



OCIO Exchange Record Keeping Process Improvements As of April 2007

Executive Office of the President • Office of Administration • DRAFT Discussion Document



Exchange / Outlook Emails In Storage for EOP

Executive Office of the President • Office of Administration

DRAFT Discussion Document

Total email messages captured for the EOP enterprise.

Year	Exchange / Outlook	Lotus Notes	Total
2002	1,667,669	13,817,141	15,484,810
2003	17,733,388	12,713,888	30,447,276
2004	34,816,617	3,287,227	38,103,844
2005	56,117,413	49,746	56,167,159
2006	48,425,982	10,393	48,436,375
Total	158,761,069	29,878,395	188,639,464

- 2002 – 2005 statistics presented are primarily from 2005 audit. Updated statistics currently being compiled.
- Decline from 2005 to 2006 can be attributed to reduction of inbound SPAM.

DRAFT



2005 analysis

Executive Office of the President • Office of Administration

DRAFT Discussion Document

	WHO	OVP
Results of 2005 Analysis - Latest data and message counts are still being compiled and will be included once complete – DRAFT		
Total Days with Zero - 50 captured message counts:	12	24
Total Days with Low captured message counts:	16	8

WHO - 28 days (**12** + **16**) missing or with low message counts out of 1477 total days or **1.9%**

OVP - 32 days (**24** + **8**) missing or with low message counts out of 1337 total days or **2.3%**

DRAFT



Listing of Red Days

Executive Office of the President • Office of Administration

DRAFT Discussion Document

WHO	OVP
2003:	2003:
12/17	9/12
12/20 and 12/21	10/01, 10/02, 10/03, 10/05
2004:	2004:
1/9, 1/10, 1/11	1/29, 1/30, 1/31
1/18	2/7, 2/8
1/29	
2/1	
2/7, 2/8	
2/15	
2005:	2005:
n/a	2/15, 2/16, 2/17
	5/21, 5/22, 5/23
	11/17, 11/18, 11/19, 11/20
2006:	2006:
n/a	3/26
	8/29

Note: Aug-Oct 2005 (6 Week Timeframe), emails from several components were lumped into OA storage

DRAFT



Highlights of Issues & Improvements

Executive Office of the President • Office of Administration

DRAFT Discussion Document

2002	2003	2004	2005	2006	2007
<p>Effort outlined to move EOP from Notes to a world class solution leveraging Microsoft's Exchange Outlook for emails.</p> <p>Highly complex environment. Growing pains of new processes and technology. Industry lacked technology maturity for comprehensive record keeping solutions.</p> <p>Outlook was new to EOP as a solution and did not include an automated solution for audit processes.</p> <p>Process was created to move journal entries manually to storage, also called a "PST". Inventory tracked manually.</p> <p>The audit process was created and refined as learning came in during the conversion from Lotus Notes to Exchange Outlook.</p> <p>The manual processes handled as part of overall operations processes. Staffed mainly by contractors.</p>	<p>EOP engaged Microsoft to write custom code called FindIt to assist with searches.</p> <p>All audit, inventory, and storage processes continued to be manual.</p> <p>The manual processes handled as part of overall operations processes. Staffed mainly by contractors.</p>	<p>Manual, complex processes continue.</p> <p>EOP continued to identify opportunities to improve people, process and technology to improve record keeping and management.</p> <p>Installed technology to improve the automation of storage creation. New product called MailAttender is installed to automatically move journals to PSTs.</p> <p>Standard approaches and processes were improved including a database created to automate inventory tracking.</p> <p>The manual processes handled as part of overall operations processes. Staffed mainly by contractors.</p>	<p>August 2005, a configuration change merged component emails, including WHIO and OVP, into OA storage. Those emails are not lost and are in OA storage. This took place over a 6 week period.</p> <p>Additional EOP investments made in people, process and technology to improve record keeping and management.</p> <p>Processes and procedures were modified. People are trained on the new improvements. This included a process for Daily Check and Independent Verification and Validation (IV&V) processes for newly created PST files.</p> <p>Technology improvements were made to the inventory process with two custom built products. The program (CMDHFI) was created to scan PST files to obtain dates and quantities of emails within the PST files.</p> <p>PST file management database (PFMS database) was created to contain the results of the CMDHFI scans.</p> <p>Performed data analysis to determine potential gaps in data from 2002 - 2005. Developed a "red / yellow" day report by component.</p> <p>A search team charter is established to clearly define the expanded roles, responsibilities, and enhanced processes for contractors and staff to follow. The team has other responsibilities due to budget constraints but they are trained for searches and have responsibility to validate email inventories.</p>	<p>Audit initiated 4Q, 2006 to review the complex environment. Many procedures entail manually intensive tasks. (Audit still underway through May 2007)</p> <p>Key learnings: the manual processes were applied inconsistently during contractor / staff turnovers.</p> <ul style="list-style-type: none"> <input type="checkbox"/> train staff so the intended procedures are followed consistently and accurately. <input type="checkbox"/> the manual nature of maintaining the PFMS (PST inventory) database requires continuous communication and training. <input type="checkbox"/> leverage CMDHFI and PFMS to insure manual processes are working. 	<p>As of March 2007:</p> <p>PFMS database has been updated and can support search requests.</p> <p>New information on PST files are uploaded to the database regularly.</p> <p>After audit is complete, next steps:</p> <ul style="list-style-type: none"> <input type="checkbox"/> New procedures and documentation to support database updates are under review, then will be submitted for approval. <input type="checkbox"/> Team is reviewing database and file information in order to improve search efficiencies. <input type="checkbox"/> Team completes validating database results against inventory to ensure accurate information. <input type="checkbox"/> Review the COTS product used to create PST files (MailAttender). <input type="checkbox"/> Work with COTS product vendor's technical support to ensure most efficient configuration is applied to Journal mailboxes. <input type="checkbox"/> Develop new procedures and documentation to support daily and yearly PST creation processes. <input type="checkbox"/> Develop enhanced IV&V process to ensure consistent product configurations.
<p>Migration from Notes to Outlook 2002 - 2004</p>					

