

Using the Ryft ONE with your own data requires getting comfortable with two things:

1. RyftONE Record Definition File or the RDF
2. RyftONE searchable data must be located on the “/ryftone” file system (“/ryftone”)
 - Add Data: RDF must be defined PRIOR to placing data onto “/ryftone”.
 - Remove Data: data must be removed from “/ryftone” prior to removal of RDF definition.

Ryft ONE Record Definition File (RDF)

The Ryft ONE distinguishes between unique file types/contents by mapping the filename suffix to a Record Definition File (RDF.) The Ryft ONE supplies two “(system)” default RDFs which can’t be modified or deleted: “*.bin” and “*.txt.”

Each file suffix will match the system default “*.bin” RDF or can be overridden by a more specific suffix such as the “*.txt” RDF, or a user defined RDF. If no specific RDF file matches the file’s suffix, the default “*.bin” RDF is automatically associated with the data file.

User-defined RDFs are encouraged to leverage Ryft ONE’s strengths and flexibility. (Run the `rhfsctl -R` command to see a listing of all RDF definitions on your Ryft ONE.)

For more thorough coverage of RDF files, see the *Ryft Open API Library User Guide*, and the *Ryft ONE: User Guide*. Both files are located in the “~ryftuser/Documents” directory on the Ryft ONE.

For the following examples, we’ll focus on a few of the RDF fields:

- `data_type` = unstructured | XML | JSON
- `file_glob` = a unique file suffix match such as “*.bin”, “*.txt”, “*.censusxml”
- `rdf_name` = the name to be associated with this RDF.

The `rhfsctl` command line tool is used to manage RDF definitions. Issue the command `rhfsctl` to see options.

Here are three example RDFs, listed by RDF sequence number. Two RDFs are system supplied (#1 and #2) and one is user-defined (#3).

```
$ rhfsctl -R
 1: RawBinary => *.bin (system)
 2: PlainText => *.txt (system)
 3: censusxml => *.censusxml
```

Here is how the RDFs compare:

1: RawBinary => *.bin (system)	2: PlainText => *.txt (system)	3: censusxml => *.censusxml
Display source for “*.bin” RDF for unstructured data. All filename suffixes will map to “*.bin” RDF #1 unless a more specific suffix exists.	Display source for “*.txt” RDF for unstructured data, which matches filenames with suffix “*.txt”	Display source for “*.censusxml” RDF for structured data, which matches filenames with suffix “*.censusxml”
<pre>\$ rhfsctl -r 1 data_type = "unstructured"; file_glob = "*.bin"; chunk_size_mb = 64; rdf_name = "RawBinary"; rdf_description = "Default format with no content-based splitting.";</pre>	<pre>\$ rhfsctl -r 2 data_type = "unstructured"; file_glob = "*.txt"; chunk_size_mb = 64; rdf_name = "PlainText"; rdf_description = "Content is split on ASCII whitespace characters."; delimiter = "(\t \n \f \v \r)";</pre>	<pre>\$ rhfsctl -r 3 data_type = "XML" file_glob = "*.censusxml"; chunk_size_mb = 64; rdf_name = "censusxml"; record_start = "<Record>"; record_end = "</Record>"; data_type = "XML"; fields : { Name = ("<Name>", "</Name>"), Gender = ("<Gender>", "</Gender>"), BirthCount = ("<Count>", "</Count>"), Year = ("<Year>", "</Year>") };</pre>

Add and Remove Data Files

The proper RDF must be defined BEFORE placing data onto the “/ryftone” file system. If no specific RDF file matches the file’s suffix, the default “*.bin” RDF is automatically associated with the data file

NOTE: Ryft ONE data files may be added or removed at any time, and must be placed onto the “/ryftone” filesystem in order to be searchable.

When completely removing data files and RDF mappings, both the data file and the RDF registration must be removed. If the RDF registration remains, it will continue to be available for data file to RDF mapping.

NOTE: The data file suffix to RDF mapping is immutable. This means that a file’s RDF association cannot be broken without complete removal of the data files AND the RDF definition from the Ryft ONE.

