

```

{
  "name" : "IMAGE sample metadata rules",
  "further_details_iri" : "https://github.com/bioinformatics-ptp/IMAGE-metadata/blob/master/README.md",
  "description" : "Validation rules for the IMAGE project.",
  "rule_groups" : [
    {
      "name" : "standard",
      "rules" : [
        {
          "Allow Multiple" : "no",
          "Description" : "the id used in the original data source to identify the record",
          "Example" : "SAMEA4450075, ANIMAL:::ID:::132713,SAMPLE:::ID:::19483",
          "Required" : "mandatory",
          "Name" : "Data source ID",
          "Type" : "text"
        },
        {
          "Type" : "limited value",
          "Name" : "Alternative id",
          "Required" : "mandatory",
          "Example" : "Siems_0722_393449",
          "Allow Multiple" : "no",
          "Description" : "the other identifier used in the original data source, e.g. accession, secondary id etc."
        },
        {
          "Description" : "A brief description of the sample",
          "Allow Multiple" : "no",
          "Example" : "Pig sperm sample from 3 year old duroc",
          "Required" : "optional",
          "Type" : "text",
          "Name" : "Description"
        },
        {
          "Example" : "IMAGE",
          "Description" : "All samples are stated to be part of the IMAGE project, while still belonging to each individual gene bank. This tag is required to identify them in public databases",
          "Allow Multiple" : "no",
          "Name" : "Project",
          "Type" : "text",
          "Valid values" : [
            "IMAGE"
          ],
          "Required" : "mandatory"
        }
      ]
    }
  ]
}

```

```

    },
    {
      "Type" : "text",
      "Name" : "Submission title",
      "Required" : "mandatory",
      "Example" : "German National Sperm Bank",
      "Allow Multiple" : "no",
      "Description" : "A descriptive title for the submission"
    },
    {
      "Example" : "A collection of cattle samples from local breeding stocks",
      "Allow Multiple" : "no",
      "Description" : "A brief description of the submission",
      "Type" : "text",
      "Name" : "Submission description",
      "Required" : "recommended"
    },
    {
      "Name" : "Material",
      "Valid terms" : [
        {
          "term" : "OBI_0100026",
          "term_iri" : "http://purl.obolibrary.org/obo/OBI\_0100026",
          "include_root" : 1,
          "allow_descendants" : 0,
          "ontology_name" : "OBI"
        },
        {
          "term" : "OBI_0001479",
          "term_iri" : "http://purl.obolibrary.org/obo/OBI\_0001479",
          "include_root" : 1,
          "allow_descendants" : 0,
          "ontology_name" : "OBI"
        }
      ],
      "Example" : "Specimen from Organism",
      "Description" : "type of the sample, either animal(organism) or sample (specimen from organism)",
      "Type" : "ontology_id",
      "Valid values" : [
        "Organism",
        "Specimen from organism"
      ],
      "Required" : "mandatory",
      "Allow Multiple" : "no"
    },
  ],

```

```

{
  "Required" : "mandatory",
  "Type" : "text",
  "Name" : "Person last name",
  "Description" : "the last name of the submitter or contact for these samples,
multiple persons can be provided",
  "Allow Multiple" : "yes",
  "Example" : "Harrison"
},
{
  "Example" : "W",
  "Description" : "the initial of the submitter or contact for these samples, multiple
persons can be provided",
  "Allow Multiple" : "yes",
  "Type" : "text",
  "Name" : "Person initial",
  "Required" : "optional"
},
{
  "Required" : "recommended",
  "Name" : "Person first name",
  "Type" : "text",
  "Allow Multiple" : "yes",
  "Description" : "the first name of the submitter or contact for these samples,
multiple persons can be provided",
  "Example" : "Peter"
},
{
  "Allow Multiple" : "yes",
  "Description" : "email of contact person",
  "Example" : "mailto:peter@ebi.ac.uk",
  "Required" : "mandatory",
  "Name" : "Person email",
  "Type" : "uri_value"
},
{
  "Required" : "mandatory",
  "Type" : "text",
  "Name" : "Person affiliation",
  "Description" : "affiliation of person",
  "Allow Multiple" : "yes",
  "Example" : "EMBL-EBI"
},
{
  "Valid terms" : [
    {

