



## DOCUMENT INGESTION/EXTRACTION PoC – MLC/GS/SC

January 19, 2021



## Agenda

---

- Executive Summary
- Extraction Results
- Architectural Considerations
- Financials
- Risks/Considerations
- Next Steps

## EXECUTIVE SUMMARY

---

- Extraction Results
  - Broker Specs
    - Groundspeed outperformed Capgemini on accuracy scoring 83% to 44% respectively.
    - Capgemini hand picked the samples to work on where Groundspeed delivered data on all submissions.
    - Groundspeed seems to be more in tune with our vision of a touchless model and pricing is competitive.  
**Recommend move forward and pursue contracting and technical next steps.**
  - Midterm Service
    - Both vendors performed fairly similar with their test results for determining intent (85% CapGemini vs. 84% Cognizant).
    - CapGemini's results deteriorated as respects to field level extraction, while Cognizants remained flat or showed small improvements.
    - CapGemini, while we did pause their POC to concentrate on Broker Specs, they were given opportunity to take additional time, although were comfortable with their results.  
**Recommend move forward and pursue contracting and technical next steps.**
- Architectural/Technical
  - Cognizant utilized Indico for Extraction providing a mature, targeted solution with very robust feedback and re-training capabilities already built into solution
  - Groundspeed was accurately delivering too much data, need to parse it down for consumption
  - Cap Gemini did not seem to utilize its best technology for POC
- Financial
  - Broker Specs have a break even ongoing tool cost, 100% resource reduction and faster TAT
  - Endorsements require further information/discussion with vendor in order to conclude

## EXTRACTION RESULTS



## BROKER SPECS SUMMARY

Scope	
Vendors	Capgemini & Groundspeed
Business Units	MLC & Global Specialty Domestic
Size (samples)	350 training; 50 test
Documents	Emails, Broker Specs (Word/PDF), Excel, and Acord forms (mixed)

Final internal review of results still in progress (highly manual process). Validated these metrics from the completed reviews so far.

### Capgemini

- Very disappointed with test sample results – only extracted data from 60% of the samples shared.
- Hand picked data elements they wanted to deliver to us rather than the full scope.
- Poor results with the test samples – 44% accuracy.
- Did not use OCR technology on the POC test.
- Question their automated technology/process - encountered a huge swing in result metrics comparing the training results vs. test results.

### Groundspeed

- Successfully extracted data from all 50 samples shared.
- Good results on the test samples – 83% accuracy.
- Out of the box technology works – refinement work necessary to fully meet our rating needs.
- Better understanding of the data we need for rating.
- Partnership and working relationship already in place (loss runs).

## EXTRACTION RESULTS– BROKER SPECS

Samples Delivered		Quality per Unit (file)*		Quality per Opportunity*	
Capgemini	Groundspeed	Capgemini	Groundspeed	Capgemini	Groundspeed
60% (30 samples)	100% (50 samples)	69%	83%	41%	99%

DATA ATTRIBUTE	Extraction Accuracy %	
	Capgemini*	Groundspeed
<b>Account Information</b>		
Agency	100%	80-85%
Insured	100%	95-100%
<b>Policy Info</b>		
Effective Date	0%	95-100%
Lines of Business	0%	90-95%
Quote need by date	0%	90-95%
<b>Limits/Deductibles</b>		
WC	100%	95-100%
GL	0%	80-85%
Auto	0%	75-80%
MLC v. Global Specialty	0%	60-65%
• Different LOBs but the data was extracted for both	0%	
<b>Rating Exposures</b>		
• Did not see major differences in quality.		
• WC - payroll	79%	80-85%
• More refinement work on the Global Specialty side.		
GL - sales/area/payroll	0%	75-80%
Auto - vehicle schedule	100%	75-80%
Property - SOV	96%	80-85%
<b>Overall Average</b>	<b>44%</b>	<b>80-85%</b>

\*Vendor metrics

## BROKER SPECS– NEXT STEPS

---

- Refinement Work - additional work will be required for rules/requirements to meet HIG needs
  - Critical for the Global Specialty side due to the Specialty lines of business (Financial lines, Professional, Contractors Equip, Cargo)
  - Hierarchy of documents/data – needed when multiple sources of data are included within a single submission
  - Testing specific to Large Property and LMT needs
- Fully develop future vision for all MLC business segments – PolicyCenter, CUE & NADB
  - PolicyCenter vision developed
  - NADB needs work
- Future vision for Global Specialty
  - How is the data going to be ingested?
  - Does Global Specialty have systems that can ingest rating data?
- System integration – ensure that the extraction data can be easily ingestible into our rating platforms (PC, CUE/CLA)
- Engage procurement for contracting/negotiating

