

R as a Web Service

Balasubramanian Narasimhan
Douglas J. Wood, Solomon Henry
Stanford University
Stanford, CA 94305

June 15, 2006

Introduction

❖ Introduction

- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

- R: a powerful, free, open-source, reliable, statistical computing environment; large community of developers; language used for prototyping much modern methodology

Introduction

❖ Introduction

- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

- R: a powerful, free, open-source, reliable, statistical computing environment; large community of developers; language used for prototyping much modern methodology
- Web: ubiquitous, convenient, simple-to-use infrastructure

Introduction

❖ Introduction

- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

- R: a powerful, free, open-source, reliable, statistical computing environment; large community of developers; language used for prototyping much modern methodology
- Web: ubiquitous, convenient, simple-to-use infrastructure

Web Service

- **W3C** A software system designed to support interoperable machine-to-machine interaction over a network Interface described in Web Services Description Language (WSDL)
- Interaction via messages (perhaps enclosed in a SOAP envelope)
- Transport layer is HTTP comprised of XML and other standards

Why R as a web service?

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

- Data often resides in a separate place from where R is available. People can find downloading, formatting, and setting up can become chores

Why R as a web service?

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

- Data often resides in a separate place from where R is available. People can find downloading, formatting, and setting up can become chores
- Many routine computations can be done using R (data quality checks, trends, recruitment breakdowns) and easily provided to users if we could have R as web service

Why R as a web service?

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

- Data often resides in a separate place from where R is available. People can find downloading, formatting, and setting up can become chores
- Many routine computations can be done using R (data quality checks, trends, recruitment breakdowns) and easily provided to users if we could have R as web service
- Large systems doing complex tasks can be built using distributed web services; these become necessary for high-throughput analyses in genomics, proteomics etc.

Why R as a web service?

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

- Data often resides in a separate place from where R is available. People can find downloading, formatting, and setting up can become chores
- Many routine computations can be done using R (data quality checks, trends, recruitment breakdowns) and easily provided to users if we could have R as web service
- Large systems doing complex tasks can be built using distributed web services; these become necessary for high-throughput analyses in genomics, proteomics etc.
- Initiatives such as Grid Computing, especially the caBIG (Cancer Bioinformatics Grid) initiative, make use of web services

Other Approaches

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ **Other Approaches**
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

R need not be used only as web service. Other possibilities are available.

- Linking R with your own C, Fortran code. There is a well-established API for calling R from these languages.

Other Approaches

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ **Other Approaches**
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

R need not be used only as web service. Other possibilities are available.

- Linking R with your own C, Fortran code. There is a well-established API for calling R from these languages.
- Embedding R in your favorite language; R has been embedded in Perl, Python, Java etc.

Other Approaches

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ **Other Approaches**
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

R need not be used only as web service. Other possibilities are available.

- Linking R with your own C, Fortran code. There is a well-established API for calling R from these languages.
- Embedding R in your favorite language; R has been embedded in Perl, Python, Java etc.
- CGI scripts can be used with R in a web server and R embedded inside Apache is also available. Educational use of R has been made with these technologies for teaching purposes.

Other Approaches

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ **Other Approaches**
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

R need not be used only as web service. Other possibilities are available.

- Linking R with your own C, Fortran code. There is a well-established API for calling R from these languages.
- Embedding R in your favorite language; R has been embedded in Perl, Python, Java etc.
- CGI scripts can be used with R in a web server and R embedded inside Apache is also available. Educational use of R has been made with these technologies for teaching purposes.

Embedding R in a language couples R to the language somewhat closer than what a web service suggests.

Other Approaches

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

R need not be used only as web service. Other possibilities are available.

- Linking R with your own C, Fortran code. There is a well-established API for calling R from these languages.
- Embedding R in your favorite language; R has been embedded in Perl, Python, Java etc.
- CGI scripts can be used with R in a web server and R embedded inside Apache is also available. Educational use of R has been made with these technologies for teaching purposes.

Embedding R in a language couples R to the language somewhat closer than what a web service suggests.

The lingua franca

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

The first question: what language will the web services speak?

The lingua franca

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

The first question: what language will the web services speak? Answer: XML (eXtensible Markup Language) XML is a (text-based) way of describing data. Here is an example:

```
<?xml version="1.0"?>
<data>
<array type="integer">
<values>
1 2 3 4 5 6 7 8 9 10
</values>
</array>
</data>
```

Key details: arbitrary but matching tags, attributes, free format. XML documents must be well-formed! In addition, they must also be valid. This leads to XML DTDs and Schemas.

