

Technical Information

Created: July 18, 2005
Updated: February 22, 2008

LinkOut is a service of Entrez that allows users to link directly from PubMed and other Entrez databases to a wide range of information and services extending beyond the Entrez system. LinkOut aims to facilitate access to relevant online resources to extend, clarify, and supplement information found in the Entrez databases. Participating LinkOut providers, including publishers, aggregators, libraries, and other web resources, can display links to their sites on records from the Entrez databases. A current list of LinkOut providers is available.

Resources Description

LinkOut requires two types of files, Identity and Resource files, to describe online resources. These files are specified in the LinkOut DTD, and are provided by a participating provider. These files include the necessary elements for the Entrez system to construct an appropriate URL to access specific resources.

Publishers and Other Full-Text Providers

Publishers and other full-text providers can use LinkOut to provide links from PubMed citations to their web-accessible full-text journals. For publishers who supply their data to PubMed electronically and providers of free full text, an icon link will appear in PubMed's AbstractPlus, Abstract, and Citation formats by default (more about default icons). Other providers should see Activating Icons for LinkOut Providers for information about promoting their icon to these display formats.

Additional information and file examples for publishers and other full-text providers are available in Information for Full-Text Providers.

Libraries

LinkOut provides libraries the means to link patrons from a PubMed citation directly to the full-text of an article from the library's full-text provider. To do this, libraries must submit holdings information to LinkOut. Libraries should use the LinkOut Files Submission Utility to generate their LinkOut files. The holdings data should include those online journals that the library has a subscription to under a prior agreement. When a library submits their holdings data to NCBI, the providers of the holdings (publishers or aggregators) must be LinkOut participants, i.e. the providers have already submitted linking information for their online journals to NCBI.

Additional information is available in Information for Libraries.

Other Resources

LinkOut can also be used to provide links from Entrez databases to non-bibliographic Web resources, such as factual databases (i.e., organism-specific, taxonomy, structure and other databases), catalogs of research materials (clones, strains, probes, etc.), funding opportunities, clinical resources, research groups, and others. Additional information and file examples are available in Information for Other Resource Providers.

LinkOut DTD

The LinkOut Document Type Definition (DTD) specifies the XML elements used in LinkOut files. The LinkOut DTD is available for download here: [LinkOut.DTD](#).

The URL for the DTD must be included in the prolog of a LinkOut file.

In the identity file, use:

```
<!DOCTYPE Provider "-//NLM/DTD LinkOut 1.0//EN" "http://www.ncbi.nlm.nih.gov/projects/linkout/doc/LinkOut.dtd">
```

In a resource file, use:

```
<!DOCTYPE LinkSet "-//NLM/DTD LinkOut 1.0//EN" "http://www.ncbi.nlm.nih.gov/projects/linkout/doc/LinkOut.dtd">
```

```
<!-- $Id: LinkOut.dtd, v 1.23 2007/09/20 21:21:56 Exp $ -->
<!-- LinkOut DTD version 1.1
```

This document is still under revision and may change.

Changes:

version 1.1:

- SubjectType and Attribute lists updated
- DTD is XML

version 1.6:

- <apad> tag added

version sequence refreshed to 1.1 on 2001/09/20

version 1.2:

- SubObjectSelector updated to

```
<!ELEMENT SubObjectSelector (Database, SubProvider)>
```

version 1.3:

- Entities added: lo.name, lo.genus, lo.species, lo.subsp

version 1.6:

- Entities added: lo.month, lo.authln

version 1.8:

- Entities added: lo.eyear, lo.eyr, lo.eyl, lo.emonth, lo.emon, lo.emo, lo.eday

version 1.11:

- added Attribute: preference

version 1.12:

- dropped lo.muid

version 1.16

- Entities added: lo.clusterid for UniGene

version 1.22

- SubjectType added: locus-specific

```
-->
```

```
<!-- Typical usage:
```

```
<!DOCTYPE LinkSet PUBLIC "-//NLM/DTD LinkOut//EN" "LinkOut.dtd">
```

```
<LinkSet>
```

```
...
```

```
</LinkSet>
```

```
or:
```

```
<!DOCTYPE Provider PUBLIC "-//NLM/DTD LinkOut//EN" "LinkOut.dtd">
```

```
<Provider>
```

```
...
```

```
</Provider>
```

```
-->
```

```
<!--
```

Rule based URL generation. In general, to build the URL both base and rule are required, where base is the HTTP base address, for example:

```
http://www.sciencemag.org/cgi/content/full/
```

and rule is in the following format:

```
&lo.vol;/&lo.iss;/&lo.page;
```

LinkOut will replace the keywords in rule with the actual value for a retrieved citation. Therefore, rule will be translated into: 281/5384/1863

The program will concatenate base with rule:

```
http://www.sciencemag.org/cgi/content/full/281/5384/1863
```

The following keywords are supported for any database:

lo.id - Unique identifier (PMID, GI, TaxID, etc.)

For PubMed only:

lo.pii - Publisher Item Identification. Must be submitted by the publisher.

For example, 6847, in the PubMed DTD this ID is an attribute of the ArticleId element.