```

```

    "ontology_name" : "EFO",
    "allow_descendants" : 1,
    "term" : "EFO_0002012",
    "include_root" : 0,
    "term_iri" : "http://purl.obolibrary.org/obo/%22/%22EFO_0002012"
  }
],
  "Allow Multiple" : "yes",
  "Description" : "the role (type of the contribution) made by the person",
  "Example" : "Submitter",
  "Required" : "mandatory",
  "Name" : "Person role",
  "Type" : "text"
},
{
  "Allow Multiple" : "yes",
  "Description" : "Organization name, multiple can be supplied, but at least one must be the name of the gene bank",
  "Example" : "EMBL-EBI",
  "Required" : "mandatory",
  "Type" : "text",
  "Name" : "Organization name"
},
{
  "Required" : "recommended",
  "Type" : "text",
  "Name" : "Organization address",
  "Allow Multiple" : "yes",
  "Description" : "The address of the organisation, multiple can be supplied",
  "Example" : "European Bioinformatics Institute, Wellcome Trust Genome Campus, CB10 1SD, UK"
},
{
  "Example" : "www.ebi.ac.uk",
  "Description" : "the websites of organization, multiple can be supplied",
  "Allow Multiple" : "yes",
  "Type" : "uri_value",
  "Name" : "Organization uri",
  "Required" : "optional"
},
{
  "Required" : "recommended",
  "Valid values" : [
    "the controlled vocabulary for countries is from INSDC http://www.insdc.org/documents/country-qualifier-vocabulary"
  ],

```

```

    "Type" : "text",
    "Name" : "Organization country",
    "Description" : "Country in which organization resides, the controlled vocabulary
    for countries is from INSDC http://www.insdc.org/documents/country-qualifier-
    vocabulary",
    "Allow Multiple" : "yes",
    "Example" : "United Kingdom"
  },
  {
    "Type" : "ontology_id",
    "Name" : "Organization role",
    "Required" : "mandatory",
    "Example" : "Submitter",
    "Allow Multiple" : "yes",
    "Description" : "the role of the organization, e.g. biobank, funder",
    "Valid terms" : [
      {
        "ontology_name" : "\"EFO",
        "allow_descendants" : 1,
        "term" : "\"EFO_0002012",
        "term_iri" : "http://purl.obolibrary.org/obo/%22/%22EFO\_0002012",
        "include_root" : 0
      }
    ]
  },
  {
    "Example" : "doi://10.1139/o09-005",
    "Description" : "Publication associated with sample, supplied as the digital
    object identifier which can unambiguously point to the publication",
    "Allow Multiple" : "no",
    "Type" : "doi",
    "Name" : "Publication DOI",
    "Required" : "optional"
  },
  {
    "Required" : "mandatory",
    "Name" : "Gene bank name",
    "Type" : "text",
    "Allow Multiple" : "no",
    "Description" : "The name of the gene bank that the data of this submission
    belongs to, ideally use the value stored in EUGENA",
    "Example" : "CryoWeb"
  },
  {
    "Allow Multiple" : "no",
    "Description" : "Country in which the gene bank resides, particularly

```

important when different countries use the same database infrastructure, e.g. CryoWeb. The controlled vocabulary for countries is from INSDC <http://www.insdc.org/documents/country-qualifier-vocabulary>,

```

“Example” : “Italy”,
“Required” : “mandatory”,
“Name” : “Gene bank country”,
“Type” : “limited value”
},
{
“Example” : “CryoWeb”,
“Description” : “the type of the data source the current Inject tool can handle”,
“Allow Multiple” : “no”,
“Name” : “Data source type”,
“Type” : “limited value”,
“Required” : “mandatory”,
“Valid values” : [
“CryoWeb”,
“CRB-Anim”,
“Template”
]
},
{
“Example” : “2018-04-27, version 1.5”,
“Allow Multiple” : “no”,
“Description” : “the version of the data source, suggested to be either the dump date or the version number used by the source”,
“Type” : “text”,
“Name” : “Data source version”,
“Required” : “mandatory”
},
{
“Example” : “Sus scrofa”,
“Description” : “the species of the sample record”,
“Allow Multiple” : “no”,
“Valid terms” : [
{
“ontology_name” : “NCBITaxon”,
“allow_descendants” : 1,
“include_root” : 0,
“term_iri” : “http://purl.obolibrary.org/obo/%22/%22NCBITaxon\_1”,
“term” : “NCBITaxon_1”
}
],
“Type” : “ontology_id”,
“Name” : “Species”,
“Required” : “mandatory”
}

```

```

    }
  ]
},
{
  "name" : "organism",
  "condition" : {
    "attribute_value_match" : {
      "Material" : "Organism"
    }
  },
  "rules" : [
    {
      "Type" : "BioSample ID or Alternative animal ID",
      "Name" : "Child of",
      "Required" : "optional",
      "Example" : "SAMEA4450075 or 11",
      "Description" : "Father and/or Mother's animal ID whose record must also be present as part of the subission or BioSamples ID if sample has been previously submitted",
      "Allow Multiple" : "max 2"
    },
    {
      "Type" : "text",
      "Name" : "Supplied breed",
      "Required" : "mandatory",
      "Example" : "Bunte Bentheimer",
      "Allow Multiple" : "no",
      "Description" : "the breed information contained in the original data source, could be a local language or breed variation"
    },
    {
      "Example" : "Germany",
      "Description" : "the country where the breed locates, the country name should be in the INSDC list",
      "Allow Multiple" : "no",
      "Type" : "text",
      "Name" : "EFABIS Breed country",
      "Required" : "mandatory"
    },
    {
      "Valid terms" : [
        {
          "term" : "\"LBO_0000000\"",
          "include_root" : 0,
          "term_iri" : "http://purl.obolibrary.org/obo/%22LBO_0000000\"",
          "ontology_name" : "\"LBO\"",