XML Bindings

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

Many modern languages have XML Bindings.

XML Bindings

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca

- ❖ **XML Bindings**
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT

- Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

Many modern languages have XML Bindings.
A Binding allows the XML data to be treated as a native object in the language.

XML Bindings

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca

- ❖ **XML Bindings**
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT

- Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

Many modern languages have XML Bindings.
A Binding allows the XML data to be treated as a native object in the language.
For example, Java has JAXB (Java for XML Binding) which allows XML data to be treated as objects.
Python has PyXML that allows XML data to be treated as Python Objects.
With a binding, one can
unmarshal read XML data and create an object/methods to access the contents

XML Bindings

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca

- ❖ **XML Bindings**
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT

- Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

Many modern languages have XML Bindings.
A Binding allows the XML data to be treated as a native object in the language.

For example, Java has JAXB (Java for XML Binding) which allows XML data to be treated as objects.
Python has PyXML that allows XML data to be treated as Python Objects.
With a binding, one can

unmarshal read XML data and create an object/methods to access the contents
marshal write out an object as XML that others can consume

XML Bindings

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca

- ❖ **XML Bindings**
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT

- Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

Many modern languages have XML Bindings.
A Binding allows the XML data to be treated as a native object in the language.

For example, Java has JAXB (Java for XML Binding) which allows XML data to be treated as objects.
Python has PyXML that allows XML data to be treated as Python Objects.
With a binding, one can

unmarshal read XML data and create an object/methods to access the contents
marshal write out an object as XML that others can consume

Important: Bindings are based on an XML DTD/Schema

R Data XML Binding

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings

❖ R Data XML Binding

- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT

- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

R has a package called `StatDataML` that describes some common data types that it understands: vectors, lists, matrices.

Creating StatDataML

```
library(StatDataML)
x <- c(1,2,3)
writeSDML(x, file="int.xml")
```

useR! 2006 Vienna 2006/06/15 – 7 / 23

R Data Schema (contd.)

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding

❖ R Data Schema (contd.)

- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT

- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

```
<?xml version="1.0"?>
<StatDataML xmlns="http://www.omegahat.org/StatDataML/">
<description>
<title>x</title>
<source>R</source>
<date> Wed Jun 7 23:50:58 2006 </date>
<version> </version>
<comment> </comment>
<creator>R-2.3.0</creator>
</description>
<dataset>
<array>
<dimension>
<dim size="3"></dim>
</dimension>
<type>
<numeric><integer><min>1</min><max>3</max></integer></numeric>
</type>
<textdata>
1 2 3
</textdata>
</array>
</dataset>
</StatDataML>
```

useR! 2006 Vienna 2006/06/15 – 8 / 23

R Data Schema (contd.)

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding

❖ R Data Schema (contd.)

- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT

- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

To read the XML back in:

```
library(StatDataML)
x = readSDML("int.xml")
x
[1] 1 2 3
```

useR! 2006 Vienna 2006/06/15 – 9 / 23

StatDataML DTD

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding

❖ StatDataML DTD

- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT

- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

```
<!-- StatDataML DTD version="1.0" -->
<ELEMENT StatDataML (description?, dataset?)>
<!ATTLIST StatDataML xmlns CDATA #FIXED "http://www.omegahat.org/StatDataML/">
<!-- document description tags -->
<ELEMENT description (title?, source?, date?, version?,
comment?, creator?, properties?)>
<ELEMENT title (#PCDATA)>
<ELEMENT source (#PCDATA)>
<ELEMENT date (#PCDATA)>
<ELEMENT version (#PCDATA)>
<ELEMENT comment (#PCDATA)>
<ELEMENT creator (#PCDATA)>
<ELEMENT properties (list)>
<!-- basic elements -->
<ELEMENT dataset (list | array)>
<ELEMENT list (dimension, properties?, listdata)>
<ELEMENT listdata (list | array | empty)*>
<ELEMENT empty EMPTY>
<ELEMENT array (dimension, type, properties?, (data | textdata))>
<!-- dimension elements -->
<ELEMENT dimension (dim*)>
<ELEMENT dim (e*)>
<!ATTLIST dim size CDATA #REQUIRED>
<!ATTLIST dim name CDATA #IMPLIED>
```

useR! 2006 Vienna 2006/06/15 – 10 / 23

Other StatDataML Bindings

To appear on Omegahat:

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ **Other StatDataML Bindings**
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

Other StatDataML Bindings

To appear on Omegahat:

JStatDataML Java StatDataML binding

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ **Other StatDataML Bindings**
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

Other StatDataML Bindings

To appear on Omegahat:

JStatDataML Java StatDataML binding
PyStatDataML Python StatDataML binding

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ **Other StatDataML Bindings**
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

Other StatDataML Bindings

To appear on Omegahat:

JStatDataML Java StatDataML binding
PyStatDataML Python StatDataML binding

Java can grok StatDataML.

