

Minutes of the 1st DDI Developers Meeting

Date 6th and 7th December 2010, respectively from 9 am to 5 pm
Location SURFfoundation, Graadt van Roggenweg 340,
Utrecht, The Netherlands

Participants

Jeremy Iverson (Chair)	jeremy@algenta.com
Achim Wackerow	joachim.wackerow@geis.org
Wendy Thomas	wlt@pop.umn.edu
Nikos Askitas	askitas@iza.org
Jannik Vestergaard Jensen	jvj@dda.dk
Dan Kristiansen	dak@dda.dk
Ørnulf Risnes	ornulf.risnes@nsd.uib.no
Samuel Spencer	samuel.spencer@abs.gov.au
Douglas Kieweg	dkieweg@ku.edu
Dan Smith	dan@algenta.com
Johanna Vompras	johanna.vompras@uni-bielefeld.de
Alerk Amin	a.amin@uvt.nl
Andias Wira-Alam	andias.wira-alam@geis.org
Wanja (Johann) Schaible	johann.schaible@geis.org
Thomas Bosch	thomas.bosch@geis.org
Alexander Mühlbauer (Minutes)	alexander.muehlbauer@geis.org

Version history v1.0, 20.12.2010, Mühlbauer, Initial version

1 Short presentations of current works with DDI3

1.1 Nesstar

Ørnulf Risnes gave a short overview of the Nesstar software with its main three parts Publisher, Server and Web. He also showed the ESS Multiwave Download Wizard.

To improve performance every object is divided in a head and a body. Information which belongs to the body of an object is lazy loaded. With indexing id columns in the relational database they experienced big performance gains.

For more information about the inner workings of the Nesstar API, see [NESSTAR: A Semantic Web Application for Statistical Data and Metadata](#) and [NEOOM: A Web and Object Oriented Middleware System](#).

1.2 STARDAT project at GESIS

Wanja Schaible showed to first steps which are being made by the STARDAT project at GESIS. This new project aims to integrate several existing tools for metadata documentation of the GESIS Data Archive. He mentioned the used technologies (GRAILS) and a first draft of architecture for a prototype. He also addressed the challenges how to map DDI3.x XML to relational databases.

1.3 SPSS DDI Write tool

Achim Wackerow demonstrated the SPSS DDI Write tool. He defined own commands - written in Python - which then can be integrated into the SPSS Python Plugin. With a quite intuitive statement the user inputs a SAV document and gets a DDI3 XML file with metadata as output. His next objective is to improve the stability of the tool.

1.4 Exanda

Achim Wackerow also showed his tool Exanda for online analysis and how users can select tables and charts on variables. He gave a short technical overview, showed how to generate PDFs out of SVGs.

The visualization is made by Adobe Flex which is open source for a short time. He prefers to use it because a huge majority of browser supports it. Finally he went into the import and export with DDI3.

1.5 Project of DSZ-BO, University of Bielefeld

Johanna Vompras introduced a their approach of DDI 3 storage: The idea is not to map xml contents against relational databases but to store contents in a native XML database (BaseX) and do the mappings with stored XQuery statements to retrieve wanted information. Performance is not considered a problem because indexing is possible and data volume is supposed to be not too big. In that storage and processing are loosely coupled by the usage of XQuery, there is great flexibility in the conjunction with DDI 3 version evolution. But certainly, XPath statements needs to be updated.

1.6 DDA DDI 3 Editor

Jannik Jensen showed a new feature of the DDA DDI3 editor: It is enriched by a wiki syntax which makes data entry more easy.

There is also a SPSS import with overwrite functionality. For this import DEX T SPSS Reader, an open-source software by UKDA and Open Data Foundation, some changes at ID generating and validating are made; these modifications will be published.

In this context of metadata import from statistical data files, it was stated that SPSS does not support the lifecycle. Nowadays, these files are still taken as reference in workflows. But in the future, DDI3 based tools should and are able to be the reference for workflows which include the data lifecycle. Statistical data files are then only output formats.

Jannik Jensen also shortly addressed the usage of DDI Light to cope with performance bottlenecks which he had faced. DDI Light is a schema based subset of DDI3 with only Ids and labels of some items to accelerate data access.

1.7 CodePlay

Samuel Spencer presented CodePlay which is a competition run by the Australian Bureau of Statistics as a Government 2.0 initiative to help drive collaboration between students, developers and national and international statistical agencies. CodePlay aims to reward innovative ways of making statistical data more appealing, relevant and useful. In this context, he emphasized the relevance of DDI for his affiliation and interest in collaboration. For more information, see [here](#).

1.8 Work of Douglas Kieweg, University of Kansas

Douglas Kieweg has started to evaluate DDI, is interested in the big picture of DDI and searches for advice how to work with it. He would like to reuse libraries as much as possible because he has no time for own DDI developments. He prefers as thin as possible clients in a SOA environment.

1.9 Questasy and LISS Panel

Alerk Amin gave a short introduction in the tool Questasy - developed by CentERdata - and its application in the LISS Panel project.

