

DATA FORMATS FOR MACHU



**Archaeology
(Cultural Heritage Underwater)**

09 January 2013

Content

- 1. WHY USE MACHU DATA FORMATS? 3
- 2. FORMAT CHARACTERISTICS 4
- 3. DATA FORMAT DESCRIPTION OF THE ARCHAEOLOGY LAYER..... 5
- 4. METADATA FORMATS 8

Appendix

- A. DOMAIN TABLE ARCHAEOLOGY..... 9
- B. LIST OF NATIONAL PERIOD DEFENITIONS.....11

DATA FORMATS FOR MACHU

This document contains a brief explanation of why data formats 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 MACHU website for more background information.

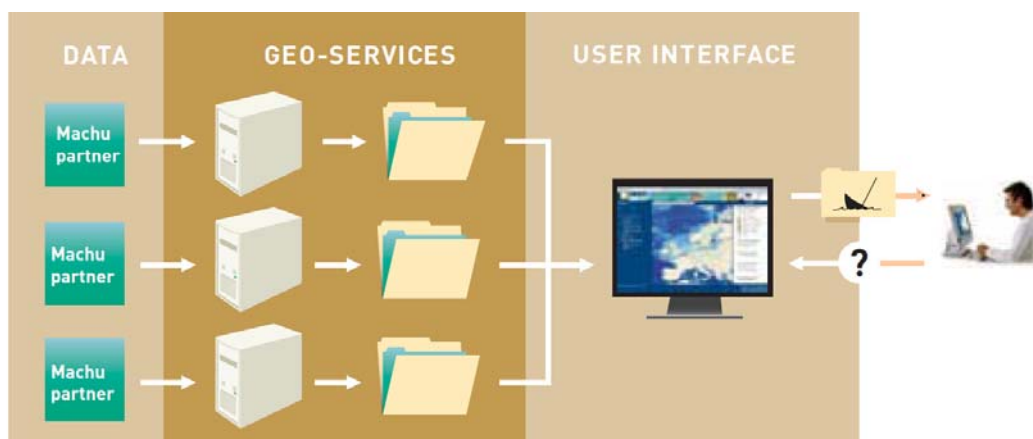
1. WHY USE MACHU DATA FORMATS?

Data formats are used for regulating the flow of information. Data formats 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 formats 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 project¹, 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 formats.

The data formats 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 formats 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 standards². 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 Monitoring, safeguarding and visualizing North-European shipwreck sites (MoSS); <http://www.mossproject.com/>.

² Open Geospatial Consortium (OGC); <http://www.opengeospatial.org/>

2. FORMAT CHARACTERISTICS

The data format descriptions are based on the ESRI³ 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 formats 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 format). 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.

³ ESRI; <http://www.esri.com/>

3. DATA FORMAT 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 mainly based on a collection of information elements originating from the MoSS management plan. The format now consists of 30 attributes, including descriptive information, assessment information, competent authorities, geographical coordinates, protection, threats, and information references (for instance to relevant information on MACHU's wrecks in situ content management system). Each site location has been given a unique identification number that should support communication but also create the opportunity to establish a relationship with other data sources.

Domain table archaeology

See appendix A.

Attribute table Archaeology

Field	Description	Type	Optional (o) Required (r)	Domain (if yes (y), consult domain table)
OBJ_IDENT	Management_ID Used to uniquely identify the object (or site). Proposal: This id could be a 2 letter country code (ISO3166-1) combined with 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 position of the wreck. In practise this is the name how it is usually described in the databases. E.g. Burgzand Noord 3	Text (50)	r	
OBJ_POP	Object poplar name	Text (50)	o	
OBJ_ORGN	Object original name The name can be given when it is known e.g. De Rooswijk	Text (50)	o	
OBJ_TYPE	Object type You can choose between: e.g. shipwreck, prehistoric site, 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)	o	
PERIOD_MAX	Last year dated	Number	o	

	e.g. 1255 (meaning 1255 AD) Number may be used to select object by age	signed (8)		
PERIOD_CO	Archaeological period Period according to country where object is located e.g. Viking age 800-1050 AD	Text (50)	o	Use local national domain values (see appendix B)
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 or transported (e.g. wood; metal)	Text (100)	o	y
ARCH_VALUE	Archaeological value High, low or unknown	Text (25)	o	y
UP_DATE	Last update Last update, changes in data about information on the site (e.g. 2003-05-03)	Year, month, day (yyyy-mm-dd)	r	
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 (100)	r	Use local domain values
LAST_VISIT	Last visit e.g. 2005-06-04	Yyyy-mm-dd	o	
NAT_REG	National registration code e.g. 41204 (example number from Dutch archaeological Information system Archis)	Text (50)	r	
LOC_OBJ	Object location e.g. Wadden Sea Burgzand, The Netherlands	Text (100)	r	
OWN_TER	Owner terrain e.g. Municipality of Texel, The Netherlands	Text (100)	r	
OWN_OBJ	Owner object If known	Text (100)	o	
LEG_STAT	Legal status Is it protected by law? e.g. protected site or not protected site	Text (25)	r	y
DEG_STAT	Degradation status e.g. Well preserved, partly damaged, destroyed, unknown	Text (25)	r	y
PHYS_PRO	Physical protection yes or no	Text (10)	r	y
ACCESS_RES	Access restrictions Restricted access for public (diving) e.g. yes/no/unknown	Text (10)	r	y
THREATS	Threats	Text (100)	o	y

	e.g. looting, fishing, erosion of seabed, abrasion, biological deterioration, other			
DEPTH	Depth (meters LAT) Minimal Dive depth as known (positive number), in meters LAT e.g. 9,0 or 10,5	Number (5)	o	
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 short names in English as given in ISO 316-1	Text (25)	r	Use ISO 316-1
POS_X	Position longitude (East or West in degrees, WGS84) e.g. written like 4.562279 or negative -4.562279.	String (10)	r	
POS_Y	Position latitude (North or South in degrees, WGS84) e.g. written like 53.025038	String (10)	r	
VER_CON	Verifiable connections To other countries. Use country names	text (255)	o	
REFERENCES	References Link to extra documentation of the object	url	o	

Alterations to version august 2009 (end of project):

UP_DATE

The name of attribute UPDATE is changed in UP_DATE because the name UPDATE is not accepted when storing the data in (some) geo-databases.

