

Including Data in the Standardized Markup for Journal Articles

Deborah Aleyne Lapeyre
Mulberry Technologies, Inc.

17 West Jefferson Street, Suite 207

Rockville, MD 20850

Phone: 301/315-9631

Fax: 301/315-8285

info@mulberrytech.com

<http://www.mulberrytech.com>

Version 1.0 (June 2015)

©2015 Mulberry Technologies, Inc.



Including Data in the Standardized Markup for Journal Articles

JATS: The Journal Article Tag Suite	1
JATS (ANSI/NISO Z39-96-2012) is	1
JATS Names XML Elements for Publishing	2
How Publishers Cite Data	2
How JATS Tags References	2
Force11 Recommends JATS Mixed Citation	3
What is needed to Cite Data?	3
Dataset Description Metadata	4
How Publishers Want Data Cited	4
New JATS Elements Requested by Force11	5
JATS Elements for Citing Data (1)	5
JATS Elements for Citing Data (2)	6
New Attributes Values (1)	7
New Attributes and Values for @pub-id	7
Machine Resolvable Problem Not Solved	8
What Else is Needed?	8
Data Citation Examples	8
Dryad Digital Repository, referenced through a DOI	9
GenBank Protein	9
RNA Sequence	10
Protein Data Bank in Europe sample	10
Data in figshare, referenced through a DOI	11
Data Curator	12
Assigning Authority	13
New @pub-id-type Values	14
External Media: Database on CD-ROM, DVD, or Disk	14
Record from a Web Data Repository	15
<i>Add Health</i> Data Set	15
GigaScience Sample	16
Colophon	16

Appendixes

Appendix A: Possible Elements in a JATS <mixed-citation>	1
Appendix B: Mapping Data Citing Components to JATS Elements	1

JATS: The Journal Article Tag Suite

- The article publishing piece of the data citing story
- JATS enables publishers to cite data sources *in journal articles*
- Tagging allows:
 - human readability
 - machine discoverability
 - flexibility to express different types of data citations

JATS (ANSI/NISO Z39-96-2012) is

- XML for tagging journal articles
- Used by:
 - STM journal publishers (production tag set and/or interchange)
(US, England, Japan, Korea, Australia, Canada, Brazil, China, Germany, Norway, Sweden, Switzerland, France, Croatia, Russia, Belgium, Egypt, Oman, United Arab Emirates, etc.)
 - National Libraries (US, UK, Australia)
 - Archives (PubMed Central, JSTORE/ITHAKA)
 - Aggregators and web-hosts (Highwire, Silverchair, Atypon)
 - Standards bodies to produce standards (ISO, IEEE)

JATS Names XML Elements for Publishing

- JATS available in DTD, XSD, and RNG XML model formats
- The Tag Set names and describes the content of:
 - metadata elements (contributor, surname, abstract)
 - textual elements (paragraph, figure, verse)
 - Tables (XHTML and OASIS models)
 - MathML elements (2.0 or 3.0)
 - bibliographic reference elements (article title, publisher, publication year)

How Publishers Cite Data

- In the narrative text
- In the bibliography (references list)
- In an additional reference list just for data

Force11 recommends tagging them as references are tagged

How JATS Tags References

- Bibliographic reference lists (<ref-list>) are in the back of:
 - articles
 - sections
 - boxed-text
- Reference lists contain references (<ref>)
- References contain citations (<mixed-citation>) each of which contains the description of one cited source

Force11 Recommends JATS Mixed Citation

<mixed-citation> is

- a bag-of-text with all punctuation and spacing preserved
- some elements inside can be tagged
- how much tagging is up to the publisher

Lapeyre, Deborah Aleyne, *Poodles of the World*. Journal of Big Dogs, 2015
vol: 13, pages: 2525-2535 DOI: 10.1165/JCM.02419-05

```
<ref id="B45">
<mixed-citation publication-type="journal">
<string-name>
  <surname>Lapeyre</surname>,
  <given-names>Deborah Aleyne</given-names>
</string-name>,
<article-title>Poodles of the World</article-title>.
<source>Journal of Big Dogs</source>,
<year>2015</year> vol: <volume>13</volume>,
pages: <fpage>2525</fpage>-<lpage>2535</lpage> DOI:
<pub-id pub-id-type="doi">10.1165/JCM.02419-05</pub-id>
</mixed-citation>
</ref>
```

What is needed to Cite Data?

