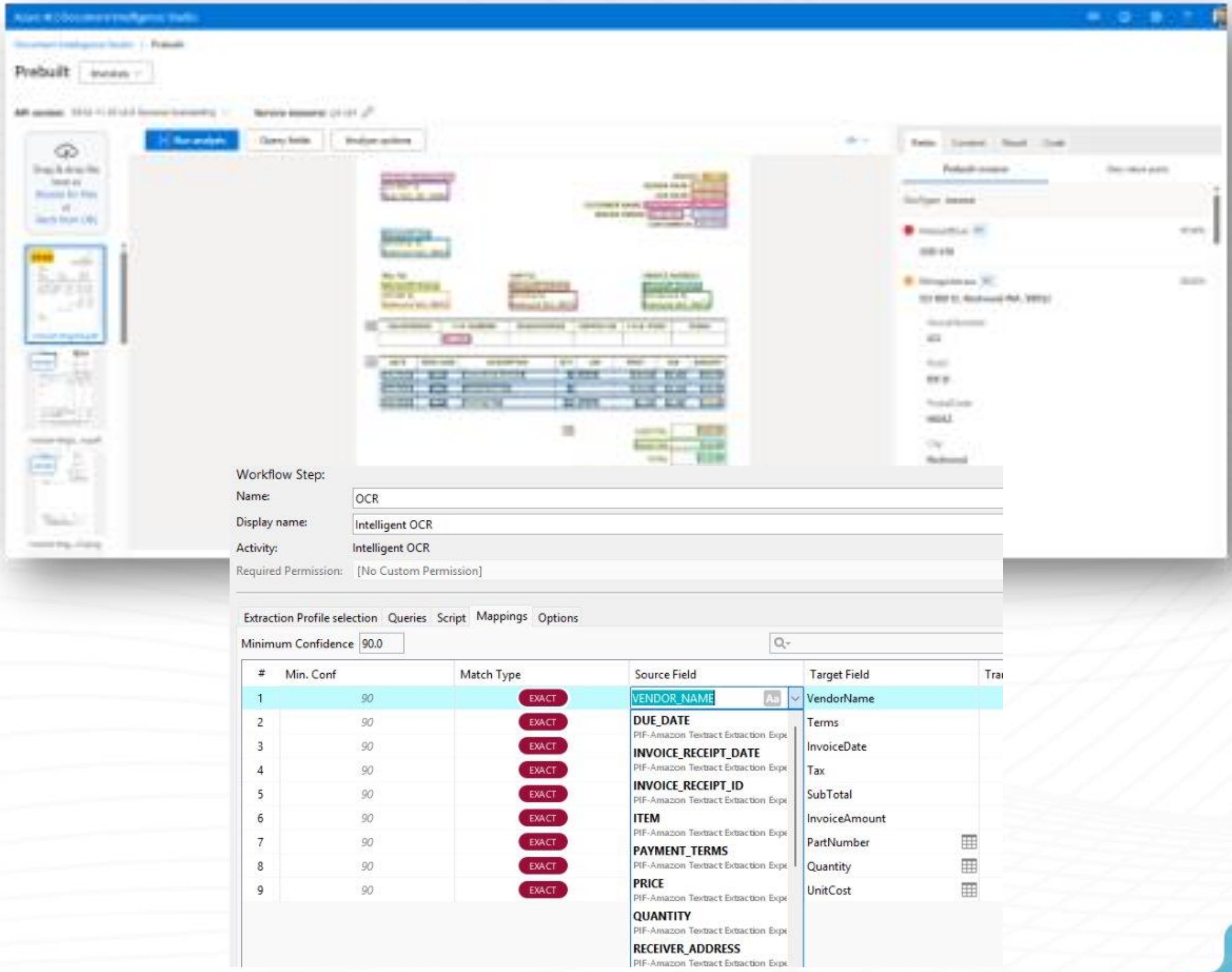
- **Azure AI Document Intelligence** now replaced with **Azure AIFoundry/Content Understanding** using models
- Relying on OCR versus CONTEXT
- Cycling permutations of Key Value Pairs (KVP)
- Multi-model is required for most transactional processing
- Prevalence of Open Source models and self hosting for compliance
- Access to ground truth data for Fine tuning
- Concept of neural templates and auto learning vs fine tunning and modelbuilding





How did we used to do it?

- **Neural template learning**
  - Azure AI Document Intelligence
  - AWS Textract
  - Google Document AI
- **Define key value pairs**
  - Parse the JSON payload
  - Add training and data set
- **Mapping Key value pairs**
  - Default fields show in Source Fields from engine
- **If engine does not find a match**
  - Requires significant JavaScript scripting



Taylor believes that companies building their own models is **not a wise use of capital** unless they are an AGI research lab.

