NRT Data Format

Introduction

The O2A components INGEST and DASHBOARD (and its Data Web Service) use a common data format to exchange time-series data: The NRT (Near Real Time) Data Format.
It uses the SENSOR's parameter codes (URNs, Uniform Resource Names), e.g. `vessel:mya_ii:dgps:lat` to describe the data columns in the files.
The data is stored in a simple tabular plain-text format which is describe below.
If a sensor cannot provide data NRT Data Format but a binary or plain-text format instead, a *driver* (converter) will have to be programmed to convert the data either to a intermediate NRT RAW Data format or to this NRT Data Format.
A number of drivers is already available (see subchapter Drivers).

The following O2A components are involved in / use NRT Data Format.

Web-app	Web-app's task
Sensor (sensor.awi.de)	Provides parameter code as URN, e. g. `vessel:mya_ii:dgps:lat`
NRT Ingester (http://ingestsrv2.dmawi.de/)	Auto-ingests and stores sensor's raw data, executes drivers and forwards to Data-WS
Data-WS (dashboard.awi.de/data/)	Data Web Service having all time-series data and mean values, visualization frontend for time series (DASHBOARD) and REST interfaces for data in/output

Format version 2

Format version 2 has been introduced on 01-03-2019 and replaces format version 1.
The component Data-WS only uses format version 2 why it is highly recommended to use this as the default format.
Format version 1 is deprecated. It is still supported by NRT Ingester for old ingest jobs but should not be implemented in new projects.

Example

```
datetime    vessel:polarstern:tsk1:salinity [psu]    vessel:polarstern:tsk1:sbe38:temperature [°C]
2019-02-28 15:50:00.000    34.1234    2.443
2019-02-28 15:50:01.000    34.1345    2.564
2019-02-28 15:50:02.000    34.1456    2.544
```

Format

Plain-text table having columns for each parameter and rows for each individual time record.
A sample file is attached: nrtformat2-sample_adcp_20190217_235334.nrt

What	How
Character encoding	UTF-8
Field / column separator	Tabulator "\t"
End of record marker	New line "\n"
Header	The header must consist of a column named "datetime" and 1-n columns having a valid parameter URN (from Sensor.awi.de) as column header field. The "datetime" column must always be the first column. `datetime    type:dev:parameter1    type:dev:parameter2`
Header - Units	Units may be given in square brackets [], following each parameter URN, separated by a space " ": `datetime    type:dev:parameter1 [°C]    type:dev:parameter2 []    type:dev:stationname [text]` Units are optional depending on the O2A component and application. • Data-WS exports the unit in a data request output (or [] if not defined). • Data-WS ignores a passed unit when importing data • NRT Ingester passes units to the Data-WS. Only if a new (not existing) parameter URN is created in the Data-WS the unit is added to that. Existing parameter URNs and their existing units are not updated by a change of the unit.

Quality Flags	Quality flags can be added to each data value. These are numbers from 0 to n defined in the flagging scheme.
	Each data column needs an additional column titled with the same parameter URN followed by “ `(quality_flag)`”:
	
	`datetime  vessel:mya:temp [°C]  vessel:mya:temp (quality_flag)`
	`2019-02-28 15:50:00  56.00000  1`
	`2019-02-28 15:50:01  566.0000  2`

Date time format	Expected time zone: UTC
	
	Allowed date formats:
	
	`yyyy-mm-dd HH:MM:SS.fff`
	`yyyy-mm-dd HH:MM:SS`
	`yyyy-mm-ddTHH:MM:SS.fff`
	`yyyy-mm-ddTHH:MM:SS`

Decimal symbol	Dot "."
Data values	Decimal values, integer numbers or characters (just no tabulators!) `34.56556  23  334.43E-2  SAMPLE1` For use of Data-WS only: The Data-WS only accepts numbers as input for each column except when its unit is "text"! For columns which do not have the unit "text" all characters or words (such as “inf”, “NaN”, “SAMPLE1”) will result in an import error.
Error values	Empty field (simply give a null string, but do not forget to add tabs as data field separator)

Recent changes

29.06.2020

- NRT Format version 2 now supports strings/texts as data values and also the Data-WS does now!
  Pass "`[text]`" as the column's unit to make it a text column.

Format version 1

Format version 1 is deprecated since 28-02-2019. It is still supported by NRT Ingester for old ingest jobs but should not be implemented in new projects.

Example

time;vessel:heincke:phins:latitude,[deg];vessel:heincke:tsg:salinity,[psu]

2019-02-28 15:50:00;56.00000;-49.05454

2019-02-28 15:50:01;56.34300;-49.00122

2019-02-28 15:50:02;56.452;-49.31232

Example without full parameter URN:

In this case the NRT Ingester can prepend a URN prefix (set in the configuration, example “vessel:heincke”) to all parameter URNs before exporting the data to the NRT Database.

time;phins:latitude,[deg];tsg:salinity,[psu]

2019-02-28 15:50:00;56.00000;-49.05454

2019-02-28 15:50:01;56.34300;-49.00122

2019-02-28 15:50:02;56.452;-49.31232

Format

Plain-text table having columns for each parameter and rows for each individual time record.

Field separator: Semicolon “;”

End of record marker: New line “\n”

Header: The header must consist of a column named “time” and 1-n columns having a valid parameter URN (from Sensor.awi.de) as column header field, or an incomplete parameter URN prepended by a URN prefix before importing it to NRT Database. A unit may be given as well.

time;type:dev:param1,[unit1];type:dev:param2,[unit2]

Date time format: Expected time zone: UTC

yyyy-mm-dd HH:MM:SS

Decimal symbol: Dot (.)

Data values: Decimal values, integer numbers or character (without containing tabulators)

34.56556;23;334.43E-2;SAMPLE1

Error values: Empty field (simply give a null string, but do not forget to add enough semicolons)

Changes from version 1 to version 2

- Field separator tabulator „\t“ instead of semicolon „;“
- Unit format changed from “parameterURN;[unit1]” to “parameterURN [unit1]”
- The column „datetime“ has to be the first column and has to be “datetime” instead of “time”
- Quality flag support added

Drivers

Currently the following drivers are available to convert sensor's raw data to NRT Data Format:

Name	Sensor types	Language
adcpToNrt	RDI ADCP instruments (format pd0 data format such as contained in files .enc, .enx, ...)	Python
ctdToNrt	Sea & Sun Technologies's CTD (format provided by STA acquisition software)	Python
ferryboxToNrt	4h-Jena Ferrybox	Python
sbe38ToNrt	Seabird SBE38 temperature probe (format provided by probe itself via serial line)	Python
snowbuoyToNrt		Python
toa5ToNrt		Python

Drivers can be programmed and installed in the NRT Ingester environment. Python is currently the prefered language.