

How Do I Send/Upload Data for CloudNine Review

When you choose CloudNine Review, you have the flexibility to decide how you want to deliver your data to CloudNine. There are two ways to submit your data:

- **Discovery Portal:** A self-service gateway to upload your data. Using Discovery Portal, you can:
 - Upload native (unprocessed) data.
 - Upload already processed data to CloudNine's client services team.
 - Send data that exists in other CloudNine applications.
- **Client Services Team:** Data may be sent directly to CloudNine's client services team via FTP or shipment, The client services team can:
 - Process raw data then upload to CloudNine Review.
 - Upload processed data to CloudNine Review.

Raw data

Simply put raw data is data in its original state (or collected state) that has not undergone any type of processing.

Processed Data

The data has been processed to expand containers, messages, and embedded files and to extract text and metadata to be used to investigate the data fully.

Whether you processed the data yourself or received processed data from a third party, typically data is received in the following manner.

Data Volume

Generally, processed data is delivered in Volumes based on the data size. Each volume is named uniquely and will often be made up of the following sub-folders:

- **Images:** Either TIF or PDF files, based on the agreed-upon delivery format.
- **Native:** Output of corresponding native files. This may be limited to specific file types such as Excel files or files that could not be processed to an image format.
- **Text:** The corresponding OCR or extracted text.
- **Data:** Either on the root of the volume or in a DATA folder you will see the data files. These files are necessary for uploading the data and corresponding files into CloudNine Review. We recommend: **.DAT** file for metadata and **.LFP** for Images.

Image Formats

CloudNine Review supports black and white, color, or mixed images in multiple formats.

- Single-Page TIF
- Single-Page JPG
- Multi-Page PDF

OCR Text Protocol

- Multi-Page text files with OCR text file named the unique doc ID.
- Single-Page text files with OCR named with its corresponding TIF image.
- Text path in the data load file that provides the relative path to the text file.

Naming Protocol

- Files should be uniquely named with an incremental numbering scheme.
 - Single-page Image files are typically named by the unique Page ID.
 - Multi-page PDF files are named the unique BegDoc.
- Native files and text files following the same naming convention as image files.

Load Files

MetaData File: The data file is used to add information about the data and link images, natives, and text in the database. The data file is a delimited text file such as a .DAT or .CSV separated by delimiters to ensure data is mapped to the appropriate field. At the very least the data file should consist of the following:

- Header Row of Field Identifiers.
- Fields separated by delimiters
 - A **.DAT** file using the standard
 - Qualifier: I (020)
 - Delimiter: þ (254)
 - Newline: ® (174)
- Required Fields for TIF, Text, and Native file loads
 - BEGDOC# (DOCID, CONTROL NUMBER, ETC): The unique identifier that is used to link metadata, text, natives, and image files.
 - NativeFile: The relative path to native files.
 - TextFile: The relative path to the corresponding text file.
 - Parent ID or BegAttach: Maintains family relationships.
- The metadata (DAT or CSV) file should **not** contain:
 - OCR or Extracted Text.
 - Message Body.
 - Conversation String

Metadata Example

```
þBegDoc#þþParentIDþþFilenameþþDocExtþþFilesizeþþNativeFileþ  
þABC00000001þþþþþMy File.msgþþmsgþþ20992þþNative\001\ABC00000001.msgþ  
þABC00000002þþþABC00000001þþþMy other file.pdfþþpdfþþ29696þþNative\001\ABC00000002.pdfþ  
þABC00000003þþþABC00000001þþþThis file.docþþdocþþ22016þþNative\001\ABC00000003.docþ
```

Image Load File Example

Image Load File: If your data has been converted to an image format, you will need an image load file to link the images to the corresponding metadata record in CloudNine Review. While we can accept other image load file types, we strongly recommend an **LFP** load file. Below are examples of the LFP load file.

Single-Page Image Load File

```
IM,SINGLE-00000001,D,0,@IMAGES;IMAGES\00000001;SINGLE-00000001.Tif;2  
IM,SINGLE-00000002, ,0,@IMAGES;IMAGES\00000001;SINGLE-00000002.Tif;2  
IM,SINGLE-00000003, ,0,@IMAGES;IMAGES\00000001;SINGLE-00000003.Tif;2  
IM,SINGLE-00000004, ,0,@IMAGES;IMAGES\00000001;SINGLE-00000004.Tif;2  
IM,SINGLE-00000005,D,0,@IMAGES;IMAGES\00000001;SINGLE-00000005.Tif;2  
IM,SINGLE-00000006, ,0,@IMAGES;IMAGES\00000001;SINGLE-00000006.Tif;2
```

Example of Multi-Page PDF load file

```
IM,PRODTEST1-00000001,D,0@IMAGES;IMAGES\00000001;PRODTEST1-00000001.Pdf;7  
IM,PRODTEST1-00000005,D,0@IMAGES;IMAGES\00000001;PRODTEST1-00000005.Pdf;7  
IM,PRODTEST1-00000007,D,0@IMAGES;IMAGES\00000001;PRODTEST1-00000007.Pdf;7
```