He views software development as an ongoing process that **requires continuous attention**, comparing it to a lawn that needs regular care.

"**Software isn't something you can create once and expect to work forever**. In contrast, pre-training a model involves a significant, one-time cost, but once it's done, the model can be used repeatedly for tasks like generating content."

Models: Build vs. Buy

Ex-OpenAI Bret Taylor Says AI Is a Bubble — Companies Shouldn't Build Their Own Models

OpenAI Chairman just said what we were all thinking…

- Building your own LLM data models and neural networks
  - Initial development costs and resources
  - Ability to get sample documents
  - Pivot with change in technology availability
- Wrapping IDP products around cloud engines
  - Focus on differentiators, UI and tools vs. underlying engine
  - OpenAi, AWS, Azure, Google

Reference: 10/3/24. Story by Giuseppe Ciccomascolo on CCN

AI+IM GLOBAL SUMMIT

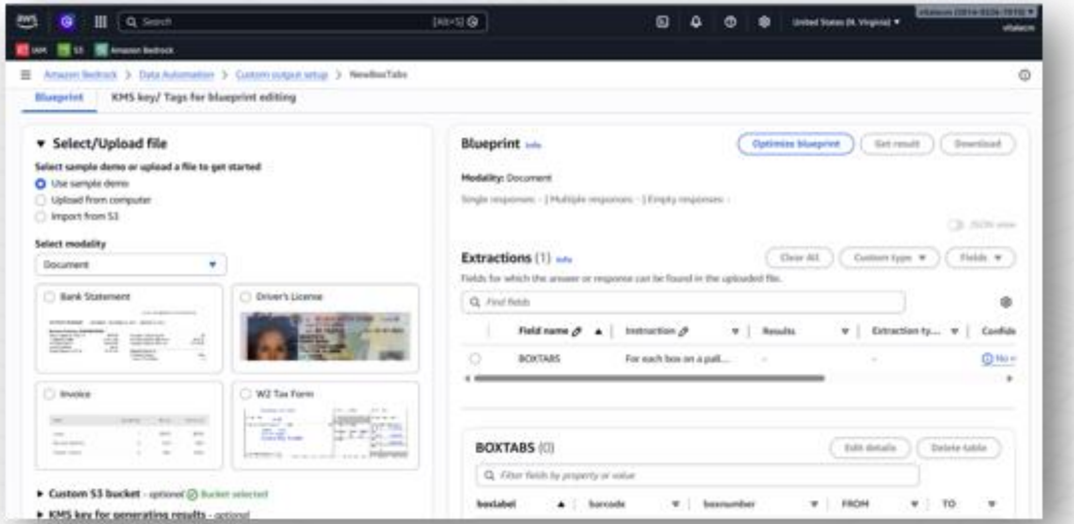
Why be boxed into a model?

- **Speed of Change**
  - Models are changing monthly
  - Lots of deprecation just when you launch a new project
- **Performance**
  - Certain models perform better than other (e.g. Gemini 3.1)
  - Certain models perform better with text, handwriting, video or audio differently, find the best model for the use case
  - Model Routers – let the router pick the correct model
- **Commercials & Costs**
  - Preview models are a lot more expensive, limited RPM (request per min)
  - Open-source models are more cost effective and can be selfhosted
- **Scalability**
  - Don't compete with hypervisor model building – they own the cloud infra and have billions of $ to invest in R&D

