Application development in XML

eXist-db & XQuery

Alexander Czmiel

17.04.2015

What do you know by now?

- HTML, CSS, JavaScript
  - to build beautiful and informative digital resources for humanities scholarship → digital scholarly editions
- XML & TEI
  - to model and markup your sources
- XPath & XSLT
  - to generate nice looking HTML-output of your TEI documents
- Oxygen
  - to edit and author your documents and scripts

What is missing?

- An easy way to analyze and ask questions across any or all of your TEI documents
- A search engine and database for querying your content
- A web server for publishing your TEI documents

Solution:

Plan for today

- Introduction to eXist-db
- Learn (some) XQuery
  - W3C XML Query Language
- Build a web application

...and everyting without using Oxygen XML Editor...

Why XML databases?

- Analyze XML documents – one, many or fragments
- Easy and efficent processing of XML documents with X-technologies
- Treat semi-structured information in a natural way – don't force your documents into the relational model (tables)
  - NoSQL database
- Supports validation
- Extensive list of native XML databases: http://www.rpbourret.com/xml/ProdsNative.htm

What is eXist-db?

- Native XML database
  - optimized to store XML documents - well-suited to complex, nested, 'semi-structured' documents like TEI
  - able to store any other file types
- Web server
  - To serve XML, HTML, Images, JSON etc. to a client (web browser)
- XPATH, XSLT, XQuery, XUpdate, XPROC processor
- Open Source Software (Java)
  - Use it for free
  - Contribute

→ Web application framework (including development)!

Many digital scholarly editions are powered by exist-db

eXist-db features I

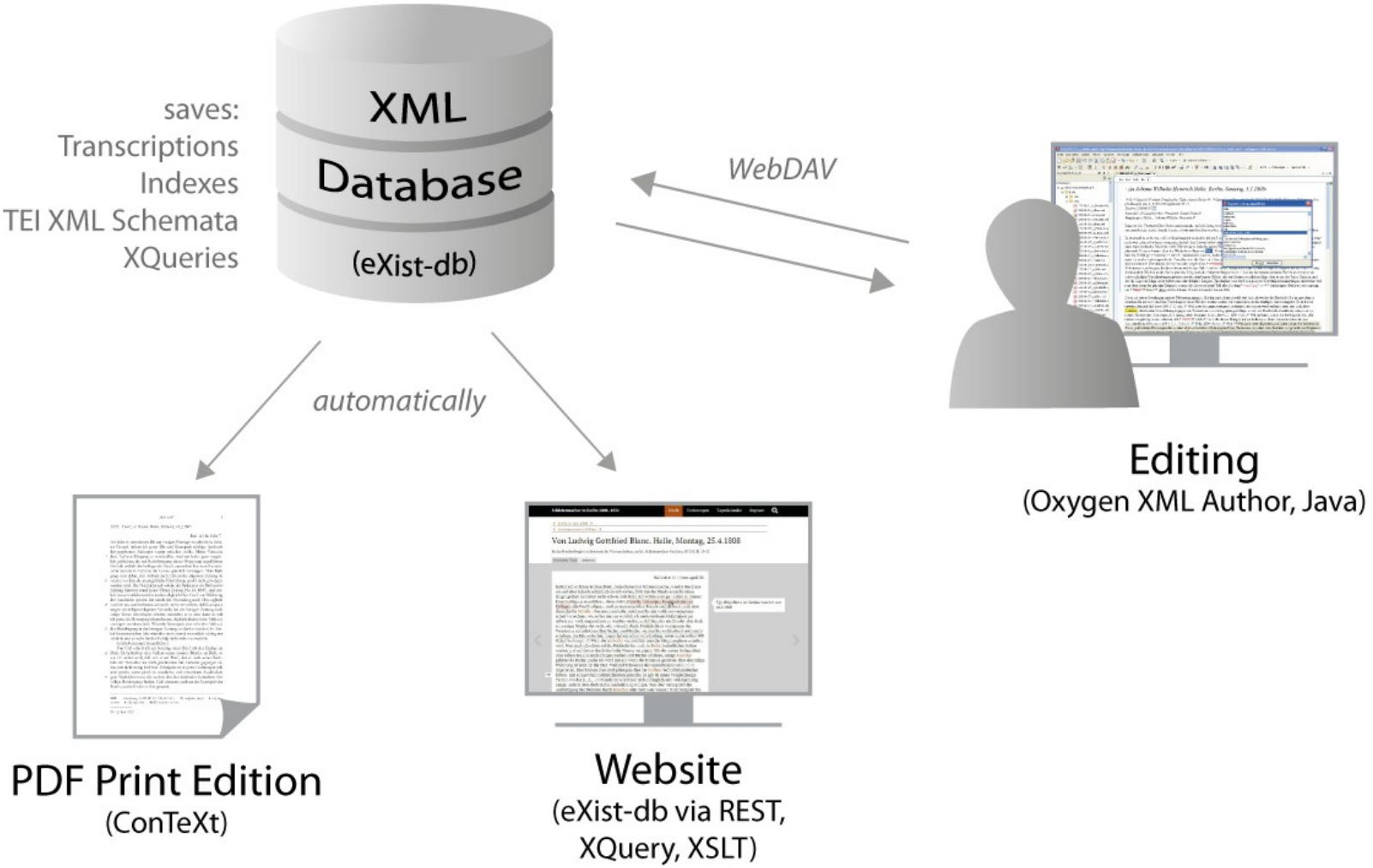
- Stores XML-documents in an efficient way
- Easy to install and to deploy (and to use)
  - Webserver
  - embedded
- Many interfaces to access data
  - REST → http://localhost:8080/exist/rest/db
  - Webdav → http://localhost:8080/exist/webdav/db
  - XML-RPC → xmldb:exist://localhost:8080/exist/xmlrpc
  - XML:DB API
  - SOAP
  - …
- Extensive documentation!