lo.doi - Article DOI
 lo.issn - Journal ISSN code
 lo.essn - Journal Electronic ISSN code
 lo.issnl - Journal ISSN code without the dash
 lo.jtit - Journal title (MEDLINE abbreviation)
 lo.vol - Volume
 lo.iss - Issue
 lo.page - First page
 lo.year - Four digit year. For example, 1998
 lo.yr - Last two digit of year. For example, 98; 00
 lo.yl - Last digit of year. For example, for 1999 use 9; for 1990 use 0
 lo.eyear - Four digit year of electronic publication date. For example, 1998
 lo.eyr - Last two digits of year of electronic publication date. For example, 98; 00
 lo.eyl - Last digit of year of electronic publication date. For example, for 1999 use 9; for 1990 use 0
 lo.month - The month. For example, September
 lo.mon - The 3 letter month abbreviation. For example, Sep
 lo.mo - Two digit month. For example, 01; 12
 lo.emonth - The month of electronic publication date. For example, September
 lo.emon - The 3 letter month abbreviation of electronic publication date. For example, Sep
 lo.emo - Two digit month of electronic publication date. For example, 01;12
 lo.day - Two digit day. For example, 01; 31
 lo.eday - Two digit day of electronic publication date. For example, 01; 31
 lo.otit - Article title
 lo.auth - First author
 lo.authln - Last name of the first author
 For Sequence databases (Nucleotide, Protein, Structure, Genome):
 lo.pacc - Primary accession for sequences
 For Taxonomy only:
 lo.scientificname - Full scientific name. For example: "Homo sapiens neanderthalensis"
 lo.genus - Genus name. For example: "Homo"
 lo.species - Species epithet. For example: "sapiens"
 lo.subsp - Sub-species epithet. For example: "neanderthalensis"
 For UniGene only:
 lo.clusterid - Cluster Id
 -->
 <!ENTITY lo.id "lo.id">
 <!ENTITY lo.pii "lo.pii">
 <!ENTITY lo.doi "lo.doi">
 <!ENTITY lo.issn "lo.issn">
 <!ENTITY lo.essn "lo.essn">
 <!ENTITY lo.issnl "lo.issnl">
 <!ENTITY lo.jtit "lo.jtitle">
 <!ENTITY lo.vol "lo.vol">
 <!ENTITY lo.iss "lo.iss">
 <!ENTITY lo.page "lo.page">
 <!ENTITY lo.year "lo.year">
 <!ENTITY lo.yr "lo.yr">
 <!ENTITY lo.yl "lo.yl">
 <!ENTITY lo.eyear "lo.eyear">
 <!ENTITY lo.eyr "lo.eyr">
 <!ENTITY lo.eyl "lo.eyl">
 <!ENTITY lo.month "lo.month">
 <!ENTITY lo.mon "lo.mon">
 <!ENTITY lo.mo "lo.mo">
 <!ENTITY lo.emonth "lo.emonth">
 <!ENTITY lo.emon "lo.emon">
 <!ENTITY lo.emo "lo.emo">
 <!ENTITY lo.day "lo.day">
 <!ENTITY lo.eday "lo.eday">
 <!ENTITY lo.otit "lo.otit">
 <!ENTITY lo.auth "lo.auth">
 <!ENTITY lo.authln "lo.authln">
 <!ENTITY lo.pacc "lo.pacc">
 <!ENTITY lo.scientificname "lo.scientificname">
 <!ENTITY lo.genus "lo.genus">
 <!ENTITY lo.species "lo.species">
 <!ENTITY lo.subsp "lo.subsp">
 <!ENTITY lo.clusterid "lo.clusterid">

```

<!-- Entities for special characters -->
<!ENTITY reg      "®" >
<!ENTITY copy     "©">
<!ENTITY trade    "™">
<!-- internal DTD entities -->
<!ENTITY % url.data      "#PCDATA">
<!ENTITY % rule.data     "%url.data; | pad | apad | subs | toupper | tolower
                           | strip | normalize">
<!ENTITY % iso.lang.codes "(DA|DE|EN|EL|ES|FR|IT|IW|JA|NL|NO|RU|SV|ZH)">
<!ELEMENT SubjectType    (#PCDATA)>
<!-- can be one of the following; see Special Elements: SubjectType for a description of these
elements:
//CHEMICAL INFORMATION
    Biological Properties
    Chemical Libraries
    Imaging Agents
    Metabolism
    Molecular Interactions
    Physical Properties
    Reactions
    Theoretical Properties
    Toxicology
    Vendors
//EDUCATION
    conferences/meetings/workshops
    glossaries/dictionaries
    online tutorials/courses
//FUNDING SOURCES
    funding sources
//LITERATURE
    abstracts/indexes/summaries
    aggregators
    books
    individual online article
    images
    libraries
    patent databases
    publishers/providers
//MEDICAL
    clinical trials
    consumer health
    diagnostics
    disease organizations
    medical equipment and devices
    pharmacology
    treatment guidelines
//MOLECULAR BIOLOGY DATABASES
    DNA/protein sequence
    gene/protein/disease-specific
    gene expression
    locus-specific
    mapping
    organism-specific
    population/variation
    protein interactions/pathways
    structure
    taxonomy/phylogenetic
//RESEARCH MATERIALS
    clones/clone libraries
    culture/stock collections
    laboratory equipment
    oligonucleotides
    other reagents
//RESEARCHERS
    colleges/universities
    companies/research institutes
    directories
    individuals
    societies/associations
//TOOLS
    3D structure prediction/functional modeling

```

```

    primer design
    protein identification/characterization
    restriction mapping
    sequence screening/similarity/alignment
    sequence viewer
    translation
-->
<!ELEMENT Attribute    (#PCDATA)>
<!-- can be one of the following; see LinkOut SubjectTypes and Attributes at
http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=helpinkout.section.files.Special_Elements_Att for a
description of these elements:
//BARRIERS
    registration required
    subscription/membership/fee required
//OWNERSHIP
    author of URL
    publisher of information in URL
//Resource Form
    author manuscript
    electronic full-text
    full-text online
    full-text PDF
    full-text PostScript
    order form
    print collection
//Administrative
    library-local
    preference
-->
<!-- This is the top level element for Provider -->
<!ELEMENT Provider    (ProviderId, Name, NameAbbr, SubjectType*,
                        Attribute*, Url*, IconUrl*, Brief?)>
<!ELEMENT ProviderId    (#PCDATA)>
<!ELEMENT Name          (#PCDATA)>
<!ELEMENT NameAbbr     (#PCDATA)>
<!ELEMENT Url          (%url.data;)>
<!ATTLIST Url
    LNG      %iso.lang.codes;  "EN"  >
<!ELEMENT IconUrl     (%url.data;)>
<!ATTLIST IconUrl
    LNG      %iso.lang.codes;  "EN"  >
<!ELEMENT Brief       (#PCDATA)>
<!-- End of Provider group -->
<!ELEMENT LinkSet     (Link+)>
<!ELEMENT Link        (LinkId, ProviderId, IconUrl*,
                        ObjectSelector|SubObjectSelector), ObjectUrl+)>
<!ELEMENT LinkId      (#PCDATA)>
<!ELEMENT ObjectSelector (Database, ObjectList)>
<!ELEMENT Database    (#PCDATA)>
<!-- All Entrez databases may be configured to include LinkOut. Please write to
linkout@ncbi.nlm.nih.gov for the current list of databases available for LinkOut -->
<!ELEMENT ObjectList  (FileName|Query|ObjId)+>
<!ELEMENT FileName    (#PCDATA)>
<!ATTLIST FileName    fieldname  CDATA  #REQUIRED>
<!-- FileName is reserved for processing by the LinkOut team -->
<!ELEMENT Query       (#PCDATA)>
<!ELEMENT ObjId       (#PCDATA)>
<!-- End of ObjectList group -->
<!-- End of ObjectSelector group -->
<!-- Libraries must use SubObjectSelector to refer to the sub providers -->
<!ELEMENT SubObjectSelector (Database, SubProvider)>
<!ELEMENT SubProvider (NameAbbr, (InclQuery|ExclQuery)*)>
<!ELEMENT InclQuery   (#PCDATA)>
<!ELEMENT ExclQuery   (#PCDATA)>
<!-- End of SubObjectSelector group -->
<!ELEMENT ObjectUrl   (((Base, (Rule|RuleToMany)?)
                        | (Rule|RuleToMany)), UrlName?,
                        SubjectType*, Attribute*)>
<!ATTLIST ObjectUrl   LNG      %iso.lang.codes;  "EN"  >
<!ELEMENT Base        (%url.data;)>
<!ELEMENT pad         (%rule.data;)*>
<!-- used to mark up strings that require padding to create fixed-length string