Questasy is an Apache/PHP/MySQL web application which is based on DDI3 with easy access for researchers without DDI3 knowledge. To get metadata information no authentication is needed; to query statements on data, users need to register and log in. The backend is a relational database which is quite fast. The structure is made by hand and does not implement the entire standard, only a needed subset is supported.

2 DDI 3.2 Updates & TIC Overview

The slides to this topic can be found [here](#).

2.1 Schedule and Process Changes

Questions about the announced summer/winter schedule arose: Is it really necessary to release as often? Every new version takes time to be understood and adopted.

2.2 DDI 2.5

This version aims to capture a number of elements found in a simple study in 3.0 and to provide a means of capturing information that will aid in round-trip translation between DDI 2 and DDI 3 and make use of controlled vocabularies. DDI 2.5 will be released soon with a review version. There is a short time for feedback. Please check own assets in DDI 2.x.

2.3 Proposed changes for ID and URN in DDI 3.2

2.4 New Types: MissingValues, DataElement

3 DDI Sticky Points and Coping Strategies

3.1 Handling DDI version changes

Changes in namespaces entail changes of XQuery statements, not only in structure but also in hierarchy. Perhaps namespace managers can help in some cases. But they are not assumed to solve every problem in this context.

Whether XSL transformations between versions are a possible approach to solve the problem, stayed unanswered.

A recommendation to TIC is to work more model-based and not to think about where to put what information into the XML file.

3.2 Record Relationships

3.3 Infinite Nesting Issues

MultiQuestionItem is used differently to the documentation.

3.4 References and Resolution

How do applications handle references which are not in the loaded DDIInstance. The idea to resolve all references into the XML file was supposed to not be useful because there are also references which points to references and so on. It makes sense to restrict reference resolution. But to what depth?

3.5 Extensions: Notes, schema extensions

Jeremy demonstrated his kind of extension, he uses the note element.

User defined namespaces are possible. New elements can be defined with XML schema. The big pro is that it is easy to integrate the new ideas into the standard. The con is schema definition is difficult if you only want to add some elements.

Extensions are problematic because “flavours of DDI” are very probable. The misuse of tags like it happened in DDI2 must be avoided, extensions should be only exceptions.

3.6 Multilingual issues

- Two or three letter country code
Achim proposed ISO639, Dan Smith suggested <http://www.rfc-editor.org/rfc/bcp/bcp47.txt>
- Object model for multilingual text
Alerk Amin and Ørnulf Risnes showed how they integrated multilingual text support in their existing systems with some limitations and work arounds. For the future, it is intended to strictly normalize structures. One challenge is to always identify the default language.
- Ambiguities in International Strings
It is not clear how to handle texts in different languages. One possible convention could be that only one string for one language is allowed. Perhaps it makes sense to add the contributor. A final solution is not found. Validation should point to this problem.
- IdentifiedStructureString
How to cope with? What is in the future?
- ControlConstruction expressions
How to identify variables inside code snippets?

4 Framework for Ensuring DDI 3 Interoperability

4.1 Ambiguity in URN Version Numbers

Are version numbers in URNs numeric or non-numeric values? How can I find a certain or the last version? It depends on individual interpretation:

Numeric	1, 2, 10, 11, 12
String	1, 10, 11, 12, 2

One idea to solve this problem is to make versioning in a subversion style: Every change leads to an increment of one (!) numeric digit with the possibility to label this specific revision. This could be the the DDI 3 version or some other text.

TIC is asked to resume the discussion and specify the appropriate usage.

4.2 Proposal on DDI Query Protocol

Achim gave an introduction how such a DDI Query Protocol could look like and what features it should have. Two topics were discussed:

- What to do with cascading references?
 - Is there a timeout at the requester, at the responder or even at both?
 - Is it necessary to distinguish between external and internal references? The server delivers only skeletons for external references which clients have to resolve.
- What do I actually get back?
 - A “Transport element” which states “this is DDI of that version” (see <http://mantis.ddialliance.org/view.php?id=434>)
 - An elegant solution would be possible with RDF.
 - Nothing special needed: We have URNs for objects, the client knows what is asking for and that’s it what it expects.
 - A server should be able to say which DDI version it supports and which not (harvesting style).

5 Technical Infrastructure for DDI: Storage and Web Service Strategies

A vital discussion about storage led to these results:

- Mapping DDI3 to relational databases can become quite complex, native XML database should be cleverly used.
- It is recommendable to use text search engines like Lucene for indexing and searching. Even Nesstar with a relational database delegates this requirement to Lucene.
- Some made good experiences with indexed XML files, native XML database Berkley DB and BaseX scale satisfying.
- The applicable solution may depend on different requirements like preservation, storage and performance.
- From the user’s perspective only the service is relevant for communication, not the underlying technique.

To reach interoperability these conclusions are made

- If a provider does not support a certain tag, but consumer expects it, the loss of data needs to be avoided. But how?
- DDI profiles, which use XPath, are the recommended configuration mechanism to define used or not used ddi elements in a machine actionable way.
- The discussion whether XPath is expressive enough to supply all wants of configuration concluded with that it is. There may be some issues with cardinalities, but seem to be also very expressive with nesting issues.
- A recommendation to TIC is to develop “branded profiles” which developers can use.
- Samuel Spencer agrees to provide a platform like a wiki so that the DDI developers’ community can intensify its collaboration. One first objective is that developers start sharing DDI 3 XML files, test interoperability, give feedback and perhaps find best practices for the usage of DDI profiles.

