

Terminology Loader and Remover

Overview

This page documents the use of admin tools for loading and removing snapshots of terminologies.

Prerequisites

- MySQL database must already exist (e.g. "mappingservicedb").
- MySQL database connection parameters must be defined in the properties file specified by "run.config".
- Create/Update DB mojo must also have already run so that the database schema exists.
- If you have already loaded a terminology, remove it before reloading it again with the same version.
- **IMPORTANT:** The Tomcat server must not be running while this is run, otherwise indexes can get corrupted.

Details

There are two RF2 terminology loaders (Snapshot and Delta), a ClaML terminology loader, and a general terminology remover.

The RF2 snapshot loader will load a collection of RF2 Snapshot files (e.g. from a SNOMEDCT International release or an NRC release). This is used to either load an initial version of SNOMEDCT or to load the next published version after a version update. The default version value is "latest", which in theory means the snapshot loader only needs to be run once, and "delta" files can be processed from that point on to update content to later "publication ready" or "published" states.

Mojo: TerminologyRf2SnapshotLoaderMojo.java (in admin/mojo/src/main/resources/java/org/ihtsdo/otf/mapping/mojo)

Project: admin/loader

Configuration Parameters

- A profile indicating RF2-snapshot must be specified
 - e.g. -PRF2-snapshot
 - See profiles in the pom.xml
- run.config - the standard configuration file specified as a -D parameter
- input.dir - the input directory for the RF2 data
- terminology - the terminology value to use when loading
- version - the (optional) terminology version, the value "latest" is used by default.
- loader.defaultPreferredNames.typeId - specified in the properties file specified by run.config
 - This is the RF2 preferred description typeId to use for computing concept preferred names.
- loader.defaultPreferredNames.refSetId - specified in the properties file specified by run.config
 - This is the RF2 preferred description language refSetId to use for computing concept preferred names
- loader.defaultPreferredNames.acceptabilityId - specified in the properties file specified by run.config
 - This is the RF2 language refset acceptability typeId to use for computing concept preferred names

The RF2 delta loader is used to process a daily build of SNOMEDCT, or a definitive RF2 delta for a following release. It only pays attention to concept, relationship, description, and language refset entries as all other data types are irrelevant for the mapping tool. When running this, the last publication date must be explicitly provided in case the base terminology was loaded with a version of "latest". This parameter allows the delta loader to know against which "effective time" values a data element should be compared when determining if it has changed since the last processing of a delta.

Remember, when processing daily builds, each day's "delta" is actually a delta relative to the previous release of SNOMEDCT and not relative to the previous day's daily build. Thus, many/most of the entries in a processed RF2 delta have already been processed (from the prior run). For more information on the "drip feed" process see [Drip Feed](#).

Mojo: TerminologyRf2DeltaLoaderMojo.java (in admin/mojo/src/main/resources/java/org/ihtsdo/otf/mapping/mojo)

Project: admin/loader

Configuration Parameters

- A profile indicating RF2-delta must be specified
 - e.g. -PRF2-delta
 - See profiles in the pom.xml
- run.config - the standard configuration file specified as a -D parameter
- input.dir - the input directory for the RF2 data
- terminology - the terminology value to use when loading
- lastPublicationDate - the last published terminology version from which this delta is relative..
- loader.defaultPreferredNames.typeId - specified in the properties file specified by run.config
 - This is the RF2 preferred description typeId to use for computing concept preferred names.
- loader.defaultPreferredNames.refSetId - specified in the properties file specified by run.config
 - This is the RF2 preferred description language refSetId to use for computing concept preferred names

- loader.defaultPreferredNames.acceptabilityId - specified in the properties file specified by run.config
 - This is the RF2 language refset acceptability typed to use for computing concept preferred names

The ClaML loader is used to load ClaML-based terminologies into the mapping tool. This includes ICD10 which is natively in ClaML as well as other terminologies like ICD9CM and ICPC that have been converted to ClaML.

Mojo: TerminologyClamLoaderMojo.java (in admin/mojo/src/main/resources/java/org/ihtsdo/otf/mapping/mojo)

Project: admin/loader

Configuration Parameters

- A profile indicating ClaML must be specified
 - e.g. -PClaML
 - See profiles in the pom.xml
- run.config - the standard configuration file specified as a -D parameter
- input.file - the input ClaML file
- terminology - the ClaML-based terminology being loaded (e.g. ICD10)
- version - the (optional) terminology version, the value "latest" is used by default

The terminology remover is used to remove any terminology (and version) from the database, regardless of what loader was used to load it.

