

CalSAWS | CalSAWS Imaging Committee



May 18, 2023

Agenda

- ① OCR Updates
- ② Performance Test Updates



OCR Updates

OCR Updates

CA-261407 OCR Corrections for No-Split Classification and SSA

We made the following adjustments to OCR processing, targeting the main reasons why documents route to Exception. These only apply to No-Split scenarios (OCRSplitOverride is set to True).

Prior Behavior

- A barcode on the second page of a document was not read.

New Behavior

- If the barcode is read on the second page of a document, the barcode info will be used to index the document.
Note: Documents with barcodes entered in QA & Indexing will continue to skip OCR processing. All information pulled for the barcode will be used for indexing.

Prior Behavior

- Documents with different form numbers on multiple pages would fail OCR and end up in Exception.

New Behavior

- If documents with different form numbers are not split in QA & Indexing, OCR will use the form number on the first page.

Prior Behavior

- SSA Award Letter classifications rates were low.

New Behavior

- Additional OCR training was completed of the SSA Award Letter to classify it accurately more frequently.



Performance Test Updates

Hyland Performance Testing

February 14, 2023 Status

- Hyland Performed 58 county load tests to help prevent additional hiccups going into the Wave 2 Go-Live and succeeding Go-Lives.
- The reason for testing was due to the External Message Agent backlog encountered a week after the Wave 1 Go-Live.
- The tests were centered around known bottlenecks we've encountered to date.
 - External Message Agent (EMA) (Solution Configuration)
 - Inserver Workers (Solution Configuration)
 - Workflow Agent Workers (Solution Configuration)
- Based on the testing, we identified an area that needed solution configuration tuning. Below is what we put in Production 2/17 before Wave 2, which help prevent a backlog and improve throughput timings for all counties (Current State and Future State).
 - Splitting out EMA Messages
 - iScriptBatch
 - iScriptGenerateTask
 - iScriptPersonUpdate
 - iScriptPreOCR
 - iScriptPrepforScan
 - iScriptRCBatchWF
 - iScriptRoutingAllGenerateTask
 - iScriptUpdateOnArchive
- Effort is tracked under CA-256388 - External Agency - EMA Load Balance Configurations for 58 County.
- Previous Average Message Times: 30-50 Seconds
- Current Average Message Times: 4-20 Seconds
- 10-46 seconds document processing improvements per EMA call which overall helps end users process and retrieve documents

Load on Solution by Wave

Wave 1, 2, 3 Number of Users & New Documents Captured

- Number users added and the average new documents captured per month

Wave	Users	New Docs Captured (Monthly)
1 (Placer, Yolo)	616	88,205
2 (Contra Costa, Santa Clara, Tulare)	4,365	429,000
3 (Orange, Santa Barbara, Ventura)	5,267	385,749

Production Metrics

Response Times Pre vs. Post Wave 2 Go-Live

2/21/2023 (Pre Wave 2 Go-Live)	Count	Avg
Capture - Create Capture Group	4,731	0.440
Capture - Add Page	38,378	0.865
Document - Create Document	8,816	0.139
Document - Add Page	9,326	0.419
Document - Get Page	18,526	0.597
Document - Get Thumbnail	82,718	0.903
Document - Rendition	124,093	0.732
View - Run View	60,922	0.208
View - Run View with Filter	5,417	0.282

3/7/2023 (Post Wave 2 Go-Live)	Count	Avg
Capture - Create Capture Group	6,047	0.402
Capture - Add Page	51,285	0.838
Document - Create Document	11,889	0.125
Document - Add Page	11,051	0.407
Document - Get Page	19,715	0.513
Document - Get Thumbnail	118,413	0.814
Document - Rendition	151,153	0.712
View - Run View	77,306	0.187
View - Run View with Filter	6,238	0.301