## MIDTERM SERVICE SUMMARY

Scope	
Vendors	Capgemini & COGNIZANT
Business Units	MLC ( <i>Genl Ind/Tech/Marine/Constr</i> ) & SC
Size (samples)	350 training; 50 test
Documents	Emails and Attachments

### Capgemini

- Paused 3 weeks short to allow focus on Broker Specs; offered to provide them additional time to then take back up with their training and that offer was not accepted; felt they were comfortable with moving forward with testing.
- Test results were less than their final training results. Final overall: Intent=; Extract=43%
- Current state Intent is determined ~70% of the time vs. test results at %.
- Did not use OCR technology during POC.
- Of the 50 Test cases, overall % of documents with data extracted was 91%.
- Straight through processing results not shared throughout POC, did need to request.
- Expectation communicated: 1 hour TAT, 95% Quality was communicated; Vendor proposed: 24 hour TAT, 95% Quality

### Cognizant

- Test results improved over final training results. Final overall: Intent=84%; Extract=72%
- Current state Intent is determined ~70% of the time vs. test results at 84%.
- Validation of test results were completed and were consistent.
- Of the 50 Test cases, overall % of documents with data extracted was 84%.
- Throughout POC straight through processing results were shared.
- Expectation communicated: 1 hour TAT, 95% Quality was communicated; Vendor proposed:



## MIDTERM SERVICE INTENT & EXTRACTION COMPARISON

Intent					
Cognizant	Sample Size	%	Capgemini	Sample Size	%
Vehicle Replace	50	92%	Vehicle Replace	50	92%
Vehicle Add	50	90%	Vehicle Add	50	86%
Location	50	82%	Location	50	82%
Quote	50	88%	Quote	50	90%
Endorsement	50	70%	Endorsement	50	74%

Extraction					
Cognizant	Sample Size	%	Capgemini	Sample Size	%
Vehicle Replace	50	62%	Vehicle Replace	50	33%
Vehicle Add	50	78%	Vehicle Add	50	34%
Location	50	70%	Location	50	39%
Quote	50	83%	Quote	50	48%
Endorsement	50	69%	Endorsement	50	58%

There was found to be no nuances between SC and MLC data. Sample may have contained a larger % of SC vs. MLC, although this would be a fair depiction of volumes.

## MIDTERM SERVICE STRAIGHT THROUGH PROCESSING COMPARISON

Straight Through/Zero Touch					
Cognizant	Sample Size	%	Cap Gemini	Sample Size	%
Vehicle Replace	50	20%	Vehicle Replace	50	10%
Vehicle Add	50	58%	Vehicle Add	50	10%
Location	50	27%	Location	50	22%
Quote	50	48%	Quote	50	20%
Endorsement	50	20%	Endorsement	50	6%

**Straight through processing is the identification and extraction of all data elements for the intent pushed through to completion of the service request.**

### COGNIZANT

Day 1 would achieve 10% straight through and at steady state increase to %.

### CAPGEMINI

Day 1 would achieve 65% straight through and at steady state increase to 80%

## Technical Vendor Considerations

---

- Endorsements
  - Cognizant utilized a vendor, Indico for extraction which provided a mature, targeted solution
  - Cognizant comes out of the box with a very robust feedback loop for re-training capabilities already built into solution
  - Cap Gemini was reusing an existing model for extraction that for elements it was trained on did well, but Cap Gemini did not work to improve for new elements
  
- Broker Specs
  - Groundspeed technology just much better suited to the technical challenge of unstructured documents
  - Groundspeed was accurately delivering too much data in situations, need to parse it down for consumption
  - Cap Gemini did not seem to utilize its best technology for POC