*Personal Storage Table (PST) is a Microsoft Outlook data file called a personal folder and has a .pst file extension.

DRAFT



Records Management

Executive Office of the President • Office of Administration

DRAFT Discussion Document

- Over 2/3 of the EOP environment does not fall under the PRA.
 - According to Federal General Records schedules, these records captured by the system *may be considered expired and are no longer needed.*
 - Each EOP component is responsible for its own records schedules.
 - OA has been working with the components to determine the definitive requirements.
- ECRMS project was discontinued due to concerns that the project would not fully meet records management and operational requirements.
 - A new, fully compliant solution is under review by OA to fulfill these requirements.

DRAFT

6



Current Limitations

Executive Office of the President • Office of Administration

DRAFT Discussion Document

- Current PST file management and search processes are resource-intensive (people, process, and technology).
 - Manual and complex steps.
 - All email is stored whether or not it is an active record. (no “Record” vs. “Non-record” function as with Lotus Notes)
 - Exchange Outlook does not provide end-to-end technology for overall records management, records schedules, inventory, audit, and searches. Project underway to provide improved technology to support processes.
 - Resource constrained. No dedicated staff.
 - Time consuming for both management and searching of PSTs.

DRAFT

7



Planned Work to Improve People, Process and Technology

Executive Office of the President • Office of Administration

DRAFT Discussion Document

- An extensive effort was initiated by OA in mid December 2006 to re-analyze OCIO exchange record keeping processes and current captured email inventories.
- This process is still in progress and is expected to be completed by May 2007. Some delays in the project occurred due to resources assigned to high priority search requests.
- Audit deliverables include:
 - Accurate inventory and statistics of current inventory of captured email*
 - Validation of Record Keeping process and procedures
 - Validation of Search and Verification processes and procedures
- Work with OA General Counsel's office to address issues:
 - Resolve duplicate PST files (mark copies as inactive so they do not require being searched), results files, combined Federal/Presidential files.
- Implement improved database. *The review process is still in progress and is expected to be completed by May 2007.*
 - Consolidate and verify PST and IV&V information.
 - Solution based on documented requirements.

