

UNITED STATES DISTRICT COURT  
SOUTHERN DISTRICT OF NEW YORK

- - - - - X

UNITED STATES OF AMERICA :

- v - :

**STIPULATION AND ORDER  
REGARDING DISCOVERY**

ANDREW ANTONUCCI, :  
DOMINIC ANTONUCCI, :  
SEAN BEARD, :  
STEPHEN BLOOD, :  
EUGENE CABRERA, :  
RICHARD DEARCOP, and :  
MARCOS MARTINEZ, :

04 Cr. 828 (SAS)

Defendants. :

- - - - - X

IT IS HEREBY STIPULATED AND AGREED, by and between the United States of America, David N. Kelley, United States Attorney, Diane Gujarati, Assistant United States Attorney, and William J. Stellmach, Special Assistant United States Attorney, of counsel; and the defendants, Charles A. Ross, Esq., on behalf of defendant Andrew Antonucci; Carl F. Schoeppl, Esq., on behalf of defendant Dominic Antonucci; Deirdre Von Dornum, Esq., on behalf of defendant Sean Beard; Louis V. Fasulo, Esq., on behalf of defendant Stephen Blood, that

1. The bulk of discovery in this matter consists of approximately 45 boxes of documents which the Government has obtained from, among other sources, customers of Platinum Investment Corporation, banks, and broker dealers.

2. It would be beneficial to the parties for discovery of these documents to be made in electronic format, specifically through production of scanned images and optical character recognition ("OCR") processed text, together with associated databases and load files necessary for viewing the documents.

3. Upon the agreement of the defendants to share costs as specified below, the Government will arrange for the documents to be scanned and OCR-processed as set forth below, and to produce the documents in electronic format to the defendants.

4. **Cost sharing.** The cost of electronic processing will not be known until a bid has been accepted by the Government under Government procurement procedures. The parties understand that a rough estimate of the total cost of electronic processing the documents in this case is \$0.25 per page, or approximately \$34,000 in total for this job. The parties understand that a rough estimate of the time within which electronic processing of these documents will be completed is approximately October 30, 2004, or approximately six weeks from entry of this Stipulation and Order, whichever is later.

- a. The defendants agree to pay a total of 50% of the total cost, each defendant accepting responsibility for his pro rata share of the costs. Defendants who are financially unable to pay their share will make timely application to this Court, as specified below, for funding for that share under the Criminal Justice Act. See Guide to Judiciary Policies and Procedures, Guidelines for the Administration of the Criminal Justice Act and Related Statutes ("CJA Guidelines") §3.02.A, available at <http://www.fd.org>, setting forth the procedures for approval by the United States District Judge or Magistrate and for subsequent certification by the Chief Judge of the Circuit (or his delegate) for authorization of payments in excess of \$1,000.
- b. The Government will pay the remaining 50% of the total cost.
- c. The defendants will contract individually with the vendor for payment of their share of the costs and will pay their share of the costs directly to the vendor. The defendants understand that this timing and this method of payment are **essential** to effect the sharing of costs in this matter.

5. **Electronic format.** The parties understand that, pursuant to detailed specifications provided to the vendor by the Government, documents will be scanned as single-page .tif images; and electronic output from the vendor will consist of a fully functional Concordance EX database for use with an IPRO 7.6 viewer. In addition, the Government will require the vendor to produce all load files, including a .dii file for images, necessary to permit convenient loading of the electronic output into Summation. Copies of the specifications submitted by the Government dated September 14, 2004 and by defense counsel dated September 14, 2004 are attached; however, the parties understand that the Government may modify those specifications as necessary to effect this task. The Government will provide the discovery on a single hard drive, to be shared by defendants, or a single set of DVDs, to be shared by defendants.

6. **No coding of documents.** The parties understand that the electronic processing envisioned herein does **not** include any coding of documents (e.g., there will be no recording of date, author, recipient, document type, and or subject for each document) other than as necessary and incidental to the scanning process to enable viewing of documents and pages and the generation of OCR text. The parties understand that in its format as delivered, the electronic discovery in this case will be text-searchable only via the OCR text generated for each

document, and that OCR text generally does not perfectly replicate the actual text of the underlying document, particularly if the underlying document is of poor quality, and that handwritten notes are generally not recognized during OCR processing.