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ **Other StatDataML Bindings**
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

Other StatDataML Bindings

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

To appear on Omegahat:

JStatDataML Java StatDataML binding
PyStatDataML Python StatDataML binding

Java can grok StatDataML.

- In Java, you can create R data objects such as lists, vectors, data frames, matrices.

Other StatDataML Bindings

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

To appear on Omegahat:

JStatDataML Java StatDataML binding
PyStatDataML Python StatDataML binding

Java can grok StatDataML.

- In Java, you can create R data objects such as lists, vectors, data frames, matrices.
- These objects can be *marshalled* into XML.

Other StatDataML Bindings

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

To appear on Omegahat:

JStatDataML Java StatDataML binding
PyStatDataML Python StatDataML binding

Java can grok StatDataML.

- In Java, you can create R data objects such as lists, vectors, data frames, matrices.
- These objects can be *marshalled* into XML.
- In Java, you are able to read StatDataML and represent them as pedigreed Java objects. In other words, StatDataML can be *unmarshalled* into Java objects.

Other StatDataML Bindings

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

To appear on Omegahat:

JStatDataML Java StatDataML binding
PyStatDataML Python StatDataML binding

Java can grok StatDataML.

- In Java, you can create R data objects such as lists, vectors, data frames, matrices.
- These objects can be *marshalled* into XML.
- In Java, you are able to read StatDataML and represent them as pedigreed Java objects. In other words, StatDataML can be *unmarshalled* into Java objects.

You can do the same in Python, but the focus here is Java.

JStatDataML

- Introduction
- Why R as a web service?
- Other Approaches
- The lingua franca
- XML Bindings
- R Data XML Binding
- R Data Schema (contd.)
- R Data Schema (contd.)
- StatDataML DTD
- Other StatDataML Bindings
- JStatDataML**
- API
- API (contd.)
- API (contd.)
- Transformations
- XSLT Language
- XSLT Example Input
- XSLT Transform
- XSLT Output
- R as an XSLT Transform
- R as a web service
- A pipeline
- Continuing work

- JStatDataML is layered on top of StatDataML so that any changes to StatDataML can be easily reworked

JStatDataML

- Introduction
- Why R as a web service?
- Other Approaches
- The lingua franca
- XML Bindings
- R Data XML Binding
- R Data Schema (contd.)
- R Data Schema (contd.)
- StatDataML DTD
- Other StatDataML Bindings
- JStatDataML**
- API
- API (contd.)
- API (contd.)
- Transformations
- XSLT Language
- XSLT Example Input
- XSLT Transform
- XSLT Output
- R as an XSLT Transform
- R as a web service
- A pipeline
- Continuing work

- JStatDataML is layered on top of StatDataML so that any changes to StatDataML can be easily reworked
- Handles missing values, factors etc.
- Helper Java classes are available to take results of a database query and marshall it into StatDataML for R computations

JStatDataML

- Introduction
- Why R as a web service?
- Other Approaches
- The lingua franca
- XML Bindings
- R Data XML Binding
- R Data Schema (contd.)
- R Data Schema (contd.)
- StatDataML DTD
- Other StatDataML Bindings
- JStatDataML**
- API
- API (contd.)
- API (contd.)
- Transformations
- XSLT Language
- XSLT Example Input
- XSLT Transform
- XSLT Output
- R as an XSLT Transform
- R as a web service
- A pipeline
- Continuing work

- JStatDataML is layered on top of StatDataML so that any changes to StatDataML can be easily reworked
- Handles missing values, factors etc.
- Helper Java classes are available to take results of a database query and marshall it into StatDataML for R computations
- Simple Application Programming Interface that is easy to use

JStatDataML

- Introduction
- Why R as a web service?
- Other Approaches
- The lingua franca
- XML Bindings
- R Data XML Binding
- R Data Schema (contd.)
- R Data Schema (contd.)
- StatDataML DTD
- Other StatDataML Bindings
- JStatDataML**
- API
- API (contd.)
- API (contd.)
- Transformations
- XSLT Language
- XSLT Example Input
- XSLT Transform
- XSLT Output
- R as an XSLT Transform
- R as a web service
- A pipeline
- Continuing work

- JStatDataML is layered on top of StatDataML so that any changes to StatDataML can be easily reworked
- Handles missing values, factors etc.
- Helper Java classes are available to take results of a database query and marshall it into StatDataML for R computations
- Simple Application Programming Interface that is easy to use

Caveat: The StatDataML data types represent a subset of R data objects. In particular, you don't want to use this for large expression arrays, proteomics data etc.

