

**Table 1.** Consensus list of fields for describing data sets. A short description and the EML tags are included. Different colors are used to distinguish information relative to each data set (black), to the site (blue), and initial information to be provided by the administrator of the online tool (red)

FIELD	CONTENT	EML MOST RELEVANT TAGS
<b>Source</b>	Information about the provider (ministry, National Park, Research institute) in the regular form (i.e. electronic and physical address, fax and telephone number)	creator
<b>Site</b>	A hierarchical classification (e. g. Continent/Country/Name of the park, etc) of the site where the information was gathered + map where the polygon of the study site could be delineated	description, geographicCoverage
<b>Person of contact + deputy</b>	Information about user (up loader/creator) is normally provided in the form of e-mail address, but including a deputy person could be useful	contact, creator, associatedParty
<b>Name</b>	The name used by the provider to identify the dataset	title
<b>Description/ Abstract/ Keywords</b>	The description given by the provider. Note that machines do not understand texts, but look for keywords	title, description, funding, studyAreaDescription, designDescription, abstract, keywordSet, keyword
<b>Human/ Sensor/ Model/GIS</b>	To distinguish between these ways of gathering data	Custom metadata tag/keywordSet methods/DwC import
<b>Geometry</b>	Whether information comes from point-based or polygon-based monitoring	geographicCoverage, spatialVector, spatialRaster
<b>Beginning of data</b>	First date of data collection in the data set	eml-coverage module -> beginDate
<b>End of data</b>	Last date of data collection in the data set	eml-coverage module -> endDate
<b>Number of data</b>	Rows in the database	numberOfRecords
<b>Frequency</b>	Whether data are collected annually, monthly, etc.	dataset->MaintUpFreqType
<b>Target (individual/ population/site)</b>	The study object: ecosystem, population, individual...	description
<b>Precision</b>	Whether the information is gathered in a particular spatial context (e.g. UTM 1x1 km for censuses). Leave empty if not	precision, attributeDefinition, eml-spatialReference module
<b>Resolution</b>	Whether the information was originally taken at one point, at several ones, the one of the satellite imagery, etc..	precision, attributeDefinition, eml-spatialReference module, geographicCoverage
<b>Protocol (Y/N)</b>	Whether the data come from well-described monitoring protocols in each of the sites. A subordinated question to answer "Y" may be whether the study matches exactly a referenced protocol. In this case the reference of this study should be provided (the doi if available)	methods, protocols
<b>Protocol (description)</b>	A brief description of the protocol is required + the upload of a document in English where the protocol is properly described if answer to protocol was "N"	methods, protocols
<b>Type of Variable</b>	Whether observations are counts, presence/absence, measurements	dateTime, measurementScale, numericDomain, nominal, ordinal, interval, ratio
<b>Fields</b>	Name of the fields included in the data set	

<b>Fields attributes. Description</b>	Brief description in English.	
<b>Field attributes. Format</b>	DD/MM/YYYY for date, etc.	
<b>Field attributes. Units</b>	If weight is measured, units will be grams	attribute->unit->standardUnit
<b>Original format</b>	Despite information will be stored in a sharable format (txt, csv...), it may come from different formats (xls, dbf, odt...)	physical
<b>Availability</b>	Whether data could be shared	resource->distribution
<b>Accessibility</b>	Whether data are accessible: online, upon request, use constraints, etc	resource->online/offline/inline
<b>Intellectual rights</b>	Self explanatory	eml access module, resource->intellectualRights
<b>Metadata date</b>	Self explanatory	resource->pubData
<b>Dataset language</b>	Self explanatory	xml:lang?, entity->additionalInfo?
<b>Data set</b>	File to be browsed in user's computer for uploading	entity, dataTable, physical