Getting Started with Data Extraction and Al





Poll: Are you familiar with Generative AI and use it?



Proprietary and confidential information. Do not distribute

What is Generative AI?



Most people are familiar with Generative AI (or GenAI) in the form of ChatGPT or other AI chatbots.

Put simply, GenAl are machine learning models that can **generate text, images, video, or other content** from end-user queries based on data they've been trained on.

Who are the players in the space?



For the sake of today's discussion we will be covering **Azure OpenAI** and **AWS Textract** Custom Queries in our demonstrations.



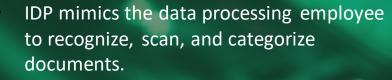
Both Cloud tools have <u>no user interface (UI)</u> to process documents in batch



Proprietary and confidential information. Do not distribute.

So, how do we consume these Cloud services?

Intelligent Document Processing.

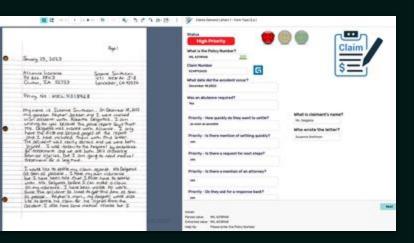


- Relevant information is extracted using cognitive AI technologies like Large Language Models (LLMs) and neural transformers
- These pinpoint exact data points within the document
- IDP organizes the extracted data and presents it in a readily accessible format for routing, review, and approval within an ERP (Enterprise Resource Planning) system or applicable line of business system.



Introducing: Kodak Alaris Info Input

- A <u>batch capture wrapper</u> around the top 5 hypervisor intelligent OCR engines
- Allows API level integration with the top 6 cloud engines:
 - o Azure Al Document Intelligence
 - o Azure OpenAi
 - o Google Document Al
 - o AWS Textract
 - o Hyperscience
 - o ABBYYVantage



What's the Difference Between GenAl and Document Al?

Gen Al Use Cases

- Use pre-existing LLM's like OpenAl, Claude, Llama, Grok
- DO NOT build models based on knowledgebase of samples and training data

Document Al Use Cases



- Must build models
- Must continue to train model based on variations in documents
- Building models takes time



What's the Difference Between GenAl and Document Al?

Gen Al Use Cases



- Unstructured documents
- Extracting data from paragraphs of text
- Speed of deployment
- No uptraining
- Prompt engineering
 - Curate a prompt that works across all documents
- Less expensive (depends on model)

Document Al Use Cases



- Transactional documents
- Key-value pair extraction
- Prebuilt use case specific models (invoices, ID's)
- Supports uptraining
- Few shot historical predictions
- More expensive



Prebuilt vs. Custom Models vs. Queries

The way you consume models impacts commercials and outcome

- Prebuilt Small language models (SLM) for popular use cases
 - \circ Returns a default set of extracted field names
 - Cannot uptrain
 - \circ Less expensive than custom 0.01/page vs.
 - No need to train knowledgebase and build models

• Custom - customer specific use case data models

- \circ Multiple document types in a batch can be recognized
- \circ \qquad Full uptraining and customization including tables
- Does not support pagination of tables
- Access to build models share and import
- \circ $\;$ Need to build new models for each change
- Queries (LLM's)
 - Azure OpenAi and AWS Textract Queries
 - Building adapters vs. autonomous
 - Great for unstructured documents where no key value pairs exist



- General prompts that can extract data across diverse document set
- Classify by asking the type of document

