Utilities

1. WMECO/CLP
2. NSTAR (BECO/COMELEC/CMBRDG)
3. PSNH

Goal

To automate the process of receiving IDR data for Eversource utilities. This will involve:

- Scraping PR-specific username and login info from utility emails
- Creating a GAHR file to send to ESG containing the login info and LDC account numbers
- Sending GAHR file to an FTP
- Receiving and parsing USHI files for each LDC account

Current Process

1. MML is validated, LOA files are processed and sent out to utility through email
2. GAAH requests are sent out for EDI data, USGH is parsed upon return (*2-4 days*)
3. IDR data is uploaded to a utility portal. Login credentials, LDC account information and available timeframe is then emailed from the utility to Engie (*3-7 days*). Example below.
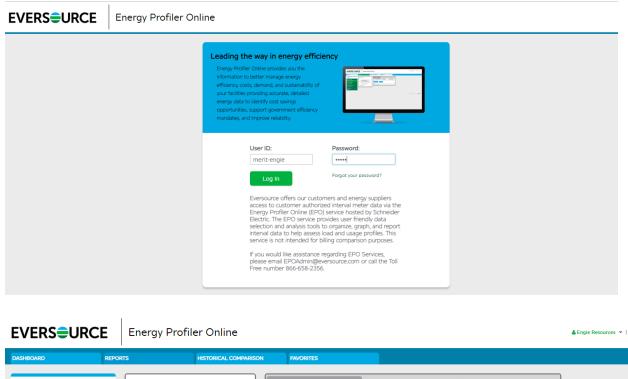
Hello Margaret,

ENGIE Resources Inc has requested access to your load data, for which you signed authorization. The data is provided via EPO Online, and you are welcome to use the service. The subscription automatically renews each year.
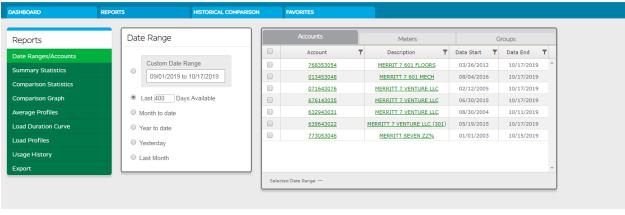
ENGIE Resources Inc,

The Merritt 7 Venture EPO Online account is now operational and ready for immediate use. The start of the annual subscription service is today, **10/17/2019** and will automatically renew each year. Below is the user id and password to begin accessing the data immediately.

Customer	BA Account Number	SA Number	User ID
Merritt 7 Venture	51-639643022	804432003	merit-engie
	51-676143035	702132006	
	51-071643076	501332007	
	51-632943031	972432002	
	51-013453048	706832009	
	51-773053046	469532007	
	51-768353054	427832005	
PASSWORD			
7877h			

4. Analysts manually log in to the portal (http://eversource.epo.schneider-electric.com/eversource) and download account data. Example below.

5. Analysts manually filter, adjust & drop the data (for each account) into /DATARESPONSE.
6. General volume is 2-5 emails per day with 1-10 accounts per email.

Proposed Process Enhancements

1. **Create GAHR file, send to ESG**

   Engie would create a GAHR file populated with portal logins, passwords and LDC accounts, for each utility. This would be sent via FTP to ESG, where they would return USHI files to be parsed by Engie.

   Example GAHR file:

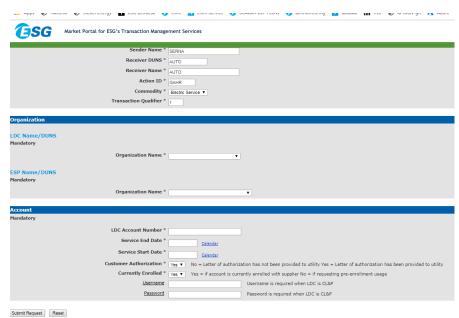## 2. Individually request through ESG portal

