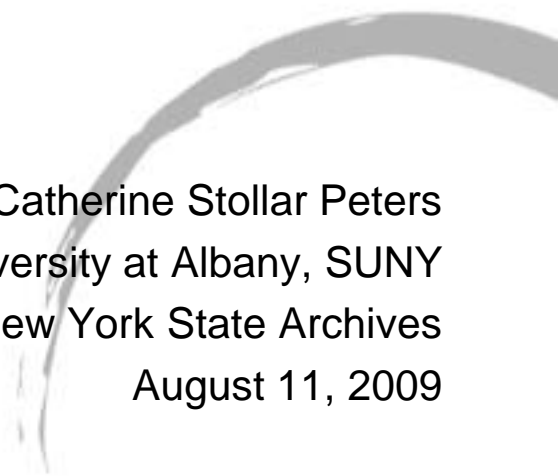




Searching for “Spitzer” in the Archives: Using Search Strategies in Conjunction with Social Network Analysis of Troopergate Records



Catherine Stollar Peters
University at Albany, SUNY
New York State Archives
August 11, 2009

Troopergate Scandal

- Spitzer administration officials and New York State Police acted inappropriately in investigating travel of then-Senate Majority Leader Joseph Bruno
- Involvement of Eliot Spitzer is unclear



Troopergate Timeline

- July 1, 2007
 - Times Union Article alleges then-Senator and Majority Leader Joseph Bruno misused state aircraft for travel relating to fund-raising
- July 2, 2007
 - Office of Attorney General launches investigation into Bruno's actions
- July 5, 2007
 - Bruno requests OAG investigate the Governor's "misuse of State resources in connection with State Police surveillance of Senator Bruno's activities"
- July 23, 2007
 - OAG issues report

- July 23, 2007
 - State Office of Inspector General issues report with similar findings to Office of Attorney General
- Shortly after July 23, 2007
 - Daren Dopp is suspended without pay
- September 21, 2007
 - Albany County District Attorney's Office publishes first report absolving Executive Chamber staff of any crimes
- March 28, 2008
 - Albany County District Attorney's Office published second report called "Investigation D" (called for by Commission on Public Integrity's request to determine if Darren Dopp perjured himself in a sworn statement)
- July 24, 2008
 - Commission on Public Integrity (Formerly Ethics Commission) Issues report declaring Baum, Dopp, Howard and Felton violated Public Office's law (only Dopp and Felton knowingly violated the law)

- July 23, 2007
 - State Office of Inspector General issues report with similar findings to Office of Attorney General
- Shortly after July 23, 2007
 - Daren Dopp is suspended without pay
- September 21, 2007
 - Albany County District Attorney's Office publishes first report absolving Executive Chamber staff of any crimes
- March 28, 2008
 - Albany County District Attorney's Office published second report called "Investigation D" (called for by Commission on Public Integrity's request to determine if Darren Dopp perjured himself in a sworn statement)
- July 24, 2008
 - Commission on Public Integrity (Formerly Ethics Commission) Issues report declaring Baum, Dopp, Howard and Felton violated Public Office's law (only Dopp and Felton knowingly violated the law)

- July 23, 2007
 - State Office of Inspector General issues report with similar findings to Office of Attorney General
- Shortly after July 23, 2007
 - Daren Dopp is suspended without pay
- September 21, 2007
 - Albany County District Attorney's Office publishes first report absolving Executive Chamber staff of any crimes
- March 28, 2008
 - Albany County District Attorney's Office published second report called "Investigation D" (called for by Commission on Public Integrity's request to determine if Darren Dopp perjured himself in a sworn statement)
- July 24, 2008
 - Commission on Public Integrity (Formerly Ethics Commission) Issues report declaring Baum, Dopp, Howard and Felton violated Public Office's law (only Dopp and Felton knowingly violated the law)

- 
- 
- October 2008
 - Senate Committee on Investigations and Government Operations investigation dropped
 - January 2009
 - Commission On Investigation
 - May 2009
 - State Inspector General
 - Current
 - Office of Attorney General

People

- New York Times Graphic from July 27, 2007
- <http://www.nytimes.com/imagepages/2007/07/24/nyregion/24spitzer.graph.web.html>

