



E-utilities Quick Start

Eric Sayers, PhD^{✉1}

Created: December 12, 2008; Updated: October 24, 2018.

Release Notes

Please see our [Release Notes](#) for details on recent changes and updates.

Announcement

On December 1, 2018, NCBI will begin enforcing the use of new API keys for E-utility calls. Please see [Chapter 2](#) for more details about this important change.

Introduction

This chapter provides a brief overview of basic E-utility functions along with examples of URL calls. Please see [Chapter 2](#) for a general introduction to these utilities and [Chapter 4](#) for a detailed discussion of syntax and parameters.

Examples include live URLs that provide sample outputs.

All E-utility calls share the same base URL:

`https://eutils.ncbi.nlm.nih.gov/entrez/eutils/`

Searching a Database

Basic Searching

`esearch.fcgi?db=<database>&term=<query>`

Input: Entrez database (&db); Any Entrez text query (&term)

Output: List of UIDs matching the Entrez query

Example: Get the PubMed IDs (PMIDs) for articles about breast cancer published in Science in 2008

[https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?db=pubmed&term=science\[journal\]+AND+breast+cancer+AND+2008\[pdat\]](https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?db=pubmed&term=science[journal]+AND+breast+cancer+AND+2008[pdat])

Storing Search Results

```
eSearch.fcgi?db=<database>&term=<query>&usehistory=y
```

Input: Any Entrez text query (&term); Entrez database (&db); &usehistory=y

Output: Web environment (&WebEnv) and query key (&query_key) parameters specifying the location on the Entrez history server of the list of UIDs matching the Entrez query

Example: Get the PubMed IDs (PMIDs) for articles about breast cancer published in Science in 2008, and store them on the Entrez history server for later use

```
https://eutils.ncbi.nlm.nih.gov/entrez/eutils/eSearch.fcgi?db=pubmed&term=science\[journal\]+AND+breast+cancer+AND+2008\[pdat\]&usehistory=y
```

Associating Search Results with Existing Search Results

```
eSearch.fcgi?db=<database>&term=<query1>&usehistory=y
```

eSearch produces WebEnv value (\$web1) and QueryKey value (\$key1)

```
eSearch.fcgi?db=<database>&term=<query2>&usehistory=y&WebEnv=$web1
```

eSearch produces WebEnv value (\$web2) that contains the results of both searches (\$key1 and \$key2)

Input: Any Entrez text query (&term); Entrez database (&db); &usehistory=y; Existing web environment (&WebEnv) from a prior E-utility call

Output: Web environment (&WebEnv) and query key (&query_key) parameters specifying the location on the Entrez history server of the list of UIDs matching the Entrez query

For More Information

Please see [ESearch In-Depth](#) for a full description of ESearch.

Sample ESearch Output

```
<?xml version="1.0" ?>
<!DOCTYPE eSearchResult PUBLIC "-//NLM//DTD eSearchResult, 11 May 2002//EN"
  "https://www.ncbi.nlm.nih.gov/entrez/query/DTD/eSearch_020511.dtd">
<eSearchResult>
<Count>255147</Count> # total number of records matching query
<RetMax>20</RetMax># number of UIDs returned in this XML; default=20
<RetStart>0</RetStart># index of first record returned; default=0
<QueryKey>1</QueryKey># QueryKey, only present if &usehistory=y
<WebEnv>0193yIkBjmM60UBXuvBvPfbIq8-9nIslDXuMP0hhuMH-
8GjCz7F_Dz1XL6z@397033B29A81FB01_0038SID</WebEnv>
  # WebEnv; only present if &usehistory=y
  <IdList>
<Id>229486465</Id> # list of UIDs returned
<Id>229486321</Id>
<Id>229485738</Id>
<Id>229470359</Id>
<Id>229463047</Id>
<Id>229463037</Id>
<Id>229463022</Id>
<Id>229463019</Id>
```

```

<Id>229463007</Id>
<Id>229463002</Id>
<Id>229463000</Id>
<Id>229462974</Id>
<Id>229462961</Id>
<Id>229462956</Id>
<Id>229462921</Id>
<Id>229462905</Id>
<Id>229462899</Id>
<Id>229462873</Id>
<Id>229462863</Id>
<Id>229462862</Id>
</IdList>
<TranslationSet>          # details of how Entrez translated the query
  <Translation>
    <From>mouse[orgn]</From>
    <To>"Mus musculus"[Organism]</To>
  </Translation>
</TranslationSet>
<TranslationStack>
  <TermSet>
    <Term>"Mus musculus"[Organism]</Term>
    <Field>Organism</Field>
    <Count>255147</Count>
    <Explode>Y</Explode>
  </TermSet>
  <OP>GROUP</OP>
</TranslationStack>
<QueryTranslation>"Mus musculus"[Organism]</QueryTranslation>
</eSearchResult>