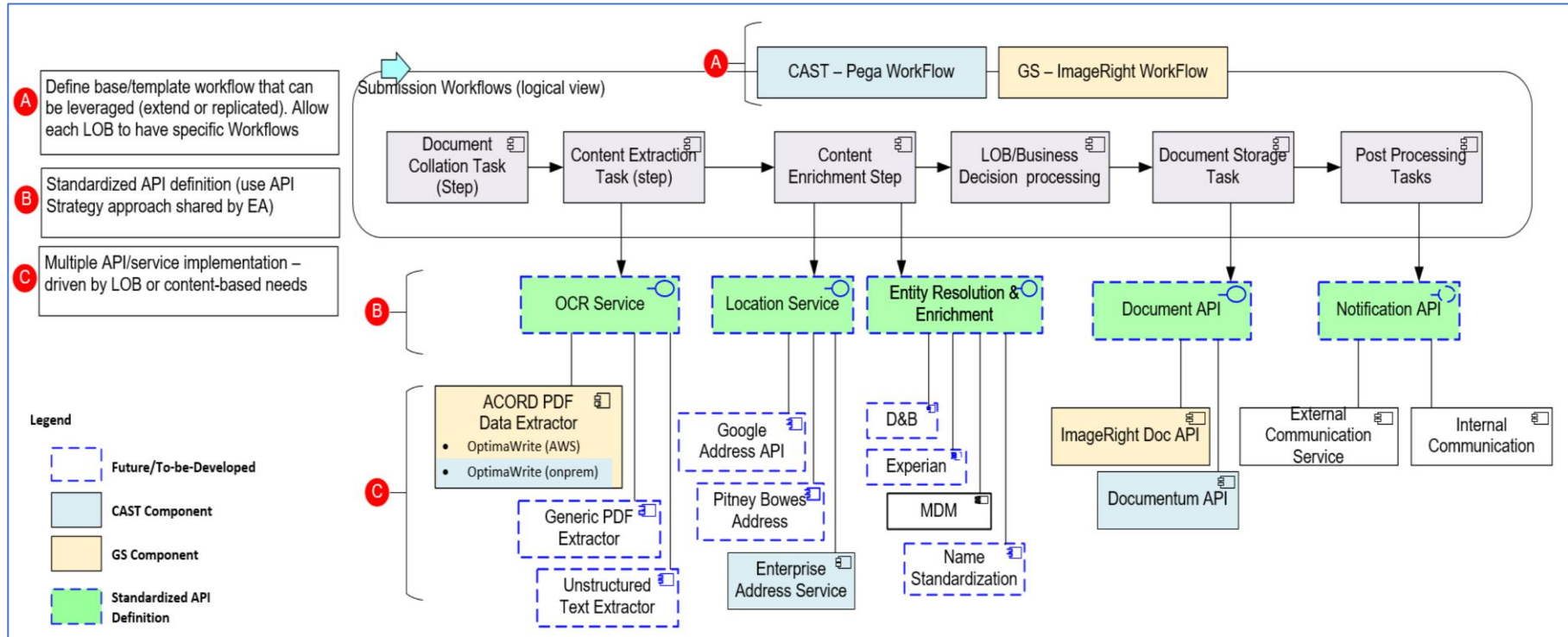
## ARCHITECTURAL CONSIDERATIONS



## VENDOR TECHNICAL ANALYSIS

Criteria/Vendor	Capgemini	Cognizant	GroundSpeed
Integration Architecture	✗ No integration pattern shared – plan was to monitor HIG mailbox	✓ Service based integration	✓ Service based integration
Availability	✗ Not applicable	✓ High proposed	✓ High proposed
Flexibility wrt doc type/size	✗ Not applicable	- TBD	✓ Flexible per business needs

# INTENTIONAL ARCHITECTURE– SUBMISSION INTAKE WORKFLOW



## TECHNICAL SUMMARY

---

### GS and MLC Should Focus on Creating Services that are Reusable Across Each Line by Decoupling LOB Specific Dependencies at the API Level

- **Maximize Business Opportunity by Maximizing Re-useability & Owning Market-Differentiating Capabilities**
  - *OCR and Ingestion*
  - *3rd Party Lookup*
  - *Document Management*
  - *and Others*
- **Workflow Capabilities and Integrations are Business Specific and Should not be Rationalized**
- **Invest in Market Differentiating Intellectual Properties to Drive Market Standing**
  - Look at OCR/Automation as strategic capabilities and thus invest in HIG assets and less lock-in with vendor offerings.

