







As an example, consider the type specification for the integers from 1 to 10:

```
<type>
  <numeric>
    <integer>
      <min>1</min> <max>10</max>
    <integer/>
  <numeric/>
</type>
```

The content of 'min' max <neg-inf> respectively. data

If 'data

<e> </e>ce

---



---

<textdata na.string="N/A" null.string="EMPTY">

...continue in MATLAB ...

---

< M A T L A B >  
Copyright 1984-2000 The MathWorks, Inc.  
Version 6.0.0.88 Release 12



## A session started with R ...

---

```
R : Copyright 2002, The R Development Core Team  
Version 1.6.1 (2002-11-01)
```

```
## load the StatDataML package  
> library(StatDataML)  
Loading required package: XML
```

```
## Create a simple data structure  
> X <- list(matrix(c(1, 2, 3, 4), 2, 2))  
> X[[2]] <- 12 + 3i  
> X[[3]] <- "Test"  
> X[[4]] <- list(a = "Test 2", b = 33.44)  
> dim(X) <- c(2, 2)
```

```
## Show what we got  
> X  
      [,1]      [,2]  
[1,] "Numeric,4" "Character,1"  
[2,] "Complex,1" "List,2"
```

```
> X[[1,1]]  
      [,1] [,2]  
[1,]    1    3  
[2,]    2    4
```

```
> X[[1,2]]  
[1] "Test"
```

```
> X[[2,1]]  
[1] 12+3i
```

...and finished in MATLAB:

---

< M A T L A B >  
Copyright 1984-2000 The MathWorks, Inc.  
Version 6.0.0.88 Release 12





### 3 Implementation

Currently we have support for R, MATLAB and Octave, and converters for SPSS and Gnumeric are under development. The complete package including all implementations can be found at the homepage of the “Omegahat” project, which aims at providing a variety of open-source software for statistical applications (focusing on web-based software, Java, the Java virtual machine, and distributed computing): <http://www.omegahat.org/StatDataML/>.

The software for the R system alone is also provided as an R package from the software archive of the R project for statistical computing (<http://www.R-project.org/>).

## Appendix: the *StatDataML* .dtd file

```
<!-- StatDataML DTD version="1.0" -->

<!ELEMENT StatDataML (description?, dataset?)>
<!-- document description tags -->
<!-- ATTLIST StatDataML xmlns CDATA #FIXED "http://www.omegahat.org/StatDataML/" -->

<!-- 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)>

<!-- list tags -->
<!ELEMENT list (dimension, properties?, listdata)>
<!-- listdata tags -->
<!ELEMENT listdata (list | array | empty)*>
<!ELEMENT empty EMPTY>

<!-- array tags -->
<!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 -->

<!-- type elements -->
<!ELEMENT type (logical | categorical | numeric | character | datetime)>
<!ELEMENT logical EMPTY>
```