- Best practices for dataset description
(what an archive should keep)
- Data citing recommendations from style guides, publishers, archives, researchers, consortia

Dataset Description Metadata

Force11 minimum elements that should be present in a dataset description:

1. Dataset Identifier
2. Title of the dataset
3. Creator
4. Publisher/Contact
5. Publication Date/ Year / Release Date
6. Version of the dataset
7. *Description (longer explanation than the title)*

Nearly all can/should be part of a bibliographic citation

How Publishers Want Data Cited

- Over 55 sources were polled on what data elements to use to cite data
- Here are the top 10 (mentioned by most, mandatory in many)
 1. Persistent global dataset Identifier
 2. Title/Name of the dataset
 3. Author/Creator
 4. Publisher/Distributor/Repository
 5. Publication Date / Year / Release Date
 6. Version of the dataset
 7. Resource Type
 8. Location of publisher/distributor
 9. Access date and time
 10. Additional URI/location/bridge service

New JATS Elements Requested by Force11

- `<data-title>`
 - the formal title or name of a cited data source (or a component of a cited data source)
 - equivalent to `<article-title>`
 - may be used with `<source>` for hierarchical relationships`</source>`
- `<version>`
 - full version statement (maybe only a number) for cited data or software
 - `@designator` attribute can hold the simple version number:
`<version designator="16.2">16th version, second release</version>`

JATS Elements for Citing Data (1)

1. Persistent Global Identifier
 - `<pub-id pub-id-type='doi'>`
2. Title/Name of the dataset
 - `<data-title>` (similar to `<article-title>`)
 - `<source>`
3. Author/Creator
 - `<name>` or `<string-name>`
 - `<collab>`
4. Publisher/Distributor/Repository
 - `<publisher>`
5. Publication Date / Year / Release Date
 - `<date>`
 - `<year>`

JATS Elements for Citing Data (2)

1. Version of the dataset
 - <version>
 - <edition>
 - <date-in-citation content-type="update">
2. Resource Type
 - @publication-format
(print, electronic, video, audio, ebook, online-only)
3. Location of publisher/distributor
 - <publisher-loc>
4. Access date and time
 - <date-in-citation content-type="access-date">
 - <year>
5. Additional URI/location/bridge service
 - <ext-link>
 - <uri>

New Attributes Values (1)

- @publication-type on citations
 - typically “book”, “journal”, “standard”
 - new value “data”
 - defined as “a dataset or other research collection such as a spreadsheet”
- @person-group-type on <person-group>
 - typically “author”, “editor”, “compiler”
 - new value “curator”
 - used for citing datasets and art

New Attributes and Values for @pub-id

- New attribute @assigning-authority
 - says who assigned the ID (such as an ARK or DOI)
 - values are organizations such as “crossref”, “figshare”, “pdb”, “genbank”, “pubmed”
- The @pub-id-type (“doi”, “archive”, “isbn”) gets new values for citing data:
 - “accession” (Bioinformatics: a unique identifier given to a DNA or protein sequence record for tracking the sequence record and the associated sequence over time in a data repository.)
 - “ark” (Archival Resource Key: a Uniform Resource Locator (URL) containing the word "ark" that is a multi-purpose identifier for information objects of any type)
 - “handle” (HDL: Handle identifier, part of the Handle System for assigning, managing, and resolving persistent identifiers for digital objects and other resources on the Internet)

Machine Resolvable Problem Not Solved

- JATS enables; it does *not enforce*
- JATS was designed for interchange among:
 - publishers and their partners
 - archives and libraries
 - aggregators and hosting services
- There is no one right way to cite data
 - different publishers different styles
 - *how much* to record is a business decision

What Else is Needed?