6. Counsel for the defendants expect to make effective use of the electronic discovery in the following manner. Specifically, Charles Ross, Esq., expects to install WebBlaze (the Internet interface for Summation) on computers at his office and to permit co-counsel to have Internet access via WebBlaze to the discovery. It is specifically **not** anticipated that any counsel will need to apply for CJA funds to purchase computer hardware or software under the procedures of CJA Guidelines §3.16. Should Mr. Ross become unable to provide the above access to co-counsel, the defendants understand that they may arrange for the database to be hosted by an internet service provider or independent litigation support company and will cooperate in making such arrangements should the need ever arise.

NOW, THEREFORE, upon the request of the parties, IT IS  
HEREBY ORDERED that

1. Discovery of the documents referred to in paragraph  
1 of the foregoing stipulation shall be made in the format  
specified in the stipulation.

2. Each defendant shall pay his pro rata share of the  
discovery in the manner specified in paragraph 4 of the foregoing  
stipulation. Any defendant who needs to apply for CJA funds to  
pay his pro rata share shall do so immediately after a vendor's  
bid on this task has been accepted, which is expected to occur on  
or about October 1, 2004.

3. Should the total cost of this task exceed 115% of  
the estimated total cost set forth in paragraph 4 of the  
stipulation any party may apply to the Court for reconsideration  
of this Order. The parties may also agree in writing to  
modifications in the details of this Order without further  
recourse to the Court.

DAVID N. KELLEY  
United States Attorney

By:   
Diane Gujarati  
William J. Stellmach  
Assistant United States Attorneys

On behalf of defendant Andrew Antonucci:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Charles A. Ross, Esq.

On behalf of defendant Dominic Antonucci:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Carl F. Schoeppl, Esq.

On behalf of defendant Sean Beard:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Deirdre Von Dornum, Esq.

On behalf of defendant Stephen Blood:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Louis Fasulo, Esq.

On behalf of defendant Eugene Cabrera:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
John Bergendahl, Esq.

On behalf of defendant Richard Dearcop:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
John Parrinello, Esq.

On behalf of defendant Marcos Martinez:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Vincent Ancona, Esq.

SO ORDERED.

DATED: New York, New York  
October 15, 2004



---

SHIRA A. SCHEINDLIN  
UNITED STATES DISTRICT JUDGE

On behalf of defendant Andrew Antonucci:

By: Charles A. Ross Date: 10-21-04  
Charles A. Ross, Esq.

On behalf of defendant Dominic Antonucci:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Carl F. Schoepl, Esq.

On behalf of defendant Sean Beard:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Deirdre Von Dornum, Esq.

On behalf of defendant Stephen Blood:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Louis Fasulo, Esq.

On behalf of defendant Eugene Cabrera:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
John Bergendahl, Esq.

On behalf of defendant Richard Dearcop:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
John Parrinello, Esq.

On behalf of defendant Marcos Martinez:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Vincent Ancona, Esq.

On behalf of defendant Andrew Antonucci:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Charles A. Ross, Esq.

On behalf of defendant Dominic Antonucci:

By: Carl F. Schoeppl Date: 10/25/04  
Carl F. Schoeppl, Esq.

On behalf of defendant Sean Beard:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Deirdre Von Dornum, Esq.

On behalf of defendant Stephen Blood:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Louis Fasulo, Esq.

On behalf of defendant Eugene Cabrera:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
John Bergendahl, Esq.

On behalf of defendant Richard Dearcop:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
John Parrinello, Esq.

On behalf of defendant Marcos Martinez:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Vincent Ancona, Esq.

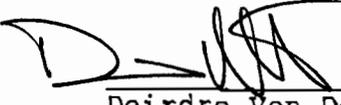
On behalf of defendant Andrew Antonucci:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Charles A. Ross, Esq.

On behalf of defendant Dominic Antonucci:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Carl F. Schoeppl, Esq.

On behalf of defendant Sean Beard:

By:  DAVID PATTON FOR DEIRDRE VON DORNUM Date: 10/22/04  
Deirdre Von Dornum, Esq.

On behalf of defendant Stephen Blood:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Louis Fasulo, Esq.

On behalf of defendant Eugene Cabrera:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
John Bergendahl, Esq.

On behalf of defendant Richard Dearcop:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
John Parrinello, Esq.

On behalf of defendant Marcos Martinez:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Vincent Ancona, Esq.

On behalf of defendant Andrew Antonucci:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Charles A. Ross, Esq.

