



DDI Working Paper Series -- Best Practices, No. 6

Subject

Creating a DDI Profile (2009-02-15)

Document identifier:

<http://dx.doi.org/10.3886/DDIBestPractices06>

Authors:

Sanda Ionescu

Editors:

Mary Vardigan

Abstract:

This document outlines recommended best practices for creating a local DDI 3.0 Profile, which is a subset of DDI 3.0 fields to be used by an organization or shared by a community of users.

The DDI 3.0 specification is extensive and designed to cover a multiplicity of use cases. However, not all of DDI's possible applications will be relevant at the level of specific organizations or user communities. By creating and implementing user-specific profiles, organizations will:

- Ensure that DDI documentation is suited to their particular requirements
- Expedite and simplify DDI production and processing
- Optimize interoperability and facilitate document sharing with other DDI users

The guidelines included in this document are intended to assist potential users in building a technically accurate and complete profile that will serve as an effective tool in managing DDI instances as well as data sharing operations.

Status:

This document is updated periodically on no particular schedule. Send comments to editor -- ddi-bp-editors@icpsr.umich.edu



30

31 **Table of Contents**

32

33	INTRODUCTION	3
34	1.1 Problem statement	3
35	1.2 Terminology	3
36	2 BEST PRACTICE SOLUTION	4
37	2.1 Definitions	4
38	2.2 Best Practice behavior	4
39	2.3 Discussion	7
40	2.4 Example	8
41	3 REFERENCES	ERROR! BOOKMARK NOT DEFINED.
42	3.1 Normative	9
43	APPENDIX A. ACKNOWLEDGMENTS	10
44	APPENDIX B. REVISION HISTORY	12
45	APPENDIX C. LEGAL NOTICES	13
46		



47

48 **Introduction**

49 Working with a pre-established subset of DDI tags rather than the entire specification will
50 greatly help organizations in streamlining and increasing the efficiency of internal
51 operations. In DDI 3.0 subsets of used/unused tags may be described using the DDI Profile
52 module.

53 Building the DDI Profile according to the best practice recommendations provided in this
54 document will greatly facilitate document sharing among organizations or communities of
55 users working with different profiles. If profiles are built consistently, following the same
56 guidelines, they will be machine-actionable, which will make it possible to automate their
57 processing as well as the exchange and processing of the actual DDI Instances that they
58 document.

59 **1.1 Problem statement**

60 The DDI 3.0 schema allows a certain degree of flexibility in building Profiles. While some
61 entries are required (for example, the XPath expression for each used and/or unused
62 element or attribute), others are optional and therefore may be missing from a valid Profile.
63 Moreover, a Profile may be constructed listing, for example, only the unused tags, while
64 others may only list the used tags, or both used and unused. Such variations will be difficult
65 to process by systems programmed to expect a particular structure.

66 On the other hand, if organizations or other DDI users create their Profiles using the same
67 sequence of steps, as recommended in this best practice, they will ensure interoperability of
68 their systems based on their Profiles' predictable content.

69 **1.2 Terminology**

70 The key words *must*, *must not*, *required*, *shall*, *shall not*, *should*, *should not*, *recommended*,
71 *may*, and *optional* in this document are to be interpreted as described in **[RFC2119]**.
72 Additional DDI standard terminology and definitions are found in
73 <http://www.ddialliance.org/bp/definitions>

74

75 2 Best Practice Solution

76 2.1 Definitions

77 XPath Syntax:

78 An XML Path Language (XPath) expression uses a path notation, like those used in URLs,
79 for addressing parts of an XML document. The expression is evaluated to yield an object of
80 the node-set, Boolean, number, or string type. For example, the expression book/author will
81 return a node-set of the <author> elements contained in the <book> elements, if such
82 elements are declared in the source XML document. In addition, an XPath expression can
83 have predicates (filter expressions) or function calls. For example, the expression
84 book[@type="Fiction"] refers to the <book> elements whose type attribute is set to "Fiction"
85 (http://www.w3schools.com/xpath/xpath_syntax.asp)
86