Production Statistics Post Go-Live	Count	Avg Delta
Capture - Create Capture Group	1,316	-0.038
Capture - Add Page	12,907	-0.027
Document - Create Document	3,073	-0.013
Document - Add Page	1,725	-0.012
Document - Get Page	1,189	-0.084
Document - Get Thumbnail	35,695	-0.089
Document - Rendition	27,060	-0.020
View - Run View	16,384	-0.021
View - Run View with Filter	821	0.019

Enhancing Your Imaging Experience

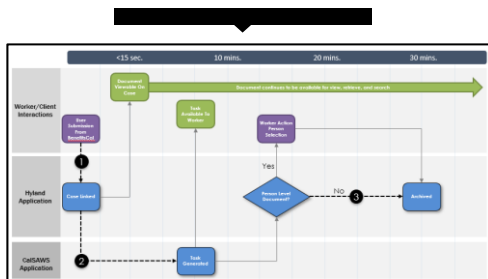
Imaging Performance Workflow Diagrams (April 12, 2023)

WORKER / KIOSK SUBMISSION

PORTAL SUBMISSION

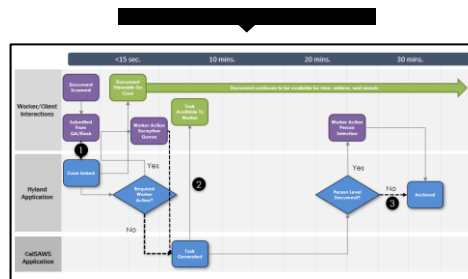
Documents submitted via BenefitsCal

- Documents linked to case **within 15 seconds**
- Images viewable **near real time**
- Task generated **on average 6.5 minutes.**



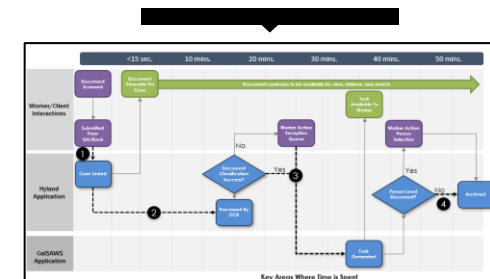
OCR Bypass & Barcode Workflow: documents that have OCR bypass enabled or are barcoded

- Documents linked to case **within 15 seconds**
- Images viewable **near real time**
- Task generated **on average 6 minutes.**



