

The Material Assay Data Format

V 2.01

James C. Loach
Shanghai Jiaotong University

Revised 6 February 2013

1 Introduction

This document specifies a data format for encoding measurements (assays) of material radiopurities. The data format consists of a JSON schema, restrictions on the content of certain fields and rules that allow the schema to be extended.

2 The data format

2.1 Overview

The data structure is divided into three top level sub-structures known as `sample`, `measurement` and `data_source`. The first captures information on physical characteristics and origin the thing being counted. The second deals with the actual measurement and its results. The third concerns the origin and input of the data contained in the document. This is the structure in outline:

```
{
  "type":          "measurement",
  "grouping":      "Experiment name or similar",
  "sample":        {...},
  "measurement":  {...},
  "data_source":   {...},
  "specification": "X.XX"
}
```

where the { . . . } indicate the three JSON sub-structures that will be described in detail below. Other fields are:

grouping The group to which the document belongs. Usually this will be the name of the experiment that made the measurement.

Conventions: less than 100 characters; no period; required.

Example: *Majorana Demonstrator*

specification The version number of the Material Assay Data Format specification used to encode the document.

Conventions: [MAJOR VERSION] . [MINOR VERSION]; required.

Example: *2.01*

2.2 Sample

The sample sub-structure has this form:

```
"sample": {  
  
  "name":          "Short description",  
  "description":  "Detailed description",  
  "id":           "Identification number",  
  "source":       "Where it came from",  
  
  "owner":       {  
    {  
      "name":      "Who owns it",  
      "contact":   "Institution or email/postal address"  
    }  
  }  
}
```

name A concise description of the sample, typically indicating the kind of material and its origin. It can be considered the heading of the document.

Conventions: less than 100 characters; single line; no period; required.

Example: *Corning 7980 ArF fused silica*

Example: *Homestake rhyolite, 4850 ft level, sample A3*

description A detailed description of the sample, typically including information on its form, appearance and processing.

Conventions: written as a sentence or comma-separated list of items; single line; no limit on length and level of detail; optional period; can be left blank.

Example: *Polished wafer of Corning 7980 ArF high purity fused silica, disc shape, 1 mm thick, 20 mm diameter, lot number 12345.*

id An identification number associated with the sample, typically used by a counting facility, or other organization, to track the sample.

Conventions: an alphanumeric identification string; can be left blank.

Example: *3128a*

Example: *2010-AF-0000004567-b*

source Information on where the sample came from, typically consisting of a company name and purchase date.

Conventions: can be left blank.

Example: *Mark Optics Incorporated, purchased April 30th 2011.*

owner Information on who owns the sample. It may refer to ownership at the time of the measurement or at some later date.

Conventions: can be left blank.

Example: *John Smith // jsmith@gnl.gov*

Example: *John Smith // Government National Laboratory*

2.3 Measurement

The measurement sub-structure has this form:

```
"measurement": {  
  
  "institution":      "Where the count was done",  
  "technique":       "The technique that was used",  
  
  "date":            [],  
  
  "requestor":      {  
    {  
      "name":          "Who managed the measurement",  
      "contact":       "Institution or email/postal address"  
    },  
  
  "practitioner":   {  
    {  
      "name":          "Who did the measurement",  
      "contact":       "Institution or email/postal address"  
    },  
  
  "description":    "Detailed multi-line description of  
                    the procedure and results",  
  
  "results":        [  
    [  
      {  
        "isotope":     "II-AAA or II or description",  
        "type":        "measurement or limit or range",  
        "value":       [],  
        "unit":        "Unit"  
      },  
      ...  
    ],  
  ]  
}
```

institution The name of the institution where the measurement was made.

Conventions: can be left blank.

Example: *Oak Ridge National Laboratory*

Example: *LBNL LBF*

technique The technique used to make the measurement.

Conventions: can be left blank.

Example: *ICP-MS*

Example: *Gamma*

date The date on which the measurement took place. This array can contain either a single value indicating the date on which the experiment took place (or began), or a two values indicating the period over which the measurement was made.

Conventions: entered in the format YYYY-MM-DD or YYYY-MM or YYYY; can be left empty.

Example: *1980-04*

Example: *1980-04-26*

requestor Information on who commissioned, funded, managed or was otherwise responsible for the measurement taking place.

Conventions: can be left blank.

Example: *John Smith // jsmith@gnl.gov*

practitioner Who was responsible for actually making the measurement.

Conventions: can be left blank.

Example: *John Smith // jsmith@gnl.gov*

description A detailed description of the measurement. Typically this is a substantial piece of text duplicated from the practitioner's report. It may repeat data broken out into other fields.

Conventions: multi-lined block of text; can be left blank.

Example:

Of the 50 packages of HYSOL Ag-epoxy you brought to the LBNL Low Background Facility, 25 packages have been analysed at the Berkeley Facility as a single sample of unopened packages, consisting of all items in Lot 1 and Lot 2 (your designation).

The total weight of all 25 unopened packages is 143 grams, while the weight of the unmixed Ag-epoxy (at 2.65 grams per package) is 66 grams. Results are listed below, where all values are based on the 143 gram total weight of the sample.

