EEA data service metadata definitions

The metadata definitions in this document are designed to help the European Environment Agency (EEA) internally, ETCs, and external contractors delivering data resources [1] used in EEA products or uploading data resources to EEA data service.

The table below lists all metadata tags available within EEA data service web interface. Following information is used to define tags:

Description:
Short explanation of metadata tag and potentialy a suggestion of what should be its content.

Obligation/Condition:
This indicates whether a metadata shall always be documented or sometimes be documented. This may have the following values:
- mandatory – the metadata has to be documented
- optional – the metadata may be documented if relevant
- conditional – the metadata should be documented if some condition(s) is fulfilled. Condition(s) is defined.

Format:
Formal properties that metadata should follow.

Note:
Additional information, usualy pointed to some metadata tag speciality or to recommended filling and editing practice.

Example:
Concrete example of the content of metadata tag

Visibility:
Data service has two distinct levels of visibility:
- administrators only - metadata tag is visible only for users with administrator privilege
- public - metadata tag can be seen by everybody

Most of listed metadata tags can be linked to the fields in the draft ISO 19115 metadata standard used by EEA for GIS data [2]. Last column of the table shows relation to this draft.

---

[1] Data resources are all types of resources provided by the EEA data service. It includes for example tabular data, maps or graphs.
[2] In this case defined as GIS specific formats (e.g. shape file, coverage, geodatabase)

Metadata tag	Definition (**Formats are printed in green**, **Examples are printed in blue**)	ISO
Title	**Description:** The name given to the resource. Typically, a *title* will be a name by which the resource is formally known. **Obligation/Condition:** mandatory **Format:** (Subset from ISO-8859-1 Characterset){1,200} HTML not allowed. **Note:** The data resource *title* is positioned at the top of the page for a specific data resource. The name of the metadata tag *title* is not visible through the data service edit tool and cannot be chosen from the metadata list. A data resource *title* is assigned during the creation of the data resource. **Example:** Nationally designated areas (National - CDDA) **Visibility:** public	Dataset identification – Title
Abstract	**Description:** Short description of the data resource. **Obligation/Condition:** mandatory **Format:** (Subset from ISO-8859-1 Characterset){1,2000} HTML allowed. **Note:** The data resource *abstract* is positioned below the *title*. The name of the metadata tag *abstract* is not visible through the data service edit tool and cannot be chosen from the metadata list. A data resource *abstract* is compiled during the creation of the data resource. **Example:** Data on anthropogenic emissions of acidifying pollutants (NH3, NOX, SO2) sent by countries to CLRTAP/EMEP with copies to EEA and ETC/ACC. Data compiled and held by ETC/ACC are annual national total emissions of NH3, NOX and SO2 from individual countries. Sectoral data is provided… **Visibility:** public	Dataset identification – Abstract
Contact person for EEA	**Description:** Outside person to be contacted by EEA if questions regarding the data resource arise at a later date, responsible project manager in EEA and the operator who uploaded the data resource and edited metadata. All three roles should be enlisted here. **Obligation/Condition:** mandatory **Format:** (Role of the person: First name Last name, e-mail adress) HTML allowed. **Note:** Rule "One person, one record" should be used. It means each role should be entered as new record. **Example:** Outside contact: Ruth Cullingford, ruth.cullingford@wrcplc.co.uk Inhouse contact: Niels Thyssen, Niels.THYSSEN@eea.europa.eu Operator: David Simoens, David.SIMOENS@eea.europa.eu **Visibility:** administrators only	Metadata on metadata – Point of Contact
Disclaimer	**Description:** A *disclaimer* is a statement which generally states that the entity authoring the *disclaimer* is not responsible for something in some manner. **Obligation/Condition:** optional **Format:** (Subset from ISO-8859-1 Characterset){recommended max. length = 600 characters} HTML allowed. **Example:** \Disclaimer: The data in Waterbase are sub-samples of national data assembled for the purpose of providing comparable indicators of pressures, state and impact of waters on a Europe-wide scale and the data sets are not intended for assessing compliance with any European Directive or any other legal instrument. Information on the national and sub-national scales should be sought from other sources.\ **Visibility:** public	

Geographic accuracy	**Description:** Geographic accuracy of location, ground distance as a value in meters.
	Obligation/Condition: conditional – data resource is GIS dataset
	Format:
	   ([0-9]){1,6} meters
	   (Subset from ISO-8859-1 Characterset){recommended max. length = 300 characters}(=Possible additional information)
	   HTML not allowed.
	Example: 150 meters - 25 ha minimum mapping unit.
	Visibility: public
		Other dataset information – Geographic accuracy

Geographic box coordinates	**Description:** Geographic position bounding box of the dataset. **Obligation/Condition:** conditional – data resource is GIS dataset **Format:** See note HTML not allowed. **Note:** In the future it will be automaticaly harvested from the xml **Example:** West bound longitude:-31.269751; East bound longitude: 120.131332; North bound latitude: 81.851929; South bound latitude: 17.4151. **Visibility:** public	Other dataset information – Geographic box

