User Manual

Utfärdad av:	Dokumentnamn:		Avd:
Jan Svensson	User Manual ECDS Data Portal 2.1		B
Fastställd datum:	Fastställd av:	Dnr:	Sekretess:
2011-11-30	-	-	-

User Manual

ECDS Data Portal 2.1

Version	Date	Author	Comments
1.0	2011-05-27	Jan Svensson	
1.1	2011-10-18	Jan Svensson	Updated for "ECDS Data Portal 2.1" and some clarifications
1.2	2011-11-30	Jan Svensson	Updated for "ECDS Data Portal 2.1.2" and some clarifications

Contents

1 Introduction

1.1 ECDS

Environment Climate Data Sweden, ECDS, is a Swedish service facilitating the search, publication and long-term accessibility of data for research in the fields of environment and climate. The ECDS-initiative was previously known under the name SND-KM, Swedish National Data Service for Climate and Environment. ECDS is hosted by the Swedish Meteorological and Hydrological Institute (SMHI).

1.2 The ECDS Data portal

The ECDS data portal gives environmental and climate researchers the possibilities to
- Describe the metadata of one or more datasets
- Search for datasets from other researchers.

The data portal is based on GeoNetwork opensource (Ref. 1). The GeoNetwork software includes a lot of functionality, but may not be so easy to use for those who are unfamiliar with geodata portals or metadata. Therefore the functionality has been reduced, but can be included later on, if there is a need from the research community (see chapter 4).

A dataset is an identifiable collection of data. A dataset may be a smaller grouping of data which is physically located within a larger dataset. Theoretically, a dataset may be as small as a single feature within a larger dataset. The purpose of this data portal is not to have such a small granularity.

To get an idea about a suitable granularity for your dataset, we recommend visiting some general purpose data portals, such as the IPY metadata portal (http://gcmd.gsfc.nasa.gov/KeywordSearch/Home.do?Portal=ipy&MetadataType=0).

The standard ISO 19115 and the GeoNetwork software give the possibility to describe subsets of datasets (or subsets of subsets of datasets) in a hierarchy. We have not exploited this possibility, but it can be included later on.

1.3 The ECDS Metadata profile

The ECDS metadata profile is based on the international standard ISO 19115. It is a rather new standard from 2003 and most of the newly developed geographical metadata profiles follow ISO 19115.

We have also been influenced by the Swedish national geographical metadata profile which follows the European INSPIRE metadata profile. INSPIRE also follows ISO 19115.

SMHI is the Swedish national repository for IPY data. Swedish IPY metadata has been migrated from the IPY data portal to ECDS data portal. This means that we are also influenced by the IPY metadata profile. The IPY metadata profile was built from the Global Change Master Directory (GCMD) Directory Interchange Format.

Metadata describe the dataset and geographical metadata includes information about What-Where-When-Who-How. Our metadata profile includes the following main groups

What:
- Dataset title
- Dataset description

- Graphic illustration
- Keywords from a wide-spread thesaurus – GCMD Science keywords
- Topic category code; an ISO-standard

Where:
- Geographic bounding box

When:
- Revision of the dataset
- Time period covered by the dataset

Who:
- Who to cite regarding the dataset
- Contact information regarding the dataset
- Contact information regarding the distribution of the dataset
- Contact information regarding the metadata

How:
- Limitations regarding the use and access of the dataset
- Data format
- Web address to access the dataset
- Quality report

1.4 References

1. Geonetwork opensource, the complete manual; v. 2.4.
2. ECDS profile of ISO 19115:2003 Version 2.0; 1.0/2011-10-13

1.5 Glossary

Metadata	Metadata is "data about data". The term commonly refers to the attributes of computer data, though it may by definition refer to any information which describes other information. Geographical metadata includes information about What-Where-When-Who-How.
ISO 19115	A standard for Geographic information - Metadata
IPY	International Polar Year 2008-2010
INSPIRE	Infrastructure for Spatial Information in Europe
Thesaurus	Another name for keywords list.