*Note: Average daily counts equal 50k-60k emails to EOP per day

DRAFT



Market Research

Executive Office of the President • Office of Administration

DRAFT Discussion Document

- “Meanwhile, storage management specialist BridgeHead Software released survey results last week indicating that 28 percent of UK firms do not archive data such as email. The survey of over 300 IT directors also revealed that **15 percent of firms were unable to estimate how long it would take to retrieve a file lost three months ago, and two percent admitted they would not be able to find the file.**” Excerpt from “Archiving smartens up,” by James Murray.
- “Records management and document management vendors have been merging to provide more-comprehensive offerings from larger companies...**Because few document management and records management vendors have robust e-mail archiving today,** customers should view point solutions from e-mail archiving vendors as tactical purchases with the goal of integrating the records into a broader solution in the future...Although most records management systems have a way to selectively pull e-mail, **most have been late to recognize the need for a more-automatic way to ensure the capture of all required e-mail** message and related attachments. Thus, capturing e-mail messages has been poorly handled, if at all.” Gartner, 2003: “Compliance Legislation Hits the Storage and Records Management Vendors”.
- “Federal Rules of Civil Procedure...Rule 37(F), a new safe harbor provision, acknowledges that **in the course of routine IT operations information is sometimes altered or lost.** However, the organizations must make documented good faith efforts to preserve any electronic data that may be relevant to the case.” Excerpt from “E-Mail Discovery as of Dec. 1, 2006: Why IT Should Care”, Lynn Haber, September 6, 2006.
- “Searching through volumes of backup tapes for requested emails is **extremely costly and time-consuming, with no guarantee** all requested records are recoverable...backup systems are not designed for information discovery, where responding to a request means finding specific emails and attachments based upon the context (e.g., date, sender, recipient) and content (e.g., keywords, subject line, attachments) of the information requested.” Excerpt from “Email Discovery: Worst-Case Scenarios vs. Best-Practices”, by Nick Mehta, Senior Director, Product Management, Symantec Corp

DRAFT



Market Research

Executive Office of the President • Office of Administration

DRAFT Discussion Document

GARTNER: A Primer on the Legal Discovery of Electronically Stored Documents by Debra Logan, John Bace

- “Through 2007, more than half of IT organizations and in-house legal departments will lack the people and the appropriate skills to handle electronic discovery requirements (0.8 probability). Through 2007, more than half of organizations will have inadequate (or poorly enforced) e-mail and records retention policies, which will complicate electronic discovery processes and drive up costs, or put the outcome of litigation in jeopardy (0.8 probability).”
- “As electronic evidence comes to predominate the legal landscape, more legal professionals will need to have deep technical knowledge to understand what is offered and how it can help them solve evidential challenges. IT professionals must also understand the process of evidence collection so that they can instruct their companies and help make decisions regarding service and product selection. We are already seeing the emergence of a job role category called “legal IT”...
- How to prepare your for the inevitability of ongoing discovery requests. These include:
 - *Defining relevant data.* This often comes down to identifying individuals who may have been involved in the case and the dates of their involvement.
 - *Determining data’s location.* This means determining what information systems the individuals involved in the case touched or had access to, and the location of the relevant records on those systems.
 - *Preserving data.* This has many nuances and can prove difficult. **Many corporations have opted for the “save everything” strategy, which Gartner believes is not a workable long-term solution.**
 - *Setting up a system for reviewing and producing data.* Even the largest corporations do not always have the expertise or resources to do this, and this work often ends up being done by outside counsel. This is the area in which the IT department can be of greatest assistance: helping legal professionals choose the best technical solutions and controlling costs.
- *The Electronic Discovery Reference Model (EDRM)* — In May 2005, a group of 29 electronic discovery products and service providers, law firms, and corporate counsels launched the Electronic Discovery Reference Model (EDRM) Project. The group’s intention is to develop a reference model describing the concepts and relationships that comprise the e-discovery process. The EDRM aims to provide a common, flexible and extensible framework for the development, selection, evaluation, and use of e-discovery products and services. The group is planning a public comment period to solicit feedback on the draft EDRM in early 2006. The primary result of the project will be a documented reference model placed in the public domain by May 2006.

DRAFT

10



Market Research

Executive Office of the President • Office of Administration

DRAFT Discussion Document

GARTNER: How to Manage Your E-Mail to Avoid Liability by Maureen Caplan Grey, Mark R. Gilbert

