

**Georgia
Tech**



**Research
Institute**



An Analysis of the Knowledge Required to Perform FOIA and PRA Review

Marlit Hayslett-Keck
William Underwood

PERPOS Technical Report ITTL/CSITD 04-1
January 2004
Revised March 2004

Georgia Tech Research Institute
Georgia Institute of Technology
Atlanta, Georgia

The Army Research Laboratory (ARL) and the National Archives and Records Administration (NARA) sponsored this research under Army Research Office Cooperative Agreement DAAD19-03-2-0018. The findings in this paper should not be construed as an official ARL or NARA position unless so indicated by other authorized documentation.

ABSTRACT

Review of Presidential electronic records for access restrictions is an intellectually demanding task that requires page-by-page review of Presidential records. Due to the increasing volume of Presidential electronic records, the need to review these records, and the cost of the limited human resources that can be applied to the review process, the review process is an archival processing bottleneck. The objective of this research is to determine the kinds of knowledge that archivists use to review Presidential Records for Presidential Record Act (PRA) restrictions and Freedom of Information Act (FOIA) exceptions and to use this information to develop and evaluate an automated tool to support archivist's decisions in reviewing Presidential Records.

The first step in our research was to become acquainted with the FOIA and PRA statutes. We attended a Freedom of Information Act and Privacy Act Workshop conducted at the USDA Graduate School. The knowledge that archivists use to judge whether records are Federal, Presidential or personal records is discussed. Some of the knowledge that archivists use to determine whether records or passages in them are subject to access restrictions is also reported.

Sample Presidential and Federal records are being analyzed that have PRA restrictions and FOIA exemptions. The sample also includes some personal records and some Presidential records that were opened because they had not PRA or FOIA access restrictions.

The kinds of knowledge identified include the knowledge that is needed to identify document type (e.g., memo, letter, agenda), to interpret the content of a record, the knowledge (concepts) characterizing a FOIA exemption or PRA restriction, and the knowledge acquired from experience in reviewing Presidential, Federal and Personal records.

Table of Contents

1. INTRODUCTION.....	1
<i>Background.....</i>	<i>1</i>
<i>Purpose.....</i>	<i>3</i>
<i>Scope.....</i>	<i>3</i>
2. KNOWLEDGE OF PRA AND FOIA REVIEW	3
2.1 DISTINGUISHING PRESIDENTIAL RECORDS FROM PERSONAL RECORDS	4
2.2 PRA EXEMPTIONS AND FOIA EXCEPTIONS	6
<i>PRA Restriction a(1) and FOIA Exemption b(1).....</i>	<i>7</i>
<i>PRA Restriction a(2).....</i>	<i>8</i>
<i>PRA Restriction a(3) and FOIA Exemption b(3).....</i>	<i>9</i>
<i>PRA Restriction a(4) and FOIA Exemption b(4).....</i>	<i>12</i>
<i>PRA Restriction a(5).....</i>	<i>14</i>
<i>PRA Restriction a(6) and FOIA Exemption b(6).....</i>	<i>16</i>
<i>FOIA Exemption b(2).....</i>	<i>17</i>
<i>FOIA Exemption b(5).....</i>	<i>18</i>
<i>FOIA Exemption b(7).....</i>	<i>19</i>
<i>FOIA Exemption b(8).....</i>	<i>21</i>
<i>FOIA Exemption b(9).....</i>	<i>21</i>
3. INITIAL ANALYSIS OF SOME SAMPLE DOCUMENTS.....	22
<i>Personal Records.....</i>	<i>22</i>
<i>Presidential Records with a(2) and/or a(5) Restrictions.....</i>	<i>25</i>
<i>Records with b(6) and b(7) Exemptions.....</i>	<i>27</i>
<i>Presidential Records that have no Restrictions.....</i>	<i>29</i>
4. SUMMARY AND CONCLUSIONS	30
REFERENCES.....	33

LIST OF FIGURES

Figure 1. The APT's User Interface for Review of Electronic Records.....	2
Figure 2. An Example of the First Lady's Personal Correspondence, 1990	22
Figure 3. XML Tagged Text File.....	23
Figure 4. Template for the Documentary Form of White House Letters.....	24
Figure 5. Record of a Nomination to Federal Judgeship.	26
Figure 6. Sample Incident Report.	28
Figure 7. Scenario Template for an Incident Report.....	29
Figure 8. Example of a Press Release - Statement by the Press Secretary.	30

1. Introduction

Background

Review of Presidential electronic records for access restrictions is an intellectually demanding task that requires page-by-page review of Presidential Library accessions. Due to the increasing volume of electronic records from all branches of government, the need to review these records, and the cost of the limited human resources that can be applied to the review process, the review process is an archival processing bottleneck. The objective of this study is to determine the kinds of knowledge that archivists use to review Presidential Records for Presidential Record Act (PRA) restrictions and Freedom of Information Act (FOIA) exceptions and to use this information to develop and evaluate an automated tool to support archivist's decisions in reviewing Presidential Records. Such a tool is not a replacement for the judgment of archivists, whose responsibility it is to review the records.

NARA estimates that about 40,000,000 unclassified Clinton administration e-mail messages have been accessioned by NARA. Assume that an email message is on average one page in length (including attachments). A DoD report of costs of declassification indicated that on average 59.4 pages could be reviewed per hour [1]. Let us assume one minute to review and make access restriction decisions on a single page. With a 2000 hour work-year (120,000 minute work-year), it would take 333.33 work-years to review the Clinton e-mail the first time. NARA's Annual Performance Report for FY 2001 referred to 28,925 cubic feet of textual paper records included in the Presidential Records of the Clinton Administration. A cubic foot of textual records translates to 2500 pages. This translates to 72,312,500 pages of textual records that will need to be reviewed. Using the assumptions above, this translates into about 602.6 work-years for review of the textual records.

The Presidential Electronic Records Pilot System (PERPOS) project is developing a system to assist archivists in processing electronic records created by office applications on personal computers. The system, called the Archival Processing Tool (APT), supports archivists in accessioning, arranging, preserving, reviewing, and describing record series. It also stores the records in a repository and supports the creation of reference collections for public access [2].

Experiments are being conducted in which archivists use the tool for processing electronic records created on White House personal computers during the administration of President George H. W. Bush. During the review activity, a reviewer can view records in a file system and review them for access restrictions. The records can be opened or closed in their entirety, redacted, marked a Personal Record Misfile, or transferred to a software library, because they are misfiled software that was used to create the records, rather than being a record.



Figure 1. The APT's User Interface for Review of Electronic Records.

Fig. 1 shows the APT user interface during the review activity. The Tools pull-down menu shows where the APT could support checking for access restrictions. There are many potential benefits of such a tool.

- 1) It might identify an access restriction not identified by the reviewer, thus reducing the risk of opening a record or passage of a record whose access should have been restricted.
- 2) It might be used as a tutor during training of review archivists.
- 3) Novice reviewers could use the tool to check their work.
- 4) The tool might provide additional evidence in case a reviewer's judgment was uncertain, or point out uncertainties, where the reviewer thought the decision was certain.
- 5) It might give a rapid review to records responsive to a FOIA request to estimate the workload in terms of the number of restrictions and types of restrictions likely to apply.
- 6) It might estimate which unprocessed electronic record series are likely to have many restrictions, and which are likely to have few or no restrictions. The systematic review of those with no or few restrictions could result in more records being opened to the public at an earlier date.
- 7) Experienced reviewers are eventually promoted or retire and NARA and Presidential Libraries lose their expertise. The tool might accumulate review knowledge so that the knowledge resource is not lost.
- 8) The tool will support PRA and FOIA review decisions for Presidential records, so it would also support review of Federal Records for FOIA exemptions.
- 9) Since most states have Open Record Acts, state records need to be reviewed for access restrictions before release to the public. The technology might be transferred to support archivists performing review of state government records.
- 10) Although the records being considered in this study are unclassified records, the technology might transfer to declassification review.

Purpose

The purpose of this report is to describe the results of an investigation to identify the kinds of knowledge that archivists use to reach their review decisions. Then to analyze some sample records to identify the functions that a tool would need to perform to discriminate Federal, Presidential and personal records, and to identify probable access restrictions.

Scope

In the next section, the Freedom of Information Act (FOIA) and the Presidential Records Act (PRA) are discussed. The distinction is drawn between Federal and Presidential Records and Personal Papers. The paragraphs of the FOIA and PRA that relate to access restrictions are described and the kinds of knowledge needed to determine whether these restrictions apply are discussed. In the third section we show some examples of the kinds of documents to which specific FOIA exemptions and PRA restrictions apply, and discuss the kinds of information that would need to be extracted from them in order to determine whether an access restriction might apply.

