

# Archive - BiomedGT Publishing to LexBIG with Protege

## Archive - BiomedGT Publishing to LexBIG with Protégé

### Contents of this Page

- [About this page](#)
- [Process Overview](#)
  - [Run a Prompt baseline comparison \(workflow manager\)](#)
  - [Operations tasks](#)
    - [Eliminate unpublishable properties and roles](#)
    - [Generate Ontylog formatted file and flat text file](#)
    - [Load OWL file into LexBIG with cumulative history; compare production to QA](#)
    - [Load Ontylog version into DTS](#)
    - [Tag QA version as Production; publish vocabulary](#)

### About this page

This page explains the process and the procedural steps for publishing BiomedGT content to LexBIG.

### Process Overview

The following sections explain the steps for each procedure in the publishing process.

#### Run a Prompt baseline comparison (workflow manager)

Using the Prompt plug-in, a workflow manager follows these steps:

1. Runs a bi-weekly comparison of the current Protégé database against the baseline from the last comparison.
2. Exports a copy of the master Protégé project file and uses it as a baseline during the next comparison cycle.

#### Operations tasks

Approximately once a month, Operations publishes the baseline file to LexBIG. This section describes each task.

##### Eliminate unpublishable properties and roles



##### Note

Currently the OWL version of the TDE baseline is loaded into Protégé and classified solely for QA purposes. This is to verify that the removal of roles did not create any classification issues. Since we are going to [EVS:classify the Ontylog output](#), this may no longer be necessary.

To eliminate unpublishable properties and roles, Operations runs a Prompt comparison. This procedure accomplishes the following:

- Compares the current and the previous baseline against the edit history as recorded in the evs\_history table, revealing such editing errors as
  - concepts created and retired in the same publishing cycle;
  - concepts that have no history; and
  - history records that have no matching concepts.
- Cleans up the history and exports it to the concept\_history table for editing. Cleanup tasks include
  - combining multiple modifies on a concept into a single modify record; and
  - eliminating modify records on concepts that have been created, merged, split, or retired.
- Exports a number of history files from the concept\_history table for publication, including
  - a cumulative history for publication in LexBIG;
  - a monthly history showing all publishable history records for the cycle; and
  - a monthly history highlighting only creations and retirements (including splits and merges) for use by the caDSR.

##### Generate Ontylog formatted file and flat text file

Operations generates an Ontylog-formatted file for input into the DTS and for download. The baseline processing also generates a flat file for download.



##### Note

This process supports caCORE 3.2. Once caCORE 3.2 is retired, the process will be discontinued.

##### Load OWL file into LexBIG with cumulative history; compare production to QA

1. Load the OWL file into LexBIG on the Dev server, along with the cumulative history up to the date the baseline file was created.
2. Tag baseline as **QA**.

This resides side by side with the previous PRODUCTION version.

1. Run a series of scripts to compare the data in the PRODUCTION version to the data in QA.

These scripts will most likely be run against the LexBIG API, since that is the form in which the user will see the data.

### **Load Ontylog version into DTS**

Operations loads the Ontylog version of the vocabulary into DTS. The method of loading requires that the data be loaded into TDE, classified, then transferred into DTS using Apelon-created applications.

The classification step can be considered a QA marker, as some bad transformations can cause classification to fail. Bad formatting can also result in a failure to load to TDE or to transfer to the DTS.

### **Tag QA version as Production; publish vocabulary**

Once the tests return satisfactory results, Operations follows these steps:

1. Removes the old PRODUCTION version.
2. Tags the QA version is tagged as PRODUCTION.

The vocabulary enters the promotion schedule and is given an expected publication date in the EVS schedule.

1. Loads the final output files, both data and history, into the FTP server in preparation for final publication.
2. Publishes the history files as soon as possible so that EVS-reliant data can be verified for use in the caDSR.

#### **Notes:**

The data on Dev will be transferred up the tiers along parallel tracks. We will push the DTS data to QA and make it available on the nciterms-qa website. We will push the LexBIG data to the software QA server and make it available on the bioportal-dataqa website.

After the data has been on QA for one to two weeks, we will send deployment requests to move both DTS and LexBIG up to Stage. After a day there, it can be moved to production.