SBOL Serialization

Chris J. Myers

University of Utah

SBOL Workshop 2012

January 6, 2012

libSBOL: Current Status

- SBOL core data model has been implemented as a java object.
- Special thanks to Allan Kuchinsky, Trevor Smith, and Agilent for their support of this effort.
- Initial version has partial support for both an XML as well as an XML/RDF serialization.
- XML serialization is currently used by tools from Boston University, Newcastle, and the University of Utah.

libSBOL: Current Status

- SBOL core data model has been implemented as a java object.
- Special thanks to Allan Kuchinsky, Trevor Smith, and Agilent for their support of this effort.
- Initial version has partial support for both an XML as well as an XML/RDF serialization.
- XML serialization is currently used by tools from Boston University, Newcastle, and the University of Utah.
- **This has enabled the first SBOL v1.0.0 data exchange from Newcastle's Parts Repository to Utah's iBioSim to BU's Clotho.**

XML Serialization

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<Collection description="This was generated during the test run."
    name="Complete Collection" displayId="collection-id-1234" uri="/col/1">
    <component description="A Component" name="DnaComponent 4321"
        displayId="component-id-4321" uri="/comp/1"/>
    <component description="A Component" name="DnaComponent 4444"
        displayId="component-id-4444" uri="/test/2">
        <sequence uri="/seq/1">
            <nucleotides>gataca</nucleotides>
        </sequence>
        <annotation subComponentURI="/test/2" bioEnd="50" bioStart="1"
            uri="/anno/1"/>
        <annotation subComponentURI="/test/2" strand="+" bioEnd="900"
            bioStart="50" uri="/anno/2">
            <precede uri="/anno/1"/>
            <precede uri="/anno/2"/>
        </annotation>
    </component>
</Collection>
```

XML/RDF Serialization

```
RDF NAMESPACES
<rdf:RDF>
<s:DnaComponent rdf:about="http://partsregistry.org/part/BBa_T9002">
  <s:displayId>BBa_T9002</s:displayId>
  <rdfs:label>T9002</rdfs:label>
  <rdfs:comment>GFP Producer Controlled by 3OC6HSL Receiver Device</rdfs:comment>
  <rdf:type rdf:resource="http://partsregistry.org/type/signalling"/>
  <s:dnaSequence>
    <s:DnaSequence rdf:about="http://partsregistry.org/seq/partseq_5591">
      <s:nucleotides>
        tccctatcagtgatagagattgacatccctatcagtgatagagatact...
      </s:nucleotides>
    </s:DnaSequence>
  </s:dnaSequence>
  <s:annotation>
    <s:SequenceAnnotation rdf:about="http://partsregistry.org/anot/an_5591_187_1">
      <s:precedes rdf:resource=""/>
      <s:subComponent>
        <s:DnaComponent rdf:about="http://partsregistry.org/part/BBa_R0040">
          <s:displayId>BBa_R0040</s:displayId>
          <rdfs:label>p(tetR)</rdfs:label>
          <rdfs:comment>TetR repressible promoter</rdfs:comment>
          <rdf:type rdf:resource="http://purl.obolibrary.org/obo/SO_0005836"/>
        </s:DnaComponent>
      </s:subComponent>
    </s:SequenceAnnotation>
  </s:annotation>
...
```

The Grand Unification

- XML and XML/RDF serializations are really not that different.
- Therefore, it would not be too difficult to adapt the XML serialization routines to produce an XML file in an RDF style (thanks to Evren Sirin of Clark & Parsia for agreeing to do this adaptation).
- Advantages:
  - RDF allows nesting without creating explicit child relationships.
  - Use of RDF style would enable RDF query tools to be used on these files.
  - SBOL would have a single standard serialization.