

Databázové systémy a SQL

Lekce 12 - Úvod do XML

Daniel Klimeš

- Formát pro přenos strukturovaných dat
- Text se značkami (tagy)
- Připomíná HTML
- Samotný standard specifikuje jen minimum značek
- Položky (uzly - nodes)
 - Elementy
 - Atributy
 - ...
- Zanořená struktura
 - 1 kořenový element (root)
 - N zanořených elementů

- Hlavička (Processing Instruction)
 - `<?xml version="1.0" encoding="UTF-8"?>`
- Uzly
 - Elementy
 - `<tag>data</tag>`
 - `<question>Datum narození</question>`
 - Atributy
 - `<element atribut=„text“>`
 - `<name lang="cz">`
 - Komentář
 - `<!-- no need to escape <code> & in comments -->`
- Hierarchie elementů
 - Root element
 - Zanořené elementy
 - `<el_root>`
 `<el_nest>text</el_nest>`
 - `</el_root>`
 - `<cluster>`
 `<question>Datum vyšetření</question>`
 - `</cluster>`

Speciální znaky

- < místo "<"
- > místo ">"
- & místo "&"
- ' místo "'"
- " místo ""

- Textový editor - Notepad
- XML editor - XML Spy
- Moderní databáze - POSTGRES, ORACLE
 - Nadstavba SQL
- Programovací jazyky
 - Parsery

- XPATH - formát vyhledávacích dotazů do XML dokumentů
- XML schéma - šablona XML dokumentu
- XSLT - transformace XML do jiného formátu (jiné XML, prostý text , HTML)
- Xquery - pokročilé vyhledávání v XML

- Internet
 - *www.w3.org*
 - *www.w3schools.com*

<bookstore>

<book>

<title>Harry Potter</title>

<author>J K. Rowling</author>

<year>2005</year>

<price>29.99</price>

< /book>

</bookstore>

- Uzly (nodes) – elementy, atributy, text, komentář,..
- Parent, children, siblings, ancestors, descendants

Výraz	Popis
/	Kořenový uzel
//	Uzel kdekoliv
.	Aktuální uzel
..	Rodičovský uzel
@	Atribut
text()	Obsah elementu

<code>/bookstore/book[1]</code>	První element kniha
<code>/bookstore/book[last()]</code>	Poslední element kniha
<code>/bookstore/book[last()-1]</code>	Předposlední kniha
<code>/bookstore/book[position()<3]</code>	První 2 knihy
<code>//title[@lang]</code>	Jakýkoliv element title s atributem lang
<code>//title[@lang='eng']</code>	Jakýkoliv element title s atributem lang s hodnotou 'eng'
<code>/bookstore/book[price>35.00]</code>	Knihy s cenou větší než 35
<code>/bookstore/book[price>35.00]/title</code>	Názvy knih s cenou větší než 35

<code>*</code>	Libovolný element
<code>@*</code>	Libovolný atribut
<code>node()</code>	Libovolný uzel

- Podpora ukládání XML dat
 - Datový typ XML
 - CREATE TABLE xmlobjekty (
 - a XML
)
 - INSERT INTO xmlobjekty VALUES (
 - XMLPARSE (DOCUMENT '<?xml
 - version="1.0"?><book><title>Manual</title><chapter>I</chapter></book>
 ')
)
 - SELECT XPATH('//title/text()', xmlobjekty.a) FROM xmlobjekty
- Speciální SQL funkce - skládání XML stromu
 - XMLELEMENT
 - XMLATTRIBUTES
 - XMLAGG
 - ...

➤ XMLELEMENT

➤ XMLATTRIBUTES

➤ XMLAGG

```
SELECT XMLELEMENT (NAME studie, XMLELEMENT (NAME id, study_id),
XMLLEMENT (NAME jmeno, study_name), XMLELEMENT (NAME nazev,
study_title)) FROM studies
```