1.6 News in ECDS Data portal 2.1.2 2011-11-30

Easier to select several science keywords when publishing your data
An error when searching for Where? and When? is corrected
It is now possible, for the owner of the metadata, to delete metadata for a dataset
The element "description of graphics file" is now visible
The Check-button (from Geonetwork) has been deleted (the Check-test was not easy to understand)
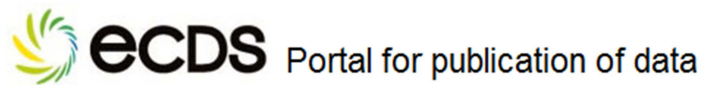Some (3) implementation errors of the ISO 19115 standard have been corrected
Possibility to send metadata information via social networks (mail, Facebook,….)

And we have updated the User manual with some more clarifications.

2 Publishing

Publishing requires a username. When you enter the Publish-page you will see the following information

When you have a username then write it and the password. If you have forget your username (userid) but remember the mailadress you used, the system can give you the username. If you know the username but have forgot your password the system can give you the possibility to create a new password. If you don't have a username you click on Register and get the following page:

User

Title:	
* Given name:	
Middle name:	
* Surname:	
* E-mail address:	
Organization/authority:	

Telephone

Address

Time & Place

Country:	Sweden
Language:	English
Time zone:	Europe/Stockholm

Acceptance & Verification

☐ I accept and approve the agreement. ⊳

* Write the text in the image:

Fb8GoT iL

Register

Mandatory elements is Given name, Surname and E-mail but we suggest you also include Organisation and Telephone. The rather strange text below (called Captcha) is used to ensure that the attempt to create a username is generated by a person.

When you have logged you have the possibility to insert metadata for a new dataset or edit your old metadata

Metadata
New metadata
Edit my metadata

We now describe how you create metadata for a new dataset (edit old metadata is quite similar)
In the first page you are required to state
- Template
- Group
-

Metadata creation

Some hints:

Template: The metadata profile to use as a template when creating your metadata. We only have one template to start with. The template is the official name of our metadata profile – ISO19115:2003 ECDS 2.0. But you may also create your own personal templates (see information later on) and then you may select this personal template.
Group: The security group to create the metadata in. We have only one group – ECDS – to start with.

Template   ISO19115:2003 ECDS 2.0 ▾

Group   ECDS ▾

Back   Create

We only have one template to start with. The template is the official name of our metadata profile – ISO19115:2003 ECDS 2.0. There is an official description of our metadata profile (see Ref 2).
You may also create your own personal templates (see below) and then you can select this personal template.

We only have one group – ECDS – to start with.

After selecting template and group (quite easy, not so many choices) click on Create to show the publishing page with all metadata elements. Before you fill in the metadata elements, read the following general information about buttons and other notifications.

Buttons:

Reset – resets all metadata elements in this metadata record.

Save – your metadata record is saved, but you can continue to work.

Save and close – your metadata record is saved and you leave the editing page and can view your saved metadata.

Save and send to reviewer – your metadata is saved and a mail is sent to the reviewer, who will review your metadata before it is made public. **During the review the metadata record is not available to you.**

Check – a formal check of your metadata record is done (a rather formal check result – don't bother)

Cancel – means that you leave the publishing page without saving your metadata record

Type (at the bottom of the page) – Type is either Metadata or Template, where Metadata is default. **Select Template to create your personal template which will be available only for you**. Useful information including in this template could be contact person information, which may be the same for most of your datasets.

Notifications:

When you move the cursor over the name of a metadata element a **?** is shown. Click the left mouse button to show help text. Click again and the help text disappears.

A red asterix means that the metadata element is Mandatory. Note: It is however possible to store metadata without every mandatory elements. You can regard it as "highly mandatory" and the reviewer will probably require that all mandatory fields ar filled in.