On behalf of defendant Dominic Antonucci:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Carl F. Schoeppl, Esq.

On behalf of defendant Sean Beard:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Deirdre Von Dornum, Esq.

On behalf of defendant Stephen Blood:

By:  \_\_\_\_\_ Date: 10/22/04  
Louis Fasulo, Esq.

On behalf of defendant Eugene Cabrera:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
John Bergendahl, Esq.

On behalf of defendant Richard Dearcop:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
John Parrinello, Esq.

On behalf of defendant Marcos Martinez:

By: \_\_\_\_\_ Date: \_\_\_\_\_  
Vincent Ancona, Esq.

**Scan/OCR Specs**  
***United States v. Antonucci***  
**September 14, 2004**

---

**1. Job Description:**

- a. **Job Summary:** Bids on a per-page basis are sought on the following work:
  - i. Scan and OCR documents. *See* specifications for Concordance EX File, IPRO Cross-Reference File, and Summation load files.
  - ii. Perform stringent quality assurance for all deliverables.
- b. **Source Material.** Consists of approximately 45 boxes of documents (approximately 135,000 pages.) Approximately one half of the documents are photocopies; approximately one-half are original business records. 90% or more of all documents are in 8½ x 11" format, stapled, clipped and/or rubber-banded. Almost all documents are one-sided. Documents in landscape format are estimated to be 3% or less of the total. Documents originally in landscape (*e.g.*, spreadsheets) should be scanned in landscape. Double sided documents, estimated at 5%, should be converted to single page format. The documents are in numbered boxes. The documents are already Bates-stamped. Vendors may briefly inspect all material prior to bidding if they desire. Visits for inspection should be arranged through the contract officer. Documents should be re-stapled/rubber-banded/clipped after scanning, and returned in the same or equivalently packed and numbered boxes. A box inventory will be provided and will indicate special instructions in the event there are any unusual items in a given box.
- c. **Electronic output** received from the vendor will consist of: .tif image files, a Concordance EX database file, and an IPRO 7.6 .lfp file; as well as load files sufficient to permit construction of an image database using Summation. The IPRO 7.6 .lfp file specs and Concordance EX database file specs are set forth below.
- d. **Delivery Media.** Electronic output will be on DVDs or hard drives to be provided by vendor.
- e. **File Uploading and Quality Assurance.** After the images, Concordance EX database and the IPRO 7.6 .lfp files are loaded onto the delivery medium, as well as the Summation load files, the vendor shall perform quality assurance on the electronic output. At a minimum, quality assurance checks should include: 1) no duplication of records in .lfp or Concordance EX database files; 2) no gaps in bates numbering other than as may occur in source materials; 3) all images are included in the .lfp file; and 4) all filepaths in the .lfp file are accurate; and 4) similar verification for the Summation load files.
- f. **Delivery Schedule.** A DVD containing the electronic output for the first box of documents shall be delivered to SDNY via overnight courier within three days after delivery of the first shipment of boxes. Work should not proceed until SDNY has OK'd the sample delivery. Thereafter, DVDs/hard drive containing the electronic output from the first batch of approximately 22 boxes shall be delivered on/before

the 10<sup>th</sup> day following award of contract, and DVDs/hard drive containing the electronic output from the remaining approximately 23 boxes shall be delivered on/before the 11<sup>th</sup> day after the preceding delivery. Commitment to an earlier delivery will be evaluated favorably in awarding the contract. There should be no overlap in the contents of the electronic output for the first batch and for the second batch. File names for each delivery should reflect which delivery they pertain to.

**g. Rolling Return of Boxes.** Boxes of documents will be delivered and returned on a rolling basis and not held until the entire project is completed. Vendor will return boxes to USAO SDNY.

**2. Image Specifications.** Image files (the output of the scanning process) shall be scanned at 300 dpi and loaded onto the delivery media in accordance with the following:

- a.** File format: single-page, CCITT Group 4 (.tif format)
- b.** File names: Documents are already Bates-stamped. Image .tif file names are to be the same as the bates number label for each sheet, **except:** (i) if the Bates numbers for a segment of documents are not all of the same length, leading 0's must be inserted between the alpha prefix and the number to make them all a uniform length; and (ii) if the Bates numbers for a segment of documents do not have an alpha prefix, a unique alpha prefix must be generated by vendor and inserted. **NB:** vendor must be alert to gaps in the sequence of Bates numbers in the hard copy documents.
- c.** Folder names: Images should be in appropriately-named subfolders under a master folder uniquely entitled XIMAGES. (Images files should NOT all be placed in one master directory.) Do not use the term XIMAGES in the name of any sub-folder.
- d.** Files per sub-folder: There is no strict maximum on number of files per sub-folder, but retrieval is quicker if the number of files is limited to a few hundred or less. Each sub-folder should be 600 MB or less in size.

