

# **DOE Document Imaging**

(202) 586-5251 – Document Imaging Office

(202) 586-2832 – Ralph Freedman, Technical Monitor

<http://ma.mbe.doe.gov/admin/Imaging.htm>

## **Document Imaging Processing**

### **General:**

The process converts paper documents into electronic files. The recommended file format is Acrobat Image + Text, an Acrobat file that contains the scanned pages for viewing and printing, and Optically Character Recognized (OCR'd) text behind it for indexing and searching. By containing the scanned page no information is lost from the original document regardless of issues such as handwriting, charts and graphics, odd fonts and formatting.

### **What the customer will need to do is:**

1. Prepare the documents by separating them so our staff will know which documents create which files. For example, do you want the documents provided made into 1 file, 10 files, or 23 files, and where does one document end and the next begin? If you want to process documents but are unable to organize them for processing please contact the Imaging office to see what can be arranged to facilitate processing.
2. Use our Document Separator Page to identify who the customer is (name, organization code and badge number), and what file name the documents should have when finished (up to 61 characters). Place these sheets at the beginning of the document(s) they refer to. Per the example above, if you want your documents made into 23 files then place a Separator Page where each of the 23 documents begin. This Document Separator Page can be obtained from our Imaging office or our web site.
3. Our default processing is for black and white scanning, if there are any pages or documents that are desired to be scanned in grayscale or color please identify them.
4. Make sure the documents are nicely contained in boxes or other reliable containers for transportation between the customer's office and the Imaging office. If you are unable to do so please contact our office for assistance.

### **What the Imaging group will do:**

1. Prepare documents for scanning by removing staples, clips, and any binding including spiral or glue binding. The documents will be prepared for feeding into a sheet-fed scanner.
2. Scan the documents. This includes Quality Control to make sure all the pages are scanned and as readable as possible.
3. Run Optical Character Recognition (OCR) software to create the text of the pages. Again, our recommended file format is Acrobat Image + Text, but if so desired the documents can also be converted to .rtf format for use in word processors, text-only format, and even html format for quick web posting.
4. Write a CD with the final files. When possible this will be held until there is enough data to fill a CD, and one CD can hold about 10,000 pages of Acrobat Image + Text files.
5. Bind, clip, or otherwise group the pages to make sure the documents stay in order after processing and during transport. The staff will not re-staple documents or re-bind documents.

### **Logistics:**

Whenever possible we ask the customer to provide transportation of their documents to and from our office. Please make sure that the documents are stored in boxes or other containers that are reliable for transport. The Imaging group can provide transportation if necessary.

### **Turnaround times:**

Turnaround times are negotiable based on the customer needs. If you need your files by a certain date just let us know and we will take care of it. Also, if you need daily or weekly processing of files that need a quick turnaround let us know and we will make sure our schedule meets your needs. Otherwise, due to the nature of the work being large-scale, we will process jobs primarily on a first-come basis. Small jobs containing under 1,000 pages will be worked into the daily processing to give them as quick a turnaround as possible.