DATA MODELS FOR MACHU

Archaeology
(Cultural Heritage Underwater)

15th of June 2020

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 archaeology (or Cultural Heritage Underwater).

See also the MACHU reports on the website
https://english.cultureelerfgoed.nl/publications/publications/2019/01/01/machu-documentation .

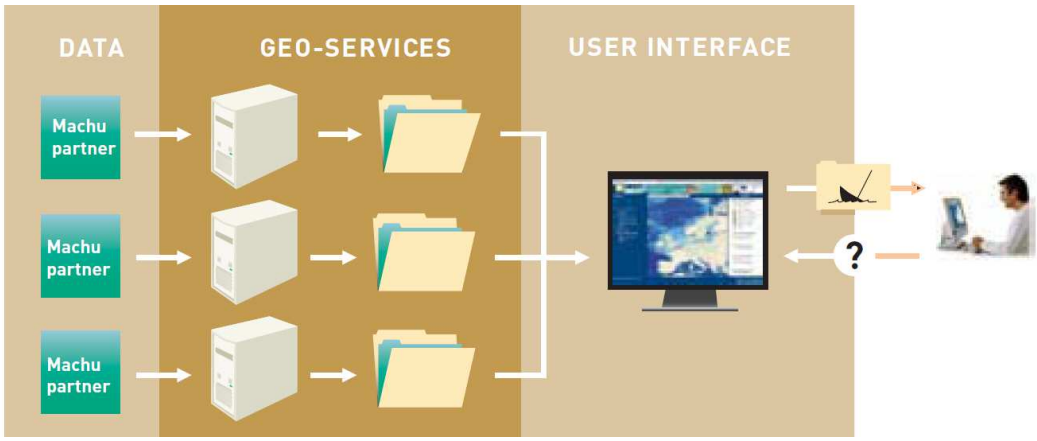
1. WHY USE MACHU DATA MODELS?

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 underwater cultural heritage. 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 enables handling data on the same subject but from different sources in MACHU GIS, as if they originated from a single source. It becomes possible, for instance, to search through many data sources at once in a single search operation, and to display such data selections.

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 website (see above).



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

[1] Open Geospatial Consortium (OGC); https://www.ogc.org

2. FORMAT CHARACTERISTICS

The data model descriptions are based on the example ESRI[2] shape file format (for vectorized data), available at the MACHU website. Note that this not necessarily means that data should also be stored as ESRI shapefiles. It is important that the data contains the described spatial representation, attributes, and that it is available as OGC web service.

MACHU data models are available for the layers:

- Archaeology (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.

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; https://www.esri.com

3. DATA MODEL DESCRIPTION OF THE ARCHAEOLOGY LAYER

Shape
Point feature

Dataset exchange name
ARCH_[country code] (e.g. ARCH_NL)

Description Archaeology Layer
The archaeology layer (or Cultural Heritage Underwater layer) contains information on archaeological sites or objects (e.g. shipwrecks). These sites are geographically recorded as point features, based on a xy-coordinate pair, using WGS84 as reference coordinate system. A point represents the location of the centre of the site. The attributes of the archaeology format are originally based on a collection of information elements originating from the MoSS management plan. Major alterations to the format have been initiated by the Dutch Cultural Heritage Agency in 2015, based on new insights after years of exploring the data.

Domain table archaeology
See appendix A.

Attribute table Archaeology

Field	Description	Type	Optional (o) Required (r)	Domain (if 'y', consult domain table)
obj_ident	**Management_ID** Used to uniquely identify the object or site. This id should be a 2 character country code (ISO3166-1) combined with a unique number (could also be NATREG number or code) e.g. NL_41204	Text (25)	r	
obj_name	**Object descriptive name** Name usually a toponym, given in reference to the location of the site. In practice this is a common name used in databases. E.g. Burgzand Noord 3.	Text (50)	r	
obj_type	**Object type** e.g. shipwreck; paleo-landscape; other.	Text (50)	r	y
period_min	**First year dated** e.g. –700 (meaning 700 BC) Number may be used to select object by age.	Number signed (8)	r	
period_max	**Last year dated** e.g. 1255 (meaning 1255 AD) Number may be used to select object by age.	Number signed (8)	r	
dendro	**Dendrochronology** Dendrochronical research available yes or no.	Text (10)	r	y

