AD15 Bulk Importer (Admin only)

I. Purpose:

The Bulk Importer enables us to import Document concepts and Requirement concepts and their respective associations (relationships) in bulk. The Bulk Importer is a specially configured Excel spreadsheet comprised of 3 sheets: Requirement, Document, Links. Portions of these are illustrated in the Figures below. Instructions on obtaining an updated template, entering data, and then importing are provided below.

The Bulk Importer is intended for entering a lot of data all at once – it is NOT intended for circumventing the normal data entry process (via Cases).

II. Who Can Do This:

This is for Administrators only. I think there are no safeguards to prevent others from doing this, if they could get hold of the links and know the procedure.

III. Warning

Using the Importer to modify previously imported data will OVERWRITE the data – including eliminating any changes that may have been made between the original import and the proposed one. ***STRONG RECOMMENDATION: Use the Importer to enter NEW data only.***

IV. Caveats on Use

The bulk importer was intended to be used for initial population of the database, and then retired from use. It is not user-friendly. It does not always issue error messages, and the ones it does issue are not always understandable. The agreement with Ovitas was that there would be very little effort in the development and support of this utility. It is not worth additional money or time making the tool any better than it is. ***Advice: use at your own risk and be ready to exercise great patience.***

V. Procedure Details:

The following describes

- How to find the template
- How to populate the Excel template
- How to import
- Activating deliverable docs
- Troubleshooting hints.

Figures below show portions of the Excel Bulk Importer Template:

Requirement Sheet

REQUIREMENTS					
Import Type	Concept VsId	Instruction	Is this a modification or is it new?*	Contract Part*	Contract Section*
Update	20120627_req_C31_I.021		new	Contract Section	I
Update	20120627_req_FedApproprLaw		new	Other	n/a
New	20120627_deliv_924167		modification	Contract Clause - DEAR	I
New	20120627_deliv_924191		modification	Contract Appendix	J.09 ESH - Appendix I

Document Sheet:

DOCUMENTS				
Import Type	Concept VsId	Instruction	Document Number (xx.yy.zzz.aaa)	Document Title
New	20120627_doc_LETSUserGuide		11.05.001.001	LETS User Guide
New	20120627_pol_TimeLabor		11.05.001.000	Time and Labor Reporting

Link Sheet:

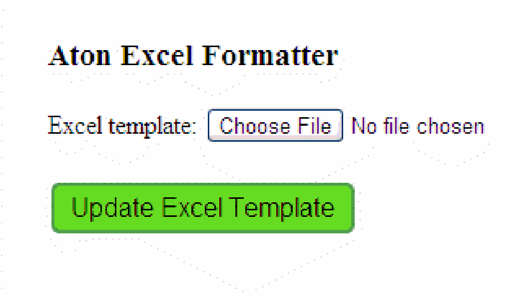
Link Category	Parent Role	Parent Concept VsId	Child Role	Child Concept VsId
Requirement to Supporting Document (Supporting Docs)	Requirement	20120627_req_C31_I.021	Document	20120627_pol_TimeLabor
Requirement to Supporting Document (Supporting Docs)	Requirement	20120627_req_FedApproprLaw	Document	20120627_pol_TimeLabor
Document to Document (Related Documents)	Document	20120627_pol_TimeLabor	Document	20120627_doc_LETSUserGuide

A. Template Preparation

1. Obtain a blank template at:
   http://rms.xx.lbl.gov:8080/lbl-aton-server/spring/mvc/importer/template
   and "xx." = *dev.* or *qa.* or is blank for production.
2. Templates must match the model.

B. Verifying that a Template conforms to the model:

1. NOTE: do this only if the model has been updated, or if you are not about the template.
2. Go to https://rms.xx.lbl.gov:8080/lbl-aton-server/spring/mvc/importer/update and "xx" = *dev* or *qa* or is blank for production.
3. Upload the Excel template
4. Update the template
5. The result is a new template that can be saved locally.



