TRS Helper v2.0

Design Doc

- TRS Helper filters data from TRS to facilitate creation of custom reports

Reports Currently Supported

- Dump - Dumps the entire contents of TRS, no filter, into a JSON file

- All - Displays all available 'SYSTRAN' available that are stored in TRS ready for release

- Best – Given an input language list, displays the 'Best' technology engine available for that language pair defined by this priority list:
  - NMT v9.2, NMT v9.0, NMT v8-lua, SPE, SMT, RBMT

- TR ID – For a given input list, displays TR ID for all available technologies

- v9 – Lists all v9.0, and v9.2 TRs
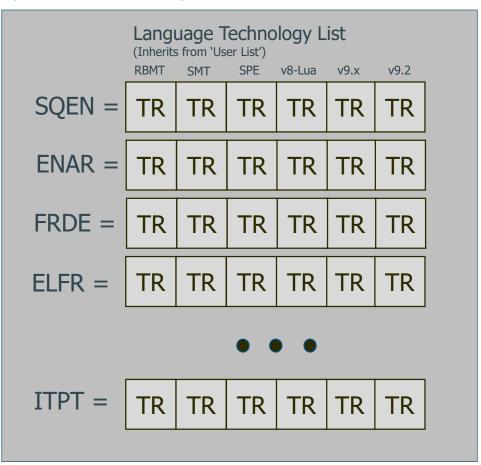 - Option to display updates to current TRS 'snapshot' by providing a 'previous' JSON TRS dump to

- trs_helper.py – main driver code, validates user input, and starts the various functionalities

- classes.py – implementation of LanguagePairMatrix, LanguageTechnologyList, TranslationResource

- reports.py – generates reports, data presentation layer

- config.py – user config

- language_lists.py – language lists used internally by 'Best' report

- language_lists_v9.json – user modifiable list of language lists for the 'v9' report

- test_module_pytest.py – pytest unit tests

Language Pair Matrix
(Inherits from 'UserDict')

Class: LanguagePairMatrix

- Inherits from 'UserDict'

- Behaves like a STD Lib Dictionary
    - Key = <language_pair_code>
    - Value = LanguageTechnologyList()

- Insert function is overloaded to store only the latest version of each Translator
    - WARNING – Should not insert to this object other than through the overloaded __setitem__() function

Class: LanguageTechnologyList

- Inherits from 'UserList'

- Behaves like a STD Lib List
    - WARNING – Should not insert to this object other than through the overloaded __setitem__() function

- Indexed by enumeration value for current SYSTRAN technology types
    - RBMT = 0, SMT = 1, SPE = 2, v8-lua = 3, v9.0 = 4, v9.2 = 5

Class: Translation Resource (TR)

- Implements Comparable

- Can be evaluated using standard comparison operators (<,>,<=,>=,==)

1. Reads current TRS snapshot
   - Either from a live TRS JSON dump or from a previously store simulated data

2. Populates LanguagePairMatrix with latest version of each available TR
   1. Must satisfy 'master', 'runnable', 'validated', 'stable', 'Generic', 'SYSTRAN'

3. Generates Report

Questions?