2. Knowledge of PRA and FOIA Review

The Freedom of Information Act (FOIA) is a law enacted by Congress in 1966 that provides for access to records maintained by the Executive branch of the U.S. Government. Amendments in 1974 established nine types of information exempt from release. Amendments in 1996 included a requirement that agencies search for records in electronic form [3].

The Presidential Records Act (PRA) of 1978 [4] governs the official records of Presidents and Vice Presidents created or received after January 20, 1981. The PRA changed the legal ownership of the official records of the President from private to federal records.

The Presidential Records Act establishes a process for public access to these records. Specifically, the PRA allows for public access to Presidential records through the Freedom of Information Act (FOIA) beginning five years after the end of the Administration, but allows the President to invoke as many as six specific restrictions to public access for up to twelve years. The PRA also establishes procedures for Congress, courts, and subsequent Administrations to obtain special access to records that remain closed to the public, following a notice period to the former and current Presidents. The PRA also requires that Vice-Presidential records be treated in the same way as Presidential records.

Presidential records are defined as:

Documentary materials, or an reasonably segregable portion thereof, created or received by the President, his immediate staff, or a unit or individual or the Executive Office of the President whose function is to advise and assist the President, in the course of conducting activities which relate to or have and effect upon the carrying out of the constitutional, statutory, or other official or ceremonial duties of the President [4].

Documentary material is defined as:

All books, correspondence, memorandums, documents, papers, pamphlets, works of art, models, pictures, photographs, plats, maps, films, and motion pictures, including, but not limited to, audio, audiovisual, or other electronic or mechanical recordings [4].

Presidential Records do not include agency records. The PRA refers to the definition of agency record as given in the Freedom of Information Act [3].

"(1) "agency" as defined in section 551(1) of this title includes any executive department, military department, Government corporation, Government controlled corporation, or other establishment in the executive branch of the Government (including the Executive Office of the President), or any independent regulatory agency; and

(2) "record" and any other term used in this section in reference to information includes any information that would be an agency record subject to the requirements of this section when maintained by an agency in any format, including an electronic format."

2.1 Distinguishing Presidential Records from Personal Records

In reviewing records, the first judgment an archivist makes is whether or not a document is a Presidential Record or a personal/political record.

The Presidential Records Act defines personal records.

“The term "personal records" means all documentary materials, or any reasonable segregatable portion thereof, of a purely private or nonpublic character which do not relate to or have any effect upon the carrying out of the constitutional, statutory, or other official or ceremonial duties of the President. Such term includes

- a) Diaries, journals, or other personal notes serving as the functional equivalent of a diary or journal which are not prepared or utilized for, or, circulated or communicated in the course of, transacting Government business.

- b) Materials relating to private political associations, and having no relation to or direct effect upon the carrying out of constitutional duties of the President; and
- c) Materials relating exclusively to the President's own election to the office of the Presidency; and materials directly relating to the election of a particular individual or individuals to Federal, State or local office which have no relation to or direct effect upon the carrying out of constitutional, statutory, or other official or ceremonial duties of the President."

The following are examples of documentary materials related to the President's and White House Staff members' personal and political activities.

1. A staff resume for a post-administration job,
2. A job application,
3. A recipe,
4. A Christmas card mailing list,
5. A grocery list,
6. Correspondence with friends expressing sentiments, e.g., "Congratulations," "Happy Birthday," "Thanks for the Gift."
7. If the material is on political party letterhead, it may indicate political/personal material.
8. If the record is from the Chairman of the RNC, it may be political/personal.
9. Abbreviations such as RNC or DNC might indicate political materials.
10. Phrases such as "I was saying to Lee Atwater, ..." "I sent so and so to Lee Atwater," may indicate political/personal material.
11. An RNC staff register would be political/personal material.
12. If a record concerns the President's plans for attending a funeral of a personal friend, rather being a state occasion, it is a personal record.
13. If the material pertains to a personal appointment, rather than an appointment for Presidential business, it may be a personal record.
14. A communication with family members may be personal if it does not involve a constitutional duty.
15. Most of the First Lady's records are presidential, though some that do not have to do with carrying out the constitutional duties of the President may be personal.
16. Polls on political issues by private pollsters, e.g., "How many Republicans in a predominately Democratic District?" "What can we do to get Women or Blacks to vote for you?" are probably personal/political records.

2.2 PRA Exemptions and FOIA exceptions

Paragraph (a) of section 2204 of the PRA states

“Prior to the conclusion of his term of office or last consecutive term of office, as the case may be, the President shall specify durations, not to exceed 12 years, for which access shall be restricted with respect to information, in a Presidential record, within one or more of the following categories:”

This is followed by a list of six restrictions on access to records under the PRA.

Four of the six PRA restrictions are identical to four of the nine FOIA exceptions. Those PRA restrictions that are different from FOIA exemptions are (a)(2), Appointments to Federal Office, and (a)(5), Confidential Advice. The paragraph identifiers a(n) are used for PRA restrictions and the identifiers b(n) for FOIA exemptions. Taken together there are eleven different restrictions.

- a(1), b(1) national security and foreign policy,
- a(2) appointments to Federal offices,
- a(3) b(3) exempted by statute,
- a(4) b(4) confidential commercial information,
- a(5) confidential advice,
- a(6) b(6) personal privacy,
- b(2) personnel rules and practices of an agency
- b(5) deliberative process privilege,
- b(7) law enforcement investigations,
- b(8) financial institution reports, and
- b(9) geological information about wells.

Though the PRA restrictions a(1), a(3), a(4) and a(5) are lifted after twelve years, the corresponding FOIA exemption remains and file continues to be withheld.

One of the FOIA categories, FOIA exemption (b)(5), which protects "inter-agency or interagency memorandums or letters which would not be available by law to a party ... in litigation with the agency", is not available as a basis for restricting access to Presidential Records [44 USC 2204(c)]. However, a Presidential Library may acquire Federal Records with the acquisition of Presidential Records, and so may need to consider that exemption for those records.

In the following sections each of the restrictions/exceptions will be discussed. Conditions under which each PRA restriction/FOIA exception might apply to a record are enumerated. These criteria were acquired from a FOIA workshop [5] and discussions with review archivists. Most of the FOIA exemption guidance was obtained from the US Department of Justice *Freedom of Information Act Guide & Privacy Act Overview* [6] and the Department of Navy's *FOIA: Basic Training* [10]

PRA Restriction a(1) and FOIA Exemption b(1)

This exemption allows for the withholding of information "(A) specifically authorized under criteria established by an Executive order to be kept secret in the interest of national defense or foreign policy and (B) are in fact properly classified pursuant to such Executive order." The relevant Executive Order is E.O. 12958. [9]

With regard to the second conditions of this exemption, EO 12958 § 1.3 & 1.7 specify that: The following security markings, usually at the top and bottom of each page, indicate that a document is classified: Confidential, Secret, Top Secret.

Material marked "Atomic," "Cosmic," and "NATO Restricted" are the NATO equivalents of U.S. classified markings.

Material with administrative markings such as Personal and Confidential, Official Use Only, FOUO, Administratively Confidential, Limit to Official Use Only, Eyes Only, Limited Distribution, or Limited Official Use, although they may be on sensitive documents, do not indicate a national security document.

With regard to condition A of this exemption, EO 12958 § 1.5 specifies that information may not be considered for classification as national security information unless it concerns:

- a. military plans, weapons, or operations.
- b. foreign government information.
- c. intelligence activities (including special activities), intelligence sources or methods, or cryptology.
- d. foreign relations or foreign activities of the United States, including confidential sources.
- e. scientific, technological, or economic matters relating to the national security.
- f. United States Government programs for safeguarding nuclear materials or facilities.
- g. vulnerabilities or capabilities of systems, installations, projects, or plans relating to the national security.

"Foreign Government Information" means:

- a. information provided to the United States Government by a foreign government or governments, an international organization of governments, or any element thereof, with the expectation that the information, the source of the information, or both, are to be held in confidence;
- b. information produced by the United States pursuant to or as a result of a joint arrangement with a foreign government or governments, or an international organization of governments, or any element thereof, requiring that the information, the arrangement, or both, are to be held in confidence; or

- c. information received and treated as "Foreign Government Information" under the terms of a predecessor order.

"Confidential source" means any individual or organization that has provided, or that may reasonably be expected to provide, information to the United States on matters pertaining to the national security with the expectation that the information or relationship, or both, are to be held in confidence.

If a classified document has not been reviewed in the last 2 years, then a researcher can request a mandatory review.

The corpus of records that we are analyzing, and that will be used for testing a review tool, does not contain classified documents. Hence, review of classified documents will not be considered further in the current investigation.

PRA Restriction a(2)

This category restricts from release records "relating to appointments to Federal office."