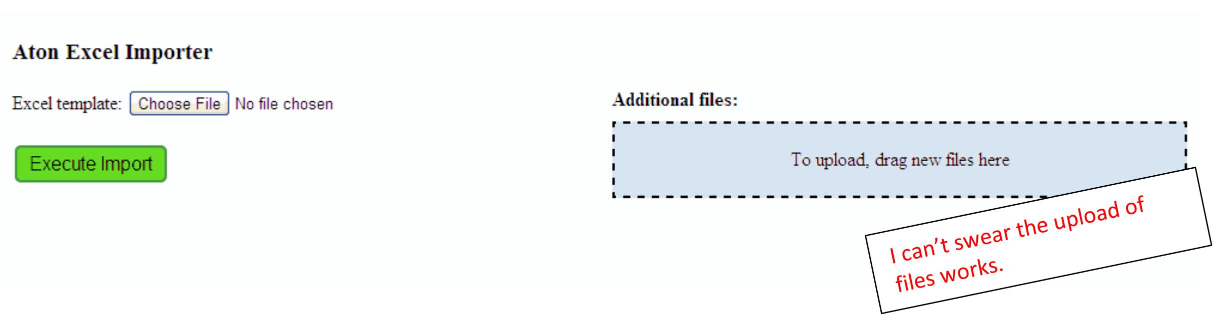
C. Entering Data

1. Best to prepare and manipulate data within a separate Excel sheet.
2. The user is expected to be intimately familiar with all the definitions of the Documents and the Requirements metadata fields. In particular, the user must know the parent-child relationships between the requirement, embedded requirements, deliverables, and so forth.
3. The user is expected to know the naming conventions for the picklists as well as the citations (main, sections, etc.)
4. In the Template, "red font" and the red asterisk next to the field title means a value must be non-blank.
5. To preserve Importer Template (the picklists!!), MUST USE "Paste Values" from Excel menu. Do NOT use quick keys (Ctl-V).
6. Persons' names <u>must be</u> converted to LDAP ID numbers – this is one reason for building a separate Excel sheet.
7. "Pipe" format is used for parameters that have cardinality >=1. For example, "aaa|bbb|ccc" under "additional owners" would mean Owner "aaa", Owner "bbb" and Owner "ccc".
    a. Unfortunately, when pipe separators are applicable, there is no picklist. This makes listing of detailed Policy Areas particularly agonizing to get right. Go to the Workflow Design Specification for the picklists.
8. Ignore any columns marked "Instructions" – the macro pulls in all fields in the App, including the ones that don't need any data.
9. <u>Column 1: Import Type (new/update):</u> This refers to whether or not the imported data is new or a modification. (Modifications may be uploaded in bulk! BUT this is not advised except as noted below.)
10. <u>Column 2: Concept VS ID:</u>
    a. This is a **unique** user-created label which (1) enables the system to identify a previously bulk-imported concept so that it can be modified again by the bulk import process, and (2) enables the user to more quickly establish links between concepts. Item (2) drives definition of a label that can help the user quickly identify the concept – that is, the label should relay the context of the concept.
        - If the Concept VS ID is <u>not</u> unique (i.e. there's a previously uploaded value), the importing process will <u>not</u> continue unless Column 1 says "update."
    b. The following definitions yield a unique label for Concept VS ID. In these, "yyyymmdd" refers to the original import date and "#" is a sequential number, where the number of digits is reflected by the number of "#" symbols. For Requirements and Documents, the "###" refers to a specific portion (section, chapter, page, etc.) of the requirement or document.
        - Deliverable: yyyymmdd_deliv_(LabID#) [The "LabID#" is the same identifier used in BSO database – "924xxx"
        - Requirement: yyyymmdd_req_(Brief Title or citation)_###
        - Policy document: yyyymmdd_pol_(Brief Title)
        - Document: yyyymmdd_doc_(Brief Title or citation)_###
    c. Make sure there are NO spaces in the assigned Concept VS ID.
    d. Examples:

- 20120315_pol_travel
- 20121104_req_O_4141D
- 20120131_req_10CFR851.3
- 20140205_prog_PPE
- 20140205_doc_PPE_WFA
- 20130825_form_reimbursement
- 20110921_deliv_924123