- In theory, the record retention and categorization rules that an enterprise applies to its paper documents applies equally to e-mail messages. . . . however, the casual nature with which an e-mail message is sent — along with its largely unstructured format — seems to defy many records management practices. In today's highly regulated and litigious world, an enterprise that fails to manage e-mail as a record is testing fate. Most enterprises have no experience in managing e-mail as anything other than an IT operation. Not surprisingly, executives typically view the "problem" as e-mail retention, so they put the burden of fixing it on the IS organization. This shortsighted approach promotes false starts, wastes time and does little to decrease liability.*
- ...it is imperative to define exactly what is a record. The enterprise should not defer developing a compliance process while waiting for the "right" strategic path to emerge.*
- ...those in regulated industries and the public sector — should take these steps:*
 - *Assign an owner to the records management issue. Involve the legal, human resources, compliance and records management organizations (and, possibly, key contacts in the business units) in developing the business policies for e-records. These must align with records management policies for physical records.*
 - *Update the electronic communications policy (often called the e-mail policy) to include retention scheduling. Paper-retention policies already in place should drive policies around e-document retention in terms of what types of content should be retained and for how long.*
 - *Involve the IS organization in transferring the business policies into technology policies for e-mail, IM and other messaging applications.*
 - *Investigate technologies that will meet near-term tactical, as well as strategic, needs. For example, consider active archiving, document management or records management applications. One trend is to invest in an active archiving system that addresses operational (growing e-mail storage) and business (legal) issues but that covers only email and, increasingly, IM. Enterprises can use this approach until they've developed a strategy for managing all types of digital and physical records.*
 - *Train users on records handling. For true e-records management to work, the end user's ability to categorize a document quickly and correctly is key. If a user takes more than five seconds to handle the record, the implementation will likely fail.*

DRAFT

11



Email Recovery Options

Executive Office of the President • Office of Administration

DRAFT Discussion Document

- A high level approach was drafted in 2005 with options to potentially recover email not captured by the system.
 - OA has backup tapes that could be reviewed for potential restoration of data from Oct 2003 - Present.
 - Plan was utilized to recover non-captured messages as it related to a DOJ investigation.
 - Decision was not made on course of action for other dates determined by the 2005 analysis.
- Approach – 3 Options:
 1. Restore file servers used to host PST files during that time period in question to search for and verify if PST files may have been misplaced. (Due to space limitations, OCIO had to constantly manually move files to make room for newly captured email. Processes were not fully defined and was heavily dependant on specific personnel.)
 - Estimated 24 days per time period that needed to be restored.
 2. Restore Journal servers used to capture messages during the time period in question. Enter the journal mailbox to determine if mail messages are present for the time period. If present, proceed with manually creating PST files from the data.
 - Estimated 24 days per time period that needed to be restored.
 3. Restore Exchange servers that hosted the inboxes of users during the time period in question. Enter each individual mailbox and manually create PST files from the data.
 - Estimated timeframe unknown (Several months per time period that needed to be restored). Extensive effort required to accomplish this.

DRAFT

12



Investments Enhanced Search Capabilities

Executive Office of the President • Office of Administration

DRAFT Discussion Document

Resources		Estimated Cost
<p>Upgrade Search Capacity Current dedicated staff (3) CFTE for 24 X 5. Goal is to move to a 24 X 7 operation with increased concurrent search capability and capacity. Current allocated staff (3).</p>	<p>Four (4) Additional CFTE Additional search hardware/software</p>	<p>\$625K</p>
<p>Upgrade & Maintain Tracking Inventory Database Validation of database results against inventory to ensure accurate message count information and to improve search efficiencies. Current allocated staff (0).</p>	<p>One-half database Senior Database Administrator (\$200K) One-quarter database administrator (\$65K) Automation tools (\$35K)</p>	<p>\$300K</p>
<p>Upgrade Messaging Stability Additional staff are required to maintain operational stability and respond to new requirements. Current allocated staff (4) CFTE.</p>	<p>Four (4) Additional CFTE</p>	<p>\$800K</p>
<p>Independent Verification and Validation The independent verification and validation (IV&V) process seeks to ensure that the results provided to the White House Counsel in response to an email search request are accurate, comprehensive, and valid. Current allocated staff (0).</p>	<p>Seven (7) Additional CFTE</p>	<p>\$840K</p>

DRAFT



Investments Backup Tape Restoration

Executive Office of the President • Office of Administration

DRAFT Discussion Document

Task	Resources	Estimated Cost
Restore Data From Backup Tapes Estimated effort and cost to recover email not captured by the system.	Independent Forensic Contractor Approximately \$500K (Setup) Approximately \$50K for each day requiring restoration Estimated cost based on restoration of 60 component days	\$3.5M

Does not include:

- Full blown forensic review
- OA GC costs
- Procurement costs
- Conducting a search of the restored tape

Caveats:

- Tapes might be corrupted
- Key dates might not be retrievable
- No guaranteed success