A red frame means EITHER that the metadata element is Mandatory and you have not written anything yet OR you have written something which is syntactically wrong, e.g. a date is written in the wrong format.

A black frame means that the metadata element is correctly entered.

Text without a frame means that a value is fixed – you can't change it

The sign **+** means that you can add another metadata element, e.g. you can have several ISO topics. To cancel – if you don't want another metadata element – click on the **x** and it will disappear. Note: Right now there is a bug in GeoNetwork. If you add an element, then cancel and delete it you cannot add an element again.

Below is a list of the metadata elements. We have chosen to use other names than the official ISO19115-names. Ref. 2 shows a cross-reference table to the corresponding ISO19115-elements.

Note: The examples are fictitious. When we get a good example from a real dataset, we will – with permission from the author – include this.

Title
The title of the dataset, which is described by the metadata. A short text describing What-Where-When might give a good title. Example: "Automatic Weather Station Data from Vestfonna Ice Cap during 2009".

Title is mandatory.

Example:

Title * | Weather Station from Vestfonna Ice Cap during 2009 |

 Revision

This is the latest date when there was a major change of the dataset. "Revision" consists of "Revision date" and "Revision type".

Revision date **is expressed as YYYY-MM-DD or YYYY-MM or YYYY**. You can also use the calendar – except for Internet Explorer where the calendar does not work.

Revision type is the event used for revision date. That is what happened to the dataset at the revision date. Can be either Creation (when dataset was created), Publication (when dataset was public) or Revision (a major change in the dataset).

Revision is mandatory.

Example:

Revision

| Revision date * | 2010-02-01 | Clear |
| Revision type * | Creation ▾ | |

Dataset version

It is good practice to have a version number of the dataset, if the dataset is changed now and then. This means that someone using your data knows what dataset (what version) has been used in their research.
You can use your own notation of the version number; e g 1, 2, 3… or A, B, C…

Dataset version is optional.

Example:

| Dataset version | 1.0 |

Dataset citation

This is the reference that is used if someone has used your dataset and it is cited in an article. References may be provided in styles used by professional scientific journals. The organisation DataCite deals with citation of datasets; see http://www.datacite.org

Dataset citation is optional.

Example:

| Dataset citation | Karlsson, K., Measurements at Vestfonna, Polar Observation Journal, 2010 |

Dataset description

Brief description of the dataset **(Note: Description of the dataset – not the project result).**
You should use words that you think will be found by someone using the Free-text-search.

Dataset description is mandatory
Example:

| Data set description * | Automatic weather station data from the ablation zone Vestfonna ice cap including Air temperature, Relative humidity, Ice temperature and Snow temperature. Observations are done each hour. |

Contact information regarding the dataset.

This is the first one of three required contact information. The three contacts are
- regarding the dataset
- regarding distribution
- regarding metadata

It may seem overambitious to have three different kinds of contacts, but ISO19115 clearly separates this kind of contact information. However if you are the one and only contact person, regardless if it is about the dataset, distribution or metadata, then we suggest that you create a personal template with you as the contact person for every kind of contact.

"Contact information regarding the dataset" is who to contact regarding questions about the dataset. Each "Point of contact" consists of Individual name, Organisation, Electronic mail and Role. You may have several points of contact.

Name (surname and first name) of the person to contact.

Organisations are more permanent than individuals so please insert the organisation name.

E-mail to the organisation/person to contact.

Several roles are allowed. We suggest you use pointOfContact, but it might also be interesting to show the principle investigator, so you could add Principle investigator and thus have two points of contact regarding the dataset. **If you use Principle investigator the Organisation should be the Project name.** The reason to include the project name is that anyone can use the project name in the free text search to find the datasets from your project.

