



XML-Datenbanken

Grundlagen, eXist: Installation und Handhabung

Ulrike Henny

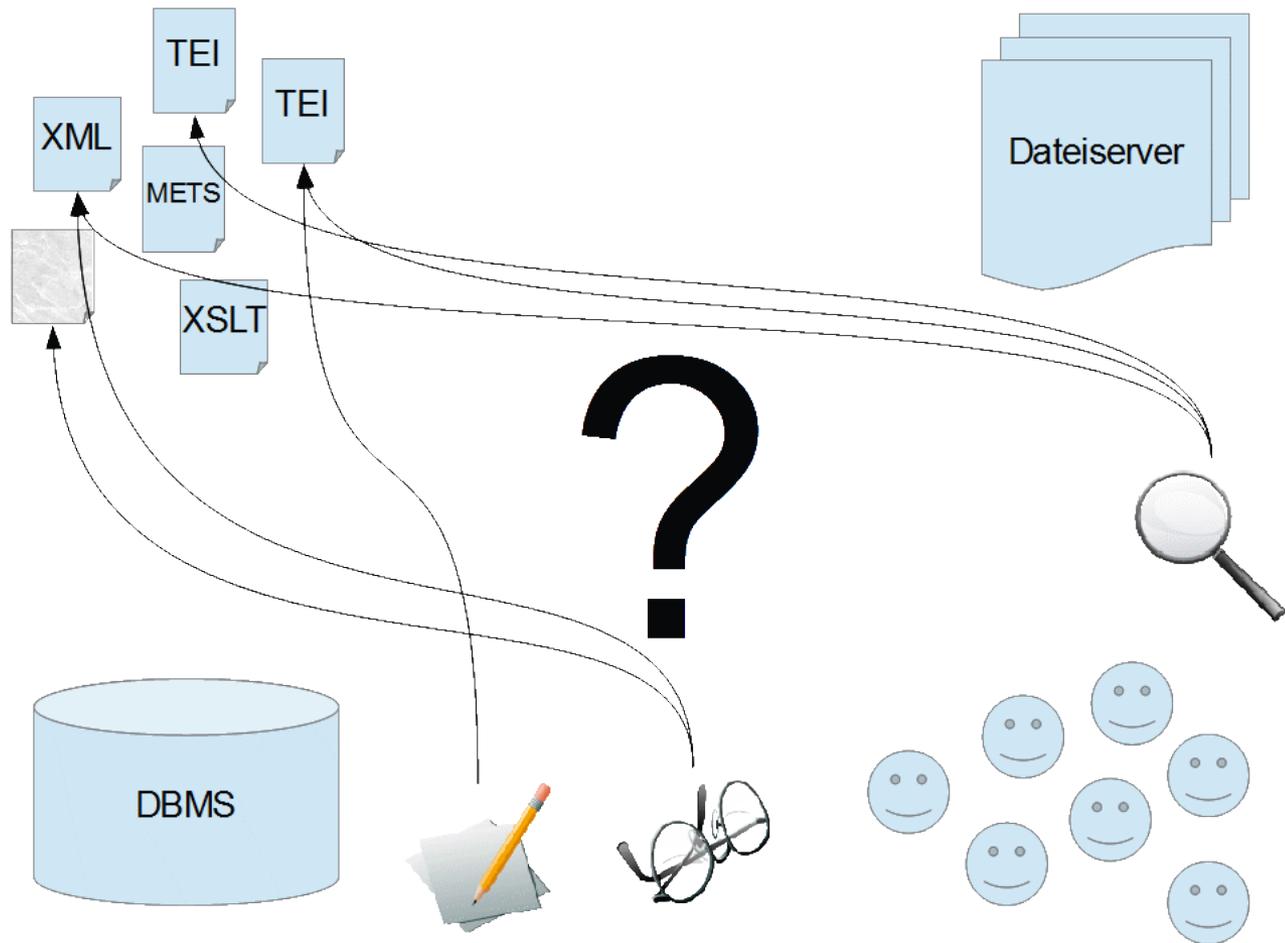
ulrike.henny@uni-koeln.de



Themen:

- Digitale Editionen und Datenbanken
- XML-Datenbanken
- eXist
 - Installationen
 - Features
- Übungen

Digitale Editionen & Datenbanken?



Digitale Editionen & Datenbanken?

- Vorteile von Datenbanken im Allgemeinen:
 - Mehrbenutzerbetrieb
 - Sicherheit
 - Authentifizierung
 - Autorisierung
 - Skalierbarkeit
 - Effizienz/Performanz
 - Volltextsuche

Digitale Editionen & Datenbanken?

- Digitale Editionen
 - Textorientierte XML-Dokumente
 - Transkriptionen, Handschriftenbeschreibungen, Kommentare,...
 - Datenorientierte XML-Dokumente
 - Register, Schlagwortlisten, Glossare, Bibliografien
 - Semistrukturierte Daten
 - Z.B. TEI-Header
 - Nicht-XML-Daten
 - Faksimiles
 - Audio/Videomaterial

Digitale Editionen & Datenbanken?