```

```

    attributes. -->
<!ATTLIST pad
    with          CDATA          #REQUIRED
    width         CDATA          #REQUIRED
    align         (left|right)   "right">
<!-- with      a character to pad with (required)
    width      result string size (required) (integer)
    align      the text should align to (left|right) (default:right)
-->

<!ELEMENT apad          (%rule.data;)*>
<!-- used to mark up strings that require padding to create fixed-length string
    attributes. It skips all preceding alpha characters before padding -->
<!ATTLIST apad
    with          CDATA          #REQUIRED
    width         CDATA          #REQUIRED
    align         (left|right)   "right">
<!-- with      a character to pad with (required)
    width      result string size (required) (integer)
    align      the text should align to (left|right) (default:right)
Examples:
    <apad with = "0" width = "6">E32</apad>  => "E00032"
    <apad with = "0" width = "6">640</apad> => "000640"
-->

<!ELEMENT subs          (%rule.data;)*>
<!-- substitute one string for another in the element's content
-->
<!ATTLIST subs
    for          CDATA          #REQUIRED
    with         CDATA          #REQUIRED>
<!-- for      the string to replace
    with      the string to substitute
-->

<!ELEMENT toupper      (%rule.data;)*>
<!-- convert element content into upper case -->
<!ELEMENT tolower     (%rule.data;)*>
<!-- convert element content into lower case -->
<!ELEMENT strip       (%rule.data;)*>
<!-- strip off spaces/letters/digits in the element's content -->
<!ATTLIST strip
    what (spaces|letters|nondigits|digits) #REQUIRED>
<!ELEMENT normalize   (%rule.data;)*>
<!-- are used for normalization of &lo.voll , &lo.iss; elements.
Examples:
    "Pt 5"             ==> "5"
    "3 Suppl"         ==> "3"
    "2A Pt 3"         ==> "2A"
    "10 Suppl 2 Pt 1" ==> "10"
    "2/3"             ==> "2"
-->

<!ELEMENT Rule          (%rule.data;)*>
<!ELEMENT UrlName      (#PCDATA)>
<!ELEMENT RuleToMany   (Rule, Separator)>
<!ELEMENT Separator    (#PCDATA)>
<!-- End of ObjectUrl group -->
<!-- End of Link group -->
<!-- End of LinkSet group -->

```

Databases Available for Linking

Links can be made from the following Entrez databases to outside resources. When entering the database in the <Database> element, use the Db abbreviation, rather than the database name.

Database Name	Db Abbreviation
http://www.ncbi.nlm.nih.gov/sites/entrez?db=PubMed	PubMed
http://www.ncbi.nlm.nih.gov/sites/entrez?db=Protein	Protein
http://www.ncbi.nlm.nih.gov/sites/entrez?db=Nucleotide*	Nucleotide
Core Nucleotide	NucCore
Nucleotide EST	NucEST
Nucleotide GSS	NucGSS
http://www.ncbi.nlm.nih.gov/sites/entrez?db=Structure	Structure
http://www.ncbi.nlm.nih.gov/sites/entrez?db=Genome	Genome
http://www.ncbi.nlm.nih.gov/sites/entrez?db=genomeprj	GenomePrj
http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene	Gene
http://www.ncbi.nlm.nih.gov/sites/entrez?db=gensat	GENSAT
http://www.ncbi.nlm.nih.gov/sites/entrez?db=geo	Geo
http://www.ncbi.nlm.nih.gov/sites/entrez?db=gds	GDS
http://www.ncbi.nlm.nih.gov/sites/entrez?DB=homologene	HomoloGene
http://www.ncbi.nlm.nih.gov/sites/entrez?db=nlmcatalog	NlmCatalog
http://www.ncbi.nlm.nih.gov/sites/entrez?db=omim	OMIM
http://www.ncbi.nlm.nih.gov/sites/entrez?db=popset	PopSet
http://www.ncbi.nlm.nih.gov/sites/entrez?db=pcassay	PCAssay
http://www.ncbi.nlm.nih.gov/sites/entrez?db=pccompound	PCCompound
http://www.ncbi.nlm.nih.gov/sites/entrez?db=pcsubstance	PCSubstance
http://www.ncbi.nlm.nih.gov/sites/entrez?db=snp	SNP
http://www.ncbi.nlm.nih.gov/sites/entrez?db=Taxonomy	Taxonomy
http://www.ncbi.nlm.nih.gov/sites/entrez?db=unigene	UniGene
http://www.ncbi.nlm.nih.gov/sites/entrez?db=unists	UniSTS

*Use "Nucleotide" to apply links to records in all three Nucleotide databases. To apply links to records in a specific Nucleotide database only, use the Db Abbreviation for that database.

Icons

Icons should:

- Be 100 x 25 pixels. Maximum size allowed is 150 x 35 pixels.
- Be in one of the following image formats: GIF, JPEG, or PNG
- Identify the resource being linked. For libraries, this could be the library name and the format of the resource (full-text/print collection). For resource providers, this could be the name of the resource and/or the format or language in which the resource is being delivered. If the item is freely available, or free with a registration, the icon should indicate this.
- Look like an actionable button. The icon should have a border and/or a colored background. Icons and text should be easy to read and should not blend in to the PubMed display.
- Not be animated.

The icon file name should use only letters from the English alphabet and numbers. Do **not** use special characters or spaces. Examples of file names are: libft.gif, libprint.gif, freeft.jpg, notfree.jpg.

Examples of icons are shown below.

Library icons



Other provider icons



Identity File

The identity file contains the information about a provider that NCBI needs to list the provider in LinkOut. This file must be named **providerinfo.xml**; the file name is case sensitive.

Section Contents

- Format
- Prolog
- Elements

Identity File Format

Note: The following format description shows the basic structure of an Identity file and includes all possible elements. Not all elements will be used in all files. Additional examples of Identity files are available for full-text providers, other resource providers, and libraries.

```
<?xml version="1.0"?>
<!DOCTYPE Provider PUBLIC "-//NLM//DTD LinkOut 1.0//EN"
"http://www.ncbi.nlm.nih.gov/projects/linkout/doc/LinkOut.dtd">
<Provider>
<ProviderId>ProviderId assigned by NCBI(required)</ProviderId>
  <Name>Name of Resource Provider or Library(required)</Name>
  <NameAbbr>NameAbbr assigned by NCBI(required)</NameAbbr>
  <SubjectType>SubjectType (optional, repeatable)</SubjectType>
  <Attribute>Attribute (optional, repeatable)</Attribute>
  <Url>URL of provider's website(optional, repeatable)</Url>
  <IconUrl>Currently not in use. (optional, repeatable)</IconUrl>
  <Brief>Short description of resource provider.(optional)</Brief>
</Provider>
```

Identity File Prolog

The Prolog occurs once at the beginning of the file.