1. Internal memos that have lists of candidates who are not the final choice should be closed.
2. External recommendations from Federal Judges or congressmen, unless that person is a personal advisor, or has been asked for advice, are not closed.
3. Unsolicited letters of recommendation from organizations including the American Bar Association (ABA) are not closed under a(2).
4. External, unsolicited job applications or letters of recommendation are not closed under a(2), but may be closed under a(6) and b(6).
5. Telephone numbers on a resume are redacted under a(6), b(6).
6. Cue phrases that might indicate a possible restriction include "appointment to Supreme Court," and "appointment of Federal Judge."

President Reagan waived this restriction except for two subcategories of records covered by the restriction.¹

- (a) the document contains clearly confidential personal information, such as a social security number, salary, medical history or other information that would otherwise require it to be restricted under category a(6), or
- (b) the document comes from, or contains any substantive notation thereon from a member of the White House staff or other member of the Administration or other person who should be regarded as an "adviser" to the President [such as a Senator or Congressman who writes to the President as a personal friend or trusted adviser, rather than in his or her official capacity), or would otherwise require it to be restricted under a(5).

¹ Letter from President Reagan to the Archivist of the United States, September 10, 1991.

The following are categories of documents for which President Bush waived his restriction rights under a(2).

- Routine requests for positions;
- Resumes (excepting personal data such as social security and home telephone numbers)
- Routine ministerial materials conveying a final action dealing with appointments to office; [When the final outcome is evident, a formal recommendation can be opened.]
- Announcements of appointments to office and accompanying press releases;
- Routine letters of congratulations, resignations and acceptances of resignations;
- Routine letters of recommendations for positions.

Twelve years after the expiration of the Presidents term in office, this restriction no longer applies. However, representatives of the former and current president currently request to see copies of any record that would have been closed under either or both a(2) and a(5) were the PRA restrictions still applicable. Hence, archivists need to be able to "tag" files that would have been closed under a(2) or a(5) without closing them.

PRA Restriction a(3) and FOIA Exemption b(3)

This category of records is "specifically exempted from disclosure by statute (other than section 552b of this title), provided that such statute (A) requires that the matters be withheld from the public in such a manner as to leave no discretion on the issue, or (B) establishes particular criteria for withholding or refers to particular types of matters to be withheld."

Statutes exempting records from disclosure include, but are not limited to:

1. *Federal Rules of Criminal Procedure* Rule 6(e)(6) exempts matters occurring before a Grand Jury from disclosure. This is often detected by the mention of a Grand Jury.²
2. *Federal Rules of Criminal Procedure* Rule 32(e)(1) governs the disclosure of presentence reports.³
3. 5 USC §7114(b)(4) of the *Civil Service Reform Act* restricts disclosure of materials that constitute guidance, advice, counsel, or training provided for management officials or supervisors, relating to collective bargaining.⁴
4. 7 USC §12, the *Commodity Exchange Act*, restricts disclosure of any data or information concerning or obtained in connection with any pending investigation of any person and any information or data obtained from a foreign futures authority.⁵

² <http://www.law.cornell.edu/rules/frcrmp/Rule6.htm>

³ <http://www.law.cornell.edu/rules/frcrmp/Rule32.htm>

⁴ <http://www4.law.cornell.edu/uscode/5/7114.html>

⁵ <http://www4.law.cornell.edu/uscode/7/12.html>

5. While 8 USC §1202 (f), *Immigration and Naturalization Act*, authorizes the withholding of information pertaining to the issuance or refusal of visas and permits to enter the United States, these are actually restricted under a(6) and b(6).⁶ Such documents are often detected by the occurrence of a passport number.
6. 10 USC §130, the Department of Defense "technical data" statute, restricts the release of unclassified, military critical technology with military or space application that is in the possession of -- or under control of -- the Department of Defense and that may not be exported lawfully without an approval, authorization or license under the Export Administration Regulations (EAR) or the International Traffic in Arms Regulations (ITAR).⁷ DoD Directive 5230.25, "Withholding Unclassified Technical Data from Public Disclosure," implements this law. DoD Directive 5230.24 establishes a distribution marking procedure for technical data. Based on the technical content of the data, one of seven distribution statements is to be applied to indicate the extent of secondary distribution that is permissible without further authorization or approval of the controlling DoD office. Without such markings foreign corporations and others acting on behalf of foreign governments may otherwise file requests for this information under the Freedom of Information Act. A document marked with the first of the seven markings, APPROVED FOR PUBLIC RELEASE. DISTRIBUTION IS UNLIMITED authorizes unlimited distribution, and so can be opened to the public. A document with one of the other six markings should be closed. Such closed documents should be referred to the controlling DoD office for proper disposition.
7. 10 USC §130b restricts the disclosure of duty station address information regarding defense employees in sensitive, routinely deployable, and/or overseas units.⁸
8. 10 USC §130c, Nondisclosure of information: certain sensitive information of foreign governments and international organizations, restricts access to information from a foreign government that the foreign government considers confidential.⁹
9. 10 USC §1102, Confidentiality of Medical Quality Assurance Records: Qualified Immunity for Participants, protects military medical quality assurance records.¹⁰
10. 13 USC §8(b) and 9(a) prohibits the use of population, agriculture and housing information furnished under the Census Act for any purpose other than the statistical purpose for which it was supplied.¹¹
11. 13 USC §301(g), Collection and Publication, of the *Foreign Commerce Act* pertains to shipper's export declarations.¹² (Similar to restriction on disclosure in 50 USC 2411(c)).
12. 15 USC §2055(a)(2) of the *Consumer Product Safety Act* protects from release confidential information such as trade secrets, processes, operations, style of work, or apparatus, or to the identity, confidential statistical data, amount or

⁶ <http://www4.law.cornell.edu/uscode/8/1202.html>

⁷ <http://www4.law.cornell.edu/uscode/10/130.html>

⁸ <http://www4.law.cornell.edu/uscode/10/130b.html>

⁹ <http://www4.law.cornell.edu/uscode/10/130c.html>

¹⁰ <http://www4.law.cornell.edu/uscode/10/1102.html>

¹¹ <http://www4.law.cornell.edu/uscode/13/ch1schI.html>

¹² <http://www4.law.cornell.edu/uscode/13/301.html>

- source of any income, profits, losses, or expenditures of any person, firm, partnership, corporation, or association. This information is also protected under a(4) and b(4).¹³
13. 15 USC §2055(b)(1), *Consumer Product Safety Act*, requires prior notice to manufacturers or labelers of information submitted to the Consumer Product Safety Commission that is to be publicly disclosed that might disclose the identity of the manufacturer or labeler.¹⁴
 14. 18 USC §2510-20 of the *Omnibus Crime Control & Safe Streets Act of 1968* exempts from disclosure the contents of electronic surveillance (wiretap intercepts) conducted in violation of the Act.¹⁵
 15. 19 USC §1677f of the *Tariff Act* restricts disclosure of proprietary information submitted to the Tariff Commission.¹⁶
 16. 22 USC §1461(a), the *Smith-Mundt Act*, limits the dissemination of USIA materials domestically.¹⁷
 17. 22 USC §3104(c), *International Investment Survey Act of 1976*, restricts disclosure of Investment Survey Information.¹⁸
 18. 26 USC §6103 of the *Internal Revenue Code*, governs disclosure of tax return information, but this is actually exempted under a(6) or b(6).¹⁹
 19. 35 USC §122, *Patent Act*, prohibits the disclosure of patent applications and information concerning them.²⁰
 20. *Federal Property and Administrative Services Act of 1949*, as amended, 41 U.S.C. §253b(m), prohibits the release or disclosure of a technical, management, or cost proposal submitted by a contractor in response to the requirements of a solicitation for a competitive proposal.²¹
 21. 42 USC §2161-66, *Atomic Energy Act of 1954, as amended*, prohibits the disclosure of information concerning Department programs that is classified as Restricted Data and Formerly Restricted Data.²²
 22. 45 USC §362(d) of the *Railroad Unemployment Insurance Act* restricts the disclosure of personally identifiable material. [Association of Retired Railroad Workers v. United States Railroad Retirement Bd., 830 F. 2d 331,334 (D.C. Cir. 1987)].²³
 23. 49 USC §40115 requires nondisclosure of information that would prejudice formulation and presentation of positions of United States in international negotiations or adversely affect the competitive position of any air carrier in foreign air transportation.²⁴

