

(Deprecated) O2A specification for GeoCSV (.sdi.tab)

- [Version 2.0](#)
 - [Data Files \(.sdi.tab\)](#)
 - [Examples](#)
 - [Metadata Files \(.sdi.meta.json\)](#)
 - [Examples](#)
 - [File Naming, File Linking](#)
 - [Examples](#)

Warning: This specification is deprecated and is only further used for already existing data products and workflows.

The commonly known [GeoCSV](#) format extends the CSV format with a geometry column (usually with [WKT](#) notation). The O2A GeoCSV specification extends the GeoCSV format with special requirements enabling it to be used in the automated O2A dataflow. Originally it was built upon (but should not be confused with) the [NRT Data Format](#) but has undergone major changes since then. Also it's heavily influenced by PANGAEA formats (Geocodes, .tab format).

Version 2.0

When storing data in O2A GeoCSV format, data and metadata go different ways. Data (including their spatio-temporal location) go into a data file and metadata go into a metadata file. Data files and metadata files are linked together by file names. Multiple data files can be linked to the same metadata file.

The concept of parameter URNs (version 1.x) has been dropped.

Data Files (.sdi.tab)

Data files are [GeoCSV](#) files using the [WKT](#) notation for the description of the geometry. However, there are additional requirements.

An O2A GeoCSV data file starts with five columns holding most of the data's spatio-temporal information. The sixth column holds an event reference. Seventh to last but one column can contain actual data. The last column holds the horizontal coordinates as WKT notation. The following table gives an overview and detailed information.

Also some more general requirements and notes

- requirements
 - file extension: .sdi.tab
 - decimal separator: . (point)
 - column separator: \t (tab) no tab in values or column headers
 - column names need to be unique
- notes
 - white spaces are allowed both in cells and column names
 - columns can be left out entirely if they do not keep any values (e.g. date_time_end)
 - empty cells in non-value-mandatory columns are fine, if information is unknown
 - will default to NULL (0) internally
 - rows with missing or invalid mandatory values will be ignored
 - column order matters (although some columns might be left out)

column type /group	column order	column header	value is mandatory	description	example values
spatio-temporal location	1	date_time_start	yes	Date and time of measurement in ISO 8601 format notation, using UTC time zone, without fractions of seconds. Or start of time range.	valid: 2019-02-28T15:50:00 invalid: 2019-02-28T15:50:00.000 invalid: 2019-02-28 15:50:00
	2	date_time_end	no	End of time range of measurement in ISO 8601 format notation, using UTC time zone, without fractions of seconds.	see above
	3	elevation [m]	no	Elevation in meter. A negative value means below sea level, while positive value means above sea level. See Pangaea Geocode definition . Note: This is not the height/depth of the measurement (unless it's taken on earth's surface) but the topographical elevation at the lat/lon position.	
	4	z_value [m]	no	Vertical position of the measurement, in meter (third spatial dimension).	

	5	z_type	yes, if z_value [m] is given	Pangaea Geocode to describe the type of z_value [m].	valid: "DEPTH, water" valid: "DEPTH, sediment/rock" valid: "HEIGHT above ground" invalid: "HEIGHT above aeroplane"
metadata reference	6	event_name	yes	Name of event. Reference key for metadata. event_name must match one event name in metadata file (if metadata file is used).	PS1010-1
data		<parameter> [<unit>]	no	Arbitrary amount of columns (at least one) with (measurement) data. Each column name has to start with the parameter/phenomenon name followed by a unit in square brackets. <parameter> and [<unit>] are separated by a single whitespace. Reference key for metadata. <parameter> must match the parameter name in metadata file (if metadata file is used).	
spatio- temporal location	last	geometry	yes	Geometry in WKT notation without third spatial dimension. The reference system needs to be EPSG:4326 and the unit is decimal degrees. Longitude comes first, latitude second. The geometry type can be chosen freely. However, a simple POINT is usually the best choice.	POINT (123.45678 -20.12345) MULTILINESTRING ((8.58 53.55, 8.58 53.56, 8.57 53.55), (8.0 53.0, 9.0 54.0, 8.0 54.0))

Examples

PANGAEA-inspired example

simple example with multiple parameters (smoothed for readability)				
date_time_start	z_value [m]	z_value_type	event_name	Pressure, at
given altitude [hPa]	Temperature, air [°C]	geometry		
1982-12-29T11:02:00	10	Altitude	PS01/00001	
1035.0		8.3		
POINT(-4.3 49.6)				
1982-12-29T11:45:00	956	Altitude	PS01/00001	
921.4		0.9		
POINT(-4.3 49.6)				
1982-12-29T13:21:00	1035	Altitude	PS01/00001	
912.4		0.2		
POINT(-4.3 49.6)				