XML Declaration (optional): <?xml version="1.0"?>

Document Type Declaration (required): <!DOCTYPE Provider PUBLIC "-//NLM//DTD LinkOut 1.0//EN" "http://www.ncbi.nlm.nih.gov/projects/linkout/doc/LinkOut.dtd">

Identity File Elements

The following lists all elements that may be included in the Identity File. Not all elements will be used in every Identity file.

Provider (*required*): Root element of the file

ProviderId (*required*): A unique ID assigned by NCBI

Name (*required*): The full name of the resource provider or library. For libraries, this is the name of the library or institution; for other providers, this is the name of the organization.

NameAbbr (*required*): A short one word name for the resource provider or library. NameAbbr may only include alpha and numeric characters; spaces and special characters such as hyphens are not allowed. Because this element must be unique, NCBI reserves the right to adjust it as necessary. Resource providers who are submitting, or may submit, links on behalf of multiple publishers should select a neutral NameAbbr.

SubjectType (*optional, repeatable*): SubjectTypes are used to display resources by Subject Category. SubjectTypes included in the identity file apply to all resources provided by a provider. See Special Elements: SubjectType for a list of SubjectTypes and descriptions. If no SubjectType is given, the SubjectType "miscellaneous" will be assigned automatically.

Attribute (*optional, repeatable*): Attributes describe resources independent of content and describe any ownership of the information that is being claimed by the individual or organization providing the link. Attributes included in the Identity file apply to all resources provided by a provider. See Special Elements: Attribute for the list of Attributes and descriptions.

Url (*optional, repeatable*): URL of the library or resource provider's website, used in the LinkOut Providers list. *Note:* The URL may be different for different languages; see the LNG attribute in the LinkOut.DTD.

IconUrl (*optional, repeatable - currently not being displayed. The library or resource provider icon being displayed is derived from the holdings file.*) The icon should meet the specifications described in Icons.

Note: The IconUrl may be different for different languages; see the LNG attribute in the LinkOut.DTD. The LNG attribute is not currently displayed.

Brief (*optional - displayed in My NCBI*): Short description of the resource provider. Maximum is 255 characters.

The **providerinfo.xml** file is specified in the LinkOut DTD.

Resource File

The resource file describes the Entrez records the provider will link from and contains the information that LinkOut needs to generate the links. Links described in the resource file should link directly to the resource; users should not have to perform any additional searching to access the resource after clicking the provider's link.

Resource files which link to online full text must be submitted as XML files. Libraries may use the Library Submission Utility to generate XML files on their behalf. Links to resources other than full text may be submitted as an XML file or as a simple text file.

Note: Resource files should be composed in a text editor, such as NotePad, rather than in a word processing program.

Resource File: XML

Files must have a file extension **.xml**; the file extension is case sensitive. The file name is assigned by the provider. File names should include alpha and numeric characters only. Spaces and special characters such as hyphens are not allowed. Examples of file names are **journals.xml**, **journaltitle.xml**, **resourcename.xml**. File size may not exceed 20 MB. To help with file management, a provider may supply more than one resource file.

Section Contents

- Format

- Prolog
- Elements

Resource File Format

Note: The following format description shows the basic structure of a Resource file and includes all possible elements. Not all elements will be used in all files. Additional examples of Resource files are available for full-text providers, providers submitting holdings on behalf of libraries, other resource providers, and libraries.

Using ObjectSelector (standard format)

```
<?xml version="1.0" ?>
<!DOCTYPE LinkSet PUBLIC "-//NLM//DTD LinkOut 1.0//EN"
"http://www.ncbi.nlm.nih.gov/projects/linkout/doc/LinkOut.dtd">
<LinkSet>
  <Link>
    <LinkId>Identifier assigned to link by provider(required)</LinkId>
    <ProviderId>ProviderId assigned by NCBI(required)</ProviderId>
    <IconUrl>URL of icon to be displayed on Entrez records(required)</IconUrl>
    <ObjectSelector>
      <Database>Entrez database in which links will appear(required)</Database>
      <ObjectList>
        <Query>Entrez search retrieving records to be linked(repeatable,
          required unless <ObjId> is specified</Query>
      </ObjectList>
    </ObjectSelector>
    <ObjectUrl>
      <Base>Stable portion of URL for provider's resource(required)</Base>
      <Rule>Variable portion of URL for resource (required)</Rule>
      <SubjectType>SubjectType (optional, repeatable)</SubjectType>
      <Attribute>Attribute (optional, repeatable)</Attribute>
      <UrlName>Short text string describing resource (optional)</UrlName>
    </ObjectUrl>
  </Link>
</LinkSet>
```

Using SubObjectSelector (for providers submitting links on behalf of libraries only)

```
<?xml version="1.0" ?>
<!DOCTYPE LinkSet PUBLIC "-//NLM//DTD LinkOut 1.0//EN"
"http://www.ncbi.nlm.nih.gov/projects/linkout/doc/LinkOut.dtd">
<LinkSet>
  <Link>
    <LinkId>Identifier assigned to link by provider(required)</LinkId>
    <ProviderId>ProviderId assigned by NCBI(required)</ProviderId>
    <IconUrl>URL of icon to be displayed on Entrez records(required)</IconUrl>
    <SubObjectSelector>
      <Database>Entrez database in which links will appear(required)</Database>
      <SubProvider>
        <NameAbbr>LinkOut NameAbbr of the SubProvider(required)</NameAbbr>
        <InclQuery>Title Abbreviation or ISSN(required unless ExclQuery is
          specified, repeatable)</InclQuery>
      </SubProvider>
    </SubObjectSelector>
    <ObjectUrl>
      <Base>Stable portion of URL for provider's resource(required)</Base>
      <Rule>Variable portion of URL for resource (required)</Rule>
      <SubjectType>SubjectType (optional, repeatable)</SubjectType>
      <Attribute>Attribute (optional, repeatable)</Attribute>
      <UrlName>Link Resolver or Service Name</UrlName>
    </ObjectUrl>
  </Link>
</LinkSet>
```

Resource File Prolog

The Prolog occurs once at the beginning of the file.

```

XML Declaration (optional): <?xml version="1.0"?> (optional)
Document Type Declaration (required): <!DOCTYPE LinkSet PUBLIC "-//NLM/DTD LinkOut 1.0//EN" "http://www.ncbi.nlm.nih.gov/projects/linkout/doc/LinkOut.dtd"
[ <!ENTITY icon.url "http://www.goodpublisher.com/icon/goodpublisher.jpg">
<!ENTITY base.url "http://www.goodmedical.org/cgi/content/" > ]>

```

The Entity Declaration, included in the Document Type Declaration, is optional. **ENTITY** may be specified to be used in the body of the file. In the above example, **ENTITY** icon.url is defined with the value "http://www.goodmedical.org/images/icon.gif", and **ENTITY** base.url is defined with the value "http://www.goodmedical.org/cgi/content/".