Data miners and machine resolvers need

- As much uniformity as possible
- Common agreements
- Best practices

JATS4R (JATS for Reuse: <http://jats4r.org>)

Data Citation Examples

As we have time and desire to geek

With thanks to Daniel Mietchen, Johanna McEntyre, Jeff Beck, Chris Maloney, and the Force11 Data Citation Implementation Group

Dryad Digital Repository, referenced through a DOI

Dubuis JO, Samanta R, Gregor T (2013). Data from: *Accurate measurements of dynamics and reproducibility in small genetic networks*. Dryad Digital Repository doi:10.5061/dryad.35h8v

```
<mixed-citation publication-type="data">Dubuis JO, Samanta R,
Gregor T (<year iso-8601-date="2013">2013</year>). Data from:
<data-title>Accurate measurements of dynamics and reproducibility
in small genetic networks</data-title>. <source>Dryad Digital
Repository</source> doi:<pub-id pub-id-type="doi">10.5061/dryad.35h8v</pub-id>
</mixed-citation>
```

GenBank Protein

Homo sapiens cAMP responsive element binding protein 1 (CREB1), transcript variant A, mRNA. GenBank NM_004379.3.

```
<mixed-citation publication-type="data">
<data-title>Homo sapiens cAMP responsive element binding protein 1
(CREB1), transcript variant A, mRNA</data-title>. <source>GenBank</source>
<ext-link ext-link-type="genbank" xlink:href="NM_004379.3">NM_004379.3</ext-
link>.
</mixed-citation>
```

RNA Sequence

Xu, J. et al. *Cross-platform ultradeep transcriptomic profiling of human reference RNA samples by RNA-Seq*. *Sci. Data* 1:140020 doi: 10.1038/sdata.2014.20 (2014).

```
<mixed-citation publication-type="data">Xu, J. <etal/>
<data-title>Cross-platform ultradeep transcriptomic profiling
of human reference RNA samples by RNA-Seq</data-title>.
<source>Sci. Data</source> <volume>1</volume>:
<elocation-id>140020</elocation-id>
doi: <pub-id pub-id-type="doi">10.1038/sdata.2014.20</pub-id>
(<year iso-8601-date="2014">2014</year>).
</mixed-citation>
```

Protein Data Bank in Europe sample

Kollman JM, Charles EJ, Hansen JM, 2014, *Cryo-EM structure of the CTP synthetase filament*, <http://www.ebi.ac.uk/pdbe/entry/EMD-2700>, Publicly available from The Electron Microscopy Data Bank (EMDB).

```
<mixed-citation publication-type="data">Kollman JM, Charles EJ, Hansen JM,
<year iso-8601-date="2014">2014</year>, <data-title>Cryo-EM structure of
the CTP synthetase filament</data-title>, <ext-link ext-link-type="uri"
xlink:href="http://www.ebi.ac.uk/pdbe/entry/EMD-2700">
http://www.ebi.ac.uk/pdbe/entry/EMD-2700</ext-link>, Publicly available
from <source>The Electron Microscopy Data Bank (EMDB)</source>.
</mixed-citation>
```

Data in figshare, referenced through a DOI

Mulvany, Ian, *citing-dataset-elements*. FigShare, 2014/06/30, 10.6084/m9.figshare.1088363.

```
<mixed-citation publication-type="data">
<name><surname>Mulvany</surname><given-names>Ian</given-names></name>,
<data-title>citing-dataset-elements</data-title>. <source>FigShare</source>,
<date-in-citation content-type='pub-date' iso-8601-date='2014-06-30'>
<year>2014</year>/<month>06</month>/<day>30</day></date-in-citation>,
<pub-id pub-id-type='doi'
  xlink:href='http://dx.doi.org/10.6084/m9.figshare.1088363'
  assigning-authority='figshare'>10.6084/m9.figshare.1088363</pub-id>.
</mixed-citation>
```