87 XML editing software, or XML “editors”:

88 Applications that facilitate the creation of XML documents by providing prompts regarding
89 the appropriate use of tags based on the XML schema which can be pre-loaded into the
90 software. XML editors also validate XML documents and assist in producing valid
91 documents by pointing to existing errors and usually indicating how the errors might be
92 corrected. Examples of commercial XML editors are XMLSpy, oXygen, XMetaL. Free
93 editors are also available. For a more complete discussion, see
94 <http://ahds.ac.uk/creating/information-papers/xml-editors/#section2>
95
96

97 2.2 Best Practice Behavior

98 Determine how you will be creating DDI 3 Instances: “from scratch”, or by conversion from
99 DDI 2.x. [If using both methods, and planning to include more fields in the DDI 3 “from
100 scratch” than in the converted files, use two profiles as a single all inclusive profile might
101 contain mandatory fields that cannot be generated from converting 2.x to 3...]

102 2.2.1 Creating DDI 3 Instances “From Scratch”

103 2.2.1.1 Evaluate the specific needs of your community or organization, the anticipated
104 uses and performance expectations for DDI 3 instances. For example, will you
105 be using DDI 3 just for study descriptions (then only study-level information will
106 be needed), or will you be describing data, or both? Will you document single
107 studies or groups of studies, or both? If documenting groups, will they be formal,
108 or informal, or both? Will you describe microdata, or aggregate data, or both? If



Data Documentation Initiative

- 109 documenting aggregate data, what is its physical representation? Will you be
110 using DDI 3 for data transport? If so, what type of data will you include inline,
111 microdata, aggregate data, or both? Depending on your anticipated uses,
112 different sections/elements of DDI 3 will need to be included in the profile.
- 113 2.2.1.2 Determine whether there are specific standards that your DDI 3 instances need
114 to be compatible with (Dublin Core, SMDX, GIS, etc.)
- 115 2.2.1.3 Examine the DDI 3 specification to select those modules, sections, and
116 elements which are needed to cover your specific needs. Alternatively, examine
117 and edit the DDI 3 Core
118 (http://www.ddialliance.org/ddi3/DDI3_CR3_Core.xml), adding and/or
119 deleting elements to produce a selection of fields that will satisfy your specific
120 needs. (Examples: for microdata transport, add the dataset module; for
121 describing aggregate data, add the nCube sections that are appropriate for your
122 type of data; for describing groups, add the Group module; for DC compatibility,
123 add DC citation; for compliance with SDMX, add the
124 PhysicalDataProduct_ncube_inline module, etc., etc.)
- 125 2.2.1.4 When you have done your selection, proceed to add those elements and/or
126 sections that are mandatory in DDI 3 in order to produce a valid instance
- 127 2.2.1.5 It will also be necessary to include some elements that are not mandatory per
128 se, but for which there are mandatory references in DDI 3 (for example the
129 Universe Scheme is not mandatory, but the Universe reference is mandatory in
130 the Study Unit module); both 2.2.1.4 and 2.2.1.5 are best achieved in
131 conjunction with 2.2.1.6 below.
- 132 2.2.1.6 Using XML editing software (see 2.1), upload the DDI 3.0 Instance schema
133 http://sourceforge.net/projects/ddi-alliance/files/Data%20Documentation%20Initiative/DDI%203.0%20%282008-04-28%29/DDI_3_0_2008-04-28_Documentation_XMLSchema.zip/download
134 and produce a sample
135 DDI 3.0 instance containing your entire selection of fields.
136
- 137 2.2.1.7 Use of an XML editor for task 2.2.1.6 will also assist in completing tasks 2.2.1.4
138 and 2.2.1.5 above, as the software will prompt for mandatory entries.
- 139 2.2.1.8 Validate your sample instance. Make sure to insert appropriate content within
140 those fields that require special/ fixed formats or embedded controlled
141 vocabularies for validation (i.e. ISO-type date, or NCName for IDs or ID
142 references). The free-form fields may contain brief comments on their
143 recommended use and/or type of content expected.