**3. Concordance EX File**

Delivery should include all Concordance-generated files associated with a database file. Following are field specifications for the Concordance EX database file. (System generated fields may also be added during creation of database.)

FieldName	Type	Size	Comments
StartBates	Text	20	Bates number for first page of a document. <b>NB:</b> This will be the Image Key field. This must match the image key in the IPRO .lfp file.
EndBates	Text	20	Bates number for last page of a document. (Documents include all attachments.)
AttachStart	Text	20	Bates number of first page of any attachment to a document.
AttachEnd	Text	20	Bates number of last page of any attachment to a document.

	Pages	Num	4.0	Number of pages in a document. (Documents include all attachments.) Machine-generated.
	BoxNo	Text	20	Box number of document box. Will be constant over box. Will appear on box or on inventory provided to vendor.
	ProdSource	Para	16	The entity that produced the documents. To be supplied by SDNY. Will be constant over a large volume of documents.
	OCRText	Free Text		The OCRText from the OCR processing for each document. If necessary, create additional OCRText fields if amount of text warrants their creation (e.g. OCRText2, OCRText3, etc.).  <b>NB:</b> the OCR text files must contain machine-generated page headers corresponding to each sheet in a document that will contain page numbers in the form "page X of Y".

#### 4. IPRO Cross-Reference File

An IPRO version 7.6 .lfp file shall be provided in the general format explained below.

**Format:**

ImportCodeIdentifier, ImageKey, BoundaryFlag, RotationParameter, VolumeName, DirectoryPath, Filename, ImageType

**Example for a two-page document with a starting Bates number of BN000325:**

```
IM,BN000325,D,0,@XIMAGES;subfolder1\subsubfolder1;BN000325.tif;2
IM,BN000326,,0,@XIMAGES;subfolder1\subsubfolder1;BN000326.tif;2
```

Field Name	Meaning
ImportCodeIdentifier	This code tells IPRO how to evaluate the line. If the .lfp file is created via IPRO or a program that converts log files to IPRO format, leave the code(s) as generated by IPRO or the conversion program. If the file is generated otherwise, the value <b>IM</b> should be inserted.
ImageKey	A unique identifier for the page, preferably the BatesNo for the page
BoundaryFlag	Indicates the beginning of a page, document, folder, box, or source. A document boundary is essential. Values are as follows: S = Source B = Box F = Folder D = Document (blank) = Page

RotationParameter	Indicates rotation of document. Value should generally be 0. Values are as follows: 0=0 1=90 2=180 3=270
VolumeName	The name of the parent folder under which all images are stored. This should be the same in all .lfp files. This should preferably be XIMAGES, with that name not being used anywhere else in path and file names.
DirectoryPath	The path under the VolumeName under which the particular image file is stored.
FileName	the .tif file name
ImageType	Indicates the type of image file. Should be 2 indicating standard single-image TIFF (.tif) format. Values are as follows: 1 = IPRO Tech image (.IMG) file 2 = standard single-image TIFF (.TIF). 3 = IPRO Tech Stacked TIFF (.STF) file 4 = color images (.BMP, .PCX, .JPEG, .PNG)

#### 4. Summation Load Files

Certain end-users of the electronic output will be using Summation rather than Concordance and IPRO. As noted, all load files, including a .dii file for images, necessary to permit convenient loading of the electronic output into Summation must accompany each delivery. As set forth in paragraph 1.e above, checking Summation load files is part of quality assurance for this job.

**TO:** SAUSA Bill Stellmach  
**FROM:** Chris Padurano  
**DATE:** September 7, 2004  
**RE:** Scan/OCR Specs for *United States v. Antonucci, et al.*

Chuck asked me to send you our comments/suggestions for the "Summation Load Files" portion of the draft you previously provided. We want to insure that whichever vendor performs the scanning and OCRing for this case understands the following:

This table displays the components in the recommended order in which they should appear in the DII file.<sup>1</sup> Note that Comment (;) lines can appear anywhere in the DII record, but most often appear at the beginning of a record.

