

# NATIONAL ARCHIVES AND RECORDS ADMINISTRATION



## CONTINGENCY PLAN TO ASSURE ACCEPTANCE AND INGESTION OF PRESIDENTIAL ELECTRONIC RECORDS

for the

NATIONAL ARCHIVES AND  
RECORDS ADMINISTRATION

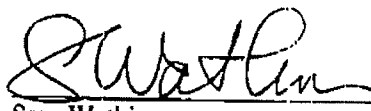
Final  
November 6, 2008

FINAL

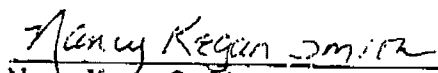
# CONTINGENCY PLAN FOR PRESIDENTIAL RECORDS (CPR)

## Signature Page

I recommend approval of the Contingency Plan for Presidential Records (CPR).


  
\_\_\_\_\_  
Sam Watkins,  
Transition Liaison for Electronic Records

11/6/2008  
Date

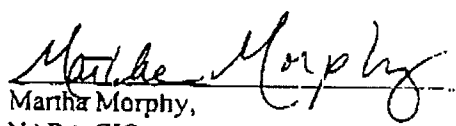
  
\_\_\_\_\_  
Nancy Kegan Smith,  
Transition Liaison for Presidential Records

11/7/08  
Date

Approved,

  
\_\_\_\_\_  
Sharon Fawcett,  
Assistant Archivist for Presidential Libraries

11-7-2008  
Date

  
\_\_\_\_\_  
Martha Morphy,  
NARA CIO

11/7/08  
Date



FINAL

## Table of Contents

Signature Page .....	i
Document Change Control Sheet.....	ii
1.0 BACKGROUND .....	1
2.0 SCOPE .....	1
3.0 PRIORITIES .....	2
4.0 PLANS .....	2
5.0 INVOCATION .....	4

FINAL  
**CONTINGENCY PLAN FOR  
BUSH PRESIDENTIAL ELECTRONIC RECORDS (CPRR)**

**1.0 Background**

On January 20, 2009, legal custody of all the Presidential records of the George W. Bush administration (Bush 43) will transfer to the National Archives and Records Administration (NARA). NARA has the responsibility to secure these records, and to search them in response to special access requirements immediately after the end of the administration. To that end, NARA is working with the staff of the Executive Office of the President (EOP) to prepare for the transition of these records to NARA's physical custody by that date.

Electronic records present unique problems to NARA in meeting the physical transfer and search and access requirements. While the move of non-electronic Presidential records is a massive and important undertaking for NARA, NARA has done this successfully for many Administration's and proven its ability to retrieve the records quickly and accurately. The transfer of Presidential electronic records requires substantial knowledge of the records' formats, structures, and nature to prepare NARA systems for ingestion into a system that can provide the search of and access to the records. The EOP Electronic Records Archive (EOP ERA) is the primary and final intended repository for the Bush 43 Presidential records. The system is currently being deployed by Lockheed Martin under the ERA contract will receive these records for search and access by the staff of the George W. Bush Presidential Library.

There are several risks associated with the transfer of the Bush records into the ERA system. The first is that ERA is a new system and there are IT components that are being developed specifically for Presidential Records. Secondly, because records are being created by the White House staff until the end of the administration, there is always uncertainty regarding an exact volume of data that NARA will receive from the EOP. While NARA has experience through the last three (3) Presidential administrations accepting and making available these records, this transition presents unique challenges because the agency is faced with 50-100 times the volume of electronic materials and formats not previously dealt with. Lastly, although NARA has been working closely with the EOP to understand the types of data that will be transferred, there is always a possibility that some electronic records may be overlooked and NARA will have to do some analysis after January 20, 2009 to prepare the records for transfer into EOP ERA. Because of these new challenges, over a year ago, NARA prepared an alternate plan for ingestion of these records into the Presidential Electronic Records Library (PERL), the IT system that contains the presidential electronic records of the Reagan, Bush 41, and Clinton administrations. Because the current PERL system architecture was not scalable to actually support the volume of records that are expected from the current Presidential administration, the alternative plan called for the acceptance of specific datasets by NARA, and later ingestion into ERA. The methodology for the alternative is described below.

**2.0 Scope**

This contingency plan provides specific action plans for those unclassified datasets that NARA determines cannot be ingested and made available for search through EOP ERA within the time

FINAL

constraints required. It is organized by those datasets so that portions of the plan may be invoked without affecting other transition activities, i.e. other dataset ingestions. The datasets addressed in this plan are critical to the agency's ability to respond to special access requests from the current and former administrations, the Congress, and the courts as required by the Presidential Records Act (PRA). These priorities are determined by the Office of Presidential Libraries (NL), based on experience with the records of previous administrations, and apply both to the importance of ingestion into ERA and the invocation of the contingency plan. Classified datasets are not covered by this contingency plan.

### 3.0 Priorities

The datasets listed below represent the first Bush administration Presidential electronic records that need to be available for search and access.