- Anforderungen an Datenbanken, die mit XML umgehen sollen:
 - Standardkonformität
 - XML 1.0, Namensräume, XML-Schema, DOM
 - Verwaltung ganzer Dokumente
 - Schema-Unabhängigkeit
 - Strukturunabhängigkeit
 - Beliebig tiefer XML-Baum
 - Kommentare, Verarbeitungsanweisungen
 - Web-Anbindung
 - Kodierung
 - XML-Abfragesprache

Typen von XML-Datenbanken

- XML-enabled*
 - XML in Tabellen
 - Mapping

- Hybride XML-Datenbanken
 - Relationale und objektorientierte Datenbanken
 - Erweiterungen für XML-Dokumente

- Native XML-Datenbanken
 - XML-Dokument als zentrale Einheit
 - XML-Datenmodell



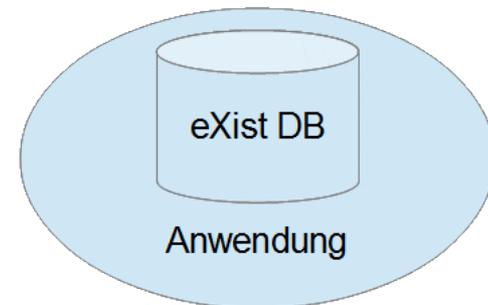
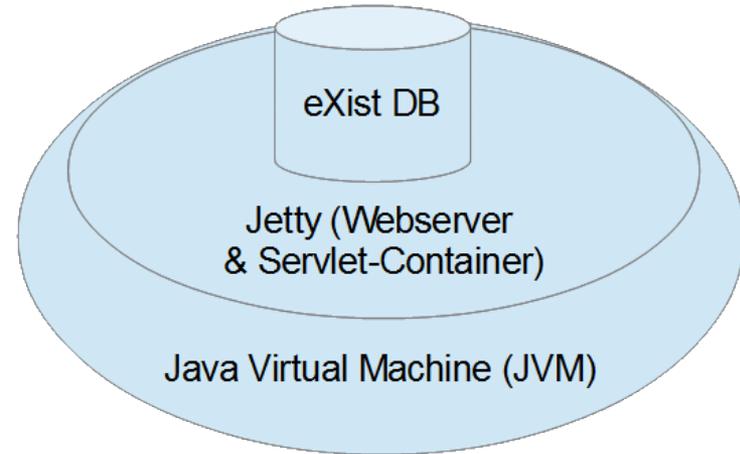
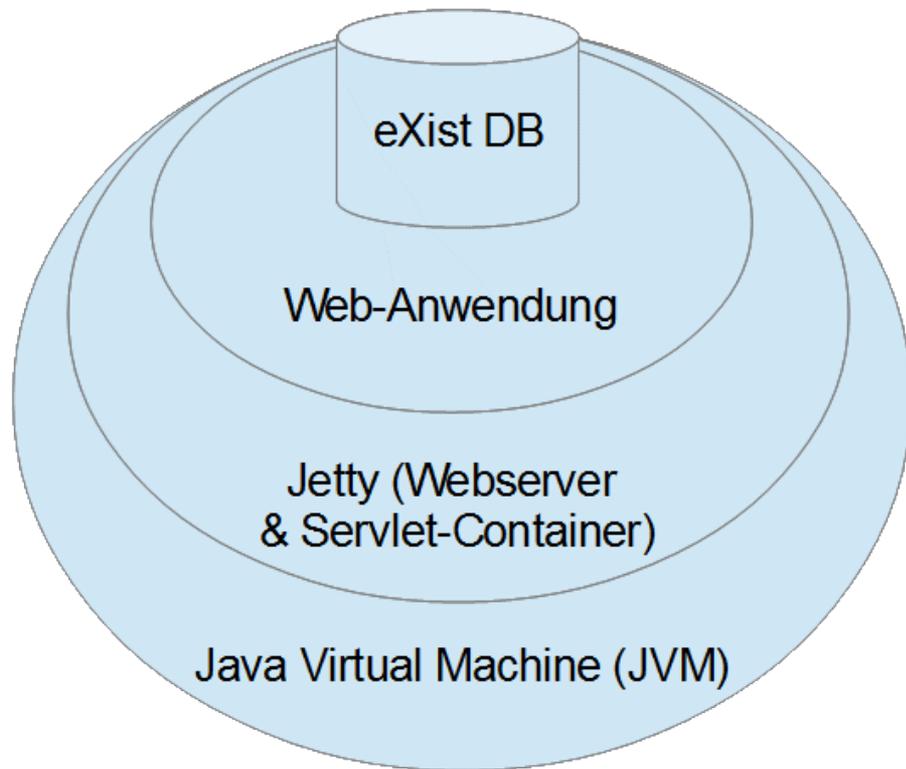
eXist

<http://www.exist-db.org>



- Open Source
- native XML-Datenbank
- 2000 von Wolfgang Meier entwickelt
- aktuelle Version: 1.4.2 (Eindhoven)
- Java

