



**Overview of current resources and update on
DAS Meeting Cambridge 2010**

Jonathan Warren

Content

What is DAS?

What resources are available?

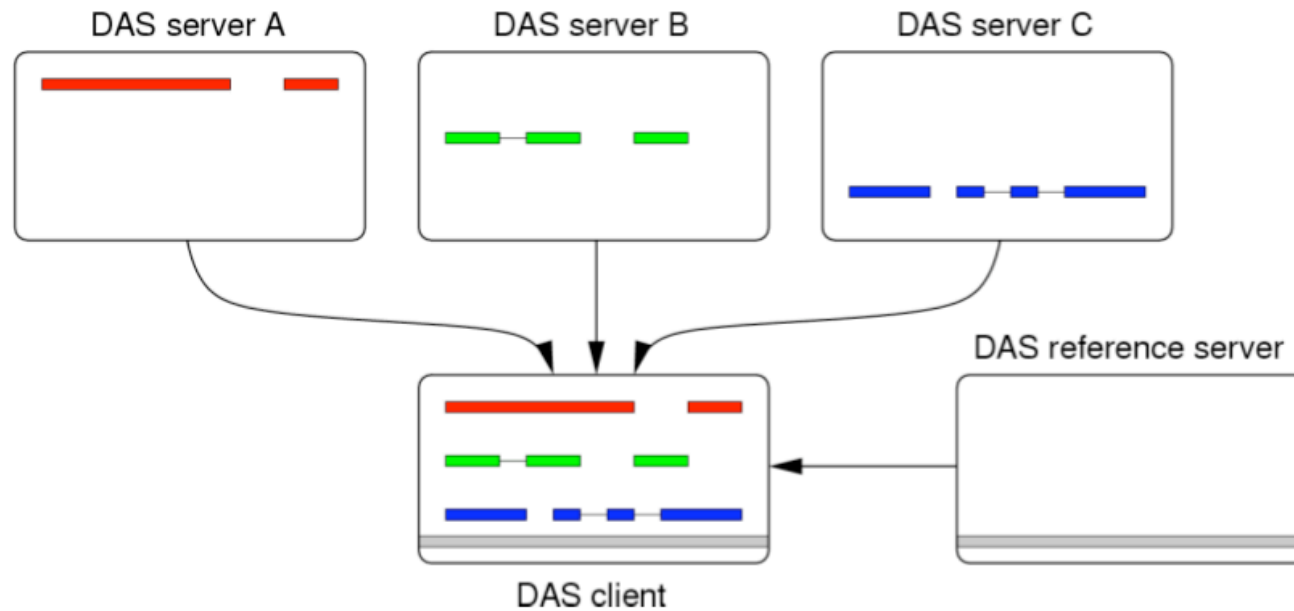
User scenario:

Distributed Annotation System Based on HTTP and XML

User perspective

- Client
- Choose coordinate system
- Connects to one registry for DAS server list
- Request a region of interest from the reference and many annotations from the DAS servers.

Vizualization of Distributed Annotation



Some DAS 1.5/1.6 Commands:

Sources
Features
Sequence
types
Stylesheet
Structure
Alignment
Interaction

Why use DAS 1.6 over 1.5?

http://www.ebi.ac.uk/~aj/1.6_draft7/documents/spec.html#coordinates

Clarification of the way DAS is being used - should promote interoperability

Represent features with more than two levels 1.6

- Represent Genes->Transcript->Exons

- GFF3 will be a supported format (Adapters for servers and databases).

- MyDAS server will support this without the need for a database

Reliably relate feature types to a more structured ontology

- cvld attributes in the xml for SO: or ECO ids - use of these may become mandatory in a future specification.

```

- <SEGMENT id="1" version="" start="113387" stop="6757161">
- <FEATURE id="ENSG00000112941">
  <TYPE id="Gene" cvId="SO:0000704">Gene</TYPE>
  <METHOD id="inferred from reviewed computational analysis" cvId="ECO:0000053">inferred from 1
  <START>140373</START>
  <END>190087</END>
  <SCORE>-</SCORE>
  <NOTE/>
  <LINK href=""/>
  <PART id="ENST00000398036"/>
  <PART id="ENST00000283426"/>
</FEATURE>
- <FEATURE id="ENSG00000112941">
  <TYPE id="Gene" cvId="SO:0000704">Gene</TYPE>
  <METHOD id="inferred from reviewed computational analysis" cvId="ECO:0000053">inferred from 1
  <START>6714718</START>
  <END>6757161</END>
  <SCORE>-</SCORE>
  <NOTE/>
  <LINK href=""/>
  <PART id="ENST00000230859"/>
</FEATURE>

