DATA MODELS FOR MACHU

Research Areas

CONCEPT

01 November 2016

Content

Appendix

DATA MODELS FOR MACHU

This document contains a brief explanation of why data models are used in MACHU GIS and a detailed description of the MACHU formats for research areas.

See also the MACHU reports on the MACHU website for more background information.
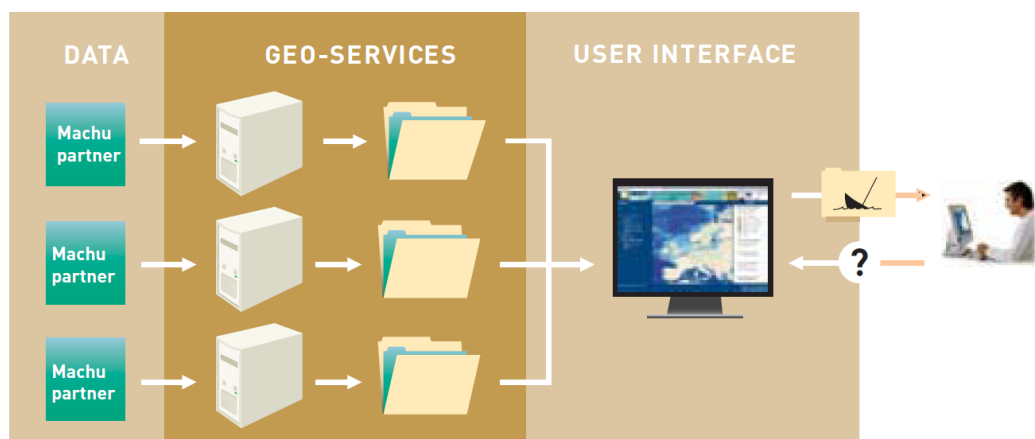
1. WHY USE MACHU DATA MODELS?

Data models are used for regulating the flow of information. Data models make it possible to harmonize the content of exchanged information and implement the technical requirements necessary to process the data in a Geographical Information System (GIS) such as MACHU GIS.

The use of the MACHU data models makes it possible to register information that is commonly felt to be of importance to the management of the cultural heritage underwater. In 2004, the Culture 2000 MoSS project, set up with the aim of monitoring, safeguarding and visualizing shipwrecks, provided a template for storing relevant management information. This template has served as an important source of information for defining the content of the MACHU data models.

The data models of MACHU are, different from those in the MoSS-project, set up with the intention to use them in a GIS-environment. This means that the formats provide information by which the data can share common spatial representation and by which the content of the data is comparable within a GIS. Using these data models makes it possible to handle data on the same subject but from different sources in MACHU GIS as if they were originated from a single source. For instance, it becomes possible to search and display data selections in a single search operation through many data sources at once.

Before the data can be used in MACHU GIS, it has to be served as a web service, according to OGC standards[1]. For a description of the process of creating a web services, see the corresponding documentation on the MACHU website.

MACHU GIS principle model: using different sources as a single source.

---

[1] Open Geospatial Consortium (OGC); http://www.opengeospatial.org/

2. FORMAT CHARACTERISTICS

The data model descriptions are based on the ESRI[2] shape file format (for vectorized data) and GeoTiff (for images). Note that this not necessarily means that data should also be stored as ESRI shapefile or GeoTiff. Important is that the data contains the described spatial representation, attributes and is available as web service.

MACHU data models are available for the layers:
- Archaeology (or Underwater Cultural Heritage)
- Research areas (including related images)
- Legislation

Examples of formatted empty shape files are available at the MACHU website.

Explanation of the components used to describe each attribute table:

Field
Contains the name of the attribute field, which is mostly an abbreviation of the content. ESRI-shape file attribute field names come with a maximum of 10 characters. In MACHU GIS an alias will be used to create readable attribute field names.

Description
Description of the content. The bold text is used as alias for the attribute field names. If more values have to be added in one field, they should be separated by commas.

Type
Description of notion (like number of characters or digits).

Optional/required
When marked 'r', adding information is required, when marked 'o' adding information is optional.

Domain
When marked 'y', attribute values should be taken from the domain list. (See appendix).
The domain lists only contain domain values that apply to values that represent common subjects. Values that represent definitions that refer to subjects on a (sub) national level are not added. It is recommended however that the suppliers of data register values that refer to subjects on (sub) national level, locally as domain values. This should prevent the use of different descriptions for a single subject, which could cause problems when performing searches through MACHU GIS. An exception is made regarding the archaeological period definitions (see archaeology data model). These values based on national periods can be used by different data suppliers. National period definitions should therefore be shared with other MACHU users.

ESRI shape files consist out of a number of data files with different extension like .shp, .dbf, .prj, .shx. When ESRI-shape files are created, attributes FID (Internally generated identification number for each feature (e.g. polygon)) and Shape (Internally generated text, indicating feature-type (e.g. polygon)) are automatically created in the database file (.dbf) of the ESRI-shape file. These files are not visible when opening the dbf-file in Excel.

