
# TRS Helper 2.0.0

***Release Notes***

# *Table of Contents*

# 1 Introduction

The 'TRS Helper' is a tool designed retrieve and display data from SYSTRAN's TRS system.

Currently, it provides the ability to run the following reports:

- Dump – dumps the entire contents of TRS, no filter, into a JSON file
- All – Displays all 'master', 'runnable', 'validated', 'stable', 'Generic', 'SYSTRAN' available that are stored in TRS
- Best – Given a list of Language Pairs, this report displays the 'Best' technology available. Best is defined by the descending priority list:
  - NMT v9.2, NMT v9.0, NMT v8-lua, SPE, SMT, RBMT
- TR ID – Given a list of LPs, display the latest version available across all available technologies, along with the TR ID
- NMT v9 – Given an input list of Language Pairs, displays all NMT v9.0 and v9.2 Translators
  - Diff option – the option exists to provide a previously dump raw TRS JSON file for comparison to show TRs that are added/deleted/upgraded/downgraded/equal

# 2 Requirements:

- Python Version >= 3.6.0
- 'requests' module installed:
  - `>> py -m pip install requests`
- 'moment' module installed:
  - `>> py -m pip install moment`

# 3 Testing

TRS Helper has been vigorously and rigorously tested using automated Unit and System tests, with Functional tests of all reporting capabilities performed in both a Windows, and CentOS 7 environment.

## 3.1 Unit Tests

- x61 Unit tests - **PASS**

```
D:\projects\TRS Helper\dev\2.0.0>python test_module_pytest.py
============================================================ test session starts ============================================================
platform win32 -- Python 3.6.0, pytest-4.3.0, py-1.8.0, pluggy-0.9.0
rootdir: D:\projects\TRS Helper\dev\2.0.0, inifile:
collected 61 items

test_module_pytest.py ..............................................................                                                  [100%]

========================================================= 61 passed in 0.75 seconds =========================================================
```

## 3.2 System Tests

- x33 System tests - **PASS**

## 3.3 Functional Tests

- Tests of all reporting capabilities performed in both a Windows, and CentOS 7 environment - **PASS**

# 4 Usage:

## 4.1 Dump TRS

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

```
>> python trs_helper.py -d
```

```
D:\projects\odni_languages_availability_script\development\common>python trs_helper.py -d
2019-04-01 16:20:00,275 - INFO - Starting TRS Helper
2019-04-01 16:20:00,276 - INFO - API Key is valid UUIDv4 - 417d                              1b4a
2019-04-01 16:20:00,280 - INFO - Reading live data from TRS - 'https://trs.systran.net/api/translationResources'
2019-04-01 16:20:24,465 - INFO - Dumping TRS data to JSON file '20190401_16-20-00_trs_dump.json'
```

## 4.2 Report – All

Displays all 'master', 'runnable', 'validated', 'stable', 'Generic', 'SYSTRAN' available that are stored in TRS

```
>> python trs_helper.py -a
```

```
D:\projects\odni_languages_availability_script\development\common>python trs_helper.py -a
2019-04-03 11:43:29,156 - INFO - Starting TRS Helper
2019-04-03 11:43:29,157 - INFO - API Key is valid UUIDv4 - 417d                              1b4a
2019-04-03 11:43:29,157 - INFO - Reading live data from TRS - 'https://trs.systran.net/api/translationResources'
2019-04-03 11:43:49,452 - INFO - Running Report: All SYSTRAN Generic LPs
2019-04-03 11:43:49,453 - INFO - Writing report to '20190403_11-43-29_all.txt/.csv'
```

### *4.2.1 Output*

Using the included

'TEMPLATE_SYSTRAN_generic_language_pairs_all_TEMPLATE.xlsx' workbook, the following report can be generated:

**SYSTRAN Generic MT Language Pairs**

| Source Language | Target Language | Src Code | Tgt Code | RBMT | SMT | Hybrid | v8 | v9.x | v9.2 |
|---|---|---|---|---|---|---|---|---|---|
| Albanian | English | SQ | EN | ✓ | | | | | |
| Arabic | English | AR | EN | ✓ | ✓ | ✓ | ✓ | ✓ | |
| Arabic | French | AR | FR | ✓ | ✓ | | | ✓ | |
| Bengali | English | BN | EN | | ✓ | | | | |
| Bulgarian | English | BG | EN | ✓ | ✓ | | | | ✓ |
| Chinese (Simplified) | English | ZH | EN | ✓ | | ✓ | ✓ | ✓ | ✓ |
| Chinese (Simplified) | Japanese | ZH | JA | | | | ✓ | | |
| Chinese (Simplified) | Korean | ZH | KO | ✓ | | | ✓ | | |
| Chinese (Traditional) | English | ZT | EN | ✓ | | | ✓ | | ✓ |
| Croatian | English | HR | EN | ✓ | ✓ | | ✓ | | ✓ |
| Czech | English | CS | EN | ✓ | | ✓ | ✓ | ✓ | |
| Danish | English | DA | EN | | ✓ | | ✓ | | ✓ |
| Dari | English | DR | EN | ✓ | | ✓ | | | |
| Dutch | English | NL | EN | ✓ | | ✓ | ✓ | ✓ | ✓ |
| Dutch | French | NL | FR | ✓ | | ✓ | | ✓ | |
| English | Arabic | EN | AR | ✓ | ✓ | ✓ | ✓ | ✓ | |
| English | Bengali | EN | BN | | ✓ | | | | |
| English | Bulgarian | EN | BG | | ✓ | | | | ✓ |
| English | Chinese (Simplified) | EN | ZH | ✓ | | ✓ | ✓ | ✓ | ✓ |
| English | Chinese (Traditional) | EN | ZT | ✓ | | | | | ✓ |
| Swedish | English | SV | EN | ✓ | | ✓ | ✓ | ✓ | ✓ |
| Tajik (Farsi) | English | TG | EN | ✓ | | | | | |
| Thai | English | TH | EN | | ✓ | | ✓ | | |
| Thai | Korean | TH | KO | | | | ✓ | | |
| Turkish | English | TR | EN | | ✓ | | ✓ | ✓ | |
| Ukrainian | English | UK | EN | ✓ | | ✓ | | ✓ | |
| Urdu | English | UR | EN | ✓ | | ✓ | ✓ | | |
| Vietnamese | English | VI | EN | | ✓ | | ✓ | | |
| Vietnamese | Korean | VI | KO | | | | ✓ | | |
| Welsh | English | CY | EN | | ✓ | | | | |

## 4.3 Report – Best

Given a list of Language Pairs, this report displays the 'Best' technology available. Best is defined by the descending priority list:

- NMT v9.2, NMT v9.0, NMT v8-lua, SPE, SMT, RBMT

```
>> python trs_helper.py -b
```

```
D:\projects\odni_languages_availability_script\development\common>python trs_helper.py -b
2019-04-03 11:46:01,756 - INFO - Starting TRS Helper
2019-04-03 11:46:01,759 - INFO - API Key is valid UUIDv4 - 417d:                                    i1b4a
2019-04-03 11:46:01,767 - INFO - Reading live data from TRS - 'https://trs.systran.net/api/translationResources'
2019-04-03 11:46:23,724 - INFO - Running Report: Best TRs Available
2019-04-03 11:46:23,724 - INFO - Writing report to '20190403_11-46-01_best_trs.txt/.csv'
Bi-Dir LPs English -> Foreign
                              RBMT    SMT     SPE     v8-LUA  v9.x    v9.2
```

### 4.3.1 Output

Using the included 'TEMPLATE_odni_languages_availability_EXTERNAL_TEMPLATE.xlsx' workbook, the following report can be generated

| Bi-Directional LPs | | | | |
|---|---|---|---|---|
| Language | RBMT | Hybrid (SMT/SPE) | NMT | |
| | | | v9.x | v9.2 |
| Arabic | | | ✓ | |
| Chinese (simplified) | | | | ✓ |
| Chinese (traditional) | | | | ✓ |
| Dutch | | | | ✓ |
| French | | | | ✓ |
| German | | | | ✓ |
| Greek | | | ✓ | |
| Italian | | | | ✓ |
| Japanese | | | | ✓ |
| Korean | | | | ✓ |
| Polish | | | | ✓ |
| Portuguese | | | | ✓ |
| Russian | | | ✓ | |
| Spanish | | | | ✓ |
| Swedish | | | | ✓ |

| Mono-Directional LPs | | | | |
|---|---|---|---|---|
| Language | RBMT | Hybrid (SMT/SPE) | NMT | |
| | | | v9.x | v9.2 |
| Albanian | ✓ | | | |
| Croatian | | | | ✓ |
| Czech | | | ✓ | |
| Dari | | ✓ | | |
| Farsi | | | ✓ | |
| Hindi | | ✓ | | |
| Pashto | | ✓ | | |
| Punjabi (Gurmukhi) | | | | |
| Punjabi (Shahmukhi) | ✓ | | | |
| Urdu | | ✓ | | |
| Serbian (Cyrillic) | | | | ✓ |
| Serbian (Latin) | | | | |
| Slovak | | | ✓ | |
| Tajik (Cyrillic) | | | | |
| Tajik (Farsi) | ✓ | | | |
| Ukrainian | | | ✓ | |

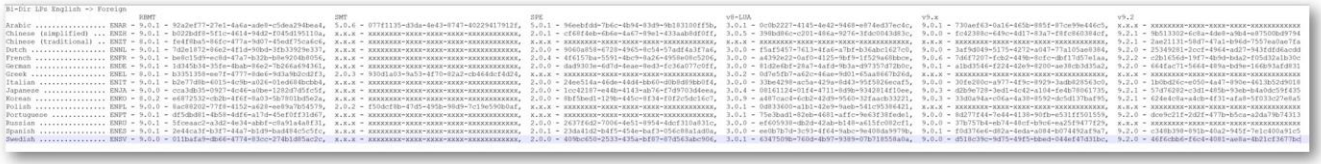
## 4.4 Report – TR ID

Given a list of LPs, display the latest version and TR ID available across all available technologies.

```
>> python trs_helper.py -id
```

```
D:\projects\odni_languages_availability_script\development\common>python trs_helper.py -id
2019-04-03 11:50:43,913 - INFO - Starting TRS Helper
2019-04-03 11:50:43,913 - INFO - API Key is valid UUIDv4 - 417d:                                    51b4a
2019-04-03 11:50:43,915 - INFO - Reading live data from TRS - 'https://trs.systran.net/api/translationResources'
2019-04-03 11:51:08,464 - INFO - Running Report: Version and TR ID
2019-04-03 11:51:08,465 - INFO - Writing report to '20190403_11-50-43_tr_id.txt/.csv'
Bi-Dir LPs English -> Foreign
```

### 4.4.1 Example output showing bi-directional ODNI languages for English > Foreign LPs



## 4.5 Report - NMT v9 Translator

NMT v9 – Given an input list of Language Pairs, displays all NMT v9.0 and v9.2 Translators

- Diff option – the option exists to provide a previously dump raw TRS JSON file for comparison to show TRs that are added/deleted/upgraded/downgraded/equal

### *4.5.1 Run report from live data on TRS*

```
>> python trs_helper.py -v9
```

```
D:\projects\odni_languages_availability_script\development\common>python trs_helper.py -v9
2019-04-01 16:21:36,084 - INFO - Starting TRS Helper
2019-04-01 16:21:36,084 - INFO - API Key is valid UUIDv4 - 417d                          1b4a
2019-04-01 16:21:36,089 - INFO - Reading live data from TRS - 'https://trs.systran.net/api/translationResources'
2019-04-01 16:21:54,540 - INFO - Running Report: NMT v9.x LPs
2019-04-01 16:21:54,540 - INFO - Writing report to '20190401_16-21-36_nmt_v9.txt/.csv'
```

### *4.5.2 Run report using simulated input data (a previously stored raw TRS .json dump)*

```
>> python trs_helper.py -in=< path_to_previous_dump>.json -v9
```

```
D:\projects\odni_languages_availability_script\development\common>python trs_helper.py -in=20190401_15-33-06_trs_dump.json -v9
2019-04-01 16:25:37,689 - INFO - Starting TRS Helper
2019-04-01 16:25:37,689 - INFO - Reading from simulated input data file - '20190401_15-33-06_trs_dump.json'
2019-04-01 16:25:38,416 - INFO - Running Report: NMT v9.x LPs
2019-04-01 16:25:38,417 - INFO - Writing report to '20190401_15-33-06_trs_dump_nmt_v9.txt/.csv'
```

### *4.5.3 Run report comparing live data on TRS against previously dumped raw TRS .json*

```
>> python trs_helper.py -v9 -=<path_to_previous_dump>.json
```

```
D:\projects\odni_languages_availability_script\development\common>python trs_helper.py -v9 -p=20190401_15-33-06_trs_dump.json
2019-04-01 16:30:06,055 - INFO - Starting TRS Helper
2019-04-01 16:30:06,055 - INFO - API Key is valid UUIDv4 - 417d2                         51b4a
2019-04-01 16:30:06,055 - INFO - Reading live data from TRS - 'https://trs.systran.net/api/translationResources'
2019-04-01 16:30:30,413 - INFO - Running Report: Compare NMT v9.x LPs against previous TRS JSON dump file.
2019-04-01 16:30:31,222 - INFO - Comparing against previous JSON TRS dump file - '20190401_15-33-06_trs_dump.json'
2019-04-01 16:30:31,223 - INFO - Writing report to '20190401_16-30-06_nmt_v9_diff.txt/.csv'
```

### *4.5.4 Run report comparing two (2) previously dumped raw TRS .json*

```
>> python trs_helper.py -
in=<path_to_previous_dump_new>.json -v9 -
p=<path_to_previous_dump_old>.json
```

- NOTE: conceptually, the '-in' option is considered to be 'current' data while '-p' is considered to be the 'previous', but in terms of function, the script doesn't care which input file is 'newer'

```
D:\projects\odni_languages_availability_script\development\common>python trs_helper.py -in=20190401_16-20-00_trs_dump.json -v9 -p=20190401_15-33-06_trs_dump.json
2019-04-01 16:33:03,153 - INFO - Starting TRS Helper
2019-04-01 16:33:03,153 - INFO - Reading from simulated input data file - '20190401_16-20-00_trs_dump.json'
2019-04-01 16:33:03,841 - INFO - Running Report: Compare NMT v9.x LPs against previous TRS JSON dump file.
2019-04-01 16:33:04,664 - INFO - Comparing against previous JSON TRS dump file - '20190401_15-33-06_trs_dump.json'
2019-04-01 16:33:04,665 - INFO - Writing report to '20190401_16-20-00_trs_dump_nmt_v9_diff.txt/.csv'
```

# 5 User Configuration

## 5.1 API Key

API Key can be entered as a positional argument at the cli, or for permanent storage, can be stored in the 'API_KEY' global variable located in 'utils\config.py'

- NOTE: API Key entered as positional argument will take precedence over the stored version in the case they are entered in both locations

```
    ##################################################
    ##  Description: User may set API here for convenience not having to type into cli argument
    ##      If API is stored here AND entered at cli, cli input will take precedence
    ##
    ##################################################
API_KEY = ''
```

## 5.2 Language Lists

- FileName: 'language_lists_v9.json'
  - Configurable in the 'LANGUAGE_LISTS_V9_FILE_NAME' global variable in 'utils\config.py':

```
    ##########################################################
    ##  Description: Language lists for NMT v9 Report
    ##
    ##########################################################
LANGUAGE_LISTS_V9_FILE_NAME = 'language_lists_v9.json'
```

- Contents: JSON Dictionary consisting of
  - Key = List Name
    - This is the name that will be used in the display report
    - List Names can be Created, Replace, Update, Delete
  - Value = List of Language Strings
    - Language String Format: [<language_name>, <SRC_language_code>, <TGT_language_code>]

- Language Lists can be Created, Replace, Update, Delete

```
{"PNMT Cloud LPs":[
      ["French", "FR", "EN"],
      ["French","EN","FR"],
      ["Italian","IT","EN"],
      ["Italian","EN","IT"],
      ["German","DE","EN"],
      ["German","EN","DE"],
      ["Spanish","ES","EN"],
      ["Spanish","EN","ES"],
      ["Portuguese","PT","EN"],
      ["Portuguese","EN","PT"],
      ["Arabic","AR","EN"],
      ["Arabic","EN","AR"],
      ["Russian","RU","EN"],
      ["Russian","EN","RU"],
      ["Chinese (simplified)","ZH","EN"],
      ["Chinese (simplified)","EN","ZH"],
      ["Japanese","JA","EN"],
      ["Japanese","EN","JA"],
      ["Korean","KO","EN"],
      ["Korean","EN","KO"]],
"ODNI Bi-Dir LPs": [
      ["Arabic","AR","EN"],
      ["Arabic", "EN", "AR"],
      ["Chinese (simplified)","ZH","EN"],
      ["Chinese (simplified)","EN","ZH"],
      ["Chinese (traditional)","ZT","EN"],
      ["Chinese (traditional)","EN","ZT"],
```

## 5.3 Logging Level

Logging Level can be configured in the in the 'LOG_FILE_LOG_LEVEL' global variable located in 'utils\config.py'

```
    # NOTSET    0
    # DEBUG     10  Detailed information, typically of interest only when diagnosing problems.
    # INFO      20  Confirmation that things are working as expected.
    # WARNING   30  An indication that something unexpected happened, or indicative of some problem in the near future
    #               (e.g. 'disk space low'). The software is still working as expected.
    # ERROR     40  Due to a more serious problem, the software has not been able to perform some function.
    # CRITICAL  50  A serious error, indicating that the program itself may be unable to continue running.
LOG_FILE_LOG_LEVEL = logging.INFO
```