6 Advocating DDI Tools and Development

Brainstorming lead to these ideas:

- support of "non-survey data"
- simplify management of DDI 3 versioning
- Decoupling of identity and model: Composition of identifiers is made with ID and parent IDs of an object. So, these relations are embedded in the identifier. The proposal is not to tie identity to the model.

- Questionnaire creation not only printed but also in forms. There are possibilities to store design information in DDI3.
- See academic researchers' proposals working with DDI3. To win them there should be resource packages suggested by applications.
- Generally usable Java plug-in library or web service for second level validation with URN checking. Perhaps Colectica makes its validation accessible via a command line tool.
- A tool to write DDI profiles: It shows a big tree with what's all possible, so that the user just clicks on the items to create a profile.
- Visualization tool of flowchart of the questions, the main sequence and branches of it. A former version of Colectica had such a zoom feature, which was quite comfortable. Also a print out of these visualizations would be good.

7 Licensing

LGPL seems to be the most favored one. One problem with GPL is that if you use software which is published under GPL you also have to publish under GPL. Some stated that the licensing also depends on the funding agency. Final question was, whether it is possible to agree on one license model with which most willingly work.

8 Future of the DDI Developers Meeting

All participants agree that a meeting like this is very useful. They hope that it will continue.

- In the context of funding there is consensus among all that two days are neither too short nor too long.
- For non-European attendants certainly traveling is a challenge. Thus it would be good on the one hand to combine the meeting with another conference. On the other hand conferences like IASSIST are already surrounded by many other meetings, so to stick the dev meeting to at least IASSIST is not preferred. One idea is to meet after (!) the EDDI conference so that developers can pick up the presentations and discussions there.
- Achim Wackerow makes a proposal to Mary Vardigan how this developers meeting could be mentioned on the DDI homepage so that new developers can join us.
- Although there has yet been low traffic following resources are recommended to be used:
<http://groups.google.com/group/ddi-developers> and
<http://www.ddialliance.org/community/listserv>
- Conference calls are possible. Please ask Achim Wackerow.

Minutes of the 2nd DDI Developers Meeting

Date 31st May 2011, 9 am to 5 pm
Location 470 Hamber Boardroom, Morris J. Wosk Centre for Dialogue,
580 W Hastings St, Vancouver

Participants

Alerk Amin (chair)	a.amin@uvt.nl
Ingo Barkow	barkow@dipf.de
Jeremy Williams	jw568@cornell.edu
Olof Olsson	olof.olsson@snd.gu.se
Johan Fihn	johan.fihn@snd.gu.se
Leif-Jöran Olsson	ljo@exist-db.org
Benjamin Clark	work@npongo.com
Achim Wackerow	joachim.wackerow@gesis.org
John Shepherdson	jwshep@essex.ac.uk
Oliver Hopt	oliver.hopt@gesis.org
Wanja (Johann) Schaible	johann.schaible@gesis.org
Thomas Bosch	thomas.bosch@gesis.org
Adam Brown	Statistics New Zealand
Bradley Westbrook	bdwestbrook@ucsd.edu
J Gager	j.b.gager@gmail.com
Jeremy Iverson	jeremy@algenta.com
Dan Smith	dan@algenta.com
Stefan Kramer	sk2349@cornell.edu

1 Topics

1.1 DDI 3.2 Identification

1.2 Subversion for DDI3.2 schema online

The DDI3.2 schema is available through a subversion server. Jeremy Iverson can be contacted for more information and account to the server.

Discussion whether the subversion server should be publicly available or only available to the developers group. Not publicly available seems to be the most favored option as it otherwise could be a source of “bad” DDI. Making it available under a dev namespace is not an option as there’s a TIC decision to have a constant namespace.

1.3 Implementing DDI

XML/RDB breakout sessions

1.3.1 RDB

Discussions about versioning.

Discussions about importing DDI3. Usage of DDIProfiles should make implementing importing easier. What to do with metadata you can't import? Using DDI as an intermediary standard to bring things together from different sources can be an only option as everyone is using different RDBs.

Problems with repeatable elements. Database structure explodes causing performance issues and deep subselects. Can be solved with e.g. stored procedures or replacing queries with code.

Mapping of different versions of DDI needed.

1.3.2 XML

1.4 Metadata integration/discovery

Discussions on how to get metadata from different standards searchable in a common portal.

Some suggestions on how to do this was mentioned

- Overall approach in MODS with cv/keywords for the specified subject
- Extension schemas in DDI like in METS. Support for extension schemas is built into DDI.
- DDI Profiles
- RDF
- Sharing and harvesting data with e.g. Google. This could be difficult as Google doesn't support Dublin Core, but is developing their own metadata format, DSPL - Dataset Publishing Language

1.5 DDI & Search

1.6 Multilinguality

1.6.1 Standalone DDI files

The use of Genericcode and CodeValueType for controlled vocabularies makes it impossible to have stand-alone multi-lingual DDI files. Achim will discuss a possible change in CodeValueType element with TIC and CV group to allow for inclusion of translations in the CodeValueType element.

