

Manipulating XML 2: XML Query Language

XML Query Language (XQuery) is...

- A query language for XML documents and databases.
- Useful for transforming files into other formats, especially XML files. (e.g. TEI to HTML)
- A powerful language.

XQ vs. XSLT

As a rule of thumb, XQ is better at extracting data from a document, XSLT is better at preserving its structure.

That said, advanced users can do either with both.

Overview of Important XQ Concepts

Variables

- Store values
- Can store different datatypes, path expressions, and most other types of input.

FLWOR (“Flower”) Expressions

- The core of the XQuery language. (More on the next slide.)

Constructors (Direct | Computed)

- Nodes can be made directly, or computed with functions.

Path Expressions (XPath in XQ)

Regular Expressions

Functions

- There are quite a few functions available to XPath and XQuery. More can be added via modules, and custom functions can be written.

Modules

- Like libraries in Python or packages in R.
- Namespaces

Sequences

- Comma separated
- No nested lists... but uses XML nodes

Example 1: Hello World

```
xquery version “3.0”;  
  
let $variable := “Hello World!”  
return $variable
```

Example 2: Simple Query with HTML Output

```
xquery version “3.0”;  
  
let $document := doc(‘my-file.xml’)  
let $stuff := $document//p/text()  
return  
<html>  
    <head/>  
    <body>  
        {$stuff}  
    </body>  
</html>
```

XQuery 2: FLWOR (Flower) Expressions

Clause	What it does
for	Sets up a loop over a sequence. For each thing in a sequence of things, ...
let	Sets a variable. (i.e. Binds a sequence of one or more items to a variable.)
where	Reduces a set of values according to some criteria.
order by	Orders a set of values.
return	Ends the expression and defines the output.

There are more clauses used in FLWOR expressions, but these are the main ones.

Among these, only *for* and *let* may be used to begin FLWOR expressions, and a FLWOR expression must end with a *return* clause.