

Recommendation regarding the description of environmental data published on SLU's web

Data are an increasingly strategic societal resource – for individuals as well as companies, organisations, and authorities. Publishing data with free access (i.e., making data openly available) provides the basis for new opportunities and innovations in science, planning, decision-making etc. as well as in a variety of other processes that benefit society.¹

In order for SLU's open data to be understood and reused, it has to be described carefully and properly. Data that is described carefully and properly allows for an appropriate and correct utilisation thereof. The template presented as part of this document aims to standardise what type of information (metadata elements) is to be included in an overall description of data and services (Table 1). This information is to be seen as a minimum of metadata provided on SLU's open data website² according to the guidelines in SLU's quality guide for the management of environmental monitoring data³. A more developed publishing system (e.g., a SLU-wide data repository or equivalent) may, however, in the long term replace this entire process.

¹ <https://www.regeringen.se/regeringsuppdrag/2019/05/uppdrag-att-oka-den-offentliga-forvaltningens-formaga-att-tillgangliggora-oppna-data/>

² <https://www.slu.se/oppnadata>

³ <https://internt.slu.se/stod-service/fortlopande-miljoanalys/verksamhetsstod/ingangssida/kvalitetsguide/>

Table 1. Recommended metadata elements for overall description of data and services within SLU's environmental monitoring and assessment work. This table is partly based on Sweden's interpretation of the EU's DCAT-AP⁴ and on SND's (Swedish National Data Service) metadata profile based on the DDI⁵.

Metadata Element	Content
Title	Title or naming of data(set). Use as specific a name as possible. <i>The 'Title' may be one of the few elements providing a hit when searching in a data portal, such as Google's Dataset Search portal⁶ that retrieves (limited) information from national and international data portals.</i>
Principal	Swedish University of Agricultural Sciences <i>This element should not be changed.</i>
Responsible institution/unit	Institution/unit
Contact person	Contact information of person in charge (e.g., project leader, coordinator, principal investigator, or equivalent) or information owner. <i>Format: First name, surname, e-mail address</i>
Contact person w.r.t technical support	Contact information of technical support with regard to data access. <i>Format: First name, surname, e-mail address</i>
Description	Free-text description of data(set) – remember to describe the data for users without expert knowledge in mind. The description should include (max. 200 words): purpose of data collection, possible target group, methods (e.g., observations made via citizen science [CS]), type of measurements performed or included variables (e.g., water quality, tree growth, plant biomass) and geographical area and time period covered by the data. <i>The 'Description' may be one of the few elements providing a hit when searching in a data portal, such as Google's Dataset Search portal that retrieves (limited) information from national and international data portals.</i>
Keywords	Free-text keywords or tags describing the data(set). Examples of lists of keywords: INSPIRE data themes, GEMET, GCMD, EnvThes
Access to data	A web page that can be navigated to in a Web browser to gain access to the data(set), its distribution and/or additional information. Alternatively, a persistent identifier (DOI or equivalent) can be specified.

⁴ https://opnadata.se/wp-content/uploads/2018/06/DCAT-AP1.1-Svensk-rekommendation_2018-06-20_slutlig__.pdf

⁵ <https://www.ddialliance.org/>

⁶ <https://datasetsearch.research.google.com/>

File format	<p>Format of data files.</p> <p>We suggest using formats that are commonly utilised, supplier-independent, and preferably with an open technical specification (i.e., a format that is not controlled by an individual person or organisation), such as CSV (',' or ';' separation) or text (HTML, XML, JSON).</p>
Data format	<p>Format/structure of the data material such as numeric, text, video, software, etc. Note that data format is not the same as file format.</p>
Website of project	<p>Link to the project's website.</p>
Description of service	<p>Description of data access service (applies only in case API exists). State purpose and type of service as well as the link to access the service. Also provide a link to the service's documentation describing service request with data response as well as any potential parameters. The documentation should also describe authentication and error handling, preferably providing examples.</p>
Language	<p>Language used in data or of variables.</p>
Release date	<p>Date of formal issuance (e.g., publication on the web) of the data.</p> <p><i>Format: YYYY-MM-DD</i></p>
Frequency	<p>The frequency at which the published data is updated (e.g., specific dates, yearly, unknown).</p>
Update/modification date	<p>Most recent date on which the published data was changed, updated, or modified.</p> <p><i>Format: YYYY-MM-DD</i></p>
Limitations w.r.t usage	<p>Limitations that affect the use of data (e.g., "do not use GIS-map for navigation", "data cannot be extrapolated").</p>
Data citation	<p>Reference of the described dataset.</p> <p>A data citation should include sufficient information so that the correct version of the data can be found. The reference should contain the following recommendations according to SLU's guidelines⁷: contact person/principal investigator, release date (year), title, responsible institution/unit, update/modification date, access to data.</p> <p>Example of a data citation: Nilsson, P., The Swedish University of Agricultural Sciences (2017). <i>Productive forest land area by year (five-year average), county, table content, and age class, Table 3.2. Riksskogstaxeringen</i>. Created 2017-05-18. URL: http://skogsstatistik.slu.se/pxweb/sv/OffStat/OffStat__ProduktivSkogsmark__Areal/PS_Areal_%c3%a5lderklasser_tab.px/table/tableViewLayout2/?rxid=7f983a45-34fb-40cf-8255-739eec62ef6b</p> <p>On the project's website, remind the user to mention the source of the data in accordance with this metadata element.</p>

⁷ <https://www.slu.se/site/bibliotek/skriva-och-referera/skriva-referenser/referenslista-enligt-harvard/#Dataset>

Update of data description	Date <i>Format: YYYY-MM-DD</i>
---------------------------------------	-----------------------------------

Please note that if there are any discrepancies between the English and Swedish version of this document, the Swedish version will take precedence.