# Metathesaurus - Rich Release Format (RRF)

[3]

Metathesaurus users may select from two relational formats: the Rich Release Format (RRF), first introduced in 2004, and the Original Release Format (ORF). Both are available as output options of MetamorphoSys, the installation and customization program.

Developers are encouraged to use the RRF, which offers significant advantages in source vocabulary transparency (that is, ability to exactly represent the detailed semantics of each source vocabulary); in the ability to generate complete and accurate change sets between versions of the Metathesaurus; and in more convenient representations of concept name, source, and hierarchical context information.

Neither Metathesaurus format is fully normalized. By design, there is duplication of data among different files and within certain files. In particular, relationships between different Metathesaurus concepts appear twice (e.g., from entry A to entry B and from entry B to entry A). Developers will need to make their own decisions about the extent to which this redundancy should be retained, reduced, or increased for their specfic applications.

All files except MRRANK.RRF are sorted by row.

## 3.1 Data Files

The data in each Metathesaurus entry may be represented in more than 20 different relations, or files. These files correspond to the four logical groups of data elements described in Sections 2.3 - 2.6 and the indexes described in Section 2.7 as follows:

- Concepts, Concept Names, and their sources (2.3) = MRCONSO.RRF
- Attributes (2.5) = MRSAT.RRF, MRDEF.RRF, MRSTY.RRF, MRHIST.RRF
- Relationships (2.4) = MRREL.RRF, MRCOC.RRF, MRCXT.RRF, MRHIER.RRF, MRMAP.RRF, MRSMAP.RRF
- Data about the Metathesaurus (2.6) = MRFILES.RRF, MRCOLS.RRF, MRDOC.RRF, MRRANK.RRF, MRSAB.RRF, AMBIGLUI.RRF, AMBIGSUI.RRF, CHANGE/MERGEDCUI.RRF, CHANGE/MERGEDLUI.RRF, CHANGE/DELETEDCUI.RRF, CHANGE/DELETEDLUI.RRF, CHANGE/DELETEDSUI.RRF, MRCUI.RRF
- Indexes (2.7) = MRXW_BAQ.RRF, MRXW_DAN.RRF, MRXW_DUT.RRF, MRXW_ENG.RRF, MRXW_FIN.RRF, MRXW_FRE.RRF, MRXW_GER.RRF, MRXW_HEB.RRF, MRXW_HUN.RRF, MRXW_ITA.RRF, MRXW_NOR.RRF, MRXW_POR.RRF, MRXW_RUS.RRF, MRXW_SPA.RRF, MRXW_SWE.RRF, MRXNW_ENG.RRF, MRXNS_ENG.RRF

## 3.2 Columns and Rows

Each file or named table of data values has by definition a fixed number of columns; the number of rows depends on the content of a particular version of the Metathesaurus.

A column is a sequence of all the values in a given data element or logical sub-element. In general, columns for longer variable length data elements will appear to the right of columns for shorter and/or fixed length data elements. The information for all columns in the files is

UMLS® Reference Manual

UMLS® Reference Manual

described in MRCOLS.RRF and on the Columns and Data Elements page of the current release documentation.

A row contains the values for one or more data elements or logical sub-elements for one Metathesaurus entry. Depending on the nature of the data elements involved, each Metathesaurus entry may have one or more rows in a given file. The values for the different data elements or logical sub-elements represented in the row are separated by vertical bars (|). If an optional element is blank, the vertical bars are still used to maintain the correct positioning of the subsequent elements. Each row is terminated by a vertical bar and line termination.

## 3.3 Descriptions of Each File

The descriptions of the files appear in the following order:

1. Key data about the Metathesaurus: Files; Columns or Data Elements; documentation that explains the meaning of abbreviations that appear as values in Metathesaurus data elements and attributes
2. Concept names and their vocabulary sources
3. Attributes
4. Relationships
5. Other data about the Metathesaurus
6. Indexes

Each file description lists the columns or data elements that appear in the file and includes sample rows from the file.

### 3.3.1 Files (File = MRFILES.RRF)

There is exactly one row in this file for each physical segment of each logical file. Data elements that appear in multiple files, e.g., CUI, AUI, will have multiple rows in this file.

| Col. | Description |
|---|---|
| FIL | Physical FILENAME |
| DES | Descriptive Name |
| FMT | Comma separated list of column names (COL), in order |
| CLS | # of COLUMNS |
| RWS | # of ROWS |
| BTS | Size in bytes in this format (ISO/PC or Unix) |

**Sample Records**

MRCOC.RRF|Co-occurringConcepts|CUI1,AUI1,CUI2,AUI2,SAB,COT,COF,COA,CVF|9|13939548|786509996|
MRSTY.RRF|Semantic Types|CUI,TUI,STN,STY,ATUI,CVF|6|1146352|64528811|

### 3.3.2 Data Elements (File = MRCOLS.RRF)

There is exactly one row in this file for each column or data element in each file. Data elements that appear in multiple files, e.g., CUI, AUI, will have multiple rows in this file.