Once an **ENTITY** is defined in the **Prolog**, it can be used in the holding file by placing the **ENTITY** name between an ampersand (&) and semicolon (;). This alleviates the need to replicate long, textual data. In the above example, '&icon.url;' is used to represent the URL of the provider icon. Using **ENTITY** icon.url will make it easier to change the icon URL if necessary, because it will only need to be changed in one place.

Resource File Elements

The following lists all elements that may be included in the Resource file. Not all elements will be used in every Resource file.

LinkSet (required): Root element of the holdings file

Link (required, repeatable): An element that describes a specific set of resources grouped together by access characteristics or for convenience. A holdings file can have multiple **Link** elements. Resources that can be accessed via a single URL **Rule** are usually grouped in one **Link**. To help with file management, you may choose to put holdings in different **Link** elements, even if the same URL **Rule** applies to all holdings.

LinkId (required): An identifier assigned by a participating provider for its own reference. The **LinkId** can be any alpha-numeric string. Each **Link** within a **LinkSet** or file must have a unique **LinkId**.

ProviderId (required): The identifier number assigned to the publisher by NCBI and listed in the providerinfo.xml file.

IconUrl (required, repeatable): The URL of the icon that will be displayed on the Entrez records described within the **Link** element. The icon should meet the specifications described in Icons. IconUrl is repeatable as defined in the LinkOut DTD. However, this functionality is not implemented. Icons appear only in PubMed.

ObjectSelector (required unless *SubObjectSelector* is specified): An element containing other sub-elements in which a provider will specify which Entrez records are being linked from by a **Link** element.

SubObjectSelector (required unless *ObjectSelector* is specified): An element containing other sub-elements in which journals and full-text providers for those journals are specified. For use by libraries and providers submitting information on behalf of libraries.

Database (required): A sub-element of **ObjectSelector** or **SubObjectSelector** that specifies the Entrez database in which the links will appear.

ObjectList (required if *ObjectSelector* is specified, repeatable): A sub-element of **Object Selector**, containing **Query** or **ObjId** elements, that specifies which Entrez records the URL(s) defined in **ObjectUrl** will link from.

SubProvider (required if *SubObjectSelector* is specified): A sub-element of **SubObjectSelector**. **SubProvider** contains **NameAbbr** and **InclQuery** elements. This element describes the publishers or full-text providers with whom a library has a pre-arranged access agreement.

Query (required sub-element of *ObjectList* unless *ObjId* is specified, repeatable): A sub-element of **ObjectList** that contains any valid Entrez search used to select the records in the database specified in **Database** that will be linked from.

ObjId (required sub-element of *ObjectList* unless *Query* is specified, repeatable): A sub-element of **ObjectList** that contains the Unique Identifier for an Entrez record.

InclQuery (required sub-element of *SubProvider* unless *ExclQuer* is specified, repeatable): A sub-element of **SubProvider** that includes a journal supplied by the **SubProvider** as a part of the online journal holdings of a library. Use the MEDLINE Journal Title Abbreviation or ISSN. Both are available in the list of LinkOut Journals.

ExclQuery (required sub-element of *SubProvider* unless *InclQuery* is specified, repeatable): A sub-element of **SubProvider** that excludes a journal supplied by the **SubProvider** from the online journal holdings of a library. Use the MEDLINE Journal Title Abbreviation or ISSN. Both are available in the list of LinkOut Journals.

ObjectUrl (*required*): An element that describes the information necessary for the Entrez system to construct URLs to access the full-text for Entrez records specified in the **ObjectList**. **ObjectUrl** contains two required sub-elements, **Base** and **Rule**.

Base (*required*): A sub-element of **ObjectUrl** that is the stable portion of the URL for the provider's web-accessible resource. This can be defined as an ENTITY in the Entity Declaration.

Rule (*required*): A sub-element of **ObjectUrl**. The **Rule** will be concatenated with the **Base** to create a valid URL when a record is retrieved from an Entrez search. Allowable Keywords in the **Rule** will be translated into the correct values for the specific Entrez record. The **Rule** is based on the provider's specifications for access to the resource. Links should point directly to the resource.

SubjectType (*optional, repeatable*): SubjectTypes are used to display resources by subject category. SubjectTypes apply to all resources identified within a **<Link>**. See Special Elements: SubjectType for the list of SubjectTypes and descriptions. If no SubjectType is given, the resource will be displayed under the category "Miscellaneous".

SubjectType is defined as repeatable in the LinkOut DTD, but in practice repetition of this element is more problematic than useful and is strongly discouraged.

Attribute (*required, repeatable*): Attributes describe resources independent of content and describe any ownership of the information that is being claimed by the individual or organization providing the link. Attributes apply to all resources identified within a **<Link>**. See Special Elements: Attribute for the list of Attributes and descriptions.

UrlName (*optional*): A short text string that describes a resource. UrlName is typically used when the allowed SubjectTypes and Attributes cannot meet the needs of a provider.

The resource file is specified in the LinkOut DTD.

Resource File: Simple Text

Files linking to resources other than full text may be submitted as a simple text file, rather than an XML file.

Files must have a file extension **.ft**; the file extension is case sensitive. The file name is assigned by the provider.

File names may only include alpha and numeric characters, spaces and special characters such as hyphens are not allowed. Example file names might be **resources.ft**, **resourceName.ft**.

File size may not exceed 20 MB. To help with file management, a provider may supply more than one resource file.

Files submitted as text files will be translated into valid XML by LinkOut. Translated files will conform to the LinkOut DTD.

File Syntax

The file consists of blocks separated by comment lines beginning with "-":

```
----- lines starting with "-" are comments -----
----- information in the first block is global-----
prid: [required]
dbase: [required]
stype: [optional]
attr: [optional]
!base: [optional]
-----
linkid: [required - LinkOut will add a suffix to make sure it is unique]
uids: [required if query is not present]
query: [required if uids is not present]
base: [required]
rule: [required]
icon: [optional]
name: [optional]
stype: [optional]
attr: [optional]
-----
```

The first block holds global information that will be used throughout the file.

```

prid: LinkOut Provider ID
dbase: Entrez database that will be hosting the links, e.g., pubmed, nucleotide, taxonomy
stype: SubjectType. See Special Elements: SubjectType for all available SubjectTypes.
attr: Attribute. See Special Elements: Attribute for all available Attributes.