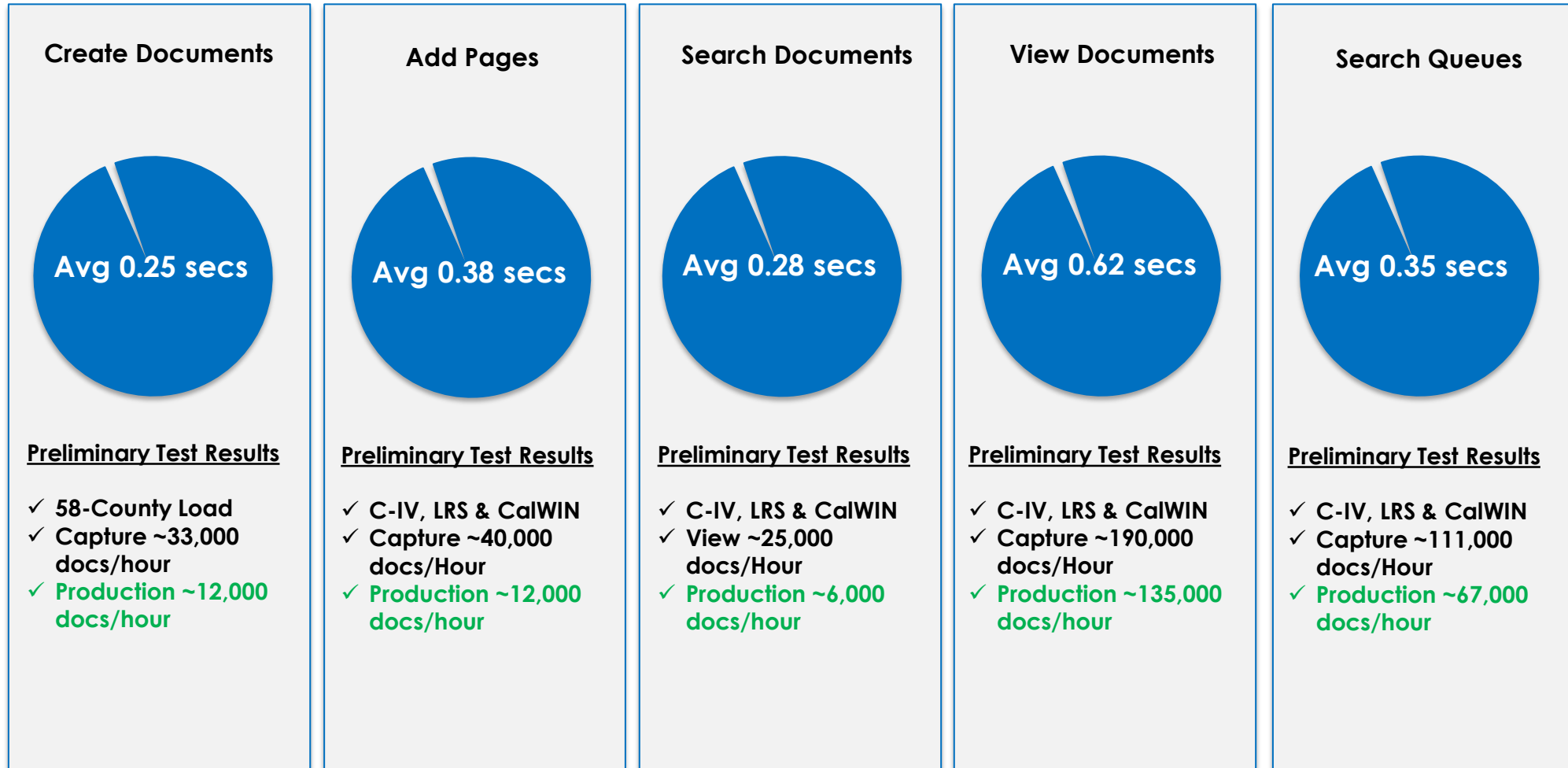
OCR Workflow: documents submitted requiring additional classification via OCR

- Documents linked to case **within 15 seconds**
- Images viewable **near real time**
- Task generated **on average 36 minutes.**



Performance Testing Overall Results

58-County Performance Test (April 12, 2023)

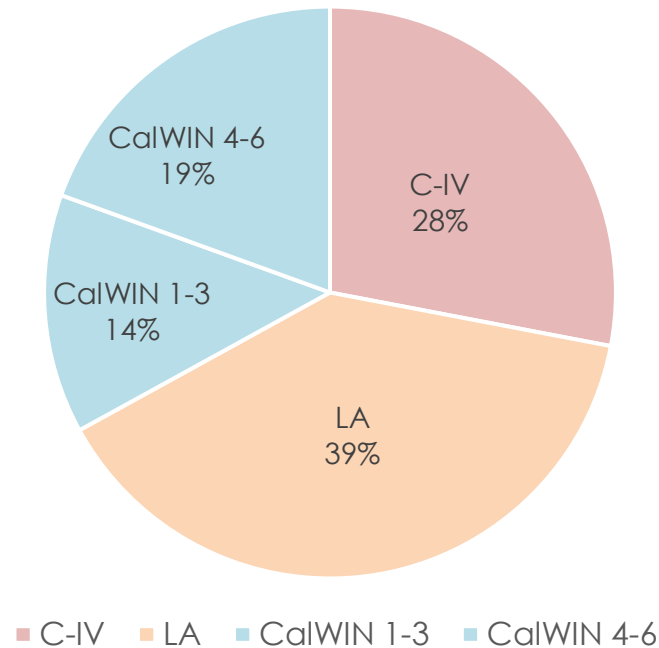


Enhanced Performance Test

58-County Projected Load

- Based on an analysis of the current post-wave 3 production metrics, we are proposing to run an additional endurance test, one at 100% projected 58 county load and another at 122% projected 58 county load.
- Using the performance testing report data, Hyland will develop a scaling plan for Production. This will be based on timing thresholds that will trigger scaling prior to impacting the end-user's experience. This plan will be delivered prior to Wave 4 go-live.