¹³ <http://www4.law.cornell.edu/uscode/15/2055.html>

¹⁴ <http://www4.law.cornell.edu/uscode/15/2055.html>

¹⁵ <http://www4.law.cornell.edu/uscode/18/p1ch119.html>

¹⁶ <http://www4.law.cornell.edu/uscode/19/1677f.html>

¹⁷ <http://www4.law.cornell.edu/uscode/22/1461.html>

¹⁸ <http://www4.law.cornell.edu/uscode/22/3104.html>

¹⁹ <http://www4.law.cornell.edu/uscode/26/6103.html>

²⁰ <http://www4.law.cornell.edu/uscode/35/122.html>

²¹ <http://www4.law.cornell.edu/uscode/41/253b.html>

²² <http://www4.law.cornell.edu/uscode/42/ch23dAschXI.html>

²³ <http://www4.law.cornell.edu/uscode/45/362.html>

²⁴ <http://www4.law.cornell.edu/uscode/49/40115.html>

24. 50 USC §402 Notes (P.L. 86-36 sect. 6) exempts the disclosure of the organization or function of the National Security Agency or any information with respect to the activities of the NSA, or of names, titles, salaries or numbers of persons employed by such agency.²⁵
25. 50 USC §403g of the *CIA Act* requires the protection of information about the internal structure of the CIA, including disclosures related to the organizations, functions, names, official titles, salaries, or numbers of personnel employed by the agency.²⁶ Agent's names are restricted based on this statute.
26. 50 USC §403-3 (c)(7) of the *National Security Act* gives the Director of the CIA the responsibility for protecting "intelligence sources and methods" from disclosure.²⁷ Intelligence sources and methods are actually closed under a(1) b(1).
27. 50 USC Appdx. §2411(c)(1) of the *Export Administration Act* governs the disclosure of export licenses and applications.²⁸

The US Department of Justice's publication *Freedom of Information Act Guide & Privacy Act Overview* is useful in determining whether these statutes apply. Another source of statutes that are relevant to FOIA exemption b(3) are the Annual FOIA Reports of Federal Agencies[8]. In FY 2002, under FOIA exemption b(3), agencies withheld information pursuant to a total of 142 different nondisclosure statutes, all of which are specified by agencies in their annual reports.

PRA Restriction a(4) and FOIA Exemption b(4)

This category relates to "trade secrets and commercial or financial information obtained from a person and privileged or confidential." In other words, information submitted by an outside party that is trade secret or proprietary in nature and submitted with an understanding of confidentiality.

This restriction applies to two broad categories of information: (1) trade secrets, and (2) information which is (a) commercial or financial, (b) obtained from a person, and (c) privileged or confidential. A *trade secret* is defined as "a secret commercially valuable plan, formula or process or device that is used for the making, preparing, compounding, or processing of trade commodities that can be said to be the end product of either innovation or substantial effort."

Commercial or financial information is any information that relates to business, trade or commerce. Commercial information can include material submitted by a nonprofit entity. Examples of items regarded as commercial or financial include: business sales statistics, research data, technical designs, customer and supplier lists, profit/loss data, overhead and operating costs, and information on financial condition.

²⁵ <http://www4.law.cornell.edu/uscode/50/402.notes.html>

²⁶ <http://www4.law.cornell.edu/uscode/50/403g.html>

²⁷ <http://www4.law.cornell.edu/uscode/50/403-3.html>

²⁸ <http://www4.law.cornell.edu/uscode/50/appendix-2411.html>

Protection for financial information is not limited to economic data generated solely by corporations or other business entities. It has been held to apply to personal financial information as well.

The term *person* refers to corporation, a financial institution, an insurance company, a state government or an agency of a foreign government. However, it does not apply to the Federal government.

Critical Mass Energy Project v. NRC, 975 F.2d 871 (D.C. Cir. 1992) established a distinction between voluntary and required submissions. Voluntarily submitted information is categorically protected if not customarily disclosed to the public. A submission is voluntary if the submission was not required.²⁹

If a submission was required, it is "confidential" if either of the conditions below (known as the National Parks Test) hold:

1. Impairment Prong: disclosure would impair the Governments ability to obtain necessary information in the future.
2. Competitive Harm Prong: disclosure would cause substantial harm to the competitive interest of the submitter.

Examples of (a)(4), (b)(4) exempt materials [10]:

1. Trade secrets and commercial or financial information contained in loans, bids, contracts, and/or proposals.
2. Statistical data concerning contract performance, income, profit, losses, and/or expenditures.
3. Personal statements in inspections, investigations, or audits that reveal company trade secrets or proprietary information.
4. Financial data in locality wage surveys used to adjust pay schedules.
5. Scientific and manufacturing data submitted with a research grant application or with a report while research is still in progress.
6. Technical or scientific information developed at private expense.
7. Technical or scientific data developed in part with federal funds and private expense if the contractor retains proprietary interest in accordance with 10 USC 2320- 2321 and the Department of Defense Federal Acquisition Regulation Supplement (DFARS). LIMITED RIGHTS.
8. Computer software that is copyrighted under the Copyright Act of 1976.

This exemption does not apply to routine business information provided voluntarily by persons or corporations and information that corporations by law must file with the Securities and Exchange Commission that is publicly available, such as corporate registration statements, periodic reports to shareholders, tender offers, and ownership reports.

²⁹ FOIA Update, Spring 1993, Exemption 4 Under Critical Mass: Step-By-Step Decision-making.
http://www.usdoj.gov/oip/foia_updates/Vol_XIV_2/page4.htm

PRA Restriction a(5)

This restriction applies to "confidential communications requesting or submitting advice, between the President and his advisers, or between such advisers." This includes, but is not limited to, policy or legal advice. It includes all documentary forms containing or requesting advice including final memoranda, draft memoranda, notes from meetings, letters, etc.

The President's advisors include counselors and assistants to the President, Deputy Assistants, Special Assistants to the President, and the Director of Media Affairs. It could include a Senator or Congressman who writes to the President as a personal friend or trusted adviser, rather than in his or her official capacity. It could also include anyone in the Executive Branch providing advice, including interagency groups and committees generating options or advice.

1. Administrative markings e.g., Administratively Confidential, Limited Official Use Only, Eyes Only, while not conclusive, may assist in determining whether the advice falls within the confidential advice restriction.
2. Some advice is not confidential, for example, "We think you should serve chicken," "I don't think we need to meet with them this week."
3. Anything from the Counsel's office that says: "I recommend ...," "I advise...," or "It is my opinion ..." may indicate possible confidential advice. However, this is not a concrete rule. The entire file must be reviewed to determine if there is any confidential advice within.

Both President Reagan and President Bush waived this restriction for some subcategories that would otherwise fall under the restriction. It is informative to see some examples of these waivers, because they are indicative of the kinds of documents to which the exemption might apply, and one might want to have the capability in tool that checks for restrictions to support waiver of these subcategories. The following are examples of categories of documents that could be opened under President Reagan's waiver.

- Informational or factual materials provided to the President or staff member.
- Routine position papers, reports, or studies that examine an issue, but do not advocate a particular decision or convey advice.
- Routine scheduling recommendations and plans that do not include advisory comments concerning the proposed visitor.
- Routine administrative requests for information, studies, or action, as well as information or studies submitted, from a staff member or agency official to another staff member or agency official.
- Memorandums from staff members to other staff members that present proposals on routine administrative matters for approval or disapproval by a staff member.
- Documents reflecting conversations among staff members with other staff members, agency officials, or members of the public concerning administrative or factual matters.

- Request from White House staff members to other staff members or agency officials for information and guidance on preparing routine responses to inquiries on specific subjects, and information received.
- Recommendations for approval of routine bills or resolutions.
- Staffing memorandum worksheets, action tracking worksheets, and similar document cover sheets when these sheets give no substantive information.
- Routine editing of speeches, statements, or reports that do not result in a substantive change in the content of the speech, statement or report.

The following are categories of documents for which President Bush waived his restriction rights under a(5).

- Presidential scheduling files including daily schedules, drafts, and proposals; routine background materials relating to schedules; movement logs, telephone logs; and other routine briefing materials;
- speech-related materials (speech drafts, background, research, clearance process and speech cards);
- draft press releases;
- "fact sheets," including drafts of and background for fact sheets;
- memoranda and reports provided to the President and his staff which are purely informational or factual in content, including position papers, reports and studies that examine an issue;
- memoranda between staff members which are purely informational or factual in content;
- requests and referrals from White House staff to other staff or agency officials for purely factual information and guidance in the preparation of routine responses to inquiries and the information received.
- Memoranda and recommendations regarding routine congressional resolutions;
- Routine memoranda analyzing and/or making recommendations regarding Administration positions on legislation or recommending Presidential approval or disapproval of legislation;
- Routine White House Staffing Memorandums and comments;
- White House Correspondence Tracking Worksheets and routine annotations;
- Background materials, preparatory materials and drafts regarding routine Presidential proclamations and Special Messages;
- Background materials, preparatory materials and drafts regarding declarations, executive orders and other exercises of Presidential legal authority, prepared by the White House Executive Clerk's Office.