disc_date	**Discovery date** When first discovered e.g. 1985-07-05 (use January 1st for day and month when only the year is known).	Year, month, day (yyyy-mm-dd)	o	
material	**Main materials** Most important materials used e.g. wood, other metal. (if more than one, use a comma to separate).	Text (100)	r	y
arch_value	**Archaeological value** e.g. high archaeological value.	Text (50)	r	y
com_auth	**Competent authority** Full original (national) name of who is approved authority and can decide about the future of the site (e.g. Rijksdienst voor het Cultureel Erfgoed).	Text (254)	o	Use local domain values
last_visit	**Last visit** e.g. 2005-06-04	Year, month, day (yyyy-mm-dd)	o	
nat_reg	**National registration code** e.g. 41204	Text (50)	o	
loc_obj	**Object location** e.g. Burgzand Noord, Wadden Sea, The Netherlands.	Text (100)	r	
own_ter	**Owner terrain** e.g. private.	Text (100)	r	y
own_obj	**Owner object** If known.	Text (100)	o	
leg_stat	**Legal status** e.g. protected.	Text (25)	r	y
deg_stat	**Degradation status** e.g. partly damaged.	Text (25)	r	y
phys_pro	**Physical protection** e.g. yes.	Text (10)	r	y
threats	**Threats** e.g. looting, fishing (if more than one, use a comma to separate).	Text (100)	o	y
depth	**Depth (meters LAT)** Minimal dive depth as known (positive number), in meters LAT (Lowest Astronomical Tide) e.g. 9.0 or 10.5	Number (5)	o	
salinity	**Salinity** Salinity of the water, e.g. brackish.	Text (10)	r	y
current	**Current** Main water flow, e.g. tidal current.	Text (50)	r	y
reas_date	**Reassessment date** When should the site be re-assessed? This is part of the planning.	Year, month, day (yyyy-mm-dd)	o	

country	**Country** In which country is the site lying? Use official codes as given in ISO 3166_1. (XZ for international waters.)	Text (2)	r	Use ISO (3166-1)
pos_x	**Position longitude** (East or West in degrees, WGS84) e.g. written like (-)4.562279	Number signed (10)	r	
pos_y	**Position latitude** (North or South in degrees, WGS84) e.g. written like (-)53.025038	Number signed (10)	r	
r95	**Position accuracy (R95)** Position accuracy within radius (meters) equals 95%. Use positive integer, e.g. 100. Unknown accuracy values should be registered as 9999.	Number signed(4)	r	
ver_con	**Verifiable connections** To other countries. Use official codes as given in ISO 3166_1. (XZ for international waters.)	Text (25)	o	
references	**References** Link to website with extra documentation of the object e.g. a location on the MACHU WIS.	URL (254)	o	
notes	**Notes** Free text field for additional information.	Text (254)	o	
source	**Source** Specifying the source where the object is listed.	Text (254)	o	

Alterations to version 1 November 2016:

Names of the attributes – lowercase letters instead of capital letters

dendro (Dendrochronology) – number of characters: 10 instead of 4
com_auth (Competent authority) – number of characters: 254 instead of 100
current (Current) – number of characters: 50 instead of 10
ver_con (Verifiable connections) – number of characters: 25 instead of 254
notes (Notes) – Added
source (Source) – Added

Modified domain lists:

obj_type (Kind of object)
arch_value (Archaeological value)
own_ter (Owner terrain)
leg_stat (Legal status)
salinity (Salinity)

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 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 https://inspire.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 https://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.

APPENDIX

DOMAIN TABLE ARCHAEOLOGY

obj_type - Kind of object/site

maritime infrastructure
drowned settlement
paleo-landscape
shipwreck
plane wreck
loose object
fish trap/ fish weir
other

dendro – Dendrochronology

yes
no
unknown

material – Materials

bone/antler
ceramic
flint
stone
glass
iron/steel
other metal
wood
other
unknown

arch_value – Archaeological value

no archaeological value
archaeological value
high archaeological value
unknown

own_ter – Owner terrain

government
private
international waters
unknown

leg_stat - Legal status

unprotected
protected
unknown

deg_stat - Degradation status

removed/destroyed
fragment
partly damaged
unknown
well preserved

phys_pro - Physical protection

yes
no
unknown

threats - Threats

anchoring
dredging
salvaging
biological deterioration
chemical deterioration
looting
infrastructural development
shipping
mechanical erosion
fishing
explosives
other
unknown
none

salinity - Salinity

brackish
fresh
salt
unknown

current – (Main) Current

tidal current
still water
sea current
river current
unknown