Data format proposal for CRP-J02017

Tamás Holczer, PhD, Hungary

Roland Nádor, Hungary

I Putu Susila, Indonesia

Date: June 2024.

# Goal of the document

This document describes the standardized data format for radiation detection measurements used in the CRP-J02017. The standardized format can be used to share data between participants, develop interoperable code, or integrate systems developed by different participants.

Version history

|  |  |  |  |
| --- | --- | --- | --- |
| Version | Date | Contributor | Comment |
| V1 | March 2024 | Tamás Holczer | Initial version |
| v2 | June 2024 | Tamás Holczer  Roland Nádor  I Putu Susila | New optional fields added, some fields renamed with more detailed description |
| v2.1 | June 2024 | Expert mission participants | Optional label added |
| v2.2 | October 2024 | Expert mission participants | Optional coefficients and comments added |

# Requirements for the format

The proposed data format should meet the following requirements:

1. should be flexible format
2. should be in human-readable format
3. should be human writable
4. should be supported in different programming languages
5. should be able to hold one or many measurements
6. should be able to hold background radiation measurements
7. should be able to hold spectrum measurements
8. the time and measurement device should be recognizable from the format

# Data format

The data format is inspired by the format used by Safecast (https://safecast.org/) but is slightly modified to our needs. The data is stored in JSON objects, which meet the first four requirements.

One measurement is stored in the following format (some fields are optional):

**measurement = {**

**"when\_captured": ts,**

**"device": id,**

**"loc\_lat": llat,**

**"loc\_lon": llon,**

**"type": t,**

**"reading": value,**

**"detector\_type": dtype,**

**"energy\_bin": ebin,**

**"coefficients": [coeff1, coeff2, …, coeffN],**

**"live-time": ltime,**

**"real\_time": rtime,**

**"waypoints": [{"wtime": wtime1, "loc\_lat": llat1, "loc\_lon": llon1}, …, {"wtime": wtimeN, "loc\_lat": llatN, "loc\_lon": llonN}],**

**"annotations": [{"atime": atime1, "avalue": avalue1}, …, {"atime": atimeN, "avalue": avalueN}],**

**"labels": [label1, label2, …, labelN],**

**"modifications": [ {tag1, text1}, {tag2, text2}, …, {tagN, textN}],"unit": unit**

**}**

where,

* when\_captured: device timestamp when the measurement was taken in ISO8601 time format
* device: identifier of the device
* loc\_lat: device latitude in decimal degree format e.g. **47.497913**, 19.040236 for Budapest (optional)
* loc\_lon: Device longitude in decimal degree format e.g. 47.497913, **19.040236** for Budapest (optional)
* type: type of measurement, currently supported types with default unit of measurement
  + rad\_dr radiation dose-rate measured in [μSv/h] by default
  + rad\_cr: radiation count-rate measured in [cps] by default
  + rad\_nc: neutron count measured in [cps] by default
  + spectrum: spectrum measurement measured in [list of keV-counts pairs] by default
  + env\_temp: external temperature measured in [°C] by default
  + env\_humid: external humidity measured in [RH%] by default
  + env\_press: external air pressure measured in [hPa] by default
  + env\_ws: wind speed measured in [m/s] by default
  + env\_wd: the direction from which the wind originates measured in [degree] by default
  + env\_rain: external rainfall measured in [mm] by default
* reading: the measured value
  + floating point number for scalar measurements
  + list of count measurements in case of spectrum type, [count1, … , countN]
* detector\_type: type of detector, e.g. HPGe (optional)
* energy\_bin: width of sampling bin, floating value measured in [keV] (optional, only meaningful with spectrum type)
* coefficients: defines how to interpret the channel data in case of spectrum type (optional)
* live\_time: time duration in milliseconds, the time duration during which the device is actually measuring data (not including processing times). This value is less-than or equal-to real\_time (optional)
* real\_time: time duration in milliseconds, the total time required to collect the measurement (optional)
* waypoints: list of timestamps, latitude and longitude that has been traveled during the measurement (optional, only meaningful with spectrum type)
* annotations: list of annotations consisting of time comment pairs (optional)
* labels: description of the circumstances of that particular measurement (e.g. source was moved, rained etc.) (optional)
* modifications: list of modifications done to the measurement, the tag is mandatory, the text is optional; possible tags: modified, injected, shifted, algorithm etc (optional)
* unit: unit of measurement, overrides the default unit defined in type if present (optional)

The uploaded dataset can contain one or more measurements along with a version number:

**dataset = {**

**"version": v,**

**"labels": [label1, label2, …, labelN],**

**"payload": [measurement1, measurement2, …, measurementN],**

**"comment": text**

**}**

Some example measurements and a schema validator can be found in the appendix. The version is fixed to 2 in this release.