```

Searching PubMed with Citation Data

```
ecitmatch.cgi?db=pubmed&rettype=xml&bdata=<citations>
```

Input: List of citation strings separated by a carriage return (%0D), where each citation string has the following format:

```
journal_title|year|volume|first_page|author_name|your_key|
```

Output: A list of citation strings with the corresponding PubMed ID (PMID) appended.

Example: Search PubMed for the following citations:

Art1: Mann, BJ. (1991) *Proc. Natl. Acad. Sci. USA*. 88:3248

Art2: Palmenberg, AC. (1987) *Science* 235:182

```
https://eutils.ncbi.nlm.nih.gov/entrez/eutils/ecitmatch.cgi?db=pubmed&retmode=xml&bdata=proc+natl+acac+sci+u+s+a|1991|88|3248|mann+bj|Art1|%0Dscience|1987|235|182|palmenberg+ac|Art2|
```

Sample Output (the PMIDs appear in the rightmost field):

```
proc natl acad sci u s a|1991|88|3248|mann bj|Art1|2014248
science|1987|235|182|palmenberg ac|Art2|3026048
```

Please see [ECitMatch In-Depth](#) for a full description of ECitMatch.

Uploading UIDs to Entrez

Basic Uploading

```
epost.fcgi?db=<database>&id=<uid_list>
```

Input: List of UIDs (&id); Entrez database (&db)

Output: Web environment (&WebEnv) and query key (&query_key) parameters specifying the location on the Entrez history server of the list of uploaded UIDs

Example: Upload five Gene IDs (7173,22018,54314,403521,525013) for later processing.

<https://eutils.ncbi.nlm.nih.gov/entrez/eutils/epost.fcgi?db=gene&id=7173,22018,54314,403521,525013>

Associating a Set of UIDs with Previously Posted Sets

```
epost.fcgi?db=<database1>&id=<uid_list1>
```

epost produces WebEnv value (\$web1) and QueryKey value (\$key1)

```
epost.fcgi?db=<database2>&id=<uid_list2>&WebEnv=$web1
```

epost produces WebEnv value (\$web2) that contains the results of both posts (\$key1 and \$key2)

Input: List of UIDs (&id); Entrez database (&db); Existing web environment (&WebEnv)

Output: Web environment (&WebEnv) and query key (&query_key) parameters specifying the location on the Entrez history server of the list of uploaded UIDs

For More Information

Please see [EPost In-Depth](#) for a full description of EPost.

Sample EPost Output

```
<?xml version="1.0"?>
<!DOCTYPE ePostResult PUBLIC "-//NLM//DTD ePostResult, 11 May 2002//EN"
  "https://www.ncbi.nlm.nih.gov/entrez/query/DTD/ePost_020511.dtd">
<ePostResult>
<QueryKey>1</QueryKey>
<WebEnv>NCID_01_268116914_130.14.18.47_9001_1241798628</WebEnv>
</ePostResult>
```

Downloading Document Summaries

Basic Downloading

```
esummary.fcgi?db=<database>&id=<uid_list>
```

Input: List of UIDs (&id); Entrez database (&db)

Output: XML DocSums

Example: Download DocSums for these protein GIs: 6678417,9507199,28558982,28558984,28558988,28558990

<https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esummary.fcgi?db=protein&id=6678417,9507199,28558982,28558984,28558988,28558990>