1. Records Management System (RMS). Operated for the White House Office of Records Management, RMS provides the key index to the bulk of Presidential textual records received by NARA. The availability of this system is critical to the ability of the Library staff to find and retrieve records handled by the White House Office of Records Management.
2. White House Photo System (PHOTO). Operated for the White House Photo Office, the PHOTO system contains all the Presidential photographs (in three (3) different formats/resolutions) taken during the Bush 43 administration, and all the information (metadata) regarding the photographs. It is critical to the library staff's ability to respond to requests from many stakeholders. Access to this dataset is one of the most active immediately after the end of the administration.
3. Presidential Diary System (Diary). The Diary system contains a daily record of the President's activities such as travel, meetings, and phone calls.
4. WARDS. WARDS is a search and access database used by the White House Office of Records Management and includes data from both the WAVES and ACR systems operated by the U.S. Secret Service. These systems control and record access to the White House complex by both White House and EOP staff (ACR), and by those with appointments within the complex (WAVES).
5. Email (ARMS and MS Xchange). The email of the administration constitutes one of the largest and most heavily used bodies of electronic records to come out of the administration.

### 4.0 Plans

NARA is confident that the ERA EOP system as currently developed by LMC can ingest all of the systems listed above except for the RMS system and the PHOTO System. These two systems present problems with ingestion into the EOP ERA system because they are currently in proprietary systems. These two systems, therefore, require a NARA contingency plan. Additionally, while NARA is confident that it has the ability to ingest the Bush emails, the

FINAL

timing of this ingestion is dependent on the completion of any ongoing restoration project for the MS Exchange emails undertaken by the EOP.

1. The RMS presents unique difficulties effecting a timely ingest into EOP ERA. The system consists of an implementation of a customized older version of Documentum running on Oracle, with image files (including copies of scanned records) incorporated as objects in the database. The volume of the system, while manageable, is not insignificant. The difficulties for ingestion into EOP ERA arise from the fact that the system operates on an older (5.2) version of Documentum that does not support the 5015.2 compliant records management features, and the problems inherent in trying to preserve the links between the descriptive metadata and the images of the correspondence in the system without the use of the Documentum product.

NARA and the EOP are currently testing an extract for ingestion into ERA. If NARA cannot effectively ingest either the metadata or the image files from RMS, we will need to acquire the software and equipment to stand up an instance of Documentum on Oracle and take a full copy of the customized dataset from the White House Office of Records Management. NARA has accomplished this type of transfer during the transition of previous administrations, although not for this voluminous or complicated system of records.

2. The PHOTO system which is the largest and uses proprietary photo management software called MerlinOne, running on Microsoft SQL as the database engine. There are two (2) main challenges to ingesting the data into ERA. The first is to establish the linkages between the metadata and the photo images outside the Merlin product. NARA and the EOP have been working on this for a couple of months and have completed an extract of the data that has been accepted by NARA. The second problem is the sheer volume of the data. After establishing the linkage, it will take weeks to extract, transform, and load the data into EOP ERA. The length of this process and the fact that the database may be in active use up until inauguration day, will necessitate that the data be transported in segments for ingestion into EOP ERA in order to ensure that the data is available in the system immediately after January 20.

The large volume of data in the PHOTO system limits our options if we cannot ingest the data into EOP ERA in time. The only viable option is to acquire a MerlinOne license and stand up the system with all the data at NARA until we can accomplish the extract, transform, and load process, and make the data available in ERA. The large volume of data limits our options here as well. It will require several days or weeks just to copy the data onto another system, so NARA and the EOP are working together to ensure that the system is operational, with all the data in place. In effect, NARA would stand up the same system that was used at the White House to support the system.

3. The email (both ARMS and MS Xchange files) is another dataset that NARA has a high level of confidence will ingest into EOP ERA without difficulty. The email formats have been a primary focus of the EOP ERA effort to ingest the EOP data since the project began. The EOP instance of ERA is ready to accept data and NARA and the EOP are testing data formats and ingestion processes now. The volume of the data (consisting of hundreds of millions of messages) makes the development of an alternative strategy very

FINAL

difficult. The EOP cannot provide NARA with the storage media or system to service the records, and the copy process alone will require weeks to complete. NARA would need to develop an entirely separate system, using neither the technology of the existing PERL system (which is much too small), nor the technology of the EOP repository (which lacks necessary controls). This substantial development and acquisition effort does not seem to be efficacious because NARA has a high level of confidence that the data can be ingested into EOP ERA and the fact that some delay (up to a few weeks) in delivering a completed email dataset to the library users is probably acceptable.

**5.0 Invocation**

The decision to invoke a contingency plan of action will be made by the NARA CIO and Assistant Archivist for Presidential Libraries based on the recommendation of Transition Liaison for electronic records (Sam Watkins, NHM) and the Records Owner (Nancy Kegan Smith, NLMS). The Transition Liaison for electronic records will also consult with the Acquisitions director, ERA management, and key personnel in the EOP before recommending the invocation of a contingency plan of action.

NARA is developing project timelines with key dates of performance that will serve as triggers to initiate contingency actions.