Data Documentation Initiative

- 144 2.2.1.9 The sample instance may serve as a template for future DDI production, and as
145 a guide for proper use of the fields in conjunction with the Profile module.
- 146 2.2.1.10 Evaluate your selection/sample Instance to determine which fields you want to
147 make mandatory in addition to those that are required in DDI 3.
- 148 2.2.1.11 Using an XML editor, upload the schema for a DDI Profile and produce a Profile
149 module, listing all the fields in the sample instance.
- 150 • Each element should be listed within the “Used” tag.
- 151 • Enter all used elements, even if child-elements are assumed to be used; a
152 comprehensive listing will make it possible to include documentation and
153 instructions for each individual element, whereas listing only parent-
154 elements would preclude it.
- 155 • Do not list used attributes if usage is the same as in the DDI 3.0 schema;
156 attributes of used elements will be assumed to be supported. List used
157 attributes only if you institute a change in its usage (mandatory nature or
158 value). List only attributes that are not supported in NotUsed fields.
159 Information on the proper use of attributes (content, format, potential CV,
160 etc.) may be included in the Instructions field for each listed element, if the
161 attributes themselves are not listed.
- 162 • Using XPath expressions (DDI 3.0 supports versions 1.0 or 2.0 of XML Path
163 Language), specify a path for each used element. Example: <Used
164 path="/ns1:DDIInstance/s:StudyUnit/r:FundingInformation/r:AgencyOrganiza
165 tionReference/r:ID">. “Path” is a required attribute.
- 166 • Indicate, as appropriate, the mandatory nature of those fields that are not
167 required in DDI 3 by changing the “required” attribute to “true” (“false” is the
168 default on this mandatory attribute, so no changes need be made if their
169 usage remains the same as in the main DDI 3.0 schema).
- 170 • Using the defaultValue and fixedValue attributes, specify default values (for
171 objects included in the Profile but missing from the DDI Instance(s)) and
172 fixed values as appropriate.
- 173 • Insert DDI-published documentation
174 [http://sourceforge.net/projects/ddi-alliance/files/Data%20Documentation](http://sourceforge.net/projects/ddi-alliance/files/Data%20Documentation%20Initiative/DDI%203.0%20%282008-04-28%29/DDI_3_0_2008-04-28_Documentation_XMLSchema.zip/download)
175 [%20Initiative/DDI%203.0%20%282008-04-28%29/DDI_3_0_2008-04-28](http://sourceforge.net/projects/ddi-alliance/files/Data%20Documentation%20Initiative/DDI%203.0%20%282008-04-28%29/DDI_3_0_2008-04-28_Documentation_XMLSchema.zip/download)
176 [_Documentation_XMLSchema.zip/download](http://sourceforge.net/projects/ddi-alliance/files/Data%20Documentation%20Initiative/DDI%203.0%20%282008-04-28%29/DDI_3_0_2008-04-28_Documentation_XMLSchema.zip/download). for each used



Data Documentation Initiative

177 element in the “Description” field, and your own instructions for proper use in
178 the “Instructions” field.

179 • The Alternate Name element may be used to give a human-readable name
180 to the tag that is being described. This will be useful especially for tags
181 whose names do not provide sufficient information about the content they
182 carry and therefore need to be associated with their parent elements (for
183 example, Name – may be Series Name, Variable Name, Question Name,
184 etc.; Content – may be Abstract Content, or Purpose Content, etc.)

185 • In the NotUsed tags enter only large sections (modules and/or schemes) that
186 your organization has decided not to support (for example, the Group
187 module, or the nCube modules for aggregate data)

188 2.2.1.12 The Profile module will be used by software to validate Instances and manage
189 data sharing.

190 2.2.1.13 Apply XSLT to display the DDI Profile module in viewer friendly format, to serve
191 as a human-readable guide for creating and evaluating DDI Instances.

192 2.2.2 Producing DDI 3 Instances by Conversion from DDI 1-2.x

193 2.2.2.1 List DDI 1-2.x fields used (may be all fields, or DDI Lite, or local selection).