Downloading Data From a Previous Search

```
esearch.fcgi?db=<database>&term=<query>&usehistory=y
```

```
# esearch produces WebEnv value ($web1) and QueryKey value ($key1)
```

```
esummary.fcgi?db=<database>&query_key=$key1&WebEnv=$web1
```

Input: Web environment (&WebEnv) and query key (&query_key) representing a set of Entrez UIDs on the Entrez history server

Output: XML DocSums

Sample ESummary Output

The output of ESummary is a series of XML “DocSums” (Document Summaries), the format of which depends on the database. Below is an example DocSum for Entrez Protein.

```
<?xml version="1.0"?>
<!DOCTYPE eSummaryResult PUBLIC "-//NLM//DTD eSummaryResult, 29 October
 2004//EN" "https://www.ncbi.nlm.nih.gov/entrez/query/DTD/eSummary_
041029.dtd">
<eSummaryResult>
<DocSum>
<Id>15718680</Id>
<Item Name="Caption" Type="String">NP_005537</Item>
<Item Name="Title" Type="String">IL2-inducible T-cell kinase [Homo
 sapiens]</Item>
<Item Name="Extra"
Type="String">gi|15718680|ref|NP_005537.3|[15718680]</Item>
<Item Name="Gi" Type="Integer">15718680</Item>
<Item Name="CreateDate" Type="String">1999/06/09</Item>
<Item Name="UpdateDate" Type="String">2009/04/05</Item>
<Item Name="Flags" Type="Integer">512</Item>
<Item Name="TaxId" Type="Integer">9606</Item>
<Item Name="Length" Type="Integer">620</Item>
<Item Name="Status" Type="String">live</Item>
<Item Name="ReplacedBy" Type="String"></Item>
<Item Name="Comment" Type="String"><![CDATA[  ]]></Item>
</DocSum>
</eSummaryResult>
```

Sample ESummary version 2.0 Output

Version 2.0 of ESummary is an alternate XML presentation of Entrez DocSums. To retrieve version 2.0 DocSums, the URL should contain the &version parameter with an assigned value of ‘2.0’. Each Entrez database provides its own unique DTD for version 2.0 DocSums, and a link to the relevant DTD is provided in the header of the version 2.0 XML.

```
esummary.fcgi?db=<database>&id=<uid_list>&version=2.0
```