	Décement analysis biograms allos atorias ascass on	and the local sectors in a second				
		12	1			
	III Comple Court	p-return +	designation		× C. bert	
	tear () became a	 Process play 				
	and the same of the same	-	Q Institution			
	A.4.	Postarea	Decument DOR	Furn Parser	Lapout Parser	
	a summer	Automatical Inc.	deally an advantage of a sphere.	time from determine of an last	and entertain instant	
	Public a reserves	Tarata A	Q To a	And Practices III	Sector and Audio	
	Anno alla di Calut Passante	144				
	12	lanest 1	100412-0002000	Mark Personal	10.11.700000	
		Specified 12				
		Access status	Specialized Mental process in American			
	and a	Auto 11				
	And the second s	Analas I	Balk Submet Parael	Expense Parser	Identity Document Proofing	involue Parser
		Ingen	Extract from Balls Statements Including ranks, francactions, 40.	Established Receipt, including segurar, they amount / (p. 40)	Padd for addit of 8 decidence units to 500 republic	Drive Dr. Market
	Tel me		6	6	4	e
engle Churt De Heiner			00011100000	DELT PERSONN	10.11.10.000	00.011
to de cront de terres						
	5 (S) (Landau Dec	Day Taylor Davies	Recomment for Solitor	100 Delawer Laters
+ 373,2013,38amily,ot	5 (S) (Landing Doc Spirtue: Classifier	Pag Sig Paran	Procurement Doc Spiller	US Sher Loss
 STL2013, Shareby, H Wiley Sources 	5 (S) (Spite Center and Social Center	And fur factors and a	Manufact Income	US Once Low
			Landing Sor Spinse Charles Anti-Landier Station Station	And fur factors and a	Internet Des lutter	Ut Sher Low
• the			SHOANY	And fur factors and a	et of Accounts	Ut Shar Loan
• the			THEOREM AND	15/19/0 15/19/0 15/19/0		Ut Drive Low
• the			The Courts of th	1000 100 100 100 1000		ULDING LOG
• the			THEOREM AND			UE Dirac Los
• the			The Courts of th	1547340 1547340 4444440 4444440 4444440 4444440 444444		US Shar Los
• the			The Courts of th	1547340 1547340 4444440 4444440 4444440 4444440 444444		UE Shar Los Bhar No fao
• the			The Courts of th	1547340 1547340 4444440 4444440 4444440 4444440 444444		UE Shar Law
• the			The Courts of th	1547340 1547340 4444440 4444440 4444440 4444440 444444		US Shar Los
• the			The Courts of th	1547340 1547340 4444440 4444440 4444440 4444440 444444		She'NB for
V Day (1997) - 1997 - 1			The Courts of th	1547340 1547340 4444440 4444440 4444440 4444440 444444		Shar hit for a
V De voerse senderer des Des des			Productional and a second seco	1547340 1547340 4444440 4444440 4444440 4444440 444444	Annual A	
V Dr. Solver		an - 1960 - 1969 8 - 1960 - 1969	Productional and a second seco	1547340 1547340 4444440 4444440 4444440 4444440 444444		
		(/ mar) () m		1547340 1547340 4444440 4444440 4444440 4444440 444444	Annual A	
Vile units and				1547340 1547340 4444440 4444440 4444440 4444440 444444		
		1 / 1000 - 0 00		1547340 1547340 4444440 4444440 4444440 4444440 444444		
Vite 100 Section 100 Wei 100 Vite 100				1547340 1547340 4444440 4444440 4444440 4444440 444444		
Vite 100 Section 100 Wei 100 Vite 100		1 / 1000 - 0 00		1547940 1547940 4444440 4444440 4444440 4444440 444444		
Vite 100 Section 100 Wei 100 Vite 100		1 / 1000 - 0 00		1547940 1547940 4444440 4444440 4444440 4444440 444444		
Vite 100 Section 100 Wei 100 Vite 100		1 / 1000 - 0 00		1547940 1547940 4444440 4444440 4444440 4444440 444444		
Vite 100 Section 100 Wei 100 Vite 100		1 / 1000 - 0 00		1547940 1547940 4444440 4444440 4444440 4444440 444444		
Vite 100 Section 100 Wei 100 Vite 100				1547940 1547940 4444440 4444440 4444440 4444440 444444		
Vite 100 Section 100 Wei 100 Vite 100				1547940 1547940 4444440 4444440 4444440 4444440 444444		

Poll: What's your perception on the cost to implement GenAI in your business? Very inexpensive or expensive?



Azure OpenAI – What Does it Cost?

What is Azure OpenAl compared to Azure Al Document Intelligence?

• Azure OpenAi is the generative ai offering for unstructured document extraction which differs from Azure AI Document Intelligence in that models are not pre-built, prompt engineering is used to structure common questions that extract the correct data.

How is pricing configured?

- Pricing is based on input and output "tokens"
- A token = ~4 characters
- Assumption below is 3,248 characters per page or 812 tokens/page

How does that compare with Azure Al Document Intelligence?

• \$0.01/page vs. \$0.00061/page - 93% difference

How much does it cost to extract 1 page?