1.6.2 Versioning of controlled vocabularies

A discussion about versioning of controlled vocabularies arose. Leif-Jöran stated that there shouldn't be a need to have a three-level versioning of vocabularies.

- Major version = change in vocabulary contents
- Minor version = typo etc.

CV group needs a policy regarding new translations and versioning. Can you add a translation of a controlled vocabulary without changing the version? What happens if you are in need of a translation and are on a tight deadline. Publish an unofficial translation? How long to get a translation approved by CV group?

1.7 Interoperability between applications

1.7.1 DDI Developers extensions workspace

If we're working on a problem it would be useful if everyone works on it in the same way. An xml schema with the extensions would be the right approach according to Achim Wackerow. Alerk Amin pointed out that it isn't possible to store extensions in a RDB. A solution to this could be to use notes instead, but these shouldn't be used to store metadata.

1.7.2 Best way to use/define profiles

GESIS has just started collection profiles from users to see how they are used. Suggestion to have each application declare a DDIProfile for input(import)/output(export).

1.7.3 Secondary validation

Colectica provides validation tools for DDI. These perform a secondary validation other than schema validation. The secondary validation includes ambiguous references and missing versions. A commandline tool for batch validation will soon be added to the Colectica web site.

1.7.4 REST protocol

Achim demonstrated a REST Service for a DDI repository. The initial structure will have a simple core and will be published by the DDI Alliance. Local services will be able to promote use cases for possible inclusions. Proposed structure is `server/DDIURN/{ddiurn}`.

Olof stated that there's a problem with the structure as javascript can't send header parameters. These should therefore be changed into query parameters or as both header and query parameters and use get as override.

1.8 DDI Web Services - OAI-PMH

Decision to include version in metadata prefix. Not to include version would create problem if you're providing several schema versions.

2 DDI Tools Catalogue

Stefan Kramer gave a presentation of the new DDI Tools Catalogue. The new catalogue lets users submit and change info for their tools. Stefan requested suggestions and comments on the new catalogue. One suggestion was to include the possibility to comment on tools.

3 DDI Developer year-round

3.1 Google Group and promotion

The usage of the Google Group (ddi-developers) has been poor. Other options were discussed as regionalized developers groups and a searchable email list, but a decision was made to continue with the Google Group and promote it to increase usage.

Alerk will write a few paragraphs for the DDI News Letter.

3.2 Relationship between DDI Developers & DDI Alliance

Should the DDI Developers be a subcommittee of the DDI Alliance? Being a subcommittee would add weight to decisions but create expectations to produce between meetings. Several opinions were raised expressing support for staying separate from the alliance, since we represent most of the adopters of the standard.

Achim will ask the alliance about their view of the DDI Devs and what's desired from their point of view.

More feedback from and to the developers are needed from TIC.

- More feedback on bug fixes before deadlines
- Increase the visibility of the Mantis bug system
- Interest groups for different topics

3.3 Minutes

Decision was made that the minutes of the meetings should be published on the Google Group. A link to the minutes will be published on the DDI Alliance web site.

4 Plans for next meeting - Gothenburg

Next meeting is to be held in connection to EDDI 2011 in Gothenburg. The conference is to be held 5-6 December and planned dates for the developers meeting is 7-9 December.

A discussion regarding the structure of the next meeting led to the following suggestions

- one day for specification issues
- one day for development issues
- more use cases
- a coding day with testing of interoperability between tools and fixing
- alternatively a half day for use cases and a half day for coding/testing

Organizing team: Jeremy(chair), Alerk and Olof.

Minutes of the 3rd DDI Developers Meeting

Date 7th to 9th of December 2011
Location Swedish National Data Service, Gothenburg, Sweden

Participants

Jeremy Iverson (chair)	jeremy@colectica.com
John Payne	jmpayne@essex.ac.uk
Marcel Hebing	marcel.hebing@googlemail.com
Dan Smith	dan@colectica.com
Edwin De Vet	e.c.j.m.devet@uvvt.nl
Thorsten Busert	busert@dipf.de
Ingo Barkow	barkow@dipf.de
Jeannette Chin	jschin@essex.ac.uk
Randy Banks	randy@essex.ac.uk
Jason Ostergren	jostergr@isr.umich.edu
Peter Sparks	zebulon@isr.umich.edu
Johan Fihn	johan.fihn@snd.gu.se
Leif-Jöran Olsson	ljo@exist-db.org
Olof Olsson	olof.olsson@snd.gu.se
Samuel Spencer	theodore.therone@gmail.com
Jannik Jensen	jvj@dda.dk
Alexander Mühlbauer	alexander.muehlbauer@geis.org
Joachim Wackerow	joachim.wackerow@geis.org

DDI Developers Mailing List: <http://groups.google.com/group/ddi-developers>
DDI Users Mailing List: <http://www.icpsr.umich.edu/mailman/listinfo/ddi-users>

1 DDI Tools Demonstrations

1.1 Colectica Demonstration - Jeremy Iverson

- Jeremy discussing DDI2 import issues
- Discussing how to manage sharing 1 element in a small way without having to wrap objects in schemes in instances, etc... (Issue 434)
 - <http://mantis.ddialliance.org/view.php?id=434>
- What is a popular fragment element name to hold individual objects (DDIFragment?)
- See Appendix 1 for a possible solution
- In 3.2 Citation and Purpose are not going to be required anymore.