UMLS® Reference Manual

UMLS® Reference Manual

| Col. | Description |
|---|---|
| COL | Column or data element name |
| DES | Descriptive Name |
| REF | Documentation Section Number |
| MIN | Minimum Length, Characters |
| AV | Average Length |
| MAX | Maximum Length, Characters |
| FIL | Physical FILENAME in which this field occurs |
| DTY | SQL-92 data type for this column |

**Sample Records**

AUI|Unique identifier for atom||8|8.00|8|MRCONSO.RRF|char(8)|
CODE|Unique Identifier or code for string in source||1|6.4|21|MRCONSO.RRF|varchar(50)|

### 3.3.3 Documentation for Abbreviated Values (File = MRDOC.RRF)

There is exactly one row in this table for each allowed value of selected data elements or attributes that have a finite number of abbreviations as allowed values. Examples of such data elements include TTY, ATN, TS, STT, REL, RELA.

| Col. | Description |
|---|---|
| DOCKEY | Data element or attribute |
| VALUE | Abbreviation that is one of its values |
| TYPE | Type of information in EXPL column |
| EXPL | Explanation of VALUE |

**Sample Records**

ATN|DDF|expanded_form|Drug Doseform|
ATN|DHJC|expanded_form|HCPCS J-code|

*Note: The MRDOC file produced by MetamorphoSys contains metadata about the release itself. Here is an example of the records:

RELEASE|mmsys.build.date|release_info|2006_01_31_17_30_34|
RELEASE|mmsys.version|release_info|7.7|

### 3.3.4 Concept Names and Sources (File = MRCONSO.RRF)

There is exactly one row in this file for each atom (each occurrence of each unique string or concept name within each source vocabulary) in the Metathesaurus, i.e., there is exactly one row for each unique AUI in the Metathesaurus. Every string or concept name in the Metathesaurus appears in this file, connected to its language, source vocabularies, and its concept identifier. The values of TS, STT, and ISPREF reflect the default order of precedence of vocabulary sources and term types in MRRANK.RRF. (Table 1)

**Sample Records**

C0001175|ENG|P|L0001175|VO|S0010340|Y|A0019182||M0000245|D000163|MSH|PM|
D000163|
Acquired Immunodeficiency Syndromes|0|N||
C0001175|ENG|S|L0001842|PF|S0011877|N|A2878223|103840012|62479008||
SNOMEDCT|PT|62479008|AIDS|4|N||
C0001175|ENG|P|L0001175|VC|S0354232|Y|A2922342|103845019|62479008||
SNOMEDCT|SY|62479008|
Acquired immunodeficiency syndrome|4|Y||
C0001175|FRE|P|L0162173|PF|S0226654|Y|A0248753||||INS|MH|d000163|SIDA|3|N||
C0001175|RUS|P|L0904943|PF|S1108760|Y|A1165232||||RUS|MH|D000163|SPID|3|N||

### 3.3.5 Simple Concept and Atom Attributes (File = MRSAT.RRF)

There is exactly one row in this table for each concept, atom, or relationship attribute that does not have a sub-element structure. All Metathesaurus concepts and a minority of Metathesaurus relationships have entries in this file. This file includes all source vocabulary attributes that do not fit into other categories. (Table 2)

**Sample Records**

C0001175|L0001175|S0010339|A0019180|AUI|D000163|AT15797077||FX|MSH|AIDS
Dementia Complex|N||
C0001175|L0001175|S0354232|A2922342|SAUI|62479008|AT34794876||
DESCRIPTIONSTATUS|SNOMEDCT|0|N||
C0001175|L2810384|S3645548|A3814219|SCUI|62479008|AT33494582||CTV3ID|
SNOMEDCT|XE0RX|N||
C0001175|L2810384|S3645548|A3814219|SCUI|62479008|AT33652930|\ISPRIMITIVE|
SNOMEDCT|0|N||
C0001175|||R19334287|SRUI||AT37098279||REFINABILITY|SNOMEDCT|1|N||

### 3.3.6 Definitions (File = MRDEF.RRF)

There is exactly one row in this file for each definition in the Metathesaurus. A definition is an attribute of an atom (an occurrence of a string in a source vocabulary). A few approach 3,000 characters in length. (Table 3)

**Sample Records**

C0001175|A0019180|AT15060425||MSH|An acquired defect of cellular immunity associated with infection by the human immunodeficiency virus (HIV), a CD4-positive T-lymphocyte count under 200 cells/microliter or less than 14% of total lymphocytes, and increased susceptibility to opportunistic infections and malignant neoplasms. Clinical manifestations also include emaciation (wasting) and dementia. These elements reflect criteria for AIDS as defined by the CDC in 1993.|N||

C0001175|A0021048|AT14042185||CSP|one or more indicator diseases, depending on laboratory evidence of HIV infection (CDC); late phase of HIV infection characterized by marked suppression of immune function resulting in opportunistic infections, neoplasms, and other systemic symptoms (NIAID).|N||

C0001175|A0021055|AT18420297||PDQ|Acquired immunodeficiency syndrome. An acquired defect in immune system function caused by human immunodeficiency virus 1 (HIV-1). AIDS is associated with increased susceptibility to certain cancers and to opportunistic infections, which are infections that occur rarely except in individuals with weak immune systems.|N||

UMLS® Reference Manual

UMLS® Reference Manual

### 3.3.7 Semantic Types (File = MRSTY.RRF)

There is exactly one row in this file for each Semantic Type assigned to each concept. All Metathesaurus concepts have at least one entry in this file. Many have more than one entry. The TUI, STN, and STY are all direct links to the UMLS Semantic Network.