• GPT-40:

- o Input: (812\1,000,000 tokens x \$5 = \$0.00406)
- o Output: (812\{1,000,000 tokens x \$15 =
 \$0.01218)
- o Total: \$0.01624 per page

- GPT-40 mini:
 - o Input: (812\1,000,000 tokens x \$0.15 =
 \$0.00012)

 - o Total: \$0.00061 per page

AWS Textract - What Does it Cost?

AWS Textract pricing is much easier to understand

Consumption model by page

- \$0.015/page up to 1M/month
- \$0.010/page over 1M pages/month

AWS compared to Azure OpenAi

- \$0.015 vs. \$0.00061/page
- 96% difference

nalyze Document API		First million pages in a month	Over 1 million pages in a month
Queries	Per 1,000 Pages	\$15.00	\$10.00
Tables	Per 1,000 Pages	\$15.00	\$10.00
Tables + Queries	Per 1,000 Pages	\$20.00	\$15.00
Forms	Per 1,000 Pages	\$50.00	\$40.00
Forms + Queries	Per 1,000 Pages	\$55.00	\$45.00
fables + Forms	Per 1,000 Pages	\$15.00 + \$50.00	\$10.00 + \$40.00
Tables + Forms + Queries	Per 1,000 Pages	\$70.00	\$55.00
Signatures	Per 1,000 Pages	\$3.50	\$1.40
Layout	Per 1,000 Pages	\$4.00	\$3.00



How to Configure Azure OpenAl and AWS



Proprietary and confidential information. Do not distribute.

Azure OpenAI Configuration

Configuring Azure OpenAl Studio

- Using chat to test out queries
- Content filter settings
- Tokens per minute settings
- How do we configure Kodak to send and receive data from Azure OpenAi?
 - Query and key-value pair extraction at the same time
 - Modifying questions and adding new questions

Appre Al Foundry Appr	OpenAl Service pif-asurespecial	a 🗇 👘 Welcome to Apure OpenAl Service
(2)		
n mome		Welcome to Azure OpenAl Service
let darted		Explore the generative AI models and craft unique prompts for your use cases.
D Model catalog		
Tagenetic C		Resource configuration View NON
Co Churt		and a standard and a standard and a standard
Association concerns		Name Subscription Subscription
Baitne min		Aureopena C View acress control Microsoft Penner Nerwork [2] C Control (1)
2 traps		API kay 1 Revenue Group Policing for
Completions		
helt A		
II. Ene-turing		API key 2 Asses OpenAl Sedpoint Location
Distance OpenAl more		
B Batch John		Get started
Daniel monarces		Get started
C Deployments		Assistants playaround Ohat alavaround
Qi Deployments Ch. Country		Assistants playground Chat playground a playground a playground a get some t
	Image: Second	A definition of the second of the secon
	Image Image El: Prove El: Prove <	A definition of the second of the secon
	Porter Porter	A pharamagana 1 Sepanjawa 1 get A search 1 A commin de la search de l



vitalrecordscontrol.com

AWS Textract Configuration

Custom Query Adapters vs. Textract Prompts

• Compiling adapters with defined document and data sets vs. autonomous queries

AWS Bedrock vs. AWS Textract Custom Queries

- **AWS Textract Custom Queries** is specialized for extracting specific data from documents
- **AWS Bedrock** provides a broader range of AI capabilities, making it suitable for more complex natural language processing tasks.

How do we configure Kodak to send and receive data from AWS?

- Query match type
- Query prompts



B Barn Barn		100.0				
Anase feeture X	· Novi I works					
NAM Andres Texanol Andres Texanol Andres W Andres Status Andres Status Andres Status	None sease 1 deae sease 1 deae to observe at 1 deae to observe	A freedomente A freedomente Marchigen en la constanti da constanti Marchigen en la constanti da constanti da constanti Marchigen en la constanti da constanti da constanti Marchigen en la constanti da constanti da constanti da constanti Marchigen en la constanti da constanti da constanti da constanti da constanti Marchigen en la constanti da constanti Marchigen en la constanti da	L Molecturer A Molecturer Marine and a Molecturer Marine and a Molecturer Marine and Amolecturer		1. Cost articles of our factors of the factor of the factors of the factor of the factors of the	A seguera adapte a seguera adapte del constitución e que aconst a selecerador a seguera estas a selecerador a seguera estas a selecerador a seguera estas
Loss have been	for the definition	free parts	moly documents	from unique	Robust object	10107), 1010000
	Adapter Bitally on					
			tot offset at		Antonio angle anton	
Series gates Series (series (series Series)	1000		terrapea.		-	
and have been been	Control or Ania 76, 2010, 71 (10, 2010) (10, 2010)		Namber of Angles and an B			
terra, service and a	Si backet averaiser on					
Temps mare (Entered Incohor			Transformation (Contraction)		