Geographical coverage	**Description:** The geographical extent of the content of the data resource. Geographical coverage will typically include spatial location.
	Obligation/Condition: conditional – data resource is dataset or graph
	Format: (Subset from ISO-8859-1 Characterset){1, 4000}
	   Recommended max. length for free text = 300 characters
	   See note.
	   HTML allowed.
	Note: Two possibilities are given to the metadata editors for this tag. If the "automatic" procedure is chosen then checked countries appear as highlighted from a map. Otherwise, the non-automatic procedure is equivalent to free text editing. Recommended practice is to choose the automatic procedure. Only geographical regions, which are not listed (e.g. seas), or some additional or detailed information should be inserted as free text.
	Visibility: public	

Keywords	**Description:** List of keywords assigned to the data resource for identification purposes.
	Obligation/Condition: mandatory
	Format: (keyword_1 [, keyword_2, keyword_3, …])
	   HTML not allowed.
	   See also note.
	Note: All keywords already used for existing resources are displayed in the dynamically created list and can be selected. New keywords which are not listed there can be added as free text.
	Rules for the new keywords:
	- keywords should as a general rule be in lower case letters, including the first letter;
	- keywords should as a general rule be noun, singular;
	- names of places/persons/organisations should start with a capital letter and the rest lower case (see EEA Writing Style);
	- acronyms as a general rule should be in capital letters (see EEA Writing Style) and should be added to the EEA Glossary (http://glossary.eea.eu.int/EEAGlossary/) if they are not there already
	- keywords in one row should be separated by comma followed by space
	Example: CO_2, CH_4, N_2O, climate, ETC/ACC, HFC, PFC, SF_6, UNFCCC, Kyoto
	Visibility: public
		Dataset identification – Keyword

Last upload	**Description:** Shows the date when the data resource was last uploaded in EEA data service. **Obligation/Condition:** mandatory **Format:** ([1-9]){2}(=DD) ([A-z]){3}(=Mmm) ([1-9]){4}(=YYYY) HTML not allowed. **Example:** 07 Sep 2004 **Visibility:** public	Dataset identification – Reference date
Methodology	**Description:** Description of how the resource was compiled. See also note. **Obligation/Condition:** optional **Format:** (Subset from ISO-8859-1 Characterset){1,4000} HTML allowed. See also note. **Note:** *Metodology* should have at least one of the following parts. Each part has to be labeled by its name: Used tools: List of the software tools used to compile resource. Applied procedures: Description of used methods and procedures. Content specification: Definition of the different key entities that can be found within the data resource. Note: Usualy some exception in *methodology* or some additional information important to understand the content of resource. References: Usualy link to official (published) *methodology* or to the entity, which is able to provide detailed information about the used *methodology*. **Example:** Used Tool: ESRI ARCGIS Applied procedures: Filter applied from the source: DEG_URB: Total. GEO: all at national level only. TIME: All (no filter applied). 1999 - 2002. Content specification: Flow: - Total Final Consumption: Total final consumption (TFC) is the sum of consumption by the different end-use sectors. Product: - Total: the total of all energy sources. Note: Employment in agriculture, employment in industry, employment in services may not sum to 100 percent because of workers not classified by sectors. References: Further information and detailed explanatory notes by country may be found at the following URL... **Visibility:** public	Other dataset information – Methodology description, Process steps
Originator	**Description:** Information about the intellectual creator. Examples of an *originator* include a person, an organisation or a convention. **Obligation/Condition:** optional **Format:** See note. HTML allowed. **Note:** The *originator* should be selected from the list which dataservice edit tool provide for this metadata tag. New *originator* should be entered also to the meta_sources table (ask dataservice manager how to do this). In order to avoid duplicated information, when *originator* is identical to the *owner* it is not necessary to add information about *originator*. **Example:** The European Topic Centre on Air and Climate Change (ETC/ACC) **Visibility:** public	Distribution information – Originator

Other information	**Description:** Other relevant information **Obligation/Condition:** optional **Format:** (Subset from ISO-8859-1 Characterset){1,4000} 　　　HTML allowed 　　　See also note. **Note:** In the future it will be deleted **Example:** … **Visibility:** public	
Owner	**Description:** An entity that owns the data resource. Examples of an owner include a person, an organisation or a convention. **Obligation/Condition:** mandatory **Format:** 　　　See note. 　　　HTML allowed. **Note:** The *owner* should be selected from the list which dataservice edit tool provide for this metadata tag. New *owner*s should be entered also to the meta_sources table (ask dataservice manager how to do this). 　　　Rule for the datasets: 　　　If the dataset is listed in the ROD (Reporting obligation database) the *owner* information should be identical to the information field called "Report to" provided there. **Example:** United Nations Framework Convention on Climate Change **Visibility:** public	Distribution information – Owner
Processor	**Description:** The technical producer or processor of the data. Examples of a *processor* include a person, an organisation, or a service. **Obligation/Condition:** optional **Format:** 　　　See note. 　　　HTML allowed. **Note:** The *processor* should be selected from the list which dataservice edit tool provide for this metadata tag. New *processor* should be entered also to the meta_sources table (ask dataservice manager how to do this). 　　　In order to avoid duplicated information, when *processor* is identical to the *owner* it is not necessary to add information about *processor*. **Example:** The European Topic Centre on Air and Climate Change (ETC/ACC) **Visibility:** public	Distribution information – Processor
Reference system	**Description:** Definition of the reference system used for the dataset. **Obligation/Condition:** conditional – resource is GIS dataset **Format:** 　　　See note 　　　HTML not allowed. **Note:** In the future it will be automaticaly harvested from the xml. **Example:** Lambert Azimutal **Visibility:** public	Reference system
Relation	**Description:** Relation is used to describe parent/child relationships between data sets. When a data set X is derived from another data set Y, X is the child of Y while Y is the parent of X. **Obligation/Condition:** optional **Format:** ((Parent resource | Derived resource)(=Resource type):(Resource name)) 　　　HTML allowed. 　　　See also note. **Note:** Rule "One relationship, one record" should be used. It means if particular datasets have both parent and child status, each status should be entered as a new record.	