The labels are optional free text strings. Some recommended labels are ‘synthetic’, ‘attacked’, ‘injected’, ‘NCBJ’ etc. The label can refer to the way the dataset was produced, the origin of the dataset, or any other important circumstance that may influence the usage of the dataset. Longer descriptions could go to the optional comment field.

## Format validation

JSON objects can be validated against a given JSON schema. The validation ensures that the tested object is synthetically correct and meets the expectations of other developers. The required schema is implemented and tested (see Appendix B for details). Validators can be found for various programming languages. A long list can be found here:

https://json-schema.org/implementations

Online validation is also possible. The recommended tool is:

https://www.jsonschemavalidator.net/

# Development possibilities

The data format's current form is incompatible with any other data sources. It would be useful to be compatible with M-INSN in the future. Unfortunately, the data format used by M-INSN is not available to us when writing this proposal. If our request for that format is accepted in the future, this format should be updated.

A new proposal version should be developed if we need authenticated readings. This can be based on the idea of RFC 5848 (Signed Syslog Messages), while encrypted messages should be similar to RFC 5425 (Transport Layer Security (TLS) Transport Mapping for Syslog).

# Appendix A

Some examples of JSON objects are given to help understand the format. They are NOT real measurements, just made-up values.

**Dataset** 1: single minimum radiation measurement

{

    "version": 2,

    "payload": [

        {

            "type": "rad\_dr",

            "device": "27",

            "when\_captured": "2024-01-30T23:12:54Z",

            "reading": 0.04012

        }

    ]

}

**Dataset** 2: single full radiation measurement

{

    "version": 2,

    "payload": [

        {

            "type": "rad\_dr",

            "device": "27",

            "when\_captured": "2024-01-30T23:12:54Z",

            "reading": 40.12,

            "loc\_lat": 47.47900,

            "loc\_lon": 19.05832,

            "unit": "nSv/h"

        }

    ]

}

**Dataset** 3: list of single type measurements

{ {

    "version": 2,

    "payload": [

        {

            "type": "rad\_dr",

            "device": "27",

            "when\_captured": "2024-01-30T23:12:54Z",

            "reading": 40.12,

            "loc\_lat": 47.47900,

            "loc\_lon": 19.05832,

            "unit": "nSv/h"

        },

        {

            "type": "rad\_dr",

            "device": "27",

            "when\_captured": "2024-01-30T23:12:58Z",

            "reading": 50.12,

            "loc\_lat": 47.47900,

            "loc\_lon": 19.05932,

            "unit": "nSv/h"

        },

        {

            "type": "rad\_dr",

            "device": "27",

            "when\_captured": "2024-01-30T23:13:00Z",

            "reading": 60.12,

            "loc\_lat": 47.47900,

            "loc\_lon": 19.06000,

            "unit": "nSv/h"

        }

    ]

}

**Dataset** 4: list of multiple type measurements

{

    "version": 2,

    "payload": [

        {

            "type": "rad\_dr",

            "device": "27",

            "when\_captured": "2024-01-30T23:12:54Z",

            "reading": 40.12,

            "loc\_lat": 47.47900,

            "loc\_lon": 19.05832,

            "unit": "nSv/h"

        },

        {

            "type": "env\_temp",

            "device": "27",

            "when\_captured": "2024-01-30T23:12:54Z",

            "reading": 30.5,

            "loc\_lat": 47.47900,

            "loc\_lon": 19.05932,

            "unit": "°C"

        },

        {

            "type": "env\_ws",

            "device": "27",

            "when\_captured": "2024-01-30T23:12:54Z",

            "reading": 5.5,

            "loc\_lat": 47.47900,

            "loc\_lon": 19.06000,

            "unit": "m/s"

