

**ADDENDUM TO:
SUMMARY DOCUMENTATION
FOR EMAP DATA:
GUIDELINES FOR THE
INFORMATION MANAGEMENT DIRECTORY**

Prepared for

U. S. Environmental Protection Agency
Atlantic Ecology Division
27 Tarzwell Drive
Narragansett, RI 02882-1197

Draft Prepared by

Melissa M. Hughes
Signal Corporation
c/o U.S. Environmental Protection Agency
27 Tarzwell Drive
Narragansett, RI 02882-1197

2 December 1996

Preface

This draft addendum to “Summary Documentation for EMAP Data: Guidelines for the Information Management Directory” (Frithsen and Strebel, 1995) provides additional metadata components for the Directory. These components were added based upon work conducted for the Columbia River Basin Project (Frithsen, 1996) and for compliance with NASA’s Global Change Master Directory and the Federal Geographic Data Committee (FGDC) Metadata Standard (FGDC, 1996). Some of these fields are not applicable to all data sets, and therefore, are not mandatory.

The addendum provides three tables. Table 1 provides information on the addition, renaming or a new format for attributes related to a data set. The attribute, an explanation, example entry and format are provided. Also indicated is whether or not the entry is mandatory. Table 2 provides a list of new tables and their proposed attributes. A brief description of the table content is provided and the attribute characteristics and specifications. The Oracle tables in Table 2 were incorporated to comply with the FGDC Metadata Standard. These tables primarily relate to ‘satellite’ data, but have been interpreted for ‘earth’ datasets. Table 3 presents the Oracle table structure of the Directory database and thus shows the placement of the attributes.

Principal Contact

Melissa M. Hughes
OAO Corporation
TELE: 401-782-3184
EMAIL: hughes.melissa@epamail.epa.gov

References

- Federal Geographic Data Committee. 1996. “Content Standards for Digital Geospatial Metadata.” June 8, 1994.
- Frithsen, J.B. and D.E. Strebel. 1995. Summary Documentation for EMAP Data: Guidelines for the Information Management Directory. 30 April 1995. Report prepared for U.S. Environmental Protection Agency, Environmental Monitoring and Assessment Program (EMAP), Washington, D.C. Prepared by Versar, Inc., Columbia, MD.
- Frithsen, J.B. 1996. Suggested Modifications to the EMAP Data Set Directory and Catalogue for Implementation in U.S. EPA Region 10. 11 June 1996. Report prepared for U.S. Environmental Protection Agency, National Center for Environmental Assessment, Washington, D.C. Prepared by Versar, Inc., Columbia, MD.

NEW METADATA COMPONENTS

Table 1. New/renamed/reformatted attributes in the Data Directory.

Attribute	Definition/Explanation	Format	Example	Mandatory	Null Values
New attribute: Organization_name.	A principal government agency or department, corporation, industry or private group. Each data center (an entity tracking and storing a data set) belongs to an organization.	Maximum 200 characters	U.S. Environmental Protection Agency	Yes	No
Rename Resource_Group_name to ' Sub_organization_name '	Each sub-organization must belong to a specific organization. Each data center must belong to a specific sub-organization.	Maximum 200 characters	U.S. EPA Region 1	Yes	No
Rename Resource_Group_code to Org_code (Organization code)	The sub-organization would be identified by a concatenation of 'Org_code Organization_name Sub_organization_name'	NNN	001	Yes	No
New attributes: Sampling_start_year, Sampling_start_month, Sampling_end_year, Sampling_end_month	Allows partial entry of dates where day and/or month not known. Full date entry is also maintained to comply with FIPS/FGDC.	NNNN NN	Year: 1995 Month: 03	No	Yes
Revise format: Phone numbers	Format revised for ease of entry.	NNN-NNN- NNNN	401-782-3000	No	Yes

Attribute	Definition/Explanation	Format	Example	Mandatory	Null Values
Rename Implementation_type to DS_format (Data set format)	Specific format for each data set; new name better reflects attribute values	Maximum 40 characters	Comma delimited ASCII formatted, tab delimited ASCII formatted, tab delimited SAS data set, semi-colon delimited SAS data set, Oracle table, etc. List of values.	No	Yes
Revise format: Zip_code	Format revised for ease of entry.	NNNNN-NNNN	02882-1197	No	Yes
New attribute: Sample_frequency	Frequency of sampling during the temporal period of the data set	Maximum 12 characters	Once, hourly, monthly, yearly, daily, etc. Incorporate a list of values.	No	Yes
New attribute: Num_sample_stat	Number of unique stations sampled during the temporal period of data set	NNNNN	100	No	Yes
New attribute: DS_quality_comments	A qualitative description of data quality. This is a comment field which will be VARCHAR(2000).	Maximum 2000 characters	Data quality comments if do not have metadata(?). Otherwise point to metadata.	No	Yes
Rename availability_code to DS_access_code (data set access code)	Change definition to reflect distribution of directory entries and/or data sets to a Home Page	Maximum 2 characters	10 - Unlimited availability 20 - Limited to MAIA Home Page (20 would be excluded from the EMAP Home Page)	Yes	No
New attribute: Data_set_progress	Status of data set completion	Maximum 11 values	Completed, in progress, planned. List of values.	No	Yes
New attributes: Minimum_altitude Maximum_altitude	Altitudes representing lower and upper limits, respectively, of data coverage. Units may vary	NNNNN	0 1500	No	Yes

Attribute	Definition/Explanation	Format	Example	Mandatory	Null Values
New attribute: Altitude_units	Units of minimum/maximum altitude	VARCHAR(5)	m, km	No	Yes
New attributes: Minimum_depth Maximum_depth	Lower and upper-most depth level, respectively, of data coverage. Units may vary.	NNNNN	0 35	No	Yes
New attribute: Depth_units	Units of minimum/maximum depth	VARCHAR(5)	m, km	No	Yes
New attribute: DS_URL	The location of the data set on the Internet.	Maximum 100 characters	http://www.epa.gov	No	Yes
New attributes: Date_data_file_created Date_data_file_revised	Enter date when data file was created; date for revision would be identical until the data file was modified.	DATE YYYYMMDD	19960609	No	Yes

Table 2. New data added to new tables, primarily related to satellite data.