Extractions Profiles - Additions

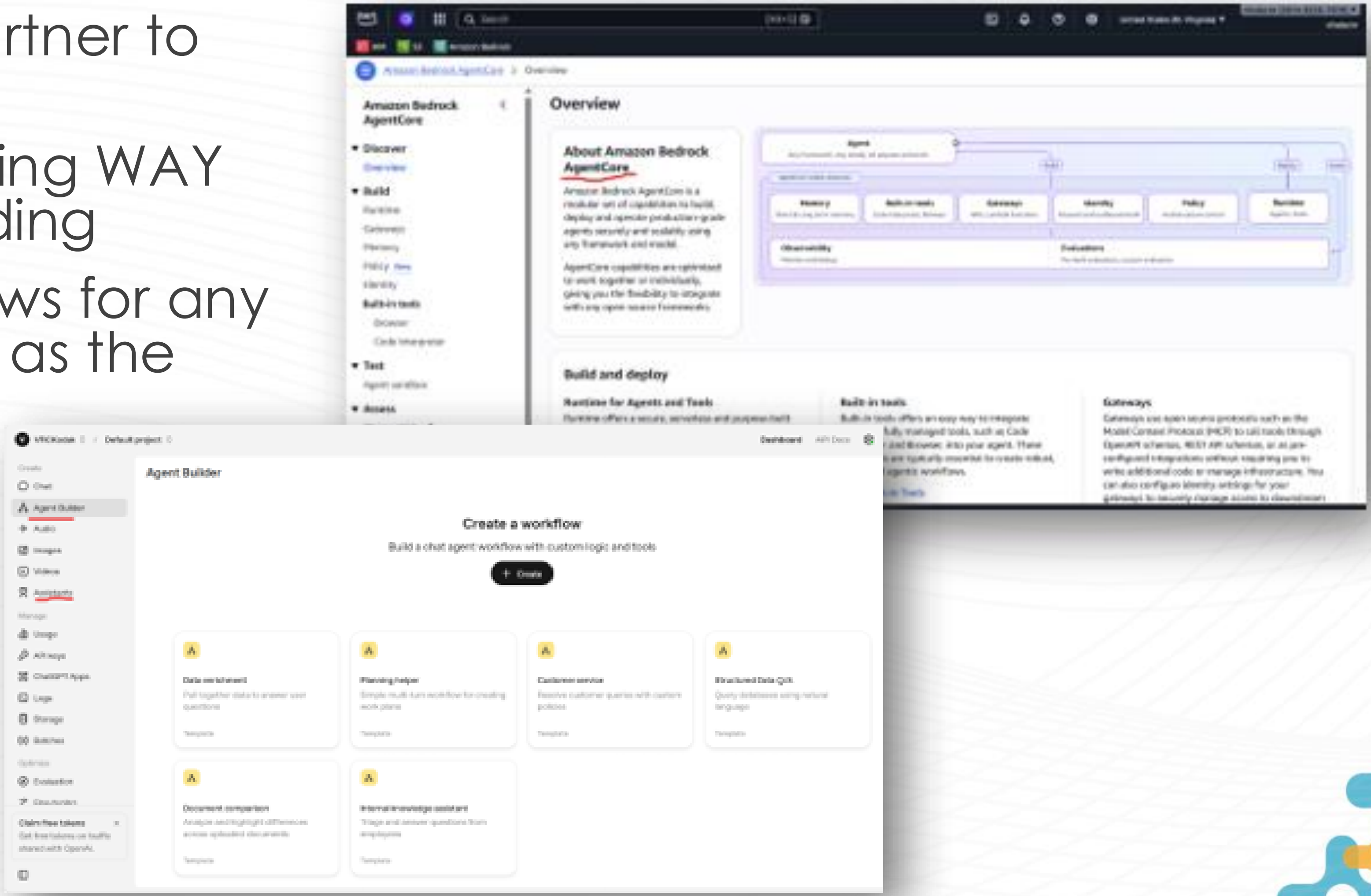
Difference in Cloud setup configuration

- **AWS BDA**
  - More difficult to config in AWS
  - Requires S3 buckets and IAM user get configured prior to BDA setup
  - Requires access to BDA projects and blueprints in AWS to make changes
- **OpenAI**
  - Very easy to setup
  - Simply create an organization and a project object
- **Google Gemini**
  - Must enable Gemini for Google Cloud API
- **Difference in setup configuration**
  - OpenAI – uses prompts for extraction
  - Gemini– uses prompts for extraction
  - AWS BDA – must create blueprints that generate key value pairs



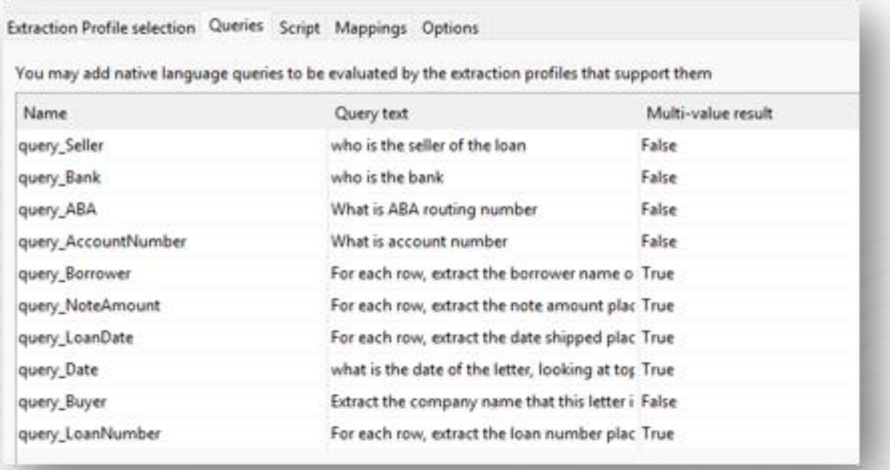
Ecosystem Interoperability

- Pick the right cloud partner to extend beyond IDP
- The hypervisors are doing WAY more than model building
- Open Intelligence allows for any LLM model to be used as the "Intelligent Engine"
- Add-on products for:
  - Agentic AI
  - Reporting
  - Chat AI
  - AI Coding



How do we do it NOW?

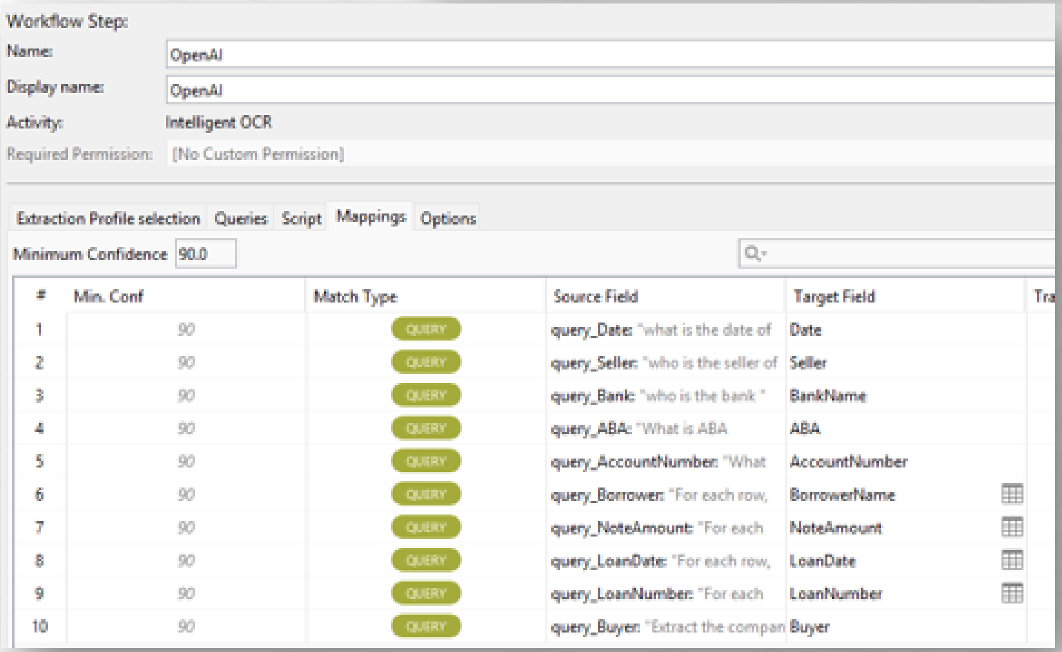
- **Generative AI/LLM Engines:**
  - OpenAI
  - AWS BDA
  - Google Gemini
  - Box AI
- **Define prompts:**
  - No training sets
  - Using NLP to understand human questions into extractable content
  - Visual context
- **If no match**
  - Tweak the prompt



Extraction Profile selection | Queries | Script | Mappings | Options

You may add native language queries to be evaluated by the extraction profiles that support them

Name	Query text	Multi-value result
query_Seller	who is the seller of the loan	False
query_Bank	who is the bank	False
query_ABA	What is ABA routing number	False
query_AccountNumber	What is account number	False
query_Borrower	For each row, extract the borrower name o	True
query_NoteAmount	For each row, extract the note amount plac	True
query_LoanDate	For each row, extract the date shipped plac	True
query_Date	what is the date of the letter, looking at to|	True
query_Buyer	Extract the company name that this letter i	False
query_LoanNumber	For each row, extract the loan number plac	True



Workflow Step:

Name:	OpenAI
Display name:	OpenAI
Activity:	Intelligent OCR
Required Permission:	[No Custom Permission]

Extraction Profile selection | Queries | Script | Mappings | Options

Minimum Confidence 90.0

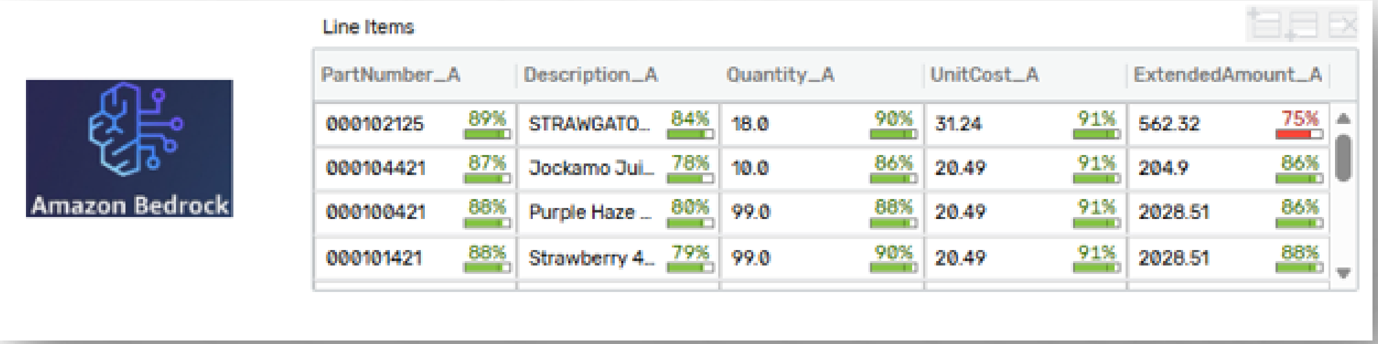
#	Min. Conf	Match Type	Source Field	Target Field	Tra
1	90	QUERY	query_Date: "what is the date of	Date	
2	90	QUERY	query_Seller: "who is the seller of	Seller	
3	90	QUERY	query_Bank: "who is the bank "	BankName	
4	90	QUERY	query_ABA: "What is ABA	ABA	
5	90	QUERY	query_AccountNumber: "What	AccountNumber	
6	90	QUERY	query_Borrower: "For each row,	BorrowerName	
7	90	QUERY	query_NoteAmount: "For each	NoteAmount	
8	90	QUERY	query_LoanDate: "For each row,	LoanDate	
9	90	QUERY	query_LoanNumber: "For each	LoanNumber	
10	90	QUERY	query_Buyer: "Extract the compan	Buyer	

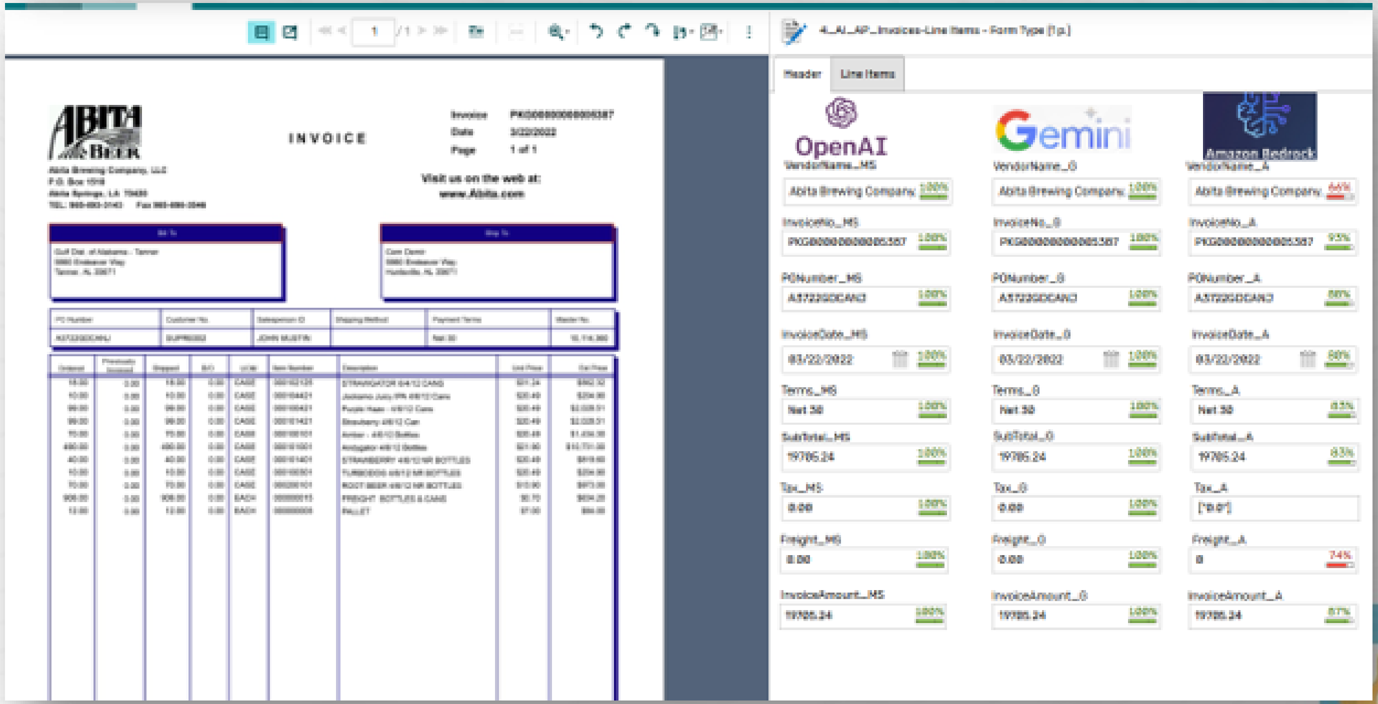
AI+IM GLOBAL SUMMIT

Invoice Processing

Generative AI Voting

- **Compare Results From:**
  - OpenAI
  - AWS BDA
  - AWS Textract
  - Google Gemini
  - ABBYYVantage
  - Hyperscience
  - Box.com AI
  - Azure OpenAI