ESG currently has a portal that supports Eversource non-EDI data requests and returns USHI files. This involves using the account number (one at a time), login username and password. This could potentially be automated or used for "one-off" requests (i.e. when the utility replies to an LOA request saying Engie already has login credentials).
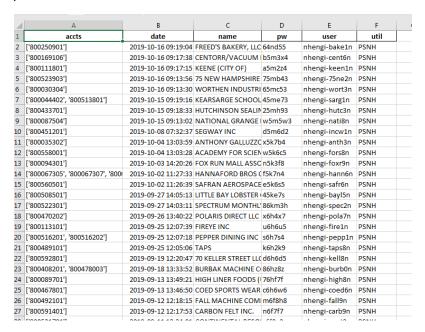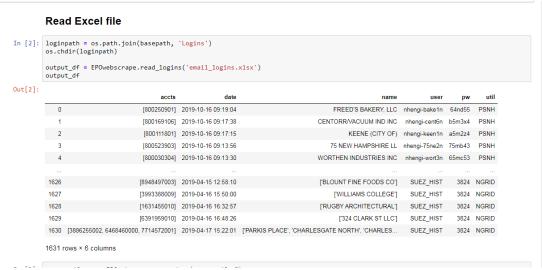
Example:

3. In-house webscrape

Upon receiving login credentials and account numbers, Engie could scrape the portal directly, filter the data and drop into /DATARESPONSE. This is currently a workaround implemented in Python but could potentially be done in a more efficient way by IT.

Key requirements with the portal is that the last 400 days is used, resulting in four or less accounts being able to be downloaded at once. Downloaded files are then parsed into Raw IDR, which is filtered into IDR data and dropped into /DATARESPONSE.

This process exists on a server, but requires a local scrape of email logins into an Excel file. Filtered data then exists on the server but must be manually downloaded one-by-one and OBSSed or dropped into /DATARESPONSE.

	A	B	C	D	E	F
1	accts	date	name	pw	user	util
2	['800250901']	2019-10-16 09:19:04	FREED'S BAKERY, LLC	64nd55	nhengi-bake1n	PSNH
3	['800169106']	2019-10-16 09:17:38	CENTORR/VACUUM	b5m3x4	nhengi-cent6n	PSNH
4	['800111801']	2019-10-16 09:17:15	KEENE (CITY OF)	a5m2z4	nhengi-keen1n	PSNH
5	['800523903']	2019-10-16 09:13:56	75 NEW HAMPSHIRE	75mb43	nhengi-75ne2n	PSNH
6	['800030304']	2019-10-16 09:13:30	WORTHEN INDUSTRI	65mc53	nhengi-wort3n	PSNH
7	['800044402', '800513801']	2019-10-15 09:19:16	KEARSARGE SCHOOL	45me73	nhengi-sarg1n	PSNH
8	['800433701']	2019-10-15 09:18:33	HUTCHINSON SEALIN	25mh93	nhengi-hutc3n	PSNH
9	['800087504']	2019-10-15 09:13:02	NATIONAL GRANGE I	w5m5w3	nhengi-nati8n	PSNH
10	['800451201']	2019-10-08 07:32:37	SEGWAY INC	d5m6d2	nhengi-incw1n	PSNH
11	['800035302']	2019-10-04 13:03:59	ANTHONY GALLUZZC	x5k7b4	nhengi-anth3n	PSNH
12	['800558001']	2019-10-04 13:03:28	ACADEMY FOR SCIEN	w5k6c5	nhengi-fors8n	PSNH
13	['800094301']	2019-10-03 14:20:26	FOX RUN MALL ASSC	n5k3f8	nhengi-foxr9n	PSNH
14	['800067305', '800067307', '800(	2019-10-02 11:27:33	HANNAFORD BROS (	f5k7n4	nhengi-hann6n	PSNH
15	['800560501']	2019-10-02 11:26:39	SAFRAN AEROSPACE	e5k6s5	nhengi-safr6n	PSNH
16	['800508501']	2019-09-27 14:05:13	LITTLE BAY LOBSTER (	45ke7s	nhengi-bayl5n	PSNH
17	['800522301']	2019-09-27 14:03:11	SPECTRUM MONTHL'	86km3h	nhengi-spec2n	PSNH
18	['800470202']	2019-09-26 13:40:22	POLARIS DIRECT LLC	x6h4x7	nhengi-pola7n	PSNH
19	['800113101']	2019-09-25 12:07:39	FIREYE INC	u6h6u5	nhengi-fire1n	PSNH
20	['800516201', '800516202']	2019-09-25 12:07:18	PEPPER DINING INC	s6h7s4	nhengi-pepp1n	PSNH
21	['800489101']	2019-09-25 12:05:06	TAPS	k6h2k9	nhengi-taps8n	PSNH
22	['800592801']	2019-09-19 12:20:47	70 KELLER STREET LL(	d6h6d5	nhengi-kell8n	PSNH
23	['800408201', '800478003']	2019-09-18 13:33:52	BURBAK MACHINE C	86hz8z	nhengi-burb0n	PSNH
24	['800089701']	2019-09-13 13:49:21	HIGH LINER FOODS (I	76hf7f	nhengi-high8n	PSNH
25	['800467801']	2019-09-13 13:46:50	COED SPORTS WEAR	c6h6w6	nhengi-coed6n	PSNH
26	['800492101']	2019-09-12 12:18:15	FALL MACHINE COMI	m6f8h8	nhengi-fall9n	PSNH
27	['800591401']	2019-09-12 12:17:53	CARBON FELT INC.	n6f7f7	nhengi-carb9n	PSNH