"Contact information regarding the dataset" is mandatory.
Example:

Graphic illustration

Consists of
- Link to the graphic file
- Description of the graphic file
- Graphic format

Provides a graphic which illustrates the dataset. Could be a photo from the site or a graph showing a time series of the data or anything which you think illustrates the dataset
- Write the name of the file which contains the files. It should be a public URL.
- Include a text description of the graphic that illustrates the dataset
- Describe the format in which the illustration is encoded. Examples: JPG, TIFF

The graphic illustration is assumed to have an approximative size of 800x800 pixels. Other sizes may give a distorted graph.

"Graphic illustration" is optional

Example:

Graphic illustration

Link to the graphic file	http://vestfonna.site.se/picture
Description of the graphic file	The tower where the instruments are located
Graphic format	JPG

Descriptive keywords from thesaurus

Provides keywords. Right now we only have one thesaurus - "GCMD Science keywords" – see http://gcmd.nasa.gov/Resources/valids/archives/GCMD_Science_Keywords.pdf. "GCMD Science keywords" is a hierarchical keyword list which can be expanded by clicking on the +. If you want to use several keywords, then press down the CTRL-key and select several keywords. The selected keywords have a light blue background.
Example:

Descriptive keywords from thesauris

GCMD Science Keywords *

- GCMD
  - Agriculture
  - Atmosphere
  - Biological Classification
  - Biosphere
  - Climate Indicators
  - Cryosphere
    - Frozen Ground
    - Glaciers/Ice Sheets
    - Sea Ice
    - Snow/Ice
  - Human Dimensions
  - Land Surface
  - Oceans
  - Paleoclimate

Thesaurus name

Title	GCMD - Science keywords
Revision date	2008-02-05
Revision type	Revision

Here you have selected
  - Atmosphere (highest level)
  - Cryosphere>Frozen ground
  - Cryosphere>Sea ice
If you change your mind and do not want the key word Cryosphere>Frozen Ground just click on this line (still with CTRL key pressed) it will be deleted and the light blue background will disappear.

Title is fixed = GCMD - Science keywords
Reference date is fixed = 2008-02-05 (which was the last time the GCMD Science keywords were revised)
Reference date type is fixed = Revised

"Descriptive keywords from thesaurus" is mandatory.

Use limitation

Limitation affecting the fitness for use of the dataset, for example "The quality of the data is not good enough to be used as input observation data in climatological models".
If there is nothing relevant to say, then write "No known limitations".

"Use limitation" is mandatory.

Example:

Use limitation * ⊞   No known limitations

Access constraints

Several choices e.g. Copyright or Patent. We suggest you use Other restrictions and use the next metadata element (that is "Other constraints").

Access constraints is mandatory.

Example:

Access constraints * ⊞   Other restrictions ▾

Other constraints

We suggest you use the Creative Commons License and write "Creative Commons Attribution (see http://creativecommons.org/licenses)". This means (in short) that this license lets others distribute, remix, tweak and build upon your work, even commercially, as long as they credit you for the original creation. There are in total six Creative Commons Licenses to choose between. If neither fits then you may write something like "Constrained; please contact the contact person for the dataset for more information".

"Other constraints" is mandatory.

Example:

Other constraints * ⊞   Creative Commons Attribution

Language

Language used in the dataset.

Dataset language is mandatory. Default is English.

Example:

Language * ⊞   English ▾

Topic category code

High-level geographic theme from http://gcmd.nasa.gov/User/difguide/iso_topics.html. This is central to the standard ISO 19115. Select one or several topics from the list.

Topic category code is mandatory.

Example:

Topic category code * ⊞   ▾

## Geographic bounding box

Geographic position for the dataset. The bounding box is expressed as

- North bound (-90, +90)
- East bound (-180, +180)
- South bound (-90, +90)
- West bound (-189, 180)

Use the map to create the box or write the bounding latitudes/longitudes (use decimals, not minutes and seconds). You can move around and zoom in the map or use the roll-bar to the right and select country/region. If a country/region is selected, a bounding box, covering the country/region, is drawn on the map.