- Notes
- Remodels everything as items for the rdb (postgresql or mssql currently) with additional extracted text fields for searching.
- Documentation vs transport exchange

1.2 Dataset Documentation Manager - Alexander Mühlbauer

- Ingestion of questionnaires is what the tool does.
- They want to have smooth interchange
- Access databases
- Problems with persistent identifiers. Don't let the users change them... or make the algorithm clear to everyone.
- Exports to DDI 2

1.3 Coffee Break

1.4 Virgil and Code Scheme viewer - Sam Spencer

<http://code.google.com/p/virgil-ui>

<http://code.google.com/p/ddi-codescheme-viewer>

Virgil c2d - CSV to DDI converter

Language issues

- isTranslatable, translated attributes are useful
- being able to specify the original text would be useful too
- there is a DDI language working group, hopefully with progress on this in 3.2

Using uuid for ID seems to be the consensus?

isPublished would be good to implement to support publishing and finalisation of classifications
versioning has not been in focus for this 2*6 pack project

1.5 Full text indexing and xForms in eXist-db - Leif-Jöran Olsson

<http://exist-db.org>

<http://ddi-exist.info>

also: <http://code.google.com/p/ddi-exist/>

<http://code.google.com/p/ddixslt/>

- Lucene engine
- Search specific text fields, languages; use lucene operators (AND, OR, etc)

- xForms shows questionnaire. User can submit response, could add a handler to store data

SND uses eXist-db and builds web services on top. Transformations and search. xForms is the next step.

1.6 DdiEditor - Jannik Jensen

DDA uses the DdiEditor in production

<http://code.google.com/p/ddieditor/>

- Import questionnaires from wiki-like syntax
 - This also allows the questionnaire to easily be transformed into a nice HTML view
- Variable browser and editor
- Curation: quickly assign questions and concepts to variables
- Sequence designer and visualization
- DeXT enhanced to import SPSS into ddi-3.1

1.7 DDI, RDF and SPARQL - Dan Smith

Run down of the information available at: <http://dan.smith.name/2011/10/ddi-3-meets-rdf-and-sparql-with-colectica-repository/>

1.8 MQDS - Peter Sparks and Jason Ostergren, ISR, Michigan

- Blaise to DDI
- Health and Retirement Study (HRS)
- Custom XML format, but can convert it to DDI
- Relational database model corresponding to DDI elements
- Gotos in the logical flow of an instrument
 - Code them with if-then-else structures

2 Topic ideas

- 3.1 to 3.2
- IASSIST Presentations (due Friday)
- DDI to/from DataCite
- URN resolution, agency registry
- Version management
- SPSS: import and export
- Researchers allegedly don't want DDI. What to do about that?

Topic - Open Discussion on Users and Improving use of the standard
Simple challenges to solve using DDI

- DDI is great for archives, but not so much for researchers
 - Wrappers / fragments / etc...
 - An entry point for limiting parts of the DDI
 - Easier to start out than learning about all the maintainables and such
- DDI Developers Community Needs a better way to share tools and information

3 Tools Catalog - Jannik Jensen

- It's available as a sort of a beta version
- Needs to be some more activity
- Ideas
 - Grouping or filtering tools by type would be useful
 - Show changelog for updated tools
 - RSS or email subscriptions for updates
 - RSS already exists (or was just added in the last hour?)
- Misc DDI site thoughts
 - Lots of information is spread out on different sites
 - More examples would be good. And should be coming.

4 DDI Examples

The examples of ddi in two directions

- Training material, examples of context eg QuestionScheme
 -
- Full instances representing the output of the tools placed in eq in a subversion repos at Assembla
- Small examples, just XML fragments would be useful
 - Sam is willing to help out
- documentation as php, on ddi field level documentation, eg.: http://php.net/str_shuffle
 - to be able to facilitate discussion around a topic
 - Achim will join us at 16.00 for an outline of the future documentation strategy of DDI-TIC

5 Evening Program

Pool+beer at Biljardpalatset, Ekelundsgatan 9-11

6 Plan for Thursday

1. DDI Developments: What's coming in 3.2 and Keeping up with TIC
2. IASSIST planning

3. Topics from list above
 - a. Breakout groups?

7 Tools RSS

<http://www.ddialliance.org/ddi-tools.xml>

8 DDI 3.2

- Svn repo assembla.com (ddi-tic-temp) Send mail to Jeremy for access
- DDI 32-dev namespace for testing
- issue management: mantis.ddialliance.org

Question about MultipleQuestionItem status from Jannik

Longwinding discussion about grid or not in questionnaire creation

Unresolved issues for 3.2?