```

```

    "allow_descendants" : 1
  },
  ],
  "Example" : "LBO_0000347",
  "Description" : "If possible breed will be mapped to a standardised breed
ontology, both supplied and mapped breed will be available to end users",
  "Allow Multiple" : "no",
  "Name" : "Mapped breed",
  "Type" : "ontology_id",
  "Required" : "recommended"
},
{
  "Valid terms" : [
    {
      "ontology_name" : "PATO",
      "allow_descendants" : 0,
      "term" : "PATO_0000384",
      "include_root" : 1,
      "term_iri" : "http://purl.obolibrary.org/obo/PATO\_0000384"
    },
    {
      "allow_descendants" : 0,
      "ontology_name" : "PATO",
      "term_iri" : "http://purl.obolibrary.org/obo/PATO\_0000383",
      "include_root" : 1,
      "term" : "PATO_0000383"
    }
  ],
  "Description" : "the sex of the animal",
  "Allow Multiple" : "no",
  "Example" : "PATO:0000384 (male)",
  "Required" : "mandatory",
  "Name" : "Sex",
  "Type" : "ontology_id"
},
{
  "Name" : "Birth location",
  "Type" : "text",
  "Required" : "recommended",
  "Example" : "Germany",
  "Description" : "name of the place where the animal was given birth",
  "Allow Multiple" : "no"
},
{
  "Example" : "47.86591",
  "Description" : "the longitude of the place where the animal was given birth in

```



```

    decimal degrees",
    "Allow Multiple" : "no",
    "Type" : "number",
    "Name" : "Birth location longitude",
    "Required" : "recommended",
    "Valid units" : [
    "decimal degrees"
    ]
  },
  {
    "Example" : "12.19115",
    "Description" : "the latitude of the place where the animal was given birth in
    decimal degrees",
    "Allow Multiple" : "no",
    "Type" : "number",
    "Name" : "Birth location latitude",
    "Required" : "recommended",
    "Valid units" : [
    "decimal degrees"
    ]
  },
  {
    "Allow Multiple" : "no",
    "Name" : "Birth location accuracy",
    "Type" : "limited value",
    "Required" : "mandatory",
    "Valid values" : [
    "missing geographic information",
    "country level",
    "region level",
    "precise coordinates",
    "unknown accuracy level"
    ]
  }
],
},
{
  "name" : "specimen from organism",
  "condition" : {
    "attribute_value_match" : {
    "Material" : "Specimen from Organism"
    }
  },
  "rules" : [
  {
    "Name" : "Derived from",

```

```

    "Type" : "text",
    "Required" : "mandatory",
    "Example" : "SAMEA4450079",
    "Description" : "The data source id or alternative id of the animal from which
the sample was collected",
    "Allow Multiple" : "no"
  },
  {
    "Required" : "recommended",
    "Type" : "uri",
    "Name" : "Specimen collection protocol",
    "Description" : "A URL link to the protocol used to collect this sample",
    "Allow Multiple" : "no",
    "Example" : "ftp://ftp.faaang.ebi.ac.uk/ftp/protocols/samples/WUR\_SOP\_
animal\_sampling\_20160405.pdf"
  },
  {
    "Required" : "mandatory",
    "Type" : "date",
    "Name" : "Collection date",
    "Valid units" : [
      "YYYY-MM-DD",
      "YYYY-MM",
      "YYYY"
    ],
    "Description" : "Date when the sample was collected",
    "Allow Multiple" : "no",
    "Example" : "1992-01-17"
  },
  {
    "Valid units" : [
      "decimal degrees"
    ],
    "Name" : "Collection place latitude",
    "Type" : "number",
    "Required" : "recommended",
    "Example" : "47.86591",
    "Allow Multiple" : "no",
    "Description" : "latitude of the place where the sample was collected"
  },
  {
    "Valid units" : [
      "decimal degrees"
    ],
    "Required" : "recommended",
    "Name" : "Collection place longitude",