eXist-Installationen



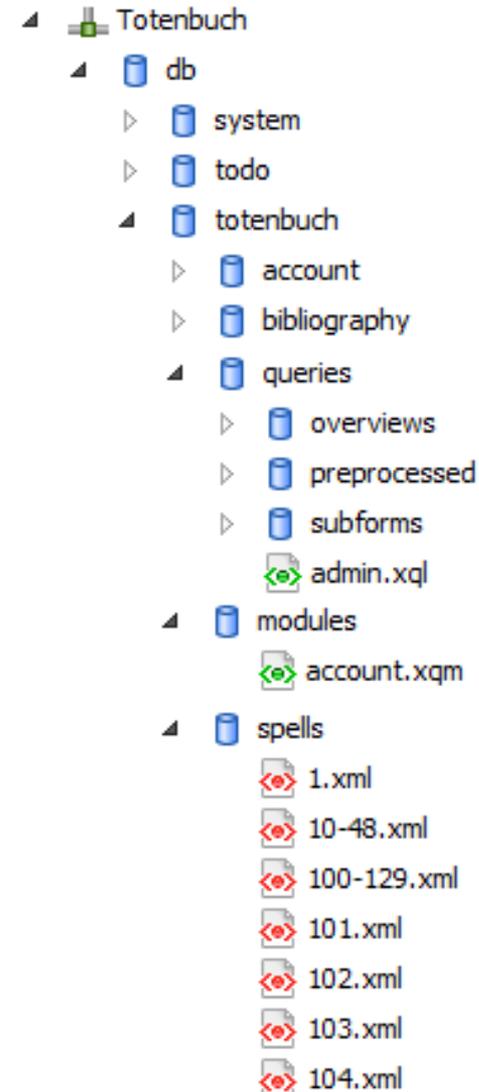
eXist-Features

- *Features:*
 - Collections
 - XPath, XSLT, XQuery
 - XMLDB, XUpdate, XQuery Update
 - HTTP-Schnittstellen
 - URL-Rewriting
 - Indices, z.B. Lucene Volltext-Index
 - Administrationsoberflächen
 - Benutzerverwaltung
 - Backup-Funktionalität
 - 2^{31} Dokumente, beliebig groß?



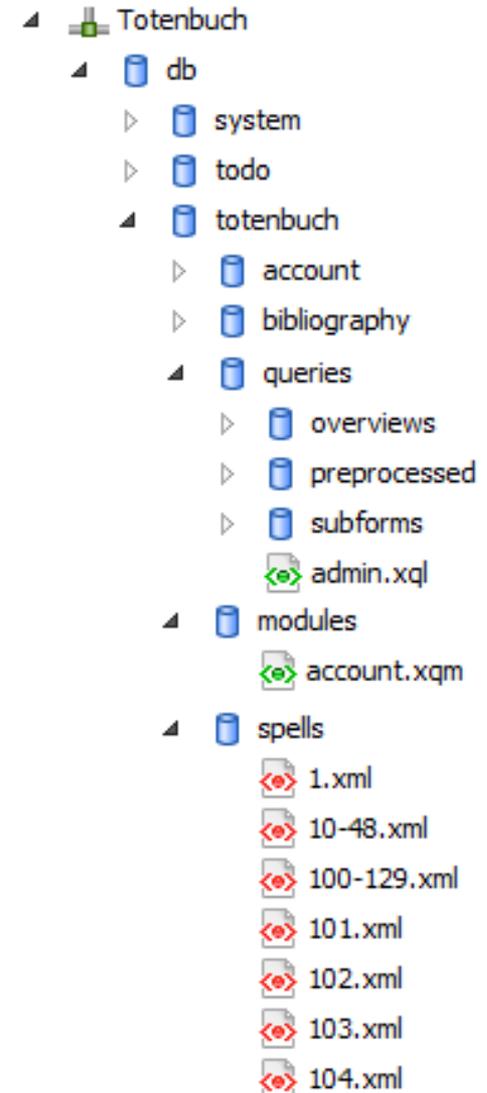
eXist – Features

- **Collections**
 - Dateien in Sammlungen
 - hierarchisch organisiert
 - Vorteile:
 - Übersichtlichkeit
 - Suchen, Abfragen, Bearbeiten, Löschen auf der Ebene einzelner Sammlungen
 - Rechtevergabe



eXist - Features

- Collections
 - Superstruktur
 - Zugriff:
 - `collection($collection-uris as xs:string*, ...) node()*`
 - `doc($document-uri as xs:string?) node()?`
 - Beispiele:
 - `collection("/db/totenbuch/spells")`
 - `collection("xmldb:exist:///db/totenbuch/spells")`
 - `collection("xmldb:exist:///db/totenbuch")//spruch`
 - `doc("/db/totenbuch/spells/1.xml")`



eXist - Features

- **XPath, XSLT, XQuery**
 - Navigation
 - Abfrage/Bearbeitung
 - Transformation
 - Darstellung
- XPath
 - Verwendung innerhalb von XSLT und XQuery
- XSLT/XQuery
 - Ablage von XSLT/XQuery-Dateien in der Datenbank oder im Dateisystem
 - verschiedene Möglichkeiten zum Aufruf

eXist - Features

XSLT-Aufrufe

- aus einer XQuery-Datei heraus

- `transform:transform(doc("/db/totenbuch/spells/1.xml"),
"xmldb:exist:///db/totenbuch/spells.xsl", ())`

- über die REST-Schnittstelle

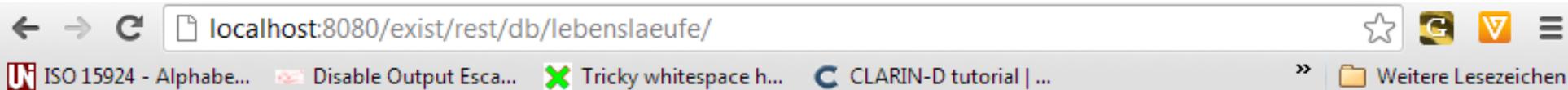
- `http://localhost:8080/exist/rest/db/totenbuch/spells/1.xml?_xsl=spells.xsl`

Exkurs: REST

- **Representational State Transfer**
 - Richtlinie für Webanwendungen/Webservices
 - Prinzipien:
 - Adressierbarkeit
 - eine URL soll genau ein Ergebnis liefern
 - alles hat eine eindeutige Adresse
 - Unterschiedliche Repräsentationen
 - Zustandslosigkeit
 - eine einzelne Anfrage enthält alle nötigen Informationen
 - Operationen
 - HTTP: GET, POST, PUT, DELETE
 - Verbindungshaftigkeit
 - Links zu anderen Ressourcen



Exkurs: REST



This XML file does not appear to have any style information associated with it. The document tree is shown below.

