Purpose

As per feature request JIRA RUC-237[1]; at ingest items should be added to portal with minimal delay. Currently after ingest there is a wait period of 4-6 hours before a newly ingested resource is added to search portal. Until then the item is not discoverable nor are any non-embargo file restrictions applied that are based on collection membership.

Proposed Solution

During the ingest process in WMS after which the resource is ingested and indexed by Solr the WMS will add a new parameter when calling the Solr indexing service. That parameter will tell the Solr indexing service that this is a newly ingested resource. The Solr indexing service, after initially indexing the resource, calls the Portal Tool web service that will then add the resource to any "high-level" portals with which the resource may be associated.

Portal associations usually are defined by the collection the resource is a member of, but associations may also be defined by metadata criteria.

High-level portals are portals in which immediate inclusion is desired. Not all portals are given this designation, since the portal association process could become very process intensive and lengthy. All portal associations that are not part of the immediate high-level analysis will still be created during the currently established indexing period that occurs 4 times a day.

The portal tool will provide an interface for managing portals that are designated as high-level. That interface will be available to all portal tool administrators. Typically, high-level portals should include the main repository portal, SOAR, RUresearch, ETD, NJDH, NJEDL.

As part of the post-ingest process in the WMS the high-level portal association will be applied when calling the Solr indexing service with an added parameter telling the Solr indexing service that the request is for a newly ingested resource. The Solr indexing service will make a web service call to the portal tool. A new method will be added to the portal tools web service framework that will add any high-level portal associations to the resource. During that process, any non-embargo file restrictions can be applied as well, also reducing the wait period for enforcement. The non-embargo file restriction will occur regardless of whether the resource in question is part of a high-level portal or not.

Adding this association step with the Solr index service will provide greater flexibility and broader availability to all applications that ingest in the repository. The analytic tool can also use this same feature, with very little effect on the Analytic tools codebase.

Testing Scenarios

1. Ingest a resource that belongs to an NJDH collection. Browse the NJDH site to find the resource in its NJDH collection.
2. Ingest a resource that belongs to a Rutgers University or Rutgers University Libraries collection. Browse the RUcore site to find the resource in its RU/RUL collection.
3. Ingest a resource that belongs to a SOAR collection. Browse SOAR to find the resource in its SOAR collection.
4. Ingest a resource that is in a restricted collection (e.g., limited to a location or Rutgers NetID). Browse the high level portal to find the resource in its collection, and confirm that the restriction is in place.
5. Ingest a resource that is not intended to be part of a high-level portal. After ingest, confirm that the resource is not discoverable until after the regularly scheduled indexing period.

References

[1] https://jira.ess.rutgers.edu/browse/RUC-237