

LexEVS Indexer Refactor

Document Information

Author: Craig Stancl, Scott Bauer, Cory Endle
Email: Stancl.craig@mayo.edu, bauer.scott@mayo.edu, endle.cory@mayo.edu
Team: LexEVS
Contract: S13-500 MOD4
Client: NCI CBIIT
National Institutes of Health
US Department of Health and Human Services

Table of Contents

- [Indexer Refactoring](#)

Indexer Refactoring

The indexer project has a fair amount of dead or legacy code. Some of this code is obviously on the code path and some is called by live classes but some appears to be called in dead end or unused methods. Some time will need to be taken to refactor any live code to lexevs-dao while evaluating and disposing of the questionable code as necessary.

Indexer Packages

```
//Refactor and move down into the lexevs-dao
edu.mayo.informatics.indexer.lucene.analyzers
edu.mayo.informatics.indexer.lucene.filters

//Bitset using classes may well go away as will other classes that
//provide support for Single Index/Boundary docs implementations
edu.mayo.informatics.indexer.lucene.hitcollector

//Appears to load xml based metadata -- move to lgConverter or lexevs-dao
edu.mayo.informatics.indexer.lucene.xmlParser
edu.mayo.informatics.indexer.lucene.xmlParser.exceptions

//Supporting classes for the metadata.xml file and a variety of other things that can mostly go away
edu.mayo.informatics.indexer.utility

//Supporting code for the older version of the multi-index search.
//Will review but largely re-implement in lexevs-dao
//These classes are called by the classes in the dao project, but many method calls may not be called
//or otherwise affect API level calls.
edu.mayo.informatics.indexer.api

//This package looks to have been replaced to some extent by LuceneLoaderCodeIndexer in the lexevs-dao project
//where it builds documents from entities.
//but it has quite a lot of calling code so we'll have to make sure all the paths are dead ends before
eliminating it.
edu.mayo.informatics.indexer.api.generators

//Quite a bit of the methods and classes have a presence in lexevs-dao
//Many seem not be be called by upper level API classes and appear to be dead execution paths
//These will take some time to evaluate, refactor and test.
edu.mayo.informatics.indexer.lucene

//Used by the WhiteSpace analyzer. We will move this into the lexevs-dao
edu.mayo.informatics.indexer.lucene.tokenizers

//This appears to be a regex query wrapper. We'll want to move it to lexevs-dao
edu.mayo.informatics.indexer.lucene.query

//Exceptions package. We'll consider simplifying or removing it in favor of rethrowing runtime exceptions.
edu.mayo.informatics.indexer.api.exceptions
```