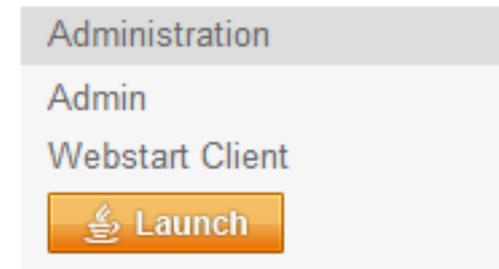
```
<?xml version="1.0" encoding="UTF-8" ?>
<exist:result xmlns:exist="http://exist.sourceforge.net/NS/exist">
  <exist:collection name="/db/lebenslaeufe" created="2012-10-07T11:14:27.003+02:00" owner="admin"
    group="dba" permissions="rwur-ur-u">
    <exist:resource name="lebenslaeufe2.xql" created="2012-10-10T22:52:30.764+02:00" last-modified="2012-10-
      10T23:08:31.313+02:00" owner="admin" group="dba" permissions="rwur-ur-u"/>
    <exist:resource name="lebenslaeufe.xql" created="2012-10-07T13:32:14.873+02:00" last-modified="2012-10-
      07T17:40:13.043+02:00" owner="admin" group="dba" permissions="rwur-ur-u"/>
    <exist:resource name="personen.xql" created="2012-10-07T11:23:23.732+02:00" last-modified="2012-10-
      07T13:30:19.506+02:00" owner="admin" group="dba" permissions="rwur-ur-u"/>
    <exist:resource name="andreae.xml" created="2012-10-07T17:51:14.825+02:00" last-modified="2012-10-
      07T17:51:14.825+02:00" owner="admin" group="dba" permissions="rwur-ur-u"/>
    <exist:resource name="beulwitz.xml" created="2012-10-07T11:21:27.929+02:00" last-modified="2012-10-
      07T11:21:27.929+02:00" owner="admin" group="dba" permissions="rwur-ur-u"/>
    <exist:resource name="untitled.xquery" created="2012-10-10T22:52:05.902+02:00" last-modified="2012-10-
      10T22:52:05.902+02:00" owner="admin" group="dba" permissions="rwur-ur-u"/>
    <exist:resource name="lebenslaeufe.xsl" created="2012-10-10T22:51:31.286+02:00" last-modified="2012-10-
      10T22:51:31.286+02:00" owner="admin" group="dba" permissions="rwur-ur-u"/>
  </exist:collection>
</exist:result>
```

eXist - Features

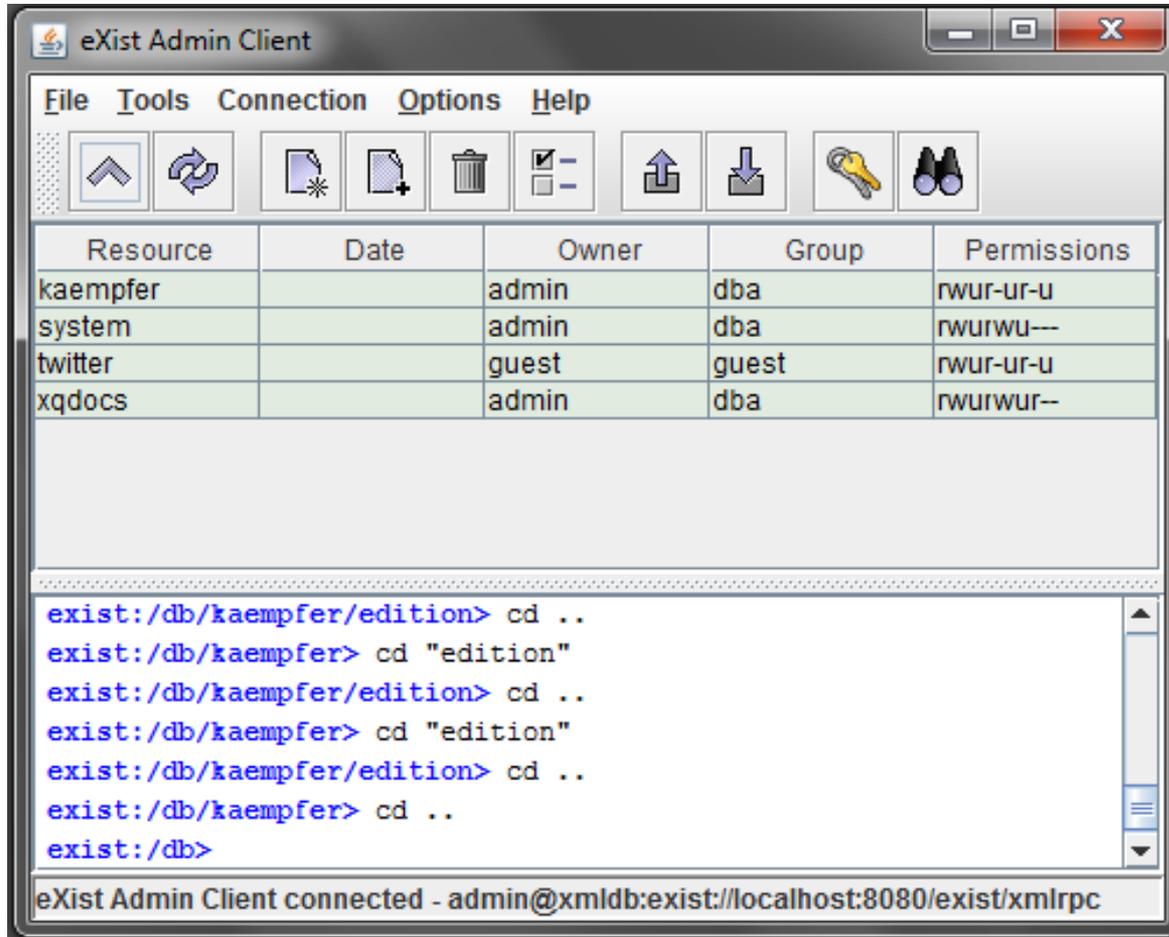
- **XQuery-Aufrufe**
 - über das XQuery-Servlet
 - <http://localhost:8080/exist/totenbuch/queries/spells.xql>
 - über den REST-Server
 - <http://localhost:8080/exist/rest/db/totenbuch/queries/spells.xql>
 - über die REST-Schnittstelle
 - [http://localhost:8080/exist/rest/db/totenbuch?_query=//spruch\[1\]](http://localhost:8080/exist/rest/db/totenbuch?_query=//spruch[1])