Below is an example version 2.0 DocSum from Entrez Protein (the same record as shown above in the default DocSum XML).

```
<?xml version="1.0"?>
<!DOCTYPE eSummaryResult PUBLIC "-//NLM//DTD eSummaryResult//EN" "https://
www.ncbi.nlm.nih.gov/entrez/query/DTD/eSummaryDTD/eSummary_protein.dtd">
<eSummaryResult>
  <DocumentSummarySet status="OK">
    <DocumentSummary uid="15718680">
      <Caption>NP_005537</Caption>
      <Title>tyrosine-protein kinase ITK/TSK [Homo sapiens]</Title>
      <Extra>gi|15718680|ref|NP_005537.3|</Extra>
      <Gi>15718680</Gi>

      <CreateDate>1999/06/09</CreateDate>
      <UpdateDate>2011/10/09</UpdateDate>
      <Flags>512</Flags>
      <TaxId>9606</TaxId>
      <Slen>620</Slen>

      <Biomol/>

      <MolType>aa</MolType>
      <Topology>linear</Topology>
      <SourceDb>refseq</SourceDb>
      <SegSetSize>0</SegSetSize>
      <ProjectId>0</ProjectId>
      <Genome>genomic</Genome>

      <SubType>chromosome|map</SubType>
      <SubName>5|5q31-q32</SubName>
      <AssemblyGi>399658</AssemblyGi>
      <AssemblyAcc>D13720.1</AssemblyAcc>
      <Tech/>
      <Completeness/>
      <GeneticCode>1</GeneticCode>

      <Strand/>
      <Organism>Homo sapiens</Organism>
      <Statistics>
        <Stat type="all" count="8"/>
        <Stat type="blob_size" count="16154"/>
        <Stat type="cdregion" count="1"/>
        <Stat type="cdregion" subtype="CDS" count="1"/>
        <Stat type="gene" count="1"/>
        <Stat type="gene" subtype="Gene" count="1"/>
        <Stat type="org" count="1"/>
        <Stat type="prot" count="1"/>
        <Stat type="prot" subtype="Prot" count="1"/>
        <Stat type="pub" count="14"/>
        <Stat type="pub" subtype="PubMed" count="10"/>
        <Stat type="pub" subtype="PubMed/Gene-rif" count="4"/>
        <Stat type="site" count="4"/>
        <Stat type="site" subtype="Site" count="4"/>
        <Stat source="CDD" type="all" count="15"/>
        <Stat source="CDD" type="region" count="6"/>
        <Stat source="CDD" type="region" subtype="Region" count="6"/>
        <Stat source="CDD" type="site" count="9"/>
        <Stat source="CDD" type="site" subtype="Site" count="9"/>
        <Stat source="HPRD" type="all" count="3"/>
        <Stat source="HPRD" type="site" count="3"/>
        <Stat source="HPRD" type="site" subtype="Site" count="3"/>
      </Statistics>
    </DocumentSummary>
  </DocumentSummarySet>
</eSummaryResult>
```

```

    <Stat source="SNP" type="all" count="31"/>
    <Stat source="SNP" type="imp" count="31"/>
    <Stat source="SNP" type="imp" subtype="variation" count="31"/>
    <Stat source="all" type="all" count="57"/>
    <Stat source="all" type="blob_size" count="16154"/>
    <Stat source="all" type="cdregion" count="1"/>
    <Stat source="all" type="gene" count="1"/>
    <Stat source="all" type="imp" count="31"/>
    <Stat source="all" type="org" count="1"/>
    <Stat source="all" type="prot" count="1"/>
    <Stat source="all" type="pub" count="14"/>
    <Stat source="all" type="region" count="6"/>
    <Stat source="all" type="site" count="16"/>
  </Statistics>
  <AccessionVersion>NP_005537.3</AccessionVersion>
  <Properties aa="2">2</Properties>
  <Comment/>
  <OSLT indexed="yes">NP_005537.3</OSLT>
  <IdGiClass mol="3" repr="2" gi_state="10" sat="4" sat_key="58760802"
owner="20"
          sat_name="NCBI" owner_name="NCBI-Genomes" defdiv="GNM" length="620"
extfeatmask="41"
          />
</DocumentSummary>

</DocumentSummarySet>
</eSummaryResult>

```

Downloading Full Records

Basic Downloading

```
efetch.fcgi?db=<database>&id=<uid_list>&rettype=<retrieval_type>
&retmode=<retrieval_mode>
```

Input: List of UIDs (&id); Entrez database (&db); Retrieval type (&rettype); Retrieval mode (&retmode)

Output: Formatted data records as specified

Example: Download nuccore GIs 34577062 and 24475906 in FASTA format

```
https://eutils.ncbi.nlm.nih.gov/entrez/eutils/efetch.fcgi?db=nuccore&id=34577062,24475906&rettype=fasta&retmode=text
```

Downloading Data From a Previous Search

```
esearch.fcgi?db=<database>&term=<query>&usehistory=y
```

esearch produces WebEnv value (\$web1) and QueryKey value (\$key1)

```
efetch.fcgi?db=<database>&query_key=$key1&WebEnv=$web1&rettype=
<retrieval_type>&retmode=<retrieval_mode>
```

Input: Entrez database (&db); Web environment (&WebEnv) and query key (&query_key) representing a set of Entrez UIDs on the Entrez history server; Retrieval type (&rettype); Retrieval mode (&retmode)

Output: Formatted data records as specified

Downloading a Large Set of Records

Please see [Application 3 in Chapter 3](#)

Input: Entrez database (&db); Web environment (&WebEnv) and query key (&query_key) representing a set of Entrez UIDs on the Entrez history server; Retrieval start (&retstart), the first record of the set to retrieve; Retrieval maximum (&retmax), maximum number of records to retrieve

Output: Formatted data records as specified

For More Information

Please see [EFetch In-Depth](#) for a full description of EFetch.