```

Each subsequent block specifies a link to the provider's resource. This has two basic parts, an Entrez query and a URL pointing back to the provider's site. Each Entrez query will be evaluated, and any Entrez records that are retrieved will be assigned a link to the URL specified by the provider. An example of simple blocks is shown below.

```

-----
query: <species1> [name]
rule: <http://...my_species1_page.html>
-----
query: <species2> [name]
rule: <http://...my_species2_page.html>
-----

```

Any valid Entrez query may be used to select records. See Entrez Help for information on constructing Entrez queries. Each query should appear on a single **query** line. Multiple query lines in one block will be OR-ed together. The following two link descriptions are equivalent:

Link Description 1:

```

-----
query: <species1> [name] OR <species2> [name]
rule: <http://...my_species1or2_page.html>
-----

```

Link Description 2:

```

-----
query: <species1> [name]
query: <species2> [name]
rule: <http://...my_species1or2_page.html>
-----

```

Objects in each Entrez database are indexed by a numerical unique identifier. For example, taxid in taxonomy, gi in nucleotide and protein, pmid in PubMed. You can specify the Entrez records you would like to link from by listing them directly with the **uids**: line.

Link description with uids:

```

--- separate unique identifiers (UIDs) with a blank space.---
---Each new line of UIDs should start have the "uids" label---
uids: 123456 123469 3847559
uids: 4755983 2197866
rule: <http://...my_species1or2_page.html>
-----

```

The optional fields allow you to customize your links in several ways.

```

icon: URL of an icon file that you would like to represent your link and resources. Only applicable to links in PubMed.
name: Additional description about the link.
stype: SubjectType. See Special Elements: SubjectType for all available SubjectTypes.
attr: Attribute. See Special Elements: Attribute for all available Attributes.

```

Special Elements: SubjectType

The <SubjectType>, <Attribute>, and <UrlName> elements are used to organize and describe resources in the Entrez displays and to facilitate access to these resources by defining them as groups.

SubjectType is used to group resources in the LinkOut Display, making it easier to browse through available LinkOut resources. In most cases, resources are grouped by Subject Category (described below).

Only one SubjectType should be applied to each Link. If no SubjectType is specified for a Link, it will be displayed under the heading "Miscellaneous".

SubjectTypes are also used in My NCBI, allowing users to display records with links to resources from a specific SubjectType or Subject Category as a filter tab, and to activate icons for these links.

SubjectType can be used in Entrez searches to retrieve records with links to resources from a specific SubjectType.

To do this, enter the following the in the search box.

!osubjtSubjectTypeName [sb]

Replace **SubjectTypeName** with the SubjectType name from the list below.

This filter can be used as a part of any Entrez search.

SubjectType Categories

SubjectTypes are grouped into the following categories:

- Chemical Information
- Education
- Funding Sources
- Literature
- Medical
- Molecular Biology Databases
- Research Materials
- Researchers
- Tools

Note: Category headings CANNOT be used in the the <SubjectType> element.

Chemical Information

biological properties: Information about chemicals with biological activity.

chemical libraries: Link to libraries of chemicals provided by depositors, for example screening centers.

imaging agents: Information on radionuclide-containing chemicals which bind to specific sites/molecules and help to detect motion/binding.

Metabolism: Information about chemicals participating in metabolism.

molecular interactions: Information about chemicals which interact or bind to other molecules (for example, protein-protein or small molecule-protein interactions)

physical properties: Information about chemicals with determined physical properties.

reactions: Information on chemicals which used as reactants/products/intermediates/catalysts.

theoretical properties: Information about chemicals with calculated properties.

toxicology: Information on chemicals investigated for their toxic effect.

vendors: Information from commercial distributors of chemicals.

Education

These SubjectTypes describe links from Entrez records to resources offering educational content, as well as those providing information about educational opportunities.

conferences/meetings/workshops: Resources offering announcements of, or information on upcoming scientific conferences, meetings, and workshops relevant to the subjects of specific Entrez records.

glossaries/dictionaries: Resources offering glossaries or dictionaries of technical terminology or other specialized vocabularies used in research relevant to specific Entrez records.

online tutorials/courses: Resources offering information about or providing Web-based primers, tutorials, or courses on specialized areas of research relevant to specific Entrez records.

Funding Sources

The SubjectType in this category describes links to resources within information on funding for research.

funding sources: Resources providing information on locating, applying for, or obtaining funding for particular areas of research relevant to the subjects of specific Entrez records.

Literature

These SubjectTypes describe links to literature databases and resources from Entrez records (typically PubMed records).

SubjectTypes under the "Literature" category appear in the LinkOut Display as follows:

- Resources from "aggregators", "individual online article", "publishers/providers" are grouped under the heading "Full Text Sources".
- Links with SubjectType "libraries" are displayed under the heading "Libraries".
- The SubjectTypes: "author profiles", "books", "document delivery", "images", and "supplemental materials" are displayed under the heading "Other Literature Sources".

aggregators: Services that provide a single point of online access to multiple full-text publications cited in Entrez records. "aggregators" are differentiated from "publishers/providers" (see below) by their aggregation of online access to the publications of multiple publishers and/or providers. See LinkOut: Information for Full-Text Providers for details on LinkOut and full-text.

author profiles: Resources offering biographical information and /or research interest of authors of a PubMed citation.

books: Links to relevant information in a book.

commentaries/discussion: Links to commentaries or discussion to extend/continue the topic discussed in the article cited in a PubMed citation.

document delivery: Links to vendors that deliver electronic or print copies of documents to users offline.

images: Links to online images from the Entrez records that they are relevant to or referenced by.

individual online article: Links from Entrez records to individual online copies of the publications that they reference (i.e., a link from a PubMed record to a copy of a research article that the author has made available on her/his personal website or on a depository).

libraries: Links to the online full text and print holdings of a library. Links to online holdings will point directly to full-text articles; links to print holdings will point to a record of the library's print holdings.

patent databases: Links to patent database records from PubMed citations of publications discussing patented products (including proteins and DNA sequences) or processes, or from relevant Entrez records.

publishers/providers: Links to online full-text articles offered by publishers or their agents.

supplemental materials: Links to the materials directly supporting the research discussed in the cited article, including data sets from experiments/studies, accessory graphics, images, sound, and multimedia files related to the article.

