Nuclear Instruments and Methods in Physics Research A 839 (2016) 6–11

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

A database for storing the results of material radiopurity measurements

J.C. Loach [a,b,*], J. Cooley [c], G.A. Cox [d], Z. Li [a], K.D. Nguyen [b], A.W.P. Poon [b]

[a] Department of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China
[b] Nuclear Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA94720, USA
[c] Department of Physics, Southern Methodist University, Dallas, TX75275, USA
[d] Institute for Nuclear Physics, Karlsruhe Institute of Technology, Karlsruhe 76131, Germany

ARTICLE INFO

Article history:
Received 22 April 2016
Received in revised form
1 September 2016
Accepted 16 September 2016
Available online 17 September 2016

Keywords:
Database
Radiation
Neutrinoless double-beta decay
WIMP dark matter
Radiopurity
Materials

ABSTRACT

Searches for rare nuclear processes, such as neutrinoless double beta-decay and the interactions of WIMP dark matter, are motivating experiments with ever-decreasing levels of radioactive backgrounds. These background reductions are achieved using various techniques, but amongst the most important is minimizing radioactive contamination in the materials from which the experiment is constructed. To this end there have been decades of advances in material sourcing, manufacture and certification, during which researchers have accumulated many thousands of measurements of material radiopurity. Some of these assays are described in publications, others are in databases, but many are still communicated informally. Until this work, there has been no standard format for encoding assay results and no effective, central location for storing them. The aim of this work is to address these long-standing problems by creating a concise and flexible material assay data format and powerful software application to manipulate it. A public installation of this software, available at http://www.radiopurity.org, is the largest database of assay results ever compiled and is intended as a long-term repository for the community's data.

1. Introduction

Searches for rare nuclear processes have led to many of the most important results in nuclear and particle physics over recent decades [1–4] and offer the prospect of many more in years to come [5–7]. Crucial to these experiments, and especially to those that search for rare weak interactions, such as neutrinoless double beta-decay and the interactions of WIMP dark matter, is the suppression of events due to radioactive backgrounds. This can be achieved through experimental design or analysis, but the most basic strategy is to directly address the source terms by shielding the experiment from environmental radioactivity and, of particular relevance here, constructing the experiment from low-radioactivity materials.

Selecting candidate low-radioactivity materials is as much an art as it is a science, relying on inference from previous measurements and on the instinct of the experimentalist. It is also a critical task because the testing and certification of candidate materials, though a well-established process, involves precision measurements that consume significant amounts of time, cost and effort. Candidates must be selected judiciously, and how well this can be done depends upon the quality of information available to the researcher. Of particular importance is their access to previous measurements of similar materials.

A central repository of material radiopurity measurements is strongly motivated by these considerations, and the authors are not the first to propose one or, even, to build one. Previous efforts such as the public material assay database created by the ILIAS collaboration [8] have been invaluable to the low-background physics community. But despite their usefulness, none have fully met the community's needs. There have been issues with the limited scope of the data sets and difficulties in augmenting them, as well as with non-portability of data and difficulties in querying it. Solving these issues was the motivation for this work, and we describe ways of storing and handling material assays that are intended to address them.

Our work comprises two parts: a data format in which assays can be encoded and a piece of software for manipulating the encoded data. In a practical sense these components are intertwined, but there is a formal separation between them that is reflected in the organization of this paper. Our material assay data format is presented in the next section, free of association with any particular database software. Following this we discuss our web application Persephone [9], which is a tool for storing, viewing and manipulating data encoded in the format. And, finally, we describe a specific installation of our software that we intend as the central repository for the community's data.

* Corresponding author at: Department of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China.