Finding Related Data Through Entrez Links

Basic Linking

Batch mode – finds only one set of linked UIDs

```
elink.fcgi?dbfrom=<source_db>&db=<destination_db>&id=<uid_list>
```

Input: List of UIDs (&id); Source Entrez database (&dbfrom); Destination Entrez database (&db)

Output: XML containing linked UIDs from source and destination databases

Example: Find one set of Gene IDs linked to nuccore GIs 34577062 and 24475906

<https://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=nuccore&db=gene&id=34577062,24475906>

‘By Id’ mode – finds one set of linked UIDs for each input UID

```
elink.fcgi?dbfrom=<source_db>&db=<destination_db>&id=<uid1>&id=<uid2>&id=<uid3>...
```

Example: Find separate sets of Gene IDs linked to nuccore GIs 34577062 and 24475906

<https://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=nuccore&db=gene&id=34577062&id=24475906>

Note: &db may be a comma-delimited list of databases, so that elink returns multiple sets of linked UIDs in a single call

Finding Links to Data from a Previous Search

```
esearch.fcgi?db=<source_db>&term=<query>&usehistory=y
```

esearch produces WebEnv value (\$web1) and QueryKey value (\$key1)

```
elink.fcgi?dbfrom=<source_db>&db=<destination_db>&query_key=$key1&WebEnv=$web1&cmd=neighbor_history
```

Input: Source Entrez database (&dbfrom); Destination Entrez database (&db); Web environment (&WebEnv) and query key (&query_key) representing the set of source UIDs on the Entrez history server; Command mode (&cmd)

Output: XML containing Web environments and query keys for each set of linked UIDs

Note: To achieve 'By Id' mode, one must send each input UID as a separate &id parameter in the URL. Sending a WebEnv/query_key set always produces Batch mode behavior (one set of linked UIDs).

Finding Computational Neighbors Limited by an Entrez Search

```
eink.fcgi?dbfrom=<source_db>&db=<source_db>&id=<uid_list>&term=
<query>&cmd=neighbor_history
```

Input: Source Entrez database (&dbfrom); Destination Entrez database (&db); List of UIDs (&id); Entrez query (&term); Command mode (&cmd)

Output: XML containing Web environments and query keys for each set of linked UIDs

Example: Find protein UIDs that are rat Reference Sequences and that are sequence similar to GI 15718680

```
https://eutils.ncbi.nlm.nih.gov/entrez/eutils/eink.fcgi?
dbfrom=protein&db=protein&id=15718680&term=rat[orgn]+AND+srcdb
+refseq[prop]&cmd=neighbor_history
```

For More Information

Please see [ELink In-Depth](#) for a full description of ELink.

Getting Database Statistics and Search Fields

```
einfo.fcgi?db=<database>
```

Input: Entrez database (&db)

Output: XML containing database statistics

Note: If no database parameter is supplied, einfo will return a list of all valid Entrez databases.

Example: Find database statistics for Entrez Protein.

```
https://eutils.ncbi.nlm.nih.gov/entrez/eutils/einfo.fcgi?db=protein
```

For More Information

Please see [EInfo In-Depth](#) for a full description of EInfo.