Geographical extent is mandatory

Example:
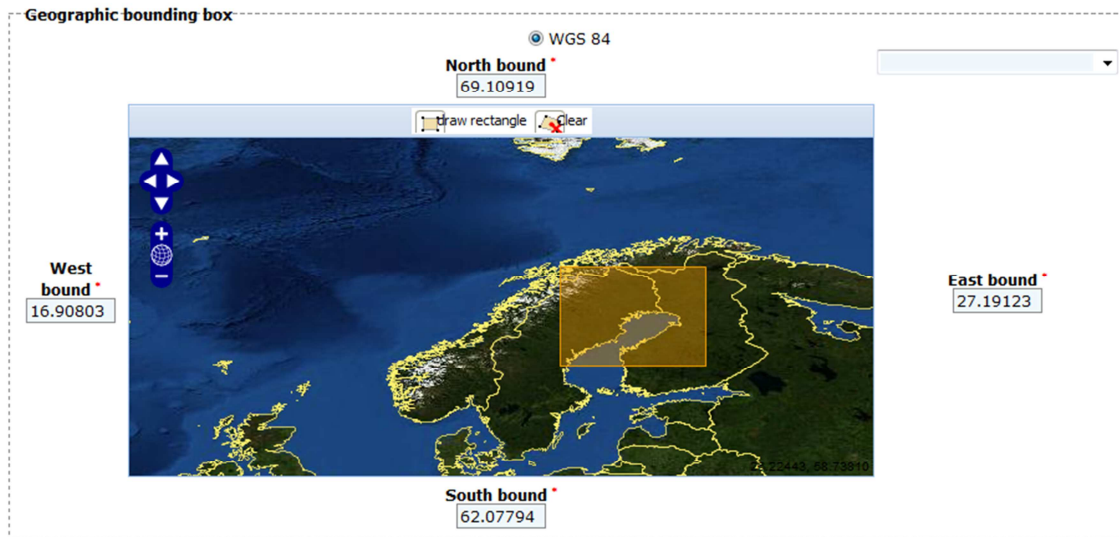
Temporal extent

Time period of the content of the dataset.

The time reference system is fixed = Gregorian.

Begin date and end date is expressed as YYYY-MM-DD or YYYY-MM or YYYY. You may use the calendar or write it as it is.
Note 1: **The calendar may not work in Internet Explorer (a bug)**
Note 2: What End date should be used when the data is continuously updated? We suggest you write the present year at the start of each new year (e.g. 2011)

Temporal extent is mandatory.

Example:

Additional information regarding geographic, temporal and vertical extent

Above you find metadata elements describing
- Geographic extent
- Temporal extent

There is an element in ISO19115 describing vertical extent. We have not used this in our metadata profile, but information about vertical extent can be described here.

You may use "additional information" when the extents above is not enough to describe the dataset. E g:
- Geographical extent: Observations from the river of Donau
- Temporal extent: Data is covering the Cenozoic era.
- Vertical extent: The measurements are made from a 10 m tower

"Additional information regarding geographic, temporal and vertical extent" is optional.

Example:

Additional information regarding geographic, temporal and vertical extent	Observations from the river of Lule älv The measurements are made from a 10 m tower

Contact information regarding distribution

This is the second of three required contact information entries, and describes who to contact regarding distribution of the dataset. Each "Point of contact" consists of Individual name, Organisation, Electronic mail address and Role. You may have several point of contacts.

Name (surname and first name) of the person to contact.

Organisations are more permanent than individuals so please insert the organisation name.

E-mail to the organisation/person to contact.

Several roles are allowed. We suggest you only use Distributor.

"Contact information regarding distribution" is optional.

Example:

Individual name *	Helpdesk at World Data Centre Kiruna
Organisation name	World Data Centre Kiruna
E-mail * ⊞	helpdesk@wdc-kiruna.se
Role *	Distributor

Data format

Consists of
- Data format name; e.g. Excel, Shape, GRIB.
- Data format version; e.g. 2003
"Data format" is optional.