| Col. | Description |
|---|---|
| CUI | Unique identifier of concept |
| TUI | Unique identifier of Semantic Type |
| STN | Semantic Type tree number |
| STY | Semantic Type. The valid values are defined in the Semantic Network. |
| ATUI | Unique identifier for attribute |
| CVF | Content View Flag. Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use. |

**Sample Record**

C0001175|T047|B2.2.1.2.1|Disease or Syndrome|AT17683839||

### 3.3.8 History (File = MRHIST.RRF)

This file tracks source-asserted history information. It currently includes SNOMED CT history only. (Table 4)

**Sample Records**

C0000294|108821000|SNOMEDCT|20001101|0|CONCEPTSTATUS|0|||
C0000294|108821000|SNOMEDCT|20020731|2|CONCEPTSTATUS|0|
FULLYSPECIFIEDNAME CHANGE||
C0000294|1185494016|SNOMEDCT|20020731|0|DESCRIPTIONSTATUS|0|||
C0000294|1461100014|SNOMEDCT|20030131|0|DESCRIPTIONSTATUS|0|||

### 3.3.9 Related Concepts (File = MRREL.RRF)

There is one row in this table (Table 5) for each relationship between concepts or atoms known to the Metathesaurus, with the following exceptions found in other files: co-occurrences found in MRCOC.RRF, and pair-wise mapping relationships between two source vocabularies found in MRMAP.RRF and MRSMAP.RRF.

Note that for asymmetrical relationships there is one row for each direction of the relationship. Note also the direction of REL - the relationship which the SECOND concept or atom (with Concept Unique Identifier CUI2 and Atom Unique Identifier AUI2) HAS TO the FIRST concept or atom (with Concept Unique Identifier CUI1 and Atom Unique Identifier AUI1).

**Sample Records**

C0002372|A0022284|AUI|RB|C0002371|A0022279|AUI||R01983351||MSH|MSH|||N||
C0002372|A0022284|AUI|SY|C0002372|A0062352|AUI||R18851331||MSH|MSH|||N||

### 3.3.10 Co-occurring Concepts (File = MRCOC.RRF)

This file includes statistical aggregations of co-occurrences of meanings in external data sources. These exist at the AUI level. There are two rows in this table for each pair of atoms that co-occur in each information source represented: one for each direction of the relationship. (Note that the COA data may be different for each direction of the relationship.) Many

Metathesaurus concepts have no entries in this file. Due to the very large number of co-occurrence relationships, they are distributed in a separate file. (Table 6)

Co-occurrences are concepts that occur together in the same entries in some information source. The relationships represented here are obtained from machine-manipulation of the information source. Co-occurrence relationships may exist between similar concepts (e.g., Atrial Fibrillation and Arrhythmia) or between very different concepts that nevertheless have some important connection in the field of biomedicine (e.g., Atrial Fibrillation and Digoxin), or between a primary concept and a qualifier (e.g., Lithotripsy and instrumentation). A co-occurrence relationship can exist between two concepts that have no other apparent relationship, although the frequency of such co-occurrences will be small.

In the current Metathesaurus, there are three sources of co-occurrence data: MEDLINE, AI/RHEUM, and CCPSS. From MEDLINE, co-occurrence data was computed for concepts that were designated as principal or main points in the same journal article i.e., the co-occurrence counts do not include articles in which either or both of the concepts were present and indexed in MEDLINE but not designated as main points. (A concept is considered to be a main point if the * is attached to the main heading or any of its subheadings.)

Two overall frequencies of MEDLINE co-occurrence are provided: one for recent MEDLINE data (MED) and one for MEDLINE data from a preceding block of years (MBD). Separate counts are provided for the frequencies with which the first concept was qualified by different MeSH qualifiers or by no qualifier at all when it co-occurred with the second concept. There are separate entries for each direction of the co-occurrence relationship. The related subheading occurrence information in each entry belongs to the first concept in the entry and is therefore different for each direction of the relationship.

In addition to the specific qualifier information associated with two co-occurring concepts, this element also includes in entries with LQ and LQB values for type of co-occurrence, totals for the number of times each main concept was qualified by a specific subheading or by no subheading.

The AI/RHEUM co-occurrence data represent the co-occurrence of diseases and findings in the AI/RHEUM knowledge base, i.e., the diseases that co-occur with a particular finding and the findings that co-occur with a particular disease. Each disease/finding pair can co-occur only once in the AI/RHEUM knowledge base.

In CCPSS, the co-occurrence data is extracted from patient records and includes problem-problem co-occurrences within a patient record as well as problem-modifier co-occurrences.

**Sample Records**

C0000294|A0085139|C0002421|A0022413|MBD|L|1|TU=1||
C0000294|A0085139|C0003968|A0026910|MED|L|1|UR=1||
C0000294|A0085139|C0006463|A0033415|MBD|L|1|<>=1||

### 3.3.11 Computable Hierarchies (File = MRHIER.RRF)

This file contains one row for each hierarchy or context in which each atom appears. If a source vocabulary does not contain hierarchies, its atoms will have no rows in this file. If a source vocabulary is multi-hierarchical (allows the same atom to appear in more than one hierarchy), some of its atoms will have more than one row in this file. MRHIER.RRF (Table 7) provides a complete and compact representation of all hierarchies present in all Metathesaurus source vocabularies. Hierarchical displays can be computed by combining data in this file with data

in MRCONSO.RRF. The distance-1 relationships, i.e., immediate parent, immediate child, and sibling relationships, represented in MRHIER.RRF also appear in MRREL.RRF.