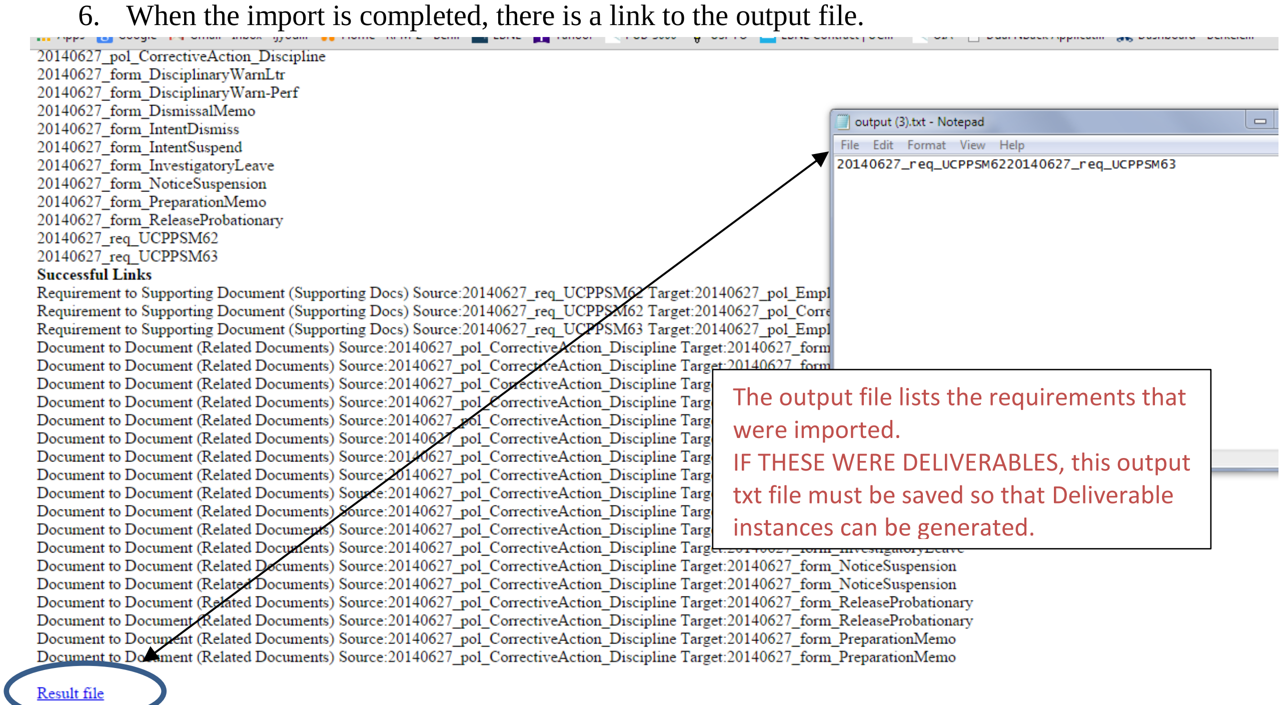
11. Setting up Links Sheet (Associations):
    a. This is quite tedious, and the User is advised to have a systematic way to set up and then check the many associations. A document tree is a helpful graphical method.
    b. The Concept VS IDs (Step 11 above) must be assigned to both Reqs and Docs before associations can be established.
    c. In the Link sheet left column, select from the provided picklist the relationship to be established. This selection automatically populates the “parent role” and the “child role” columns AND also sets up the picklists of available Concept VS ID values for the third and fifth columns.
    d. The picklist order is the same order that is in the doc or req sheet. A systematic approach (for example) would be to put the docs or reqs in alphabetical or numerical order.
    e. Select the specific parent and child items (3^{rd} and 5^{th} columns).

D. Importing Data
1. Access the Bulk Importer at:
    i. http://rms.xx.lbl.gov:8080/lbl-aton-server/spring/mvc/importer/import
    ii. and “xx” = *dev* or *qa* or is blank for production.
2. Upload the Excel template
3. Drag any documents referenced by the URL properties (for example: Link to Document) in the Excel template in the “additional files” box. [Note: I have never used this feature, and I don’t know if it works or not.]
4. Select the Execute Import button to start.



5. As the system is processing, it will provide real-time progress such as:

6. When the import is completed, there is a link to the output file.



7. The detailed information is "temporary" and can be copied and pasted into Word or other file for studying. In the example below, for the failed RM Governance I had mis-read "Document Approver approval" and inserted the Doc Approver's ID number instead of "approved". Note that the error message is reasonably suggestive. The root hang-up for the second error is not clear; but re-typing the date cleared the failure.

Import Complete

Successful Concepts
20120621_req_10CRFR830
20120621_req_H.30_ContractorAssurance
20120621_req_I.76_ManagementControls
20120621_req_O414.1D_QualityAssurance_part
20120621_req_O414.1D_QualityAssurance_full
Failed Concepts

20120621_doc_RMGovernance	The Enumeration value "496689" for Property "Document Approver approval" is not a valid Enumeration value
20120621_pol_ContractorAssurance	Could not format value "1/2/2012" as a Date

Here's another example of a failure message: Can't tell what this means at all.

