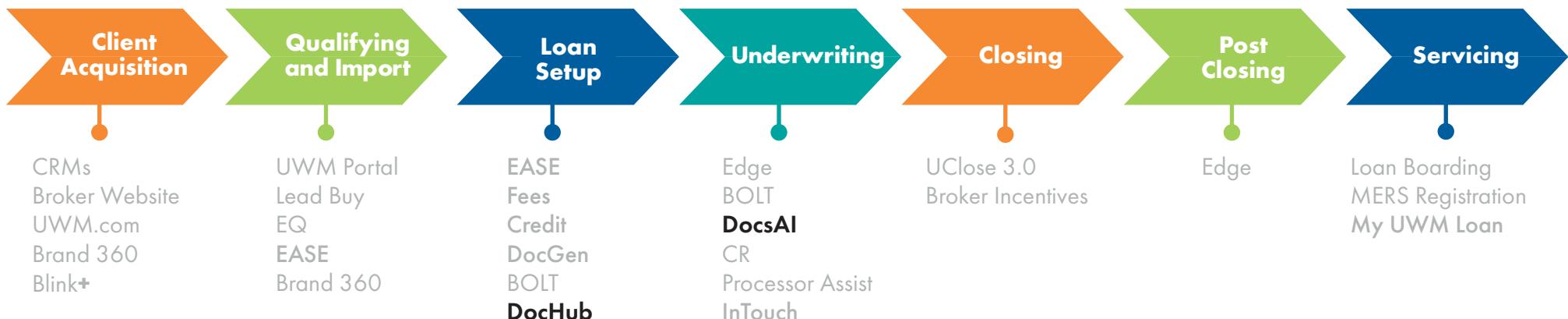


Life of a Loan Recap: DocHub and DocsAI



DocHub Overview

DocHub is a document management system that is launched on the loan level through EDGE, and allows users to receive, review, index, save, and share documents electronically. DocHub can also manage and manipulate document metadata, and share documents securely with internal teams and external clients. Any change to documents in DocHub sends out a Kafka message to let other systems know of the change.

DocsAI Overview

DocsAI uses optical character recognition to classify and automatically label documents. Once the document is labeled, the document's data is extracted and saved to the database. Lastly, it compares this information against EASE and automatically associates the document to the relevant entities in BOLT.

January 2024 Metrics

Received Mail

Acts as the inbox for uploaded documents. Doc labelers sort and review documents and file them into the loan folder.

- Total PDF files uploaded: 638,511
- RMIs created: 3,503,218
- Total documents: 18,835,537

Automated Data Classification (ADR)

The machine learning model to label documents, which drives extraction and association.

- Total documents classified: 5,043,161
- New submission average time: 23.4 seconds
- Unknown rate: 16%

Automated Data Extraction (ADE)

Documents are digitized and the data is saved and retrieved by BOLT. Incorrect data is updated by BOLT and sent back to DocsAI.

- Total loans extracted: 83,440
- Documents extracted: 847,998
- Documents associated: 402,553

For more information on DocHub, [click here.](#)

Common Terms

DMS - Document Management System

A system that collects and stores all documents related to a loan. DocHub is UWM's document management system

DM&D - Document Management & Digitalization

DocHub and DocsAI are UWM's document management and digitization systems.

RMI - Received Mail Item

Also known as an upload, which can contain more than one file.

LLM: Large Language Model

Used to achieve general purpose language comprehension and generation, and other language processing tasks such as answering questions based on documents.

OCR: Optical Character Recognition

Technology that converts handwritten and printed text into machine-encoded text. DocsAI uses OCR as part of the processes that classifies uploaded documents and extracts their data.