58-County % Breakdown



Enhanced End to End Performance Testing

Testing Schedule May 2023

Week of 5/15		
Friday, May 19	8:00 am - 12:00 pm	Integrated Full Load Pre-Test @ 90% (Wave 4), 100% (58-County)
Week of 5/22		
Monday, May 22	8:00 am - 5:00 pm	Integrated Performance Test @ 100% (58-County)
Tuesday, May 23	Investigate Issues, Resolve Errors, Modify Scripts, etc..	
Wednesday, May 24	8:00 am – 5:00 pm	Endurance Test @ 100% (58-County)
Thursday, May 25	Investigate Issues, Resolve Errors, Modify Scripts, etc..	
Friday, May 26	Investigate Issues, Resolve Errors, Modify Scripts, etc..	
Week of 5/29		
Tuesday, May 30	Investigate Issues, Resolve Errors, Modify Scripts, etc..	
Wednesday, May 31	8:00 am - 12:00 pm	Stress Test 100%-111%-122%-133%
Thursday, June 1	Investigate Issues, Resolve Errors, Modify Scripts, etc..	
Friday, June 2	8:00 am - 5:00 pm	Endurance Test @ 122% (58-County Projected Growth)
Week of 6/5		
Monday, June 5	Investigate Issues, Resolve Errors, Modify Scripts, etc..	
Tuesday, June 6	Investigate Issues, Resolve Errors, Modify Scripts, etc..	
Wednesday, June 7	8:00 am - 5:00 pm	Contingency Test Day 1
Performance Testing Complete		

