

JIO Metadata Draft

- [What is metadata?](#)
- [Why use metadata?](#)
- [How to format metadata with JIO](#)
- [List of metadata to use](#)
 - [Identification](#)
 - [Intellectual property](#)
 - [Content](#)
- [Examples](#)
 - [Posting a webpage for JIO](#)
 - [Posting JIO library](#)
 - [Posting a webpage for interoperability levels](#)
 - [Posting an image](#)
 - [Posting a book](#)
 - [Posting a video](#)
 - [Posting a job announcement](#)
 - [Getting a list of document created by someone](#)
 - [Getting all documents about JIO in the resilience project](#)
- [Tools](#)
 - [W3C Date function](#)
- [Sources](#)

What is metadata?

The word "metadata" means "data about data". Metadata articulates a context for objects of interest -- "resources" such as MP3 files, library books, or satellite images -- in the form of "resource descriptions". As a tradition, resource description dates back to the earliest archives and library catalogs. The modern "metadata" field that gave rise to Dublin Core and other recent standards emerged with the Web revolution of the mid-1990s. <http://dublincore.org/metadata-basics/>

Why use metadata?

Uploading a document to several servers can be very tricky, because the document has to be saved in a place where it can be easily found with basic searches in all storages (For instance: ERP5, XWiki and Mioga2 have their own way to save documents and to get them). So we must use metadata *for interoperability reasons*. Interoperability is the ability of diverse systems and organizations to work together.

How to format metadata with JIO

See below XML and its JSON equivalent:

```
<dc:title>My Title</dc:title>
=
{"title":"My Title"}
```

```
<dc:contributor>Me</dc:contributor>
<dc:contributor>And You</dc:contributor>
=
{"contributor":["Me", "And You"]}
```

```
<dc:identifier scheme="DCTERMS.URI">http://my/resource</dc:identifier>
<dc:identifier>Xaoe41PAPNIWz</dc:identifier>
=
{"identifier": [
  {"scheme": "DCTERMS.URI", "content": "http://my/resource"},
  "Xaoe41PAPNIWz"
]}
```

List of metadata to use

Identification

- **"_id"**, a specific JIO metadata which helps the storage to find a documents (can be a real path name, a dc:identifier, a uuid, ...). **String Only**
- **"identifier"**, {"identifier": "http://domain/jio_home_page"}, {"identifier": "urn:ISBN:978-1-2345-6789-X"}, {"identifier": [{"scheme": "DCTERMS.URI", "content": "http://domain/jio_home_page"}]}

an unambiguous reference to the resource within a given context. Recommended best practice is to identify the

resource by means of a string or number conforming to a formal identification system. Examples of formal identification systems include the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL), the Digital Object Identifier (DOI) and the International Standard Book Number (ISBN).

- **"format"**, {"format": ["text/html", "52 kB"]}, {"format": ["image/jpeg", "100 x 100 pixels", "13.2 KiB"]}

the physical or digital manifestation of the resource. Typically, Format may include the media-type or dimensions of the resource. Examples of dimensions include size and duration. Format may be used to determine the software, hardware or other equipment needed to display or operate the resource.

- **"date"**, {"date": "2011-12-13T14:15:16Z"}, {"date": {"scheme": "DCTERMS.W3CDTF", "content": "2011-12-13"}}

a date associated with an event in the life cycle of the resource. Typically, Date will be associated with the creation or availability of the resource. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [Date and Time Formats, W3C Note, <http://www.w3.org/TR/NOTE-datetime>] and follows the YYYY-MM-DD format.

- **"type"**, {"type": "Text"}, {"type": "Image"}, {"type": "Dataset"}

the nature or genre of the content of the resource. Type includes terms describing general categories, functions, genres, or aggregation levels for content. Recommended best practice is to select a value from a controlled vocabulary. **The type is not a MIME Type!**

Intellectual property

- **"creator"**, {"creator": "Tristan Cavelier"}, {"creator": ["Tristan Cavelier", "Sven Franck"]}

an entity primarily responsible for making the content of the resource. Examples of a Creator include a person, an organization, or a service. Typically the name of the Creator should be used to indicate the entity.

- **"publisher"**, {"publisher": "Nexedi"}

the entity responsible for making the resource available. Examples of a Publisher include a person, an organization, or a service. Typically, the name of a Publisher should be used to indicate the entity.

