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</METHOD>
    <START>6714718</START>
    <END>6757161</END>
    <SCORE></SCORE>
    <NOTE/>
    <LINK href=""/>
    <PART id="ENST00000230859"/>
  </FEATURE>
</SEGMENT>
Sources documents have coordinate systems which mean you are mapping annotations to the correct genomes/sequences.
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 Pro server support the use of sources and all other 1.6 specification commands and responses.
New Registry has search interface

Coordinate Systems

<table>
<thead>
<tr>
<th>Prefix</th>
<th>chromosome</th>
<th>species</th>
<th>accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGSC</td>
<td>3.4</td>
<td>Rattus norvegicus</td>
<td>10116</td>
</tr>
<tr>
<td>AGP</td>
<td>1</td>
<td>Zea mays</td>
<td>4577</td>
</tr>
<tr>
<td>VEGA</td>
<td>38</td>
<td>Homo sapiens</td>
<td>8606</td>
</tr>
<tr>
<td>ZFISH</td>
<td>7</td>
<td>Danio rerio</td>
<td>7955</td>
</tr>
<tr>
<td>BROADD</td>
<td>2</td>
<td>Canis familiaris</td>
<td>8615</td>
</tr>
<tr>
<td>BDGP</td>
<td>3.4</td>
<td>Drosophila melanogaster</td>
<td>7227</td>
</tr>
<tr>
<td>Stb</td>
<td>1</td>
<td>Serenica bicolor</td>
<td>4558</td>
</tr>
<tr>
<td>VEGA</td>
<td>35</td>
<td>Gorilla gorilla</td>
<td>9599</td>
</tr>
<tr>
<td>NASC20</td>
<td></td>
<td>Arabidopsis thaliana</td>
<td>3702</td>
</tr>
<tr>
<td>IGGP</td>
<td>1</td>
<td>Vitis vinifera</td>
<td>29760</td>
</tr>
</tbody>
</table>

Sources

<table>
<thead>
<tr>
<th>posid</th>
<th>nickname</th>
<th>status description</th>
<th>Capabilities</th>
<th>types</th>
<th>Specification</th>
<th>Project</th>
<th>coordinateSystem</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ZF-022FMOSL-1</td>
<td>Show/Hide</td>
<td>ZF-MODELS microarray oligos mapped to Zv7</td>
<td>DAS/1.53E</td>
<td>ZFISH_7.Chromosome.Danio rerio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>ZF-024nm0221</td>
<td>Show/Hide</td>
<td>Machine learning approach used with ZF-MODELS variants</td>
<td>DAS/1.53E</td>
<td>NCBI_35.Chromosome.Homo sapiens</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pos</td>
<td>nickname</td>
<td>status</td>
<td>description</td>
<td>Capabilities</td>
<td>types</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>----------</td>
<td>--------</td>
<td>-------------</td>
<td>--------------</td>
<td>-------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>hsa35pep</td>
<td>Show/Hide</td>
<td>A reference and annotation server for Ensembl peptides. Provides exon structure and...</td>
<td>☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐</td>
<td>id category cvld number exon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Homo_sapiens.VEGA-clone.translation</td>
<td>Show/Hide</td>
<td>Annotation source (returns clones) for Homo_sapiens translation</td>
<td>☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐</td>
<td>id category cvld number exon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Mus_musculus.VEGA-clone.transcript</td>
<td>Show/Hide</td>
<td>Annotation source (returns clones) for Mus_musculus transcript</td>
<td>☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐</td>
<td>id category cvld number exon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Mus_musculus.VEGA-clone.translation</td>
<td>Show/Hide</td>
<td>Annotation source (returns clones) for Mus_musculus translation</td>
<td>☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐</td>
<td>id category cvld number exon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Homo_sapiens.VEGA-clone.transcript</td>
<td>Show/Hide</td>
<td>Annotation source (returns clones) for Homo_sapiens transcript</td>
<td>☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐</td>
<td>id category cvld number exon</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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