Name								
	Pd-Anapon Textual Extraction Analysis			TextractRole				
Region	Server d			Allows AMS Textract to call other AMS services on your behalf.				
Access Key 14	And a standard to sopport							
Secretifier				Summary Creation data Nonember 13 2014 11/10 0/12 0/10 0				
API	404/155 v							
Shube	Sectors conside us and 0 Ma ⁺ /a ⁺ /a ⁺ /0.0	eter the shell her		Last activity			Reiment	
905-URL	305.0% Myschip an and Camporano.com/01402/071934			C 13 days age			Those	
SHE Topic ARM	amananeus ant 210-6020701844							
Are All a	anactive 31403773 stellar behadfule			Parminium Trust rel	ationships Tags La	Accessed Rev	oka seusions	
Peakures	 Table 							
	Farms			Permissions policies	5) ===			
	Coveries			You can attach up to 10 mana	prof policies.			
	and DCR results						liber by Type	
	Seve engine response payload		Q: Seech			All types		
				Palicy same [5]		 1 (500) 		
			~	O B Amerik	utheres .	AWS-managed		
			See		Publicies	ANT managed		
						ANT-managed		
						ANS managed		
				0 8 • treatie	and the second	AWS managed		
ne OC play name Into salty: Into traction Profile	x			ction Queries Script M	leppings Options	Q		
e Lin ne OC Reyname Into Lingy Into traction Profile Iou may add nat	R digent OCR digent OCR selection: Comins: Script: Mappings: Opti	stuction pr Mit	imum Confidence	70.0			Transform	
e Lin ne OC play name Into kulty: Into traction Profile	X Alignet OCR alignet OCR selection. Queries Script. Mappings: Opti- five language queries to be evaluated by the o	Multi-value	imum Confidence # Min. Conf	70.0 Match Type	Source Field	Target Field	Transform	
e Lin ne OC play name Into into; Into traction Profile fou may add nat Name gury_PayCate gury_Employee	X dispert DCK adeption: Comits: Script. Mappings: Opti five language queries to be excluded by the r Comprises United to the Pep Date? Dame: While to the Pep Date?	struction pr Min Multi-value	imum Confidence Min. Conf 1 70	20.0 Match Type	Source Field query_PayDate	Target Field PayOate	Transform	
e Lin ne OC Ny name Intr Linky Intr traction Public Iou may add nat Name pury_PayCate pury_Employee pury_CrossPay	X dispect OCR selection: Clarines: Societ: Mappings: Opti- fice language queries to be evaluated by the or Overy text While is the Pap Dan ² Holmen: What is the cannet Gran Pay While is the cannet Gran Pay	Multi-salue	imum Confidence Min. Conf Min. Conf 1 70 2 70	20.0 Match Type CORRE CORRE	Source Field query, PayOute query, EmployeeNa.	Target Field PayOate EmpName	Transform	
e Lin ne OC fay name les isky les traction Profile to may add nat Name pary PayCate pary Emphoyes pary Emphoto	X digget COX selection. Contrin: Script: Mappings. Opti- fice language questes to be enduated by the c Query text Tables to the Py Dank Philome. 2014 in the Complexy Stand' Name to the Complexy Stand' stat is the complex Stand' stat is the complex Stand'	Multi-salue	imum Confidence Min. Conf 1 70	20.0 Match Type	Source Field query_PayDate	Target Field PayOate	Transform	
e Lin ne OC Ny name Intr Linky Intr traction Public Iou may add nat Name pury_PayCate pury_Employee pury_CrossPay	X digget COX selection. Contrin: Script: Mappings. Opti- fice language questes to be enduated by the c Query text Tables to the Py Dank Philome. 2014 in the Complexy Stand' Name to the Complexy Stand' stat is the complex Stand' stat is the complex Stand'	Mattenion pr Min Mattenion	imum Confidence Min. Conf Min. Conf 1 70 2 70	20.0 Match Type CORRE CORRE	Source Field query, PayOute query, EmployeeNa.	Target Field PayOate EmpName	Transform	

italrecordscontrol.cor/

Discuss:

How does your organization• ROIstack rank the following when• Functionalitychoosing to invest in new• Integration withtechnology?• existing systems



Proprietary and confidential information. Do not distribute

Cost

vitalrecordscontrol.com

Document AI Classification & Extraction with Hypervisor Voting



Proprietary and confidential information. Do not distribute

Use Case #1: Financial Services

Profile of Customer

Large hedge fund, financial services company Washington DC, mortgage division

What was the pain?