- PIDs / Identity model
- RDF support

80 should be fixed until February. Could be doable.

Jannik's incentive for the question: "will be fixed" / "won't be fixed" for voting up or down.

Voiting on topic list for today's agenda. 12: Rdb/xml db, 9: use cases, 8: iassist presentations, ddi and rdf, web services, 7: urn resolution, 6: versioning items, 5: ddi to and from datacite (and other formats), 4: spss import, @action

Cutoff at 5 votes?

Variables etc internationalization

Bug number #

QuestionText

9 Agency Registry

Who is producing this metadata is shown with agencyid.

registry.ddialliance.org

Approval process is not processing the pending applications, so a lot of applications in the pipe.

resolver.colectica.com is still up, doing in principle what the registry will do.

DDI Developers REST Best Practise Group

10 Single Namespace Discussion

Pros: Useful for selection, limits conflicting elements

Cons: Elements are being made unique in DDI3.2, prettier, easier to work with

Motion: Limit or reduce the number of namespaces
For: 10
Against: 6
Abstain: 1

11 Talk about paper on DDI and RDBs - Ingo Barkow

12 SND - Leif-Jöran Olsson and Johan Fihn

Demo of eXist, se: <http://code.google.com/p/ddi-exist/>
<http://snd.gu.se/ddi/validator> - online validator

SIMS - archival information system

Keywords:

- ELSST - <http://elsst.esds.ac.uk/>

13 Evening Programme

City Pub. Map available as handout.

14 Ideas for tools

14.1 Single serve software - Sam Spencer

Task-oriented software.

Users don't want to see ddi. Create small apps to do a specific task.

Smaller apps which uses a small part of ddi produces a smaller risk as it affects a smaller part of the standard. If a part of the standard not used by the app is changed it would not be affected.

Starting a draft for a rest interface where you specify a xpath search expression for retrieval.

14.2 Use Cases

Researchers need to find out about DDI.

Researchers use a lot of different tools for collecting data. It's easy for an archive to get questions, variables and that kind of documentation. Getting more detailed information is much more difficult as the number of tools for this is limited.

No matter what tools that exist, users won't use them if they don't see the benefits.

15 TIC-Interaction Achim

15.1 DDI Documentation

Purpose: Provide better documentation and taking some of the work of TIC by letter the community build up documentation that can be moderated into the official documentation.

The DDI-Lifecycle Guide I and part II is part of the official documentation of the standard with a separate section for field level documentation. The official documentation shall be separated from a community driven effort.

15.1.1 Requirements

Official documentation:

The documentation could be on a web page that can be indexed by eg. Google

Community driven documentation/ conversation:

- Uses cases
 - search the documentation
 - linking to official documentation/ embed examples snippets
 - best practice based on elements
 - discussion feature
 - place to upload full DDI instances as examples and to be able to test DDI output by other tools and applications
 - DDI glossary eg. description of a code, scheme etc

Examples of implementations:

- Impl 1: wiki book, <http://en.wikibooks.org/>
 - There are a lot of other XML books in wiki book form Eg.: <http://en.wikibooks.org/wiki/XForms>
- Impl 2: PHP like function discussion, http://php.net/sub_str

Examples of source code in different languages for certain elements or use cases?

15.1.2 Challenges

- Linking official documentation together User Guide I and II and the field level documentation and providing paragraph and field level links to be used by community discussion
- The more fleshed examples are to 'grow' into the official documentation
- Moderator feature. Who would volunteer to do that work?
- Software management/ administration
- It is important that the community then contributes to this collaborative environment
- Create list of DDI use cases to be exemplified

- study level representation
- ddi instance
- group
- resource package
- general language usage
- note usage
- concept
- question
- universe
- variable
- ...
- Choice of platform for the alliance for development work, bug tracking, collaboration, discussion and documentation
 - Jira, <http://www.atlassian.com/software/jira/overview>
 - offering open source development licensing
 - Redmine, <http://www.redmine.org/>
 - Others ?

15.2 REST service

Reuse as many existing HTTP headers

- http://en.wikipedia.org/wiki/List_of_HTTP_header_fields
- <http://www.w3.org/Protocols/rfc2616/rfc2616-sec14.html>

Default format application/vnd.ddi+xml

Default language, en - possible investigation in Accept-language header field

- In DDI -if language is not specified, language is en-us defined by the xml-1.0 specification

Only use the wrapper if ddi xml is returned

ResolutionDepth

- Every time there is a reference the reference counter goes up. Eg. a variable scheme has references to variables [Count up] The resolved variables has references to concepts [Count up] - Equals resolution depth of 2
- ResolutionDepth defines a restricted depth-first search within the reference 'tree'
- internal references in the object is not resolved

15.3 DDI 3.1 -> 3.2

Will the alliance provide a style sheet to do the transformation between the versions? If not, will a change record will be provided?

A change record like the one provided for the 3.0 to 3.1 transition but with XPath or any other similar solution for changed elements would be preferred. Also HTML with links to the field level documentation and the bug tracker would be preferred over PDF as a format for the change record.