AI+IM GLOBAL SUMMIT

Models vs. Neural Templating

Performance at Scale

- Accuracy of Models vs. Templating
- ***KVP mapping and parsing*** hoping you programmed every permutation of a word (e.g. P.O. Number, PO #)
- ***Context based***
  Models don't leverage KVP's but rather uses visual context to extract data with better accuracy
- We see a ***10-20% plus increase*** in extraction accuracy especially with ***poor quality documents and tabular data***

Neural Template

Line Items

PartNumber		Description		Quantity		UnitCost		ExtendedAmount	
000102125	49%	STRAWGATO...	93%	18.00	92%	31.24	93%	562.32	94%
000104421	49%	Jockamo Jui...	93%	10.00	92%	20.49	94%	204.9	93%
	0%	Purple Haze ...	93%	99.00	93%	20.49	93%	2028.51	94%
	0%	Strawberry 4...	93%	99.00	93%	20.49	93%	2028.51	94%
	0%	Amber - 4/6/...	92%	70.00	93%	20.49	93%	1434.3	94%

LLM/Model

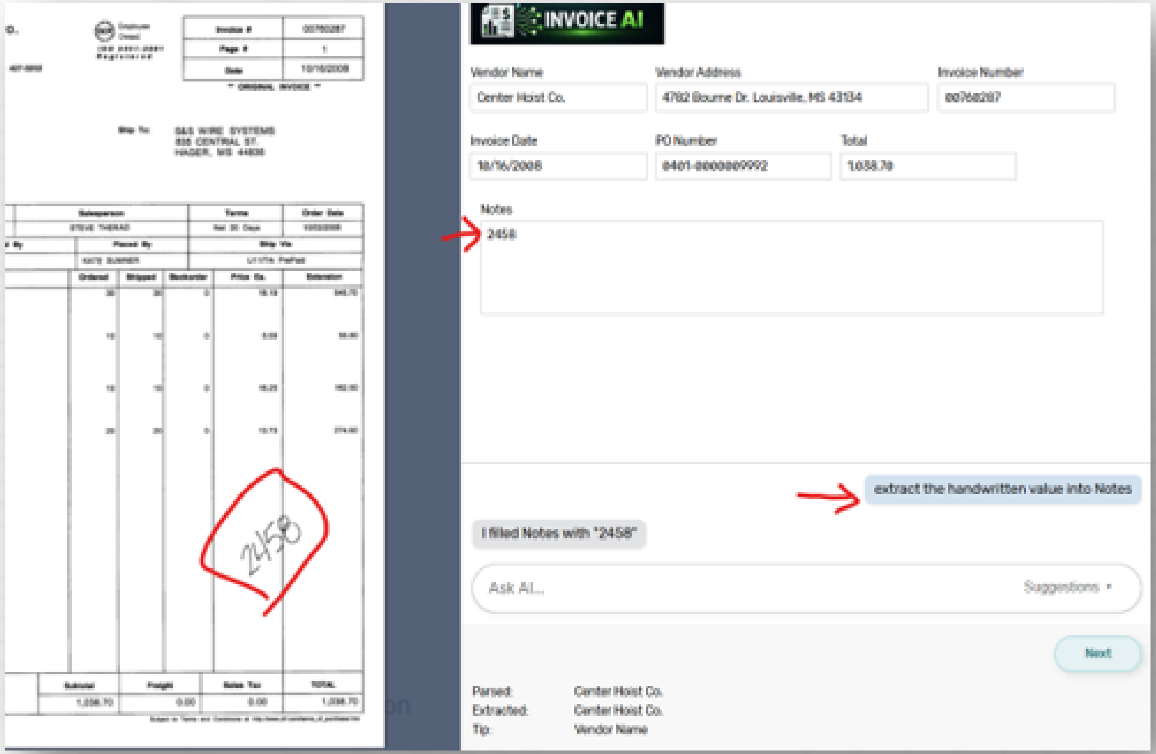
Line Items

PartNumber_A		Description_A		Quantity_A		UnitCost_A		ExtendedAmount_A	
000102125	89%	STRAWGATOR 6/4/1...	85%	18.0	90%	31.24	91%	562.32	75%
000104421	86%	Jockamo Juicy IPA 4...	79%	10.0	87%	20.49	91%	204.9	86%
000100421	88%	Purple Haze 4/6/12 ...	81%	99.0	88%	20.49	91%	2028.51	86%
000101421	88%	Strawberry 4/6/12 C...	79%	99.0	90%	20.49	91%	2028.51	88%
000100101	89%	Amber 4/6/12 Bottles	81%	70.0	91%	20.49	91%	1434.3	89%