        },

        {

            "type": "env\_wd",

            "device": "27",

            "when\_captured": "2024-01-30T23:12:54Z",

            "reading": 90,

            "loc\_lat": 47.47900,

            "loc\_lon": 19.06000,

            "unit": "degree"

        },

        {

            "type": "env\_rain",

            "device": "27",

            "when\_captured": "2024-01-30T23:12:54Z",

            "reading": 1000,

            "loc\_lat": 47.47900,

            "loc\_lon": 19.06000,

            "unit": "mm"

        }

    ]

}

**Dataset** 5: minimal spectrum measurement

{

    "version": 2,

    "payload": [

        {

            "type": "spectrum",

            "device": "32",

            "when\_captured": "2024-01-30T23:12:54Z",

            "live\_time": 30000,

            "real\_time": 40000,

            "energy\_bin": 2,

            "reading": [0, 10, 20, 30, 40, 35, 20, 30]

        }

    ]

}

**Dataset** 6: full spectrum measurement

{

    "version": 2,

    "payload": [

        {

            "type": "spectrum",

            "device": "32",

            "when\_captured": "2024-01-30T23:12:54Z",

            "live\_time": 30000,

            "real\_time": 40000,

            "energy\_bin": 2,

            "reading": [0, 10, 20, 30, 40, 35, 20, 30],

            "spectrum\_type": "HPGe",

            "waypoints": [

                {"ts": "2024-01-30T23:12:54Z", "lat": 12.300, "lon": 0.123},

                {"ts": "2024-01-30T23:12:56Z", "lat": 12.300, "lon": 0.133}

            ]

        }

    ]

}

**Dataset** 7: spectrum measurement with environmental measurements

{

    "version": 2,

    "payload": [

        {

            "type": "spectrum",

            "device": "32",

            "when\_captured": "2024-01-30T23:12:54Z",

            "live\_time": 30000,

            "real\_time": 40000,

            "energy\_bin": 2,

            "reading": [0, 10, 20, 30, 40, 35, 20, 30],

            "spectrum\_type": "HPGe",

            "waypoints": [

                {"ts": "2024-01-30T23:12:54Z", "lat": 12.300, "lon": 0.123},

                {"ts": "2024-01-30T23:12:56Z", "lat": 12.300, "lon": 0.133}

            ]

        },

        {

            "type": "env\_temp",

            "device": "27",

            "when\_captured": "2024-01-30T23:12:54Z",

            "reading": 30.5,

            "loc\_lat": 47.47900,

            "loc\_lon": 19.05932,

            "unit": "°C"

        },

        {

            "type": "env\_ws",

            "device": "27",

            "when\_captured": "2024-01-30T23:12:54Z",

            "reading": 5.5,

            "loc\_lat": 47.47900,

            "loc\_lon": 19.06000,

            "unit": "m/s"

        },

        {

            "type": "env\_wd",

            "device": "27",

            "when\_captured": "2024-01-30T23:12:54Z",

            "reading": 90,

            "loc\_lat": 47.47900,

            "loc\_lon": 19.06000,

            "unit": "degree"

        },

        {

            "type": "env\_rain",

            "device": "27",

            "when\_captured": "2024-01-30T23:12:54Z",

            "reading": 1000,

            "loc\_lat": 47.47900,

            "loc\_lon": 19.06000,

            "unit": "mm"

        }

    ]

}

# Appendix B

The datasets can be validated against the following schema.

{

    "$schema": "https://json-schema.org/draft/2019-09/schema",

    "$id": "https://iaea.org/rds.data.schema.json",

    "title": "DRS data format",

    "description": "This document describes the data format used by IAEA CRP J02017. The current version is proposed by the project members. Corresponding author: Tamas Holczer.",

    "type": "object",

    "properties": {

        "version": {

            "description": "Version of the data format",

            "type": "number"

        },

        "labels": {

            "description": "List of optional labels",

            "type": "array",

            "items": {

                "description": "Label of the dataset refering to the origin or other attributes of the dataset",

                "type": "string"