**Sample Records**

C0001175|A2878223|1|A3316611|SNOMEDCT|isa|
A3684559.A2880798.A339606.A3287869.A3316611|||
C0001175|A2878223|2|A3512124|SNOMEDCT|isa|
A3684559.A2880798.A3398606.A3287869.A3512124|||
C0001175|A2878223|3|A3696836|SNOMEDCT|isa|
A3684559.A2880798.A3398606.A3399957.A3399109.A3144217.A3696836|||
C0001175|A2878223|4|A3512124|SNOMEDCT|isa|
A3684559.A2880798.A3398606.A3399957.A3399109.A3512124|||
C0001175|A2878223|5|A3316611|SNOMEDCT|isa|
A3684559.A2880798.A3512117.A3082701.A3316611|||
C0001175|A2878223|6|A2888699|SNOMEDCT|isa|
A3684559.A2880798.A3512117.A3082701.A3398847.A3398762.A2888699|||
C0001175|A2878223|7|A3316611|SNOMEDCT|isa|
A3684559.A2880798.A3512117.A3287869.A3316611|||
C0001175|A2878223|8|A3512124|SNOMEDCT|isa|
A3684559.A2880798.A3512117.A3287869.A3512124|||
C0001175|A2988194|1|A2888699|SNOMEDCT|isa|
A3684559.A2880798.A3512117.A3082701.A3398847.A3398762.A2888699|||

To find the specific concept names used in a hierarchy, look up the atom identifiers in the AUI and STR data elements in MRCONSO.RRF.

For most source vocabularies, the value of RELA (if present) applies up the hierarchy to the top or root. In other words, it also applies to the relationship between the atom's parent and the atom's grandparent, etc. The two exceptions in this version of the Metathesaurus are GO (Gene Ontology) and NIC (Nursing Intervention Classification). Except for GO and NIC atoms, the MRHIER rows for an atom's ancestors (parent, grandparent, etc.) contain no added information except the source-asserted hierarchical number or code (HCD). If this is not of interest, there may be no reason to find MRHIER rows for an atom's ancestors.

To find an atom's siblings in a specific context, find all MRHIER.RRF rows that share its SAB, RELA*, and PTR values.

To find an atom's children in a specific context, append a period (.) and the atom's AUI to its PTR and find all MRHIER.RRF rows with its SAB, RELA*, and the expanded PTR.

*The RELA is needed to retrieve correct siblings and children for University of Washington Digital Anatomist (UWDA) hierarchies. Some UWDA atoms appear in multiple hierarchies that are distinguished ONLY by their RELA values.

### 3.3.12 Contexts (File = MRCXT.RRF)

This file is no longer created by default. It has been replaced by MRHIER.RRF which is a correct, complete, and computable representation of hierarchies. Users who require the MRCXT (Table 8) file will need to create that file after creating a subset. To create the MRCXT file use the new MRCXT Builder application, accessible from the MetamorphoSys Welcome screen. Information on the MRCXT Builder can be found at http://www.nlm.nih.gov/research/umls/implementation_resources/metamorphosys/MRCXT_Builder.html. The information below describes the content of the file when produced by the MRCXT Builder.

This very large file contains pre-computed hierarchical context information (including concept names) intended to facilitate the display of hierarchies present in UMLS source vocabularies. All of the information in this file (plus additional sibling relationships) can be computed by joining the MRHIER.RRF file with MRCONSO.RRF. There can be many rows in this file for each occurrence of an atom in a hierarchy in any of the UMLS source vocabularies - a "context in" this discussion. Many Metathesaurus concepts have many atoms with contexts while others may have none. The number of rows per context differs depending on the number of ancestor, sibling, or child terms an atom has in that context. Because some atoms have multiple contexts in the same source, e.g., MeSH, a context number (CXN - e.g., 1, 2, 3) is used to identify all members of the same context. The CXNs are not global but are created as required for each atom. Each distinct context for a single atom can be retrieved with a CUI-AUI-SAB-CXN key. The "distance-1 relationships" i.e., the immediate parent, immediate child, and sibling relationships represented in MRCXT.RRF are also present in the MRREL.RRF file.