- Unstructured bailee letters sent from banks and financial institutions with borrower, loan, and wire instructions
- Classify and extract data and feed to LMS (loan management)
- No ability to control the document format, sequence of documents, common field names and normalization of loan servicer names
- Match extracted lender, warehouse seller, loan wire ID and seller ID and alias with LMS

Who were we selling to?

VP of IT

			_		
	SECOND FISCAL BAN	Server Server	NUA MINISTANAN MINISTANAN MINISTANAN		
	ine on contra-				
	His surgaps are abulated to init actual of the Caldward' (no hast arego more payment of all using using the Buill to the down areast Lindow? Restored Lancad Results (spreamer between the bask and he there	Manager Lands' Landing webs is named. We spar-			
	Forward to not oppropried, the Collimant is not helpfup persons a feature for the propries of performing the annual solution of the basis is a such of the it is not sender advancement (is of an provide the time performance) and an its Anal for equipments or for the topping London's obligation annuality for any appropriate theorem and performance in addition.	rend, and object to the Papir's sale direction and one share country of two Columnal durate he promptly par			
	is adding, if the Collective excepted its particular by you, for againstite for fact is low-addedy wealths fault along for billioning. Wire: Transfer Enstructions:	promis of each particul air to be tractical			
	Hardware Aver, & Toutt AAA American An Chait In An Chait In Kartuel Lake In Kartuel Lake In Birl Scheiter III.				
	If the Collement's WET compatible produces by pro- trained should be compatible compatible definition for Collement should be compatible compatible definition for	And Terrorand & An Post Association, Soci			
	CONTRACTOR DAVID DAVID TRACTOR ADDRESS AND ADDRESS ADDRESS ADDRESS ADDRESS ADD				
Ver					And
🗮 😋 🚽	1 >>> B H N ちぐうかひ 1	👔 Gali Bahalathas - Pari Tan Da	i)		
SECOND FEIC	M. BANK - BARLER NOTICE Inc. 1. 100		Validad	-	
BARLINING BOUT DC	No. Inc. No. 10	Allocation of		from the last	- Derich Reise Dana
COMPACIAL COMPACT	where any part of strategies that the That The Theory is	ABC CHATGE.		pt And Add	Dend to Review Durine Oncom Trans Transit

Interio, A





Use Case #1: Financial Services

The Process

- When mortgages are sold between financial institutions, a bailee letter is sent to the buying bank
- New seller bailee letters are added frequently.
- Bailee letter formats may change within the same seller
- Validated extracted data is matched to the Loan Management System (LMS) - Azure SQL
- Current process is manual separation, copy and paste of Bailee Letters into LMS

		N/2 P3292228
	COND FISCAL BANK - BAILEE NOTICE	June 10, 2024
20100022 RANCLENDING GROUP		Justient: SUBRIDY, KDXXI
EASE LINDERGORDE		Note Actional: 775,000.00
CAPITAL		
rount of all eases awing the	iou ("the Collineed") has been uniqued and plodged to Second Fincel Dank (Dank by the showe named Lender ("Mortgage Lender") sciency ander a cette aut between the Dank and the Mortgage Lender.	
pose of perfecting the accust stud understeading that all	ral in to be held by you as a builts in promonism on behalf of and for the beer by introver of the Roads in such Collement, and adjust to the Road's sole diso provide from the parchase, sole or other transfer of the Collement should be approach to be a supervised by the Collebert. The Collebert should add to so hold by you.	tion and control. prompily paid to
. If the Collateral is accepte a increasionly available fram	d for purchase by you, the applicable proceeds of such purchase are to be with in using the following	e transformi to
ransfer Instruction	s:	
GEORDEA BANK & T ABA (RENOLIZIE	RUNT	
Fre Codit to	Second Fluent Weethouse Cheming Account #1000100000	
For Final Code to:	BASIC LENDENG GROUP INC REF: [Loudd: bring partnessed by you]	
ateral is NOT accepted for cheald be returned via ero	purchase by you, each Calledoral should be returned to the Book immedi might delivery, to:	intely. Such
ATTN: MOR	ANK II TECHT TOLADI INMERICARI LENDINI) NONE, SUETE SIA N. 14238	
by post, without the police we party. The Basia's security in a full amount of each process arise and be categoried without a from the Mongage Londer is variant instructions, or only	In so recent should the Collinear like Addisonal is any party role thas to be the research of the Ame. The Collinear line is not been sequed at standing- ranne in the Collinear d add its documed in short beam sequences and only spon the document particular short of the Deck and the Deck and the Deck and the function of the particular of a short dynamic and the Deck and the function state of the Deck and the Deck and the Deck and the Deck and the Deck and the Deck and the Deck and the Deck and the Deck and the Deck and the Deck and the Deck and the Deck Colling is the Deck law monitories the applicable proceeds films the role of such Colling aligns to the Deck Vice Prosident of Marigage Westhouse Operations, when	d by the Dank to scorigt by the Collement chall y requests or of the Dank to ursl. If you have
for your cooperation with th	la surter.	