Inspired by: <https://doi.pangaea.de/10.1594/PANGAEA.382336>

Download of proper example data file: [example-1.sdi.tab](#)

Empty file with complete header

date_time_start	date_time_end	elevation [m]	z-value [m]	z-type
event_name	<parameter> [<unit>]	<parameter> [<unit>]	geometry	

Download of template data file: [template.sdi.tab](#)

Metadata Files (.sdi.meta.json)

Metadata files are [JSON](#) files. There's a fixed structure with the possibility to add custom metadata.

On the top-level only these fixed keys/keywords are allowed: version, events, parameters, expeditions, platforms, projects, meta. version holds the used version of this specification, meta holds information valid for the whole dataset and the other keys hold lists of according elements. Those elements have their own fixed keys/keywords, including meta. Some fixed keys reference objects in other lists.

- events > expedition expeditions > name
- events > platform platforms > name
- meta > projects projects > name

Meta is a special key to hold custom metadata. It has some fixed keywords but you can add as much of your own custom key-value pairs as you like. However, this data will only be displayed but not filterable.

The whole metadata file is optional. If the only metadata you want to have attached to your data is an event name, it is totally sufficient to have that in your data file. When using a metadata file, everything is optional except a version, a list of events with at least one event. Also, every entry in all of your lists (events, parameters, expeditions, platforms, projects) needs to have a name. If one of your list entries references to an entry in one of the other lists (see bullet points above) and that one does not exist it will be interpreted as an entry with just a name (see examples). Keys with empty values (<key> : "") are fine but useless and will be interpreted as if this key-value pair would have been left out.

Also, please read the [JSON specs](#) to know about things like how to escape special characters.

top level keys	second level keys	mandatory	description	example values
version		yes	used O2A Spatial GeoCSV specification	"2.0"
events		yes		
	name	yes	event name, serves as metadata reference to data file	"PS01/00001"
	alias		event alias	
	expedition		name of expedition the event is part of, references entry in expeditions list	"ANT-I/1"
	platform		used platform, references entry in expeditions list	"Polarstern"
	device		used device	"Radiosonde (RADIO)"
	uri		URI/URL	
	meta		key-value pairs for custom event metadata (see next table)	
parameters		no		
	name	yes	parameter name (reference key)	"Temperature, air"
	alias		parameter alias	"TTT"
	unit		unit of measurement	"°C"
	method			
	uri		URI/URL	
	meta		key-value pairs for custom parameter metadata (see next table)	
expeditions		no		
	name	yes	expedition name (reference key)	"ANT-I/1"
	alias		expedition alias	"PS01"
	uri		URI/URL	" https://doi.org/10.2312/BzP_0014_1983 "
	meta		key-value pairs for custom metadata (see next table)	
platforms		no		
	name	yes	platform name (reference key)	"Polarstern"
	alias		platform alias	
	uri		URI/URL	" https://doi.org/10.17815/jlsrf-3-163 "
	meta		key-value pairs for custom metadata (see next table)	
projects		no		
	name	yes	project name (reference key)	"Meteorological Long-Term Observations @ AWI"
	alias		project alias	"AWI_Meteo"

	uri		URI/URL	" http://www.awi.de/en/science/long-term-observations.html "
	meta		key-value pairs for custom project metadata (see next table)	
meta		no	key-value pairs for custom dataset metadata (see next table)	

meta keys	description	example values
pi_name	name of principle investigator	"König-Langlo, Gert"
pi_email	mail of principle investigator	"gert.koenig-langlo[at]awi.de"
pi_url	homepage of principle investigator	" http://www.awi.de/en/about-us/organisation/staff/gert-koenig-langlo.html "
pi_orcid	ORCID of principle investigator	" https://orcid.org/0000-0002-6100-4107 "
comment		"Height of tropopause 11650 m"
citation		"König-Langlo, Gert (1983): Radiosonde PS01/00001 during POLARSTERN cruise ANT-I/1 on 1982-12-29 11:24h. Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA, https://doi.org/10.1594/PANGAEA.382336 , In: König-Langlo, G (1983): Upper air soundings during POLARSTERN cruise ANT-I/1. Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA, https://doi.org/10.1594/PANGAEA.853633 "
project	references entry in projects list	
license		"Creative Commons Attribution 3.0 Unported (CC-BY-3.0)"
metadata_url	link to metadata	" https://doi.pangaea.de/10.1594/PANGAEA.382336?format=metadata_jsonld "
data_url	link to data	" https://doi.pangaea.de/10.1594/PANGAEA.382336?format=textfile "
sop_url	link to SOP information	