eXist - Features

- **Administrationsoberflächen**
 - Java Admin Client
 - Verwaltung von Collections und Dateien
 - Benutzerverwaltung
 - Rechteverwaltung
 - Backup
 - Admin-Webseite
 - ...
 - Systeminformationen
 - Verwaltung von Dokumentation und Beispielen
 - Zusätzliche Tools (“XQuery Sandbox“)
 - Query Profiling
 - Anbindung an Oxygen
 - Verwaltung von Collections und Dateien



eXist – Java Admin Client



The screenshot shows the eXist Admin Client window. The title bar reads "eXist Admin Client". The menu bar includes "File", "Tools", "Connection", "Options", and "Help". The toolbar contains icons for home, refresh, search, add, delete, list, upload, download, key, and binoculars. Below the toolbar is a table with the following data:

Resource	Date	Owner	Group	Permissions
kaempfer		admin	dba	rwur-ur-u
system		admin	dba	rwurwu---
twitter		guest	guest	rwur-ur-u
xqdocs		admin	dba	rwurwur--

Below the table is a terminal window showing the following commands and output:

```
exist:/db/kaempfer/edition> cd ..
exist:/db/kaempfer> cd "edition"
exist:/db/kaempfer/edition> cd ..
exist:/db/kaempfer> cd "edition"
exist:/db/kaempfer/edition> cd ..
exist:/db/kaempfer> cd ..
exist:/db>
```

The status bar at the bottom of the window reads: "eXist Admin Client connected - admin@xmldb:exist://localhost:8080/exist/xmlrpc"



eXist – Admin-Webseite

```
<SPEECH>
<SPEAKER>HAMLET</SPEAKER>
<LINE>Reat, reat, perturbed spirit!</LINE>
<STAGEDIR>They enter.</STAGEDIR>
<LINE>So, gentle,
<LINE>... all my
<LINE>And let so
<LINE>... to
<LINE>... his
<LINE>... not
<LINE>... our
<LINE>... out of
<LINE>That ever I was born to set it right!</LINE>
<LINE>Nay, come, let's go together.</LINE>
</SPEECH>
```

Select a Page

- Home
- System Status
- Browse Collections
- User Management
- View Running Jobs
- Examples Setup
- Install Tools
- Install Documentation
- Backups
- Query Profiling
- Grammar cache
- Shutdown
- Logout

Logged in as: admin

Running XQueries

ID	Type	Source
<input type="checkbox"/> 28490615	File	C:\Program Files (x86)\eXist\webapp\admin\proc.xql

Running Jobs

No long running jobs are active right now.

eXist – Oxygen-Anbindung

The screenshot displays the Oxygen XML Editor interface. On the left, the 'Datenquellen Explorer' (Data Source Explorer) shows a project structure under 'Verbindungen' (Connections):

- Totenbuch
- Kaempfer
 - db
 - kaempfer
 - edition
 - facsimile.xml
 - tei-comment.xml
 - tei-transcript.xml
 - tei-translation.xml
 - system
 - todo
 - twitter
 - xqdocs

The main editor window shows the XML code for 'tei-comment.xml'. The code is as follows:

```

1 <!-- <!DOCTYPE TEI SYSTEM "schema/AmEx_M
2 <!ENTITY facsimile SYSTEM "facsimile.xml
3 ]> --><!--
4   - xml:lang: generell einfügen oder n
5   - botanische Autorenkürzel wurden mi
6   - botanische Pflanzennamen mit <rs t
7 -->
8 <TEI xmlns="http://www.tei-c.org/ns/1.0"
9   <teiHeader xml:lang="ger">
10     <fileDesc>
11       <titleStmt>
12         <title>
13           <seg xml:lang="lat">
14             Fasciculi V</seg>
15         </title>
16       <author>
17         <ptr target="#kaempf
18       </author>
19     <editor>
20     <ptr target="#kakee1

```

XML-Datenbanken, eXist - Literatur

- eXist Developer's Blog: <http://atomic.exist-db.org/blogs/eXist>
- eXist-Dokumentation: <http://www.exist-db.org/exist/documentation.xml>
- Ronald Bourret, *Consulting, writing, and research in XML and databases*: <http://www.rpbourret.com/>
- Kazakos, W./Schmidt, A./Tomczyk, P., *Datenbanken und XML*, Berlin/Heidelberg 2002.
- Meier, Wolfgang: „eXist: An Open Source Native XML Database“, in: A.B. Chaudhri et al. (Eds.): *Web Databases and Web Services 2002*, LNCS 2593, pp. 169–183, 2003.
- Schöning, Harald, *XML und Datenbanken. Konzepte und Systeme*, München/Wien 2003.