Long-

Dec. 33

Nontro Wantow Parsonal Sector of

Diane i the Pack segregate

the Bask

If the Coll Collegest

And with

any other Resk of t

nach new my quoti (HIP) 295 Thank you



<u>Use Case #1: Financial Services</u>

10 01

The Challenge

- Could have single or multiple borrowers on a single bailee Appendix table of borrower is variable
- Normalizing single and multi-borrowers
- Unstructured bailee letters, paragraph of text, tables of borrower and loan information
- Parsing punctuation
- Maintaining cloud neural models

The Initial Solution

- Microsoft Azure Al Document Intelligence & Google Document Al custom models
- Kodak Alaris Info Input batch capture "wrapper"
- Info Input Classify Designer split on bailee first page
- A lot of javascript logic to compare cloud OCR engines
- Abandoned multiple engines & Google

	ECOND FISCAL BANK - BAILEE NOTICE	Rest 11, 2024	
ton BREE		Names BERGELAUM	
ten Macuneticas	utas.	the local T1000	
ing and contract			
more paired of all only only 5	uniter ("Na Californi") inc hars sergeral and philippi in Second Frend Real. In Real for the diverse aread Lindler ("Margage Lindle") arean points and more horizon the Real and the Margage Lindle.		
for the property of performing the sec it is not include autointating the se	manife as the below part and before a provintie on Solid' of and the her her same steemen of the beat is a sufficiently and addpoint for Death's safe die di provide here the partners, and on the manife of the Calencel dual to the part Londo's influence-mentality for Calencel the Calencel dual of all of a NAM'sy part.	tion and contrill. Scientific paid to	
to field is toroubledy analytic i		a machinal to	
Wax Transfer Instruction	and Control of Control		
CONTRACTOR DAVID, S ADAL ADMITTER	teat		
An Chait In	Reveal Visial Wantimeer Charley		
for fault (sub-to-	Annual Representation (Annual Press) Reads Sciences (Annual Press) REP. Connect Annual Annual Press)		
Edu Calabarria NET resigned i Calabarri david in concentri inco	le pentosi le pia, nak faktoré skutific renenit k ilu Buk kand Kongge Adheri, te		
APPE M OFFICER TERMONY	AVANE & TRENT OF DEVENT REALINGTION (LINEWIG) SOCIETAL REPORTS OF SOCIETAL REPORTS OF		
And tell-for you, althout the prior any other party. The Bend Second Read of the Mathematic of well-pro-	d) The second density hyperbolic devices of one pays when these too written ensues of the famil. "Here "obtained here of two marging or trackets research or the "Obtained dust to Annual Parish theoretimes and upon the analytics for paralam of each Obtained. The Tardition to the term of family a significant dust to a significant dust of the second second second or family a significant dust of the second second second second second or family a significant dust of the second second second second second or family a significant dust of the second seco	a ha tha finair na naniar far fha Inflament shaft	





<u>Use Case #1: Financial Services</u>

The Challenge

- Getting both Azure and Google models to "line up"
- Normalizing extracted seller name to LMS seller name, and then populating seller ID and loan wire ID
- Creating a custom UI for triaging LMS lookups