**Sample Records**

C0001175|S1911299|A1855909|ICPC2P|B9001|1|ANC|1|ICPC2-Plus|C1140253|A1861145|||||
C0001175|S1911299|A1855909|ICPC2P|B90001|1|ANC|2|BLOOD/BLOOD FORMING ORGANS/IMMUNE MECHANISM|C0847039|A1852564|B||||
C0001175|S1911299|A1855909|ICPC2P|B90001|1|ANC|2|Diagnosis/Diseases Component|C0497531|A0916974|7|||| C0001175|S1911299|A1855909|ICPC2P|B90001|1|ANC|3|HIV-INFECTION|AIDS|C0497169|A1852069|B90|||| C0001175|S1911299|A1855909|ICPC2P|B90001|1|CCP||Acquired Immune-Deficiency Syndrome|C0001175|A1855909|B90001||||

### 3.3.13 Mappings (File = MRMAP.RRF)

This file contains sets of mappings between vocabularies. Most mappings are between codes/identifiers (or expressions formed by codes/identifiers) from two different vocabularies. At least one of the vocabularies in each set of mappings is present in the Metathesaurus; usually both of them are. The version of a vocabulary that appears in a set of mappings may be different from the version of that vocabulary that appears in the other Metathesaurus release files. The versions of the vocabularies in a map set are specified by the FROMVSAB and TOVSAB attributes of the map set concept (see below). Users should be aware that the mappings are only valid between the versions of the vocabularies specified in these attributes. The version of the map set itself is specified by the MAPSETVERSION attribute of the map set concept.

The MRMAP.RRF (Table 9) file is complex, to allow for more complicated mappings. Where possible, all mappings are also represented in the simpler MRSMAP.RRF file described below.

Each set of mappings is represented by a map set concept in MRCONSO.RRF (with TTY = ‘XM’) identified by a CUI (MAPSETCUI). Metadata of a map set are found in MRSAT.RRF as attributes of the map set concept. Each map set has three SAB values associated with it: the SAB of the map set itself (MAPSETVSAB), the SAB of the source being mapped (FROMVSAB) and the SAB of the source being mapped to (TOVSAB). Thus, a single map set asserts mappings from only one source to only one other source.

A subset of the mappings is redundantly represented as mapped_to and mapped_from relationships in MRREL.RRF. These are one-to-one mappings between two vocabularies which are both present in the UMLS. These general relationships are not as precise as the mapping files, since any differences between versions of the vocabularies in the map set and the versions of those vocabularies in the rest of the Metathesaurus files are ignored. Such

differences may affect the validity of the relationships in MRREL.RRF in a small number of cases.

There are three map sets that contain mappings from Metathesaurus concepts (represented by CUIs) to expressions formed by one or more concept names. These were formerly called associated expressions, and all have MAPTYPE='ATX'. This data is derived from earlier mapping efforts and is represented in the MRATX file in ORF.

**Sample Records**
**Map set concepts (in MRCONSO.RRF):**

|C1306694|ENG|P|L3139022|PF|S3660621|Y|A3829740||||MTH|XM|1000|MSH Associated Expressions|0|N|0|
|C1321851|ENG|P|L3502396|PF|S4036398|Y|A4363175|100046|||SNOMEDCT|XM|100046|SNOMEDCT mappings to ICD-9-CM|4|N|0|

**Map set metadata (in MRSAT.RRF):**

|C1321851|L3502396|S4036398|A4363175|CODE|100046|AT58994529||FROMVSAB|MTH|SNOMEDCT_2006_01_31|0|
|C1321851|L3502396|S4036398|A4363175|CODE|100046|AT58994530||MAPSETVSAB|MTH|SNOMEDCT_2006_01_31|N|0|
|C1321851|L3502396|S4036398|A4363175|CODE|100046|AT58994531||TOVSAB|MTH|ICD9CM_2006|N|0|

**Mappings (in MRMAP.RRF):**

|C1306694|MTH||0|AT28307527||C0011764||C0011764|CUI|||RO||2201||<Developmental Disabilities> AND <Writing>|BOOLEAN_EXPRESSION|||||ATX|||0|
|C1306694|MTH||0|AT52620421||C0010700||C0010700|CUI|||RN||1552||<Bladder>/<surgery>|BOOLEAN_EXPRESSION|||||ATX|||0|
|C1321851|SNOMEDCT|0|0|AT31496368||68995007|68995007|68995007|SCUI|||RQ|mapped_to|4751059|4751059|295.22|CODE|||||1|||0|
|C1321851|SNOMEDCT|0|0|AT31496369||404049001|404049001|404049001|SCUI|||RN|mapped_to|4388059|4388059|215.9|CODE|||||2|||0|

### 3.3.14 Simple Mappings (File = MRSMAP.RRF)

This file provides a simpler representation of most of the mappings in MRMAP.RRF (Table 10) to serve applications which do not require the full richness of the MRMAP.RRF data structure. Generally, mappings that support rule-based processing need the additional fields of MRMAP.RRF (e.g. MAPRANK, MAPRULE, MAPRES) and will not be represented in MRSMAP.RRF. More specifically, all mappings with non-null values for MAPSUBSETID and MAPRANK are excluded from MRSMAP.RRF.

**Sample Records**

|C1306694|MTH|AT28312030||C0009215|CUI|SY||<Codeine> AND <Drug Hypersensitivity>|BOOLEAN_EXPRESSION||
|C1306694|MTH|AT28312033||C0795964|CUI|RU||<Speech Disorders>|BOOLEAN_EXPRESSION||
|C1321851|SNOMEDCT|AT31496368||68995007|SCUI|RQ|mapped_to|295.22|CODE|0|
|C1321851|SNOMEDCT|AT31496369||404049001|SCUI|RN|mapped_to|215.9|CODE|0|

### 3.3.15 Source Information (File = MRSAB.RRF)

The Metathesaurus has "versionless" or "root" Source Abbreviations (SABs) in the data files. MRSAB.RRF connects the root SAB to fully specified version information for the current release. For example, the released SAB for MeSH is now simply "MSH". In MRSAB.RRF (Table 11), you will see a current versioned SAB, e.g., MSH2003_2002_10_24. MRSAB.RRF allows all other Metathesaurus files to use versionless source abbreviations, so that all rows with no data change between versions remain unchanged. MetamorphoSys can produce files with either the root or versioned SABs so that either form can be available in custom subsets of the Metathesaurus.

