Data types for IMAGE attributes

BioSamples takes sample records with a set of attributes. Each attribute has a name and a value. It can also have 'Units', or a 'Term Source' and a 'Term Source ID'. The Term Source and ID allow us to refer to entries in other databases or ontologies. This is fully described on the BioSamples help pages. The following section describes the expectations for each data type within IMAGE.

number

A number, with units specified. BioSamples recommends that units are given without abbreviations. Terms defined in the UO are encouraged to be used. For example, a birth weight could have a value of 1.3 and the units specified as 'kilogram' .

text

Text, using US English spellings. No check on this type.

limited value

Text type, but the allowed values are limited which either defined in the allowed values or point to other samples. Samples can be referred to in two ways. If the sample you need to reference is in the submission, use the sample name. If the sample was already submitted, use the BioSample ID (e.g. SAMEA2821491).

date

Dates should be reported in an ISO 8601 format, YYYY-MM-DD for dates or YYYY-MM for months. To ensure clarity, the format must be reported as the 'units'.

URI

A URI, such as 'http://www.imageh2020.eu/'. Depending on the context, http, ftp, mailto links may be appropriate. Examples:

- ftp, ftp://ftp.image.ebi.ac.uk/ftp/README
- http, http://www.imageh2020.eu/
- mailto, mailto:bob@example.org

DOI

Digital Object Identifier (DOI) can unambiguously point to one publication, which can be resolved here.

ontology id

A reference to an ontology term. The attribute value should be the term label. The term source should be the ontology used, and the term source ID should be an ID from that ontology. For example, cerebral cortex could be described with an ontology term from 'UBERON' with ontology ID of 'UBERON:0000956' and the attribute value is 'cerebral cortex'. Though in the experiment submission, direct links to ontologies cannot be submitted as attributes. The use of ontology terms is still encouraged by setting the attribute value to exactly match the term name in the ontology.

A special case is taxonomy. A species name and identifier from the NCBI Taxonomy database. For example, human would be described in the term with value of 'Homo sapiens', term source as 'NCBI Taxonomy' and term source ID as 9606.

Valid values/units/terms

Types	valid values	valid units	valid terms
number	B	Y	N
text	N	N	N
limited value	Y	N	N
ontology_id	N	N	Y
uri	B	N	N
doi	B	N	N
date	B	Y	N

In some attributes, the values/units/terms allowed for that attribute is limited, which are defined by the "Valid values", "Valid units" and "Valid terms" correspondingly. The table above describes what types of fields could have what types of limitation. For example, "limited value" type must have corresponding "Valid values" setting, while "text" type could not have any as there is no check on this type. "**B**" in the table stands for both, which means those types could either have or do not have 'Valid values" setting.