Di Stefano B, Collombet S, Graf T. Figshare <http://dx.doi.org/10.6084/m9.figshare.939408> (2014).

```
<mixed-citation publication-type="data">Di Stefano B, Collombet S,
Graf T. <source>Figshare</source> <ext-link ext-link-type="uri"
  xlink:href="http://dx.doi.org/10.6084/m9.figshare.939408">
http://dx.doi.org/10.6084/m9.figshare.939408</ext-link>
(<year iso-8601-date="2014">2014</year>).
</mixed-citation>
```

Data Curator

The value “curator” was added to the list of suggested values for the @person-group-type attribute. Here is an example of how the “curator” value might be used for @person-group-type:

Frankis, Michael, curator. "Mountain bluebird." *Encyclopedia of Life*, available from <http://eol.org/pages/1177542>. Accessed 30 Mar 2015.

```
<mixed-citation publication-type="data">
<person-group person-group-type='curator'>
<name><surname>Frankis</surname><given-names>Michael</given-names></name>
</person-group>, curator. "<data-title>Mountain bluebird</data-title>."
<source>Encyclopedia of Life</source>, available from
<ext-link ext-link-type='uri' xlink:href='http://eol.org/pages/1177542'>
http://eol.org/pages/1177542</ext-link>. Accessed
<date-in-citation content-type="access-date"
iso-8601-date="2015-03-30">30 Mar 201</date-in-citation>.
</mixed-citation>
```


Assigning Authority

A new attribute `@assigning-authority` was added to the elements `<ext-link>` and `<pub-id>`. The existing attribute `@pub-id-type` should now only be used to state how the element content is to be interpreted as an identifier. For example, a “DOI” would have the `@pub-id-type` attribute value of “doi”, and the `@assigning-authority` attribute value might be “crossref” or “figshare”. (Note that values are in lowercase for both attributes!) Another example from the life sciences would be: `@pub-id-type` value of “accession”, `@assigning-authority` of “uniprot”.

Mulvany, Ian, *citing-dataset-elements*. Figshare, 2014/06/30, 10.6084/m9.figshare.1088363.

```
<mixed-citation publication-type="data">
<name><surname>Mulvany</surname><given-names>Ian</given-names></name>,
<data-title>citing-dataset-elements</data-title>. <source>FigShare</source>,
<date-in-citation content-type="pub-date" iso-8601-date='2014-06-30'>
<year>2014</year>/<month>06</month>/<day>30</day></date-in-citation>,
<pub-id pub-id-type='doi'
xlink:href='http://dx.doi.org/10.6084/m9.figshare.1088363'
assigning-authority='figshare'>10.6084/m9.figshare.1088363</pub-id>.
</mixed-citation>
```

New @pub-id-type Values

New values for the @pub-id-type attribute (“accession”, “ark”, and “handle”) were added to JATS for tagging data sources.

Heinz D.W., Baase W.A., et al. *How amino-acid insertions are allowed in an alpha-helix of T4 lysozyme*. RCSB Protein Data Bank, accession 1021.

10.2210/pdb1021/pdb

```
<mixed-citation publication-type='data'>
<name><surname>Heinz</surname><given-names>D.W.</given-names></name>,
<name><surname>Baase</surname><given-names>W.A.</given-names></name>,
<etal>et al.</etal> <data-title>How amino-acid insertions are allowed in
an alpha-helix of T4 lysozyme</data-title>.
<source>RCSB Protein Data Bank</source>, accession
<pub-id pub-id-type='accession' assigning-authority='pdb'
xlink:href='http://www.rcsb.org/pdb/explore/explore.do?structureId=1021'>1021</
pub-id>.
<pub-id pub-id-type='doi' xlink:href='http://dx.doi.org/10.2210/pdb1021/pdb'>
10.2210/pdb1021/pdb</pub-id>
</mixed-citation>
```