Übung 1: Installation von eXist

- siehe: *Documentation -> Quick Start*
 - <http://www.exist-db.org/exist/quickstart.xml>

- Voraussetzung: Java Development Kit (JDK), Version 1.6
 - <http://www.oracle.com/technetwork/java/javase/downloads/jdk6u35-downloads-1836443.html>
 - Falls nicht vorhanden, bitte installieren!

- Aktuelle eXist-Version:
 - <http://www.exist-db.org/exist/download.xml>
 - 1.4.2 (Stable Release)

Übung 1: Installation von eXist

- Aktuelle eXist-Version herunterladen
 - <http://www.exist-db.org/exist/download.xml>
 - 1.4.2 (Stable Release)
 - [**eXist-setup-1.4.2-rev16251.jar**](#)
- Installationsvorgang starten
 - Windows/Mac: durch Doppelklick auf die Datei
 - Linux: über die Kommandozeile in das Verzeichnis wechseln, in dem die Installationsdatei liegt und Folgendes aufrufen: `java -jar eXist-1.4.2-rev16251.jar`
- Installation durchführen (1)
 - Man wird nach dem JDK-Pfad gefragt; wenn nicht anders gewählt, ist dies (bei Windows) `C:\Programme\Java\jdk...`

The installed application needs a JDK. A java runtime environment (JRE) will be not sufficient.

 Select the JDK path:

C:\Program Files (x86)\Java\jdk1.6.0_35



Browse...

Übung 1: Installation von eXist

- Installation durchführen (2)
 - Frage nach dem Verzeichnis, in dem eXist installiert werden soll: besser außerhalb von „Programme“!

On Windows Vista, please do not install eXist into the 'Program Files' directory unless you deactivated UAC or plan to always run eXist in privileged mode.

 Select the installation path:

C:\eXist

 Browse...

- Frage nach Datenverzeichnis und Admin-Passwort
 - Angabe zum Datenverzeichnis unverändert lassen
 - Admin-Passwort neu setzen!

Set Data Directory and Admin Password

Please select a directory where eXist will keep its data files. On Vista and Windows 7, this should be outside the usual 'Program Files' directory:

Data dir:

C:\eXist2\webapp\WEB-INF\data

 Browse...

Please enter a password for user 'admin', the database administrator:

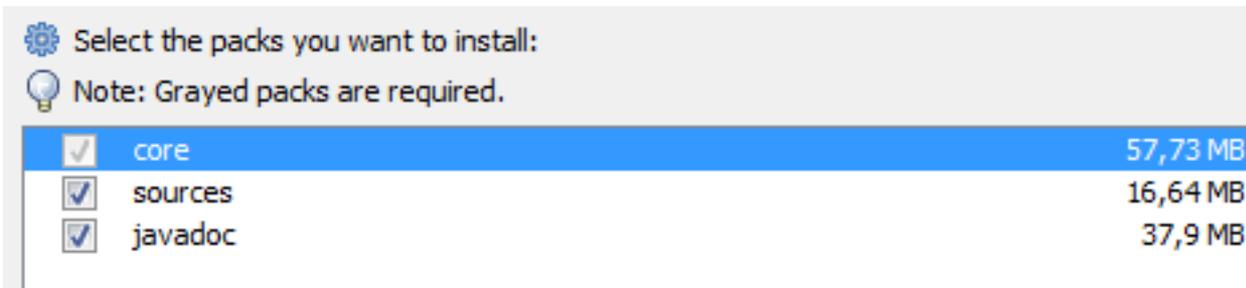
Enter password:

Retype password:

Übung 1: Installation von eXist

□ Installation durchführen (3)

- Frage nach den zu installierenden Komponenten: Auswahl belassen und weitergehen



- die weiteren Schritte durchführen lassen/bestätigen

□ eXist starten

- Windows/Mac: auf „eXist Database Startup“ klicken

- Icon auf dem Desktop (wenn dies am Ende der Installation gewünscht wurde)
- Oder über das Start-Menü -> alle Programme -> eXist XML Database -> eXist Database Startup

- Linux: über die Kommandozeile

- Im Verzeichnis, in dem eXist installiert ist, `tools/wrapper/bin/exist.sh` start aufrufen

Übung 1: Installation von eXist

- <http://localhost:8080/exist> im Browser aufrufen!
 - Läuft alles?
 - Falls nicht:
 - Webserver (Jetty) und Datenbank (eXist) müssen laufen
 - Hinweise in der Dokumentation: <http://www.exist-db.org/exist/quickstart.xml> (11. If something goes wrong...)

Übung 2: Benutzer & Collections anlegen

- über den *Java Admin Client*
 - <http://localhost:8080/exist> im Browser aufrufen
 - Java Admin Client über den „Launch“-Button starten (unten rechts auf der Seite)

Administration

Admin

Webstart Client



Übung 2: Benutzer & Collections anlegen

- Admin-Passwort eingeben

eXist 1.4.2 Database Login

?