194 2.2.2.2 Map DDI 1-2.x used fields to corresponding DDI 3 fields. If using all fields, or a
195 local selection, use the DDI 2.x to DDI 3 mapping published by the DDI Alliance
196 <http://www.ddialliance.org/sites/default/files/CorrespondenceMapping-spreadsheet.pdf> If using DDI Lite, refer to the DDI 3 Core
197 http://www.ddialliance.org/ddi3/DDI3_CR3_Core.xml which is a mapping of DDI Lite
198 <http://www.ddialliance.org/sites/default/files/ddi-lite.html> to DDI 3 with the additional
199 mandatory fields included.

200 2.2.2.3 Proceed to create a DDI 3 sample Instance and a Profile module as described
201 under 2.2.1.11 to 2.2.1.13 above.

202 2.3 Discussion

203 Choices: Using the Core as a starting point vs. using the entire specification (easier to add
204 to a smaller subset than to select from hundreds of tags); creating a sample Instance first,
205 or in conjunction with the Profile (visualizing an Instance helps assess whether the selection
206 is appropriate, correct, etc.); recommending the listing of Used elements rather than
207 NotUsed (but, listing NotUsed may make more sense in certain cases.); assuming that all
208 attributes are used, if not specified as NotUsed. Before including these recommendations in
209 a Discussion, need to determine whether they are indeed the best choices.

2.4 Example:

Creating the DDI Core Profile

The DDI Core is a list of recommended DDI fields that includes a selection of the most relevant and commonly used elements and attributes, providing a basic framework for creating DDI documentation.

Such a subset of DDI tags was first created using Version 2.1, and published as “DDI Lite”. The latter comprises all of the fields that map to several other standards or commonly used selections: the Dublin Core Metadata Element Set, the CESSDA list of mandatory and recommended elements, the ICPSR study description (catalog record), statistical packages setups (SPSS, Stata, and SAS), and Nesstar’s default template.

Using DDI Lite as the starting point for building the DDI Core thus provides a good way to preserve this basic compatibility with other commonly used standards, or field selections.

To build the DDI Core Profile:

- 2.4.1 All DDI Lite tags (using DDI Version 2.1) were mapped to their corresponding DDI 3.0 fields.
- 2.4.2 Using XMetaL Author 4.6, a sample DDI 3.0 Instance was created that included all of the fields indicated by the mapping described at 2.4.1.
- 2.4.3 Additional tags that had not been included in DDI Lite, but are mandatory in DDI 3.0, were added to the sample Instance, as prompted by XMetaL Author’s validation function (Example: “DataRelationship” sequence in the Logical Product module).
- 2.4.4 Based on the DDI 3.0 in-line documentation, a short description of its prescribed content was inserted within each tag, alongside with summary information regarding its use (mandatory, repeatable/non-repeatable, recommended or embedded Controlled Vocabulary, fixed format, XHTML formatting allowed, etc.).
- 2.4.5 Using XMetaL Author 4.6, a DDI 3.0 Profile module was created, listing all elements used in the sample Instance.
 - The XPath was documented in the “path” attribute.
 - Information about changes in the optional nature of the element was specified in the “required” attribute (example: Geography Name in Top and Lowest Level Geography Reference was made mandatory, while in DDI 3.0 it is only part of a mandatory choice).
 - The in-line documentation for each used element was included in the “Description” sub-element.



Data Documentation Initiative

- Additional instructions for use, if appropriate, were included in the “Instructions” subelement.

This file may be viewed and downloaded at:

http://www.ddialliance.org/sites/default/files/bp/CORE_Profile_FinalDDI3.xml

3.0 Normative

[RFC2119] S. Bradner, Key words for use in RFCs to Indicate Requirement Levels, <http://www.ietf.org/rfc/rfc2119.txt>, IETF RFC 2119, March 1997.