External Media: Database on CD-ROM, DVD, or Disk

Walker MM, Keith LH. EPA's Clean Air Act air toxics database [disk]. Boca Raton (FL): Lewis Publishers; 1992-1993. 4 computer disks: 3 1/2 in.

```
<mixed-citation publication-type="data" publication-format="disk">
<name><surname>Walker</surname><given-names>MM</given-names></name>,
<name><surname>Keith</surname><given-names>LH</given-names></name>.
<data-title>EPA's Clean Air Act air toxics database</data-title> [disk].
<publisher-loc>Boca Raton (FL)</publisher-loc>: <publisher-name>Lewis Publish-
ers</publisher-name>;
<date-in-citation content-type="copyright-year"
iso-8601-date="1992">1992-1993</date-in-citation>.
4 computer disks: 3 1/2 in.</mixed-citation>
```

Record from a Web Data Repository

Benz, Michael; Braband, Henrik; Schmutz, Paul; Halter, Jonathan; Alberto, Roger. *C21 H49 Al Cl7 N7 O7 Tc*, version 130981. From Crystallography Open Database, accession 1517518.

```
<mixed-citation publication-type='data'>
<name><surname>Benz</surname><given-names>Michael</given-names></name>;
<name><surname>Braband</surname><given-names>Henrik</given-names></name>;
<name><surname>Schmutz</surname><given-names>Paul</given-names></name>;
<name><surname>Halter</surname><given-names>Jonathan</given-names></name>;
<name><surname>Alberto</surname><given-names>Roger</given-names></name>.
<data-title>C21 H49 Al Cl7 N7 O7 Tc</data-title>,
version <version>130981</version>.
From <source>Crystallography Open Database</source>, accession
<pub-id pub-id-type='accession'
  assigning-authority='crystallography open database'
  xlink:href='http://www.crystallography.net/cod/1517518.html'>1517518</pub-id>.
</mixed-citation>
```

Add Health Data Set

Harris, Kathleen Mullan. 2009. *The National Longitudinal Study of Adolescent to Adult Health (Add Health), Waves I & II, 1994–1996; Wave III, 2001–2002; Wave IV, 2007–2009* [machine-readable data file and documentation]. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill. DOI: 10.3886/ICPSR27021.v9

```
<mixed-citation publication-type="data">
<name><surname>Harris</surname><given-names>Kathleen Mullan</given-names></name>.
<date-in-citation content-type="pub-date"><year>2009</year></date-in-citation>.
<data-title>The National Longitudinal Study of Adolescent to Adult
Health (Add Health), Waves I & II, 1994–1996; Wave III,
2001–2002; Wave IV, 2007–2009</data-title>
[machine-readable data file and documentation]. <publisher-loc>Chapel Hill,
NC</publisher-loc>: <publisher-name>Carolina Population Center, University of
North Carolina at Chapel Hill</publisher-name>. DOI: <pub-id pub-id-type='doi'
xlink:href='http://dx.doi.org/10.3886/ICPSR27021.v9'>10.3886/ICPSR27021.v9</pub-id>
</mixed-citation>
```

GigaScience Sample

Zheng LY, Guo XS, He B, Sun LJ, Pi CM, Jing H-C: Genome data from [http://dx.doi.org/10.5524/100012] GigaScience 2011.

```
<mixed-citation publication-type="data">Zheng LY,  
Guo XS, He B, Sun LJ, Pi CM, Jing H-C: Genome data from  
[<ext-link ext-link-type="uri" xlink:href="http://dx.doi.org/10.5524/100012">  
http://dx.doi.org/10.5524/100012</ext-link>] <source>GigaScience</source>  
<year iso-8601-date="2011">2011</year>.  
</mixed-citation>
```

Colophon

- Slides and handouts created from a single XML source
- Projected in HTML (created from XML by XSLT)
- Handouts distributed in PDF
 - source XML transformed to XHTML + CSS
 - PDF from that
 - all lights out; no pagination or tables adjusted