---

[2] ESRI; http://www.esri.com/

3. DATA MODEL DESCRIPTION FOR THE RESEARCH AREAS LAYER

Shape
Polygon feature

Dataset exchange name
RES_[countrycode] e.g. RES_NL

Description Research Areas Layer
The research areas layer contains information on research areas, being areas where research has taken place of which results are expected to be meaningful to the management of cultural heritage underwater. Research areas are recorded as polygons. Research results often exist out of huge data files (e.g. multi beam readings) that can not easily be exchanged through a web-based GIS. The research area layer should make it possible to indicate the availability of research information, more then presenting the source data itself. Data source information (metadata) should make it possible to recover the actual source data when needed. Metadata should be made available for each research area, either as metadata added to a research image, or as a metadata file linked through the attribute 'Metadata'.
The format exists out of 8 attributes, containing a brief description of the kind of research, research period and a reference to an image (if available) that can be used to (indicatively) present the results of the research in the GIS and used as a link to the research data sources.
Research images are supposed to be geo-referenced (e.g. as GeoTIFF) and made available together with the research areas dataset. The image names should be identically to the name as referred to in the attribute field RES_IMAGE of the research area layer.

Domain table research areas
See appendix A.

Attribute table Research Areas

Field	Description	Type	Optional (o) Required (r)	Domain (if yes (y), consult domain table)
RES_IDENT	**Identifier** Unique identification number of research area. Might be used to uniquely identify the research area (polygon). Use <2 letter country code (ISO3166-1)>_<ID-code> E.g. for the Netherlands: NL_RA00001. etc.	Text (25)	r	
RES_TYPE	**Type of research** Choose one type of research, e.g. archaeological survey	Text (50)	r	y
RES_METH	**Research Method** Choose one type of method, e.g. multi beam	Text (50)	r	y
RES_TECH	**Research technique** Additional technical information on research method, for example a specification of used equipment e.g. Seabat 8101 (in case of multi beam).	Text (50)	o	
RES_START	**Research start date** (First) date of he research the	Yyyy-mm-dd	r	

	polygon represent				
RES_END	**Research end date** (Last) date of he research the polygon represents (identical to RES_START if period is just one day)	Yyyy-mm-dd	r		
RES_IMAGE	**Image** Linkage from research area (polygon) to image. Representative image of the research data (Name of the image should be unique and preferably build like: RES_<2 letter country code (ISO3166-1)>_ <year_end>_<year_start>_<metho d>_<number> e.g. RES_DE_2005_mbeam_1 or RES_NL_2006_2005_sss_3 Note: add <year_start> only if different from <year_end>. For suggestions on abbreviations for <method>, see domain list!	URL (254)	r (if available)	(for usage <method>, see domain list at RES_METH)
RES_META	**Metadata** Linkage to metadata of research data	URL (254)	r (if no image available), otherwise o	
RES_REF	**References** Link to a reference, e.g. a website of the project.	URL (254)	o	

Alterations to version august 2009 (end of project):

RES_IMAGE
The usage of the field RES_IMAGE is changed. The field RES_IMAGE has been used in MACHU GIS to add an image of a particular research area to the view, as well as to unlock the metadata information of the actual research data of that specific area. The double functions of the field RES_IMAGE lead to problems when an image was not available. In those cases an empty 'No Image' had to be created to unlock the metadata. Because a field RES_META is now added to the format, the field RES_IMAGE will only be used to add the image to the view. If no image is available, the field RES_image remains empty. Creating an 'empty image' will therefore no longer be necessary.

RES_META
RES_META is a new field, added to the research area format. The field contains the URL to the metadata of the research data of a specific area. This field will be used to unlock metadata directly form the attribute table of the research area layer in MACHU GIS. When no image is available, this field should contain an URL to the metadata. If an image is available, filling in this field is optional.

Alterations to version January 2013:

RES_METH / RES_IMAGE (Research Method) – Domain modified
RES_METH (Image) - Type modified from Text (50) to URL (254).

4. METADATA FORMATS

Data in MACHU GIS is accompanied by metadata. Metadata contains source information like content description, information about data quality, restrictions on data use and contact information to owner or custodian of the data.

Each dataset should contain metadata, distributed in xml-format (Extensible Markup Language) according to the INSPIRE Metadata Implementing Rules. INSPIRE[3] stands for 'Infrastructure for Spatial Information in Europe'. It is a European Commission initiative to build a European spatial data infrastructure (ESDI) that allows a variety of users to identify and access spatial data from a wide range of sources across Europe. INSPIRE prescribes the use of ISO 19115, metadata profile for geography (and ISO 19119 metadata standard for services). See INSPIRE website http://inspire.jrc.ec.europa.eu for more information.

To create metadata one can use any available metadata editor that meets the INSPIRE implementing rules. An editor is also available at the INSPIRE GeoPortal, see http://inspire-geoportal.ec.europa.eu.

To connect metadata to data in MACHU GIS, metadata files should be renamed after the source dataset e.g. ARCH_NL.shp.xml for ARCH_NL.shp.

For data recovery purposes (through a search engine or metadata catalogue) it is recommended to add 'MACHU' as keyword in the metadata.

---

[3] INSPIRE; Infrastructure for Spatial Information in Europe; http://inspire.jrc.ec.europa.eu.

APPENDIX

A. DOMAIN TABLE RESEARCH AREAS

RES_TYPE – Type of research

archaeological survey
archaeological monitoring
archaeological assessment
archaeological excavation
other

RES_METH / RES_IMAGE

Research method	Keyword or abbreviation for image name
aerial photography	aerphoto
coring	coring
diving	diving
georadar	georadar
magnetometer/gradiometer	magneto or gradio
multibeam	mbeam
seismic investigation	seismic
side-scan sonar	sss
single beam	sbeam
sub-bottom profiling	subbottom
other	-