Automated Enterprise Architecture

Patrick Stingley

[Pstingley@blm.gov](mailto:Pstingley@blm.gov)

(202) 815-3550

4/27/2020

These programs take data from IEMS/BigFIx, clean them, sort them, count them and align them with the Federal Enterprise Architecture Framework. A spreadsheet is provided that will take this aligned data and align it with TBM as well. All the materials needed to run this are locate at: <https://github.com/pstingley/Automated-Enterprise-Architecture/tree/master> You can download them from there. If you have problems, please contact me. All of these programs run under Python 2.7, which is why there is a batch file to run them. This allows the user to direct the programs to the correct version of Python.

First, Create a directory [1 TRM/<month> <year> (i.e., January 2018)] for the scripts and data files. This allows the reports and the scripts to be adjusted and improved, without affecting previously working scripts.

Open Internet Explored (It tends to work better with ie) Log into IEMS at the following URL: https://webreportsmo.doi.net/webreports?page=DomainsList

Your network credentials should allow you to enter directly. If they don’t send an e-mail to IEMS Master and ask for access.

I have starred the important reports so they are easier to find. You might want to do this.

I’m not sure if they will work for other Bureaus, so I have provided the field names in a file titled “Just the Apps Ma’am headers.csv”

Second, choose the report titled [Just the Apps Ma'am](https://iem-reports.doi.net/webreports?page=ExploreComputers#filterManager=2d96f86a4f1f4b91b342f9d4ee25735ac049fa11&wr_computerTable=f9d660f8f32eef1ed95027928034574135b98a2e&reportInfo=795e6128eda03aaee79dfc1dc149eac806170ca8) This report (Just the Apps..) provides a list of all of the applications on all of the Operating Systems on the network. It doesn’t include the hostnames of the computers the software is loaded on because previous testing indicated that adding the host name changed the output for some reason. This report is meant to be an inventory of all of the software on all of the computers including UNIX and its derivatives. It should be about 38MB (Double click on the name, which is a hyperlink to the report.) Sometimes it takes a while…

Choose the “Export to .csv” option. When the data is exported to .csv all of the rolled up values represented by “+” signs are expanded, so don’t worry if they are rolled up on the screen display, the data will be fine.

Click on the link in the upper right that says “Export to CSV” It is a hyperlink (although the only visible indication of it is a tiny swirling ring in the tab on top of the window). This will begin the download process. Because of the delay involved, it is advised to look at the tab to see if the swirling icon indicates something’s working. If the process has started it will be swirling.

All of these reports come out with the same name “ExploreComputers.csv”, so it is essential to change the title of the file you’re saving it to, to something different each time so they don’t write over one another.

I have found that renaming the file with the name of the report helps to keep things organized, for example “ExploreComputers – Just the Apps Ma’am.csv”. Appending the report name to the ExploreComputers title, allows the files to lump together when sorted.

Sometimes (due to some unknown change in the mainframe) the report rolls up the fields instead of showing all of the information in each row of each column. To remedy this, go to the column chooser and for the column with the most data choose the minus [-] option. This makes it list each thing individually. After this, it will populate all of the other rows.

Use the back arrow to get back to the menu where you can select reports.

Next, run the report titled: “ExploreComputers - All Software on all computers with IPs and Hostname and who Used them 6-24-2019.csv”. The headers for this report should be as follows:

Computer Name,IP Address,Last Logged on user,User Name,Installed Applications - Windows,Runlevel 5 Services - Solaris,Runlevel 5 Services - Linux,Runlevel 5 Services - AIX,Installed Applications - Versions - Mac OS X

I eliminated the MAC addresses in August because they’ve never been helpful and I replaced them with the IP Address.

This will produce a report as the title describes. It includes Solaris, Linux, Windows, etc., but this one shows the hosts the software is loaded on and who used them in case you need to call someone to ask about a piece of software. For the first year or so this file was usually about 238MB, but after eliminating Run Level 3 from the report, it is now about the same size as Just the Apps (37-38K).

Copy the following files into the current directory:

run.bat

counter.py

DataScrubber(date).py

Match\_With\_FEAF\_(date).py

sortlines.py

Taxonomy(date).txt

All of these programs have been written with simplicity in mind. They are designed such that a person who is not good a Python can maintain them. There are a lot of if statements that replace various things found in the inventory. These could be simplified, but I found that the program actually took longer to run when they were removed, so I left them.

Open a command window and run run.bat. This runs all of the programs in the proper version of Python.

I adjust the dates when I edit the programs, which changes run.bat

Today, Data Scrubber scrubbed 716364 lines and identified 699666 unique products in 1 minutes and 57 seconds. word

DataScrubber will produce two different files with names like those below:

Just use the one titled something like: ExploreComputers - Just the Apps Ma'am.csv with pipes 11 11 21 26-11-2018.csv

Run.bat will now run two programs, sortlines.py and countlines.py that sort and count the data respectively, creating intermediate files in the process which I don’t delete because they can be useful for analysis at times.

Matching to the FEAF

This process matches the cleaned up list of programs to a taxonomy file where each of these programs has been sorted into the FEAF category that best matches it. New programs are constantly coming on to the network, so the taxonomy file is constantly being edited to reflect the new items.

The Taxonomy file is order sensitive. It is essential that some of the entries at the bottom remain there.

I recommend re-naming the taxonomy file when you edit it, to give some form of version control. If you do, you will need to edit the Match\_With\_FEAF\_(date).py file to read the Taxonomy file’s new name. Rename Match\_With\_FEAF to the current date to indicate that you have edited it.

Today Match with FEAF catalogued the data in 35.723 seconds. The time seems to be getting shorter, perhaps because of fewer different products, but I like to think it’s because more things are found.

Open the file named TRM Data Analysis Template 40-20-2020.xlsx and save it with the current date.

Import the data from the file named something like: “FEAF Alignment 19 04 57 25-09-2019.txt” as tab delimited data into the tab named TRM. The quotes are not important.

You will find many files under the categories of virus, policy violations, gaming and etc. This is because there are no categories in the FEAF for these kinds of programs, but a complete inventory of the products on our network will bring them to light.

The file named Not FEAF Aligned (date).txt contains the names and counts of programs found on the network that have not yet been aligned with the FEAF categories. I paste these into the tab named New Products Found. Typically, there are a couple of hundred new items added on the network every month, so this file turns out to be a list of the additions.

I look these items up, determine what FEAF category they go into and add entries for them in a new Taxonomy.txt file.

TBM

In the spreadsheet, there is a tab named TBM Taxonomy. As might be surmised from its name, this contains the TBM taxonomy. While all of the software falls under software in the Cost Pools and by definition is somewhere within the IT Towers, many products on our network have yet to be mapped to the IT Sub Towers. We are not alone. I have been working with the government leads at GSA and with the TBM council and the services these products perform simply have not been aligned yet.

There is a tab titled TRM – TBM Crosswalk that shows the correspondences between the FEAF and the TBM taxonomies.

Macros in the spreadsheet will take the products posted in the TRM, look up the correspondences using the TRM – TBM Crosswalk table and post the aligned products and associated FEAF service categories in the tab titled TBM Enabled TRM.

That’s it! That’s as far as it is possible to go.