http://dx.doi.org/10.1016/j.nima.2016.09.036
0168-9002/

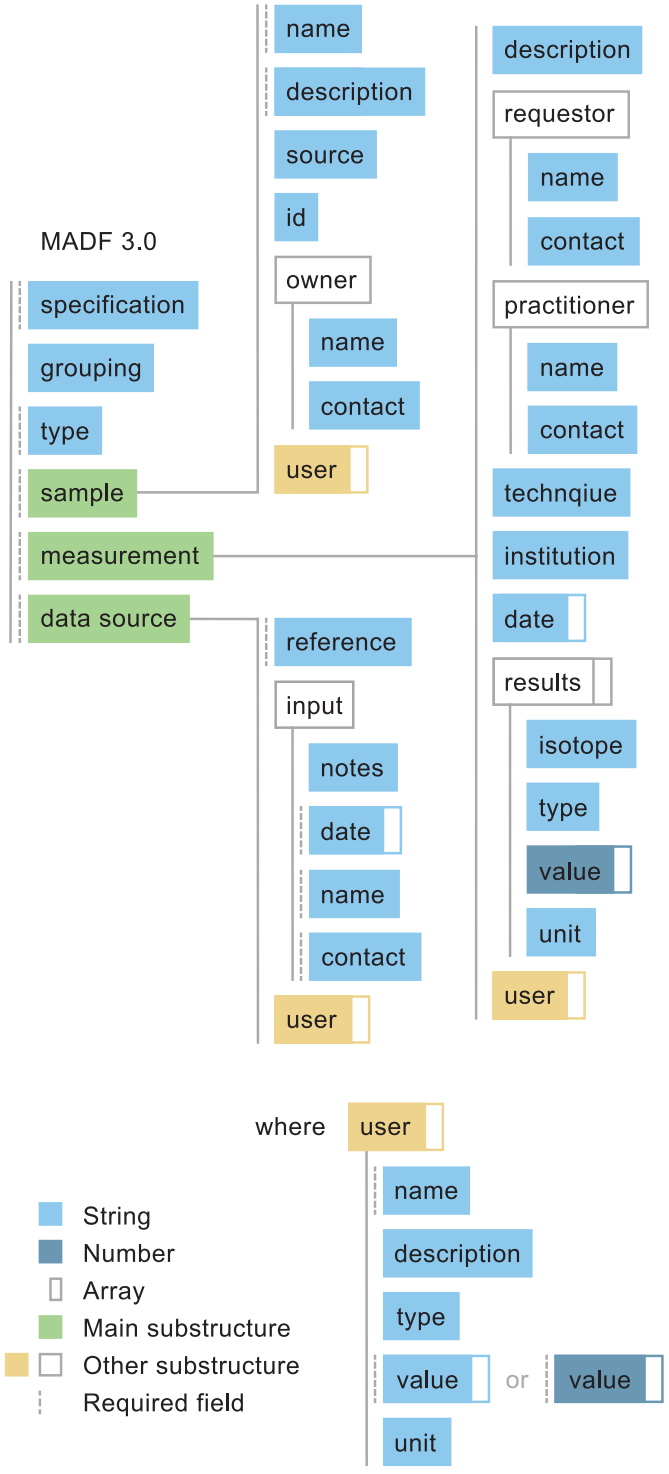

Fig. 1. The structure of the MADF 3.0 data format. Each entry in a user array has the structure given in the lower-right part of the figure. These arrays are used to add extra, use-case-specific fields to the core data structure. The data format is expressed in JSON. Further details are given in Table 1.

2. The Material Assay Data Format

The Material Assay Data Format (MADF) describes a way to encode assays of material radiopurity in JavaScript Object Notation (JSON) [10]. JSON is a language-independent, open standard for encoding structured data. It is especially widespread in Internet communications and is natively read by many advanced and popular computational tools, including those used in physics analyses.

Table 1
The structure of the MADF 3.0 data format. All fields are strings unless stated otherwise. Fields marked with a ∗ are required.

Field	Description
`specification`∗	MADF specification version.
`grouping`	Experiment name or similar.
`type`∗	Fixed value ``assay'' indicating the document type.
`sample`	
`.name`∗	Concise description.
`.description`∗	Detailed description.
`.id`	Identification number.
`.source`	Where the sample came from.
`.owner`	Who owns the sample.
`.name`	Name.
`.contact`	Email address or telephone no.
`.user`	User field array. See below.
`measurement`	
`.description`	Detailed description.
`.requestor`	Who coordinated the measurement.
`.name`	Name.
`.contact`	Email address or telephone no.
`.practitioner`	Who did the measurement.
`.name`	Name.
`.contact`	Email address or telephone no.
`.technique`	Technique name.
`.institution`	Institution name.
`.date`	Array of strings. See Appendix C.
`.results`	Array of measurements.
`.isotope`	Isotope name, usually in the format `symbol-mass number`.
`.type`	`measurement`, `limit` or `range`.
`.value`	Array of one to three numbers. See Appendix A.
`.unit`	Unit chosen from a restricted set of choices. See Appendix B.
`.user`	User field array. See below.
`data_source`	
`.reference`∗	Where the data came from.
`.input`	Data entry details.
`.notes`	Input simplifications, assumptions.
`.date`∗	Array of strings. See Appendix C.
`.name`∗	Name.
`.contact`∗	Email address or telephone no.
`.user`	User field array. See below.
`.user`	
`.name`∗	Concise description, usually a single word.
`.description`	Detailed description.
`.type`	`measurement`, `limit`, `range` or `string`.
`.value`∗	String or an array of one to three numbers. See Appendix A.
`.unit`	Unit. See Appendix B.

Each MADF JSON document represents a single assay performed on a particular sample of material. It contains three main substructures: `sample` describes the sample of material being assayed; `measurement` describes the assay process and its results, including a list of isotopes with their measured values or limits; and `data_source` gives the origin of the information encoded in the document and describes how it was encoded. Overviews of MADF 3.0 are shown in Fig. 1 and Table 1. Further details are given in the appendices.

The design principles behind the data format are summarized as follows:

Simplicity: Each assay is encoded in a single, semi-structured, human-readable JSON document and there are no structural relationships between documents. In contrast to relational databases, with their arrangements of tables, all information is encapsulated in a single piece of data. This makes the data highly portable, and assays can even be output and stored in self-contained text files.
In MADF many fields can store extended, descriptive information and only a small number of fields are mandatory. In this way it is more similar to a structured