

☒ CalSAWS M&E☐ CalWIN Migration

Distribution Date:	November 16, 2021
To:	PPOC.Alpine; PPOC.Amador; PPOC.Butte; PPOC.Calaveras; PPOC.Colusa; PPOC.Del Norte; PPOC.El Dorado; PPOC.Glenn; PPOC.Humboldt; PPOC.Imperial; PPOC.Inyo; PPOC.Kern; PPOC.Kings; PPOC.Lake; PPOC.Lassen; PPOC.Local; PPOC.Madera; PPOC.Marin; PPOC.Mariposa; PPOC.Mendocino; PPOC.Merced; PPOC.Modoc; PPOC.Mono; PPOC.Monterey; PPOC.Napa; PPOC.Nevada; PPOC.Plumas; PPOC.Riverside; PPOC.San Benito; PPOC.San Bernardino; PPOC.San Joaquin; PPOC.Shasta; PPOC.Sierra; PPOC.Siskiyou; PPOC.Stanislaus; PPOC.Sutter; PPOC.Tehama; PPOC.Trinity; PPOC.Tuolumne; PPOC.Yuba; PPOC.Local; Committee.Imaging.All; Consortium.RegionalManagers.All
CIT Name:	CalSAWS Imaging – Best Practices for OCR Success
From:	CalSAWS Project

PPOCs, please forward to the appropriate impacted staff in your county:

- | | |
|---|--|
| <input type="checkbox"/> General
<input type="checkbox"/> Policy
<input type="checkbox"/> CW
<input type="checkbox"/> CF
<input type="checkbox"/> MC
<input type="checkbox"/> CMSP
<input type="checkbox"/> FC/KG/AAP
<input type="checkbox"/> Child Care
<input type="checkbox"/> WtW
<input type="checkbox"/> Other Program(s) _____
<input type="checkbox"/> BenefitsCal
<input type="checkbox"/> Customer Correspondence
<input type="checkbox"/> Other _____ | <input type="checkbox"/> Reports
<input type="checkbox"/> Fiscal
<input type="checkbox"/> Caseload Movement
<input type="checkbox"/> Management
<input type="checkbox"/> Fiscal
<input type="checkbox"/> Security
<input type="checkbox"/> Batch and Interfaces
<input checked="" type="checkbox"/> Imaging
<input type="checkbox"/> Migration
<input type="checkbox"/> Conversion
<input type="checkbox"/> Technical
<input type="checkbox"/> Training
<input type="checkbox"/> Help Desk |
|---|--|

Description: (Including any step-by-step instructions)	<p>Purpose</p> <p>The purpose of this CIT is to provide information on best practices for Optical Character Recognition (OCR) success in the CalSAWS Imaging Solution. This CIT does not impact Los Angeles County.</p> <p>Background</p> <p>The CalSAWS Project has received feedback that OCR in the CalSAWS Imaging Solution is not working as expected, and that the volume of documents being routed to the Exception and other County-maintained workflow queues is too high.</p> <p>In response to this feedback, the CalSAWS Imaging Team is providing these best practices for OCR success.</p>
--	--

Additional Information

Add Imaging form numbers to county forms

OCR will attempt to read the form number in the bottom left of a document. Counties may incorporate the Form Numbers on the Imaging Form Name Matrix into their forms so OCR can leverage the printed numbers to assist with classification.

Location: The last line on the bottom left, with no text to the immediate right of the form number

Font: Arial, Times New Roman, or Courier

Font size: 8, 10, or 12

Text direction: Horizontal

Margins at minimum: Left .3 in, bottom .3 in (make sure to consider your county's scanner margins)

Any combination of the above requirements can be used. In general, a blank line between the imaging form number and other information on the page is recommended.

Handwritten form numbers may be used but cannot be reliably read by OCR.

Set form name at point of scan

In QA & Indexing, users may enter a form name at the point of scan. Selecting a form name adds weight to the selection, meaning if OCR is trying to decide between two names, it will lean toward the name the user selected. Even if a form name is selected, it is possible that OCR will override the user-selected value.

All forms (excluding flyers or brochures) generated from CalSAWS have an exact match in Imaging. The exact match should be used in most situations, unless dictated otherwise by County policy.

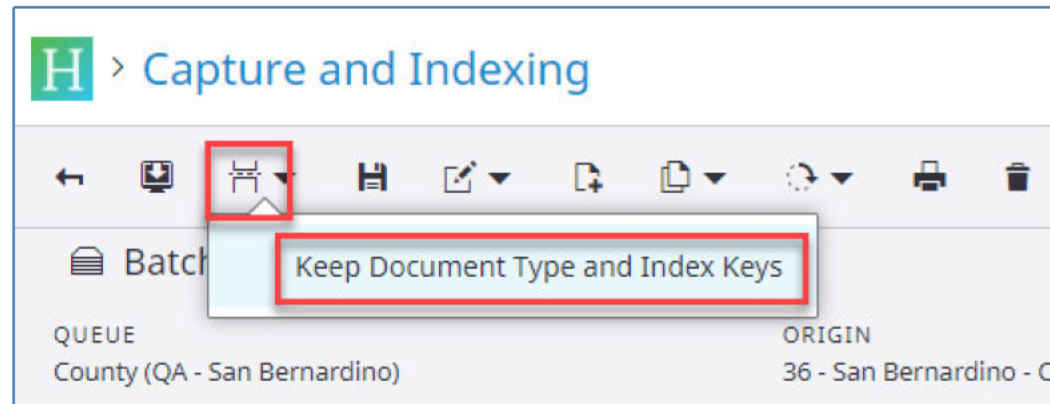
The Imaging Form Names Matrix lists all the Imaging Form Names and their associated document properties. It can be found in the [REDACTED]