Twelve years after the expiration of the Presidents term in office, this restriction no longer applies. However, representatives of the former and current president currently request to see copies of any record that would have been closed under either or both a(2) and a(5) were the PRA restrictions still applicable. Hence, archivists need to be able to "tag" files that would have been closed under a(2) or a(5) without closing them.

PRA Restriction a(6) and FOIA Exemption b(6)

This category of restriction relates to "personnel and medical files and similar files the disclosure of which would constitute a clearly unwarranted invasion of personal privacy."

This exemption has two conditions. First, the information must be "personnel", "medical" or "similar" files. Second the disclosures must constitute "clearly unwarranted invasion of personal privacy."

With regard to the first condition, the information must pertain to a specific individual, not to groups or organizations. The term "similar files" is interpreted broadly. All information that applies to a particular individual meets the first condition for protection. That is to say, any information that pertains to an identifiable person is potentially eligible for protection, regardless of the type of file in which the information is found.

With regard to the second condition, only living individuals, not corporations or associations, possess protectable privacy interests. Where the information is in the public domain, the privacy interest, while not automatically eliminated, may not exist.

Examples of information to be restricted include:³⁰

1. Home addresses, home telephone numbers, and home e-mail addresses,
2. Official duty email addresses and duty telephone numbers for personnel assigned to overseas, sensitive, and/or routinely deployable units,
3. Personnel Files: social security numbers, citizenship, religious affiliation, criminal history (criminal convictions),
4. Marital status (e.g., separated), legitimacy of children,
5. Medical files: medical condition (e.g., AIDS, HIV, gonorrhea), mental condition (e.g., psychologist, psychiatrist),
6. Employment Application data (resumes, SF-171s),
7. Security clearance records,
8. Personnel administrative or disciplinary actions,
9. Welfare payments,
10. Information about unsuccessful applicants for federal employment [6, footnote 209],
11. Information on the personal details of a Federal employee's service is also restricted, e.g., employees home address, reasons for leaving previous positions, contents of an FBI investigation concerning federal judicial appointment or appointment to the Executive Branch.
12. Federal employee's job performance evaluations or appraisals, e.g., Staff ratings such as 1, 2, 3 where 3 is negative, phrases such as "Sam isn't as good as ..." ³¹

³⁰ Items 1-8 are from Ref. 10.

³¹ OPM regulation, 5 C.F.R. § 293.311(a), allows basic information about Federal employees, such as names, present and past position titles, present and past grades, present and past annual salaries, duty stations and position descriptions to be disclosed. However, Section 293.311(b), states that information

13. Salaries of non-Federal employees,
14. Information that is practically obscure (once public but currently hard to locate),
15. External recommendation for appointment to federal office that is negative.
16. Telephone number on a resume that is part of an appointment to Federal Office.
17. Cue phrases indicating possible exemption from release: alcoholic, drunk, intoxicated, affair, mental breakdown, nervous breakdown, idiot, hate, personal.

This exemption does not apply to [6, page 43]:

1. Corporations,
2. Deceased individuals, Public records, where not practically obscure.
3. Identities of FOIA requesters, unless first-party.

While deceased individuals do not have privacy rights, the courts have upheld the concept of "survivor privacy." This holds that release of information about a deceased individual would be so upsetting to family members as to invade their privacy. Exemption b(6) has been used to protect survivor privacy of the family members of the Challenger astronauts.

In order to determine whether an invasion of privacy is warranted or not, the privacy interest in the information must be balanced against the public interest of the disclosure. The public interest that should be considered is limited to that information that will "shed light on an agency's performance of its statutory duties."

FOIA Exemption b(2)

This category is "related solely to the internal personnel rules and practices of an agency," which if released, would circumvent how an agency does business.

There are two subcategories of records [10]:

High 2: Substantial internal matters, the disclosure of which would risk circumvention of a statute or agency regulation, operating rules, guidelines, and manuals for investigators, auditors, and/ or examiners; examination questions and answers used for training, employment, and/ or promotion; security classification guides.

Low 2: internal matters of a relatively trivial nature, e.g., parking facility rules, lunch hour rules, sick leave policy, file numbers, mail routing stamps, initials, and data processing notations.

The information in the Low 2 subcategory may be disclosed. Among the kinds of records that must be withheld in the first subcategory (High 2) are:

described in paragraph (a) need not be released if it would "otherwise be protected from mandatory disclosure under an exemption of the FOIA." http://www.opm.gov/cfr/vol_1/5cfr293.htm

1. General guidelines for conducting investigation.
2. Law enforcement manuals setting out guidelines for criminal investigations, such as when to pursue an investigation, prison security techniques, and hijacker profiles.
3. Information that would reveal the nature and extent of a particular investigation.
4. Agreements between state and federal agencies concerning when to exchange information relevant to potential tax law violations.
5. Computer codes and sensitive computer programs used by law enforcement agencies.
6. Hiring plans and job applicant testing materials.
7. Examination questions and answers used for training, employment, and/or promotion.
8. Vulnerability assessments, such as computer security programs, critical systems, facilities, stockpiles, security plans, and training manuals.
9. Sensitive administrative notations, such as confidential sources numbers.
10. Identifying information relating to sources or undercover agents.
11. Information regarding agency audit or examination guidelines, such as cut off criteria. Operating rules and manuals for auditors and or examiners.
12. Security classification guidelines.

FOIA Exemption b(5)

This category relates to "inter-agency or intra-agency memorandums or letters which would not be available by law to a party other than an agency in litigation with the agency." Exemption 5 encompasses the governmental deliberative process privilege, as well as the attorney-client and attorney work-product privileges. *See, e.g., Coastal States Gas Corp. v. DOE*, 617 F.2d 854 (D.C. Cir. 1980). The deliberative process privilege shields from mandatory disclosure documents that are "predecisional" and "deliberative," *i.e.*, that were created during consideration of a proposed action and that were part of a decision making process. *Darci L. Rock*, 13 DOE ¶ 80,102 (1985). This FOIA exemption only applies to Federal records.

Examples: [from Ref. 10]

1. Non-factual parts of After-Action Reports and Lessons Learned containing evaluations, opinions, suggestions, and/or recommendations.
2. Advice, suggestions, or evaluations of consultants, boards, committees, groups, or commissions.
3. Speculative information the acquisition/disposition of real estate or facilities where release would provide unfair advantage.
4. Trade secret, confidential research, or commercial information owned by the government.
5. Information prepared for anticipated administrative or litigation proceedings.
6. Information pertaining to an attorney-client relationship or to attorney work product (covers factual material/no temporal limit).
7. Planning, programming, and budgetary information.

8. Facts inextricably intertwined with opinion and recommendation.

FOIA Exemption b(7)

This category relates to "records or information compiled for law enforcement purposes, but only to the extent that the production of such law enforcement records or information:

- a) Could reasonably be expected to interfere with enforcement proceedings,
- b) Would deprive a person of a right to a fair trial or an impartial adjudication,
- c) Could reasonably be expected to constitute an unwarranted invasion of personal privacy,
- d) Could reasonably be expected to disclose the identity of a confidential source, including a State, local, or foreign agency or authority or any private institution which furnished information on a confidential basis, and, in the case of a record or information compiled by a criminal law enforcement authority in the course of a criminal investigation or by an agency conducting a lawful national security intelligence investigation, information furnished by a confidential source,
- e) Would disclose techniques and procedures for law enforcement investigations or prosecutions, or would disclose guidelines for law enforcement investigations or prosecutions if such disclosure could reasonably be expected to risk circumvention of the law, or
- f) Could reasonably be expected to endanger the life or physical safety of any individual."

Information protected under this exemption must meet two conditions. First, the information must have been compiled for law enforcement purposes. The reviewer must be able to identify what law is to be enforced or what the purpose of the compilation is. Second, the disclosure of such must threaten one of the harms found in the exemptions six subparts.

Under the first condition, law enforcement includes civil or criminal statutes, as well as statutes authorizing administrative proceedings. Any record of a criminal law enforcement agency, if even arguably related to law enforcement should presumptively be protected. Records compiled to enforce state law as well as federal law should be protected. Law enforcement information includes (1) investigations of attorneys' professional conduct, (2) background security investigations by government units, such as FBI investigations of applicants for federal employment or CIA background investigations, (3) law enforcement manuals, and (4) law enforcement program oversight reports.

Under the second condition, exemption b (7)(a) protects current or prospective law enforcement proceedings, the disclosure of which could reasonably be expected to interfere with those law enforcement proceedings. Proceedings include law enforcement investigations, prosecutions, administrative and regulatory proceedings (of a law enforcement nature), and applies not only to federal law enforcement proceedings, but

state and foreign as well. It also includes material of an Independent Counsel investigation. This exemption can be invoked only for so long as the proceeding is pending.