API

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ **API**
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

All Classes

Packages

- [edu.stanford.hrpdcc.xml](#)
- [edu.stanford.hrpdcc.xml.r](#)

Overview Package Class T

PREV NEXT

Packages

- [edu.stanford.hrpdcc.xml](#)
- [edu.stanford.hrpdcc.xml.r](#)

API (contd.)

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ **API (contd.)**
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

Array

Categorical

Ce

Character

Complex

Data

Dataset

Datetime

Description

Dim

Dimension

E

Empty

F

I

Integer

Label

List

Listdata

Logical

Max

Min

Na

Nan

Neginf

Numeric

ObjectFactory

Posinf

Properties

R

Real

StatDataML

T

Textdata

Type

API (contd.)

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ **API (contd.)**
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

Package edu.stanford.hrpdcc.xml.r

Interface Summary

[RDataSource](#) An interface the encapsulates an R Data source

Class Summary

AbstractRDataSource	A class that most R Data source imple
DataFrame	A class to encapsulate R DataFrame o
Factor	A class to encapsulate String vectors t 2006
IntegerVector	A class to encapsulate R integer vecto
List	A class to encapsulate R lists Created:
Matrix	A class to encapsulate R matrices Cre
RDBSource	A class encapsulating an RDataSource 2 11:13:42 2006
RealVector	A class to encapsulate R real vectors C
StringVector	A class to encapsulate String vectors C
TestDBSource	A Test data source that to use for testi
TestRDataSource	An RDataSource for testing purposes C

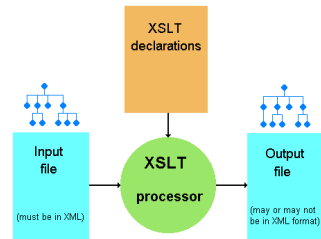
Transformations

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ **Transformations**
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

Transformations are a general approach to convert XML data to convert to other formats (including XML). They are specified by means of a language, XSLT (Extensible Stylesheet Language Transformations), an XML-based language used for the transformation of XML documents. The original document is untouched and the output is a transformed document that can be converted to HTML, plain text, or PDF.

Transformations

Transformations are a general approach to convert XML data to convert to other formats (including XML). They are specified by means of a language, XSLT (Extensible Stylesheet Language Transformations), an XML-based language used for the transformation of XML documents. The original document is untouched and the output is a transformed document that can be converted to HTML, plain text, or PDF.



useR! 2006 Vienna 2006/06/15 – 16 / 23

XSLT Language

- Declarative
- Applies transform to nodes that meet conditions
- Results are added to result tree
- Nodes not matching conditions are left untouched

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT

- Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

useR! 2006 Vienna 2006/06/15 – 17 / 23

XSLT Example Input

```
<?xml version="1.0"?>
<persons>
  <person username="naras">
    <name>Balasubramanian</name>
    <family-name>Narasimhan</family-name>
  </person>
  <person username="sherry">
    <name>Solomon</name>
    <family-name>Henry</family-name>
  </person>
</persons>
```

useR! 2006 Vienna 2006/06/15 – 18 / 23

XSLT Transform

```
<?xml version="1.0"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
  version="1.0">
  <xsl:output method="xml" indent="yes"/>
  <xsl:template match="/">
    <transform>
      <xsl:apply-templates/>
    </transform>
  </xsl:template>
  <xsl:template match="person">
    <record>
      <username>
        <xsl:value-of select="@username" />
      </username>
      <name>
        <xsl:value-of select="name" />
      </name>
    </record>
  </xsl:template>
</xsl:stylesheet>
```

- ❖ Introduction
- ❖ Why R as a web service?
- ❖ Other Approaches
- ❖ The lingua franca
- ❖ XML Bindings
- ❖ R Data XML Binding
- ❖ R Data Schema (contd.)
- ❖ R Data Schema (contd.)
- ❖ StatDataML DTD
- ❖ Other StatDataML Bindings
- ❖ JStatDataML
- ❖ API
- ❖ API (contd.)
- ❖ API (contd.)
- ❖ Transformations
- ❖ XSLT Language
- ❖ XSLT Example Input
- ❖ XSLT Transform
- ❖ XSLT Output
- ❖ R as an XSLT