Medical

These SubjectTypes describe links from Entrez records to relevant online medical resources.

clinical trials: Links from Entrez records to resources providing information about relevant ongoing and upcoming clinical trials, including information on locating, applying to, or participating in clinical trials.

consumer health: Links from Entrez records to relevant resources or materials intended to educate or inform the general public about particular health topics or issues.

diagnostics: Resources providing information on the signs, symptoms and diagnosis of diseases, syndromes, or conditions relevant to specific Entrez records.

disease organizations: Institutions, groups, or other organizations whose work is focused on the research of particular diseases or conditions, or that offer information, support, and/or other resources to people suffering from particular diseases or conditions, linked from relevant Entrez records.

medical equipment and devices: Organizations or companies providing medical equipment or other medical devices relevant to specific Entrez records.

pharmacology: Resources providing information on the pharmacological aspects of substances relevant to specific Entrez records.

treatment guidelines: Resources offering treatment guidelines or information on treatment guidelines for particular diseases, syndromes, or conditions relevant to specific Entrez records.

Molecular Biology Databases

These SubjectTypes describe links from Entrez records to corresponding or related records in other online molecular biology databases.

DNA/protein sequence: Databases providing nucleotide or amino acid sequence data.

gene/protein/disease-specific: Databases providing data specific to particular genes, proteins, diseases, syndromes, or conditions.

gene expression: Databases providing data on gene expression (e.g., RNA transcription, knockouts).

locus-specific: Databases, often described as LSDB, focusing on a single gene or a small family of genes.

mapping: Databases providing chromosomal mapping data, such as sequence tagged site (STS) mapping.

meta-databases: Databases that draw and/or synthesize information from other databases and extend or add value to the original information.

organism-specific: Databases providing data specific to a particular organism or group of organisms.

population/variation: Databases providing data on sequence variation within particular populations, including mutations or other polymorphisms.

protein interactions/pathways: Databases that describe molecular interactions and/or biological pathways.

structure: Databases providing structural or pattern data, such three-dimensional protein structures, chemical structures, or two-dimensional electrophoresis patterns.

taxonomy/phylogenetic: Databases providing data on organism taxonomy, phylogenetic trees and relationships, or biodiversity.

Research Materials

These SubjectTypes describe links to relevant research material suppliers from Entrez records that discuss the use of these materials.

clones/clone libraries: Suppliers providing DNA clones and clone libraries (e.g., DNA, cDNA).

culture/stock collections: Suppliers providing cell cultures.

herbarium/museum collections: curated plant/animal specimen collections.

laboratory equipment: Suppliers providing laboratory equipment.

oligonucleotides: Suppliers providing equipment, materials, and reagents used in researching oligonucleotides.

other reagents: Suppliers providing other reagents not described by the SubjectTypes above.

Researchers

These SubjectTypes describe links from Entrez records to the individuals or organizations that conducted or supported the research cited therein.

colleges/universities: Colleges, universities or other academic institutions that have conducted or supported research cited in Entrez records.

companies/research institutes: Private companies or research institutions that have conducted or supported research cited in Entrez records.

directories: Links to the directory entries of individuals or organizations from Entrez records that cite research conducted or supported by those individuals or organizations.

individuals: The web pages of individuals who have conducted research cited in Entrez records.

societies/associations: Societies or organizations that have conducted or supported research cited in Entrez records.

Tools

These SubjectTypes describe links to research tools, utilities or other software from Entrez records citing research that discusses or requires the use of these tools.

3D structure prediction/functional modeling: Tools used in predicting the three-dimensional structure of proteins or in modeling protein function.

primer design: Tools used in the design of nucleotide primer sequences for use in polymerase chain reactions.

protein identification/characterization: Tools used in the identification, annotation, characterization, or other description of proteins.

restriction mapping: Tools used in the mapping of restriction enzyme cutting sites.

sequence screening/similarity/alignment: Tools used in screening, comparing, matching, and alignment of sequences.

sequence viewer: Tools used in creating visual representations of nucleotide or amino acid sequences.

translation: Tools used in the translation of sequences or of sequence file formats.

Special Elements: Attribute

The <SubjectType>, <Attribute>, and <UrlName> elements are used to organize and describe resources in the Entrez displays and to facilitate access to these resources by defining them as groups.

The Attribute element is used to provide information about the resource itself, independent of content. It describes the format, barriers to access, and the ownership (if any) of the resources being linked. On the basis of information provided by Attribute, users can make an informed choice from a list of LinkOut resources available. For example, they will know if a resource is free before clicking the link.

Attributes can be used in Entrez searches to retrieve records with links to resources with a specific Attribute.

To do this, enter the following the in the search box.

loattrAttribute*Name* [sb]

Replace **Attribute***Name* with the Attribute name from the list below.

This filter can be used as a part of any Entrez search.

Special aliases are also available for these searches:

Full text [sb]

Free full text [sb]

Free resource [sb]

For each resource, the provider should select all Attribute values from the list below that apply and include them in the appropriate LinkOut files.

Attribute Categories

Attributes are grouped in the following categories:

Barriers
 Ownership
 Resource Form
 Miscellaneous

Barriers

These Attributes describe barriers to accessing the information being linked to from an Entrez record.

registration required: A free registration is required to access the information being linked. If any expense is involved, please use the Attribute “subscription/membership/fee required”. Citations with this Attribute will be marked with an  icon in the LinkOut Display.
subscription/membership/fee required: Membership in an organization, a subscription, or some other fee or expense is required to obtain access to the information being linked.

If no Barrier Attributes are used, citations will be marked with a **FREE** icon in the LinkOut Display. Free full-text sources will be also marked with the free full-text icon  in PubMed’s Summary display.

Ownership

These Attributes describe any ownership of the information that is being linked claimed by the individual or organization providing the link.

author of URL: The individual or organization providing a link is the author of the information that is being linked.
publisher of information in URL: The individual or organization providing a link is the publisher of the information that is being linked.
 Citations with this Attribute will be marked with an  icon in the LinkOut Display.

Resource Form

These Attributes describe the form of the documents being linked from Entrez records.

author manuscript: Final, peer reviewed manuscripts that have been accepted for publication. These manuscripts are typically posted on a journal website or at PubMed Central.
electronic full-text: Online full-text holdings of a library.
full-text online: Full-text documents that are in text, HTML, word processor, or other formats (besides PDF or PostScript).
full-text PDF: Full-text documents that are in Adobe Portable Document Format (PDF).
full-text PostScript: Full-text documents that are in PostScript format.
order form: Order form for the resources mentioned in the Entrez record, e.g. clones.
print collection: Print collection of a library.

Miscellaneous

Attributes not included in other categories.

library-local: Full-text documents hosted in a library or institution for local use. Links to items described as library-local will not be displayed in the LinkOut Display or in the Library Submission Utility.
preference: When duplicate URLs are submitted for the same Entrez record but different IconUrl, SubjectType, and/or Attributes are assigned, use **preference** to indicate which link should be displayed. See Duplicate Links for details.
 When a provider submits multiple, unique, URLs for a PubMed record, the provider should use **preference** to indicate which URL should be linked through their icon in PubMed. See Multiple Links for details.

