

# Haystac Robotic Document Separator



Advanced AI that transforms a large file into multiple documents. Without separator sheets.

Scanned files often contain multiple documents within. For example, a mortgage loan file package contains up to 100 documents. An insurance claims file may contain forms, invoices, receipts, ID documents. A medical file could contain dozens of documents from patient intake forms to X-rays. A box of records might contain hundreds of diverse documents.

To convert it from paper, the file is scanned as a single batch and each page becomes an image. To keep the file intact, the individual pages are saved into one large PDF.

For a digital process to make sense of the information contained within that PDF, each document within it must also be separated. This is a precondition of document classification and data extraction.

And for some files, the approval or auditing workflow will need to know if the file contains all the required documents. In the mortgage file example, federal loan forms, the appraisal, income verification, proof of insurance, etc. This cannot be verified until document separation is completed.



## Stuck in the Last Century?

Today, this document separation step is still done the same slow, expensive and mistake-prone way as it has been for over 30 years.

A worker must manually separate the file into documents, then insert a separator sheet between documents to instruct the scanning software on the beginning and end. After the batch is scanned, the sheets are removed by a worker and re-used for the next batch.

It's an incredibly tedious manual process that is slow and prone to errors. Humans get bored and miss a document here or there, and as they tire their performance degrades. And when it comes to deciding where a document begins and ends, humans are not consistent.

This is also a costly process. In addition to the obvious labour costs, some business process outsourcers and internal scanning departments have been known to spend tens of thousands of dollars each year to print separator sheets.

The total cost of this process has been estimated at 6 to 10 cents per document. That adds up fast, even for small scanning centers.

It's time for a better way. It's time for AI.

### Files that need separated

- Mortgage Loan Files
- Patient Records
- Insurance Claim Files
- Contracts
- Legal Case Files
- Many more

### Challenges with using Separator Sheets

- Slow
- Error-prone
- Inconsistent separation
- Costs from 6 to 10 cents per document

Haystac package-005.pdf p.10 / 86 Page props Complete Curation ldbd-resnet50-1.0.0-2021-07-21

At a Special Term, of the Supreme Court of the County of Nassau held in the Nassau Supreme Court in the city of Mineola, New York on the \_\_\_\_\_ day of \_\_\_\_\_, 20\_\_\_\_

PRESENT: **HON. THOMAS ANTHONY ADAMS**

SUPREME COURT OF THE STATE OF NEW YORK  
COUNTY OF NASSAU

THE BANK OF NEW YORK MELLON F/K/A THE BANK OF NEW YORK, AS TRUSTEE FOR THE CERTIFICATEHOLDERS OF CWBMS, INC., CHL MORTGAGE PASS-THROUGH TRUST 2006-S, MORTGAGE PASS-THROUGH CERTIFICATES, SERIES 2006-8

Plaintiff

vs

SURJIT BHULLAR, HARPREET BHULLAR A/K/A HARPEET BHULLAR, AMERICAN EXPRESS CENTURION BANK, CHASE BANK USA, N.A., CITIBANK (SOUTH DAKOTA), N.A., DISCOVER BANK, FIA CARD SERVICES NA, HSBC MORTGAGE CORPORATION (USA), "JANE DOE",

Defendant(s)

INDEX #: 14-004702

MORTGAGED PREMISES:  
108 PRAGUE STREET  
VALLEY STREAM, NY 11580

SBL #: 37 - 268 - 33,34, 35, 36, 37

**ORDER OF VACATUR OF JUDGMENT OF FORECLOSURE AND SALE AND DISCONTINUANCE OF ACTION AND CANCELLATION OF LIS PENDENS**

UPON the reading and filing of the Notice of Motion dated January 3, 2017, by Plaintiff's Counsel, Gross Polowy, LLC, Laura Straus, Esq., for an Order Discontinuing this Action and Cancelling Lis Pendens,

AND, upon the Affirmation of Laura Straus, Esq., dated January 3, 2017, setting forth the basis for the relief requested herein,

AND, after having heard \_\_\_\_\_ in opposition of said Motion, it is hereby

**ORDERED**, that the Plaintiff's Motion to Vacate Judgment of Foreclosure and Sale and Discontinue Action and Cancel Lis Pendens, is granted in all respects; and it is further

**ORDERED**, that the above entitled action, and all causes of action alleged therein, and any and all counterclaims and cross claims, be and the same hereby are discontinued without costs to either party as against the other; and it is further

Page Orientation  
 Up  
 Left  
 Down  
 Right

Doc Split  
 Begin  
 End

Additional  
 Inset  
 Heading  
 Garbage  
 Separator  
 Uncertain  
 Accidental  
 Duplicate  
 Clear

Update selected

Back to files  
Home  
About

0.9692 0.0138 0.2197 0.0965 0.011  
 0.0011 0.9998 0.2292 0.005 0.0025  
 0.0562 0.0075 0.0344 0.0018 0.3274  
 0.9981 0.0072 0.0242 0.0189 0.004  
 0.0000 0.0030 0.1004 0.1230 0.0016

# The Haystac Robotic Document Separator (RDS) completely replaces the need for separator sheets, and the human activity associated with them.

## How RDS Works

1. Using a small number of your own files, first fine-tune RDS.
2. Upload PDF files from your production scan workflow into the RDS cloud or the on-prem server (your choice). The file can be any size and any number of pages.
3. RDS reads the file and determines where individual documents begin and end.
4. RDS separates the documents and saves them as individual files.
5. The results are sent back to the workflow.
6. If exceptions are found, the corrections are fed back into the RDS model so it can continue to learn.

## Under The Hood

There is some serious AI at work here. RDS uses deep learning AI models that were pre-trained on tens of thousands of diverse documents.

This allows RDS to capture document experience and to accumulate knowledge, just like humans do.

The pre-trained models use transfer learning so they can work with your documents.

Good news: you don't have to know a thing about AI or data science to operate RDS. We've hidden the magic beneath the simple user interface.



# Transform Your Process with Haystac RDS

## Benefits

### **Saves time and money**

Eliminate separator sheets and the human labor associated with them.

### **Faster time to solution**

RDS trains on far fewer samples than machine learning or rules-based methods.

### **Consistency of document separation**

The machine is consistent, while humans are not.

### **Round the clock operation**

Fully automated, runs 24x7 without any breaks.

## Deployment

RDS is available as a cloud-based micro service with RESTful APIs.

It can be easily integrated into your document capture process.

An On Premise option is also available.