Example 1: Copy sample data file “censusxml” to local machine/laptop, and modify local data to make it “uscensusxml”. See script file training_copy_data_to_laptop.sh.

```
# examples use unix compatible commandline, passwordless logins via ssh-keygen, ssh-copy-id
# export RYFTONE=172.16.13.3 # done ahead of time
mkdir census
scp -r ryftuser@$RYFTONE:/ryftone/ODBC/census/censusxml.rdf census/uscensusxml.rdf
scp -r ryftuser@$RYFTONE:/ryftone/ODBC/census/census.censusxml census/census.uscensusxml
# modify uscensus.rdf contents from 'censusxml' to 'uscensusxml'
sed -i 's/censusxml/uscensusxml/' census/uscensusxml.rdf
# show uscensusxml.rdf
cat census/uscensusxml.rdf
```

Example 2: Add “uscensusxml” data file to “/ryftone” file system on target system. See script file training_add_data_to_RyftONE.sh.

```
# copy rdf
scp census/uscensusxml.rdf ryftuser@$RYFTONE:/ryftone
# register rdf
ssh ryftuser@$RYFTONE "rhfsctl -a /ryftone/uscensusxml.rdf"
# verify rdf
ssh ryftuser@$RYFTONE "rhfsctl -f \"everyfilename.uscensusxml\""
```

```
# copy data
scp census/census.uscensusxml ryftuser@$RYFTONE:/ryftone
# test search
ssh ryftuser@$RYFTONE 'ryftrest -f "*.uscensusxml" -q "(RECORD.Name EQUALS \"Larry\")" --
local --count -vv'
```

Example 3: Completely remove/purge “uscensusxml” data files from target system. See script file `training_purge_data_from_RyftONE.sh`.

```
# remove known files
ssh ryftuser@$RYFTONE "rm -f /ryftone/*.uscensusxml"
ssh ryftuser@$RYFTONE "rm -f /ryftone/uscensusxml.rdf"
# perhaps look elsewhere for files
ssh ryftuser@$RYFTONE "find /ryftone -name '*.uscensusxml' 2>/dev/null|xargs rm -f"
ssh ryftuser@$RYFTONE "find /ryftone -name 'uscensusxml.rdf' 2>/dev/null|xargs rm -f"
# remove registered rdf
ssh ryftuser@$RYFTONE "rhfsctl -R |grep \"*.uscensusxml\""
ssh ryftuser@$RYFTONE "rhfsctl -R |grep \"*.uscensusxml\"; if [ \"\ $?\" -eq 0 ] ; then echo
\"rdf found\"; fi"
ssh ryftuser@$RYFTONE "rhfsctl -R |grep \"*.uscensusxml\"; if [ \"\ $?\" -eq 0 ] ; then
rhfsctl -R|grep \"*.uscensusxml\"|cut -d':' -f1|xargs rhfsctl -d; fi"
# confirm not found
ssh ryftuser@$RYFTONE "rhfsctl -R |grep \"*.uscensusxml\"; if [ \"\ $? -eq 0 ] ; then echo
\"purge NOT complete\"; el
se echo \"purge complete\"; fi"
```

Streaming Data

The following examples show streaming files to a Ryft ONE, searching the data, and purging the data.

Example: Stream data from one Ryft ONE to target Ryft ONE. See script file “`training_stream_data.sh`.”

```
# export RYFTONE=172.16.92.10 # done ahead of time
ssh ryftuser@$RYFTONE mkdir /ryftone/load

time date; for i in 1 2 3; do echo $i; dest=${i}_wiki.pbin; scp
/ryftone/regression/wikipedia*bin ryftuser@$RYFTONE:/ryftone/$dest; ssh ryftuser@$RYFTONE "mv
/ryftone/$dest /ryftone/load/"; done; date
```

Example: Search streamed data on target RyftONE which was loaded from above command sequence. See script file “`training_stream_search.sh`.”

```
# search for string in wikipedia
ssh ryftuser@$RYFTONE ryftrest -f \"load/*\" -q \"some other unique string\" --local --count
-vv
```

Example: Remove data loaded onto target Ryft ONE. See script file “`training_stream_purge.sh`.”

```
ssh ryftuser@$RYFTONE rm -rf /ryftone/load
```

Copyright (c) 2016, Ryft Systems, Inc.

All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

All advertising materials mentioning features or use of this software must display the following acknowledgement:

This product includes software developed by Ryft Systems, Inc.

Neither the name of Ryft Systems, Inc. nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY RYFT SYSTEMS, INC. "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL RYFT SYSTEMS, INC. BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.