

# Optical Character Recognition (OCR) & Exception Queues

This communication shares information on OCR performance and provides examples of how counties can manage exception queues in the CalSAWS Imaging Solution.

## What is Optical Character Recognition (OCR)?



Optical Character Recognition (OCR) technology is a new feature in the Imaging Solution that will automatically detect and validate certain key values from scanned documents, allowing for automatic indexing.

## Document Types that OCR is still learning:

The following document types will go through OCR; however, OCR is still learning to read them, and we do not have success rates to share at this time:



- Attendance Records
- Death Certificate
- Educational Expense
- Income Tax / Tax Return Forms
- Other ID Cards
- Passport
- Pregnancy Verification
- Receipts
- Retirement
- School Registration
- Shelter Expense
- Social Security Card
- Third Party Liability
- Trust Funds
- Utility Expense
- Vehicle Registration

# OCR Performance

The graphs below show the following rates for OCR processing:

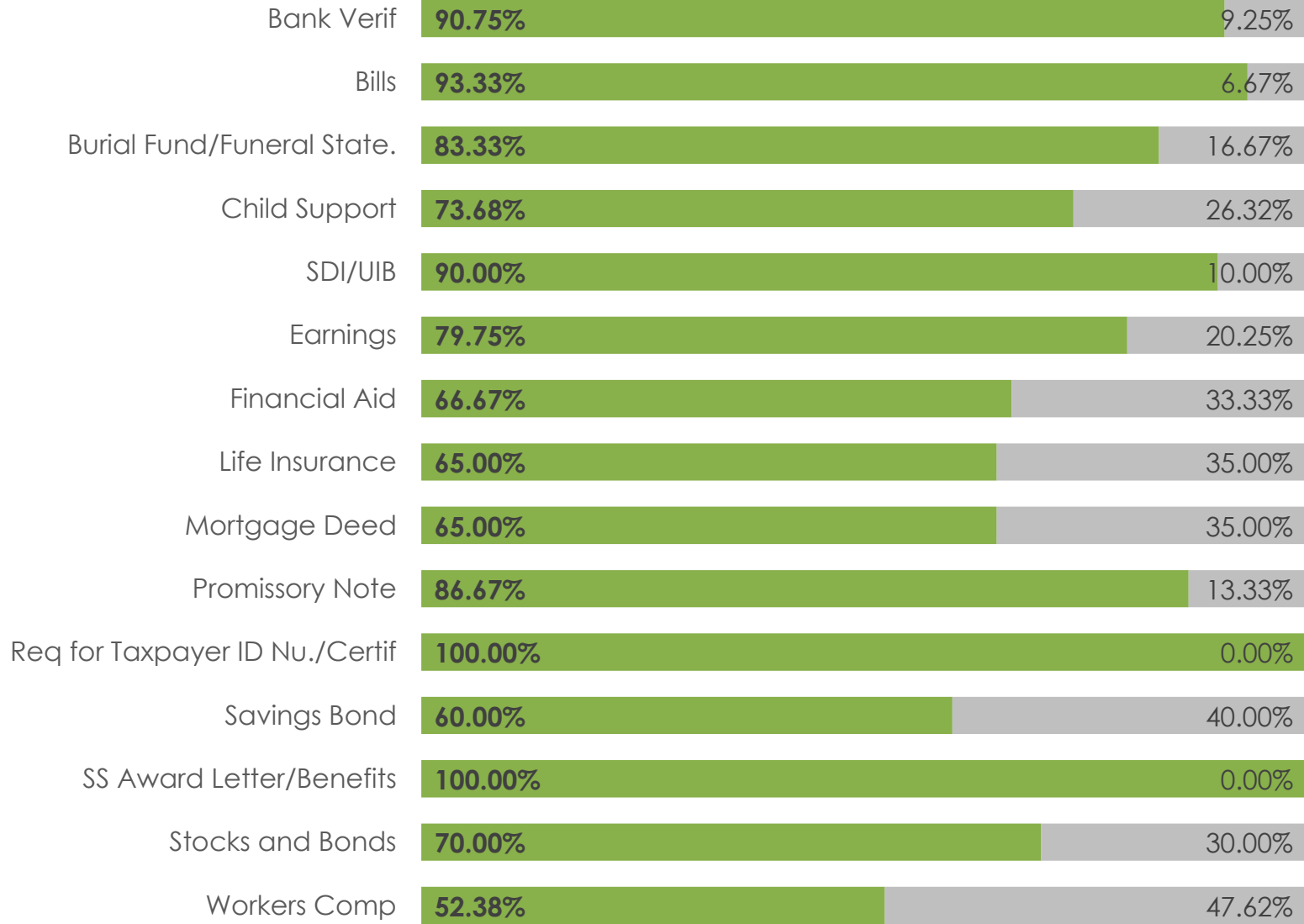
- When the correct Form Name is applied, and the document is archived successfully
- When the Form Name is not identified, and the document is routed to Exception