Records

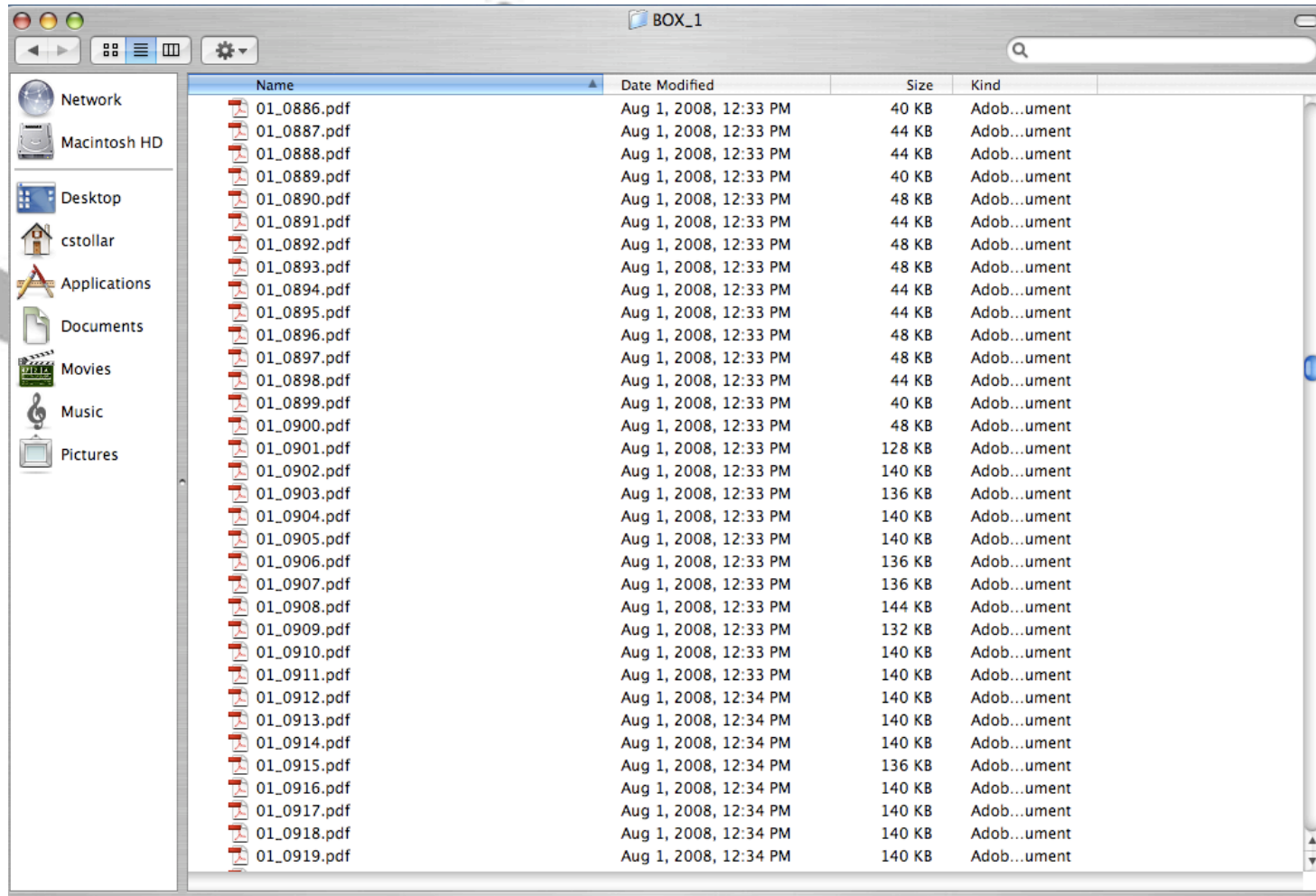
Albany County DA

- Second investigation
- Accessioned, B1989
- Scanned documents
- Single scan files
- Access via search
- Duplicates

Commission on Public Integrity

- RDA in process
- Scanned documents
- Arranged files
- Metadata Access
- Unique interviews