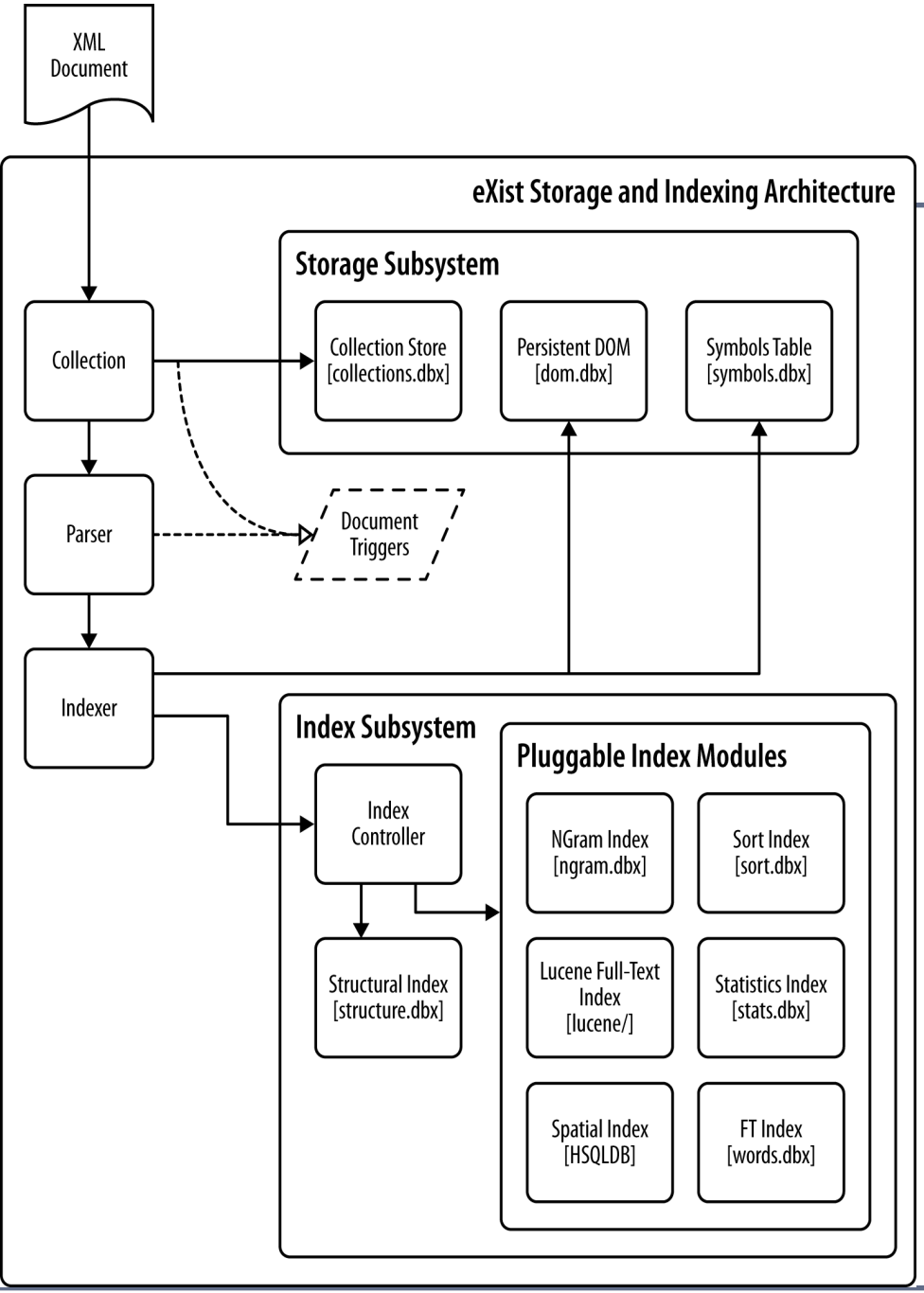
eXist-db features II

- Integrated development environment (IDE) → eXide
  - Web applications written in XQuery, XSLT, HTML, CSS and Javascript
- Integrated XPath, XQuery, XSLT etc. processors
- Different types of indexes (e.g. Lucene based full text index)
- User & permission management
- Application management with package-manager
- Versioning
- Monitoring
- Backup & restore
- Integrates very nicely with oXygen
- ….

More about eXist-db

- http://exist-db.org
- eXist documentation → http://localhost:8080/exist/apps/doc/
- eXist book by Erik Siegel & Adam Retter (O'Reilly)
- eXist mailing list → https://lists.sourceforge.net/lists/listinfo/exist-open
- Stackoverflow → http://stackoverflow.com/questions/tagged/exist-db
- Twitter @existdb
- IRC #existdb on irc.freenode.net
- Professional support at http://www.existsolutions.com/


saves:
Transcriptions
Indexes
TEI XML Schemata
XQueries
XML
Database
(eXist-db)
WebDAV
automatically
Editing
(Oxygen XML Author, Java)
PDF Print Edition
(ConTeXt)
Website
(eXist-db via REST,
XQuery, XSLT)



Where does eXist-db store documents?

- "**deconstructs**" the whole XML document
- Stores single components in an efficient data structure (B+-tree)
- automatically indexes the entire XML structure (+ additional indexes, e.g. Lucene)
- Everything goes to:
  $EXIST_HOME/webapp/WEB-INF/data/
  in your file system

file:///c:/Users/czmiel/ownCloud/IDE_ac/Graz2015/exist2.2/webapp/WEB-INF/data/

Open eXist on your Computer

- http://localhost:8080

What is XQuery?