## FINANCIALS





## BROKER SPECS/SOVS FINANCIALS

	Target State	Current State	Groundspeed
Ongoing MLC Resource Costs	\$0	10 BPO FTEs (\$250K)	\$0
Ongoing Tool Costs	\$12 per submission or less	<b>\$235K</b> (\$12 per submission charge)	\$12 per submission
SLA	1 hour	24 hours	TBD
Vendor Implementation Charge			TBD
Invest Cost	TBD*		
Maintenance Cost	TBD*		

\* HIG IT costs to be determined, not previously ROM'D, was within Account Setup scope originally.

### Global Specialty Notes

- Roughly 50% of GS Domestic submissions come in on broker specs
- Opportunity is 2-3 FTE in OSC and a notable UW capacity gain – dependent on the specific scope of implementation

## MIDTERM SERVICE FINANCIALS

**Further detail is required of Cognizant's proposal before conclusions can be drawn**

	Target State	Current State	Cognizant
Automation Assumptions	70% intent 48% processing	Zero	TBD
Resource Cost per Endt	Manual \$50 Supported \$11	Manual \$58 Supported \$24	TBD
Tool Cost per Endt.	Neutral to Smart e-mail	Smart e-mail - requested	.42 cents per e-mail
Vendor Implementation Charge	N/A	N/A	\$327K
Invest Cost (Policy Center only)	\$1,920K	N/A	N/A
Maintenance Cost (Policy Center only)	\$340K	\$192K	N/A

### Global Specialty Notes

- GS Domestic has 8 FTE processing 33k endorsements per year
- 50% of endorsements are similar to MLC given the LOBs (Property, GL, Auto)

## RISKS & CONSIDERATIONS

---

- Risks

- Glidepaths on results aren't achieved, ensure recourse is included in contracts
- Moving broker specs from Cognizant could jeopardize this ongoing relationship
- Black box solution, any change will require customization and costs

- Consideration

- Primary Vendor Choice considerations were Extraction Results, Ease of Integration, Pricing
- Cognizant we can take advantage of platform already used for intake and potential replacement of Smart E-mail is already in their EDM statement of work.
- Groundspeed we can take advantage of platform already in use for loss run extraction.
- Potential for all submissions (structured & unstructured/Acord & Broker Specs) to be extracted by Groundspeed – cost analysis required. This would simplify the technical environment with one vendor for new business intake.

## NEXT STEPS

---

- Begin procurement process with Groundspeed
- Drive to further detail and understanding within Cognizant's proposal
- Provide Capgemini with feedback and their disposition with PoC
- Determine Small Commercial Midterm Service impact
- Develop architectural design for Global Specialty
- Develop ROMs for Broker Specs/SOVs and Global Specialty Midterm Service

## APPENDIX



## SCOPE

NLP & Data Extraction capabilities required to read e-mails for intent, extract data from multiple unstructured sources, structure for ingestion within our systems to enable straight through processing.

### Broker Spec/SOV

- Vendors: Capgemini, Groundspeed
- Business Units: MLC, GS
- Transactions Tested: New business submission e-mails including Broker Spec word docs, PDFs, Excel workbooks and Statement of Values
- Size: 350 training samples. 50 test samples.

### Endorsement

- Vendors: Capgemini, Cognizant
- Business Units: MLC, SC
- Transactions Tested: Vehicle replace, Vehicle add, Location, Quote, Endorsement
- Size: 350 training samples for each of the 5 transactions. 50 test samples for each of the 5 transactions.

## Broker Specs – Training Results

Document Type		
PDF	Excel	Other
47%	21%	32%

DATA ATTRIBUTE	Extraction %		Extraction Accuracy %	
	Capgemini Capture	Groundspeed	Capgemini Precision	Groundspeed
	183 samples	32 samples*		99.06% overall
<b>Account Information</b>				
Agency	99.8%		100%	100%
Insured	91.3%		100%	100%
<b>Policy Info</b>				
Effective Date	-		-	99%
Lines of Business	-		-	98.2%
Quote need by date	-		-	100%
<b>Limits/Deductibles</b>				
WC	-		-	100%
GL	-		-	100%
Auto	-		-	100%
Property	-		-	100%
<b>Rating Exposures</b>				
WC - payroll	95.3%		59.3%	100%
GL - sales/area/payroll	100%		11%	100%
Auto - vehicle schedule	95.8%		95.7%	100%
Property - SOV	67%		66.7%	99.7%

\*Groundspeed only performed a focused audit – 10% of the total files delivered.