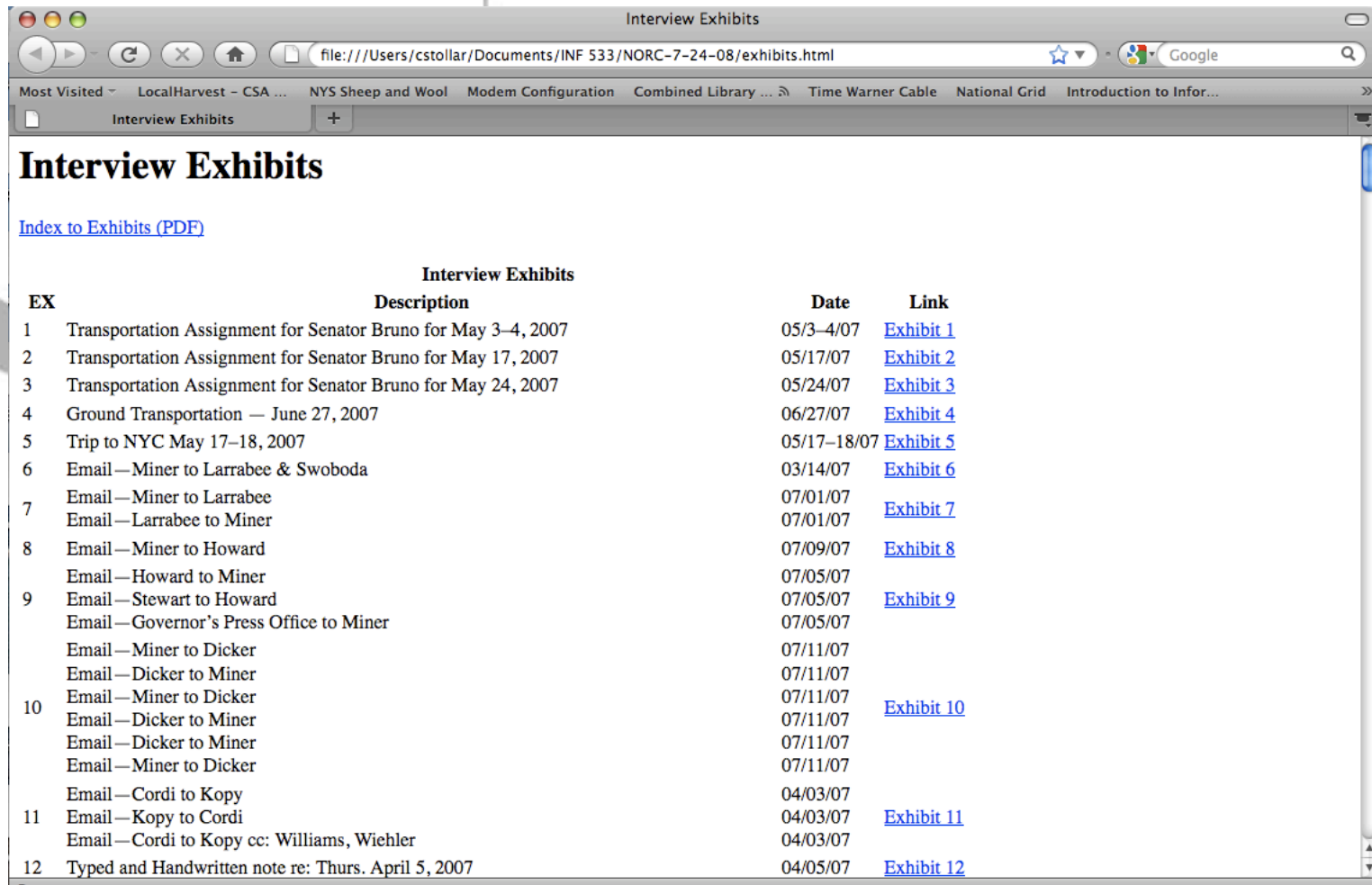
Screenshot of B1989 Records



The screenshot shows a Mac OS X file browser window titled "BOX_1". The window displays a list of PDF files in a table format. The columns are "Name", "Date Modified", "Size", and "Kind". The files are listed in chronological order from 01_0886.pdf to 01_0919.pdf. All files were modified on August 1, 2008, at 12:33 PM, except for the last five files (01_0912.pdf through 01_0919.pdf) which were modified at 12:34 PM. The file sizes range from 40 KB to 140 KB, and all are Adobe documents.