Exemption b(7)(b) protects information the disclosure of which would deprive an individual of the right to a fair trial or impartial adjudication.

Exemption b(7)(c) is similar to exception b(6), but protects law enforcement records--release of which could reasonably be expected to cause an unwarranted invasion of personal privacy. The analysis of these records is similar to that for PRA restriction a(6) and FOIA exception b(6).

Whether information falls under this exemption requires a two-step analysis. First, there must be a privacy issue at stake.

- The mention of an individual's name or other information that would be sufficient to identify an individual in a law enforcement file is likely to constitute an invasion of privacy.
- Federal, state and local law enforcement personnel have a privacy interest in having references to them in investigation files protected, as do individuals who provide information to law enforcement agencies.
- The names of witnesses, their home and business addresses, and their telephone numbers are protectable.

Secondly, in order to determine whether an invasion of privacy is warranted or not, the privacy interest in the information must be balanced against the public interest of the disclosure. The public interest that should be considered is limited to that information that will "shed light on an agency's performance of its statutory duties."

Exemption b(7)(d) protects the identities of confidential sources and in the case of a criminal investigation conducted by a criminal law enforcement authority or a lawful national security intelligence investigation, all information provided by those sources. The exemption applies to intelligence sources, criminal sources, state, local and foreign agencies and to private institutions (e.g., credit bureaus) that furnished information on a confidential basis. The exemption applies even if the identity of the source is known, the source is anonymous, the source is dead, the source has testified, or the source may be called to testify. Sources include citizens providing unsolicited allegations of misconduct. This exemption applies to intelligence sources of the US Secret Service.

Exemption b(7)(e) applies to techniques or procedures for law enforcement investigations or prosecutions or guidelines for such, if disclosure of the guidelines could reasonably be expected to risk circumvention of the law. It includes:

- Methods of law enforcement such as stakeouts,
- The FBI record regarding the relocation of witnesses under the Witness Relocation Program, because disclosure would jeopardize the effectiveness of the program,

- Leads, spares (tape backups of computerized data, spare patrol cars left in reserve, or an Emergency Operations Center (EOC) in another location that is used only if a power failure or natural disaster renders the primary communications facility inoperable),
- Movements in diagrams or schedules,
- Methods of the US Secret Service (USS).

Exemption b(7)(f) protects information from disclosure that would endanger the life or physical safety of any individual. It applies to law enforcement personnel such as FBI agents, or other law enforcement personnel, such as the US Secret Service, as well as information that might lead to their identities, where their safety would be endangered by disclosure. It is also applied to records containing information about protecting the life of the president and/or the first lady. In addition, courts have recently held that it may be used to protect the lives of citizens. (*Living Rivers, Inc. v. US Bureau of Reclamation*, 272 F. Supplement, 2d 1313; D. Utah 2003)

FOIA Exemption b(8)

This category relates to information "contained in or related to examination, operating, or condition reports prepared by, on behalf of, or for the use of an agency responsible for the regulation or supervision of financial institutions."

Stock exchanges, banks, and savings and loans are considered financial institutions. Documents covered by this exemption include bank examination reports and related records, and reports examining bank compliance with consumer laws and regulations.

There are two major purposes for this exemption, first, to protect the security of financial institutions by withholding from the public reports that contain frank evaluations of a bank's stability and secondly, to promote cooperation and communication between employees and bank examiners [10].

Presidential Library reviewers seldom if ever encounter these kinds of information in Presidential Records.

FOIA Exemption b(9)

This category relates to "geological and geophysical information and data, including maps, concerning wells." One court has held that this exemption applies only to "well information of a technical or scientific nature." [10]

Presidential Library reviewers seldom if ever encounter these kinds of information in Presidential Records.

3. Initial Analysis of Some Sample Documents

In the previous section, many of the conditions under which archivists might conclude that a record or a portion of a record were restricted from release to the public were enumerated. In this section, we show document examples from some of the categories of access restriction and analyze them to determine the kinds of knowledge needed to interpret them and to determine whether a PRA restriction or FOIA exemption might apply. We also consider how to formally represent the knowledge characterizing an access restriction so that it could be automatically applied by a tool to check for access restrictions.

Personal Records

In Fig. 2 is shown a digital image of a letter to Mary G. Roebing from Barbara Bush.³² The paper record copy of this letter would most likely be found in the files of one of the assistants to the First Lady. An electronic copy of the letter might also be found in the PC files of the same assistant.

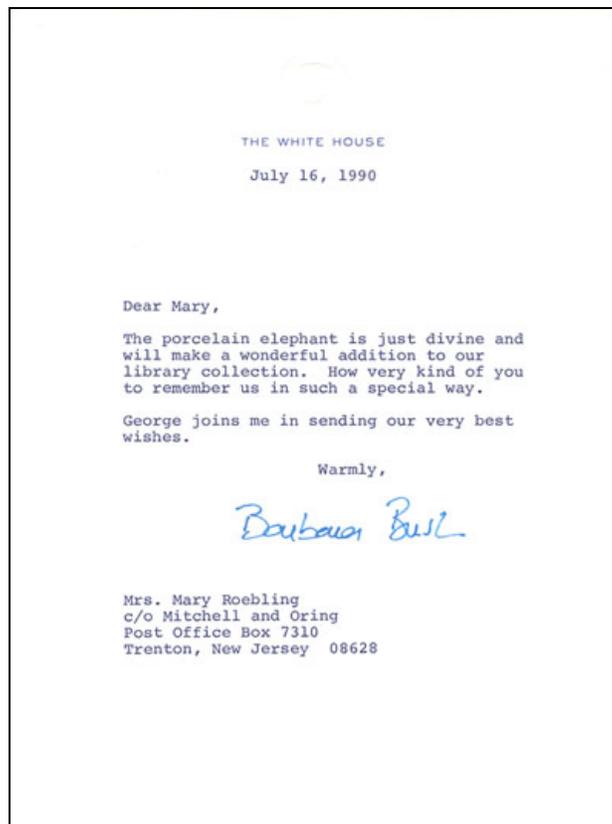


Figure 2. An Example of the First Lady's Personal Correspondence, 1990

³² The source of this letter is the Special Collections and University Archives of Rutgers University/Mary G. Roebing Papers. "Mary G. Roebing was appointed to government positions by U.S. presidents from Franklin Roosevelt to George Bush, and she maintained close contact with many of them and their families."The digital image is from the Women in Leadership Database. <http://www.scc.rutgers.edu/wild/>

As pointed out earlier, most of the First Lady's records are presidential, not personal. A copy of this letter in the files of the First Lady's Office or a digital copy of the letter would be a personal record. The research question is: What kinds of information from the document and what kinds of knowledge about personal records would an automated tool need to reach that conclusion?

First, there is the context in which the copy of the letter is filed. At the time an archivist is reviewing records, the office in which it is included is known and is a part of the metadata about the record. In this case the office would most likely be "First Lady's Office - Correspondence."

Second, it is necessary to recognize that the document is a letter. To recognize that a document is a letter, it must be possible to automatically identify such information as dates, person's names, greetings, organization names, postal addresses and location names. The form of letters, as opposed to memoranda, press releases, speeches, and many other document types, must also be known.

Figure 3 shows a manually marked up copy of the letter shown in Figure 2. Extensible Markup Language (XML) tags such as <date> </date> are used to bracket information.

```
<date>July 16, 1990</date>
<greeting>Dear</greeting><persons-given-name>Mary,</persons-given-name>
<para>
The porcelain elephant is just divine and
will make a wonderful addition to our
library collection. How very kind of you
to remember us in such a special way.
</para>
<para>
<persons-given-name>George</persons-given-name> joins me sending our very best
wishes.
</para>
<formula-of-respect>Warmly,</formula-of-respect>
<persons-name>Barbara Bush</persons-name>
<job title></job title>
<persons-name>Mrs. Mary Roebling</persons-name>, <title></title>
<organization-name>c/o Mitchell and Oring
</organization-name>
<postal-address>Post Office Box 7310</postal-address>
<location><city>Trenton</city>, <state>New Jersey</state></location>
<zipcode>08628</zipcode>
```

Figure 3. XML Tagged Text File.

The letter in Fig 2 was printed on White House stationery and signed by the author. The digital copies of letters printed on stationery may not include the letterhead or the signature. If "The White House" were included in the digital copy, it might have been tagged as <office-building>. If First Lady Barbara Bush's name had been printed on the letter, it might be tagged as <persons-name>.