- **"contributor"**, {"contributor": ["Full Name", "Full Name", ...]}

an entity responsible for making contributions to the content of the resource. Examples of a Contributor include a person, an organization or a service. Typically, the name of a Contributor should be used to indicate the entity.

- **"rights"**, {"rights": "Access limited to members"}, {"rights": "https://www.j-io.org/documentation/jio-documentation/#copyright-and-license"}

information about rights held in and over the resource. Typically a Rights element will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the rights element is absent, no assumptions can be made about the status of these and other rights with respect to the resource.

Content

- **"title"**, {"title": "JIO Home Page"}

the name given to the resource. Typically, a Title will be a name by which the resource is formally known.

- **"subject"**, {"subject": "JIO"}, {"subject": ["JIO", "basics"]}

the topic of the content of the resource. Typically, a Subject will be expressed as keywords or key phrases or classification codes that describe the topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.

- **"description"**, {"description": "Simple guide to show the basics of JIO"}, {"description": {"lang": "fr", "content": "Ma description"}}

an account of the content of the resource. Description may include but is not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.

- **"language"**, {"language": "en"}

a language of the intellectual content of the resource. Recommended best practice for the values of the Language element is defined by RFC 3066 [RFC 3066, <http://www.ietf.org/rfc/rfc3066.txt>] which, in conjunction with ISO 639 [ISO 639, <http://www.oasis-open.org/cover/iso639a.html>], defines two- and three-letter primary language tags with optional subtags. Examples include "en" or "eng" for English, "akk" for Akkadian, and "en-GB" for English used in the United Kingdom.

- **"source"**, {"source": ["Image taken from a drawing by Mr. Artist", "<phone number>"]} ,

a Reference to a resource from which the present resource is derived. The present resource may be derived from the Source resource in whole or part. Recommended best practice is to reference the resource by means of a string or number conforming to a formal identification system.

- **"relation"**, {"relation": "Resilience project"}

a reference to a related resource. Recommended best practice is to reference the resource by means of a string or number conforming to a formal identification system.

- **"coverage"**, {"coverage": "France"}

the extent or scope of the content of the resource. Coverage will typically include spatial location (a place name or geographic co-ordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity). Recommended best practice is to select a value from a controlled vocabulary (for example, the Thesaurus of Geographic Names [Getty Thesaurus of Geographic Names, <http://www.getty.edu/research/tools/vocabulary/tgn/>]). Where appropriate, named places or time periods should be used in preference to numeric identifiers such as sets of co-ordinates or date ranges.

- **"category"**, {"category": ["parent/26323", "resilience/javascript", "javascript/library/io"]}

the category the resource is associated with. The categories may look like navigational facets, they correspond to the properties of the resource which can be generated with metadata or some other informations (see [faceted search](#)).

- **"product"**, {"product": "..."}
for e-commerce use.

- **custom**, {custom1: value1, custom2: value2, ...}

Examples

Posting a webpage for JIO

```
jio.put({
  "_id"      : "...",
  "identifier" : "http://domain/jio_home_page",
  "format"    : ["text/html", "52 kB"],
  "date"      : new Date(),
  "type"      : "Text",
  "creator"   : ["Nexedi", "Tristan Cavelier", "Sven Franck"],
  "title"     : "JIO Home Page",
  "subject"   : ["JIO", "basics"],
  "description" : "Simple guide to show the basics of JIO",
  "category"  : ["resilience/jio", "webpage"],
  "language"  : "en"
}, callbacks); // send content as attachment
```

Posting JIO library

```
jio.put({
  "_id"      : "...",
  "identifier" : "jio.js",
  "date"      : "2013-02-15",
  "format"    : "application/javascript",
  "type"      : "Software",
  "creator"   : ["Tristan Cavelier", "Sven Franck"],
  "publisher" : "Nexedi",
  "rights"    :
    "https://www.j-io.org/documentation/jio-documentation/#copyright-and-license",
  "title"     : "JavaScript Input/Output",
  "subject"   : "JIO",
  "category"  : ["resilience/javascript", "javascript/library/io"]
  "description" : "jIO is a client-side JavaScript library to manage " +
    "documents across multiple storages."
}, callbacks); // send content as attachment
```

Posting a webpage for interoperability levels

```
jio.put({
  "_id"      : "...",
  "identifier" : "http://dublincore.org/documents/interoperability-levels/",
  "date"      : "2009-05-01",
  "format"    : "text/html",
  "type"      : "Text",
  "creator"   : ["Mikael Nilsson", "Thomas Baker", "Pete Johnston"],
```

```

"publisher" : "Dublin Core Metadata Initiative",
"title" : "Interoperability Levels for Dublin Core Metadata",
"description" : "This document discusses the design choices involved " +
    "in designing applications for different types of " +
    "interoperability. [...]",
"language" : "en"
}, callbacks); // send content as attachment