Failed Concepts
20140627_req_UCPPSM62 Could not retrieve the Property for the property version set ID of "versionSet618733e5-4962-11e3-857c-2b16d70d5b20"
20140627_req_UCPPSM63 Could not retrieve the Property for the property version set ID of "versionSet618733e5-4962-11e3-857c-2b16d70d5b20"
Successful Links
Document to Document (Related Documents) Source:20140627_pol_CorrectiveAction_Discipline Target:20140627_form_DisciplinaryWarnLtr

8. The Production RMS has a lot of data (as of November, 2014). If you are new or rusty at using the Importer, try first on the QA RMS. This will minimize the number of tries on the Prod tool.

E. Modifying data via Bulk Importer:

WARNING: *The Importer will overwrite the information of whatever item in the system has the same VD ID value in the Importer sheet. Do not use to update info that has resided in the system for a while. (See Section III above).*

The main use of the new vs update switch in the Importer is while troubleshooting a particular pile of data. Often a portion of an import will succeed and some will fail. For failed items, the user should try to correct them and then re-run the sheet. If there are links to be made, the items that succeeded the first (or previous) time around need to remain. However, since they succeeded, they are now considered uploaded. The way to get everything to succeed is to switch to "update" for any line item that succeeds.

If you choose to update data previously uploaded via the Bulk Importer, here's the process.

1. Find the line item that is to be modified – its Concept VD ID is required!
2. Change first column to "update".
3. Make all the necessary mods.
4. Then run the Importer per usual.

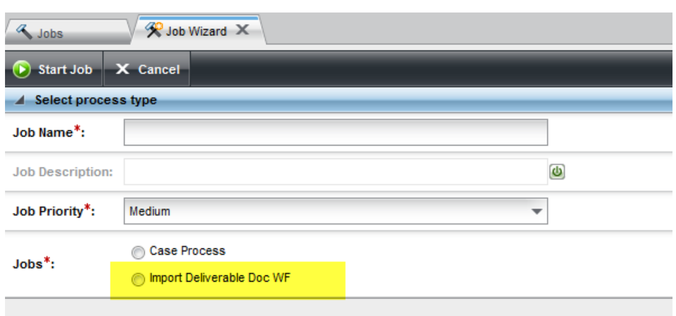
5. If the line item(s) of interest is one of a bunch of other items that are not changing:
   a. Be sure to change the first column to “update” for all those other items.
   b. Note that the other items will still be imported, and their version numbers incremented.
   c. There is huge risk if you don’t know whether or not the “other stuff” has been changed between its original upload and this import. (If running a series of imports in a row to troubleshoot and get the whole pile in, then obviously the “other” (successful items) have no changes.) Better is to remove anything that you are not interested in changing. Best is to NOT use the importer to fix just a couple of items – go through the proper way: Case and Req workflows.

The Production RMS has a lot of data (as of November, 2014). If you are new or rusty at using the Importer, try first on the QA RMS. This will minimize the number of troubleshooting tries on the Prod tool.

As of this writing, I have not tested modifying a requirement that has associations (that is, if a requirement has associations, and it is modified by the Importer that does not include info in the Link sheet, will those associations be impacted?). The reader is free to experiment at his/her own risk – do not practice on Prod RMS. [My humble opinion: it is better to modify the requirement via the usual method through Case.]

