CEOS

Working Group on Information Systems and Services
Data Stewardship Interest Group

Introduction

This Committee on Earth Observation Satellites (CEOS) Working Group on Information Systems and Services (WGISS) paper attempts to capture elements involved in preserving Earth Observation (EO) data for extended periods of time. The phrase '*long-term*' is often used in relation to the practice of archiving EO data because of their inherent value for revealing change over time. The concept of long-term is subjective, however, which is not good when dealing with the preservation of records that cannot be reproduced. The Consultative Committee for Space Data Systems (CCSDS) has defined *long-term* in a manner that is both understandable and practical for science data managers. Such a definition can be applied in practice to help ensure that EO data remain viable and usable.

> A period of time long enough for there to be concern about the impacts of changing technologies, including support for new media and data formats, and of a changing user community, on the information being held in a repository. This period extends into the indefinite future.
>
> (source: http://public.ccsds.org/publications/archive/650x0b1.pdf)

Continuity Planning and Vital Records

All U.S. federal agencies are required to have **Continuity Of Operations Plans** (COOP) in place to ensure that "Primary Mission Essential Functions (PMEFs) continue to be performed during a wide range of emergencies, including localized acts of nature, accidents and technological or attack-related emergencies." While these plans exist, the effectiveness of them, if exercised, has been questioned. The portion of these plans that ties most to long-term archive strategies, and one that CEOS WGISS can relate to is the **Vital Records** portion.

(source: http://www.fema.gov/about/org/ncp/coop/index.shtm)

The U.S. Geological Survey (USGS) defines Vital Records as:

> Those records, regardless of format (i.e., paper, photo, database, magnetic tape), that must be protected in the event of an emergency or disaster because the records are essential or critical for continuation of functions to the office or to the USGS and cannot be lost even in the most drastic situation in which access to records are lost for a prolonged period of time.

The USGS goes on to state that Vital Records are those that will be needed in time

periods ranging anywhere from a few minutes to 24 hours after a disaster to get an office or an agency up and running again. They are records that, if lost or destroyed, would be both costly and time consuming to recreate - if they can be recreated at all. They can be active (currently used by the office) or inactive (in storage). Both the USGS and the U.S. National Archive and Records Administration (NARA) consider EO data to be Vital Records.

The identification and protection of Vital Records is crucial as it allows an office or agency to:

• Minimize the disruption of normal business operations after an emergency.
• Minimize the economic impact of the disruption.
• Provide for rapid and smooth restoration of services.
• Comply with legal and regulatory requirements.

Source: http://internal.usgs.gov/gio/irm/vital.html (note this is an internal site)

The U.S. NARA further recommends that agencies have a program centered upon the Vital Records under their care. The guidance states:

> Vital Records Program is a strategic and methodical approach to the identification, selection and protection of that agency information which supports mission critical activities and would be necessary for those activities to continue to function during a crisis, emergency or disaster event.

> Vital Records mean those essential agency records that are needed in order to continue to meet operational responsibilities under national security emergencies or other emergency or disaster conditions (emergency operating records) or to protect the legal and financial rights of the Government and those affected by Government activities (legal and financial rights records).

EO data are often needed in times of emergencies and as such it would be a good practice for WGISS participants to consider their EO data as Vital Records setting up a program to adequately plan for when disasters strike.

Source http://internal.usgs.gov/gio/irm/vrnara.pdf (note this is an internal site)

Data Life Cycle Concept

A key concept to addressing long-term archive strategies is to embrace that science data, including EO data, have *Data Life Cycles* associated with them. A very simple

records management lifecycle model, originating from the records management community, consists of:

>Phase 1. Creation
>Phase 2. Maintenance and Use
>Phase 3. Disposal

This simple model could be easily adapted for EO data to look something like:

>Phase 1. Acquisition
>Phase 2. Maintenance and Use
>Phase 3. Long-Term Archive

Even without modifying the original, simple model the word disposal should not imply destruction but merely transfer to an archive within the same organization or possibly to another organization. It should be rare that an EO dataset would be destroyed, as observations cannot be recreated.

Many additional models are considered in the separate CEOS WGISS document entitled, "Data Lifecycle Models and Concepts" located on the Data Stewardship Interest Group (http://www.ceos.org/index.php?option=com_content&view=category&layout=blog&id=174&Itemid=274) page.

(source: http://www.archives.gov/about/info/whats-a-record.html)