DRAFT



Appendix – Additional Details

Executive Office of the President • Office of Administration

DRAFT Discussion Document

- 2002-2007 issues & improvements
- Current process overview

DRAFT

15



Issues and Improvements 2002 – 2003

Executive Office of the President • Office of Administration

DRAFT Discussion Document

- In late 2002:
 - EOP components started a methodical transition from Lotus Notes to Microsoft Exchange and Microsoft Outlook.
 - Exchange Outlook did not include an automated record keeping solution. Various vendor products existed but none were designed specifically with government records keeping, inventory, and audit already built into the product. A decision was made to proceed with the transition to Exchange Outlook with the focus on manual processes for record keeping and searching.
 - Record keeping of Exchange emails achieved through Exchange Journal Mailboxes and storage of messaging information within PST* files.
 - Journal information **manually** copied to PST files.
 - Budget and resource constraints meant staff juggled the priorities of day to day operations of management, inventory, and audit with other operational priorities.
- In Fall 2003:
 - OCIO worked with Microsoft to develop an application, FindIt, to provide text search capabilities of PST files in preparation for a pending, high-profile search.

* Personal Storage Table (PST) is a Microsoft Outlook data file called a personal folder and has a .pst file extension.

DRAFT

16



Issues and Improvements 2004

Executive Office of the President • Office of Administration

DRAFT Discussion Document

- In 2004:
 - Migration to Exchange from Lotus Notes continued.
 - Concept of operations and standard operating procedures developed for creation and inventory of PST files.
 - Database developed to inventory PST files.
 - Process of moving files from Journal Mailboxes to PST files automated with program named Mail Attender.

DRAFT

17



Issues and Improvements 2005

Executive Office of the President • Office of Administration

DRAFT Discussion Document

- In 2005:
 - In August, a network configuration change unexpectedly impacted Exchange Journal Mailboxes. This discovery triggered a more in-depth review of the entire PST creation and management process.
 - A Message Storage Team charter was established, consisting of members throughout OCIO.
 - To identify issues with current PST creation and inventory and management.
 - To address issues identified.
 - To provide enhanced or new processes and procedures.
 - To train the staff on the new processes and procedures.
 - To establish updated management operating rhythms around inventory and audit.

DRAFT

18



Issues and Improvements 2005

Executive Office of the President • Office of Administration

DRAFT Discussion Document

This was an extensive undertaking to determine the scope of the issue and determine where OCIO had gaps in the process. Team results:

- Determined issues, including the following findings:
 - PST file creation suffered from:
 - Incomplete processes; operating routines did not reflect the complex environment leading to inconsistent management of the process
 - Limited written documentation.
 - Personnel-dependent vs. process dependent.
 - Email messages not properly captured on certain days. Captured messages are stored by component; if messages were not captured for 10 components on a single calendar day, that single calendar day would register as 10 days.*
 - Analysis indicates information was not captured on 473 out of 13,552 days (3%).
 - Of the 473, WHO information was not captured on 12 days and OVP information was not captured on 24 days.
 - OCIO developed a three-pronged proposal to address this issue, which was presented to White House Counsel in November 2005.
 - Recommendations provided included:
 - OCIO would perform tape restores to:
 - (1) determine if PST files were on legacy servers that may have been removed,
 - (2) recover Journal servers from that time period and create PST files,
 - (3) restore individual mailboxes and create new PST files.
 - Each of the three processes are expensive and an extremely resource-intensive undertaking.**
 - Because of a software configuration error, 230 PST files combine Federal messages with Presidential messages for a six-week period between August 2005 – October 2005. These messages are not lost.

* statistics based primarily from 2005 audit. Updated statistics currently being compiled.

DRAFT



Issues and Improvements 2005

Executive Office of the President • Office of Administration

DRAFT Discussion Document

- Developed new processes and procedures to address issues.
- Developed new tools.
 - Program named CMDFI that scans PST files to obtain dates and quantities of emails within the file.
 - PST file management database (PFMS database) to contain the results of the CMDFI scans.
- Daily check and Independent Verification and Validation (IV&V) processes for newly created PST files.
- Performed data analysis to try and determine potential gaps in data.