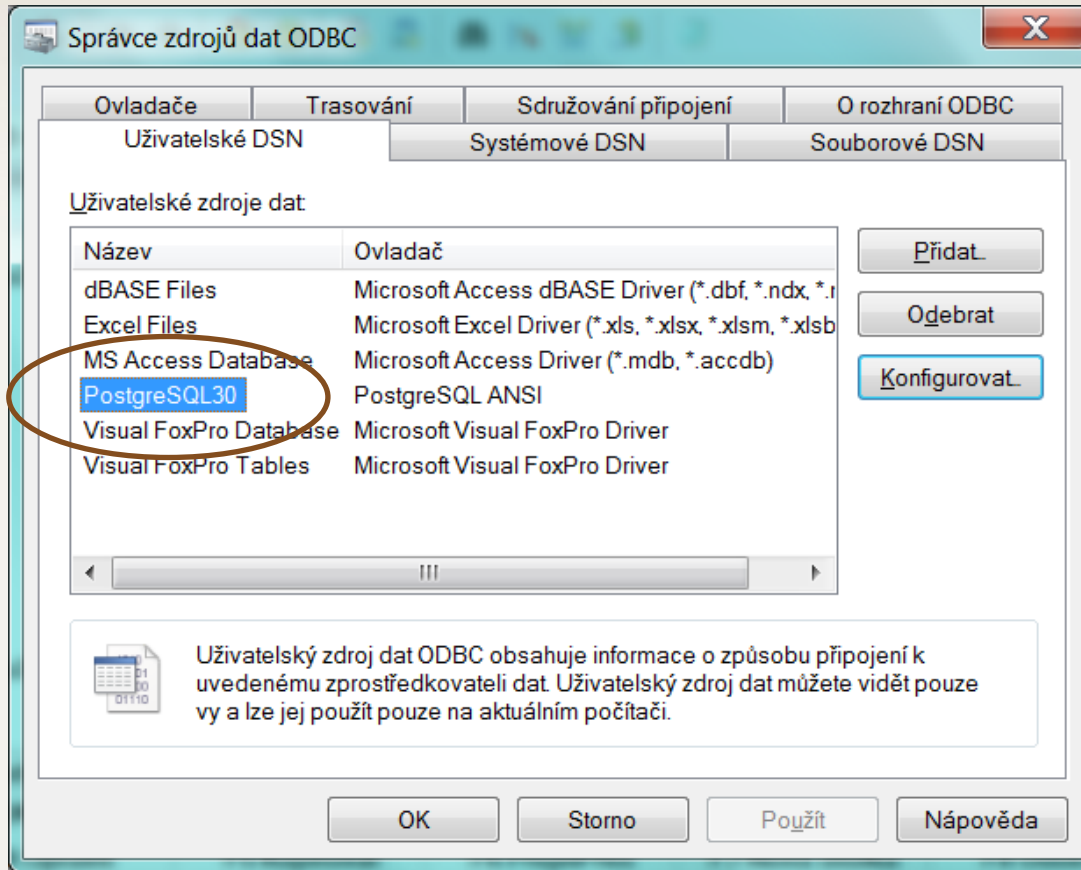
```
SELECT XMLELEMENT (NAME studie, XMLATTRIBUTES(study_id
as"id", study_name as "jmeno"), XMLELEMENT (NAME nazev,
study_title)) FROM studies
```

```
SELECT XMLELEMENT (NAME studie, XMLAGG( XMLELEMENT
(NAME nazev, XMLATTRIBUTES (study_id as "id"), study_title)))
FROM studies
```

- INSERT INTO...
- SELECT XPATH('//nazev/text()', xmltabulka.a) FROM xmltabulka
- SELECT XPATH('//studie/@jmeno', xmltabulka.a) FROM xmltabulka where XMLEXISTS('//studie[@id="3"]' PASSING BY REF a)

- Přenos dat z databáze do jiné aplikace
- Vyžaduje ODBC ovladač k dané databázi
 - psqlodbc_09_03_0400.zip
- Konfigurace propojení (ODBC zdroj)
 - c:\Windows\SysWOW64\odbcad32.exe (WIN 7, 8 - 64 bit)