Possible Elements in a JATS <mixed-citation>

A <mixed-citation> element is a bag-of-text that may contain, intermixed with the text (letters, numbers, or special characters), the following elements:

Any combination of:

- <inline-supplementary-material> Inline Supplementary Material Metadata
- Related Material Elements
 - <related-article> Related Article Information
 - <related-object> Related Object Information
- <hr> Horizontal Rule
- <string-date> Date as a String
- Emphasis Elements
 - <bold> Bold
 - <fixed-case> Fixed Case
 - <italic> Italic
 - <monospace> Monospace Text (Typewriter Text)
 - <overline> Overline
 - <overline-start> Overline Start
 - <overline-end> Overline End
 - <roman> Roman
 - <sans-serif> Sans Serif
 - <sc> Small Caps
 - <strike> Strike Through
 - <underline> Underline
 - <underline-start> Underline Start
 - <underline-end> Underline End
 - <ruby> Ruby Annotation Wrapper
- <alternatives> Alternatives For Processing
- Inline Display Elements
 - <inline-graphic> Graphic, Inline
 - <private-char> Private Character (Custom or Unicode)
- <chem-struct> Chemical Structure (Display)

Including Data in the Standardized Markup for Journal Articles

- `<inline-formula>` Formula, Inline
- `<label>` Label (of an Equation, Figure, Reference, etc.)
- Math Elements
 - `<tex-math>` TeX Math Equation
 - `<mml:math>` Math (MathML Tag Set)
- Other Inline Elements
 - `<abbrev>` Abbreviation or Acronym
 - `<milestone-end>` Milestone End
 - `<milestone-start>` Milestone Start
 - `<named-content>` Named Special (Subject) Content
 - `<styled-content>` Styled Special (Subject) Content
- `<annotation>` Annotation in a Citation
- `<article-title>` Article Title
- `<chapter-title>` Chapter Title in a Citation
- `<collab>` Collaborative (Group) Author
- `<collab-alternatives>` Collaboration Alternatives
- `<comment>` Comment in a Citation
- `<conf-acronym>` Conference Acronym
- `<conf-date>` Conference Date
- `<conf-loc>` Conference Location
- `<conf-name>` Conference Name
- `<conf-sponsor>` Conference Sponsor
- `<data-title>` Data Title
- `<date>` Date
- `<date-in-citation>` Date within a Citation
- `<day>` Day
- `<edition>` Edition Statement, Cited
- Linking Elements
 - `<email>` Email Address
 - `<ext-link>` External Link
 - `<uri>` Uniform Resource Identifier (URI)
- `<elocation-id>` Electronic Location Identifier
- `<etal>` Et Al.
- `<fpage>` First Page

Including Data in the Standardized Markup for Journal Articles

- <gov> Government Report, Cited
- <institution> Institution Name: in an Address
- <institution-wrap> Institution Wrapper
- <isbn> ISBN
- <issn> ISSN
- <issn-l> ISSN-L (Linking ISSN)
- <issue> Issue Number
- <issue-id> Issue Identifier
- <issue-part> Issue Part
- <issue-title> Issue Title
- <lpage> Last Page
- <month> Month
- <name> Name of Person
- <name-alternatives> Name Alternatives
- <object-id> Object Identifier
- <page-range> Page Ranges
- <part-title> Part Title in a Citation
- <patent> Patent Number, Cited
- <person-group> Person Group for a Cited Publication
- <pub-id> Publication Identifier for a Cited Publication
- <publisher-loc> Publisher's Location
- <publisher-name> Publisher's Name
- <role> Role or Function Title of Contributor
- <season> Season
- <series> Series
- <size> Size
- <source> Source
- <std> Standard, Cited
- <string-name> Name of Person (Unstructured)
- <supplement> Supplement Information
- <trans-source> Translated Source
- <trans-title> Translated Title
- <version> Version Statement
- <volume> Volume Number