- XML Query Language (http://www.w3.org/TR/xquery/)
- see it as extended XPath
- used for XML documents and XML databses
- no XML syntax
- useful for extraction and selection of XML fragments and construction of new elements
- Turing complete
- ...and a lot of fun to work with! :)

XQuery supports many expressions:

- Literals (string literals like 'a' and numeric literals like 1)
- Variables ($foo), to which you bind values
- Functions, either built-in like substring-before('hello','l') or your own
- Comments (: this is a comment! :)
- Comparisons: =, <, >, eq
- Conditionals: if then else
- Declarations: declare namespace tei=”http://www.tei-c.org/ns/1.0”
- FLWOR Expressions: the core of XQuery

FLWOR Expressions: the core of XQuery

- let: name a sequence, assigning the whole sequence a variable
- for: iterate through a sequence, assigning each item to a variable
  - where: filter a sequence (optional) → never use with eXist-db!
  - order by: order a sequence (optional)
- return: return the resulting sequence (required)

XQuery resources

- W3C: http://www.w3.org/standards/techs/xquery#w3c_all
- W3Schools Tutorial: http://www.w3schools.com/xquery
- Wikibook: https://en.wikibooks.org/wiki/XQuery
- Walmsley, Priscilla, XQuery. Search across a variety of XML Data, Sebastopol 2007
- XQuery-talk mailing list: http://x-query.com/mailman/listinfo/talk

- Basic XQuery examples: http://localhost:8080/exist/apps/demo/examples/basic/hello.html

Example FLWOR Expressions I

```
for $item in ('c', 'b', 'a')
    order by $item
return $item
```

Returns ('a', 'b', 'c')