```

Posting an image

```

jio.put({
  "_id" : "...",
  "identifier" : "new_york_city_at_night",
  "format" : ["image/jpeg", "7.2 MB", "8192 x 4096 pixels"],
  "date" : "1999",
  "type" : "Image",
  "creator" : "Mr. Someone",
  "title" : "New York City at Night",
  "subject" : ["New York"],
  "description" : "A photo of New York City taken just after midnight",
  "coverage" : ["New York", "1996-1997"]
}, callbacks); // send content as attachment

```

Posting a book

```

jio.put({
  "_id" : "...",
  "identifier" : {"scheme": "DCTERMS.URI", "content": "urn:ISBN:0385424728"},
  "format" : "application/pdf",
  "date" : {"scheme": "DCTERMS.W3CDTF", "content": getW3CDate()}, // see tools below
  "creator" : "Original Author(s)",
  "publisher" : "Me",
  "title" : {"lang": "en", "content": "..."},
  "description" : {"lang": "en", "content": "Summary: ..."},
  "language" : {"scheme": "DCTERMS.RFC4646", "content": "en-GB"}
}, callbacks); // send content as attachment

```

Posting a video

```

jio.put({
  "_id" : "...",
  "identifier" : "my_video",
  "format" : ["video/ogg", "130 MB", "1080p", "20 seconds"],
  "date" : getW3CDate(), // see tools below
  "type" : "Video",
  "creator" : "Me",
  "title" : "My life",
  "description" : "A video about my life"
}, callbacks); // send content as attachment

```

Posting a job announcement

```

jio.post({
  "format" : "text/html",
  "date" : "2013-02-14T14:44Z",
  "type" : "Text",
  "creator" : "James Douglas",
  "publisher" : "Morgan Healey Ltd",
  "title" : "E-Commerce Product Manager",
  "subject" : "Job Announcement",
  "description" : "...",
  "language" : "en-GB",
  "source" : "James@morganhealey.com",
  "relation" : ["TotalJobs"],
  "coverage" : "London, South East",
  "job_type" : "Permanent",
  "salary" : "£45,000 per annum"
}, callbacks); // send content as attachment
// result: http://www.totaljobs.com/JobSeeking/E-Commerce-Product-Manager_job55787655

```

Getting a list of document created by someone

With complex query:

```
jio.allDocs({"query": "creator: \"someone\""}, callbacks);
```

Getting all documents about JIO in the resilience project

With complex query:

```
jio.allDocs({"query": "subject: \"JIO\" AND category: \"resilience\"", callbacks});
```

Tools

W3C Date function

```
/**
 * Tool to get the date in W3C date format
 * - "2011-12-13T14:15:16+01:00" with use_utc = false (by default)
 * - "2011-12-13T13:15:16Z" with use_utc = true
 *
 * @param {Boolean} use_utc Use UTC format
 * @return {String} The date in W3C date format
 */
function getW3CDate(use_utc) {
  var d = new Date(), offset;
  if (use_utc === true) {
    return d.toISOString();
  }
  offset = - d.getTimezoneOffset();
  return (
    d.getFullYear() + "-" +
    (d.getMonth() + 1) + "-" +
    d.getDate() + "T" +
    d.getHours() + ":" +
    d.getMinutes() + ":" +
    d.getSeconds() + "." +
    d.getMilliseconds() +
    (offset < 0 ? "-" : "+") +
    (offset / 60) + ":" +
    (offset % 60)
  ).replace(/[0-9]+/g, function (found) {
    if (found.length < 2) {
      return '0' + found;
    }
    return found;
  });
}
```

Sources

- [Interoperability definition](#)
- [Faceted search](#)
- [DublinCore](#)
 - [Interoperability levels](#)
 - [Metadata elements](#)
 - <http://www.chu-rouen.fr/documed/eahilsantander.html>
 - http://openweb.eu.org/articles/dublin_core (French)
- [CouchDB](#)
- [Resource Description Framework \(RDF\)](#)
- [Five Ws](#)
- [Metadata](#)
- MIME Types
 - https://en.wikipedia.org/wiki/Internet_media_type
 - <https://www.iana.org/assignments/media-types>