```

Sources documents have coordinate systems which mean you are mapping annotations to the correct genomes/sequences

-<SOURCES>

-<SOURCE uri="test16genes" title="Test 1.6 sources" doc_href="http://www.myhomepage.org" description="Test for DAS 1.6 sources">

<MAINTAINER email="jw12@sanger.ac.uk"/>

-<VERSION uri="test16genes" created="2010-09-09T14:40:07+0000">

 <COORDINATES uri="http://www.dasregistry.org/dasregistry/coordsys/CS_DS311" taxid="9606" source="Chromosome" authority="GRCh" test_range="1:113387,6757161" version="37">GRCh_

<CAPABILITY type="das1:features" query_uri="http://das.sanger.ac.uk/das/test16genes/features"/>

<CAPABILITY type="das1:types" query_uri="http://das.sanger.ac.uk/das/test16genes/types"/>

<CAPABILITY type="das1:stylesheet" query_uri="http://das.sanger.ac.uk/das/test16genes/stylesheet"/>

<PROP name="label" value="Test16E"/>

<PROP name="leaseTime" value="2010-09-09T14:40:07+0000"/>

<PROP name="spec" value="DAS/1.6E"/>

</VERSION>

<SOURCE>

-<SOURCE uri="DS_112" title="structure" description="A reference server for protein structure annotations. Serves 3D coordinates of protein structures.">

<MAINTAINER email="andreas.prlc@gmail.com"/>

-<VERSION uri="DS_112" created="2010-09-09T14:40:07+0000">

<COORDINATES uri="http://www.dasregistry.org/dasregistry/coordsys/CS_DS7" source="Protein Structure" authority="PDBresnum" test_range="5pti">PDBresnum,Protein Structure</COORDIN

<CAPABILITY type="das1:structure" query_uri="http://das.sanger.ac.uk/das/structure/structure"/>

<PROP name="label" value="eFamily"/>

<PROP name="leaseTime" value="2010-09-09T14:40:07+0000"/>

<PROP name="spec" value="DAS/1.6E"/>

</VERSION>

<SOURCE>

Sources Doc Advantages:

Sources documents mean smoother running of ensembl and other DAS clients.

You can automatically load many DAS sources to the DAS registry using your sources document and the registry should keep in sync with new additions/deletions/alterations.

MyDAS and Proserver support the use of sources and all other 1.6 specification commands and responses.

New Registry has search interface

Coordinate Systems

Prev **1** 2 3 4 5 6 7 8 9 10 Next 1-10 of 139

hide / show

RGSC_3.4,Chromosome,Rattus norvegicus	10116	Chromosome	RGSC
AGP_1,Chromosome,Zea mays	4577	Chromosome	AGP
VEGA_38,Chromosome,Homo sapiens	9606	Chromosome	VEGA
ZFISH_7,Chromosome,Danio rerio	7955	Chromosome	ZFISH
BROADD_2,Chromosome,Canis familiaris	9615	Chromosome	BROADD
BDGP_5.4,Chromosome,Drosophila melanogaster	7227	Chromosome	BDGP
Sbi_1,Chromosome,Sorghum bicolor	4558	Chromosome	Sbi
VEGA_35,Chromosome,Gorilla gorilla	9593	Chromosome	VEGA
NASC29,Chromosome,Arabidopsis thaliana	3702	Chromosome	NASC29
IGGP_1,Chromosome,Vitis vinifera	29760	Chromosome	IGGP










Sources

Prev **1** 2 3 4 5 6 7 8 9 10 Next 1-10 of 536

hide / show

posid	nickname	status	description	Capabilities	types	Specification	Project	coordinateSystem
1	DS_182ZFMODELS-Zv7-microarrays		Show/Hide ZF-MODELS microarray oligos mapped to Zv7			DAS/1.53E		ZFISH_7,Chromosome,Danio rerio
2	DS_154 nscnp deleterious		Show/Hide Machine learning approach used SWISSPROT variants			DAS/1.53E		NCBI_35,Chromosome,Homo sapiens