There is one row in this file for every version of every source in the current Metathesaurus; eventually there will also be historical information with a row for each version of each source that has appeared in any Metathesaurus release. Note that the field CURVER has the value Y to identify the version in this Metathesaurus release. Future releases of MRSAB.RRF will also contain historical version information in rows with CURVER value N.

Sources with contexts have "full" contexts, i.e., all levels of terms may have Ancestors, Parents, Children and Siblings. A full context may also be further designated as Multiple, Nosib (No siblings) or both Multiple and Nosib.

Multiple indicates that a single concept in this source may have multiple hierarchical positions.

No siblings (Nosib) indicates that siblings have not been computed for this source.

The Source Vocabularies page of the current release documentation lists each source in the Metathesaurus and includes information about the type of context, if any, for each source.

**Sample Records**

C1371270|C1140284|RXNORM_04AB|RXNORM|RXNORM Project, META2004AB | RXNORM | 04AB | 2004_05_17 | | 2004AB | | Stuart Nelson, M.D., Head, MeSH Section; e-mail: nelson@nlm.nih.gov | Stuart Nelson, M.D., Head, MeSH Section; e-mail: nelson@nlm.nih.gov | 0 | 138005 | 110403 | | BN,IN,OBD,OCD,SBD,SBDF,SCD,SCDC,SCDF,SY | ORIG_CODE,ORIG_SOURCE | ENG | UTF-8 | Y | Y | RxNorm work done by the National Library of Medicine|RxNorm work done by NLM. Bethesda (MD): National Library of Medicine, META2004AB release. | |

### 3.3.16 Concept Name Ranking (File = MRRANK.RRF)

There is exactly one row for each concept name type from each Metathesaurus source vocabulary (each SAB-TTY combination). The RANK and SUPPRESS values in the distributed file are those used in Metathesaurus production. Users are free to change these values to suit their needs and preferences, then change the naming precedence and suppressibility by using MetamorphoSys to create a customized Metathesaurus. (Table 12)

**Sample Records**

0210|AIR|SY|N|
0209|ULT|PT|N|
0208|CPT|PT|N|

### 3.3.17 Ambiguous Term Identifiers (File = AMBIGLUI.RRF)

In the instance that a Lexical Unique Identifier (LUI) is linked to multiple Concept Unique Identifiers (CUIs), there is one row in this table for each LUI-CUIs pair. This file identifies those lexical variant classes which have multiple meanings in the Metathesaurus.

In the Metathesaurus, the LUI links all strings within the English language that are identified as lexical variants of each other by the luinorm program found in the UMLS SPECIALIST Lexicon and Tools. LUIs are assigned irrespective of the meaning of each string. This table may be useful to system developers who wish to use the lexical programs in their applications to identify and disambiguate ambiguous terms.

| Col. | Description |
|---|---|
| LUI | Lexical Unique Identifier |
| CUI | Concept Unique Identifier |

**Sample Records**

L0000003 |C0010504|
L0000003 |C0917995|
L0000032 |C0010206|
L0000032 |C0010207|

### 3.3.18 Ambiguous String Identifiers (File = AMBIGSUI.RRF)

In the instance that a String Unique Identifier (SUI) is linked to multiple Concept Unique Identifiers (CUIs), there is one row in this table for each SUI-CUIs pair.

This file resides in the META directory. In the Metathesaurus, there is only one SUI for each unique string within each language, even if the string has multiple meanings. This table is only of interest to system developers who use the SUI in their applications or in local data files.

| Col. | Description |
|---|---|
| SUI | String Unique Identifier |
| CUI | Concept Unique Identifier |

**Sample Records**

S0000176 |C0042266|
S0000176 |C0546846|
S0000217 |C0024817|
S0000217 |C0555026|

### 3.3.19 Metathesaurus Change Files

There are six files or relations that identify key differences between entries in the previous and the current edition of the Metathesaurus. Developers can use these special files to determine whether there have been changes that affect their applications.

The usefulness of individual files will depend on how data from the Metathesaurus have been linked or incorporated in a particular application.

Each relation or named table of data has a fixed number of columns and variable number of rows. A column is a sequence of all the values in a given data element. A row contains the

values for two or more data elements for one entry. The values for the different data elements in the row are separated by vertical bars (|). Each row ends with a vertical bar and line termination.

### 3.3.19.1 Deleted Concepts (File = CHANGE/DELETEDCUI.RRF)

Concepts whose meaning is no longer present in the Metathesaurus are reported in this file. There is a row for each concept that existed in the previous release and is not present in the current release. If the meaning exists in the current release, i.e., the missing concept was merged with another current concept, it is reported in the MERGEDCUI.RRF file (Section 3.3.19.2) and not in this file.

| Col. | Description |
|---|---|
| PCUI | Concept Unique Identifier in the previous Metathesaurus |
| PSTR | Preferred name of this concept in the previous Metathesaurus |

### 3.3.19.2 Merged Concepts (File = CHANGE/MERGEDCUI.RRF)

There is exactly one row in this table for each released concept in the previous Metathesaurus (CUI1) that was merged into another released concept from the previous Metathesaurus (CUI2). When this merge occurs, the first CUI (CUI1) was retired; this table shows the CUI (CUI2) for the merged concept in this Metathesaurus.