Including Data in the Standardized Markup for Journal Articles

- <volume-id> Volume Identifier
- <volume-series> Volume Series
- <year> Year
- <fn> Footnote
- <target> Target of an Internal Link
- <xref> X (cross) Reference
- Baseline Change Elements
 - <sub> Subscript
 - <sup> Superscript
- <x> X - Generated Text and Punctuation

Mapping Data Citing Components to JATS Elements

Prior to the June 2014 Force11 meeting, over 55 primary data sources (style guides, Archive submission guidelines, publisher's websites, schemas such as the DataCite Schema, articles on citing data by thought leaders, etc.) were reviewed to see what data fields were recommended for citing data such as genomic datasets. While dozens of data items were mentioned, most of the sources agreed on some variation of the top ten, with many making these mandatory. In the following pages, these requested data fields have been mapped to the JATS elements from JATS Committee Draft 1.1d3.

In the pages that follow:

- A numbered heading gives the data field name or names (as found in multiple sources).
- The paragraph below it will give an approximate definition. (Many definitions have been taken from ESIP Data Citation Guidelines [Ruth Duerr 2012] and the DataCite Schema documentation.)
- The bulleted item(s) that follow show JATS elements that could be used to represent this data within a citation. A tagged sample of each element is given.

1. Persistent Global Dataset Identifier / Locator / DOI / URL

Possibly a URL, but ideally a persistent identifier (DOI, PURL, Handle, ARK). The HTTP form of the DOI is preferred by some sources.

- `<pub-id>` with `@pub-id-type`

```
<pub-id pub-id-type="doi">10.1128/JCM.02410-08</pub-id>
```

```
<pub-id pub-id-type="doi">10.1099/ijs.0.039248-0</pub-id>
```

Linking attributes can be added to make the non-URL-DOI a live link:

```
<pub-id="doi" xlink:href="http://dx.doi.org/http://dx.doi.org/10.6070/H4WM1BBQ">
```

```
10.6070/H4WM1BBQ</pub-id>
```

Including Data in the Standardized Markup for Journal Articles

- `<ext-link>` with `@ext-link-type`
`<ext-link-type="uri"`
`xlink:href="http://dx.doi.org/http://dx.doi.org/10.6070/H4WM1BBQ">`
`http://dx.doi.org/http://dx.doi.org/10.6070/H4WM1BBQ</ext-link>`
- `<uri`
`<uri xlink:href="http://www.biomedcentral.com/1471-2180/13/198"></uri>`

7. Publisher/Distributor/ Repository/ Data Center /Archive

The organization distributing and curating the data (ideally over the long term) such as a Data Center or Archive

- `<publisher-name>`
`<publisher-name>Public Library of Science</publisher-name>`

8. Location of publisher/distributor

Such as city, state, country

- `<publisher-loc>`
`<publisher-loc>San Francisco, USA</publisher-loc>`

9. Access Date(s) and Time

Exactly when the online data was accessed.

- `<date-in-citation>`
`<date-in-citation content-type="access-date"`
`iso-8601-date="2014-06-13:10:00">`
Accessed on: `<year>2014</year>`, `<month>June</month>`,
`<day>13</day>` at 10:00am`</date-in-citation>`

10. Additional URI/Location /bridge service

Typically used for a URL in addition to the regular DOI

- `<ext-link>` with `@ext-link-type`
`<ext-link ext-link-type="uri"`
`xlink:href="http://r-forge.r-project.org/projects/splits">`
`http://r-forge.r-project.org/projects/splits</ext-link>`
- `<uri>`
`<uri xlink:href="http://www.biomedcentral.com/1471-2180/13/198"></uri>`

2. Title/Name of the Dataset

Formal title of the dataset (may include applicable dates). Similar to an article title in its role in the citation. Because a dataset located in a repository or inside a portion of a repository, there are two elements available. The `<source>` can be used to name repository levels.