Username

Password

Type ▼

Configuration

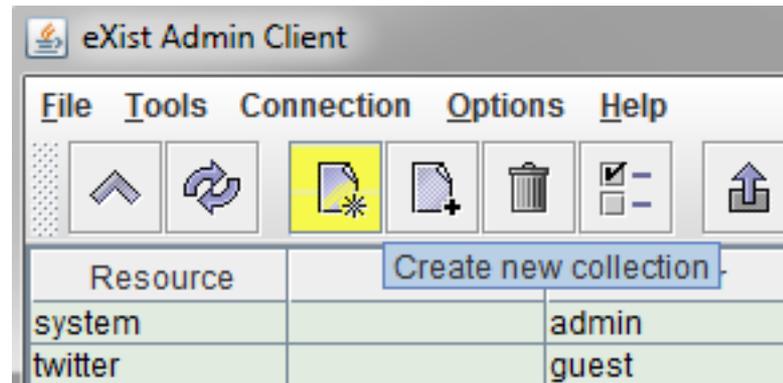
URL

Title

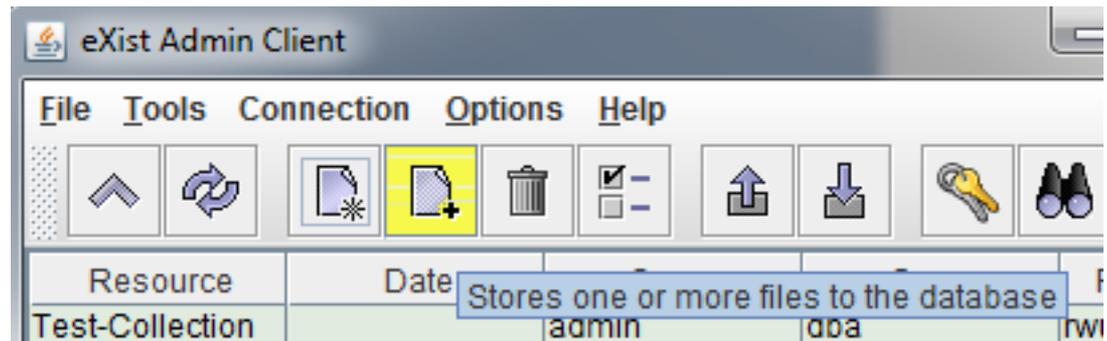
Favourites

Übung 2: Benutzer & Collections anlegen

- Neue Collection anlegen



- Ressourcen hinzufügen/
hochladen



Übung 2: Benutzer & Collections anlegen

Benutzerverwaltung

Date	Owner	Group	Manage users
	admin	dba	rwur-ur-u
	admin	dba	rwurwu---
	guest	guest	rwur-ur-u

Benutzer anlegen

Username:

Password:

Password (repeat):

Home-Collection:

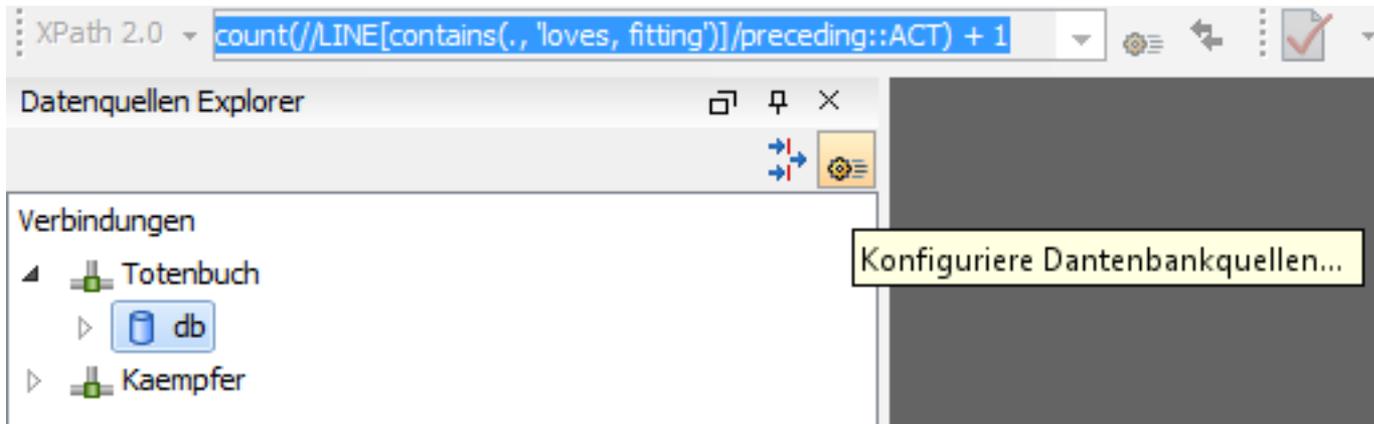
Übung 3: Anbindung an Oxygen

- Dokumentation siehe http://oxygenxml.com/xml_editor/eXist_support.html
- Dazu braucht man ein paar Treiberdateien:
 - in den Materialien: eXist.zip/exist_oxygen
- Oxygen öffnen, zur Datenbankansicht wechseln:



Übung 3: Anbindung an Oxygen

- Datenquelle konfigurieren:



Übung 3: Anbindung an Oxygen

- Neue Datenquelle hinzufügen (1):

Datenquellen

Datenquellen

Name	Typ
JDBC-ODBC Bridge	Generic JDBC
MySQL (Veraltet)	Generic JDBC
WebDAV FTP	WebDAV (S)FTP
Totenbuch	eXist
Kaempfer	eXist



Übung 3: Anbindung an Oxygen

- Neue Datenquelle hinzufügen (2):
 - Namen vergeben
 - Typ: eXist
 - Treiberdateien hinzufügen
 - Bestätigen