Item	Required/Optional	Significance
<b>i</b>	Optional	Comment lines used for organizational or informational purposes.
<b>@Fulltext</b>	(relates to loading full-text documents)	Indicates that ocrBase documents are being loaded from the Default Directory designated by the <b>@D</b> line. This means the same directory where the images are stored. See also <b>@O</b> and <b>@OCR/@OCR-END</b> .
<b>@O</b>	(relates to loading full-text documents)	Directory containing ocrBase documents, if other than the Default Directory (the directory designated by the <b>@D</b> line). See also <b>@Fulltext</b> and <b>@OCR/@OCR-END</b> .
<b>@OCR/@OCR-END</b>	(relates to loading full-text documents)	Load actual OCR text using a DII file. OCR must be placed between the <b>@OCR</b> and <b>@OCR-END</b> tokens and the tokens must reside on their own

<sup>1</sup> In order to retrieve an image and/or ocrBase document, a DII record must contain at a minimum: the **image tag**, the default directory, and document filenames. A DII record without one or more of these elements will not be loaded into the ImgInfo Table and is listed in the error file generated by the Load DII utility. If you are using the **DII file to load eMail or eDocs**, other requirements exist.

For a basic DII file, maintain the order of these elements within each record as indicated: image tag first followed by the default directory followed by the filenames. See the **sample DII files** for suggestions when other components are added to the DII record.

		lines.
<b>@C</b>	Optional	Column to Link. Unlimited number of columns. Length limited as defined by data type (e.g., a text field of 8 characters would only allow 8 characters to be used in the <b>@C</b> item)
<b>@I</b>	<b>Required</b>	Image Tag
<b>#pages</b>	Optional	Number of pages of the document
<b>@D</b>	<b>Required</b>	Default Directory
<b>@I</b>	(relates to @D)	Image Location as in Case Customize
<b>@V</b>	(relates to @D)	Volume Label of <b>CD-ROM</b> and any subdirectories. Volume labels longer than 8 characters are accepted.
<b>@L</b>	Optional	Long Name entry
<b>Image Filename</b>	<b>Required</b>	Image document files and <b>iterators</b> . Long filenames are accepted

- White space (spaces, tabs, blank lines, etc.) is ignored except as required to separate elements of a line. You can use white space to enhance readability of the **DII**. Blank lines can be used to separate records.
- It may also be helpful to include comments within the file. To do so, type a semicolon (;) in the first column. Summation then disregards the rest of that line. There is no continuation character, so you must use another semicolon in the first column of each successive line of comments.
- Summation will not load a record into the **ImgInfo** Table if the image tag or filename is missing from the **DII** record. Other components may be safely omitted but should later be added manually by editing the **ImgInfo** record within Summation. If you have no information available for a required item, include the designator but leave the rest of the line blank. Refer to the **Components** chart for information about optional vs. required components.

### **DII File Format**

The **DII** file is a formatted ASCII file that can be created within any text editor. A service bureau's data entry personnel need only be aware of the format to use when creating a **DII** file. They do not need to have a copy of Summation on-site to create the **ImgInfo** Table entries required to retrieve images and full-text documents from within Summation.

The DII file uses a non-delimited, multiple line format. Each line in a DII file corresponds to a separate field within a record of the ImgInfo Table. Each record ends with the image filenames themselves.

#### **DOCUMENT BREAKS**

Based on your draft, we are unclear how the document page breaks will be determined. We understand that document page breaks can be broken up either "physically" or "logically." We prefer that the document page breaks be broken up "logically." For example, assume there is a "document" that contains a one-page fax cover sheet, three-page letter, and six-page report. Breaking up the document pages "physically" would cause the vendor to assign a beginning and end range to that "document" of 1-10. On the other hand, breaking up the document logically would cause the vendor to assign a range of 1 to the fax cover sheet (with an "attach range" of 1-10), a range of 2-4 for the letter (with an "attach range" of 1-10), and a range of 5-10 for the report (with an "attach range" of 1-10).

Also, ask your tech people about the Concordance "AttachStart" and "AttachEnd" and its significance to a Summation Load File. If these automatically correspond to the Summation "attach range" field then it's no problem. However, the vendor may need to supply both a Concordance field and a Summation field for our documents to load properly. The answer to this question might make our concern regarding the document page breaks moot. Thanks.