BenefitsCal OCR Testing

Simulating OCR Readiness for BenefitsCal Workflow

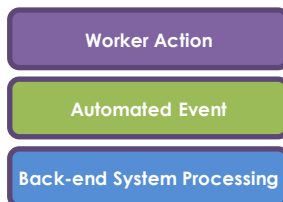
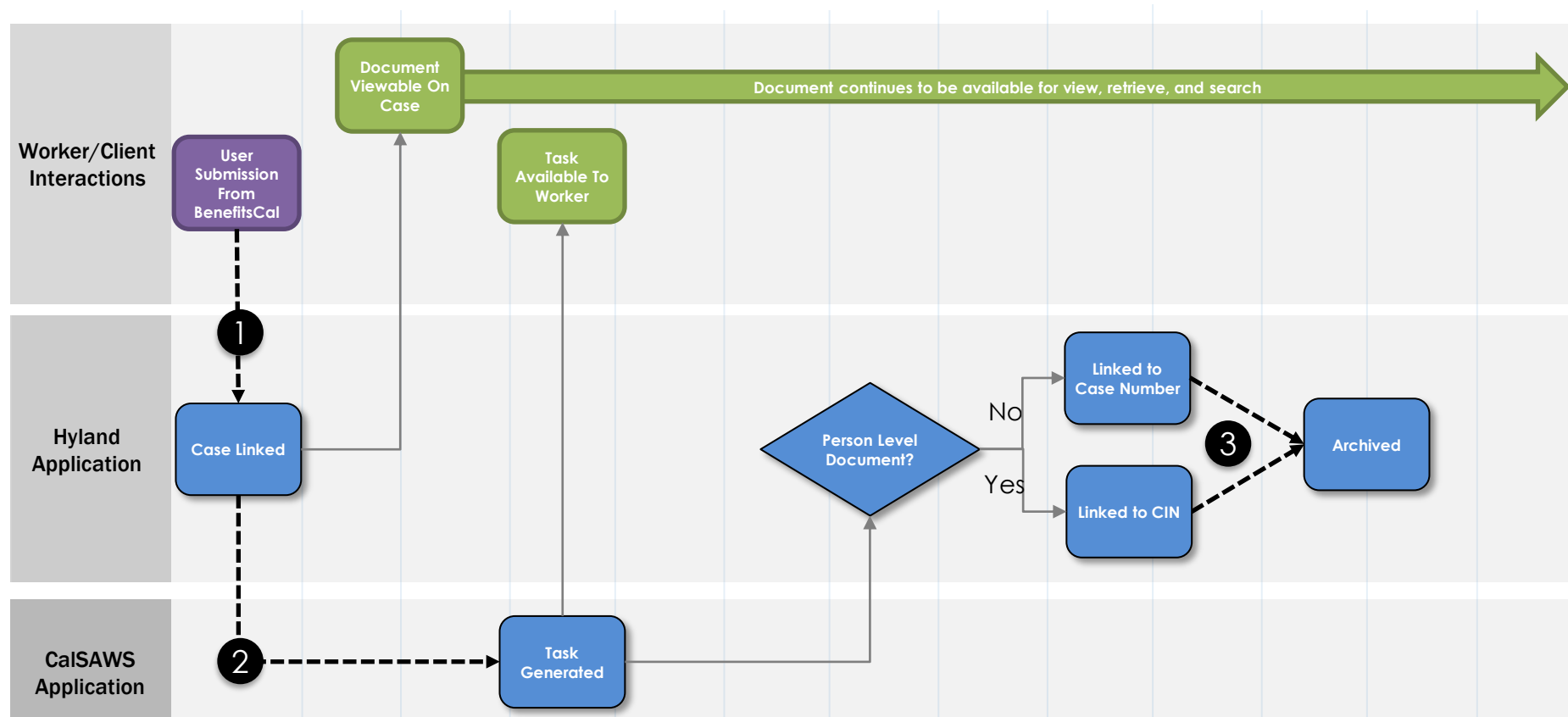
- Upon the completion of the 58-County performance test, the OCR portal functional testing can be performed.
- We met with BenefitsCal and Hyland and determined that BenefitsCal does not have a way to retain production images for OCR testing.
- We are working with Hyland to collect a sufficient test data set to use for OCR testing.

Week of TBD		
TBD	TBD	BenefitsCal Documents
TBD	TBD	BenefitsCal Documents (Optional Day 1)
OCR Testing Complete		