Mojo: TerminologyRemoverMojo.java (in admin/mojo/src/main/resources/java/org/ihtsdo/otf/mapping/mojo)

Project: admin/remover

Configuration Parameters

- A profile of "Terminology" must be passed
- run.config - the standard configuration file specified as a -D parameter
- terminology - the terminology to remove (e.g. ICPC)
- version - the terminology version to remove, the default is "latest"

The map members remover can be used to delete simple or complex map refset members from the database (by refset id). This is useful for dropping and reloading a particular set of maps (e.g. to set the initial state correctly).

Mojo: TerminologyMapRemoverMojo.java (in admin/mojo/src/main/resources/java/org/ihtsdo/otf/mapping/mojo)

Project: admin/remover

Configuration Parameters

- A profile of "MapMembers" must be passed
- run.config - the standard configuration file specified as a -D parameter
- refset.id - the refset id of the complex or simple maps to remove

Sample

Sample command line call of the admin tool to load SNOMEDCT using the version "latest":

```
% cd ~/code/admin/loader
% mvn install -PRF2-snapshot -Drun.config=/home/ihtsdo/config/config.properties -Dterminology=SNOMEDCT -
Dinput.dir=/home/ihtsdo/data/snomedct-20140731-snapshot
```

Sample command line call of the admin tool to remove ICPC:

```
% cd ~/code/admin/remover
% mvn install -P Terminology -Drun.config=/home/ihtsdo/config/config.properties -Dterminology=ICPC -
Dversion=2010
```

Sample Eclipse run configuration for loading SNOMED CT:

Name: loader - RF2 (SNOMED)

[Main](#)
[JRE](#)
[Refresh](#)
[Source](#)
[Environment](#)
[Common](#)

Base directory:

C:/Users/Brian Carlsen/workspace/OTF-Mapping-Service/admin/loader

[Browse Workspace...](#)
[Browse File System...](#)
[Variables...](#)

Goals: clean install [Select...](#)

Profiles: RF2-snapshot

☐ Offline
 ☐ Update Snapshots
 ☐ Debug Output
 ☐ Skip Tests
 ☐ Non-recursive
 ☐ Resolve Workspace artifacts

1 Threads

Parameter Name	Value
run.config	c:/mapping/config/config.properties
terminology	SNOMEDCT
input.dir	c:/mapping/data/snomedct-20140731-mini

[Add...](#)
[Edit...](#)
[Remove](#)

Maven Runtime: Embedded (3.1.1/1.5.0.20131102-1211) [Configure...](#)

Sample Eclipse run configuration for removing ICPC:

Name: remover - ICPC

[Main](#)
[JRE](#)
[Refresh](#)
[Source](#)
[Environment](#)
[Common](#)

Base directory:
 C:/Users/Brian Carlsen/workspace/OTF-Mapping-Service/admin/remover
[Browse Workspace...](#)
[Browse File System...](#)
[Variables...](#)

Goals: clean install [Select...](#)

Profiles: Terminology

☐ Offline
 ☐ Update Snapshots
☐ Debug Output
 ☐ Skip Tests
 ☐ Non-recursive
☐ Resolve Workspace artifacts
 1 Threads

Parameter Name	Value
run.config	c:/mapping/config/config.prope...
terminology	ICPC
version	2010

[Add...](#)
[Edit...](#)
[Remove](#)

Maven Runtime: Embedded (3.1.1/1.5.0.20131102-1211) [Configure...](#)

[Apply](#)
[Revert](#)

Troubleshooting

Occasionally the load of a terminology will fail for one reason or another. If the process does not run to completion, it is a recommend practice to remove the terminology before attempting to import it again.

```
# stop the server
% service tomcat7 stop

# attempt to load ICPC
% cd ~/code/admin/loader
% mvn install -PClaML -Drun.config=/home/ihtsdo/config/config.properties -Dterminology=ICPC -Dversion=2010 -
Dinput.file=/home/ihtsdo/data/icpc-2.xml

... perhaps a failure occurs here during the process

# Run the remover
% cd ~/code/admin/remover
% mvn install -PTerminology -Drun.config=/home/ihtsdoconfig/config.properties -Dterminology=ICPC -
Dversion=2010

# attempt to load ICPC again
% cd ~/code/admin/loader
% mvn install -PClaML -Drun.config=/home/ihtsdo/config/config.properties -Dterminology=ICPC -Dversion=2010 -
Dinput.file=/home/ihtsdo/data/icpc-2.xml
```

References/Links

- n/a