Examples

Minimal example

valid example with two versions of same metadata

```
// Minimal example of a metadata file (.sdi.meta.json), only containing version and one event with an expedition name
{
  "version": "2.0",
  "events": [
    {
      "name": "foo",
      "expedition": "bar"
    }
  ]
}

// Less minimal representation of the same metadata. Both versions are valid.
{
  "version": "2.0",
  "events": [
    {
      "name": "foo",
      "expedition": "bar"
    }
  ],
  "expeditions": [
    {
      "name": "bar",
      "alias": ""
    }
  ],
}
```

PANGAEA-inspired example

Inspired by: <https://doi.pangaea.de/10.1594/PANGAEA.382336>

Example #1

```
{
  "version": "2.0",
  "events": [{
    "name": "PS01/00001",
    "expedition": "ANT-I/1",
    "platform": "Polarstern",
    "device": "Radiosonde (RADIO)",
    "meta": {
      "location": "English Channel"
    }
  }
],
  "parameters": [
    {
      "name": "Pressure, at given altitude",
      "alias": "PPP",
      "unit": "hPa",
      "meta": {
        "pi_name": "König-Langlo, Gert",
        "pi_email": "gert.koenig-langlo[at]awi.de",
        "pi_orcid": "https://orcid.org/0000-0002-6100-4107",
        "pi_url": "http://www.awi.de/en/about-us/organisation/staff/gert-koenig-langlo.html",
      }
    }, {
      "name": "Temperature, air",
      "alias": "TTT",
      "unit": "°C",
      "meta": {
        "pi_name": "König-Langlo, Gert",
        "pi_email": "gert.koenig-langlo[at]awi.de",
        "pi_orcid": "https://orcid.org/0000-0002-6100-4107",
        "pi_url": "http://www.awi.de/en/about-us/organisation/staff/gert-koenig-langlo.html",
      }
    }
  ],
  "expeditions": [{
    "name": "ANT-I/1",
    "alias": "PS01",
    "uri": "https://doi.org/10.2312/BzP_0014_1983"
  }
],
  "platforms": [{
    "name": "Polarstern",
    "uri": "https://doi.org/10.17815/jlsrf-3-163",
  }
],
  "projects": [{
    "name": "Meteorological Long-Term Observations @ AWI",
    "alias": "AWI_Meteo",
    "uri": "http://www.awi.de/en/science/long-term-observations.html",
  }
],
  "meta": {
    "citation": "König-Langlo, Gert (1983): Radiosonde PS01/00001 during POLARSTERN cruise ANT-I/1 on 1982-12-29 11:24h. Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA, https://doi.org/10.1594/PANGAEA.382336, In: König-Langlo, G (1983): Upper air soundings during POLARSTERN cruise ANT-I/1. Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA, https://doi.org/10.1594/PANGAEA.853633",
    "license": "Creative Commons Attribution 3.0 Unported (CC-BY-3.0)",
    "comment": "Height of tropopause 11650 m",
    "metadata_url": "https://doi.pangaea.de/10.1594/PANGAEA.382336?format=metadata_jsonld",
    "data_url": "https://doi.pangaea.de/10.1594/PANGAEA.382336?format=textfile",
    "sop_url": "",
  }
}
```

Empty file with complete keywords

Download: [template.sdi.meta.json](#)

File Naming, File Linking

Metadata files follow this naming pattern: <basename>.sdi.meta.json.

Data files follow this naming pattern: <basename>[@<handle>].sdi.tab.

All data files with the same <basename> are associated with the corresponding metadata file. The @<handle> can be used to have multiple data files with the same <basename>. Files can only have up to one handle, and '@' cannot be used anywhere else in filenames.

Examples

three valid examples

```
./path/to/data
|-- foo.sdi.meta.json
|-- foo@1999.sdi.tab
`-- foo@2000.sdi.tab
```

```
./path/to/data
|-- foo.sdi.meta.json
|-- foo.sdi.tab
|-- bar.sdi.meta.json
`-- bar.sdi.tab
```

```
./path/to/data
|-- foo.sdi.meta.json
|-- foo@part1.sdi.tab
|-- foo@part2.sdi.tab
|-- bar.sdi.meta.json
`-- bar.sdi.tab
```