The XML tags are indicative of the kinds of information that need to be automatically extracted from the office electronic files. *Information extraction* (IE) is a procedure that selects, extracts and combines data from text in order to produce structured information. Information extraction tools can be used to identify location names, person's names, organization names, position titles, greetings, formulas of respect, dates, numeric expressions (zip codes, social security numbers), and other semantically important information in the documents [7].

To extract information from documents represented in proprietary file formats it is necessary to use a file reader that extracts just the ASCII text and not the file header or the proprietary text-formatting markup for bold, italics, font sizes, etc. We are using Stellent's Outside-In tools for converting files from legacy file formats to ASCII text.³³

Markup and style sheet languages such as SGML, HTML, XML and CSS and XSL were originally developed to support the creation of documents and the interpretation of documents so marked up. They were not designed to support the automated markup of legacy electronic records. However, we are investigating whether it is possible to create templates that summarize the essential elements of the form of different types of documents using the XML notation. Figure 4 shows a possible template for White House letters.

```
<date></date>
<greeting></greeting><persons-name></persons-name>
<para></para>*
<formula-of-respect></formula-of-respect>
<persons-name></persons-name>
<job title></job title>
<persons-name></persons-name>
<title></title>
<organization-name></organization-name>
<postal-address></postal-address>
<location><city></city>, <state></state></location> <zipcode></zipcode>
```

Figure 4. Template for the Documentary Form of White House Letters.

Templates will be constructed for other document types found in our experimental collection of Bush Presidential Records, e.g., memoranda, agenda, and press releases. Then the automatically marked up file will be compared to these templates to determine the most likely document type. The classification of the document type may need a probabilistic element as there can be slight variations in documentary form.

Knowing that the record is a letter enables determination of whom the letter is from and to whom it is addressed. The letter's use of first names, e.g., Mary, George, rather than a formal form of address, e.g., Mrs. or the President, are indicative of informal correspondence, but not necessarily of a personal record. The informal formula of

³³ <http://www.stellent.com/groups/mkt/documents/system/topnav.hcst>

respect, i.e., Warmly, rather than a formal formula of respect, e.g., Sincerely, is also indicative of an informal letter, but not necessarily a personal record. The lack of use of a job title with the signature also indicates the informal nature of the letter.

The evidence that this record is a personal record is primarily that it is correspondence with a friend expressing sentiments, "How very kind of you to remember us in such a special way." The First Lady is thanking Mary Roebling for a gift. However, the term "gift" is not in the text, and the phrase "thank you" is not in the text. A gift giving and thanks giving scenario (script) is needed to make the inference that the First Lady is thanking Mary Roebling for a gift.

The verbs thank, greet, welcome, apologize, condole, congratulate are examples of speech acts. In particular, they are illocutionary acts, an act performed by a speaker by virtue of the utterance having been made, defined with respect to speaker's purpose. Illocutionary acts are realized as Performatives, direct speech acts, e.g., "Thank you", or as indirect illocutionary acts, e.g., "How very kind of you to remember us in such a special way." Indirect speech acts are difficult to interpret.

An archivist would look at the entire folder contents, and might find a copy of a letter from Mary Roebling to President George Bush and/or First Lady Barbara Bush in which she described the gift. The combination of letters would be easier to interpret as it sets the context for the Letter from the First Lady to Mary Roebling.

Based on additional examples of the First Lady's personal correspondence, rules could be developed to recognize the First Lady's personal correspondence. Examples of the First Lady's records that involved carrying out ceremonial duties related to the President's constitutional responsibilities could be used to construct rules for discriminating those records that might be Presidential records.

Presidential Records with a(2) and/or a(5) Restrictions

Fig. 5 shows a copy of a document that has been opened by the Bush Presidential Library.³⁴ Access to it would have been restricted but for the waiver by former President Bush of specific subcases of the PRA a(2).³⁵

³⁴ Bush Presidential Library, President Bush Records-White House Staff and Office Files, Counsels Office, Lee Liberman Nominee Files, Leighton, Ronald B. (FOIA 98-0205-F).

³⁵ In a January 6, 1997 letter from President Bush to the National Archivist, the President outlined materials that he waived from restriction under PRA a(2). Included in his list were "routine ministerial materials conveying a final action dealing with appointments to office."

THE WHITE HOUSE
WASHINGTON

April 27, 1992

MEMORANDUM FOR THE PRESIDENT

FROM: C. BOYDEN GRAY
COUNSEL TO THE PRESIDENT

SUBJECT: Nomination of Ronald B. Leighton to be United States
District Judge for the Western District of Washington.

The Attorney General, Samuel K. Skinner, Nicholas E. Calio, C. Boyden Gray, and Constance Horner recommend that you nominate Ronald B. Leighton to be United States District Judge for the Western District of Washington. Since 1976, Mr. Leighton has been affiliated with the law firm of Gordon, Thomas, Honeywell, Malanca, Peterson & Daheim as an associate and then as a partner. He received a B.A. degree, cum laude, (1973) from Whitworth College and a J.D. degree (1976) from the University of California, Hastings College of Law. Mr. Leighton is 40 years old, married, has two children and resides in Tacoma, Washington. The American Bar Association has reviewed Mr. Leighton's qualifications and found him to be qualified for this position.

RECOMMENDATION

That you call Mr. Leighton (Telephone recommendation attached at Tab A.).

That you sign the nomination at Tab B.

Figure 5. Record of a Nomination to Federal Judgeship.

That the document is likely to involve a nomination can be determined from the context of the document, that is, it is in "Lee Liberman's Nominee Files." To determine that this document might be subject to restriction a(2), appointments to federal office, it must be determined that the subject is an appointment or nomination to federal office. The terms "nomination" and "nominate" are indicative. The name of the Federal Office, "United States District Judge," is also indicative.

To determine that this document might be subject to restriction a(5), confidential advice, it must be checked that the advice is to the President by one of his advisors. One needs to have an authority list of advisors, which would include "Counsel to the President."

This document is a variant of the general form of an internal memorandum. A Memorandum for the President has the following documentary form.

<office-building>The White House</office-building>
<location>Washington</location>
<date></date>
MEMORANDUM FOR THE PRESIDENT
FROM:<name></name>
<subject>SUBJECT:</subject>
<para></para>+

This pattern can be matched against a copy of the document from which information has been extracted and which has been marked up. Person's names can be determined by information extraction. Having determined that the document is a memorandum to the President, one is able to determine whether it is from an advisor to the President.

While the terms "recommend" and "recommendation" indicate that the document is advice to the President, it is not "confidential advice." It represents the final decision in the process of considering a nomination for an appointment to a federal office.

The verbs in the text such as "nominate" and the nominalized verbs such as "nomination" can also be determined by information extraction. A case frame for the verb "nominate" is "agent nominates object goal." If the "goal" is a federal office then the record is about an appointment to federal office. One of the rules that might be used to recognize that this document was about an appointment to federal office is:

If a record contains the verb nominate or appoint, and
the [direct] object of the verb is a person's name, and
the goal case [indirect object] is a federal office,
Then the record is about an appointment to Federal Office.

To determine that the record is in a subcategory for which President Bush waived the a(2) restriction, that is, "Routine ministerial materials conveying a final action dealing with appointments to office," one needs a script (Scenario Template) for the deliberative appointment process that begins with the resignation (or death) of a Federal Judge, considers a list of candidates, recommendations, and concludes with the nomination forwarded to the U.S. Senate.

Records with b(6) and b(7) Exemptions

Figure 6 shows a "simulated" summary of a law enforcement report on a domestic disturbance at on-base family housing. All information in the report is fictional and is taken from reference 11, Appendix 11. The text in italics is information that would need to be redacted.

Incident Report

At 2140, the law enforcement desk, extension 222-3456, took an anonymous call that reported a disturbance at 1234 Basestreet, quarters allegedly occupied by two military members. SRA Patrolman arrived on the scene at 2155. SRA Patrolman met Nora Neighbor, who was very agitated. Because she feared her neighbors would retaliate against her if they knew she reported their fight, she asked that her name not be released before she would talk. After she was promised her identity would remain anonymous, she stated: "I heard cursing and heard furniture and dishes breaking. They fight all the time. I've seen Betty Battle with a black eye, and I also saw Bob Battle with bruises the day after they had their last fight, last Saturday night. This time, there was a tremendous crash; I heard a man scream "My Lord NO!" then I saw Betty Battle come out of the house with dark stains on her clothes—she got into her car and drove away. I could see this really well, because the streetlight is right between our houses; I'm the wife of their NCOIC. If only Nick, my husband, was here now, he'd know what to do! I haven't heard anything from Bob Battle."

Figure 6. Sample Incident Report.