Sample EInfo Output

```
<?xml version="1.0"?>
<!DOCTYPE eInfoResult PUBLIC "-//NLM//DTD eInfoResult, 11 May 2002//EN"
"http://www.ncbi.nlm.nih.gov/entrez/query/DTD/eInfo_020511.dtd">
<eInfoResult>
<DbInfo>
<DbName>protein</DbName>
<MenuName>Protein</MenuName>
<Description>Protein sequence record</Description>
<Count>26715092</Count>
<LastUpdate>2009/05/12 04:39</LastUpdate>
<FieldList>
<Field>
<Name>ALL</Name>
<FullName>All Fields</FullName>
<Description>All terms from all searchable fields</Description>
```

```

<TermCount>133639432</TermCount>
<IsDate>N</IsDate>
<IsNumerical>N</IsNumerical>
<SingleToken>N</SingleToken>
<Hierarchy>N</Hierarchy>
<IsHidden>N</IsHidden>
</Field>
...
<Field>
<Name>PORG</Name>
<FullName>Primary Organism</FullName>
<Description>Scientific and common names
of primary organism, and all higher levels of taxonomy</Description>
<TermCount>673555</TermCount>
<IsDate>N</IsDate>
<IsNumerical>N</IsNumerical>
<SingleToken>Y</SingleToken>
<Hierarchy>Y</Hierarchy>
<IsHidden>N</IsHidden>
</Field>
</FieldList>
<LinkList>
<Link>
<Name>protein_biosystems</Name>
<Menu>BioSystem Links</Menu>
<Description>BioSystems</Description>
<DbTo>biosystems</DbTo>
</Link>
...
<Link>
<Name>protein_unigene</Name>
<Menu>UniGene Links</Menu>
<Description>Related UniGene records</Description>
<DbTo>unigene</DbTo>
</Link>
</LinkList>
</DbInfo>
</eInfoResult>

```

Performing a Global Entrez Search

egquery.fcgi?term=<query>

Input: Entrez text query (&term)

Output: XML containing the number of hits in each database.

Example: Determine the number of records for mouse in Entrez.

[https://eutils.ncbi.nlm.nih.gov/entrez/eutils/egquery.fcgi?term=mouse\[orgn\]](https://eutils.ncbi.nlm.nih.gov/entrez/eutils/egquery.fcgi?term=mouse[orgn])

For More Information

Please see [EGQuery In-Depth](#) for a full description of EGQuery.

Sample EGQuery Output

```

<?xml version="1.0"?>
<!DOCTYPE Result PUBLIC "-//NLM//DTD eSearchResult, January 2004//EN"

```

```

"https://www.ncbi.nlm.nih.gov/entrez/query/DTD/egquery.dtd">
<!--
    $Id: egquery_template.xml 106311 2007-06-26 14:46:31Z osipov $
-->
<!-- ===== -->
<Result>
  <Term>mouse[orgn]</Term>
  <eGQueryResult>
    <ResultItem>
      <DbName>pubmed</DbName>
      <MenuName>PubMed</MenuName>
      <Count>0</Count>
      <Status>Term or Database is not found</Status>
    </ResultItem>
    <ResultItem>
      <DbName>pmc</DbName>
      <MenuName>PMC</MenuName>
      <Count>3823</Count>
      <Status>Ok</Status>
    </ResultItem>
    ...
    <ResultItem>
      <DbName>nucore</DbName>
      <MenuName>Nucleotide</MenuName>
      <Count>1739903</Count>
      <Status>Ok</Status>
    </ResultItem>
    <ResultItem>
      <DbName>nucgss</DbName>
      <MenuName>GSS</MenuName>
      <Count>2264567</Count>
      <Status>Ok</Status>
    </ResultItem>
    <ResultItem>
      <DbName>nucgst</DbName>
      <MenuName>EST</MenuName>
      <Count>4852140</Count>
      <Status>Ok</Status>
    </ResultItem>
    <ResultItem>
      <DbName>protein</DbName>
      <MenuName>Protein</MenuName>
      <Count>255212</Count>
      <Status>Ok</Status>
    </ResultItem>
    ...
    <ResultItem>
      <DbName>proteinclusters</DbName>
      <MenuName>Protein Clusters</MenuName>
      <Count>13</Count>
      <Status>Ok</Status>
    </ResultItem>
  </eGQueryResult>
</Result>

```

Retrieving Spelling Suggestions

```

espell.fcgi?term=<query>&db=<database>

```

Input: Entrez text query (&term); Entrez database (&db)

Output: XML containing the original query and spelling suggestions.

Example: Find spelling suggestions for the PubMed Central query 'fiberblast cell grwth'.

<https://eutils.ncbi.nlm.nih.gov/entrez/eutils/espell.fcgi?term=fiberblast+cell+grwth&db=pmc>

For More Information

Please see [ESpell In-Depth](#) for a full description of EGQuery.