DRAFT

20



Issues and Improvements 4Q 2006 Internal Review

Executive Office of the President • Office of Administration

DRAFT Discussion Document

- Many procedures entail manually intensive tasks.
- Processes and procedures established in 2005 successfully used in searches during the first half of this year.
- Opportunities / Issues:
 - Due to the manual nature, high contractor and staff turnover in key roles negatively impacted procedures and processes.
 - Budget and resource constraints meant staff juggled the priorities of day to day operations of management, inventory, and audit with responding to legal searches.
 - Due to competing demands and the complexity of the processes the PFMS (PST inventory) database was not completely maintained; database provides incomplete inventory (this is being addressed).
 - Staff training improvement initiative identified as critical to ensure staff interprets their tasks correctly.

DRAFT

21



Issues and Improvements 2007

Executive Office of the President • Office of Administration

DRAFT Discussion Document

As of March 2007:

- ❑ PFMS database has been updated to support search requests.
 - New information on PST files are uploaded to the database regularly.
 - ❑ New procedures and documentation to support database updates are under review, then will be submitted for approval.
 - Team is reviewing database and file information in order to improve search efficiencies.
 - Team is still validating database results against inventory to ensure accurate message count information.
- ❑ The COTS product used to create PST files is under review.
 - Team is working with COTS product vendor's technical support to ensure most efficient configuration is applied to Journal mailboxes.
 - Team is developing new procedures and documentation to support daily and yearly PST creation processes.
 - IV&V process is under development to ensure consistent product configurations.

DRAFT

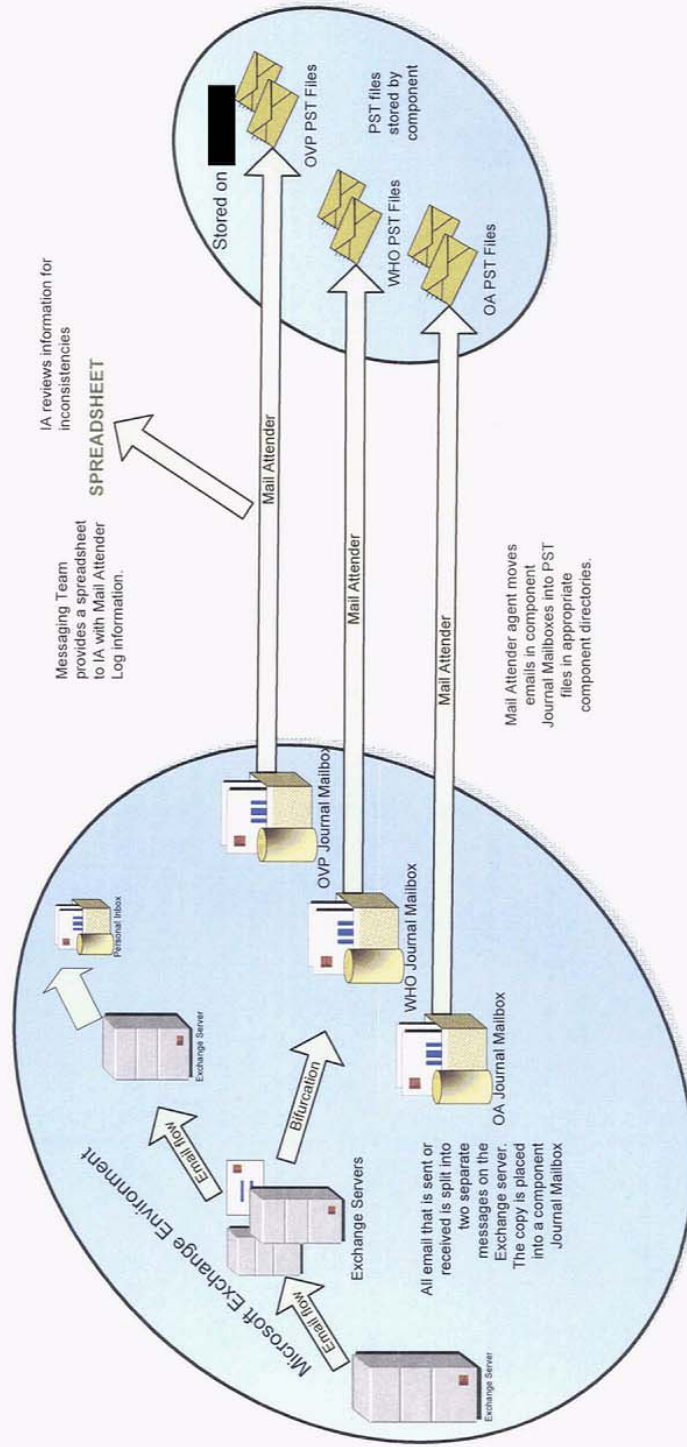
22



Current PST Creation Process

Executive Office of the President • Office of Administration

DRAFT Discussion Document



- ❑ PST files are created daily by the program Mail Attender.
- ❑ Operational support team provides statistical information on PST file creation for audit controls.