The name of the law enforcement investigator, e.g., *SRA Patrolman*, is usually withheld under FOIA exemptions 6 and 7(C). The names of witnesses, e.g., *Nora Neighbor*, are also withheld under FOIA exemptions 6 and 7(C). When *Nora Neighbor* was promised that her identify would remain anonymous and she made a statement, she became a confidential informant. Data that could identify her, and in some cases, the information she related, should be withheld from release under FOIA exemptions 6, 7(C) and 7(D). *Betty Battle's* name should be redacted under exemptions 6 and 7(C), unless she is the FOIA requestor. *Bob Battle's* name should also be redacted under exemptions 6 and 7(C), unless he is the requestor.

That this document was "information compiled for law enforcement purposes," and thus possibly contains information exempt from release under exemption b(7) could be determined from the context of the document, e.g., record series "Investigations", and from the document type, "Incident Report."

Person's names can be determined by information extraction. The person's roles, investigator, witness, confidential informant, can be determined by identifying the verbs in the text and using the case frames for the verbs to determine which agent "arrived" at the scene, the investigator, and which agent reported the incident, the witness. The agent who was promised that their identify would remain anonymous, is the confidential source. This information can be used to fill in the slots in the Scenario Template for an Incident Report.

Incident Report
Report by: anonymous
Time: 2140
Date:
Location: 1234 BaseStreet
Investigator: SrAPatolman
Witness: Nora Neighbor
Confidential Source: Nora Neighbor
Participants: Betty Battle, Bob Battle

Figure 7. Scenario Template for an Incident Report.

A rule such as the following could be formulated for exemption 7(C).

If record series is "Law Enforcement Investigation" and
Document type is "Incident Report" and
Investigator of Incident Report is X,
Then X should be redacted and exemption is b(6), b(7)(C).

There would be similar rules for witnesses and confidential sources.

Presidential Records that have no Restrictions

There are Presidential Records in the Bush Presidential Library that can be opened for access by the public because they were released to the public at the time they were created, or they were public at the time they were received by a White House staff member. For instance, electronic copies of White House press releases, newspaper articles, press feeds, speeches, and transcriptions of press conferences would not have any restrictions.

The White House Press Office distributed daily press releases to members of the press. The press releases contained transcripts of Presidential speeches and remarks, announcements of Presidential actions (such as nominations, executive orders and bill signings), fact sheets, texts of proclamations, statements by the press secretary, press pool reports, schedules, and transcripts of press briefings. Fig. 8 shows a press release that has been opened by the Bush Presidential Library.³⁶

³⁶ Bush Presidential Library, Bush Presidential Records, White House Press Office, Press Office Files - Daily Press Releases 1989-1993.

THE WHITE HOUSE

Office of the Press Secretary

For Immediate Release October 2, 1990

STATEMENT BY THE PRESS SECRETARY

The President met for 30 minutes today with UNITA leader Jonas Savimbi. The President pledged our continued support for Dr. Savimbi and his movement. The President believes that the steadfastness of that support is a key factor in pushing the current negotiations to a successful conclusion. The President expressed the hope that these negotiations will lead to peace and national reconciliation in Angola. The President pledged that we are prepared to do what we can to promote this goal.

#

Figure 8. Example of a Press Release - Statement by the Press Secretary.

That this record has no restrictions can be determined in part by its context, that is, the name of the office or staff member in whose files the record is contained. In this case it is the "White House Press Office." In addition knowing its document type, "Press Release," would increase the likelihood that it had been released to the press. The form of a White House press release can be represented as follows.

```
<office-building>The White House</office-building>
<organization-name>Office of the Press Secretary</organization-name>
For Immediate Release
<date></date>
<title></title>
<para></para>+
```

An additional condition on this document being releasable to the public is that it be the final version of the press release. It is unlikely, but possible, that a draft of a press release could contain information that was not released to the public and might have access restrictions.

4. Summary and Conclusions

The FOIA and PRA statutes were reviewed and a Freedom of Information Act and Privacy Act Workshop was attended. The Department of Justice *FOIA Guide and Privacy Act Overview* and NARA's internal guidance for PRA and FOIA review were examined. Experienced archivists described to us their experience in reviewing Presidential records for FOIA exemptions and PRA restrictions. A reviewer's decision as to whether or not a particular record (or parts of a record) is exempt from release under

one of the FOIA exemptions is an inference from the context, the documentary form and content of a record. The knowledge and guidance obtained was discussed in section 2 of this report.

Some examples of Personal and Presidential records were analyzed to determine the kinds of knowledge needed to interpret them and to determine whether a PRA restriction or FOIA exemption might apply. The knowledge of access restrictions needs to be represented in such a form that automated methods can be used to analyze and interpret a digital record and conclude whether an access restriction applies. It was explained how rules could be used to formally represent the knowledge characterizing access restrictions.

The objective of this study is to acquire and represent the knowledge that is needed to review electronic records and to use this knowledge in a knowledge-based system to support automatic checking of electronic records for possible access restrictions. The effort and cost to represent the knowledge and test its use is a potential bottleneck to the realization of such a capability. Hence, we intend to investigate rule induction or case-based learning from sample documents and passages known to have particular access restrictions.

"Review is not a science. It is an art."³⁷ Each Administration is unique and archivists become experts on the nuances of a former President's preferences, his family's and staff's names, policies formulated and legislation passed during his term in office. Furthermore, the guidance that archivists are provided does not cover all cases and they often have to rely on their common sense in order to make a decision as to whether documents do or do not fall within these restrictions.

The names of the President's family, White House staff names and titles, congressional bills and legislation, names of heads of state, and names of Presidential advisors will be collected and represented in a knowledge base. Common sense reasoning technology will be explored in combination with rule-based or case-based reasoning as a means of interpreting electronic records to make judgments as to whether an access restriction applies.

In a related research task, we are defining the form of the types of documents that occur in the White House staff member's electronic files. Experiments will be conducted to determine whether these patterns can be matched against automatically marked up documents to determine an electronic record's document type. Automatic identification of the document type of an electronic record enables identification of the parties involved and sometimes the subject of the document.

Records need to be collected that represent each of PRA restrictions and FOIA exemptions as well as records that were Personal Misfiled Records, but have been donated to the Presidential Library and are now open. The sample should also include some records that were not subject to PRA restrictions or FOIA exemptions.

³⁷ Personal conversation with Nancy Smith, NARA.

Because we do not actually have collections of unprocessed Presidential, Federal and Personal records for use in our laboratory, we are constructing a suite of such records that have been previously withdrawn because of PRA restrictions or FOIA exemptions, but are now open to the public. They are open to the public because they are: donated personal records, declassified records, records that would be restricted but for waiver of restriction by a former President, and records opened after expiration of a restriction.

References

1. Executive Classification of Information—Security Classification Problems Involving Exemption (b)(1) of the Freedom of Information Act (5 U.S.C. 552), Third Report by the Committee on Government Operations, U.S. Congress, House of Representatives, Committee on Government Operations, House Report No. 93-221, 93d Congress, 1st Session, 1973, p. 50.
2. W. Underwood, M. Hayslett-Keck and S. Laib. The Archival Processing Tool: User's Guide, Version 2.06. PERPOS Technical Report ITTL/CSITD 03-2, Information Technology and Telecommunications Laboratory, Georgia Tech Research Institute, Revised December, 2003
3. Freedom of Information Act, 5 U.S.C. § 552, As Amended By Public Law No. 104-231, 110 Stat. 3048.
http://www.usdoj.gov/oip/foia_updates/Vol_XVII_4/page2.htm
4. Presidential Records Act, (44 USC 2201-2207)
<http://www4.law.cornell.edu/uscode/44/ch22.html>
5. *Freedom of Information Act and Privacy Act Workshop, Participant Guide*, 2002 Graduate School USDA.
6. *Freedom of Information Act Guide & Privacy Act Overview*, May 2002 Edition. US Department of Justice. <http://www.usdoj.gov/oip/oip.html>
7. R. Simpson and E. Whitaker. Information Extraction Technologies and Tools, Information Technology and Telecommunications Laboratory, Georgia Tech Research Institute, Forthcoming.
8. US Department of Justice. Office of Information and Privacy. Summary of Annual FOIA Reports for Fiscal Year 2002.
<http://www.usdoj.gov/oip/foiapost/2003foiapost31.htm>
9. Executive Order 12958 - Classified National Security Information, April 17, 1995. <http://www.dss.mil/seclib/eo12958.htm>
10. DON PA/FOIA Branch, Freedom of Information Act: Basic Training,
<http://foia.navy.mil/basic-training-presentation.pdf>
11. DoD Regulation 5400.7/Air Force Supplement. DoD Freedom of Information Act Program. http://www.e-publishing.af.mil/pubfiles/af/dod/dodr5400.7_afsup1/dodr5400.7_afsup1.pdf