	<u>Important</u>:
	The manual update has to be made twice in both parent and derived data resources.
	Example:
	   Parent data set: <u>Corine land cover (CLC1990) 100 m v12/2000</u>
	   Derived data set: <u>Green urban areas 1990</u>
	Visibility: public	

Rights	**Description:** Information about rights held in and over the resource. Typically a *rights* element will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses intellectual property rights (IPR), copyright and various property rights.
	Obligation/Condition: optional
	Format: (Subset from ISO-8859-1 Characterset){recommended max. length = 300 characters}
	   HTML allowed.
	   See also note.
	Example: two distinct cases
	   <u>If owner is EEA:</u>
	   EEA grants free access to all its data/applications/maps/graphs provided that the user agrees to display a link to the EEA web site http://www.eea.europa.eu and to acknowledge the source as follows: Copyright EEA, Copenhagen, [production year].
	   <u>If owner is third party:</u>
	   Access is managed by the owner mentioned below. Please contact the owner for more information about their data policy.
	Visibility: public
		Distribution information – Access rights

Scale of the dataset	**Description:** Gives a rough value of accuracy of the dataset.
	Obligation/Condition: conditional – resource is GIS dataset
	Format: (1:([1-9]){1,8)) - ((Subset from ISO-8859-1 Characterset){recommended max. length = 300 characters}(=Possible additional information))
	   HTML not allowed.
	   See also note.
	Note: Rule "One scale, one record" should be used. It means if particular datasets are of different scales, each scale should be entered as a new record.
	Example: 1:100 000
	Visibility: public
		Other dataset information - Scale

Source	**Description:** A reference to a resource from which the present data resource is derived. Various information could be entered here:
	   - reference to for example a scientific literature where the resource was first published
	   - more detailed description of the owner or processor (exact body or department, date of delivery, original database, table or GIS layer, …)
	   - contributors
	Obligation/Condition: mandatory
	Format: (Subset from ISO-8859-1 Characterset){recommended max. length = 500 characters }
	   HTML allowed.
	   See also note.
	Note: Rule "One source, one record" should be used. It means if the present resource has more sources, each of them should be entered as a new record.
	Example: International Energy Agency (IEA): Energy Balances OECD countries (2003 edition) and Energy Balances non-OECD countries (2002 edition). Energy Balances.
	Visibility: public	

Temporal coverage	**Description:** The temporal scope of the content of the data resource. Temporal coverage will typically include a year or a time range. **Obligation/Condition:** mandatory **Format:** ([(1-9)]{4}(=YYYY))(=year or rangeyear) 　　　　(year_1 | rangeyear_1-rangeyear_2 [, year_2 | rangeyear_3-rangeyear_4, …) 　　　　HTML not allowed. **Example:** 1996, 1998, 2000 - 2004 **Visibility:** public	
Theme	**Description:** Categorisation of the data resource into EEA themes. See full list at http://themes.eea.eu.int/. **Obligation/Condition:** mandatory **Format:** 　　　　See note. 　　　　HTML not allowed **Note:** Theme or themes have to be selected from predefined list. **Example:** air, nature, tourism, environmental reporting, etc. **Visibility:** public	Dataset identification – Topic cathegory
Type	**Description:** The nature or genre of the content of the resource. Type includes terms describing general categories, functions, genres, or aggregation levels for content. **Obligation/Condition:** conditional – resource is map or graph **Format:** ([map | graph | aggregates]) 　　　　HTML not allowed. **Note:** "aggregate" type will be assigned to data resources (datasets) which represent a collection of data typically used in reports (Ex: Climate change scenarios data). These type of data should be moved to Maps&Graphs section of the EEA data service. **Example:** map **Visibility:** public	
Unit	**Description:** Describes the unit(s) taken in account for the measurement values(s) of the data resource. **Obligation/Condition:** optional **Format:** ([component_1, component_2, …][:] unit) 　　　　HTML allowed. 　　　　See also note. **Note:** Rule "One unit, one record" should be used. It means if components in the data resource are measured in more different units, each unit (with the list of the related components) should be entered as a new record. **Example:** BOD5, BOD7, COD, CODCr, Dissolved Oxygen, CODMn: mg/l O2 **Visibility:** public	