**# This is an attempt to synthesize Lindsay's draft for a processing manifest with Scott's suggestions**

**# (plus there are additional proposed revisions)**

**Manifest**:

 **manifestID**: # A unique ID for the manifest. Can be a sequential number.

 **type**: processing

 **namespace**: WE1S

 **version**: 1.0

 **resourceId**: 1

 **title**: New York Times topic model

**sourceInfo**:

 manifestID: # Refers to a corpus manifest, which will include info such as resourceLocation.

 corpusSections: # Refers to 1 or more sections of corpus (e.g., NY Times, "humanities", raw text, 2010, 2011)

 description: # Describes the source

**processingInfo**:

 ~~processingID: # An ID for the processing sequence that can be referenced by a resource or other processingInfo element.~~

 **processingRef**: # Optional reference to the processingID of another processingInfo element.

 **processingDate**: # Date when processing steps were performed.

 **processingEditor**: # Individual(s) who performed the processing step. Can contain multiple values.

 **processingWorkstation**: # Workstation or computer where processing was done.

 **package**: # Calls manifestID of package of sequenced tools, scripts, and manual steps.

 manifestID: # ID of the manifest for a package.

 **processingSequence**: # If no package, this begins a sequence of processing steps involving tools, scripts, and manual tasks.

 **seq**: 1

 **mechanisms**:

 **type**: tool

 **label**: Mallet

 **version**: 2.0.7

 **resourceLocation**: # If needed.

 **description**: # If needed

 - **option**:

 **argument**: num-topics

 **value:** 20

 - **option**:

 **argument**: remove-stopwords

 **value:** True

 - **option**:

 **argument**: stoplist-file

 **value:**  # path to stoplist.txt

 - **option**:

 **argument**: optimize-interval

 **value:** 20

 **outputInfo**:

 manifestID: # Refers to a corpus manifest, which will include info such as resourceLocation.

 corpusSections: # Refers to 1 or more sections of corpus (e.g., NY Times, "humanities", topic model, 2010, 2011)

 description: # Describes the output.