

## *more* HOBIT at the BiBiServ

Jan Krüger    Henning Mersch

Bielefeld Bioinformatics Service  
Institute of Bioinformatics  
CeBiTec Bielefeld University

[jkrueger@TechFak.Uni-Bielefeld.DE](mailto:jkrueger@TechFak.Uni-Bielefeld.DE)  
[h.mersch@fz-juelich.DE](mailto:h.mersch@fz-juelich.DE)

Berlin, October 2005



# Overview

<http://bibiserv.techfak.uni-bielefeld.de/hobit/>

- 1 Overview
- 2 Grammars
- 3 WebServices
- 4 Remarks



# SequenceML.xsd

<http://schemas.bioservices.net/2005/biotypes>

- Provides a representation of sequences like FASTA
- Based on Martin's *BioTypes*
- supported by *BioDOM*<sup>1</sup>

---

<sup>1</sup><http://biodom.sf.net>



# AlignmentML.xsd

<http://schemas.bioservices.net/2005/biotypes>

- Provides a representation of Alignments of annotated sequences
- Based on Martin's *BioTypes*, but extends them
- supported by *BioDOM* <sup>2</sup>

---

<sup>2</sup><http://biodom.sf.net>



# RNAML

<http://www-lbit.iro.umontreal.ca/rnaml/current/rnaml.xsd>

- a standard syntax for exchanging RNA information
- Schema and DTD available - no unique namespace
- API's for C,C++ and Java - last activities in 2002
- supported by *BioDOM*<sup>3</sup>

---

<sup>3</sup><http://biodom.sf.net>



# EBIApplicationResult

<http://www.ebi.ac.uk/schema/ApplicationResult.xsd>

- developed from EBI and supported by some Webservice located at the EBI mainly blast, fasts, ...
- provides representation for all blast like output
- some strange non optional tags/attributes



# PhyloML

- provides a representation of phylogenetic trees
- from phylip package (readtree)
- until now there is no schema or dtd
- supported by *BioDOM*<sup>4</sup>
- developed by Alexander Kaiser -  
akaiser@techfak.uni-bielefeld.de

---

<sup>4</sup><http://biodom.sf.net>



# Project: BioDOM

<http://biodom.sf.net>

- a Java library providing easy conversion between XML and (proprietary) ASCII formats
- also available as Webservice
- supported existing format: SequenceML, AlignmentML, RNAML and PhyloML
- not yet supported : EBIApplicationResult
- ...





# List of available Webservice

<b>TOOL</b>	<b>INPUT</b>	<b>OUTPUT</b>	<b>URL</b> <sup>5</sup>
REPuter	FASTA	EBIApplication	/reputer
e2g	FASTA	EBIApplication	/e2g
DCA(uc)	SequenceML	AlignmentML	/dca
Dialign	SequenceML	AlignmentML	/dialign
pknotsRG	RNAML	RNAML	/pknotsrg
RNAhybrid(uc)	RNAML	RNAML	/rnahybrid
RNashapes(uc)	RNAML	RNAML	/rnashapes
RNAfold	RNAML	RNAML	/rnafold
ROSE	(PhyloML)	PhyloML	/rose
SplitsTree(uc)	PhyloML	PS	/splits
BioDOM	many	many	/biodomws

<sup>5</sup><http://bibiserv.techfak.uni-bielefeld.de/>



# WebService Choreography

## Webservice Pipelines in Bioinformatics

Jan Krüger and Henning Mersch

### Introduction

We present an automatism for easily connecting bioinformatical projects based on webservice technology, which are mainly offered by the Bielefeld University Bioinformatics Server (BiBiServ - <http://bibiserv.techfak.uni-bielefeld.de>) and other collaborators of the Helmholtz Open Bioinformatics Technology project (HOBIT - <http://hobit.sourceforge.net>).

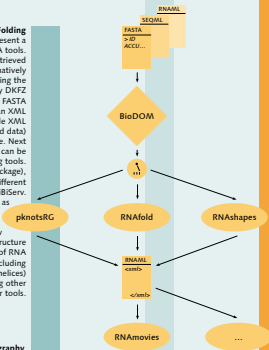
### RNA Folding

As first example, we present a pipeline of different RNA tools. Some sequences will be retrieved from local files or alternatively from the EMBL database using the SoapDB webservice offered by DKFZ Heidelberg. The sequences (in FASTA format) will be converted to an XML format (RNAAML is a suitable XML format for most RNA related data) using the BioDOM webservice. Next the converted sequences can be folded using different folding tools. RNAfold (Vienna RNA Package), pknotsRG or RNASHapes are different folding tools offered by BiBiServ.

The result - represented as RNAAML - can be visualized using RNAmovies. This webservice can be used to view a single RNA secondary structure or sequential animation of RNA secondary structures (including pseudoknots and entangled helices) or further processed using other webservices or tools.

### Webservice Choreography

Using standardized input and output formats makes it easy to combine



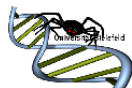
### BioDOM

Connecting webservices is mostly a problem of incompatible data formats which require more or less complex data conversion. To solve this problem using webservices in bioinformatics, we support XML formats beside proprietary formats for easy data exchange between common bioinformatics webservices. BiBiServ offers an API called „BioDOM“, which is also available as webservice. BioDOM provides a wide range of converters from and to more or less standardized grammars (XML Schemas), like SequenceML, AlignmentML, RNAML and EBIAApplication. BioDOM is an open project hosted at SourceForge (<http://biodom.sourceforge.net>).

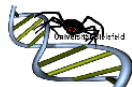
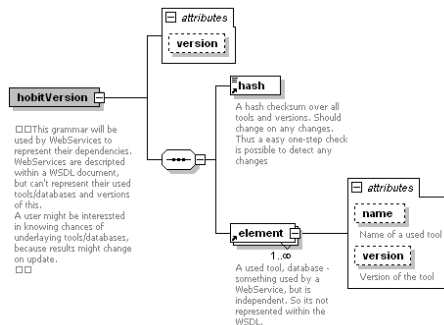


# getVersions()

- method `getVersions()` of BiBiServ-WebServices
- Information about used programs / databases / versions used by WS
- includes a (MD5) Hash for easy discovery of changes
- XSD driven
- namespace:  
`http://hobit.sourceforge.net/xsds/hobitVersion.xsd`
- new feature of BiBiWS-Library

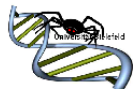


# getVersions()



# PropertiesParser

- new class `PropertiesParser` of BiBiWS Library
- validates parameters against a Java properties file
- supports all primitive data types
- supports min, max, default for numeric data types
- supports regexp for strings



## short remarks

- ReqResMap: Map request  $\leftrightarrow$  response method



## short remarks

- ReqResMap: Map request  $\leftrightarrow$  response method
- Promide (OligoSelection) NOT finished due to missing support by author  
→ Probably never will :-/



## short remarks

- ReqResMap: Map request  $\leftrightarrow$  response method
- Promide (OligoSelection) NOT finished due to missing support by author  
→ Probably never will :-/
- educational media at BiBiServ  
online course : SADR (Sequence Analysis of Distributed Resources)





# Time to say goodbye for me

- contact always welcome: `h.mersch@fz-juelich.de`



Forschungszentrum Jülich  
*in der Helmholtz-Gemeinschaft*

# Time to say goodbye for me

- contact always welcome: `h.mersch@fz-juelich.de`

- ⇒ successor: Sven Hartmeier



Forschungszentrum Jülich  
*in der Helmholtz-Gemeinschaft*