- Transform
- ❖ R as a web service
- ❖ A pipeline
- ❖ Continuing work

useR! 2006 Vienna 2006/06/15 – 19 / 23

XSLT Output

- Introduction
- Why R as a web service?
- Other Approaches
- The lingua franca
- XML Bindings
- R Data XML Binding
- R Data Schema (contd.)
- R Data Schema (contd.)
- StatDataML DTD
- Other StatDataML Bindings
- JStatDataML
- API
- API (contd.)
- API (contd.)
- Transformations
- XSLT Language
- XSLT Example Input
- XSLT Transform
- XSLT Output**
- R as an XSLT Transform
- R as a web service
- A pipeline

```
<?xml version="1.0" encoding="UTF-8"?>
<transform>
<record>
<username>naras</username>
<name>Balasubramanian</name>
</record>
<record>
<username>shenry</username>
<name>Solomon</name>
</record>
</transform>
```

useR! 2006 Vienna 2006/06/15 - 20 / 23

R as an XSLT Transform

- Introduction
- Why R as a web service?
- Other Approaches
- The lingua franca
- XML Bindings
- R Data XML Binding
- R Data Schema (contd.)
- R Data Schema (contd.)
- StatDataML DTD
- Other StatDataML Bindings
- JStatDataML
- API
- API (contd.)
- API (contd.)
- Transformations
- XSLT Language
- XSLT Example Input
- XSLT Transform
- XSLT Output
- R as an XSLT Transform**
- R as a web service
- A pipeline

Now imagine if you can extend XSLT to do R computations. Enter Sxslt, an omegahat project which is an extension of XSLT to R.

This allows us to do processing such as:

```
<xsl:template match="sqrt">
<xsl:value-of select="r:call('sqrt', number())"/>
</xsl:template>
<xsl:template match="lm">
<i class="output">
<xsl:value-of select="r:call('myfun', string(@y), string(@x))"/>
</i>
</xsl:template>
```

In an xml document, if Sxslt sees a tag like:

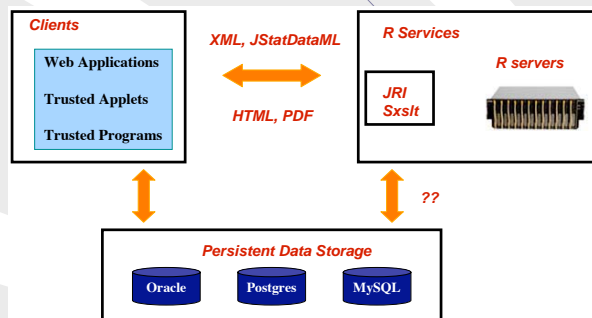
```
<lm y="creatine" x="predictors" />
```

it will replace it with the value of the R function myfun taking y and x as arguments.

useR! 2006 Vienna 2006/06/15 - 21 / 23

A pipeline

- Introduction
- Why R as a web service?
- Other Approaches
- The lingua franca
- XML Bindings
- R Data XML Binding
- R Data Schema (contd.)
- R Data Schema (contd.)
- StatDataML DTD
- Other StatDataML Bindings
- JStatDataML
- API
- API (contd.)
- API (contd.)
- Transformations
- XSLT Language
- XSLT Example Input
- XSLT Transform
- XSLT Output
- R as an XSLT Transform
- R as a web service**
- A pipeline



useR! 2006 Vienna 2006/06/15 - 22 / 23

Continuing work

- Introduction
- Why R as a web service?
- Other Approaches
- The lingua franca
- XML Bindings
- R Data XML Binding
- R Data Schema (contd.)
- R Data Schema (contd.)
- StatDataML DTD
- Other StatDataML Bindings
- JStatDataML
- API
- API (contd.)
- API (contd.)
- Transformations
- XSLT Language
- XSLT Example Input
- XSLT Transform
- XSLT Output
- R as an XSLT Transform
- R as a web service**
- A pipeline

- Being deployed in small settings at Stanford
- Technical issues being worked out
- RSXMLObjects is an omegahat project for serializing general R objects to XML. These might be more general than StatDataML
- Integration of this web services into a Struts framework that is widely used at the DCC and elsewhere
- Security issues although there are good solutions available
- Strengthening the technology for enterprise use

If you are interested in playing with these technologies, I can provide you some downloads.

useR! 2006 Vienna 2006/06/15 - 23 / 23