The Solution

- Microsoft Azure Al Document Intelligence custom model accuracy was much better than Google
- Kodak Alaris Info Input batch capture "wrapper"
- Info Input Classify Designer split on bailee first page
- Lots of Javascript logic to compare:
 - AlternativeSellerName create alias'
 - Ability to Save and Add alias or use once
 - Based on ABA, account and seller ID populate the WireID
 - Verify the loan number extracted matches a valid loan number in LMS

	SECOND FISCAL BANK - BAI	LEE NOTICE And 10, 2024	
	tem MACHINELINATIC	Survey BILLITALIA Bus Assessed 77(100-00	
	ine of colloc-		
	His partypy loss analysis is intrastor ("An Caldword") in her angest out poly eccerpanies" if all units only its field is the direct annal Lacker ("Antypy Los Mantimer Los and Neurity hyperson between the heat and the Mergan Losks		
	Formatti vi ne opanovati, the Collimant is strice (ARC) parts are forder to presende to the for party and predicting the average alternative field have a sould collected, and all of a set and and account of the strice of the strice of the strice of the strate field have dependent on the Collected or Allegation and the Collected of the party of these of the strice of the strice of the strice of the appropriate field on the party model is half by parts.	part in the Padd's acts direction and controls of the Colleman density in persently paid to	
	In addition, if the Collimant's excepted its particular by pay, for against the proceeds of the fraction for the former than a stage for billioning. Write Transider Instructionan:	and parties as to be the tradition in	
	Consequences Auroral & Normal Young Auroral Tourist Calo Announces Touristics Consequences Consequences for Constitute Consequences Consequences Auroral Mathematics Constitute Consequences for France Constitute Constitute Constitute Constitute Constitute Auroral Constitute Constitute Constitute Constitute Constitute Auroral Constitute Constitute Constitute Constitute Constitute Auroral Constitute Constitute Constitute Constitute Constitute Auroral Constitute Constitute Constitute Constitute Auroral Constitute Constitute Constitute Constitute Auroral Constitute Constitute Constitute Constitute Constitute Constitute Auroral Constitute Constitute Constitute Constitute Constitute Auroral Constitute Constitute Constitute Constitute Constitute Constitute Constitute Auroral Constitute Constitute Constitute Constitute Constitute Constitute Auroral Constitute Constitute Constitute Constitute Constitute Auroral Constitute Constitute Constitute Consti		
	With Collinear's NOT cought for produce by pice, such Collinear (should be one Collinear characteristic concerned to concerning to defense, for	and to be that been body look	
	ORDINAL AND A TRUNT AND A TRUNT		
-			bilationinale A 🔿 🚍
R 8 –	<) >>> =	Cust Balaxiations - Perm Type (Fp.)	
SECOND FEE	ALBANK - BALLEE NOTICE Inc. 1. 100	Validation	

SECOND FISCAL BANK - BAILEE NOTE	(a) her (0.10)			Validation	
- NEC	Rented TOTAL	Allertands at		Change from the last	
and contra-		ABC CHPTHL	202	AcceptAnelApp #	E Send To Review Ourup
energing on exclusion for entry "An interest" for the comparison statistical set and the	tel to fact a	Salar, N		Selection and an and a selection of the	Choose from the last
door is an advanta species in the active Wriggs (seeks		BADE LENDING ONCUP INC.	212.	N *	•
	the function of the second sec	Los disense onder es Balandis, M. Marcellin, M. Secondis, M. Secondi	22 22 101	Bes Die Hertgep Disjonation Bes d'Egypte Besteurn Hongspe Besteur Besteur Hongspe Besteur Besteur Land Besteur Hongspe Frienig Besteur Frienissehr Besteur Frienissehr	14230



Proprietary and confidential information. Do not distribute

Microsoft Power Platform Al Builder & Encodian



Use Case #2: Solar Energy

Profile of Customer

Employee owned, Solar Energy service and installation company with 1,000 employees

How Did We Get the Lead?

Previous client, recommended us to an accounting firm that conducted and RFP

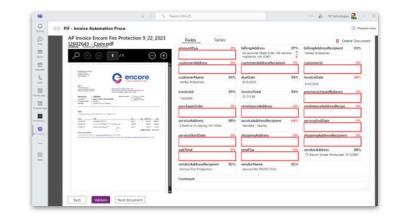
What was the pain?

- Managing field tech diagrams, work orders, notes for installation of solar.
- Was using Microsoft Sharepoint but had lots of limitations
- Support for offline file access

Who were we selling to?