Sources with Types matching keywords

posid	nickname	statusdescription	Capabilities	types								
2 DS_191   	hsa35pep	Show/Hide A reference and annotation server for Ensembl peptides. Provides exon structure and		Show/Hide Types <table border="1"> <thead> <tr> <th>id</th> <th>category</th> <th>cvld</th> <th>number</th> </tr> </thead> <tbody> <tr> <td>exon</td> <td></td> <td></td> <td>0</td> </tr> </tbody> </table>	id	category	cvld	number	exon			0
id	category	cvld	number									
exon			0									
3 DS_175   	Homo_sapiens.VEGA-clone.translation	Show/Hide Annotation source (returns clones) for Homo_sapiens translation		Show/Hide Types <table border="1"> <thead> <tr> <th>id</th> <th>category</th> <th>cvld</th> <th>number</th> </tr> </thead> <tbody> <tr> <td>exon</td> <td></td> <td></td> <td>0</td> </tr> </tbody> </table>	id	category	cvld	number	exon			0
id	category	cvld	number									
exon			0									
4 DS_176   	Mus_musculus.VEGA-clone.transcript	Show/Hide Annotation source (returns clones) for Mus_musculus transcript		Show/Hide Types <table border="1"> <thead> <tr> <th>id</th> <th>category</th> <th>cvld</th> <th>number</th> </tr> </thead> <tbody> <tr> <td>exon</td> <td></td> <td></td> <td>0</td> </tr> </tbody> </table>	id	category	cvld	number	exon			0
id	category	cvld	number									
exon			0									
5 DS_177   	Mus_musculus.VEGA-clone.translation	Show/Hide Annotation source (returns clones) for Mus_musculus translation		Show/Hide Types <table border="1"> <thead> <tr> <th>id</th> <th>category</th> <th>cvld</th> <th>number</th> </tr> </thead> <tbody> <tr> <td>exon</td> <td></td> <td></td> <td>0</td> </tr> </tbody> </table>	id	category	cvld	number	exon			0
id	category	cvld	number									
exon			0									
6 DS_178   	Homo_sapiens.VEGA-clone.transcript	Show/Hide Annotation source (returns clones) for Homo_sapiens transcript		Show/Hide Types <table border="1"> <thead> <tr> <th>id</th> <th>category</th> <th>cvld</th> <th>number</th> </tr> </thead> <tbody> <tr> <td>exon</td> <td></td> <td></td> <td>0</td> </tr> </tbody> </table>	id	category	cvld	number	exon			0
id	category	cvld	number									
exon			0									

keywords parameter to sources cmds e.g.

<http://www.dasregistry.org/das/sources?keywords=zebrafish>

* keywords parameter to coordinatesystem command

<http://www.dasregistry.org/das/coordinatesystem?keywords=gorilla>

* added total, start, end attributes to

coordinatesystem request response if rows specified

<http://www.dasregistry.org/das/coordinatesystem?keywords=gene&rows=1-10>

Easy DAS

Upload files of various formats

Hosted at the EBI

**No need for servers or
databases.**

1.6E proposals

<http://www.biodas.org/wiki/DAS1.6E>

.WibbleDibble files!??

- ../das/DSN/format

```
<DASFORMAT>
<COMMAND name="das1:features">
  <FORMAT name="das-JSON">
    ..if no types specified here then all types for this source have this format for this command
    <TYPE id="gene"/>
    <TYPE id="exon"/>
  </FORMAT>
  <FORMAT name="das-GoogleProtocolBuffers">
    ..if no types specified here then all types for this source have this format for this command
    <TYPE id="gene"/>
    <TYPE id="exon"/>
  </FORMAT>
</COMMAND>
<COMMAND name="das1:entry_points">
  <FORMAT name="das-JSON">
  </FORMAT>
</COMMAND>
</DASFORMAT>
```

Other Extensions:

- **Das Writeback (implemented)**
CRUD(Create, Read, Update and Delete)
- **Longer genomic alignments/compara**
Addition to the standard alignment
specification

New viewers supporting DAS

Dalliance Thomas Down

<http://www.biodalliance.org/>

<http://www.biodalliance.org/human/ncbi36/>

IGV Broad Institute

<http://www.broadinstitute.org/igv/>

Karyodas (Decipher, mykaryoview)

<http://code.google.com/p/karyodas/>

Apollo - new DAS DataAdapter release soon.

Jbrowse- Grant proposal

Other DAS clients:

- Ensembl uses DAS to pull in genomic, gene and protein annotations. It also provides data via DAS.
- Gbrowse is a generic genome browser, and is both a consumer and provider of DAS.
- IGB is a desktop application for viewing genomic data.
- SPICE is an application for projecting protein annotations onto 3D structures.
- Dasty2 is a web-based viewer for protein annotations
- Jalview is a multiple alignment editor.
- PeppeR is a graphical viewer for 3D electron microscopy data.
- DASMI is an integration portal for protein interaction data.
- DASher is a Java-based viewer for protein annotations.
- EpiC presents structure-function summaries for antibody design.
- STRAP is a STRucture-based sequence Alignment Program.

Acknowledgments:

Andy Jenkinson (EBI)
Rafael Jiminez
Bernat Gel
Jose Villaveces
Gustavo Salazar
James Smith
Andreas Prlic
Gregg Helt
Lincoln Stein
Tim Hubbard
The DAS community