IEPR-22-02 Gathering and organizing data, part 2

IEPR-22-02 Gathering and organizing data, part 2

To help in choosing methods for planning, as I spoke of in TN242870 IEPR-22-02 Gathering and organizing data, identifying required data and locating existing data supporting your choice is the next step.

Additional submitted attachment is included below.

IEPR-22-02 Gathering and organizing data, part 2

To help in choosing methods for planning, as I spoke of in <u>TN242870 IEPR-22-02 Gathering and organizing data</u>, identifying required data and locating existing data supporting your choice is the next step.

This can be done by making use of existing data locating tools. Most computer operating systems have data locating tools built-in.

I offer these examples of tools to use that can be run on the Microsoft Windows operating system. You can access these tools by running these batch files to automate the data locating process.

Make a batch file that uses the "dir" command to write the result to a file. Name the batch file "ctlg.bat". Its contents are as follows:

```
echo.^<!DOCTYPE html^>^<html^>^<head^> > "%2 files on %1 drive.htm"
echo.^<title^> >> "%2 files on %1 drive.htm"
echo.%2 files on %1 drive >> "%2 files on %1 drive.htm"
echo.^</title^> >> "%2 files on %1 drive.htm"
echo.^</head^>^<body^> >> "%2 files on %1 drive.htm"
echo.^<pre^> >> "%2 files on %1 drive.htm"
dir %1:\*.%2 /-c /n /on /s /x /4 >> "%2 files on %1 drive.htm"
echo.^<pre^> >> "%2 files on %1 drive.htm"
echo.^</body^>^</html^> >> "%2 files on %1 drive.htm"
```

Make a batch file that uses the "call" command to call "ctlg.bat" and add the replaceable parameters. Name the batch file "ctlg-call.bat". Its contents are as follows with the drive letter (C, D, etc.) you with to list the files of replacing "**Z**":

```
call ctlg.bat Z "CSV"
call ctlg.bat Z "PDF"
call ctlg.bat Z "PPT"
call ctlg.bat Z "PPTX"
call ctlg.bat Z "PY"
call ctlg.bat Z "TAB"
call ctlg.bat Z "TXT"
call ctlg.bat Z "XLS"
call ctlg.bat Z "XLSX"
call ctlg.bat Z "XLSM"
```

Run the "ctlg-call.bat" batch file, results will written to the directory that contains the batch file and be something like this result:

```
<!DOCTYPE html><html><head>
<title>
"XLSM" files on Z drive
```

```
</title>
</head><body>
<pre>
 Volume in drive Z is DATA
 Volume Serial Number is B89F-0FC1

 Directory of Z:\

04/20/2018  05:21 PM          7960738 RESOLV~1.XLS Resolve Fuel
Price Calculations 2017 IEPR Update.xlsm
04/20/2018  05:24 PM         27146758 RESOLV~2.XLS
RESOLVE_Dispatch_Viewer 2017-09-07.xlsm
04/20/2018  05:33 PM           802713 RESOLV~3.XLS
RESOLVE_Results_Viewer 2017-10-24.xlsm
04/20/2018  05:36 PM         42805870 RESOLV~4.XLS
RESOLVE_User_Interface 2018-04-17.xlsm
               4 File(s)       78716079 bytes

     Total Files Listed:
               4 File(s)       78716079 bytes
               0 Dir(s)     89734696960 bytes free
<pre>
</body></html>
```

This is a fast way to put a file list out to the public.

With a little additional processing you can produce web pages with lists and links sorted by data field names like this:

Field Name "Area" is found in:

prod_cost_inputs/42mmt_Ref_20180416_2017_IEPR/conventional_generators_baseline.csv

prod_cost_inputs/42mmt_Ref_20180416_2017_IEPR/generators_optimized.csv

Field Name "BTM PV" is found in:

prod_cost_inputs/42mmt_Ref_20180416_2017_IEPR/annual_load.csv

Field Name "Baseline Consumption" is found in:

prod_cost_inputs/42mmt_Ref_20180416_2017_IEPR/annual_load.csv

Field Name "Capacity" is found in:

prod_cost_inputs/42mmt_Ref_20180416_2017_IEPR/conventional_generators_baseline.csv

prod_cost_inputs/42mmt_Ref_20180416_2017_IEPR/renewable_generators_baseline.csv

prod_cost_inputs/42mmt_Ref_20180416_2017_IEPR/storage_baseline.csv

Field Name "Capacity per Unit MW" is found in:

prod_cost_inputs/42mmt_Ref_20180416_2017_IEPR/generators_optimized.csv

By gathering and organizing data as shown above, I believe CEC will not lack a place to host data and analytical tools in accessible, understandable, easily navigable way.

It may only take a week or two to build a precursor web site for the California Planning Library depending on the time it takes for batch files to process when writing the file list web pages.

It can be as simple as running the batch files from a central location on the network. There may be the need to run the batch files from a removable storage device to collect file lists from computers drives that can't be accessed from the network.

Once all batch files are processed, other existing data management tools will be used to produce the various sorted lists as needed to better understand what data exists, how it was used, and how to make better and faster use of the data for planning.

Perhaps the entire project described in the California Planning Library proceeding can be completed by the end of May 2022 by using these existing methods?

Steve Uhler
sau@wwmpd.com