## ENDORSEMENT INTENT COMPARISON – TRAINING PROGRESS

Cognizant	Intent							
	Initial Run		As of 11/19/2020		As of 12/8/2020		As of 12/22/2020	
	Training Sample Size	Intent %	Training Sample Size	Intent %	Training Sample Size	Intent %	Testing Sample Size	Intent %
Vehicle Replace	50	88%	50	92%↑	50	86%↓	50	92%↑
Vehicle Add	50	84%	50	86%↑	50	92%↑	50	90%↓
Location	50	78%	51	82%↑	51	82%	51	82%
Quote	50	76%	50	69%↓	50	76%↑	50	88%↑
Endorsement	50	42%	45	81%↑	45	69%↓	45	70%↑

Capgemini				
Vehicle Replace			350	100%
Vehicle Add			350	100%
Location	350	66%	350	87%↑
Quote	350	29%	350	99%↑
Endorsement	350	95%	350	94%↓



## ENDORSEMENT EXTRACTION COMPARISON – TRAINING PROGRESS

Cognizant	Extraction							
	Initial Run (Cap)		As of 11/19/2020 (Initial Run – COG)		As of 12/8/2020		As of 12/22/2020	
	Training Sample Size	Capture Accuracy	Training Sample Size	Capture Accuracy	Training Sample Size	Capture Accuracy	Testing Sample Size	Capture Accuracy
Vehicle Replace			50	58%	45	60%↑	45	62%↑
Vehicle Add			50		46	74%	46	78%↑
Location			50		33	58%	33	70%↑
Quote			50		49	71%	49	83%↑
Endorsement			50		33	64%	33	69%↑

Capgemini			
Vehicle Add	350	46%	350 40%↓
Vehicle Replace	350	30%	350 58%↑
Location	350	41%	350 71%↑
Quote	350	42%	350 50%↑
Endorsement	350	46%	350 59%↑

# COGNIZANT INTENT DETERMINATION



Intelligent Process A

Intelligent Process A

Intelligent Process A

Viewing Premraj Maniraj (in...

Intelligent Process A

Intelligent Process A

← → ↺

poc.hig-indico.ciia.cognizantproducts.com/explain/datasets/36/models/29/performance

🔍 ☆ 🌐 ⚙️ 👤

indico

<

DATASETS

DISCOVER

EXPLAIN

TEACH

REVIEW

ACCOUNT

API DOCUMENTATION

HELP

Model Group ID: 29

Selected Model ID: 101 | Model Training Status: Complete | about 1 hour total

PERFORMANCE

CLASS BALANCE

TRAINING

Try It Out

Labeled Text

From: Sherry Sloan

Sent: 2018-04-18 01:41:30+00:00

To: agency.services@thehartford.com

Cc:

Subject: ENDORSEMENT Quote Request 22SBABQ6779 Carolinas Oral & Facial Surgery Center

Please provide me with a QUOTE to remove the building coverage for location 1(705Jefferson Street) and location 3 (709Jefferson Street). The BPP and Liability should remain.

Thank you,

Sherry Sloan, CISR

DHN&T Insurance Agency

3604 Shannon Rd Ste 105

Durham, NC 27707

919-286-7473

Location

Quote

Endorsement

ReplaceVehicle

AddVehicle

Edit Text

Metrics

# COGNIZANT DATA FIELDS

Intelligent Process A

Intelligent Process A

Intelligent Process A

Viewing Premraj Maniraj (in...

Intelligent Process A

Intelligent Process A

← → ↺

poc.hig-indico.ciia.cognizantproducts.com/explain/datasets/27/models/25/performance

🔍 ☆ 🌐 ⚙️ 👤 ⋮

indico

DATASETS

DISCOVER

EXPLAIN

TEACH

REVIEW

ACCOUNT

API DOCUMENTATION

HELP

Try It Out

Labeled Text

From: Branch, Jeff (MMA)

Sent: 2018-04-16 19:49:53+00:00

To: agency.service@thehartford.com

Cc: Vela, Silvia (MMA)

Subject: Add Vehicle Request: EyeCare Partners, LLC #72UUNHB4923

Good Morning,

Effective March 14th, 2018, please add the following vehicle to the policy:

#	Entity	Garage	Zip Code	Year	Make	Model	VIN	Cost New	Radius of	Added/Deleted	Operations Date
AVC	35213	2017	Hyundai	Santa Fe	KM8SM4HF2HU176612	\$19 31,000	<50 Miles				

Please let me know if you need any further information to process.

