

Erlsom

Erlang XML processing library

XML

```
<?xml version="1.0"?>
<x:books xmlns:x="urn:books">
  <book id="bk001">
    <author>Writer</author>
    <title>The First Book</title>
    <genre>Fiction</genre>
    <price>44.95</price>
    <pub_date>2000-10-01</pub_date>
    <review>An amazing story of nothing.</review>
  </book>

  <book id="bk002">
    <author>Poet</author>
    <title>The Poet's First Poem</title>
    <genre>Poem</genre>
    <price>24.95</price>
    <pub_date>1990-10-01</pub_date>
    <review>Least poetic poems.</review>
  </book>

  ...
</x:books>
```

XSD - XML schema

```
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  targetNamespace="urn:books"
  xmlns:bks="urn:books">

  <xsd:element name="books" type="bks:BooksForm"/>

  <xsd:complexType name="BooksForm">
    <xsd:sequence>
      <xsd:element name="book" type="bks:BookForm" minOccurs="0" maxOccurs="unbounded"/>
    </xsd:sequence>
  </xsd:complexType>

  <xsd:complexType name="BookForm">
    <xsd:sequence>
      <xsd:element name="author" type="xsd:string"/>
      <xsd:element name="title" type="xsd:string"/>
      <xsd:element name="genre" type="xsd:string"/>
      <xsd:element name="price" type="xsd:float"/>
      <xsd:element name="pub_date" type="xsd:date"/>
      <xsd:element name="review" type="xsd:string"/>
    </xsd:sequence>
    <xsd:attribute name="id" type="xsd:string"/>
  </xsd:complexType>
</xsd:schema>
```

Erlsom modes

- SAX parser
- DOM parser
- 'data-binder'

SAX (Simple API for XML) mode

```
1> erlsom:parse_sax(Xml, [], fun(Event, Acc) -> io:format("~p~n", [Event]), Acc end).
startDocument
{processingInstruction,"xml"," version=\"1.0\""}
{startPrefixMapping,"x","urn:books"}
{startElement,"urn:books","books","x",[]}
{ignorableWhitespace,"\n "}
{startElement,[],"book",[],[{attribute,"id",[],[],"bk001"}]}
{ignorableWhitespace,"\n "}
{startElement,[],"author",[],[]}
{characters,"Writer"}
{endElement,[],"author",[]}
{ignorableWhitespace,"\n "}
{startElement,[],"title",[],[]}
{characters,"The First Book"}
{endElement,[],"title",[]}
{ignorableWhitespace,"\n "}
{startElement,[],"genre",[],[]}
{characters,"Fiction"}
{endElement,[],"genre",[]}
{ignorableWhitespace,"\n "}
{startElement,[],"price",[],[]}
{characters,"44.95"}
{endElement,[],"price",[]}
{ignorableWhitespace,"\n "}
{startElement,[],"pub_date",[],[]}
{characters,"2000-10-01"}
...
{endElement,"urn:books","books","x"}
{endPrefixMapping,"x"}
endDocument
{ok,[],[]}
```

DOM (Document Object Model) mode

```
2> erlsom:simple_form(Xml).
{ok,{{"urn:books"}books",[],
 [{"book",
  [{"id","bk001"}],
  [{"author",[],["Writer"]},
   {"title",[],["The First Book"]},
   {"genre",[],["Fiction"]},
   {"price",[],["44.95"]},
   {"pub_date",[],["2000-10-01"]},
   {"review",[],["An amazing story of nothing."]}]},
 {"book",
  [{"id","bk002"}],
  [{"author",[],["Poet"]},
   {"title",[],["The Poet's First Poem"]},
   {"genre",[],["Poem"]},
   {"price",[],["24.95"]},
   {"pub_date",[],["1990-10-01"]},
   {"review",[],["Least poetic poems."]}]},
 {"book",
  [{"id","bk003"}],
  [{"author",[],["Poet"]},
   {"title",[],["The Poet's Second Poem"]},
   {"genre",[],["Poem"]},
   {"price",[],["25.95"]},
   {"pub_date",[],["1991-10-01"]},
   {"review",[],["Least poetic poems."]}]},
 ...
 ]}]}
```

'data-binder' mode

```
3> {ok, Model} = erlsom:compile_xsd_file("books.xsd").
{ok,{model,[{type...
```

```
4> {ok, Result, _} = erlsom:scan(Xml, Model).
{ok,{#BooksForm,[],
  [{#BookForm,[],"bk001","Writer","The First Book","Fiction",
    "44.95","2000-10-01","An amazing story of nothing."},
   {#BookForm,[],"bk002","Poet","The Poet's First Poem",
    "Poem","24.95","1990-10-01","Least poetic poems."},
   {#BookForm,[],"bk003","Poet","The Poet's Second Poem",
    "Poem","25.95","1991-10-01","Least poetic poems."},
   {#BookForm,[],"bk004","Poet","The Poet's Third Poem",
    "Poem","14.95","1994-10-01","Least poetic poems."},
   {#BookForm,[],"bk005","Poet","The Poet's Fourth Poem",
    "Poem","34.95","1995-10-01","Least poetic poems."},
   {#BookForm,[],"bk006","Poet","The Poet's Fifth Poem",
    "Poem","64.95","1999-10-01","Least poetic poems."}]},
  "",\n"}
```

```
Books = Result#BooksForm.book,
[Book1 | _] = Books,
Book1_title = Book1#BookForm.title.
```

XML

```
<?xml version="1.0" encoding="UTF-8"?>
<measurements>
  <element>
    <data>
      <no2>
        <precision>1.57776563E15</precision>
        <value>1.17896127E15</value>
      </no2>
      <no2_strat>
        <precision>2.00000001E14</precision>
        <value>1.7461823E15</value>
      </no2_strat>
      <no2_trop>
        <precision>1.63598203E15</precision>
        <value>-5.6722103E14</value>
      </no2_trop>
    </data>
    <location>
      <latitude>0.0633</latitude>
      <longitude>9.8805</longitude>
    </location>
    <time>2016-12-31T12:55:57Z</time>
  </element>
  <element>
    <data>
      ...
    </element>
  </measurements>
```


XSD - XML schema

```
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <xsd:element name="measurements" type="measurements"/>
  <xsd:complexType name="measurements">
    <xsd:sequence>
      <xsd:element name="element" maxOccurs="unbounded" type="element"/>
    </xsd:sequence>
  </xsd:complexType>

  <xsd:complexType name="element">
    <xsd:sequence>
      <xsd:element name="data" type="data"/>
      <xsd:element name="location" type="location"/>
      <xsd:element name="time" type="xsd:dateTime"/>
    </xsd:sequence>
  </xsd:complexType>

  <xsd:complexType name="data">
    <xsd:sequence>
      <xsd:element name="no2" type="no2"/>
      <xsd:element name="no2_strat" type="no2"/>
      <xsd:element name="no2_trop" type="no2"/>
    </xsd:sequence>
  </xsd:complexType>

  <xsd:complexType name="no2">
    <xsd:sequence>
      <xsd:element name="precision" type="xsd:double"/>
      <xsd:element name="value" type="xsd:double"/>
    </xsd:sequence>
  </xsd:complexType>

  <xsd:complexType name="location">
    <xsd:sequence>
      <xsd:element name="latitude" type="xsd:double"/>
      <xsd:element name="longitude" type="xsd:double"/>
    </xsd:sequence>
  </xsd:complexType>
</xsd:schema>
```