

Integration Services

The [fcrepo-camel-toolbox](#) project includes a number of production-ready services that can be used to integrate Fedora with external systems, such as Solr or a Triplestore.

All of these services can be deployed in a web container such as Tomcat or Jetty. Alternately, they can be deployed in an OSGi container such as Karaf. Deployment and configuration instructions are available in the associated README files.

Solr Indexing

The [solr indexer](#) uses the [LDPath service](#) to convert RDF documents to JSON. A default transformation program can be specified in the [service configuration](#) (e.g. default or myTransformation). It is also possible to override the default transformation program by assigning an RDF property to particular documents: `<> indexing:hasIndexingTransformation "specialTransform"`. Furthermore, one can choose to index only certain documents from the repository. By identifying certain documents as `<> a indexing:Indexable` and enabling the `indexable.predicate` configuration value, only those resources will be indexed. (For Tomcat/Jetty-deployed applications, this can be enabled by setting `JAVA_OPTS="-Dfcrepo.onlyIndexableObjects=true"`)

Triplestore Indexing

The [triplestore indexing service](#) runs just like the Solr Indexing service, pushing all changes from the repository into an external triplestore. Both Fuseki and Sesame have been used successfully with this service. Like with the Solr Indexing service, it is possible to identify certain objects as "Indexable" by setting an `rdf:type` as `indexing:Indexable`. (One must also enable this filtering, as described above).

Reindexing Service

Periodically, it may be necessary to reindex some or all of a repository. In certain cases, one may wish to re-index only Solr, only the Triplestore, or both. The reindexing service exposes a RESTful endpoint where it is possible to initiate these sorts of reindexing processes. By default, the reindexing service exposes an HTTP endpoint at `localhost:9080/reindexing` (to change this, see the [documentation](#)). That endpoint accepts JSON documents like so:

```
curl -XPOST localhost:9080/reindexing/objects -H"Content-Type: application/json" \
-d '["activemq:queue:solr.reindex","activemq:queue:triplestore.reindex"]'
```

This will reindex both Solr and the external triplestore, starting at the `/objects` node in Fedora. To start at the root node in Fedora, you would POST to `localhost:9080/reindexing/`, while to start at the node `/a/b/c/d`, you would POST to `localhost:9080/reindexing/a/b/c/d`

The values in the JSON array are used to determine which endpoints to reindex.

By sending a GET request to the reindexing service, you will retrieve a short summary of its usage.

Serialization Service

The serialization service allows objects in Fedora to be written to a specified disk location in any MIME type Fedora supports. One can decide whether or not to include binaries in the serialization, but keep in mind that they may take up a considerable amount of disk space.

When this service is enabled, it will write any new or changed object's description to disk. It listens to Fedora and only operates on objects that have been created, modified or deleted. If an object was created before this service was enabled and is unchanged, it will not be written to disk. This service exposes an `activemq` queue for initiating serialization of Fedora objects that might not otherwise get written to disk.

If an object is deleted from Fedora, the serialized description and/or binary will be removed as well.

For more details see: <https://github.com/fcrepo4-exts/fcrepo-camel-toolbox/tree/master/fcrepo-serialization>