

NCIBI 2010 Annual Research Meeting

NCIBI Web Services Documentation and Demonstration

University of Michigan
20 April, 2010

Natural Language Processing (NLP) Web Service

Available Data

The NLP web service provides access to sentences, gene/protein tags, parse data, and metadata from biomedical literature in PubMed and/or PubMed Central Open Access.

Programmatic Interface

Base URL

<http://nlp.ncibi.org/fetch?>

URL Parameters

pmid

PMID is the PubMed ID for a document.

pmcid

PMCID is the PubMed Central ID for a document.

tagger

Tagger is the named entity tagger. Currently supports "NameTagger" tags.

parser

Parser is the grammatical parser. Currently supports the "Stanford" parser.

type

Type describes the output generated from tagging or parsing. For tagging, the currently supported type is "Gene" and for parsing is "Phrase."

id

Id is the canonical identifier for the named entity tag. "NameTagger" identifiers for the "Gene" type are Entrez gene Ids.

limit

Optional. Limit sets the maximum number of returned results. It defaults to 1000.

tool

Optional. Tool is a string with no internal spaces that identifies the resource using the service.

email

Optional. Email associates an email address with the request. This parameter allows us to contact users if there are problems or if the software interface changes.

metadata

Optional. When set to "all," outputs author, journal, title, date of publication, and MeSH terms.

Valid Parameter Combinations

pmid, *pmcid*, *pmid* & *tagger* & *type*, *pmid* & *parser* & *type*, *tagger* & *type* & *id*

Examples

<http://nlp.ncibi.org/fetch?pmid=17523140>

See <http://nlp.ncibi.org/about.html> for more examples.

Gene2MeSH Web Service

Available Data

The Gene2MeSH web service provides programmatic access to ranked lists of relevant gene/MeSH term pairs with associated statistical data, MeSH Qualifiers, and PubMed references. Ranking is based on p-values from Fisher Exact tests.

Programmatic Interface

Base URL

<http://gene2mesh.ncibi.org/fetch?>

URL Parameters

geneid

Geneid is an Entrez gene Id.

genesymbol

Genesymbol is an Entrez gene symbol.

taxid

Taxid is a taxonomy Id from the NCBI taxonomy database.

mesh

MeSH is a Medical Subject Headings (MeSH) descriptor.

limit

Optional. Limit sets the maximum number of returned results. It defaults to 1000.

tool

Optional. Tool is a string with no internal spaces that identifies the resource using the service. This parameter allows us to track usage of the service.

email

Optional. Email associates an email address with the request. This parameter allows us to contact users if there are problems or if the software interface changes.

Valid Parameter Combinations

geneid, *genesymbol*, *genesymbol* & *taxid*, *mesh*, *mesh* & *taxid*

Examples

<http://gene2mesh.ncibi.org/fetch?geneid=1436>

<http://gene2mesh.ncibi.org/fetch?genesymbol=csf1r&taxid=9606>

<http://gene2mesh.ncibi.org/fetch?mesh=diabetes>

See <http://gene2mesh.ncibi.org/about.html#programmatic> for more examples.

MiMI/Metabolomics Web Service

Available Data

The NCIBI MiMI/Metabolomics web service provides programmatic access to gene, interaction, and metabolomics (pathways, compounds, reactions, and enzymes) data.

Programmatic Interface

Base URL

<http://mimi.ncibi.org/MimiWeb/fetch.jsp?>

URL Parameters

search

Search term to submit a general MiMI free-text search.

geneid

Geneid is a Entrez gene Id for a gene.

cid

Cid is a KEGG compound Id.

rid

Rid is a KEGG reaction Id.

type

Type is the type of data to return. Currently supported types are “reactions,” “interactions,” and “compounds.” The default is “interactions.”

The types return the following information:

- interactions – A list of interactions and associated data for the given geneid
- reactions – A list of reactions and associated data that the given geneid participates in as an enzyme
- compounds – A list of compounds and associated data with which a given geneid is associated
- nlp – A list of literature derived interactions and relevant sentences from natural language processing

Valid Parameter Combinations

search, geneid & type, cid, rid

Examples

<http://mimi.ncibi.org/MimiWeb/fetch.jsp?search=pwp1>

<http://mimi.ncibi.org/MimiWeb/fetch.jsp?geneid=1436>

<http://mimi.ncibi.org/MimiWeb/fetch.jsp?geneid=1436&type=interactions>

<http://mimi.ncibi.org/MimiWeb/fetch.jsp?cid=C00061>

See http://mimi.ncibi.org/MimiWeb/ws/about_mimi.html for more examples.

ConceptGen Web Service

Available Data

The ConceptGen web service provides programmatic access to a ranked list of related concepts from public data sources such as Gene Ontology, KEGG, and Gene2MeSH. Ranking is based on enrichment results for each concept.

Programmatic Interface

Base URL

<http://conceptgen.ncibi.org/ConceptWeb/webservice?>

URL Parameters

search

Keyword search term for a concept (e.g. "cell").

conceptType

The type of concept (e.g. "KEGG").

ext_id

The original Id for the concept type (e.g for KEGG, "path:hsa03020").

networkType

Either "direct" or "complete." A "direct" network includes a center node and its connections. A "complete" network includes connections between all nodes.

pvalue

Optional. The p-value cutoff. The default is 0.05.

qvalue

Optional. The q-value cutoff. The default is 0.05.

output

Optional. The output type as either "text" or "xml." The default is text.

limit

Optional. Limits the maximum number of results returned. The default is 1000.

count

Optional "true" or "false" value. Sets the query to return a count of the resulting concepts. The default is false.

Valid Parameter Combinations

search, *ext_id* & *conceptType* & *networkType*

Examples

<http://conceptgen.ncibi.org/ConceptWeb/webservice?search=cancer&output=xml>

http://conceptgen.ncibi.org/ConceptWeb/webservice?ext_id=ath:hsa03020&conceptType=kegg%20pathway&networkType=direct&output=xml

Demonstration Perl Scripts

1. NLM eUtils web services query (eutilsExample.pl)

- a. Query for PMIDs of the top 20 abstracts, as ordered by PubMed, published by Matthias Kretzler in 2009.
- b. Output the PMIDs.

2. NCIBI NLP web services query (nlpExample.pl)

- a. Use a PMID from above as input to the NLP web service to retrieve all sentences from those abstracts that have been named entity tagged with at least one gene name.
- b. Output the sentences and genes.

3. NCIBI Gene2MeSH web services query (g2mExample.pl)

- a. Use a gene Id from above as input to the Gene2MeSH web service to retrieve the top 15 related MeSH terms for each gene Id.
- b. Output the MeSH terms.

4. NCIBI MiMI/Metabolomics web services queries (mimiExample.pl, mimiNlpExample.pl)

- a. Use a gene Id from above as input to the MiMI/Metabolomics web service to retrieve interaction genes and their gene Ids.
- b. Output the interaction gene's symbol and gene Id.
- c. Use a gene Id from above as input to the MiMI/Metabolomics web service to retrieve NLP derived interactions.
- d. Output the interaction gene's symbol and gene Id.

5. NCIBI ConceptGen web services query (cgExample.pl)

- a. Use a MeSH term from above as input to the ConceptGen web service to retrieve related concepts
- b. Output the related concepts

6. Query many NCIBI web services in one script (apiExample.pl) using the filtered output of one query as the input of another