Data format proposal for CRP-J02017

Author: Tamás Holczer, PhD, Hungary

Date: March 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.

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:

measurement = {"when_captured": ts, "device": id, "loc_lat": llat, "loc_lon": llon, "type": t, "reading": value}

where,

- when_captured: device timestamp when the measurement was taken in ISO8601 time format
- device: identifier of the device
- loc_lat: device latitude
- loc_lon: Device longitude
- type: type of measurement, currently supported types
  - bg_rad: background radiation measured in [μSv/h]
  - bg_cnt: background_count measured in [cps]
  - neutron: neutron count measured in [cps]
  - spectrum: spectrum measurement measured in [list of keV-counts pairs]
  - env_temp: external temperature measured in [°C]
  - env_humid: external humidity measured in [RH%]
  - env_press: external air pressure measured in [hPa]
- reading: the measured value
  - floating point number for scalar measurements

o list of count measurements along with spectrum type {"spectrum_type": stype, "count": [count1, count2, … , countN]] for spectrum measurements

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

dataset = {"version": v, "payload": [measurement1, measurement2, …, measurement N]}

Some example measurements and a schema validator can be found in the appendix. The version is fixed to 1 in this release.

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 background radiation measurement near a research reactor in Budapest

```json
{
    "version": 1,
    "payload": [
        {
            "when_captured": "2024-01-30T23:12:54Z",
            "device": "27",
            "loc_lat": 47.47900,
            "loc_lon": 19.05832,
            "type": "bg_rad",
            "reading": 0.04012
        }
    ]
}
```

Dataset 2: single temperature measurement near to IAEA headquarters in Vienna

```
{
    "version": 1,
    "payload": [
        {
            "when_captured": "2024-02-05T10:15:31Z",
            "device": "36",
            "loc_lat": 48.23448240688433,
            "loc_lon": 16.41586222177042,
            "type": "env_temp",
            "reading": 11.76
        }
    ]
}
```

Dataset 3: single spectrum measurement near Bruce Powers NPP in Canada (spectrum may contain up to 2000 channels; only the first few channels are displayed here)

```
{
    "version": 1,
    "payload": [
        {
            "when_captured": "2024-02-29T21:30:01Z",
            "device": "Ext_ident_876",
            "loc_lat": 44.31435569033905,
            "loc_lon": -81.54987472120817,
            "type": "spectrum",
            "reading": {
                "spectrum_type": "1keV",
                "count": [0, 0, 0, 0, 0, 0, 2, 6, 15, 13, 9, 18, 7, 15, 11,
18, 15, 21, 30]
            }
        }
    ]
}
```

Dataset 4: pressure readings from Jakarta

```
{
    "version": 1,
    "payload": [
        {
            "when_captured": "2024-02-08T10:15:00Z",
            "device": "Pr-1",
            "loc_lat": -6.2001828704783675,
            "loc_lon": 106.81903182570639,
```

```
            "type": "env_press",
            "reading": 1008.1
        },
        {
            "when_captured": "2024-02-08T10:20:00Z",
            "device": "Pr-1",
            "loc_lat": -6.2001828704783675,
            "loc_lon": 106.81903182570639,
            "type": "env_press",
            "reading": 1008.2
        },
        {
            "when_captured": "2024-02-08T10:25:00Z",
            "device": "Pr-1",
            "loc_lat": -6.2001828704783675,
            "loc_lon": 106.81903182570639,
            "type": "env_press",
            "reading": 1008.3
        },
        {
            "when_captured": "2024-02-08T10:30:00Z",
            "device": "Pr-1",
            "loc_lat": -6.2001828704783675,
            "loc_lon": 106.81903182570639,
            "type": "env_press",
            "reading": 1008.2
        }
    ]
}
```

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 Tamas Holczer.",
    "type": "object",
    "properties": {
        "version": {
            "description": "Version of the data format",
            "type": "number"
        },
        "payload": {
            "description": "List of measurements",
            "type": "array",
```

```json
            "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": ["bg_rad", "bg_cnt", "neutron", "spectrum",
"env_temp", "env_humid", "env_press"]
                    },
                    "reading": {
                        "anyOf": [
                            {
                                "description": "Actual measurement",
                                "type": "number"
                            },
                            {
                                "description": "Actual measurement",
                                "type": "object",
                                "properties": {
                                    "spectrum_type": {
                                        "description": "Type of spectrum",
                                        "type": "string"
                                    },
                                    "count": {
                                        "description": "List of measurements",
                                        "type": "array",
                                        "items": {
```

```
                                        "description": "Counts detected on
the given channel",
                                        "type": "number"
                            },
                            "minItems": 1
                    }
                },
                "required": [
                    "spectrum_type",
                    "count"
                ]
            }
        ]
    }
},
"required": [
    "when_captured",
    "device",
    "loc_lat",
    "loc_lon",
    "type",
    "reading"
]
        }
    }
},
"required": [
    "version",
    "payload"
]
}
```