Thank you!

Jeff Branch, CPCU, RPLU, MLIS, CRIS, ARM -E, ARE  
Client Service Executive | Property + Casualty Division  
Marsh & McLennan Insurance Agency LLC  
1 Polaris Way | Suite 300 | Aliso Viejo, CA 92656  
+1 949 900 2253  
Jeff.Branch@MarshMMA.com | MMA-West.com

property & casualty | employee benefits | compensation consulting | individual, family + students | industry expertise | risk management center | mma marketlink | map | events | vCard

PolicyNumber

EffectiveDate

VIN

Make

Model

Year

VehicleType

GaragingZip

CostNew

ComprehensiveDeductible

CollisionDeductible

LossPayee

Lessor

USE

ClassCode

# COGNIZANT TOOL TRAINING



Intelligent Process A

Intelligent Process A

Intelligent Process A

Viewing Premraj Maniraj (in...

Intelligent Process A

Intelligent Process A

← → ↺ 🔒 poc.hig-indico.ciia.cognizantproducts.com/review/queues/15 🔍 ☆ 🗄️ 👤 ⋮

Review Queue: ReplaceVehicleDataExtraction\_ReplaceVehicleDataExtractionTask

Keyboard Shortcuts

PREVIEW

Page 1 of 1

173-New Style Cabinets Inc (628 KB).pdf

🖱️ 🖋️ 🗑️

R Reject

N Next

Open All

Hide All

1

PolicyNumber

⬆

.83UECZH9408

98% 🔍 ✓

+ Add Value

2

EffectiveDate

⬆

April 13, 2020

99% 🔍 ✓

+ Add Value

3

CurrentVINNumber

⬆

5 PVN E

66% 🔍 ✓

F4S53739

70% 🔍 ✓

+ Add Value

4

NewVINNumber

⬆

WD2PD644345649849

From: Service

Sent: 2020-04-14 02:45:58+00:00

To: 'agency.services@thehartford.com'

Cc:

Subject: New Style Cabinets, Inc.

New Style Cabinets, Inc.

Policy NO: 83UECZH9408

Effective: April 13, 2020

Please REMOVE the following Vehicle from the above captioned policy:

2004 Dodge **Exempt**

VIN# 90U727644565649849

Please ADD the following Vehicle to the above captioned policy:

2015 Hino 268

VIN# 5EVNFB1V0E4553739

Thank You,

Nikki

Bulow Group

Insurance - Risk Management - Consulting

in 📧 📷 📺

2004

2004

2004

Client Service Team

p: (708) 258-5448 f: (708) 377-4178

18521 Spring Creek Drive, Tinley Park, Illinois 60477

Coverage cannot be bound, altered or cancelled without speaking to or receiving written confirmation from a licensed agent. The information contained in this electronic mail is privileged and confidential information intended only for the individual of entity named above. If the person reading this message is not the intended recipient, an agent or employee responsible to deliver it to the intended recipient, you are hereby notified that any dissemination, distribution or copying of this communication is strictly prohibited. If you have received this communication in error, please notify us immediately by telephone.

© 2021 by The Hartford. Classification: Company Confidential. No part of this document may be reproduced, published or used without the permission of The Hartford.

28

## BUSINESS CASE OPPORTUNITY

---

- Here the Business Opportunity rests with the potential value in creating re-useable services not dedicated solutions (opportunity cost)
- Re-useability requires removing business awareness where services operate in a decoupled, stateless (or even serverless) manner
- Instead, services should be broken down into simpler functions (to support re-use. i.e. "less is more"), for example:
  - Document extraction cannot be one product like "optima-write" or "SavidgeScan" hoping it can do it "all" and if not look elsewhere!
  - Instead consider all products as individual services focusing on what they do best, for example:
    - SavidgeScan supports PDF extraction when presented with a template (editable PDF)
    - Optima-write (should focus on ACORD extraction – as it was built for)
    - Unstructured Text – TBD should focus on leveraging ML/NLP