

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

CATHERINE STOLLAR PETERS

Abstract: In the Spring of 2009, I analyzed a corpus of documents related to the 2007 “Troopergate” scandal in New York State. The documents were gathered by the Albany County District Attorney’s Office during a the investigation of Darren Dopp, Governor Eliot L. Spitzer’s Director of Communications, to determine if he had perjured himself or committed other crimes during a 2007 inquiry into the “Troopergate” case. The records consist of email, deposition transcripts, travel logs, and faxes printed, redacted, and scanned by a vendor and made available to the public in the Fall of 2008. The documents are publicly available, not under copyright and contain no legally sensitive material. Using text indexing software (the Lemur Toolkit) and social network analysis techniques (and UCInet software), I analyzed 8,562 archival documents in the series.

My research fulfills three goals. First, I demonstrate improved access methods to digital archival documents through sophisticated indexing as compared to the simple posting lists created by native PDF searching. Second, I use archival documents to create visualizations of relationships over time between important actors in the “Troopergate” scandal including Eliot Spitzer, Joseph Bruno, Darren Dopp and other Executive Chamber staffers. Third, this project employs methods of document analysis that can be applied to digital records for use in archival appraisal, description, arrangement, and use. Although my research was conducted using scanned images of documents, all of the techniques I used can be applied to born-digital documents in future investigations of description and use of born-digital archival materials.

About the author:

Catherine Stollar Peters is an Archives and Records Management Specialist for the New York State Archives in the Information Services Unit specializing in information access, organization and use. She also serves on cross-organizational teams dealing with electronic records, appraisal, image creation and management, and access strategies. In her previous position as Archivist at the Harry Ransom Center at the University of Texas at Austin she specialized in electronic records preservation. Her work with the archival materials of hypertext author Michael Joyce is the subject of multiple publications including a case study published by the Society of American Archivists in the symposium proceedings of New Skills for a Digital Era. Currently, she is working on her Ph.D. in Information Science at the University at Albany with a focus in archiving digital records and manuscripts and archival forensics.