Special Elements: UriName

UriName is a short text string used to describe a resource. If possible, <SubjectType> and <Attribute> should be used, rather than UriName. UriName should be used only when the allowed terms in the SubjectType and Attribute lists do not suffice to describe the resource. UriName is not searchable.

When UriName is used, it will appear in the LinkOut Display. In the example below, the first two providers have used UriName to provide additional information about the resource being linked. The link appears as the description given in UriName. The third provider has not used UriName, so the link appears as the resource name.

The screenshot shows a LinkOut interface with the following elements:

- Display: LinkOut (dropdown)
- Show: 20 (dropdown)
- Sort by: (dropdown)
- Send to: (dropdown)
- Filters: All: 1, bacteria: 0, mRNA: 1, RefSeq: 0
- Text: "The following [LinkOut](#) resources are supplied by external providers. These providers are responsible for maintaining the links. What does the [icon](#) mean?"
- Resource 1: 1: [AM000002](#). Reports Mus musculus mRNA...[gi:67511481] Links
 - ⊖ Molecular Biology Databases
 - FREE Jackson Laboratory Mouse Strains Index
 - [inbred strain BALB](#)
 - ⊖ Research Materials
 - TaqMan® probe and primer sets from Applied Biosystems
 - [Gene Expression Assays](#)
 - ⊖ Tools
 - FREE UCSC Genome Browser

Allowable Rule Keywords

As in all XML files, entities can be defined and used in LinkOut files to facilitate file construction.

There are a number of pre-defined Linkout entities (Allowable Rule Keywords) in the LinkOut DTD. When an entity is used in the <Rule> element, the Entrez system will replace the entity with the corresponding value from each Entrez record when constructing the URL specified in your <Rule>.

The following is a list of entities and the values with which they will be substituted:

Entity	Corresponding value from Entrez record
&lo.id;	Unique identifier (PMID, GI, TaxID, etc.)

For PubMed only:

Entity	Corresponding value from Entrez record
&lo.auth;	First Author. Example: Smith JE.
&lo.authln;	First Author Last Name. Example: Smith.
&lo.day;	Two-digit day of the publication date. Example: 01; 31.
&lo.doi;	Article DOI. Must be submitted by Publisher.
&lo.eday;	Two-digit day of electronic publication date. Example: 01; 31.
&lo.emo;	Two-digit month of electronic publication date. Example: 01; 12.
&lo.emon;	Three-letter month abbreviation of electronic publication date. Example: Sep.
&lo.emonth;	Month of electronic publication date. Example: September.
&lo.eyear;	Four-digit year of electronic publication date. Example: 1998.
&lo.eyl;	Last digit of year of the electronic publication date. Example: 1999 =9.
&lo.eyr;	Last two digits of year of electronic publication date. Example: 98; 00.
&lo.essn;	Journal electronic ISSN code.
&lo.iss;	Journal issue.
&lo.issn;	Journal print ISSN code.
&lo.issnl;	Journal print ISSN code with the dash stripped.
&lo.jtit;	Journal title (NLM's Title Abbreviation).
&lo.mo;	Two-digit month abbreviation of the publication date. Example: 01.
&lo.mon;	Three-letter month abbreviation of the publication date. Example: Sep.
&lo.month;	Month of the publication date. Example: September.
&lo.otit;	Article title.
&lo.page;	First page.
&lo.pii;	Publisher Item Identifier. Must be submitted by Publisher.
&lo.vol;	Journal volume.
&lo.year;	Four-digit year of the publication date. Example: 1998.
&lo.yl;	Last digit of year of the publication date. Example: 1999 =9.
&lo.yr;	Last two digits of year of the publication date. Example: 98.

For Sequence databases (Nucleotide, Protein, Structure, Genome):

Entity	Corresponding value from Entrez record
&lo.pacc;	Primary accession for sequences.

For Taxonomy only:

Entity	Corresponding value from Entrez record
&lo.genus;	Genus name. Example: "Homo"
&lo.scientificname;	Scientific name. Example: "Homo sapiens neanderthalensis"
&lo.species;	Species epithet. Example: "sapiens"
&lo.subsp;	Sub-species epithet. Example: "neanderthalensis"

For UniGene only:

Entity	Corresponding value from Entrez record
&lo.clusterid;	Cluster Id.

Special Characters

XML requires certain special characters to be encoded in the files. When using the following characters in LinkOut files, *except in LinkOut entities*, please encode them as shown.

Character	XML Encoding
&	&
<	<
>	>
©	©
®	®
™	™

LinkOut does not support Unicode (UTF-8).

Do not use ASCII encoding in LinkOut files.

Transferring Files via FTP

Providers must write to linkout@ncbi.nlm.nih.gov to apply for inclusion in LinkOut before submitting files.

LinkOut files are submitted to NCBI via FTP to the host: <ftp-private.ncbi.nlm.nih.gov>. The following file name conventions apply:

- File names may include alpha and numeric characters only. Spaces and special characters such as hyphens are not allowed.
- Identity and Resource files must have the file extension “.xml”. The exception is resource files submitted as a simple text file.
- Identity files must be named “providerinfo.xml”.

File Placement

- Resource provider files (both full-text and other providers) are placed in the “holdings” directory of the FTP account.
- Files submitted on behalf of libraries by full-text providers are placed in the “users “ directory.
- Files submitted by libraries are placed in the “collection” directory
- “.doi” or “.pii” files are placed in the “pid” directory .

Files placed in an incorrect directory will not be processed.

How to FTP

From a standard ftp client:

1. At a command prompt type: ftp-private.ncbi.nlm.nih.gov and press enter.
2. Type your login name at the login prompt and press enter. (See how to obtain ftp account from NCBI.)
3. Type your password at the password prompt and press enter.
4. You should now be logged into the ftp server. If you receive an error message, check your login information, type 'bye' followed by enter, and retry steps 1-3.
5. Type 'bin' and press enter. This changes your ftp server to BINARY mode.
6. Type 'cd [directory, i.e. holdings, users]' and press enter. This changes your current directory.
7. Type the 'put' command, followed by your pathname or drive and filename and press enter. (For example, 'put C:\filename' or 'put /home/testfiles/journalv6n3'.)
8. Type 'dir' and press enter to display the files in the current directory.
9. Type 'bye' to disconnect from the server and close the ftp session.

From a browser window:

1. Enter: ftp://ftp-private.ncbi.nlm.nih.gov in your browser's "Address" field and press enter.
2. From the file pull-down menu, choose "Login As".
3. Type in your username and password in the dialog box. The screen will display your directory with subdirectory folders inside.
4. Open the folder into which you will be depositing files.
5. Add files in the selected directory