Appendix – Workflows

Portal Submission

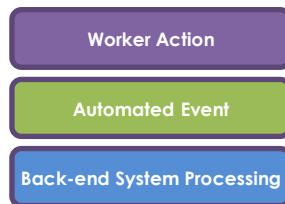
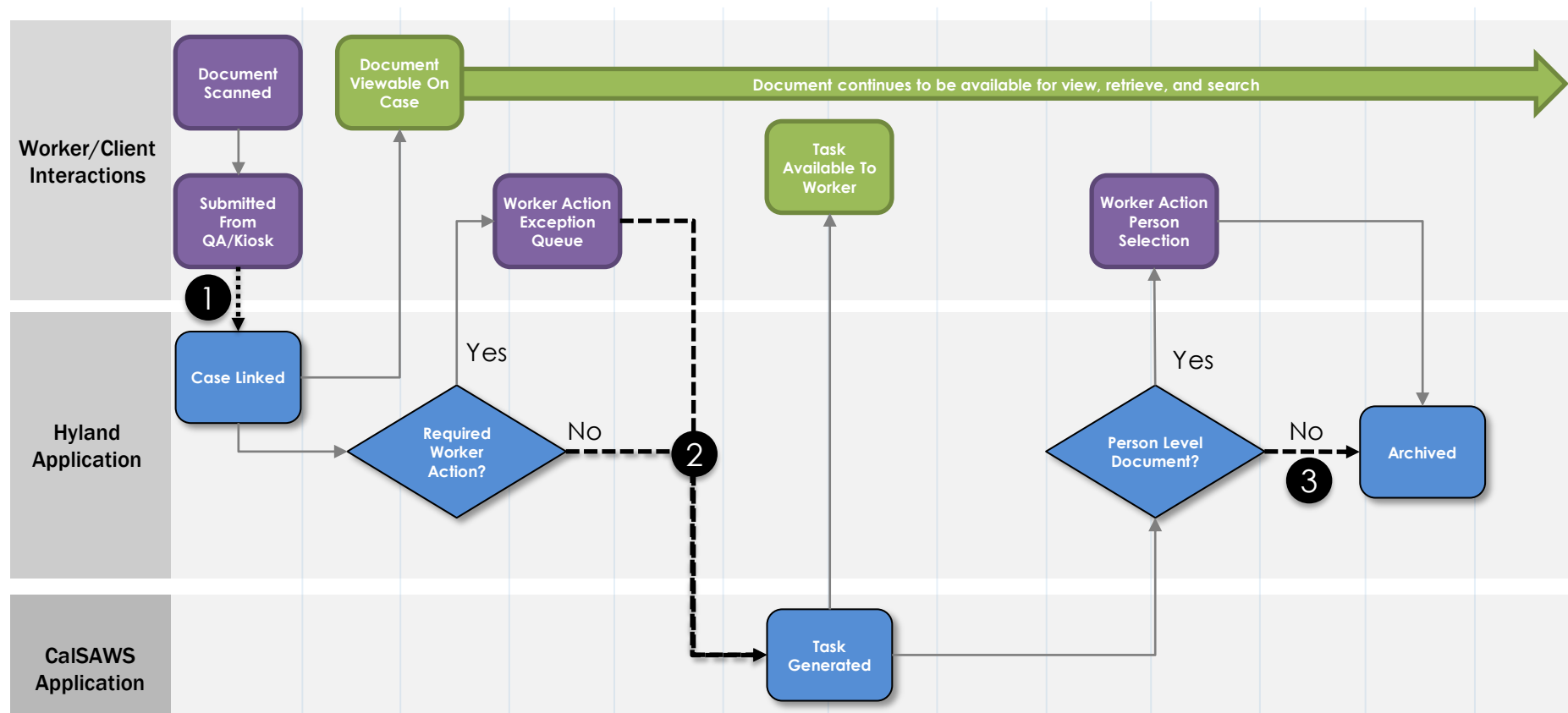


Key Areas Where Time is Spent

- 1. Linking** – Linking of submitted documents is near instantaneous, making the document available/viewable on the case.
- 2. Task** – Document keys are validated, and append logic is applied to prevent duplicate tasks, finally customer reporting is update and a task generated.
- 3. Archive** – Document Keys are re-validated and updated where necessary, then document workflow history is moved to an archive state.

