

Basic Course on the use of BioNumerics Software for the analysis of Pulsed Field Gel Electrophoresis-generated profiles of *E. coli*

Sharing data in the network: The markup language XML

Arnold Knijn
IT sector - ISS

XML in BioNumerics

Peer-to-peer data exchange

BioNumerics users can exchange information at a peer-to-peer level by simply making a selection of database entries, and clicking the information fields and experiment data to be exported in XML format. Received XML files can be imported and directly analyzed together with other database entries. BioNumerics automatically recognizes which experiments are compatible. XML exchange files can optionally be compressed and encrypted.



Received XML files can be imported as fully editable database entries.

What is XML?

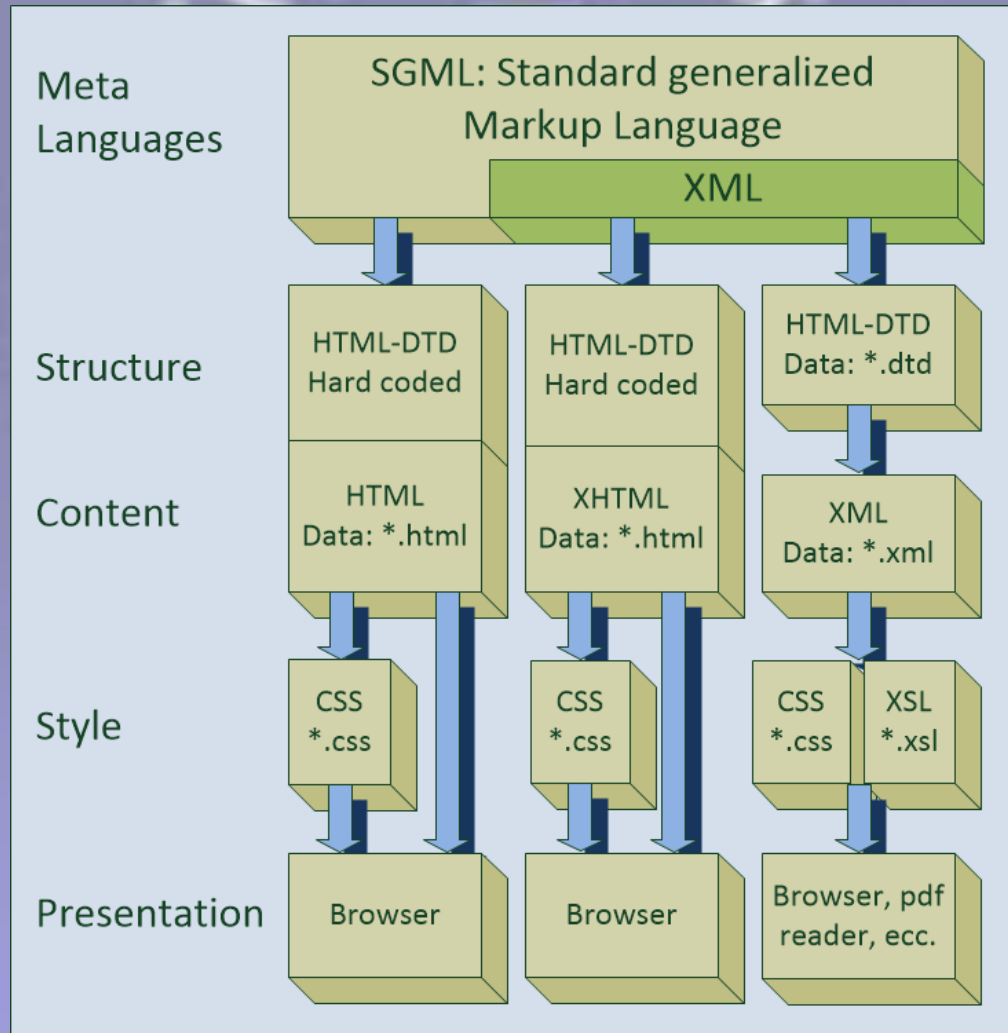
- **eXtensible Markup Language**: a meta markup language that defines a set of rules for encoding documents in a self-describing format that is both human-readable and machine-readable.

	<code><?xml version="1.0" encoding="iso-8859-1"?></code>
	<code><address></code>
ISS	<code><Destination>ISS</Destination></code>
Viale Regina Elena, 299	<code><Street number="299">Viale Regina Elena</Street></code>
00161 Rome	<code><CAP>00161</CAP><City>Rome</City></code>
Italy	<code><Country>Italy</Country></code>
	<code></address></code>

Key features

- Elements have both an opening and a closing tag
- Elements follow a strict hierarchy, with documents containing only one root element
- Elements cannot overlap other elements
- Element names must obey XML naming conventions
- XML is case sensitive

Architecture



Why XML?