Name

Testdatenquelle

Typ

eXist

Treiberdateien (JAR, ZIP)

file:/C:/Users/Ulrike/Dropbox/SpringSchool2012/Teilnehmermaterialien/eXist/exist_oxygen/exist.jar
file:/C:/Users/Ulrike/Dropbox/SpringSchool2012/Teilnehmermaterialien/eXist/exist_oxygen/ws-commons-util-1
file:/C:/Users/Ulrike/Dropbox/SpringSchool2012/Teilnehmermaterialien/eXist/exist_oxygen/xmlldb.jar
file:/C:/Users/Ulrike/Dropbox/SpringSchool2012/Teilnehmermaterialien/eXist/exist_oxygen/xmlrpc-client-3.1.2
file:/C:/Users/Ulrike/Dropbox/SpringSchool2012/Teilnehmermaterialien/eXist/exist_oxygen/xmlrpc-common-3.:

Übung 3: Anbindung an Oxygen

- Neue Verbindung hinzufügen (1):

Verbindungen

Aktivi...	Name	URL
<input checked="" type="checkbox"/>	Totenbuch	xmlldb:exist://awktotenbuch.uni-koel...
<input checked="" type="checkbox"/>	Kaempfer	xmlldb:exist://localhost:8080/exist/x...

    |  

Begrenze die Anzahl der Zellen

Maximale Anzahl der Kinder für Containerknoten

Übung 3: Anbindung an Oxygen

- Neue Verbindung hinzufügen (2):
 - Namen vergeben
 - zuvor eingerichtete Datenquelle auswählen
 - XML DB URI: xmldb:exist://localhost:8080/exist/xmlrpc
 - Benutzer: admin, PW eingeben!
 - Bestätigen

Name:

Datenquelle: 

Verbindungsdetails

XML DB URI:

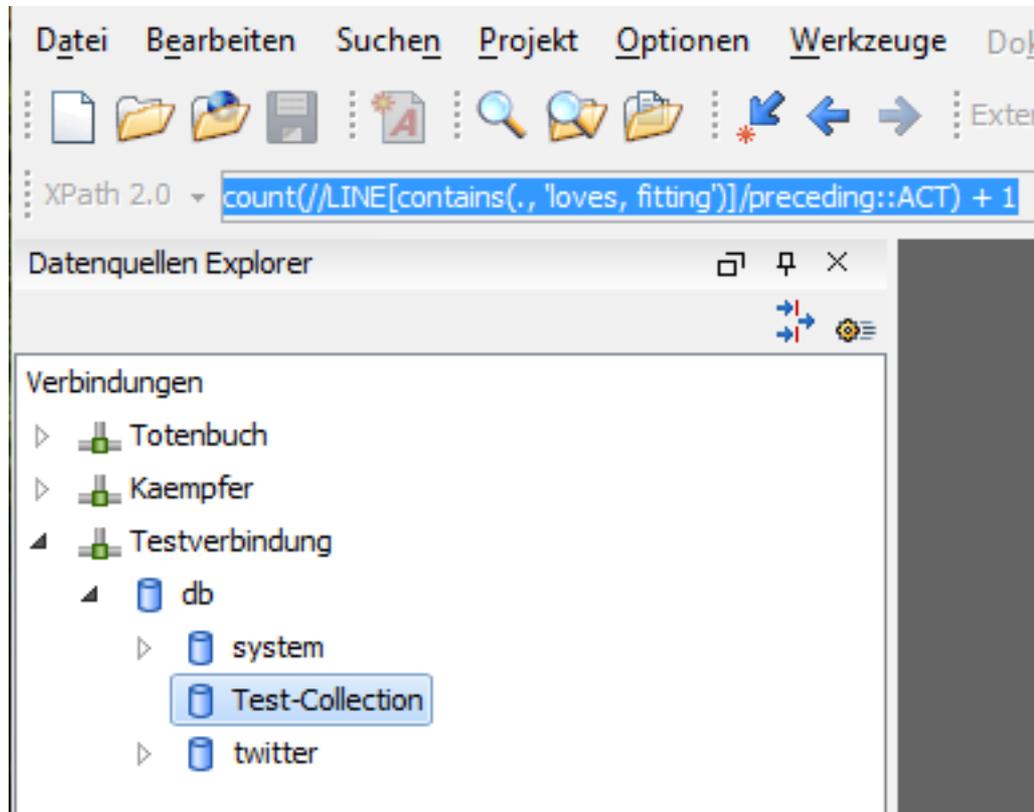
Benutzer::

Passwort::

Sammlung:

Übung 3: Anbindung an Oxygen

- Die Verbindung wird nun im Datenquellen-Explorer angezeigt!



- Es kann direkt auf Collections und Ressourcen der eXist-Instanz zugegriffen werden
- Hinzufügen neuer Sammlungen und Ressourcen durch Rechtsklick auf die Collection, in der sie angelegt werden sollen

Übung 4: Ausgabe einer Webseite

- Beispieldateien:
 - XML: exist.zip/beispieldateien/andreae.xml
 - XSLT: exist.zip/beispieldateien/lebenslaeufe.xsl
 - Hier: in einer Collection „lebenslaeufe“ ablegen!

- Über die REST-Schnittstelle (Aufruf im Browser) können...
 - Ressourcen angefordert werden (hier die XML-Datei):
 - <http://localhost:8080/exist/rest/db/lebenslaeufe/andreae.xml>
 - XSLT-Stylesheets angewendet werden:
 - http://localhost:8080/exist/rest/db/lebenslaeufe?_xsl=lebenslaeufe.xsl
 - XQuery/XPath-Anfragen gesendet werden:
 - `http://localhost:8080/exist/rest/db/lebenslaeufe?_query=declare namespace tei="http://www.tei-c.org/ns/1.0"; //tei:persName`



Danke für die Aufmerksamkeit!