Detector: MERLIN(BKY) U: <100 ppb

Sample: S6MB annulus Th: <230 ppb

Weight: 43 grams K: 84(8) ppm

Data file: 5695S

Count time: 48100 sec

results A list of measurements for each isotope included in the measurement. The four measurement types are `measurement`, `range` and `limit`. The meaning of the contents of the `value` array varies according to the measurement type and the size of the array:

type	value	Description
measurement	[0]	measurement with no error
	[0,0]	measurement with symmetric error
	[0,0,0]	measurement with asymmetric error (+, -)
range	[0,0]	range (lower, upper)
	[0,0,0]	range with confidence level
limit	[0]	limit
	[0,0]	limit with confidence level

General conventions: the isotope name should be entered as the two letter chemical symbol with or without the mass number (in *exceptional circumstances* a more detailed description, such as 'U early' may be given); permissible units are pct, ppm, ppb, ppt, ppq, mBq/kg, uBq/kg, nBq/kg, n/a; the confidence level (c.l.) on a limit or range is a number $0 < \text{c.l.} < 100$; if the confidence level on a limit or range is omitted then it will be interpreted as a 95% limit; confidence levels on measurements are assumed to be 68%; documents with zero measurements *are* permitted.

Example: `U-238 // "measurement" // [120,10] // "ppb"`

Example: `Th-232 // "range" // [40,50,95] // "uBq/kg"`

Example: `Th-232 // "limit" // [40] // "uBq/kg"`

2.4 Data source

The data source sub-structure has this form:

```
"data_source": {  
  "reference":      "Where the data came from",  
  "input":        {  
    "name":         "Who created this document",  
    "contact":      "Institution or email/postal address",  
    "date":         []  
  },  
  "notes":        "Comments on/issues with data entry"  
}
```

reference Where the information in the document was taken from.

Conventions: required.

Example: *D. Leonard et al., Nucl. Instr. Meth. A 591, 490 (2008).*

Example: *CERNEXPERIMENT internal report TECH-DOC-0004*

input Information on who entered the data or who was responsible for the data entry.

Conventions: entered in the format YYYY-MM-DD; required.

Example: *Alice Undergraduate // aundergrad@uni.edu*

notes Comments on issues that arose during the input of the data, typically focussing on the interpretation of the source data.

Conventions: multiple line block of text; can be left blank.

Example: *In the results table the paper refers to 'U conc.' and 'Th conc.' as the things being measured. But from the text it's clear that what was actually measured were concentrations of 'U-238' and 'Th-232' and so this is what I entered in the results section.*

2.5 Extendability

2.5.1 The user array

The sub-structures `measurement`, `sample` and `data_source` may each contain additional structured information in form of an optional user array. This ability to extend the core specification should only be used in *exceptional circumstances*. The user array takes this form:

```
"user":  
  [  
    {  
      "name":           "Short description",  
      "description":   "Detailed description",  
      "type":          "measurement or limit  
                      or range or string",  
      "value":         [] or "",  
      "unit":          "Associated unit"  
    },  
    ...  
  ]
```

name Display name.

Conventions: required.

Example: *Data file*

description Description.

Conventions: optional.

Example: *Size of the Marinelli beaker*

type The type of data stored in `value`.

Conventions: Following the same convention as the `measurement` sub-structure, this field describes the type of data stored in the `value` field; *string* is added option; required.

Example: *measurement*

Example: *range*

Example: *limit*

Example: *string*

value An array of numbers or a single string.

Conventions: required.

Example: *[304.1]*

Example: *[0.1, 0.2]*

Example: *[21.1, 95]*

Example: *AB444-D.spe*

unit Unit.

Conventions: optional.

Example: *cm*

2.5.2 Example of appropriate use

Below is an example of appropriate use. Here a user - a hypothetical low background counting facility - has decided to include a field to store the name of the measurement data file (containing the spectrum taken with the HPGe detector). The facility requires this number to have special prominence, and to be a separate field by which documents can be searched and sorted. At the same time it is a number of no interest to anyone outside of the facility.

```
"measurement": {
    "institution": "...",
    "technique": "...",
    "date": [""],
    "requestor": { "name": "...", "contact": "..." },
    "practitioner": { "name": "...", "contact": "..." },
    "description": "...",
    "results": [
        {
            "isotope": "...",
            "type": "measurement",
            "value": [0],
            "unit": "Unit"
        }
    ],
    "user": [
        {
            "name": "Data file",
            "type": "string",
            "value": "2010-A-0092"
        }
    ]
}
```

2.6 Changes

2013-01-14 (to v1.01) Correct insignificant typos.

2013-01-14 (to v1.01) Add comment that uncertainties on measurements are expected to be 68% c.l. and correct default confidence level on limits from 90% to 95%.

2013-01-30 (to v2.00) Change the way optional fields are handled, remove m_ and u_ prefixes.

2013-01-30 (to v2.00) Convert data to a one- or two-valued array.

2013-01-30 (to v2.00) Revise the way the results are handled. Now use a unified structure with a type. Combine values, error and limits into a variable-length value array.

2013-02-06 (to v2.01) Results type meas becomes measurement.

2013-02-06 (to v2.01) Rename opt array to user and revise it's structure.