```

```

    "Type" : "number",
    "Allow Multiple" : "no",
    "Description" : "longitude of the place where the sample was collected",
    "Example" : "12.19115"
  },
  {
    "Required" : "mandatory",
    "Type" : "text",
    "Name" : "Collection place",
    "Description" : "name of the place where the sample was collected",
    "Allow Multiple" : "no",
    "Example" : "Germany"
  },
  {
    "Allow Multiple" : "no",
    "Required" : "mandatory",
    "Valid values" : [
      "missing geographic information",
      "country level",
      "region level",
      "precise coordinates",
      "unknown accuracy level"
    ],
    "Name" : "Collection place accuracy",
    "Type" : "limited value"
  },
  {
    "Required" : "mandatory",
    "Name" : "Organism part",
    "Type" : "ontology_id",
    "Valid terms" : [
      {
        "term" : "UBERON_0001062",
        "term_iri" : "http://purl.obolibrary.org/obo/UBERON\_0001062",
        "include_root" : 0,
        "ontology_name" : "UBERON",
        "allow_descendants" : 1
      }
    ],
    "Allow Multiple" : "no",
    "Description" : "Ontology for organism part where the specimen was taken from",
    "Example" : "UBERON_0001968 (semen)"
  },
  {
    "Type" : "ontology_id",

```

```

    "Name" : "Developmental stage",
    "Required" : "recommended",
    "Example" : "EFO_0001272 (adult)",
    "Description" : "developmental stage of the animal from which the specimen is
collected",
    "Allow Multiple" : "no",
    "Valid terms" : [
    {
    "allow_descendants" : 1,
    "ontology_name" : "EFO",
    "include_root" : 0,
    "term_iri" : "http://purl.obolibrary.org/obo/EFO\_0000399",
    "term" : "EFO_0000399"
    }
    ],
    },
    {
    "Type" : "ontology_id",
    "Name" : "Physiological stage",
    "Required" : "recommended",
    "Example" : "PATO_0001701 (mature)",
    "Allow Multiple" : "no",
    "Description" : "physiological stage of the animal from which the specimen is
collected",
    "Valid terms" : [
    {
    "term_iri" : "http://purl.obolibrary.org/obo/PATO\_0000261",
    "include_root" : 0,
    "term" : "PATO_0000261",
    "allow_descendants" : 1,
    "ontology_name" : "PATO"
    }
    ],
    },
    {
    "Description" : "Animal age at the point of collection.",
    "Allow Multiple" : "no",
    "Example" : "3 years",
    "Required" : "optional",
    "Type" : "number",
    "Name" : "Animal age at collection"
    },
    {
    "Example" : "mailto:vdc-internal@ebi.ac.uk",
    "Description" : "Either an e-mail address to contact about availability or a link
to a web page giving information on who to contact. E-mail addresses should

```

be prefixed with 'mailto:', e.g. '<mailto:samples@example.ac.uk>'. In either case, long term support of the web page or e-mail address is necessary. Group e-mail addresses are preferable to individual accounts."

```

"Allow Multiple" : "no",
"Type" : "uri_value or text",
"Name" : "Availability",
"Required" : "recommended",
"Valid values" : [
  "no longer available",
  "mailto:a.valid@email"
],
},
{
  "Required" : "recommended",
  "Valid values" : [
    "ambient temperature",
    "cut slide",
    "frozen, -80 degrees Celsius freezer",
    "frozen, -20 degrees Celsius freezer",
    "frozen, liquid nitrogen",
    "frozen, vapor phase",
    "paraffin block",
    "RNAlater, frozen -20 degrees Celsius",
    "TRIzol, frozen",
    "paraffin block at ambient temperatures (+15 to +30 degrees Celsius)",
    "freeze dried"
  ],
  "Type" : "text",
  "Name" : "Sample storage",
  "Allow Multiple" : "no",
  "Description" : "How the sample was stored. Temperatures are in degree Celsius. 'Frozen, vapor phase' refers to storing samples above liquid nitrogen in the vapor"
},
{
  "Allow Multiple" : "no",
  "Description" : "How the sample was prepared for storage",
  "Required" : "recommended",
  "Valid values" : [
    "cryopreservation in liquid nitrogen (dead tissue)",
    "cryopreservation in dry ice (dead tissue)",
    "cryopreservation of live cells in liquid nitrogen",
    "cryopreservation, other",
    "formalin fixed, unbuffered",
    "formalin fixed, buffered",
    "formalin fixed and paraffin embedded",

```

```

    "freeze dried (vaiable for reproduction)",
    "freeze dried"
  ],
  "Name" : "Sample storage processing",
  "Type" : "text"
},
{
  "Name" : "Sampling to preparation interval",
  "Type" : "number",
  "Required" : "optional",
  "Valid units" : [
    "minutes",
    "hours",
    "days",
    "weeks",
    "months",
    "years",
    "minute",
    "hour",
    "day",
    "week",
    "month",
    "year"
  ],
  "Description" : "How long between the sample being taken and the assay
experiment preperations commencing. If sample preperations were then left
in intermediate stages after preperation commenced, for example as sheared
chromatin, then this should be made clear in your protocols",
  "Allow Multiple" : "no"
}
]
}
]
}

```