Example FLWOR Expressions II

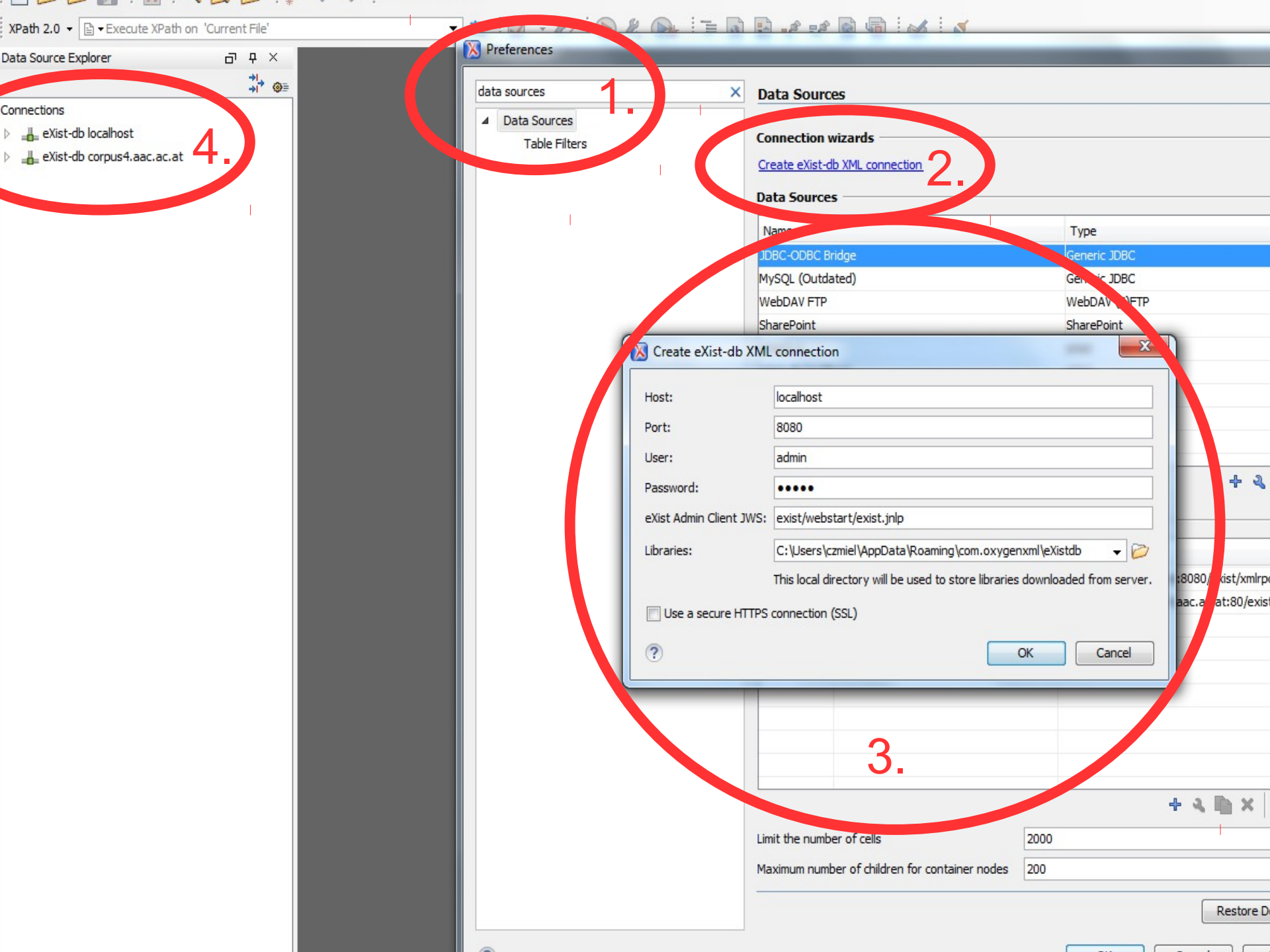
```
let $people := ('Lou', 'Sebastian', 'James')
for $person in $people
let $greeting := concat('Hello, ', $person)
return $greeting
```

Returns ('Hello, Lou', 'Hello, Sebastian', 'Hello, James')

Example FLWOR Expressions III

```
for $role in collection('/db/punch/data')//tei:role
    order by $role
return $role
```

Returns all role elements in the Punch collection in (implicitly) alphabetical order


XPath 2.0
Execute XPath on 'Current File'
Data Source Explorer
Connections
eXist-db localhost
eXist-db corpus4.aac.ac.at
4.
Preferences
data sources
1.
Data Sources
Table Filters
Data Sources
Connection wizards
Create eXist-db XML connection
2.
Data Sources
Name
Type
JDBC-ODBC Bridge
Generic JDBC
MySQL (Outdated)
WebDAV FTP
SharePoint
Create eXist-db XML connection
Host:
localhost
Port:
8080
User:
admin
Password:
eXist Admin Client JWS:
exist/webstart/exist.jnlp
Libraries:
C:\Users\czmiel\AppData\Roaming\com.oxygenxml\eXistdb
This local directory will be used to store libraries downloaded from server.
Use a secure HTTPS connection (SSL)
OK
Cancel
3.
Limit the number of cells
2000
Maximum number of children for container nodes
200
Restore D