Entries in this file represent concepts pairs that were considered to have different meanings in the previous edition, but which are now identified as synonyms.

| Col. | Description |
|---|---|
| PCUI1 | Concept Unique Identifier in the previous Metathesaurus |
| CUI | Concept Unique Identifier in this Metathesaurus in format C####### |

### 3.3.19.3 Deleted Terms (File=CHANGE/DELETEDLUI.RRF)

There is exactly one row in this table for each Lexical Unique Identifier (LUI) that appeared in the previous Metathesaurus, but does not appear in this Metathesaurus.

LUIs are assigned by the luinorm program, part of the lvg program in the UMLS SPECIALIST Lexicon and Tools; see Chapter 6.

These entries represent the cases where LUIs identified by the previous release's luinorm program, when used to identify lexical variants in the previous Metathesaurus, are no longer found with this release's luinorm on this release's Metathesaurus. This does not necessarily imply the deletion of a string or a concept from the Metathesaurus.

| Col. | Description |
|---|---|
| PLUI | Lexical Unique Identifier in the previous Metathesaurus |
| PSTR | Preferred Name of Term in the previous Metathesaurus |

### 3.3.19.4 Merged Terms (File = CHANGE/MERGEDLUI.RRF)

There is exactly one row in this file for each case in which strings had different Lexical Unique Identifiers (LUIs) in the previous Metathesaurus yet share the same LUI in this Metathesaurus; a LUI present in the previous Metathesaurus is therefore absent from this Metathesaurus.

LUIs are assigned by the luinorm program, part of the lvg program in the UMLS SPECIALIST Lexicon and Tools; see Chapter 6.

These entries represent the cases where separate lexical variants as identified by the previous release's luinorm program version are a single lexical variant as identified by this release's luinorm.

| Col. | Description |
|---|---|
| PLUI | Lexical Unique Identifier in the previous Metathesaurus but not present in this Metathesaurus |
| LUI | Lexical Unique Identifier into which it was merged in this Metathesaurus |

### *3.3.19.5 Deleted Strings (File = CHANGE/DELETEDSUI.RRF)*

There is exactly one row in this file for each string in each language that was present in an entry in the previous Metathesaurus and does not appear in this Metathesaurus.

Note that this does not necessarily imply the deletion of a term (LUI) or a concept (CUI) from the Metathesaurus. A string deleted in one language may still appear in the Metathesaurus in another language.

| Col. | Description |
|---|---|
| PSUI | String Unique Identifier in the previous Metathesaurus that is not present in this Metathesaurus |
| PSTR | Preferred Name of Term in the previous Metathesaurus that is not present in this Metathesaurus |

### *3.3.19.6 Retired CUI Mapping (File = MRCUI.RRF)*

There are one or more rows in this file (Table 13) for each Concept Unique Identifier (CUI) that existed in any prior release but is not present in the current release. The file includes mappings to current CUIs as synonymous or to one or more related current CUI where possible. If a synonymous mapping cannot be found, other relationships between the CUIs can be created. These relationships can be Broader (RB), Narrower (RN), Other Related (RO), Deleted (DEL) or Removed from Subset (SUBX). Rows with the SUBX relationship are added to MRCUI by MetamorphoSys for each CUI that met the exclusion criteria and was consequently removed from the subset. Some CUIs may be mapped to more than one other CUI using these relationships.

CUIs may be retired when (1) two released concepts are found to be synonyms and so are merged, retiring one CUI; (2) the concept no longer appears in any source vocabulary and is not 'rescued' by NLM; or (3) the concept is an acknowledged error in a source vocabulary or determined to be a Metathesaurus production error.

See Sections 3.3.19 1 through 5 for files containing changes from the last release only, without mappings.

**Sample Records**

C1313903|2004AA|SY|||C0525045|Y|
C1313909|2004AA|RO|||C0476661|Y|
C1321833|2004AA|DEL|||||
C1382264|2004AB|SY|||C0993613|Y|
C1382494|2004AB|DEL|||||

#### *3.3.19.7 AUI Movements (File = MRAUI.RRF)*

This file records the movement of Atom Unique Identifiers (AUIs) from a concept (CUI1) in one version of the Metathesaurus to a concept (CUI2) in the next version (VER) of the Metathesaurus. The file is historical. (Table 14)

**Sample Records**

A0000039|C0236824|2004AC|||move|A0000039|C1411876|Y|
A0000077|C0003477|2005AB|||move|A0000077|C1510447|Y|
A8177040|C1237728|2005AB|||move|A8177040|C1237732|Y|

### 3.3.20 Word Index (File = MRXW_BAQ.RRF, MRXW_DAN.RRF, MRXW_DUT.RRF, MRXW_ENG.RRF, MRXW_FIN.RRF, MRXW_FRE.RRF, MRXW_GER.RRF, MRXW_HEB.RRF, MRXW_HUN.RRF, MRXW_ITA.RRF, MRXW_NOR.RRF, MRXW_POR.RRF, MRXW_RUS.RRF, MRXW_SPA.RRF, MRXW_SWE.RRF)