16 DDI and DataCite

SND plans to produce a style sheet to transform DDI to DataCite xml. There are some issues regarding what publication date to use and also no way to have an organisation for a creator.

17 IASSIST Papers

There are a number of papers that have been/will be submitted by the group:

- Ingo and Alerk et.al. - Complete session about variable baskets and solutions for this
- Johan and Jeremy - Integrating DDI3 based tools using web services
- Olof and Arofan et.al. - Trouble with triples (DDI-RDF)
- Sam - Dynamic XForms DDI Transform, including standard format for Control Construct conditionals
- Jannik - DDI-Lifecycle Infrastructure at DDA, Future Roadmap for DDI-Lifecycle at DDA
- Olof and Leif-Jöran - Further developments with eXist-db (possibly a poster session)
- Johan and Olof - Creating a DDI3 variable/question/study bank - Pitfalls and Possibilities

Posters

- Sam - DDI Quick Guide using the US Census

18 DDI Developers Year Round

18.1 Google Group

Sam will post updates on questions, tool development.

We encourage people to post updates from blogs, relevant news and issues.

18.2 Minutes

Decision to promote the use of Google Docs for future meetings, and for links to meetings to be mailed to the Google Groups.

19 Plans for next meeting - When, where, who?

Where: IASSIST 2012 Washington

When: June, Parallel to IASSIST Workshops

Organisers: Sam Spencer, Johan Fihn, Jeremy Iverson

Where: RC33 - Sydney

When: July (If there is a large group unable to attend IASSIST)

Organisers: Sam Spencer, Ingo Barkow, Joachim Wackerow

Why: To introduce Australian developers to the DDI Community.

Appendix 1: DDI Fragment Element Candidate

Addressing the transfer of DDI-L of several items of same or different types. As of ddi-3.1 eg. two requested concepts that reside within a concept scheme with many other concepts are served as a concept scheme made up for the purpose of producing valid DDI-L. Corresponding to bug-434 on mantis.ddialliance.org

The solution is defined as a wrapper/ DDIFragment with the capability to hold zero to many elements in any order of type:

- Note
- OtherMaterial
- Versionable
- Maintainable

It is suggested that the wrapper/ DDIFragment is to reside in the instance.xsd schema. Below is a working implementation -suggested enhancements add content type description.

Intended structure pseudo-sample:

```
<ddi:Fragment>
  <ddi:any-versionable/>
  <Notes/>
  <OtherMaterials/>
</ddi:Fragment>
```

Idea - The DDI transport fragment

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" targetNamespace="ddi:fragment:3_1"
  xmlns="ddi:fragment:3_1" xmlns:a="ddi:archive:3_1" xmlns:cm="ddi:comparative:3_1"
  xmlns:c="ddi:conceptualcomponent:3_1" xmlns:d="ddi:datacollection:3_1"
  xmlns:ds="ddi:dataset:3_1" xmlns:dc="ddi:dcelements:3_1" xmlns:pr="ddi:ddiprofile:3_1"
  xmlns:g="ddi:group:3_1" xmlns:m3="ddi:physicaldataproduct_ncube_inline:3_1"
  xmlns:l="ddi:logicalproduct:3_1" xmlns:m1="ddi:physicaldataproduct_ncube_normal:3_1"
  xmlns:p="ddi:physicaldataproduct:3_1" xmlns:pi="ddi:physicalinstance:3_1"
  xmlns:r="ddi:reusable:3_1" xmlns:s="ddi:studyunit:3_1"
  xmlns:m2="ddi:physicaldataproduct_ncube_tabular:3_1"
  xmlns:m4="ddi:physicaldataproduct_proprietary:3_1" xmlns:co="ddi:comparative:3_1"
  xmlns:ddi="ddi:instance:3_1" xmlns:ddip="ddi:ddiprofile:3_1" elementFormDefault="qualified"
  attributeFormDefault="unqualified">

  <xs:import namespace="http://www.w3.org/XML/1998/namespace" schemaLocation="xml.xsd"/>
  <xs:import namespace="ddi:archive:3_1" schemaLocation="archive.xsd"/>
  <xs:import namespace="ddi:comparative:3_1" schemaLocation="comparative.xsd"/>
  <xs:import namespace="ddi:conceptualcomponent:3_1" schemaLocation="conceptualcomponent.xsd"/>
  <xs:import namespace="ddi:datacollection:3_1" schemaLocation="datacollection.xsd"/>
  <xs:import namespace="ddi:dataset:3_1" schemaLocation="dataset.xsd"/>
  <xs:import namespace="ddi:dcelements:3_1" schemaLocation="dcelements.xsd"/>
  <xs:import namespace="ddi:ddiprofile:3_1" schemaLocation="ddiprofile.xsd"/>
  <xs:import namespace="ddi:group:3_1" schemaLocation="group.xsd"/>
  <xs:import namespace="ddi:logicalproduct:3_1" schemaLocation="logicalproduct.xsd"/>
  <xs:import namespace="ddi:physicaldataproduct:3_1" schemaLocation="physicaldataproduct.xsd"/>
  <xs:import namespace="ddi:physicaldataproduct_ncube_inline:3_1"
    schemaLocation="physicaldataproduct_ncube_inline.xsd"/>
  <xs:import namespace="ddi:physicaldataproduct_ncube_normal:3_1"
    schemaLocation="physicaldataproduct_ncube_normal.xsd"/>
  <xs:import namespace="ddi:physicaldataproduct_ncube_tabular:3_1"
    schemaLocation="physicaldataproduct_ncube_tabular.xsd"/>
```