- `<data-title>`
`<data-title>`Monitoring the Future: A Continuing Study of American Youth (12th Grade Survey)`</data-title>`
- `<source>` `<source>`figshare`</source>` OR
`<source>`Dryad Digital Repository`</source>`

3. Creator/Author/Rightsholder/Primary Responsibility

Data creators. People or organizations responsible for developing (intellectual work) the dataset. Primary Responsibility

Potential JATS Equivalents:

- `<name>`
`<name>`
`<surname>`Edelstein`</surname>`
`<given-names>`PH`</given-names>`
`</name>`
- `<string-name>` `<string-name>`
`<surname>`Edelstein`</surname>` ,
`<given-names>`PH`</given-names>`
`</string-name>`
- `<person-group name>`
`<person-group person-group-type="author">`
`<name>`
`<surname>`Edelstein`</surname>`
`<given-names>`PH`</given-names>`
`</name>`
`</person-group>`

Including Data in the Standardized Markup for Journal Articles

- person-group/collab

```
<person-group person-group-type="author">  
<collab collab-type="compilers">The BAC Resource Consortium</collab>  
</person-group>
```

- <institution>

```
<institution content-type="university">Boston University</institution>
```

- <institution-wrap>

```
<institution-wrap>  
<institution-id institution-id-type="Ringgold">1812</institution-id>  
<institution content-type="university">Harvard University</institution>  
</institution-wrap>
```

```
<institution-wrap>  
<institution-id institution-id-type="Ringgold">1846</institution-id>  
<institution-id institution-id-type="ISNI">0000 0001 2170 1429</institution-  
id>  
<institution content-type="university">Boston University</institution>  
</institution-wrap>
```

4. Publisher

The organization distributing and curating the data (responsible for its persistence, ideally over the long term) such as a Data Center or Archive

- <publisher-name>

```
<publisher-name>Lewis Publishers</publisher-name>
```

OR

```
<publisher-name>Carolina Population Center, University of  
North Carolina at Chapel Hill</publisher-name>
```

OR

```
<publisher-name>Public Library of Science</publisher-name>
```

5. Publication Date/ Year / Release Date

When this version of the dataset was made available for citation. May be only a year. Some sources place this inside the dataset title/name.

- `<date>`
`<date iso-8601-date="2015-06">`
`<month>June</month><year>2015</year>`
`</date>`
- `<year>`
`<year iso-8601-date="2015-06">2015</year>`

6. Version

The precise version number of the data used.

- `<version>`
`<version>16.2.1</version>`

7. Resource Type

Material designator; medium; general type description The only way current JATS has to record this is @publication-format/@publication-type.

```
<mixed-citation publication-type="data"
  publication-format="online">...
<mixed-citation publication-type="data"
  publication-format="spreadsheet">...
```

8. Location of Publisher/Distributor

Location of the party publishing the data; may include such as city, state, country.

- `<publisher-loc>`
`<publisher-loc>San Francisco, USA</publisher-loc>`

9. Access Date and Time

Exactly when the online data was accessed

- `<date-in-citation>`
`<date-in-citation content-type="access-date"`
 `iso-8601-date="2014-06-13:10:00">`
Accessed on: `<year>2014</year>`, `<month>June</month>`,
`<day>13</day>` at 10:00am
`</date-in-citation>`

10. Additional URI/ Location / Bridge Service

Additional URI, location, bridge service, secondary distributor, reflector, or other institutional role such as funding. Typically holds a URL in addition to the regular DOI

- `<ext-link>` with `@ext-link-type` attribute

```
<ext-link ext-link-type="uri" xlink:href="http://  
r-forge.r-project.org/projects/splits">
```

```
http://r-forge.r-project.org/projects/splits</ext-link>
```

- `<uri>`

```
<uri xlink:href="http://www.biomedcentral.com/1471-2180/13/198"
```

```
www.biomedcentral.com/1471-2180/13/198</uri>
```