OASIS, Best Practice, <http://www.oasis-open.org/committees/uddi-spec/doc/bp/uddi-spec-tc-bp-template.doc>, 2003



261 **Appendix A. Acknowledgments**

262 The following individuals were members of the DDI Expert Workshop held 10-14 November
263 2008 at Schloss Dagstuhl, Leibniz Center for Informatics, in Wadern, Germany.

264 Nikos Askitas, Institute for the Study of Labor (IZA)

265 Karl Dinkelmann, University of Michigan

266 Michelle Edwards, University of Guelph

267 Janet Eisenhauer, University of Wisconsin

268 Jane Fry, Carleton University

269 Peter Granda, Inter-university Consortium for Political and Social Research (ICPSR)

270 Arofan Gregory, Open Data Foundation

271 Rob Grim, Tilburg University

272 Pascal Heus, Open Data Foundation

273 Maarten Hoogerwerf, Data Archiving and Networked Services (DANS)

274 Chuck Humphrey, University of Alberta

275 Jeremy Iverson, Algenta Technology

276 Jannik Vestergaard Jensen, Danish Data Archive (DDA)

277 Kirstine Kolsrud, Norwegian Social Science Data Services (NSD)

278 Stefan Kramer, Yale University

279 Jenny Linnerud, Statistics Norway

280 Hans Jørgen Marker, Danish Data Archive (DDA)

281 Ken Miller, United Kingdom Data Archive (UKDA)

282 Meinhard Moschner, GESIS - Leibniz Institute for the Social Sciences

283 Ron Nakao, Stanford University

284 Sigbjørn Revheim, Norwegian Social Science Data Services (NSD)



Data Documentation Initiative

- 285 Wendy Thomas, University of Minnesota
- 286 Mary Vardigan, Inter-university Consortium for Political and Social Research (ICPSR)
- 287 Joachim Wackerow, GESIS - Leibniz Institute for the Social Sciences
- 288 Wolfgang Zenk-Möltgen, GESIS - Leibniz Institute for the Social Sciences
- 289
- 290
- 291
- 292
- 293
- 294
- 295



Appendix B. Revision History

Rev	Date	By Whom	What
0.9	2009-02-15	Stefan Kramer	Minor reformatting.

299

300 **Appendix C. Legal Notices**

301 Copyright © DDI Alliance 2009, *All Rights Reserved*

302

303 <http://www.ddialliance.org/>

304

305 Content of this document is licensed under a Creative Commons License:

306 Attribution-Noncommercial-Share Alike 3.0 United States

307

308 This is a human-readable summary of the Legal Code (the full license).

309 <http://creativecommons.org/licenses/by-nc-sa/3.0/us/>

310

311 You are free:

- 312 • to Share - to copy, distribute, display, and perform the work
- 313 • to Remix - to make derivative works

314

315 Under the following conditions:

- 316 • Attribution. You must attribute the work in the manner specified by the author or
317 licensor (but not in any way that suggests that they endorse you or your use of
318 the work).
- 319 • Noncommercial. You may not use this work for commercial purposes.
- 320 • Share Alike. If you alter, transform, or build upon this work, you may distribute
321 the resulting work only under the same or similar license to this one. For any
322 reuse or distribution, you must make clear to others the license terms of this
323 work. The best way to do this is with a link to this Web page.
- 324 • Any of the above conditions can be waived if you get permission from the
325 copyright holder.
- 326 • Apart from the remix rights granted under this license, nothing in this license
327 impairs or restricts the author's moral rights.

328

329 **Disclaimer**

330

331 The Commons Deed is not a license. It is simply a handy reference for understanding the Legal
332 Code (the full license) — it is a human-readable expression of some of its key terms. Think of it as
333 the user-friendly interface to the Legal Code beneath. This Deed itself has no legal value, and its
334 contents do not appear in the actual license.

335

336 Creative Commons is not a law firm and does not provide legal services. Distributing of, displaying
337 of, or linking to this Commons Deed does not create an attorney-client relationship.
338 Your fair use and other rights are in no way affected by the above.

339

340 **Legal Code:**

341 <http://creativecommons.org/licenses/by-nc-sa/3.0/us/legalcode>