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. INSPIRE⁴ 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.

⁴ INSPIRE; Infrastructure for Spatial Information in Europe; <http://inspire.jrc.ec.europa.eu>.

APPENDIX

A. DOMAIN TABLE ARCHAEOLOGY

OBJ_TYPE - Kind of Object

shipwreck
prehistoric site
built structure
geo-archaeological structure
other
loose object
unknown

PERIOD_CO - Period definitions according to country where object is located

See list of national domain values, appendix 4.

MATERIAL – Materials

bone/antler
ceramic
flint/stone artefacts
glass
iron
other corroding material
other organic material
other metal
other non-corroding material
steel
wood
unknown

ARCH_VALUE – Archaeological Value

high
low
unknown

LEG_STAT - Legal status

protected by law
not protected by law
Unknown

DEG_STAT - Degradation Status

destroyed
partly damaged
removed
unknown
well preserved

PHYS_PRO - Physical protection

Yes
No
Unknown

ACCESS_RES - Access restrictions

yes
no
unknown

THREATS - Recognized threats

looting
diving
fishing
anchoring
infrastructural development
erosion of seabed
abrasion
chemical deterioration
biological deterioration
other
unknown

B. LIST OF NATIONAL PERIOD DEFINITIONS

(Version end of project 2009)

Period definitions according to country where object is located (for archaeology data format).

<i>Period definitions of Belgium</i>
Paleolithicum: 250.000 – 12.000 BP
Mesolithicum: 10.000 – 5.000 BP
Neolithicum: 5.300 – 2.000 BC
Bronze Age: 2.000 – 800 BC
Iron Age: 750 – 57 BC
Roman period: 57 BC – 476 AD
Early Middle Ages: 476 -987 AD
High Middle Ages: 987 – 1250 AD
Late Middle Ages: 1250 – 1500 AD
New Time: 1500 – now
period not known

<i>Period definitions of Germany</i>
Palaeolithicum > 8800 BC
Mesolithicum 8800 - 4000 BC
Neolithicum 4000 - 2000 BC
Bronze Age 2000 - 800 BC
Iron age 800 BC - 0
Roman period 0 - 450 AD
Early Medieval period 450 - 900
High Medieval period 900 - 1200
Late Medieval period 1200 - 1500
New Time A 1500 - 1700
New Time B 1700 - 1800
New Time C 1800 - now
period not known

<i>Period definitions of Great Britain (United Kingdom)</i>
Palaeolithic 500000-10000
Lower Palaeolithic 500000-150000
Middle Palaeolithic 150000-40000
Upper Palaeolithic 40000-10000
Mesolithic 10000-4000
Early Mesolithic 10000-7000
Late Mesolithic 7000-4000
Early Prehistoric 500000-4000
Neolithic 4000-2200
Early Neolithic 4000-3300
Middle Neolithic 3300-2900
Late Neolithic 2900-2200
Bronze age 2600-700
Early Bronze age 2600-700
Middle Bronze age 2600-700
Late Bronze age 2600-700
Iron age 800 BC - 43 AD
Early Iron age 800-300 BC
Middle Iron age 300 - 100 BC
Late Iron age 100 BC- 43 AD
Later Prehistoric 4000 BC - 43 AD
Prehistoric 500000 BC - 43 AD
Prehistoric or Roman 500000 BC - 410 AD
Roman 43 - 410 AD
Early Medieval 410-1066
Medieval 1066-1540
Post Medieval 1540 - 1901
Early Med. or later 410-1540
20th century 1901-2000
Early 20th century 1901-1932
First World War 1914-1918
Mid 20th century 1933-1966
Second World War 1939-1945
Late 20th century 1967-2000
21st century 2000-2100
Uncertain

<i>Period definitions of Netherlands</i>
Paleolithicum < 8800 BC
Mesolithicum 8800 - 4900 BC
Neolithicum 5300 - 2000 BC
Bronze Age 2000 - 800 BC
Iron Age 800 - 12 BC
Roman period 12 BC - 450 AD
Early Medieval period 450 -1050
Late Medieval period 1050 - 1500
New Time A 1500 - 1650
New Time B 1650 -1850
New Time C 1850 - now
period not known

Period definitions of Poland
period not known

Period definitions of Portugal
period not known

Period definitions of Sweden
Paleolithicum (<10000 BC)
Mesolithicum (10000-4000 BC)
Neolithicum (4000-1800 BC)
Bronze Age (1800 BC - 500 BC)
Preroman period (550 BC – 0)
Roman period (0-400 AD)
Folkvandringstid (400-550 AD)
Vendel period (550-800 AD)
Viking Age 800 - 1050 AD
Medieval period (1050-1520 AD)
New Time (1520 AD or younger)
period not known