

aivis Data Specification

Version 1, 19.04.2021

1. Data Quality

- Raw data as little processed as possible
- Especially no record should be removed or interpolated

2. File Format

Data is transferred via CSV files with file extension “.csv”.

Generally, the following rules apply:

- Filenames must be unique when sanitized (see below)
- Charset must be UTF-8
- Records must be separated by CR LF (Windows line ending) or LF (Unix line ending)
- Fields must be separated by comma
- Text values must be enclosed in quotation marks, if they contain literal line endings, commas or quotation marks
- Quotation marks inside such a text value have to be prefixed (escaped) with another question mark
- The first line of each CSV file represents the header, which must contain column headers that are file-unique when sanitized (see below)
- Every record including the header must have the same number of fields
- If a column is empty, the value is regarded to be unknown and is left out on calculations
- Values are interpreted either as 64bit Float (written as in XML Schema) or arbitrary text
- By default, all numbers are treated as numerical data and arbitrary text is treated as categorical data – but it is possible to switch numerical data to categorical (e.g., for numerical device ids)
- While reading CSV data, the filename without extension and the column header is sanitized (all non-alphanumeric characters are converted to “_”) and concatenated via “.” to be called the column id

2.1 Timeseries data

On timeseries data the following rules apply additionally:

- Each record must hold the detection time in its first field
- If availability consideration is enabled, each record must hold the availability time in its second field
- The detection and availability time is represented as milliseconds since Linux Epoch UTC and formatted as unsigned 64bit Integer (written as in XML Schema)

2.2 Tabular data

On tabular data the following rules apply additionally:

- Since there is no way to correlate multiple records from different files yet, only one file can be used.

3. Metadata

Besides the data files also a file called “metadata.json” can be added to set column settings programmatically which can be also done in the UI.

```
{
  "columnSettings": [ (type: list, optional, default: empty list)
    {
      "id": "csvfilename.force", (type: string)
      "type": "NUMERICAL", (type: "AUTO" | "CATEGORICAL" | "NUMERICAL", optional, default: "AUTO")
      "name": "Force at wheel", (type: string | null, optional, default: null)
      "description": "Force applied to the wheel ...", (type: string | null, optional, default: null)
      "unitName": "Newton", (type: string | null, optional, default: null)
      "unitSymbol": "N" (type: string | null, optional, default: null)
    }, ...
  ]
}
```