Split documents at point of scan

When scanning multiple documents at the same time, use of **separator sheets**, and/or **manually splitting** documents when performing quality assurance before submitting a scanned batch, will **greatly reduce** the likelihood of them being routed to exception.

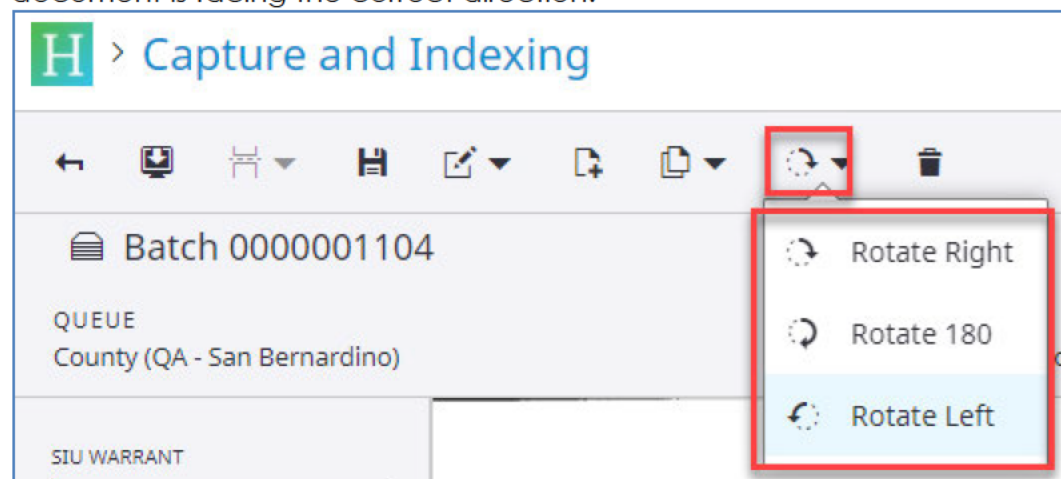
When scanning multiple documents at the same time **without separator sheets** (e.g., a shelter expense, a miscellaneous document, and a utility expense), if the System can correctly identify the form name for pages 1, 2, 4, and 5 but **cannot confidently categorize** page 3, the **entire document** will be sent to exception. Since the System does not know whether page 3 is part of the shelter expense, its own document, or part of the utility expense, it **must err on the side of caution** and route the entire 5-page document to the exception queue for staff to review and make sure it is split correctly.

You can separate documents with **barcoded case coversheets** or **generic separator sheets**; or take a moment to manually split them during quality assurance, to increase OCR's chance of classifying and indexing them correctly. The "Keep Document Type and Index Keys" action will create a document split above the selected page and retain the current case and form context for both documents.



Submit images in the correct orientation

After scanning a document, users can use the rotate buttons to ensure a document is facing the correct direction.



Organize pages in the correct order

Physical page organization should be done before a document is first scanned. If something is out of order after scanning, users should drag and drop the thumbnails of a document into the correct order while the document is in QA & Indexing.

Ensure documents generated from CalSAWS are for programs already linked to a case

When generating a document, users are prompted to select a program. Always select a program that is on the case the document is being generated for. If the document is generated for a program that is not on the case, the barcode will not be valid. When the invalid barcode is read by Imaging, information for the form and case will not be returned.

	<p>Use Photo scan for things with a background (ID, passport, birth certificate, Green card, etc.)</p> <p>Documents with a color background, such as IDs, passports, etc., can be imaged using Photo scan. This is a color capture mode that will allow OCR to better read the document.</p> <p>Scanner settings</p> <p>Scanners should be set to 300 dpi for optimal OCR extraction.</p> <p>County Action</p> <p>The CalSAWS Project recommends that counties distribute this communication to all Imaging users.</p> <p>For additional Imaging documentation, click the following URL for Imaging Fact Sheets: [REDACTED]</p>
Primary Project Contact: (Name, phone number, email address)	Rhiannon Chin ChinR@CalSAWS.org
Backup Project Contact: (Name, phone number, email address)	Chris Vasquez VasquezC@CalSAWS.org
Attachments:	None
Web Portal Link:	<p>[REDACTED]</p> <p>OR</p> <p>You may also retrieve the CIT document and attachments by following these steps:</p> <ol style="list-style-type: none"> 1. Click on the CRFIs & CITs link at the top of the page. 2. Click on the "CalSAWS Information Transmittal (CIT)" folder. 3. Click on the "2021" folder. 4. Click on the appropriate CIT # folder.