            }

        },

        "payload": {

            "description": "List of measurements",

            "type": "array",

            "items": {

                "type": "object",

                "properties": {

                    "when\_captured": {

                        "description": "Timestamp",

                        "type": "string",

                        "format": "date-time"

                    },

                    "device": {

                        "description": "Identifier of the device",

                        "type": "string"

                    },

                    "loc\_lat": {

                        "description": "Lattitude",

                        "type": "number",

                        "minimum": -90,

                        "maximum": 90

                    },

                    "loc\_lon": {

                        "description": "Longitude",

                        "type": "number",

                        "minimum": -180,

                        "maximum": 180

                    },

                    "type": {

                        "description": "Type of the measured value",

                        "type": "string",

                        "enum": ["rad\_dr", "rad\_cr", "rad\_nc", "spectrum", "env\_temp", "env\_humid", "env\_press", "env\_ws", "env\_wd", "env\_rain"]

                    },

                    "reading": {

                        "anyOf": [

                            {

                                "description": "Actual measurement",

                                "type": "number"

                            },

                            {

                                "description": "Actual measurement",

                                "type": "array",

                                "items": {

                                    "description": "Counts detected on the given channel",

                                    "type": "number"

                                },

                                "minItems": 1

                            }

                        ]

                    },

                    "detector\_type": {

                        "description": "Detector type",

                        "type": "string"

                    },

                    "energy\_bin": {

                        "description": "Width of sampling bin for spectrum measured in keV",

                        "type": "number"

                    },

                    "coefficients": {

                        "description": "List of coefficients used in the interpretation of channel data",

                        "type": "array",

                        "items": {

                            "description": "Coefficient",

                            "type": "number"

                        },

                        "minItems": 1

                    },

                    "live\_time": {

                        "description": "Time duration during which the device is actually measuring data measured in milliseconds",

                        "type": "number"

                    },

                    "real\_time": {

                        "description": "Total time required to collect the measurement measured in milliseconds",

                        "type": "number"

                    },

                    "waypoints": {

                        "description": "List of waypoints",

                        "type": "array",

                        "items": {

                            "description": "Waypoint",

                            "type": "object",

                            "properties": {

                                "ts": {

                                    "description": "Time",

                                    "type": "string",

                                    "format": "date-time"

                                },

                                "lat": {

                                    "description": "Lattitude",

                                    "type": "number",

                                    "minimum": -90,

                                    "maximum": 90

                                },

                                "lon": {

                                    "description": "Longitude",

                                    "type": "number",

                                    "minimum": -180,

                                    "maximum": 180

                                }

                            },

                            "required": [

                                "ts",

                                "lat",

                                "lon"

                            ]

                        },

                        "minItems": 1

                    },

                    "annotations": {

                        "description": "List of annotations",

                        "type": "array",

                        "items": {

                            "description": "Annotation",

                            "type": "object",

                            "properties": {

                                "time": {

                                    "description": "Time",

                                    "type": "string",

                                    "format": "date-time"

                                },

                                "comment": {

                                    "description": "Comment",

                                    "type": "string"

                                }

                            },

                            "required": [

                                "time",

                                "comment"

                            ]

                        },

                        "minItems": 1

                    },

                    "labels": {

                        "description": "List of labels",

                        "type": "array",

                        "items": {

                            "description": "Label",

                            "type": "string"

                        },

                        "minItems": 1

                    },

                    "modifications": {

                        "description": "Modifications done to the measurement",

                        "type": "array",

                        "items": {

                            "description": "Modification",

                            "type": "object",

                            "properties": {

                                "tag": {

                                    "description": "Tag of the modification",

                                    "type": "string"

                                },

                                "text": {

                                    "description": "Textual description of the modification",

                                    "type": "string"

                                }

                            },

                            "required": [

                                "tag"

                            ]

                        },

                        "minItems": 1

                    },

                    "unit": {

                        "description": "Unit of measurement, overrides the default unit defined in type",

                        "type": "string"

                    }

                },

                "required": [

                    "when\_captured",

                    "device",

                    "type",

                    "reading"

                ]

            }

        },

        "comment": {

            "description": "Long textual description of the dataset",

            "type": "string"

        }

    },

    "required": [

        "version",

        "payload"

    ]

}