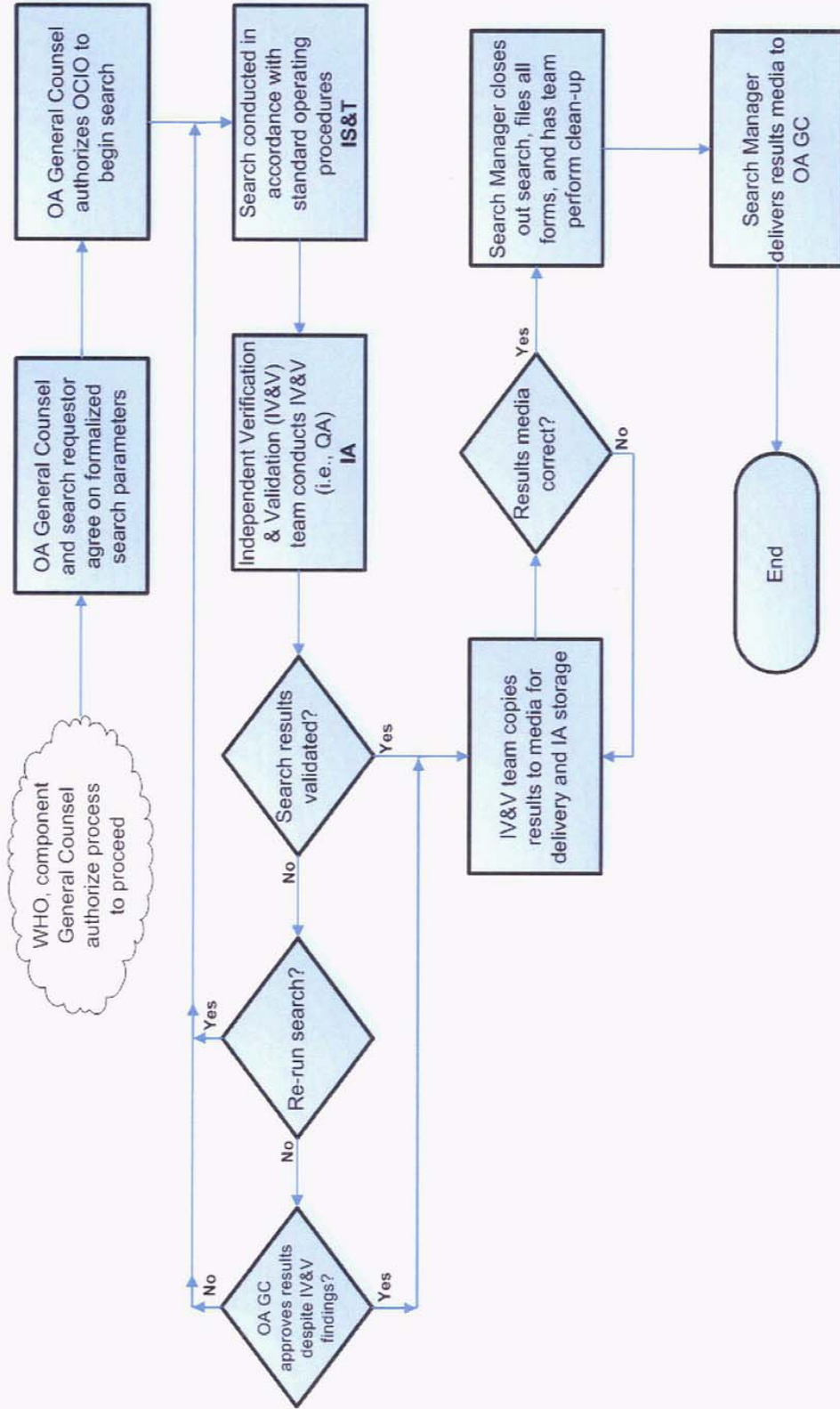
DRAFT



Current Search Process

Executive Office of the President • Office of Administration

DRAFT Discussion Document



DRAFT