DATA SET CITATION - Recommended reference to be used for the data set.

Attribute	Definition	Format	Example	Mandatory	Null Values
Originator	Name(s) of originators.	VARCHAR(200)	Goodman, S.	No	Yes
Title	Title of data set.	VARCHAR(240)		No	Yes
Series_name	Map series	VARCHAR(80)		No	Yes
Issue_identification	Volume and issue numbers	VARCHAR(40)	1989-1990	No	Yes
Publication_date	Date of publication.	NNNN	1992	No	Yes
Publication_place	Publishing house or city of publication.	VARCHAR(200)	NASA/GSFC	No	Yes
Publisher	Publisher of data set	VARCHAR(200)	NASA Climate Data System	No	Yes
Edition	Version	VARCHAR(10)	1.0	No	Yes
Data_presentation_form	Mode in which data are presented.	VARCHAR(40)	Remote-Sensing Image, Map, Model, View, etc.	No	Yes

EARTH DATA RESOLUTION

Attribute	Definition	Format	Example	Mandatory	Null Values
Latitude_resolution	The minimum difference between two adjacent latitude values	VARCHAR(80)	50 m	No	Yes
Longitude_resolution	The minimum difference between two adjacent longitude values	VARCHAR(80)	50 m	No	Yes
Altitude_resolution	The minimum difference between two adjacent altitude values	VARCHAR(80)	3 km	No	Yes
Altitude_units	Units of altitude resolution	VARCHAR(5)	m, km	No	Yes
Depth_resolution	The minimum difference between two adjacent depth values	VARCHAR(80)	100 ft	No	Yes
Depth_units	Units of depth resolution	VARCHAR(5)	m, km	No	Yes

BROWSE

Attribute	Definition	Format	Example	Mandatory	Null Values
File	Name of graphic file associated with the data (generally satellite data) file.	VARCHAR(80)	czcs.gif	No	Yes
Description	Description of graphic file.	VARCHAR(200)	NA	No	Yes
Caption	One line caption for graphic file display.	VARCHAR(80)	1978-1986 Chlorophyll Global Composite	No	Yes
Format	Format of browse graphic file.	VARCHAR(20)	GIF, JPEG	No	Yes
Graphic file URL	URL of graphic file	VARCHAR(80)	http://naraxp.nar.epa.gov	No	Yes

DISTRIBUTION

Attribute	Definition	Format	Example	Mandatory	Null Values
Distribution_media	Media options for user receiving the data.	VARCHAR(30)	Electronic, CD-ROM	No	Yes
Distribution_size	Approximate size for entire data set.	VARCHAR(20)	13.5 MB compressed, 3.5 GB uncompressed	No	Yes
Distribution_format	Standards used for receiving the data.	VARCHAR(20)	HDF, CDF	No	Yes
Fees	Cost for receiving data.	VARCHAR(20)	Free, \$100	No	Yes

Table 3. List of all attributes in the Directory database.

DATA SET IDENTIFICATION

Organization Name:
Sub-organization Name:
Data Set ID:
Version Number:
Entry Date:
Revision Date:
Data Set Progress:
Data Set URL:

DATA SET DESCRIPTION

Data Set Name:
Data Set Format:
Data Set Source:
Number of Sampling Stations:
Data Set Creation Date:
Data Set Revision Date:
Abstract:
General Keyword:
(Multiple entries allowed)
Data Set Comments:

DATA QUALITY COMMENTS

Data Quality Comments:

TEMPORAL PERIOD

Sampling Start Date:
Sampling End Date:
Sampling Start Date - Year:
Sampling Start Date - Month:
Sampling End Date - Year:
Sampling End Date - Month:
Sampling Frequency:

GEOGRAPHIC COVERAGE

Minimum Latitude:

Maximum Latitude:

Minimum Longitude:

Maximum Longitude:

Minimum Altitude:

Maximum Altitude:

Altitude Units:

Minimum Depth:

Maximum Depth:

Depth Units:

Locational Keywords:

(Multiple entries allowed)

DATA CENTER

Data Center Name:

Address1:

Address2:

Address3:

Address4:

City:

State:

Zip Code:

Country:

Voice Phone:

FAX Phone:

EMAIL Address:

Preferred Contact Position:

Originating Organization:

Originating Sub-organization:

Originating Data Center:

CONTACTS

(Multiple entries allowed)

Contact Title:

Contact Last Name:

Contact First Name:

Contact Middle Initial:

Contact Role:

Address1:

Address2:

Address3:

Address4:

City:

State:

Zip Code:

Country:

Voice Phone:

FAX Phone:

EMAIL Address:

DATA SET CITATION

(Multiple entries allowed)

Originator:

Title:

Series Name:

Issue Identification:

Publication Date:

Publication Place:

Publisher:

Edition:

Data Presentation Form:

Citation URL:

EARTH DATA RESOLUTION

Latitude Resolution:

Longitude Resolution:

Altitude Resolution:

Altitude Units:

Depth Resolution:

Depth Units:

BROWSE

File:

Description:

Caption:

Format:

Graphic File URL:

DISTRIBUTION

(Multiple entries allowed)

Distribution Media:

Distribution size:

Distribution Format:

Fees: