**Process for Uploading Water Rights Master Flat File to the California Open Data Portal**

This document describes the steps taken by OIMA staff to transform and load data from the eWRIMS water rights master flat file onto the California Open Data Portal (data.ca.gov) on a regular basis. Currently (as of 2021-06-23) these steps are coded into an R script which is run daily to maintain up-to-date data on the open data portal.

1. Download the most recent version of the “Water Rights Master Flat File” from the Water Board’s eWrims intranet site:
   1. Direct link to file: <https://intapps.waterboards.ca.gov/downloadFile/faces/flatFilesEwrims.xhtml?fileName=ewrims_flat_file.csv>
2. Convert field names to snake case (optional)
3. Drop some fields, including:
   * application\_number\_party,
   * pwss\_id,
   * fee\_due,
   * num\_comments,
   * num\_attachments,
   * last\_update\_date,
   * primary\_owner\_entity\_type\_p,
   * current\_status,
   * appl\_id,
   * permit\_permit\_id,
   * license\_license\_id
4. Check the fields that should be treated as dates or timestamps in the data portal’s API (fields listed below) for compatibility with the portal’s Data Store standards, and convert if necessary. For compatibility, date or timestamp fields should be in YYYY-MM-DD format, and missing / null values in these fields should be encoded as an empty text string. (NOTE: these standards are set be the data portal, and could change according to future guidance from the data portal administrators)
   * priority\_date
   * receipt\_date
   * rejection\_date
   * application\_recd\_date
   * application\_acceptance\_date
   * effective\_from\_date
   * effective\_to\_date
   * effective\_date
   * permit\_original\_issue\_date
   * complete\_construction\_date
   * complete\_applic\_water\_date
   * license\_original\_issue\_date
   * license\_requested\_date
   * inspection\_date
   * report\_date
   * offer\_sent\_date
   * accepted\_offer\_date
   * date\_received
   * date\_completed
   * enf\_case\_start\_date
   * enf\_case\_closure\_date
5. Check fields that should be treated as numeric in the data portal’s API (fields listed below) for compatibility with the portal’s Data Store, and convert if necessary. Numeric fields should only contain values that can be recognized as numeric (e.g. only numbers and decimal points – no text characters, spaces, or other special characters), and missing/null values in these fields should be encoded as “NaN”. (NOTE: these standards are set be the data portal, and could change according to future guidance from the data portal administrators)
   * number\_of\_protests
   * ini\_reported\_div\_amount
   * face\_value\_amount
   * fee\_received
   * appl\_fee\_amount
   * appl\_fee\_amt\_recd
   * max\_dd\_appl
   * max\_dd\_ann
   * max\_storage
   * max\_taken\_from\_source
   * year\_diversion\_commenced
   * max\_beneficially\_used
   * quantity\_of\_water\_diverted
   * quantity\_measurement\_year
   * max\_rate\_of\_diversion
   * recent\_water\_use\_min
   * recent\_water\_use\_max
   * drilled\_well\_year
   * depth\_of\_well
   * count\_npo\_or\_other
   * number\_of\_residences
   * use\_population
   * use\_population\_people
   * estimated\_use\_per\_person
   * use\_population\_stock
   * area\_for\_inci\_irrigation
   * use\_net\_acreage
   * use\_gross\_acreage
   * use\_direct\_div\_annual\_amount
   * use\_direct\_diversion\_rate
   * use\_storage\_amount
   * season\_direct\_div\_rate
   * season\_storage\_amount
   * season\_direct\_div\_aa
   * use\_count
   * pod\_number
   * direct\_div\_amount
   * direct\_diversion\_rate
   * storage\_amount
   * diversion\_rate\_to\_off\_stream
   * pod\_count
   * sp\_zone
   * north\_coord
   * east\_coord
   * latitude
   * longitude
   * section\_number
   * township\_number
   * range\_number
   * huc\_12\_number
   * huc\_8\_number
   * num\_of\_petitions
   * number\_of\_enforcement\_case
6. Ensure that missing/null values in remaining fields (i.e., all fields not treated as date/timestamp or numeric) are encoded with a meaningful value (e.g., “NA”)
7. Save the dataset to a plain text csv file (can be comma delimited, tab delimited, etc.)
8. Load the data from the file created in the previous step to the CA data portal using the portal’s Filestore API and the following information:
   1. API Key (unique code assigned to each user with an account on the data portal): OIMA staff can create new portal accounts for DIT staff if needed
   2. Resource ID (unique ID assigned to each resource/file loaded to the portal, can be found in the URL for that resource): 151c067a-088b-42a2-b6ad-99d84b48fb36
   3. Interface to the API (if needed):
      1. R scripts use the [ckanr package’s](https://cran.r-project.org/web/packages/ckanr/index.html) *resource\_update()* function
      2. Additional information and resources from the [CKAN Filestore and Datastore API documentation](https://docs.ckan.org/en/2.7/maintaining/filestore.html)