Chat AI within IDP

Generative AI Voting

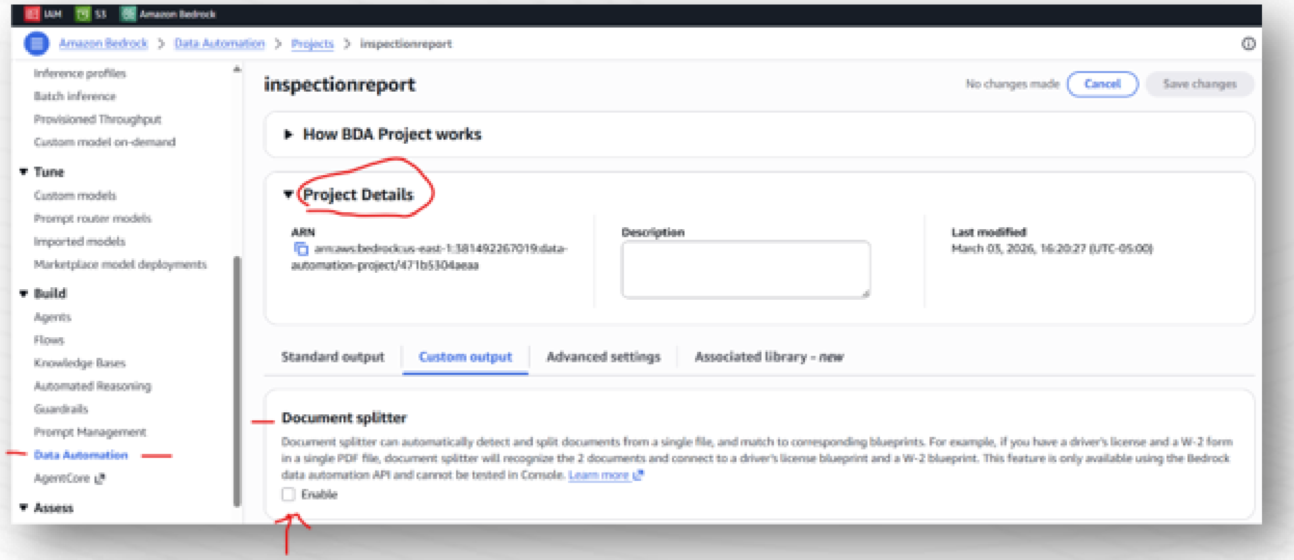
- IDP has typically always focused on extraction and classification
- BUT when AI prompting use cases start to extend
- Ask specific questions about the current document
- Take **post extraction actions** like:
  - Summarization
  - Additional document field extraction
  - Concatenation or combining fields
  - Calculations
  - Decision criteria from previous documents