ODBC správce zdrojů



PostgreSQL ODBC driver

PostgreSQL ANSI ODBC Driver (psqlODBC) Setup

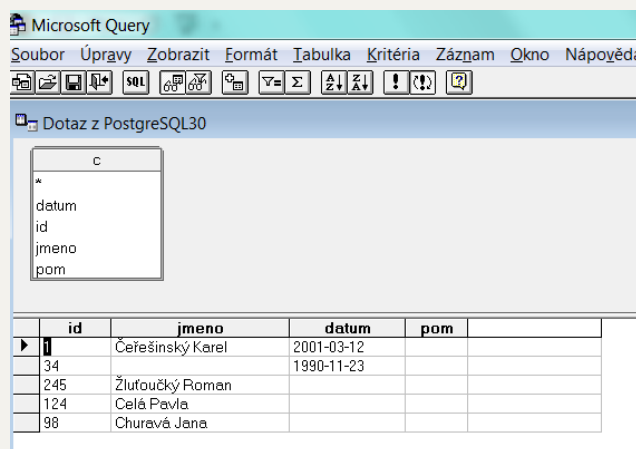
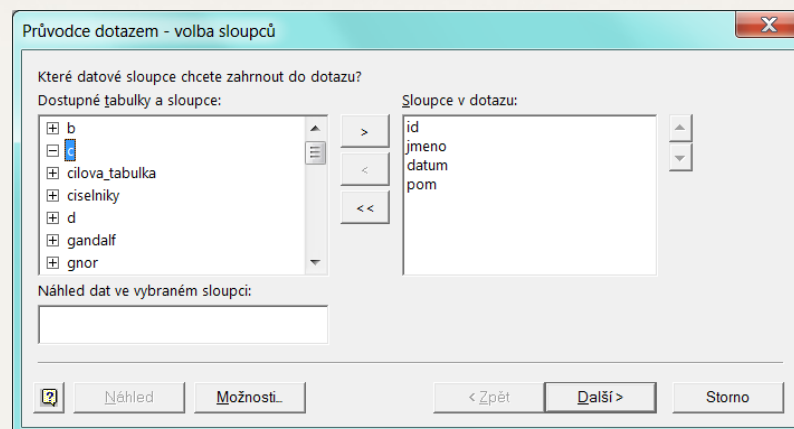
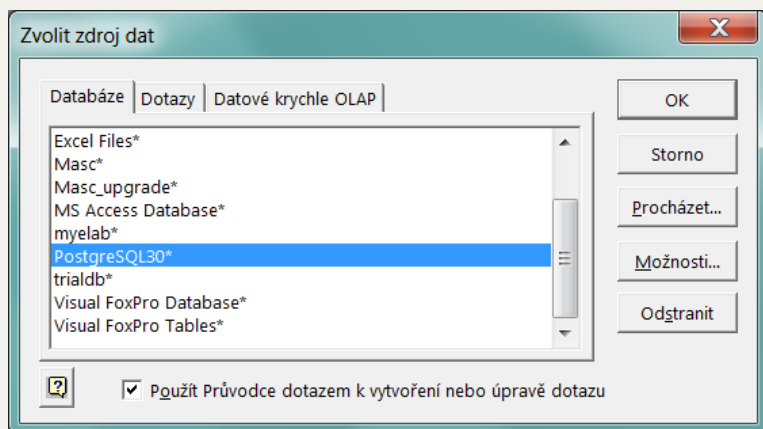
Data Source	PostgreSQL30	Description	
Database	postgres	SSL Mode	disable
Server	localhost	Port	5432
User Name	postgres	Password	••••••••

Options

Datasource Global Manage DSN

Test Save Cancel

- MS EXCEL
 - DATA->Z jiných zdrojů -> Z Microsoft query



- `library(RODBC)`
- `conn <- odbcConnect("PostgreSQL30", uid="postgres", pwd="heslo")`
- `queryResult <- sqlQuery(conn, "SELECT * FROM tabulka")`
- `odbcClose(conn)`
- `queryResult`

- Verze R musí odpovídat ODBC ovladači
 - 32 x 64 bit

- <http://www.uzis.cz/registry>
- Schéma registru

- Která studie má nejvíce pacientů? (study_name)
- Do kterých studií se zařazovali pacienti v roce 2008 (study_name)
- Vypište pracoviště, která nezařadila žádného pacienta
- Vypište pracoviště, která se účastní více studií, má nejvíce pacientů
- Vytvořte pohled study_name, počet zařazených pacientů
- Vypište věk pacientů při zařazení do studie XY
- Kumulativní počty náběru pacientů do studie
- Najděte pacienty zařazené do více než jedné studie



Zápočet