Worker/Kiosk Submission

OCR Bypass & Barcode Workflow

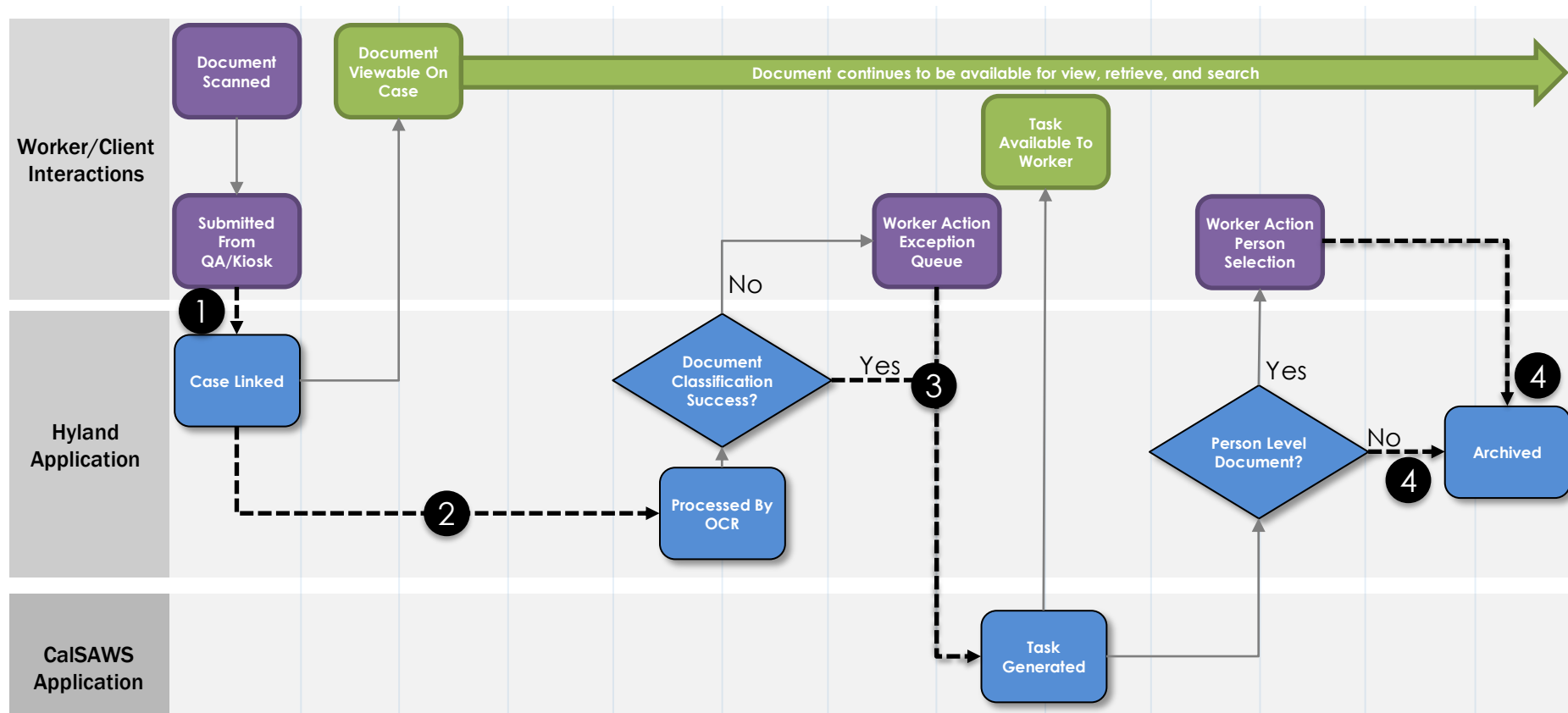


Key Areas Where Time is Spent

- 1. Linking** – Linking of submitted documents is near instantaneous, making the document available/viewable on the case.
- 2. Task** – Document keys are validated, and append logic is applied to prevent duplicate tasks, finally customer reporting is update and a task generated.
- 3. Archive** – Document Keys are re-validated and updated where necessary, then document workflow history is moved to an archive state.

Worker/Kiosk Submission

OCR Workflow



Key Areas Where Time is Spent

- 1. Linking** – Linking of submitted documents is near instantaneous, making the document available/viewable on the case.
- 2. OCR** – Documents are converted to a readable format where needed, then routed to Brainware (OCR) for analysis, once complete a reindex/split of documents is performed based on results.
- 3. Task** – Document keys are validated, and append logic is applied to prevent duplicate tasks, finally customer reporting is update and a task generated.
- 4. Archive** – Document Keys are re-validated and updated where necessary, then document workflow history is moved to an archive state.