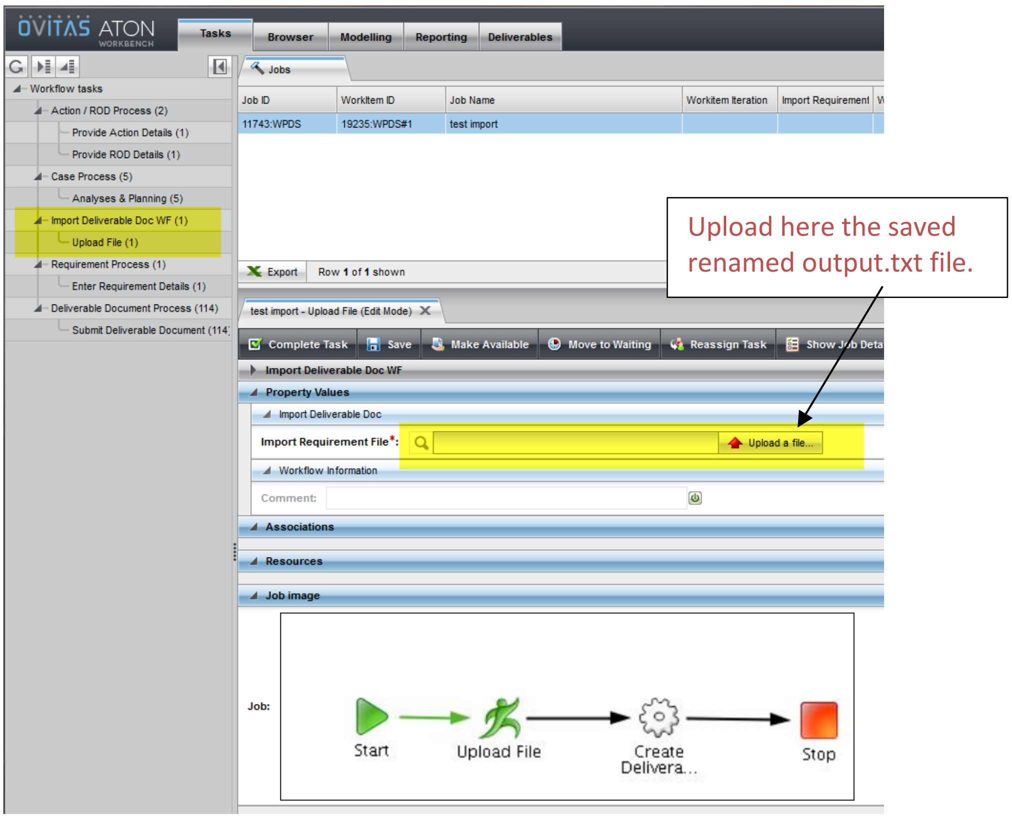
F. Activating (Generating) Deliverable Docs

1. If the Imported data includes DeliverableDef entries, then these will have to be “activated” within the application to generate DelivDoc instances.
2. When the Excel import is finished, it creates an output.txt file (see Step D.6 image above), and places it in either temporary folder or on the desk top. The file is a string of all the requirements concepts that were created successfully in the particular import process.
3. Save the output.txt file on your hard-drive. Should rename it.
4. Log into the Aton Workbench as a user with “Admin” privileges.
5. Go to the Job Wizard (Inbox column) and start a job of type “Import Deliverable Doc WF”. The Job Name should be reasonably descriptive.



6. This workflow has one activity. Find the “Import Deliverable Doc WF” jobs in the “Workflow Tasks” tree and select a Workflow to work on. (See figure below)
7. There is one property that needs to be uploaded (Import Requirement File). This is to upload the corresponding output.txt file that you saved (step 3 above).

8. Select the "Complete Task" button. Depending on how many requirements there are, it could take a while for all of the deliverable doc jobs to get created. You don't have to wait, however, as they are created in the background.



G. Archiving Imported Template Data (Recommended):

1. Since it is possible to modify the database using the Bulk Importer, it is important to archive any Importer spreadsheets.
   a. Label these files with filename: "yyyymmdd (Brief descriptor) AtonBulkImportTemplate.xlsx", where in this case the date refers to the import date.
2. Archive the various Importer spreadsheets for possible later use.
3. Archive the output.txt files – recommend a filename that matches the spreadsheet (distinguished by the suffix xlsx vs txt)

H. Troubleshooting Tips

1. Remember Section IV above. Good luck. You'll have to be patient and clever.
2. Some common errors in preparing the spreadsheet (see Section C above):
   - ID value has a space (failure to load the name)
   - "New" instead of "update" for a repeated item (failure because the line item is loaded already)
   - Unrecognized formats or items:

- Remember to "Paste Value" from the Excel Menu – don't use the general "Paste" or shortcut Control+C.
- Mis-typed value instead of using the dropdown value
- "Approved" when answer is "yes"
- Person's name instead of LDAP number
- Date – sometimes you just have re-type (cut/paste doesn't always work)
- Omitted value in mandatory (red asterisk) column
- The most painful: Policy Areas that don't match exactly the picklist wording.
- The second most painful: there's an extra space at the end of "DOE Acquisition Regulation (DEAR)." So if you cut and paste and don't use the dropdown value, you'll get a failure.
- Not using the template for the particular RMS system (Prod vs QA vs Dev). All 3 systems are at the same level (as of Nov, 2014) so it is likely that the same spreadsheet might be applicable to all three.

Revision History

Revision	Date	Who	Description
0	11-15-2014	L.J. Young	Initial