VP of IT and outsourced accounting firm





Use Case #2: Solar Energy

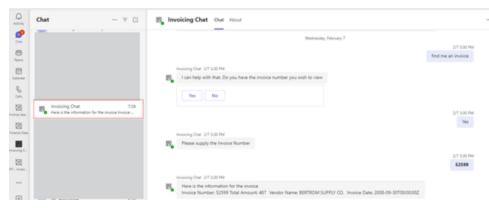
The Challenge

- Document library limits in Sharepoint
- Photo uploads from the field with offline access
- Integrationinto 3 ERP and CRM systems

The Solution

- Microsoft Power Platform Using PowerApps, Power Automate and Al builder
- Power App Base Document Kit
- Al build erint e gration for invoic e models
- Dataverse to manage all the data for Chat Ai prompting
- Encodian Filer, Trigr and Flowr to handle the document manipulation, storage, retention, searching and workflow

		C > 0, 5ee0-00-0		- 4	 Whiteships 	-
	PF - Invoice Automation Proce					(f) Present &
2	Configuration Default					
8 ie [0 [0	Model selection Fields mapping	Document automation with Al Builder model Na analysis that you're yn Ynaning mode all a builde. Ow palifele yw mode, cam bab her ard what he mode ta. Medi awediadi motor hocming Model	e jou're trede	el and		
		Models Lief	. [Most noble		
		Name Invoice Processing Model	\$13,0504	Pobult		
- 1						
0						
		Saw Not				Canoel





vitalrecordscontrol.com

Doc Al Classification & RPA for Financial Statements



Proprietary and confidential information. Do not distribute

Use Case #4: Tax & Operating Statements

Profile of Customer

• Large hedge fund, publicly traded financial services company and their Boston based property management company

What was the pain?

- Receives tax documents and operating statements (balance sheets and income statements) from each property quarterly
- Documents are manually reviewed and manually entered into their financial system, line by line
- Large volume of tax and operating statements that come from 8,100 properties, all different layouts and formats.

Who were we selling to?



Use Case #4: Tax & Operating Statements

The Challenge

- Each property operating statements are different formats
- Incoming paperwork may not have the correct documents needed to enter into financial system
- Tax documents while structured, come in as a packet and only certain pages of the tax packet need to be extracted

The Solution

- Qualify out pages through Azure classification model
- Azure Al Document Intelligence custom models to extract only those identified pages at \$0.03/page
- Azure AI custom extraction model annotating
- Alias learning based on the operating statement label
- UiPath Unattended and Attended bots for data entry
- UiPath Document Understanding cost of
 - **\$0**.25/page

Project Stats

- Number of pages per year
 - POC: 1M pages
 - PROD: 5M pages
- Number of properties
 - POC: 1,800
 - PROD: 8,300

Use Case: Classify Out Pages

 $150\,$ page documents but only $20\,$ pages need to be extracted

- 150 pages x \$0.03/page is \$4.50 OR
- 150 pages x \$0.003/page for classification at \$0.45
- 20 pages x \$0.03/page for extraction equals \$0.60 <u>RESULTS</u>
- 76% savings with this approach
- \$4.50 vs. \$1.05



Azure OpenAl Generative Al for Medical

Charts



Proprietary and confidential information. Do not distribute

Use Case #5: Vital Chart Gen Al QC Process

The Customer

• VRC performs medical scanning and coding services using manual labor

The Challenge

- Ensure that patient charts have ONLY that patient documents, no other patients
- Provider documentation is completely unstructured
- Building models would take too long and require lots of PHI samples

The Solution

- Query prompt engineering
- Kodak Info Input capture wrapper
- V Azure OpenAl

Why Azure OpenAl Over AWS?

• Azure did a betterjob identifying Dates of Service in the context of NLP

HIGHMARK. 🖗	Important Information Regarding Medical Record Request
11/8/2023	
To:	
Carepoint Health Medical Group	
316 MONASTERY PL STE A	
UNION CITY, NJ 07087-4464	
Phone: 12018673535	
Fax: 12018672433	
	E C
RE: Request for Medicare Advantage (MA) Medical Reco	rds from January 1, 2022 to Present
Dear Provider:	
As a Medicare Advantage (MA) Organization, Highmark is required	
Medicaid Services (CMS) for data submission and coding accuracy	y, we are reviewing medical records to ensure that we

Live Demonstrations



Proprietary and confidential information. Do not distribute.