Name	Date Modified	Size	Kind
01_0886.pdf	Aug 1, 2008, 12:33 PM	40 KB	Adob...ument
01_0887.pdf	Aug 1, 2008, 12:33 PM	44 KB	Adob...ument
01_0888.pdf	Aug 1, 2008, 12:33 PM	44 KB	Adob...ument
01_0889.pdf	Aug 1, 2008, 12:33 PM	40 KB	Adob...ument
01_0890.pdf	Aug 1, 2008, 12:33 PM	48 KB	Adob...ument
01_0891.pdf	Aug 1, 2008, 12:33 PM	44 KB	Adob...ument
01_0892.pdf	Aug 1, 2008, 12:33 PM	48 KB	Adob...ument
01_0893.pdf	Aug 1, 2008, 12:33 PM	48 KB	Adob...ument
01_0894.pdf	Aug 1, 2008, 12:33 PM	44 KB	Adob...ument
01_0895.pdf	Aug 1, 2008, 12:33 PM	44 KB	Adob...ument
01_0896.pdf	Aug 1, 2008, 12:33 PM	48 KB	Adob...ument
01_0897.pdf	Aug 1, 2008, 12:33 PM	48 KB	Adob...ument
01_0898.pdf	Aug 1, 2008, 12:33 PM	44 KB	Adob...ument
01_0899.pdf	Aug 1, 2008, 12:33 PM	40 KB	Adob...ument
01_0900.pdf	Aug 1, 2008, 12:33 PM	48 KB	Adob...ument
01_0901.pdf	Aug 1, 2008, 12:33 PM	128 KB	Adob...ument
01_0902.pdf	Aug 1, 2008, 12:33 PM	140 KB	Adob...ument
01_0903.pdf	Aug 1, 2008, 12:33 PM	136 KB	Adob...ument
01_0904.pdf	Aug 1, 2008, 12:33 PM	140 KB	Adob...ument
01_0905.pdf	Aug 1, 2008, 12:33 PM	140 KB	Adob...ument
01_0906.pdf	Aug 1, 2008, 12:33 PM	136 KB	Adob...ument
01_0907.pdf	Aug 1, 2008, 12:33 PM	136 KB	Adob...ument
01_0908.pdf	Aug 1, 2008, 12:33 PM	144 KB	Adob...ument
01_0909.pdf	Aug 1, 2008, 12:33 PM	132 KB	Adob...ument
01_0910.pdf	Aug 1, 2008, 12:33 PM	140 KB	Adob...ument
01_0911.pdf	Aug 1, 2008, 12:33 PM	140 KB	Adob...ument
01_0912.pdf	Aug 1, 2008, 12:34 PM	140 KB	Adob...ument
01_0913.pdf	Aug 1, 2008, 12:34 PM	140 KB	Adob...ument
01_0914.pdf	Aug 1, 2008, 12:34 PM	140 KB	Adob...ument
01_0915.pdf	Aug 1, 2008, 12:34 PM	136 KB	Adob...ument
01_0916.pdf	Aug 1, 2008, 12:34 PM	140 KB	Adob...ument
01_0917.pdf	Aug 1, 2008, 12:34 PM	140 KB	Adob...ument
01_0918.pdf	Aug 1, 2008, 12:34 PM	140 KB	Adob...ument
01_0919.pdf	Aug 1, 2008, 12:34 PM	140 KB	Adob...ument