Read Excel file

```
In [2]:  loginpath = os.path.join(basepath, 'Logins')
         os.chdir(loginpath)

         output_df = EPOwebscrape.read_logins('email_logins.xlsx')
         output_df
```

Out[2]:

	accts	date	name	user	pw	util
0	[800250901]	2019-10-16 09:19:04	FREED'S BAKERY, LLC	nhengi-bake1n	64nd55	PSNH
1	[800169106]	2019-10-16 09:17:38	CENTORR/VACUUM IND INC	nhengi-cent6n	b5m3x4	PSNH
2	[800111801]	2019-10-16 09:17:15	KEENE (CITY OF)	nhengi-keen1n	a5m2z4	PSNH
3	[800523903]	2019-10-16 09:13:56	75 NEW HAMPSHIRE LL	nhengi-75ne2n	75mb43	PSNH
4	[800030304]	2019-10-16 09:13:30	WORTHEN INDUSTRIES INC	nhengi-wort3n	65mc53	PSNH
...	...	...	...	...	...	...
1626	[8948497003]	2019-04-15 12:58:10	['BLOUNT FINE FOODS CO']	SUEZ_HIST	3824	NGRID
1627	[3993388009]	2019-04-16 15:50:00	['WILLIAMS COLLEGE']	SUEZ_HIST	3824	NGRID
1628	[1631455010]	2019-04-16 16:32:57	['RUGBY ARCHITECTURAL']	SUEZ_HIST	3824	NGRID
1629	[6391959010]	2019-04-16 16:48:26	['324 CLARK ST LLC']	SUEZ_HIST	3824	NGRID
1630	[3886255002, 6468460000, 7714572001]	2019-04-17 15:22:01	['PARKIS PLACE', 'CHARLESGATE NORTH', 'CHARLES...	SUEZ_HIST	3824	NGRID

1631 rows × 6 columns

Download files from EPO portal

```
In [4]: output_df_rec
```

Out[4]:

	accts	date	name	user	pw	util
0	[51139828032_572781001, 51480628064_092192009]	2019-10-17 15:36:18	Stamford Washington Office	sbc-engie	pag42	CLP
1	[51639643022_804432003, 51676143035_702132006,...	2019-10-17 15:18:23	Merritt 7 Venture	merit-engie	7877h	CLP
2	[51013453048_706832009, 51773053046_469532007,...	2019-10-17 15:18:23	Merritt 7 Venture	merit-engie	7877h	CLP
3	[51670597012_759661001, 51441697000_804951001,...	2019-10-17 13:06:17	Griffin Park Realty	grpark-engie	892ck	CLP
4	[51414072009_412481000]	2019-10-17 13:02:55	400 Atlantic Title LLC	400a-engie	3336t	CLP
5	[51824665004_958832004]	2019-10-17 09:56:30	SKF	skf-engie	a72cc	CLP
6	[29228980016]	2019-10-17 09:50:56	REBEL RESTAURANTS LLC	regal-nengie	tts383	NSTAR
7	[29623540027]	2019-10-17 09:41:42	TMI OF WESTWOOD	tmi-nengie	8827dt	NSTAR
8	[26242141013]	2019-10-17 09:39:35	LQ MANAGEMENT LLC	lq-nengie	2324at	NSTAR
9	[54207668054_796921003]	2019-10-16 13:32:31	South Campus BH Holdings LLC	klond-engie	g3862	WMECO
10	[51040332041_613261002, 51036842094_893271005,...	2019-10-15 14:34:25	Illinois Tool Works Inc.	itool-engie	xt386	CLP
11	[51884534009_126882000, 51003193018_182162009]	2019-10-15 14:34:25	Illinois Tool Works Inc.	itool-engie	xt386	CLP
12	[26389111001]	2019-10-15 09:12:48	AEA TECH	aea-nengie	rtg291	NSTAR

```
In [*]: results = []
```

```
output_df_rec[ 'success' ] = results

user:  sbc-engie
password:  pag42
set to last  400  days.
logging on...
looking for 2 accts.
found items in portal.
trying search & download...

submitting...
downloaded EPO data file eversource_3a5595eb-e897-4c05-81e8-927becb56804_ddeec44f_hourlycsv.csv.
done with 1 of 13.
--------------------------
user:  merit-engie
password:  7877h
set to last  400  days.
logging on...
looking for 4 accts.
found items in portal.
trying search & download...

submitting...
error, accts ['51639643022_804432003', '51676143035_702132006', '51071643076_501332007', '51632943031_972432002'].
done with 2 of 13.
--------------------------
user:  merit-engie
password:  7877h
set to last  400  days.
logging on...
looking for 3 accts.
found items in portal.
```

Batch process downloaded EPO files into Raw IDRs

```
In [13]:  readpath = os.path.join(basepath, 'Downloads')
          writepath = os.path.join(basepath, 'Raw_IDR')
          error_log = []

          for i, file in enumerate(output_df_rec.files[:3]):
              try:
                  os.chdir(readpath)
                  filedf = pd.read_csv(file, sep = ",", header = 0)
                  #accts = ast.literal_eval(output_df.accts[i])
                  IDRdrop.raw_split(filedf, readpath, writepath, output_df_rec.util[i], output_df_rec.accts[i])
                  print('success, file: ', file)

              except:
                  print('error, file: ', file)

          print('RAW split complete.')
```

```
found 4 accounts.
CLP_51441697000_804951001_IDR_RAW.csv  written.
CLP_51670597012_759661001_IDR_RAW.csv  written.
CLP_51811108075_733761006_IDR_RAW.csv  written.
CLP_51923897094_005661008_IDR_RAW.csv  written.
success, file:  eversource_55a161b0-74f8-42a6-ada9-3813478ad56f_ac7efb37_hourlycsv.csv
found 1 accounts.
CLP_51414072009_412481000_IDR_RAW.csv  written.
success, file:  eversource_e4c735a9-c66a-4687-a00a-a57797d356ce_b30b40ee_hourlycsv.csv
error, file:  None
RAW split complete.
```

Batch filter Raw IDR into IDR files to be dropped

```
In [17]:  readpath = os.path.join(basepath, 'Raw_IDR')
          writepath = os.path.join(basepath, 'IDR')
          error_log = []

          for dropfile in processfiles:
              try:
                  IDRdrop.data_drop(dropfile, readpath, writepath)
                  print('success, file: ', dropfile)

              except:
                  error_log.append(dropfile)
                  print("error, file: ", dropfile)
```

```
success, file:  CLP_51414072009_412481000_IDR_RAW.csv
success, file:  CLP_51923897094_005661008_IDR_RAW.csv
overlap removed at 9/3/2019, for file:
success, file:  CLP_51441697000_804951001_IDR_RAW.csv
success, file:  CLP_51811108075_733761006_IDR_RAW.csv
overlap removed at 3/14/2019, for file:
success, file:  CLP_51670597012_759661001_IDR_RAW.csv
```