Web Services at EMBL-EBI
(with examples)

Rodrigo Lopez Serrano
Head of Web Production

www.ebi.ac.uk
Overview

• Background to EMBL-EBI Services
• What are Web Services at EMBL-EBI?
• Why?
• Examples of EMBL-EBI Web Services
Background to Services…

- Access to data and applications using printed media (punch cards), tape, CD and DVD

Data Retrieval 1989-1996

• The EMBL Network File Server
  • This is the same year the T.B. Lee wrote the WWW proposal
  • “The EMBL Network File Server is a facility available on the computing system of the European Molecular Biology Laboratory to enable external users to access files via electronic mail (NETSERV@EMBL).”
  • Example commands:
    HELP
    GET NUC:X12399

• 1992: Data and Software distribution via ftp.
• 1995: Network file server moved to TCP and runs Fasta
1995-2005 - Public Services

- Gopher (1993) & World Wide Web
    - Gopher holes using SRS, Veronica and NCSA’s Mosaic
  - WWW/SRS, dbfetch, XEMBL, First GRIDs, Cluster and Farms @EBI
- Public Services from EBI. Nucleic Acids Res [2004, 32(web server issue):w3-9]
  - Web Servers for SSS, MSA, PSP, etc.
  - SOAP-based services provided by the European Bioinformatics Institute. (PMID:15980463)
2005 - today

- Focus on programmatic access to tools and data:
  - SOAP (Simple Object Access Protocol)
  - REST (Representational State Transfer)
- “Web Services at the European Bioinformatics Institute”
What are Web Services?

- [https://www.ebi.ac.uk/Tools/dbfetch/](https://www.ebi.ac.uk/Tools/dbfetch/)

- `wget`, `cURL`, `lynx`, [https://www.ebi.ac.uk/Tools/dbfetch/dbfetch?db=ensemblgene&id=ENSG00000012048&format=embl&style=default&Retrieve=Retrieve](https://www.ebi.ac.uk/Tools/dbfetch/dbfetch?db=ensemblgene&id=ENSG00000012048&format=embl&style=default&Retrieve=Retrieve)
What are Web Services?

• Network/web based services that provide access to data and/or methods:
  • Designed for use by programs
  • Web browser not required
  • Greater granularity and additional features

• Web Services models...
  • REST: resource based (nouns)
    • http://www.ebi.ac.uk/Tools/dbfetch/dbfetch?db=embl&id=L12345
    • http://www.ebi.ac.uk/Tools/dbfetch/dbfetch/embl/L12345
  • SOAP: operation/method/procedure based (verbs)
    • EntryData = fetchBatch(“embl”, “L12345”, “default”, “default”)
## Commonly used in Bioinformatics

<table>
<thead>
<tr>
<th>Service Type</th>
<th>Request</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>REST</td>
<td>Simple parametrised URL</td>
<td>Plain text (flat-file, fasta format, etc.) or XML</td>
</tr>
<tr>
<td>Distributed Annotation System (DAS)</td>
<td>Simple parametrised URL</td>
<td><strong>XML</strong></td>
</tr>
<tr>
<td>BioMart Web Service</td>
<td><strong>XML</strong> (GET or POST)</td>
<td>Table (HTML, TSV, CSV, XLS, etc.)</td>
</tr>
<tr>
<td>SOAP</td>
<td><strong>XML</strong> ‘soap:Envelope’</td>
<td><strong>XML</strong> ‘soap:Envelope’</td>
</tr>
</tbody>
</table>
Why use Web Services

- Local services require resource commitments:
  - Installation: hardware and software investment
  - Maintenance: patches & upgrades
  - Execution: CPU, memory & storage
  - In house requirements for specialist resources

- Move costs to service provider(s):
  - Generic services with large user base
  - Usage of service(s) reinforces commitment
Why Web Services technologies?

• Web Services work with the Internet
  • Firewall support (security)
  • Caches and proxies (performance, redundancy)
• Integration into existing pipelines and web applications
  • Web app's often implemented as SOAP/REST services
    • e.g. NCBI Entrez, EBI Search, ENSEMBL, etc.
  • Programming language neutral
  • Operating system independent
What can be done with Web Services

- In a single Web Services workflow:
  - Given an ENSEMBL coding gene
    - Obtain the CDS from ENSEMBL WS or ENA WS
    - Translate the CDS or get it from UniProt WS using Xrefs
    - Search for orthologs in UniProtKB using BLAST WS
    - Align top 500 hits using Clustal Omega WS
    - Build a phylogeny using WS for WebPrank
    - Run InterProScan WS to determine protein families and functional domains
  - Collect literature references from EuropePMC WS
EMBL-EBI Web Services (End-Point)

- ArrayExpress (REST):
  - http://www.ebi.ac.uk/arrayexpress/xml/v2/experiments?keywords=prostate
- ChEBI
  - http://www.ebi.ac.uk/chebi/webservices.do
- ChEMBL:
  - https://www.ebi.ac.uk/chembldb/index.php/ws
- EBI Search:
  - http://www.ebi.ac.uk/Tools/webservices/services/eb-eye
- ENA:
  - http://www.ebi.ac.uk/ena/about/browser
Web Services (End Point Examples)

• DBFetch:
  • http://www.ebi.ac.uk/Tools/webservices/services/dbfetch_rest

• ENSEMBL:
  • http://beta.rest.ensembl.org/

• UniProt (JAPI)
  • http://www.ebi.ac.uk/uniprot/remotingAPI/

• EuropePMC:
  • http://europepmc.org/RestfulWebService

• QuickGO:
  • http://www.ebi.ac.uk/QuickGO/WebServices.html
Web Services for common bioinformatics tools

- [http://www.ebi.ac.uk/Tools/webservices/](http://www.ebi.ac.uk/Tools/webservices/)
  - Sequence Similarity Searching (Blast+, Fasta, Exonerate), Multiple Sequence Alignment (Clustal Omega), Protein Function Analysis (InterProScan, Xfm), etc.