

Appendix

Sample fragments from the XML documents

xmlbenchmark.xml

```
<?xml version="1.0" standalone="yes"?>
<site>
<regions>
<africa>
<item id="item0">
<location>United States</location>
<quantity>1</quantity>
<name>duteous nine eighteen </name>
<payment>Creditcard</payment>
<description>
<parlist>
<listitem>
<text>
<keyword> officer embrace such fears distinction attires </keyword>
</text>
</listitem>
<listitem>
<text>
wounds gum merely raise red breaks earth god folds closet captain dying reek
</text>
</listitem>
</parlist>
</description>
<shipping>Will ship internationally, See description for charges</shipping>
<incategory category="category540"/>
<mailbox>
<mail>
<from>Libero Rive mailto:Rive@hitachi.com</from>
<to>Benedikte Glew mailto:Glew@sds.no</to>
<date>07/05/2000</date>
<text>
<keyword> reprieve </keyword> enforcement <emph> armed </emph> gold
</text>
</mail>
</mailbox>
</item>
```

conferences.xml

```
<proceedings key="95">
  <editor>Max J. Egenhofer</editor>
  <editor>John R. Herring</editor>
  <title>Advances in Spatial Databases, 4th International Symposium
    SSD'95, Portland, Maine,
    USA, August 6-9, 1995, Proceedings
  </title>
  <series href="db/journals/lncs.html"></series>
  <volume>951</volume>
  <publisher>Springer</publisher>
  <year>1995</year>
  <isbn>3-540-60159-7</isbn>
  <url>db/conf/ssd/ssd95.html</url>
</proceedings>
<inproceedings key="KuijpersPB95">
  <author>Bart Kuijpers</author>
  <author>Jan Paredaens</author>
  <author>Jan Van den Bussche</author>
  <title>Lossless Representation of Topological Spatial Data.</title>
  <pages>1-13</pages>
  <booktitle>SSD</booktitle>
  <year>1995</year>
  <url>db/conf/ssd/ssd95.html#KuijpersPB95</url>
</inproceedings>
```

journals.xml

```
<article key="DavisSY82a">
  <title>Introduction</title>
  <author>Martin Davis</author>
  <author>Carl Smith</author>
  <author>Paul Young</author>
  <pages>1</pages>
  <journal>Information and Control</journal>
  <month>January</month>
  <year>1982</year>
  <volume>52</volume>
  <number>1</number>
  <url>http://theory.lcs.mit.edu/~iandc/ic82.html</url>
</article>
```

shakespeare.xml

```
<PLAY>
  <TITLE>The Tragedy of Antony and Cleopatra</TITLE>
  <FM>
    <P>Text placed in the public domain by Moby Lexical Tools.</P>
    <P>SGML markup by Jon Bosak, 1992-1994.</P>
    <P>XML version by Jon Bosak, 1996-1998.</P>
    <P>This work may be freely copied and distributed worldwide.</P>
  </FM>
  <PERSONAE>
    <TITLE>Dramatis Personae</TITLE>
    <PGROUP>
      <PERSONA>MARK ANTONY</PERSONA>
      <PERSONA>OCTAVIUS CAESAR</PERSONA>
      <PERSONA>M. AEMILIUS LEPIDUS</PERSONA>
      <GRPDESCR>triumvirs.</GRPDESCR>
    </PGROUP>
    <PERSONA>SEXTUS POMPEIUS</PERSONA>
    <PGROUP>
      <PERSONA>DOMITIUS ENOBARBUS</PERSONA>
      <PERSONA>VENTIDIUS</PERSONA>
      <PERSONA>EROS</PERSONA>
      <PERSONA>SCARUS</PERSONA>
      <PERSONA>DERCETAS</PERSONA>
      <PERSONA>DEMETRIUS</PERSONA>
      <PERSONA>PHILO</PERSONA>
      <GRPDESCR>friends to Antony.</GRPDESCR>
    </PGROUP>
  </PERSONAE>
....
```

xpress.xml

```
<FMDataBase>
<FCCAmRadio>
<CallLetters>KB2RRD</CallLetters>
<LicenseNo>9409072814</LicenseNo>
<Flags1>A</Flags1>
<Flags2>I</Flags2>
<Name>
  <LastName>COLLADO</LastName>
  <FirstName>RAFAEL</FirstName>
  <MiddleName>P</MiddleName>
</Name>
<Address>
  <Addr1>120 GRATTON ST APT 2L</Addr1>
<City>BROOKLYN</City>
  <State>NY</State>
  <Zip>11237</Zip>
</Address>
<LicValid>
  <BeginDate>09/09/1994</BeginDate>
  <EndDate>09/09/2004</EndDate>
  <InitDate>09/09/1994</InitDate>
</LicValid>
<Info>SYS</Info>
<Record>134001</Record>
</FCCAmRadio>
```

student1.xml

```
<STUDENT SRNO="800000000">  
  <NAME> envious nameable ungrammatical </NAME>  
  <YEAR> 1999 </YEAR>  
  <PROG> M.Tech. </PROG>  
  <DEPT name="Metallurgy"/>  
</STUDENT>
```

student4.xml

```
<STUDENTLIST><STUDENT SRNO="800000000">  
  <NAME> envious nameable ungrammatical </NAME>  
  <YEAR year="1999"/>  
  <PROG pname="Master_of_Engineering"/>  
  <DEPT dname="Biochemistry"/>  
  <PREVSCHOOL sname="Indian_Institute_of_Technology_Guwahati"/>  
</STUDENT>
```