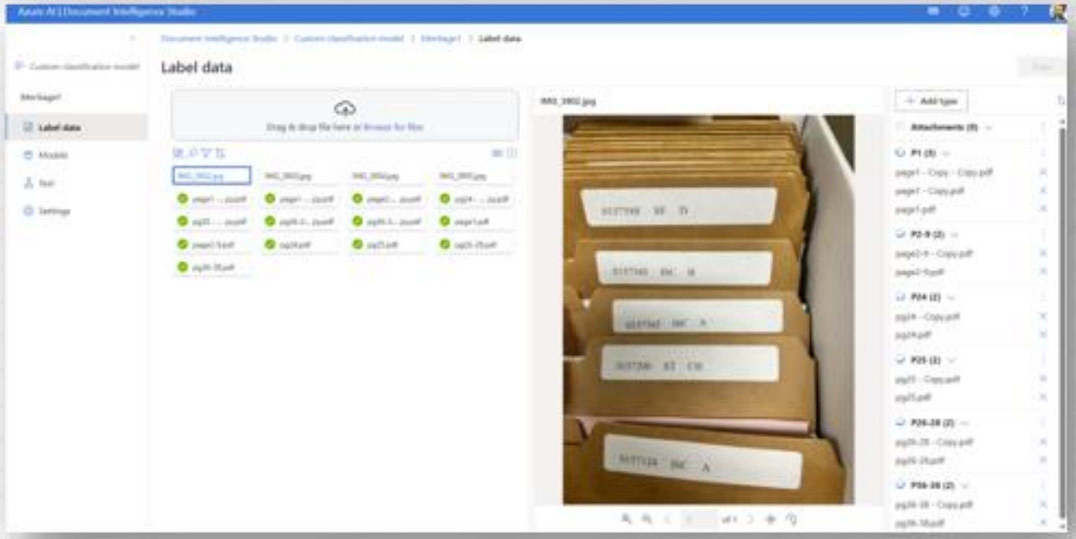


Gen-Classification

Generative AI Splitting

- **AWS BDA Document Separation**
  - AWS BDA Split feature
  - No AWS configuration – checkbox on/off
  - Enabled on the blueprint project
  - Uses same AWS BDA extraction profile parameters
- **Google DocAI**
  - Neural template splitting
  - Only works on procurement documents
  - Fails 50% of the time
- **Azure AI Document Intelligence – Custom Classifier**
  - Must create a knowledgebase of types
  - Not ideal for transactional documents where formats change
  - Few shot learning
  - Used for separation & classification

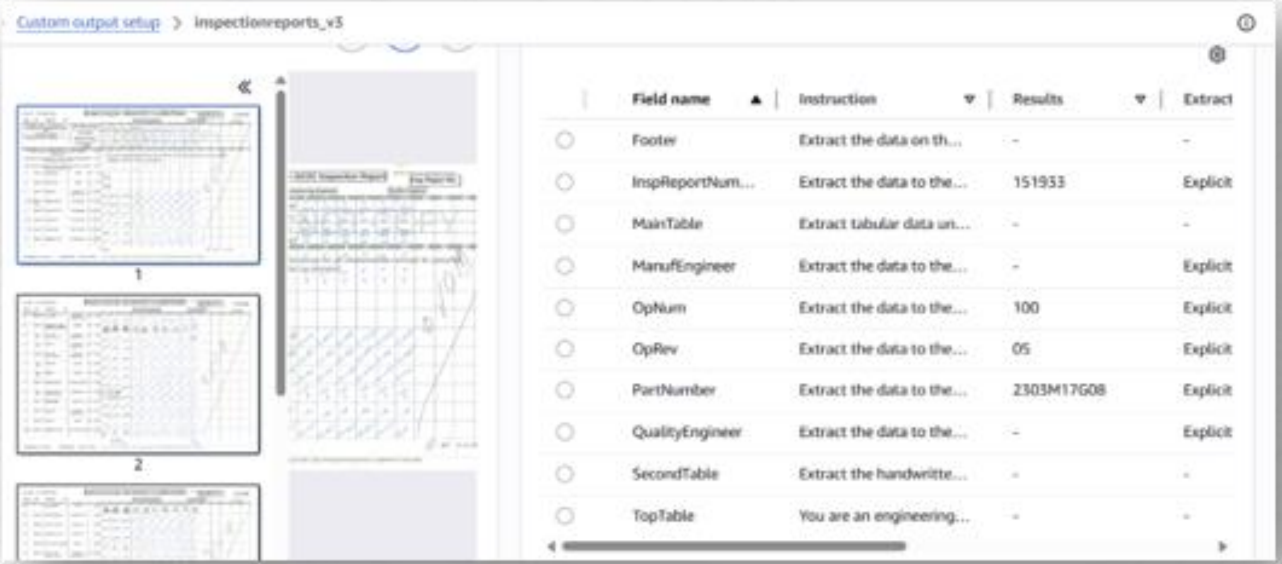




NEW ADDITIONS
USE CASES

CMMC Certified Inspection Reports

- **Customer Profile**
  - DoD Manufacturing company that stores CMMC,CUI data in our VRC hosted DocStar ECM
- **Use Case**
  - Export out all Inspection Reports from DMS, use IDP to read and extract handwritten data from multi-page inspection reports and export to SQL Server DB for reporting/compliance
- **Challenges**
  - Difficulty with misaligned tables
  - Multiple disparate tables in a single document
  - Tables that roll over the page
  - Watermarks
  - Handwriting legibility



QUESTIONS?

THANK YOU!

Brent Wesler

VP Strategic Technology &

Digital Automation

VRC Companies

Bill Holley

Sr Enterprise Solutions Manager

Kodak Alaris

AI+IM GLOBAL SUMMIT

Kodak alaris

Empower your data

WE WANT YOUR FEEDBACK

RATE THIS SESSION NOW

Open the **AI+IM Global Summit app**, find this session, and leave your rating and comments.

Open the App → Schedules → Rate this Session

AI+IM GLOBAL SUMMIT