```

schemaLocation="physicaldataprodut_ncube_tabular.xsd"/>
<xs:import namespace="ddi:physicaldataprodut_proprietary:3_1"
schemaLocation="physicaldataprodut_proprietary.xsd"/>
<xs:import namespace="ddi:physicalinstance:3_1" schemaLocation="physicalinstance.xsd"/>
<xs:import namespace="ddi:reusable:3_1" schemaLocation="reusable.xsd"/>
<xs:import namespace="ddi:studyunit:3_1" schemaLocation="studyunit.xsd"/>
<xs:import namespace="ddi:instance:3_1" schemaLocation="instance.xsd"/>

<xs:complexType name="DDIFragmentType">
  <xs:choice minOccurs="0" maxOccurs="unbounded">
    <xs:element ref="r:Note"/>
    <xs:element ref="r:OtherMaterial"/>

    <!-- versionable -->
    <xs:element ref="l:Category"/>
    <xs:element ref="l:CategoryGroup"/>
    <xs:element ref="co:CategoryMap"/>
    <xs:element ref="co:CodeMap"/>
    <xs:element ref="d:ComputationItem"/>
    <xs:element ref="c:Concept"/>
    <xs:element ref="c:ConceptGroup"/>
    <xs:element ref="co:ConceptMap"/>
    <xs:element ref="c:DataElementConcept"/>
    <xs:element ref="l:DataRelationship"/>
    <xs:element ref="ds:DataSet"/>
    <xs:element ref="r:GeographicLocation"/>
    <xs:element ref="r:GeographicStructure"/>
    <xs:element ref="d:IfThenElse"/>
    <xs:element ref="a:Individual"/>
    <xs:element ref="d:Instruction"/>
    <xs:element ref="d:Loop"/>
    <xs:element ref="d:Methodology"/>
    <xs:element ref="d:MultipleQuestionItem"/>
    <!--xs:element ref="NCube"/>
    <xs:element ref="NCubeGroup"/>
    <xs:element ref="NCubeInstance"/>
    <xs:element ref="NCubeInstance"/>
    <xs:element ref="NCubeInstance"/-->
    <xs:element ref="a:Organization"/>
    <xs:element ref="p:PhysicalStructure"/>
    <xs:element ref="m4:ProprietaryRecordLayout"/>
    <xs:element ref="d:QuestionConstruct"/>
    <xs:element ref="d:QuestionItem"/>
    <xs:element ref="co:QuestionMap"/>
    <xs:element ref="p:RecordLayout"/>
    <xs:element ref="d:RepeatUntil"/>
    <xs:element ref="d:RepeatWhile"/>
    <xs:element ref="d:Sequence"/>
    <xs:element ref="d:StatementItem"/>
    <xs:element ref="g:SubGroup"/>
    <xs:element ref="c:SubUniverse"/>
    <xs:element ref="c:Universe"/>
    <xs:element ref="co:UniverseMap"/>
    <xs:element ref="l:Variable"/>
    <xs:element ref="l:VariableGroup"/>
    <xs:element ref="co:VariableMap"/>

    <!-- maintainable -->
    <xs:element ref="c:ConceptScheme"/>
    <xs:element ref="c:ConceptualComponent"/>

```

```
<xs:element ref="d:ControlConstructScheme"/>
<xs:element ref="d:DataCollection"/>
<xs:element ref="ddi:DDIInstance"/>
<xs:element ref="ddip:DDIProfile"/>
<xs:element ref="c:GeographicLocationScheme"/>
<xs:element ref="c:GeographicStructureScheme"/>
<xs:element ref="g:Group"/>
<xs:element ref="d:Instrument"/>
<xs:element ref="d:InterviewerInstructionScheme"/>
<xs:element ref="g:LocalAddedContent"/>
<xs:element ref="g:LocalHoldingPackage"/>
<xs:element ref="l:LogicalProduct"/>
<!--xs:element ref="NCubeLogicalProduct"/>
<xs:element ref="NCubeScheme"/-->
<xs:element ref="a:OrganizationScheme"/>
<xs:element ref="p:PhysicalDataProduct"/>
<xs:element ref="pi:PhysicalInstance"/>
<xs:element ref="p:PhysicalStructureScheme"/>
<xs:element ref="d:QuestionScheme"/>
<xs:element ref="p:RecordLayoutScheme"/>
<xs:element ref="g:ResourcePackage"/>
<xs:element ref="s:StudyUnit"/>
<xs:element ref="c:UniverseScheme"/>
<xs:element ref="l:VariableScheme"/>
</xs:choice>
</xs:complexType>
<xs:element name="DDIFragment" type="DDIFragmentType"/>
</xs:schema>
```