Example:

Data format name	Excel
Data format version	2010

On-line resource

Information about on-line resources from which the dataset or other information about the dataset can be obtained; e.g. http://climatdata.se. On-line resource consists of
- URL (web address)
- Description – a short description about the on line resource.

Note: For users "with small needs and limited resources" there is the possibility to store data at ECDS. This functionality will be added later.

"On-line resource" is optional.

Example:

Data quality information

Provides overall information about the quality of the content in the dataset. Quality report consists of
- Scope of the quality report
- Quality report

ISO 19115 offers several choices for the scope. We have used Dataset and it is fixed.

Quality report is a general text about the quality of the data, such as which instruments have been used, the process when the data are created, and the quality control of the data.

Data quality information is mandatory. If it is not possible to write a quality report then write something like "Not known" or "Not documented".

Example:

Metadata information

The information about metadata is either fixed or created automatically. Metadata information consists of
- Metadata identifier – created automatically. A unique identifier (UUID) for this metadata set.
- Metadata language – fixed = English
- Scope of the metadata – fixed = Dataset.
- Date stamp – created automatically.
- Name of the metadata standard – fixed = ISO 19115:2003 ECDS
- Version of the metadata standard – fixed = 2.0

Example:

Metadata identifier	92b464c8-e957-4cd7-a28c-8beb4663dd84
Metadata language	English
Scope of the metadata information	Dataset
Date stamp	2011-05-27T05:41:08
Name of the metadata standard	ISO 19115:2003 ECDS
Version of the metadata standard	2.0

Contact information regarding metadata

This is the third of three required contact information entries, and describes who to contact regarding the metadata for this dataset. Each "Point of contact" consists of Individual name, Organisation, Electronic mail address and Role. You may have several points of contact.

Name (surname and first name) of the person to contact.

Organisations are more permanent than individuals so please insert the organisation name.

E-mail to the organisation/person to contact.

Several roles are allowed. We suggest you only use "Point of contact".

"Contact information regarding metadata" is mandatory.

Example:

```
┌─ Contact information regarding metadata ─────────────────────────────────────┐
│ Individual   [Will Smith                    ]  E-mail * ⊞   [will.smith@mail.com ] │
│ name *                                                                        │
│ Organisation [Uppsala University            ]                                 │
│ name                                                                          │
│ Role *       [ Point of contact    ▼ ]                                        │
└──────────────────────────────────────────────────────────────────────────────┘
```

Now you are ready with your dataset. You can Save it and come back later to finalize your metadata. Or you can send it to the Reviewer. During the review the metadata record is not available for you. But otherwise you can always go back and edit your metadata.

Once you have saved your metadata and edit the metadata, the there is a Delete-button you can use to delete the metadata record for a dataset.

3 Searching

Simple search
To make a simple search, start by clicking on **Search** in the menu.

Then enter text in the **Free text** input box to the left and click the **green search button** at the bottom.

Example:
Text Karin finds every dataset which contains the word Karin (which may be an author's name)
Text IPY finds every dataset which contains the word IPY (which mean every IPY-project)
Text IPY-026 finds every dataset which contains the word IPY-026 (which means every dataset from the project IPY-026)

You may use ? as a wild card.

Example: Text Creativ? finds (among others) every data set which use Creative Commons licence. However – GeoNetwork use a rather advanced algorithm for search with wild card, so the result are not always predictable!

The search result is show in a list to the left.

Advanced search

There are three ways of making a more advanced search.

By keywords:
Under the Free text field you will see a box containing a hierarchal list with all keywords that are connected to a public dataset.
To make a keyword search, click on the **plus button** to go down through the hierarchy.
Select any level in the hierarchy by **clicking on the text**. It will be marked with a light blue background. If you want to select on several keywords keep the CTRL key pressed while selecting the keywords. Then click the **green search button** to make a search for this keyword.

