



Can DDI eXist?

Johan Fihn – Swedish National Data Service
Leif-Jöran Olsson – Swedish Language Bank
Olof Olsson – Swedish Language Bank

IASSIST 2011, 2011-06-01, Vancouver, Canada

<http://ddi-exist.info/index.html>

The web site will contain most of our work with eXist-db and DDI. Demos and examples will be posted here.

eXist-db

- **eXist-db** is an **Open Source Project** licenced under LGPL and is developed in an **Agile** and **Open** manner.
- 9+ Years since inception
- 22 Stable Releases
- > 530k lines of code
- ~ 140 man years of effort
- ~ \$8.0M development cost
- Large Open Source community and user base
- Version 1.4.1 is imminent!

History - 1

Founded	5th January 2001
0.2.1	16th January 2001
0.2.2	25th January 2001
0.2.3	22nd February 2001
0.3	23rd May 2001
0.5	5th August 2001
0.6	14th September 2001
0.7	1st November 2001
0.8	5th August 2002
0.8.1	26th September 2002
0.9	24th January 2003
0.9.1	9th April 2003
0.9.2	15th August 2003

History - 2

1.0	15th October 2006
1.1	15th October 2006
1.1.1	4 th November 2006
1.0.2	24 th February 2007
1.1/1.2	(many snapshots)
1.1.2	15 th April 2007
1.2	14th January 2008
1.2.1	14 th May 2008
1.2.2	25 th May 2008
1.2.3	10 th June 2008
1.2.4	2 nd August 2008
1.2.5	24 th February 2009
1.2.6	17 th June 2009
1.4	8th September 2009

Reflections

- Recent major releases have been huge
 - Many new features
 - Sometimes initially lacking documentation
 - Many bug fixes
 - Some breaking changes, e.g. alignment to specs
- Users depend on SVN trunk between releases
 - Good for testing and feedback
 - Challenging for “Production” systems
- High quality releases

Planned releases

- Shorter release cycle of 6 months
 - 1.5 Development – to be released soon
 - 1.6 January 2012 *)
 - 1.7 Development – to be defined
 - 1.8 January 2013 *)

*) Exact timing depends on factors

Version 1.6 – January 2012

- Low-level Storage API redesign and improvements
- Native XSLT 3.0 engine
- XQuery Debugger
- JSR-225 / XQJ API
- Pluggable Authentication API / OpenId
- W3C XQuery Full-Text support
- XQuery Update, align to W3C revised specification
- Java 1.6 Generics, Bug fixes and improvements
- WebDAV – using 'Milton'

Version 1.8 – January 2013

- Rewrite HTTP APIs
 - REST
- REST and SOAP delivery of XQuery functions
 - Continued development of 'SoapServer' servlet
- ANTLR Update
 - Better feedback on syntax errors

More features - 1

- (Redesign) Clustering
- Structural Index use for local:* and pfx:*
- Backup/Restore Versioning
- EXI Support (previously Binary XML)
- Content Management Integration
 - JSR-170 / JSR-238 / CIMS
- CIFS Server

More features - 2

- XQuery threading
 - parallel and out-of-order execution
- Further uses of Statistical Index
- Automate Installer Testing
- Documentation
- **Feature Requests from the community!**

XML Validation

XML Validation support in eXist-db

- Implicit
 - Xerces
 - XML catalog support
- Explicit
 - Using extension functions
 - Wrapper around third party libraries
 - Validation Report

Explicit validation - 1

External libraries

- JAXP
 - Validate by parsing
 - javax.xml.parsers
 - Part of Java since Java5
 - XSD, catalogue files (DTD)

- JAXV
 - java.xml.validation
 - XSD

- Implemented by Xerces

Explicit validation - 2

External libraries

- Jing
 - James Clark's / George Bina
 - Hosted on Google code
 - Grammar
 - XML schema (.xsd)
 - Namespace-based Validation Dispatching Language (.nvdI)
 - RelaxNG (.rng and .rnc)
 - Schematron 1.5 (.sch)

Functions

<http://exist-db.org/xquery/validation>

```
validation:jaxp-report(xml, cache, catalogs*) node()  
validation:jaxv-report(xml, xsd+) node()  
validation:jing-report(xml, grammar) node()
```

Specials

- Jaxp-parse
 - Parse in validating mode: schema defaults
- Grammar cache management

Apache Lucene

eXist-db uses Lucene as an extension index for full text searches

This enables powerful indexing and searching on XML node level

Configuration examples for indexing DDI3 can be found on <http://ddi-exist.info>

SND Online Catalogue

Currently under development

Transformations from DDI into

- xhtml
- pdf and epub
- marc and other similar formats

Full text search

OAI-PMH

- harvesting of metadata in the catalogue

SND Question Bank

Backend

SIMS - SNDs Internal Management System

Windows based in-house developed system for managing SND's collections, metadata, contacts, disseminations etcetera

Exports to DDI2.1 and DDI3.1

Framework is C#.NET and MySQL

Uses wsdl to store DDI instances in eXist

SND Question Bank

Frontend

User interface built with jQuery. Communicates with eXist-db using json

eXist transforms QuestionItems to json array

filtered search

- study citation fields
- language

DDI Validator

<http://ddi-exist.info/#validator>

Validation of DDI2.1 and DDI3.0/3.1

Validation by

- file upload
- url
- copy/paste

Drupal module as frontend available

EXPath

<http://expath.org>

“Collaboratively Defining Open Standards for Portable XPath Extensions”

XForms

We will present the following XForms-based solution in December

Generation from model e.g. DDIPProfile

Validation of input

Presentation view of QuestionItems