Screenshot of CPI Records

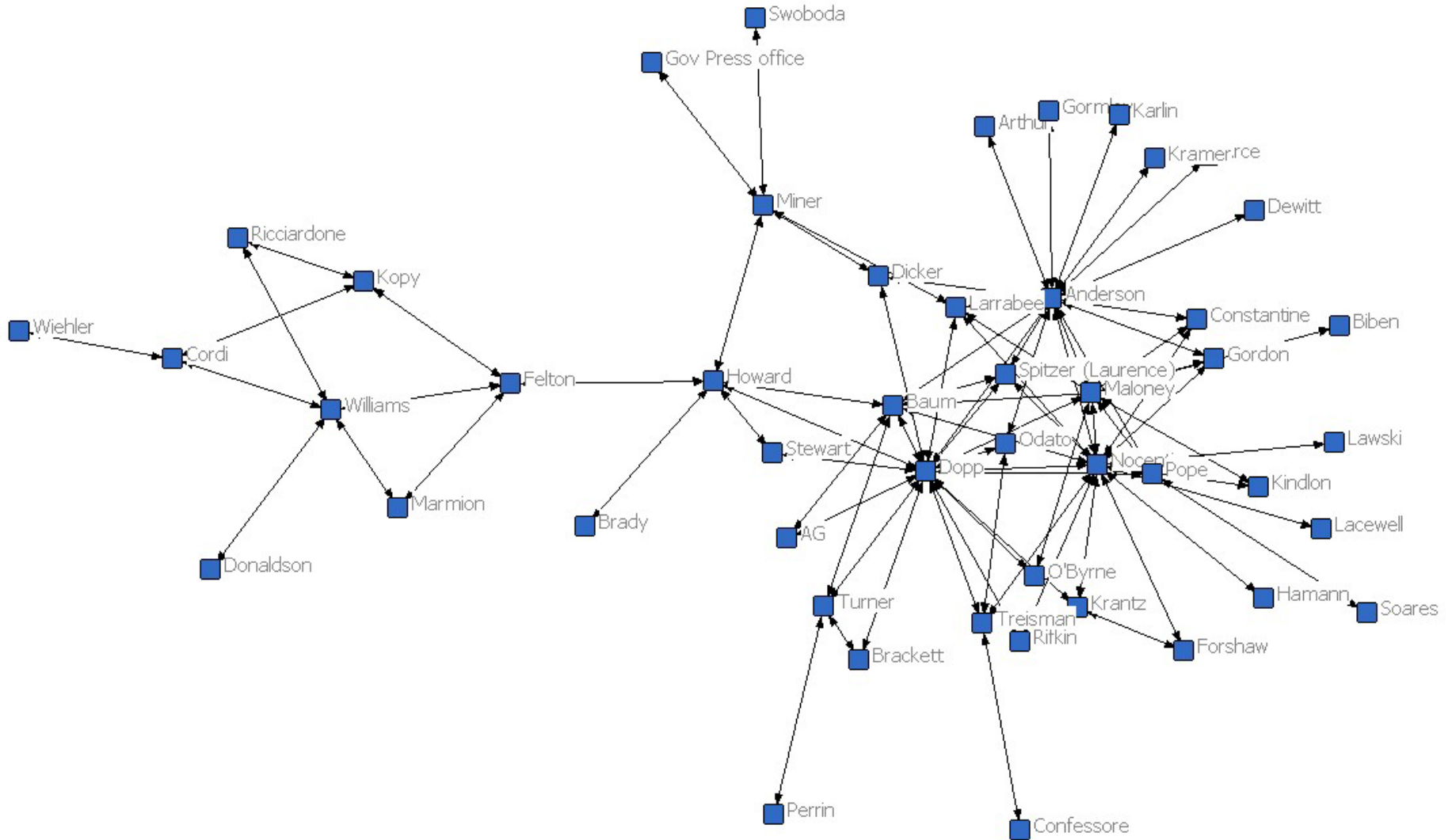


The screenshot shows a web browser window titled "Interview Exhibits". The address bar contains the file path: file:///Users/cstollar/Documents/INF 533/NORC-7-24-08/exhibits.html. The browser's bookmark bar shows several sites, including "LocalHarvest - CSA ...", "NYS Sheep and Wool", "Modem Configuration", "Combined Library ...", "Time Warner Cable", "National Grid", and "Introduction to Infor...". The main content area displays the title "Interview Exhibits" and a link "Index to Exhibits (PDF)". Below this is a table with four columns: "EX", "Description", "Date", and "Link". The table lists 12 exhibits, each with a unique identifier, a description of the document, its date, and a link to the exhibit.

EX	Description	Date	Link
1	Transportation Assignment for Senator Bruno for May 3-4, 2007	05/3-4/07	Exhibit 1
2	Transportation Assignment for Senator Bruno for May 17, 2007	05/17/07	Exhibit 2
3	Transportation Assignment for Senator Bruno for May 24, 2007	05/24/07	Exhibit 3
4	Ground Transportation — June 27, 2007	06/27/07	Exhibit 4
5	Trip to NYC May 17-18, 2007	05/17-18/07	Exhibit 5
6	Email—Miner to Larrabee & Swoboda	03/14/07	Exhibit 6
7	Email—Miner to Larrabee Email—Larrabee to Miner	07/01/07 07/01/07	Exhibit 7
8	Email—Miner to Howard Email—Howard to Miner	07/09/07 07/05/07	Exhibit 8
9	Email—Stewart to Howard Email—Governor's Press Office to Miner Email—Miner to Dicker Email—Dicker to Miner	07/05/07 07/05/07 07/11/07 07/11/07	Exhibit 9
10	Email—Miner to Dicker Email—Dicker to Miner Email—Dicker to Miner Email—Miner to Dicker	07/11/07 07/11/07 07/11/07 07/11/07	Exhibit 10
11	Email—Cordi to Kopy Email—Kopy to Cordi Email—Cordi to Kopy cc: Williams, Wiehler	04/03/07 04/03/07 04/03/07	Exhibit 11
12	Typed and Handwritten note re: Thurs. April 5, 2007	04/05/07	Exhibit 12

Search Solution

- Lemur Toolkit, Indri Search Engine
 - Faster than Adobe Reader/Acrobat
 - More than posting list, ranked listing
 - Requires more user friendly interface
- Search Issues
 - OCR'd PDF text
 - Text askew
 - Misspellings
 - Header information in email lost when printed and scanned
 - Can not rely on co-occurrence of names



Future Work

- Archives
 - Automatic comparison of duplicate records
 - Intern project to arrange B1989
 - Accession CPI records
- Research
 - Create timeline slices of social network