

# The Material Assay Data Format

V 1.0

James C. Loach  
Shanghai Jiaotong University

11 December 2012

## 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.X"  
}
```

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: *1.0*

## 2.2 Sample

The sample sub-structure has this form:

```
"sample": {  
  
  "m_name":          "Short description",  
  "m_description":   "Detailed description",  
  "m_id":            "Identification number",  
  "m_source":        "Where it came from",  
  "m_owner":        {  
    "name":          "Who owns it",  
    "contact":       "Institution or email/postal address"  
  }  
}
```

**m\_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*

**m\_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; finishes with 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.*

**m\_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*

**m\_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.*

**m\_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:

```
"m_measurement": {  
  
    "m_institution":    "Where the count was done",  
    "m_technique":      "The technique that was used",  
  
    "m_date":           "YYYY-MM-DD",  
    "m_date":           ["YYYY-MM-DD", "YYYY-MM-DD"],  
  
    "m_requestor":      {  
        {  
            "name":      "Who managed the measurement",  
            "contact":    "Institution or email/postal address"  
        },  
  
    "m_practitioner":   {  
        {  
            "name":      "Who did the measurement",  
            "contact":    "Institution or email/postal address"  
        },  
  
    "m_description":    "Detailed multi-line  
                        description of the  
                        procedure and results",  
  
    "m_results": [      {  
        {  
            "isotope":    "II-AAA or II",  
            "value":      0,  
            "error":      0,  
            "unit":       "Unit"  
        },  
        {  
            "isotope":    "II-AAA or II",  
            "limit":      0,  
            "cl":         0,  
            "unit":       "Unit"  
        },  
        ...  
    ]  
}
```

**m\_institution** The name of the institution where the measurement was made.

Conventions: can be left blank.

Example: *Oak Ridge National Laboratory*

Example: *LBNL LBF*

**m\_technique** The technique used to make the measurement.

Conventions: can be left blank.

Example: *ICP-MS*

Example: *Gamma*

**m\_date** The date on which the measurement took place or began. This is either a single value indicating the date on which the experiment took place or began, or a two-value array indicating the period over which the measurement was made.

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

**m\_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*

**m\_practitioner** Who was responsible for actually making the measurement.

Conventions: can be left blank.

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

**m\_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

**m\_results** A list measurements (definite measurements or limits) for each isotope included in the measurement.

Conventions: the isotope name should be entered as the two letter chemical symbol followed by a hyphen and the mass number; in *exceptional circumstances* it is permissible to omit the hyphen and mass number; permissible units are pct, ppm, ppb, ppt, ppq, mBq/kg, uBq/kg, nBq/kg; the confidence level (c.l.) is an integer  $0 < \text{c.l.} < 100$ ; if the confidence level is omitted (i.e. set to 0) then it will be interpreted as a 90% limit; documents with zero measurements *are* permitted.

Example (meas.): *U-238 // 120 // 10 // ppb*

Example (limit): *Th-232 // 40 // 90 // uBq/kg*

## 2.4 Data source

The data source sub-structure has this form:

```
"m_data_source": {  
  "m_reference":      "Where the data came from",  
  "m_input":         {  
    "name":           "Who created this document",  
    "contact":        "Institution or email/postal address",  
    "date":           "YYYY-MM-DD"  
  },  
  "m_notes":         "Comments on/issues with data entry"  
}
```

**m\_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*

**m\_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*

**m\_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

The sub-structures *measurement*, *sample* and *data source* may each contain extra key-value pairs in addition to those defined in the previous sections. Field names should begin with a prefix 'u\_' (to distinguish them from the mandatory fields with prefix 'm\_'), should be lowercase and contain no spaces or special characters (other than the underscore). Field values should be single-line text strings. The ability to extend the specification should only be used in *exceptional circumstances*.

Consider this example of appropriate use:

```
"m_measurement": {  
  
    "m_institution":    "...",  
    "m_technique":      "...",  
    "m_date":           "...",  
    "m_date":           ["...", "..."],  
  
    "m_requestor":      { "name": "...", "contact": "..." },  
    "m_practitioner":   { "name": "...", "contact": "..." },  
  
    "m_description":    "...",  
  
    "m_results": [        
        { "isotope": "...", "value": 0, "error": 0, "unit": "..." }  
    ],  
  
    "u_datafile":       "2010-A-0092"  
  
}
```

Here a user - a hypothetical low background counting facility - has decided to include a field to store the name of the measurement data file (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.