By geographical area:

Start by clicking on **Where?** under the keywords box, to display the map.

By date:
Start by clicking on **When?** under the keywords box, to display two input fields (Start and Stop date).
Note: **There is a calendar available. But it may not work for every web browser.**

The search algoritm is

"Free text" AND "Keywords" AND "Geographical area" AND "Date"

Example:
Here we have selected

Free text - IPY
Keywords
- Oceans
- Atmospheric Temperature
- Atmospheric Water Vapor
Area – as shown on the map
Time from 2005 to 2010

Search results

The search results are displayed to the right on the search page. If there are many results they are split into pages. Click through the pages by the navigation at the bottom of the page.

To see the full metadata for a result in the list, click the metadata button. To hide the metadata again, click the metadata button again.

Example:

Aggregate Results matching search criteria : 1 - 9 / 9 (page 1 / 1), 0 selected

Select : <u>all</u> , <u>none</u>   ⊞ actions on selection

Sort by   Title

☐ **AGAVE2007 TRACK**

Abstract   Icebreaker Oden\\\\\\\\\\\\\\\'s trackline during the Arctic Gakkel Vents Expedition (AGAVE) 2007.

Keywords   Oceans > Marine Sediments > Stratigraphic Sequence

⊞ Metadata

☐ **ASCOS WEATHER STATION ON ODEN**

Abstract   ASCOS was an interdisciplinary expedition to the central Arctic on the icebreaker Oden from 1 August to 9 September 2008. The expedition focused on processes relating to cloud formation in the summer ...

Keywords   Atmosphere > Air Quality > Visibility , Atmosphere > Atmospheric

Hint: Generally the metadata records are sorted in alphabetical order of the Title. If you sort by "Change date" those metadata records, which have changed lately, will appears at the top.

The full metadata is shown when you press the metadata button.
Example: The first lines from "Agave 2007 Track". Note that you can distribute the metadata record by different social media e g Facebook or an ordinary e-mail. You can distribute metadata from your datasets or metadata for other dataset which you want to share. **But – once again – it may be some problem with Internet Explorer.**

Metadata

Data set

Title	AGAVE2007 track

Revision

Revision date	2007-07-01
Revision type	**Creation**: Date identifies when the resource was brought into existence

Dataset version	1.0
Dataset citation	Sohn, R. A., Willis, C., Humphris, S., Shank, T., Singh, H., Edmonds, H. N., Kunz, C., Hedman, U., Helmke, E., Jakuba, M., Liljebladh, B., Linder, J., Murphy, C., Nakamura, K.-i., Sato, T., Schlindwein, V., Stranne, C., Tausenfreund, M., Upchurch, L., Winsor, P., Jakobsson, M., and Soule, A. (2008). Explosive volcanism on the ultraslow-spreading Gakkel Ridge, Arctic Ocean. Nature 453, 1236-1238.
Data set	Icebreaker Oden\\\\\\\\\\\\\\\\\'s trackline during the Arctic Gakkel Vents

4 Future functionality

Additional functionality may be added in the future, depending on the response from the research community. Some changes we foresee are:

- The possibility for users "with small needs and limited resources" to store their data at ECDS
- Include "Inspire themes" and some more data-elements from the INSPIRE metadata profile into our metadata profile.
- Include other time reference systems besides the Gregorian calendar
- Vertical extent, described by official and well-known reference systems, may be interesting for several research areas
- For some users which are already defined in the user database the authentization may not work. Until this bug is fixed we can create a user name for you manually. Please contact our Helpdesk.
- The "Graphic illustration" are taken from a web site defined by you. A more stable solution may be that we fetch your graphic and store it at ECDS. To be discussed.

More information (including templates and/or examples) may be useful regarding

- Dataset citation using practices from DataCite
- Creative Commons License
- Quality report