Sample ESpell Output

```
<?xml version="1.0"?>
<!DOCTYPE eSpellResult PUBLIC "-//NLM//DTD eSpellResult, 23 November
2004//EN" "https://www.ncbi.nlm.nih.gov/entrez/query/DTD/eSpell.dtd">
<eSpellResult>
<Database>pmc</Database>
<Query>fiberblast cell grwth</Query>
<CorrectedQuery>fibroblast cell growth</CorrectedQuery>
<SpelledQuery>
  <Replaced>fibroblast</Replaced>
  <Original> cell </Original>
  <Replaced>growth</Replaced>
</SpelledQuery>
<ERROR/>
</eSpellResult>
```

Demonstration Programs

EBot

EBot is an interactive web tool that first allows users to construct an arbitrary E-utility analysis pipeline and then generates a Perl script to execute the pipeline. The Perl script can be downloaded and executed on any computer with a Perl installation. For more details, see the EBot page linked above.

Sample Perl Scripts

The two sample Perl scripts below demonstrate basic E-utility functions. Both scripts should be copied and saved as plain text files and can be executed on any computer with a Perl installation.

ESearch-EFetch demonstrates basic search and retrieval functions.

```
#!/usr/local/bin/perl -w
# =====
#
#                               PUBLIC DOMAIN NOTICE
#                               National Center for Biotechnology Information
#
# This software/database is a "United States Government Work" under the
# terms of the United States Copyright Act. It was written as part of
# the author's official duties as a United States Government employee and
# thus cannot be copyrighted. This software/database is freely available
# to the public for use. The National Library of Medicine and the U.S.
# Government have not placed any restriction on its use or reproduction.
#
# Although all reasonable efforts have been taken to ensure the accuracy
```

```

# and reliability of the software and data, the NLM and the U.S.
# Government do not and cannot warrant the performance or results that
# may be obtained by using this software or data. The NLM and the U.S.
# Government disclaim all warranties, express or implied, including
# warranties of performance, merchantability or fitness for any particular
# purpose.
#
# Please cite the author in any work or product based on this material.
#
# =====
#
# Author: Oleg Khovayko
#
# File Description: eSearch/eFetch calling example
#
# -----
# Subroutine to prompt user for variables in the next section

sub ask_user {
    print "$_[0] [$_[1]]: ";
    my $rc = <>;
    chomp $rc;
    if($rc eq "") { $rc = $_[1]; }
    return $rc;
}

# -----
# Define library for the 'get' function used in the next section.
# $utils contains route for the utilities.
# $db, $query, and $report may be supplied by the user when prompted;
# if not answered, default values, will be assigned as shown below.

use LWP::Simple;

my $utils = "https://www.ncbi.nlm.nih.gov/entrez/eutils";

my $db      = ask_user("Database", "Pubmed");
my $query   = ask_user("Query",    "zanzibar");
my $report  = ask_user("Report",   "abstract");

# -----
# $esearch cont?ins the PATH & parameters for the ESearch call
# $esearch_result contains the result of the ESearch call
# the results are displayed and parsed into variables
# $Count, $QueryKey, and $WebEnv for later use and then displayed.

my $esearch = "$utils/esearch.fcgi?" .
              "db=$db&retmax=1&usehistory=y&term=";

my $esearch_result = get($esearch . $query);

print "\nESEARCH RESULT: $esearch_result\n";

$esearch_result =~
    m|<Count>(\d+)</Count>.*<QueryKey>(\d+)</QueryKey>.*<WebEnv>(\S+)</WebEnv>|s;

my $Count      = $1;
my $QueryKey   = $2;
my $WebEnv     = $3;

```

```

print "Count = $Count; QueryKey = $QueryKey; WebEnv = $WebEnv\n";

# -----
# this area defines a loop which will display $retmax citation results from
# Efetch each time the the Enter Key is pressed, after a prompt.

my $retstart;
my $retmax=3;

for($retstart = 0; $retstart < $Count; $retstart += $retmax) {
  my $efetch = "$utils/efetch.fcgi?" .
    "rettype=$report&retmode=text&retstart=$retstart&retmax=$retmax&" .
    "db=$db&query_key=$QueryKey&WebEnv=$WebEnv";

  print "\nEF_QUERY=$efetch\n";

  my $efetch_result = get($efetch);

  print "-----\nEFETCH RESULT(" .
    ($retstart + 1) . ".." . ($retstart + $retmax) . "): " .
    "[$efetch_result]\n-----PRESS ENTER!!!-----\n";
  <>;
}