- a widely accepted open standard
- separates content from form
- extensible
- self-describing
- universal (internationalization is no problem)
- independent from hardware and software
- provides a robust and durable format for information storage
- easily transformable
- **Data exchange between loosely coupled systems**

XML versus Flat File and Legacy

XML	Flat file	Bundle
Elaborate	Simple	Black box
Verbose	Compact	Compact
Time-consuming	Fast	Fast
Structured	Flat	Structured
Flexible	Rigid	Rigid
Self-describing	Plain data	Black box
Generic tools (editors, parsers, validators, query languages)	Ad hoc tools	Ad hoc tools
Easy to search	Ad hoc search	No search
Easy to manipulate	Ad hoc manipulation	No manipulation
Well-formed	Prone to syntax errors	Black box

XML Tools plugin

- XML Tools in BioNumerics: export/import of selected entries and experiment types

Database entries			
Key	Genus	Species	Strain number
G@Gel04@009	STANDARD		
→ G@Gel07@002	Ambiorix	sylvestris	52441
→ G@Gel07@003	Ambiorix	aberrans	52449
G@Gel07@004	Vercingetorix	palustris	42815
→ G@Gel07@005	Ambiorix	sylvestris	52416
G@Gel07@006	Ambiorix	sp.	52415
→ G@Gel07@007	Ambiorix	sylvestris	52414
G@Gel07@008	Ambiorix	sp.	52424

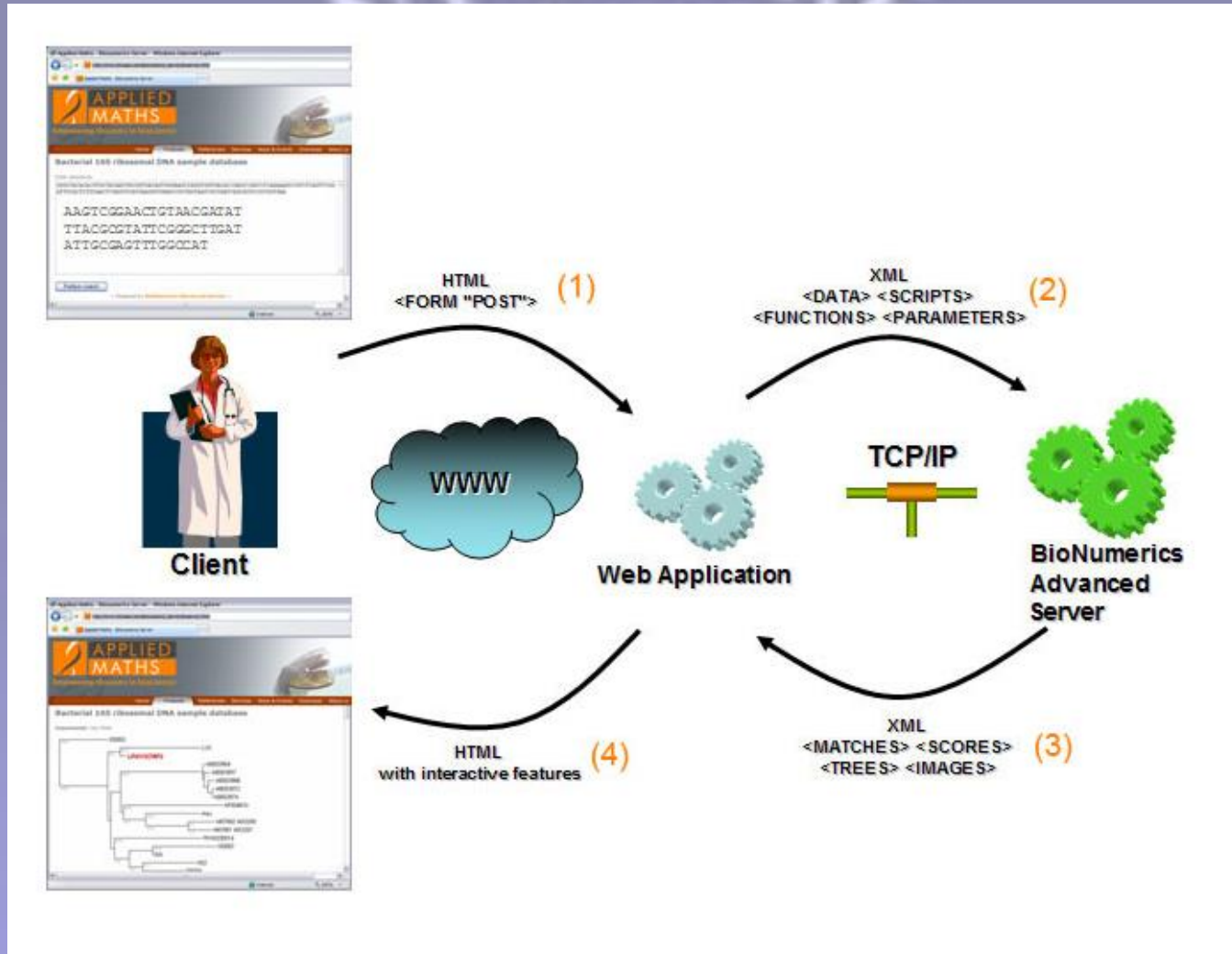


- Example: convert a local database to a connected database

Some XML technologies

- DTD and XML Schema to define XML structures and data types
- XSLT to transform XML data
- XSL-FO to describe the layout of a document
- SOAP to exchange XML data between applications
- WSDL to describe web services
- RDF to describe web resources
- XPath and XQuery to access XML data
- SMIL to define graphics

BioNumerics Server



Flexibility