**This is point in time information as of July 2021. It is subject to change as OCR training and testing continues.**

## CITIZENSHIP / IDENTITY



## FINANCIAL



**NOTE:** This is **point in time** information as of July 2021. **It is subject to change** as OCR training and testing continues.

Correct Form Name      Exception

## EMPLOYMENT SERVICES



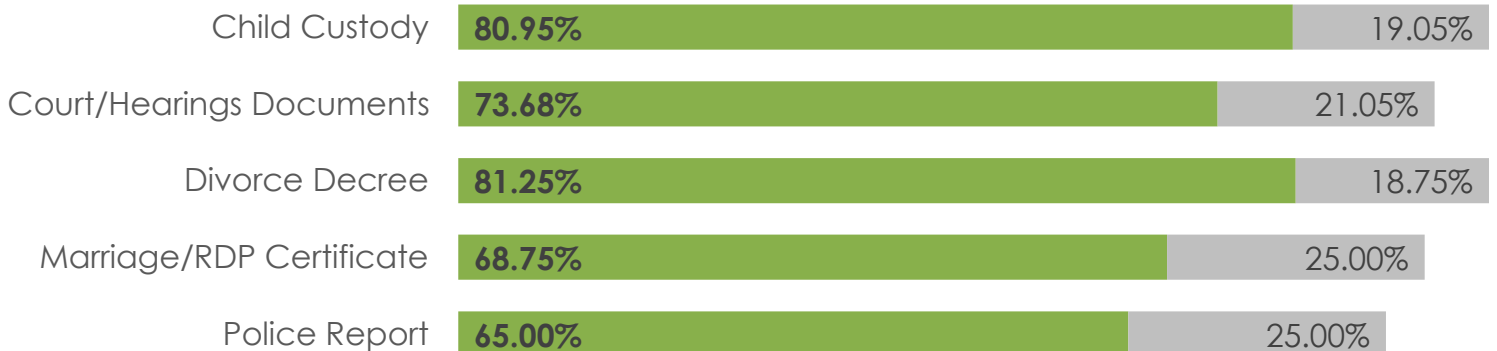
## HEALTHCARE



## EDUCATION



## LEGAL



## MISCELLANEOUS



**NOTE:** This is **point in time** information as of July 2021. **It is subject to change** as OCR training and testing continues.



Correct Form Name



Exception

## Separate Documents to Reduce Exception Routing

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!

## Documents that will not be run through OCR

These documents may be too broad to be accurately processed or conflict with other key items that do go through OCR, and they will always need to be manually indexed:



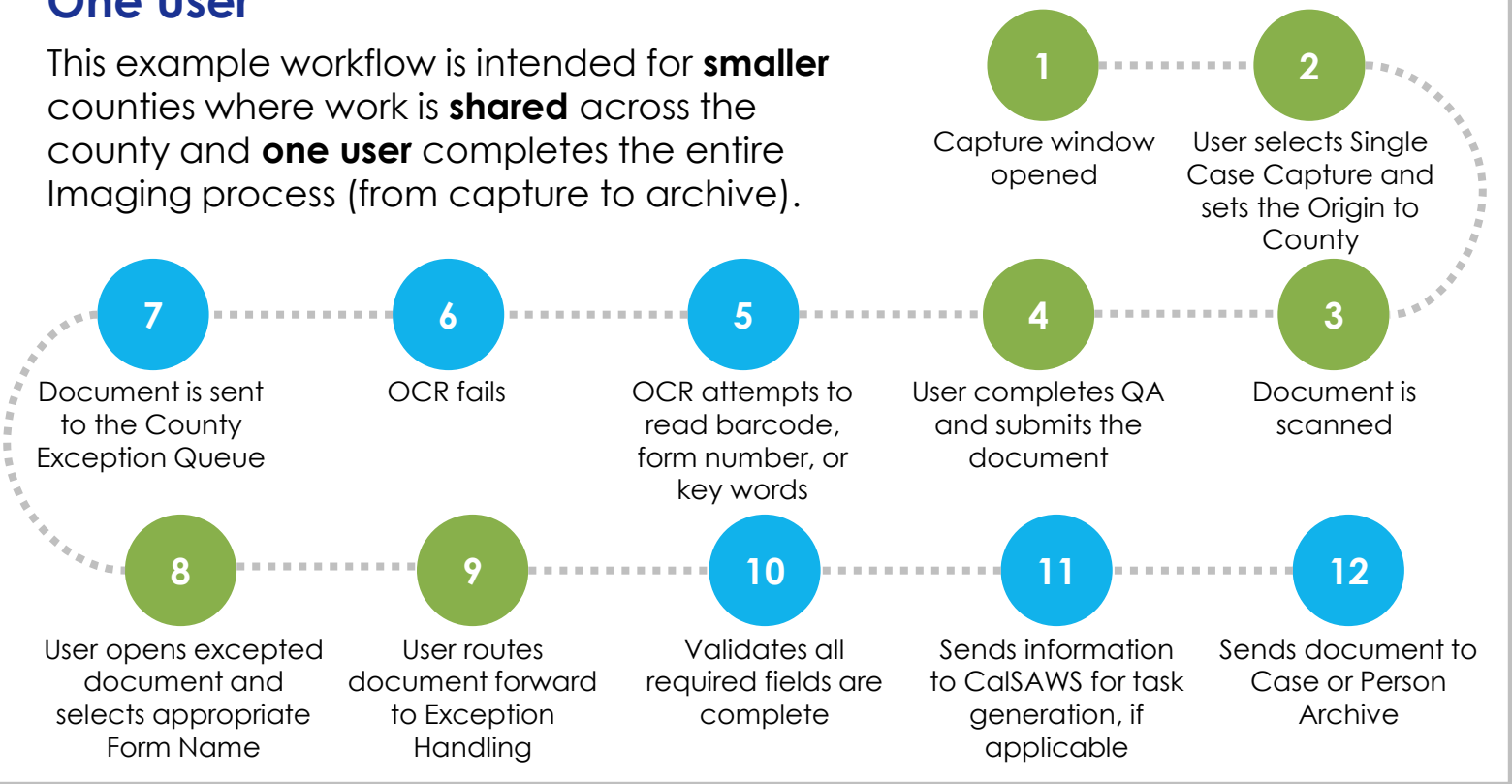
- Adoption Documents
- Assessment
- Authorized Rep Designation
- Contracts
- Disability Verification
- FC Application
- Foster Childs Data Record  
AFDC-FC Certificate
- Gifts / Loans
- Hunt v. Kizer Record of Bills
- IPV
- Long Term Care Facility  
Verification
- Lost Warrant Replacement  
Affidavit
- OP/OI Info
- RSDI – Retirement / Survivors /  
Disability Ins.
- Sanctions

# Sample Workflows for Managing Exception Queues

These are provided as **examples ONLY**. Each County needs to determine how exception queues will be managed based on their internal policies and business processes.

## One User

This example workflow is intended for **smaller** counties where work is **shared** across the county and **one user** completes the entire Imaging process (from capture to archive).



## Separate Users

This example workflow is intended for **larger** counties where work is **separated** by office and **users** have different responsibilities.