```

EPost-ESummary demonstrates basic uploading and document summary retrieval.

```

#!/usr/local/bin/perl -w
# =====
#
#                PUBLIC DOMAIN NOTICE
#          National Center for Biotechnology Information
#
# This software/database is a "United States Government Work" under the
# terms of the United States Copyright Act.  It was written as part of
# the author's official duties as a United States Government employee and
# thus cannot be copyrighted.  This software/database is freely available
# to the public for use.  The National Library of Medicine and the U.S.
# Government have not placed any restriction on its use or reproduction.
#
# Although all reasonable efforts have been taken to ensure the accuracy
# and reliability of the software and data, the NLM and the U.S.
# Government do not and cannot warrant the performance or results that
# may be obtained by using this software or data.  The NLM and the U.S.
# Government disclaim all warranties, express or implied, including
# warranties of performance, merchantability or fitness for any particular
# purpose.
#
# Please cite the author in any work or product based on this material.
#
# =====
# Author:  Oleg Khovayko
#
# File Description:  ePost/eSummary calling example
#
# -----
my $utils_root = "https://www.ncbi.nlm.nih.gov/entrez/eutils";

```

```

my $ePost_url      = "$eutils_root/epost.fcgi";
my $eSummary_url  = "$eutils_root/esummary.fcgi";

my $db_name = "PubMed";

# -----
use strict;

use LWP::UserAgent;
use LWP::Simple;
use HTTP::Request;
use HTTP::Headers;
use CGI;

# -----
# Read input file into variable $file
# File name - first argument $ARGV[0]

undef $/; #for load whole file

open IF, $ARGV[0] || die "Can't open for read: $!\n";
my $file = <IF>;
close IF;
print "Loaded file: [$file]\n";

# Prepare file - substitute all separators to comma

$file =~ s/\s+/,/gs;
print "Prepared file: [$file]\n";

#Create CGI param line

my $form_data = "db=$db_name&id=$file";

# -----
# Create HTTP request

my $headers = new HTTP::Headers(
    Accept      => "text/html, text/plain",
    Content_Type => "application/x-www-form-urlencoded"
);

my $request = new HTTP::Request("POST", $ePost_url, $headers );

$request->content($form_data);

# Create the user agent object

my $ua = new LWP::UserAgent;
$ua->agent("ePost/example");

# -----
# send file to ePost by HTTP

my $response = $ua->request($request);

# -----

print "Response status message: [" . $response->message . "]\n";

```

```

print "Response content: [" .          $response->content . "]\n";

# -----
# Parse response->content and extract QueryKey & WebEnv
$response->content =~
  m|<QueryKey>(\d+)</QueryKey>.*<WebEnv>(\S+)</WebEnv>|s;

my $QueryKey = $1;
my $WebEnv   = $2;

print "\nEXTRACTED:\nQueryKey = $QueryKey;\nWebEnv = $WebEnv\n\n";

# -----
# Retrieve DocSum from eSummary by simple::get method and print it
#
print "eSummary result: [" .
  get("$eSummary_url?db=$db_name&query_key=$QueryKey&WebEnv=$WebEnv") .
  "]\n";

```

For More Information

Announcement Mailing List

NCBI posts general announcements regarding the E-utilities to the [utilities-announce announcement mailing list](#). This mailing list is an announcement list only; individual subscribers may **not** send mail to the list. Also, the list of subscribers is private and is not shared or used in any other way except for providing announcements to list members. The list receives about one posting per month. Please subscribe at the above link.

Getting Help

Please refer to the [PubMed](#) and [Entrez](#) help documents for more information about search queries, database indexing, field limitations and database content.

Suggestions, comments, and questions specifically relating to the EUtility programs may be sent to entilities@ncbi.nlm.nih.gov.