There is one row in these tables for each word found in each unique Metathesaurus string (ignoring upper-lower case). All Metathesaurus entries have entries in the word index. The entries are sorted in ASCII order.

| Col. | Description |
|---|---|
| LAT | Abbreviation of language of the string in which the word appears |
| WD | Word in lowercase |
| CUI | Concept identifier |
| LUI | Term identifier |
| SUI | String identifier |

**Sample Records from MRXW_ENG.RRF**

ENG|anaemia|C0002871|L0280031|S0352688|
ENG|anemia|C0002871|L0002871|S0013742|
ENG|anemias|C0002871|L0002871|S0013787|
ENG|blood|C0002871|L0376533|S0500659|

**Sample Records from MRXW_FRE.RRF**

FRE|ANEMIE|C0002871|L0162748|S0227229|

### 3.3.21 Normalized Word Index (File = MRXNW_ENG.RRF)

There is one row in this table for each normalized word found in each unique English-language Metathesaurus string. All English-language Metathesaurus entries have entries in the normalized word index. There are no normalized string indexes for other languages in the Metathesaurus.

| Col. | Description |
|---|---|
| LAT | Abbreviation of language of the string in which the word appears (always ENG in this edition of the Metathesaurus) |
| NWD | Normalized word in lowercase (described in Section 2.7.2.1) |
| CUI | Concept identifier |
| LUI | Term identifier |
| SUI | String identifier |

**Sample Records**

ENG|anaemia|C0002871|L0280031|S0352688|
ENG|anemia|C0002871|L0002871|S0013742|
ENG|anemia|C0002871|L0002871|S0013787|
ENG|blood|C0002871|L0376533|S0500659|

### 3.3.22 Normalized String Index (File = MRXNS_ENG.RRF)

There is one row in this table for each normalized string found in each unique English-language Metathesaurus string (ignoring upper-lower case). All English-language Metathesaurus entries have entries in the normalized string index. There are no normalized word indexes for other languages in this edition of the Metathesaurus.

| Col. | Description |
|---|---|
| LAT | Abbreviation of language of the string (always ENG in this edition of the Metathesaurus) |
| NSTR | Normalized string in lowercase (described in Section 2.7.3.1) |
| CUI | Concept identifier |
| LUI | Term identifier |
| SUI | String identifier |

**Sample Records**

ENG|anaemia|C0002871|L0280031|S0352688|
ENG|anaemia unspecified|C0002871|L0696700|S0803315|
ENG|anemia|C0002871|L0002871|S0013787|

UMLS® Reference Manual

UMLS® Reference Manual

Table 1. Concept Names and Sources (File = MRCONSO.RRF)

| Col. | Description |
|---|---|
| CUI | Unique identifier for concept |
| LAT | Language of term |
| TS | Term status |
| LUI | Unique identifier for term |
| STT | String type |
| SUI | Unique identifier for string |
| ISPREF | Atom status - preferred (Y) or not (N) for this string within this concept |
| AUI | Unique identifier for atom - variable length field, 8 or 9 characters |
| SAUI | Source asserted atom identifier [optional] |
| SCUI | Source asserted concept identifier [optional] |
| SDUI | Source asserted descriptor identifier [optional] |
| SAB | Abbreviated source name (SAB). Maximum field length is 20 alphanumeric characters. Two source abbreviations are assigned:<br>• Root Source Abbreviation (RSAB) — short form, no version information, for example, AI/RHEUM, 1993, has an RSAB of "AIR"<br>• Versioned Source Abbreviation (VSAB) — includes version information, for example, AI/RHEUM, 1993, has an VSAB of "AIR93"<br><br>Official source names, RSABs, and VSABs are included on the Source Vocabularies page. |
| TTY | Abbreviation for term type in source vocabulary, for example PN (Metathesaurus Preferred Name) or CD (Clinical Drug). Possible values are listed on the Abbreviations Used in Data Elements page. |
| CODE | Most useful source asserted identifier (if the source vocabulary has more than one identifier), or a Metathesaurus-generated source entry identifier (if the source vocabulary has none) |
| STR | String |
| SRL | Source restriction level |
| SUPPRESS | Suppressible flag. Values = O, E, Y, or N<br>O: All obsolete content, whether they are obsolesced by the source or by NLM. These will include all atoms having obsolete TTYs, and other atoms becoming obsolete that have not acquired an obsolete TTY (e.g. RxNorm SCDs no longer associated with current drugs, LNC atoms derived from obsolete LNC concepts).<br>E: Non-obsolete content marked suppressible by an editor. These do not have a suppressible SAB/TTY combination.<br>Y: Non-obsolete content deemed suppressible during inversion. These can be determined by a specific SAB/TTY combination explicitly listed in MRRANK.<br>N: None of the above<br>Default suppressibility as determined by NLM (i.e., no changes at the Suppressibility tab in MetamorphoSys) should be used by most users, but may not be suitable in some specialized applications. See the MetamorphoSys Help page for information on how to change the SAB/TTY suppressibility to suit your requirements. NLM strongly recommends that users not alter editor-assigned suppressibility, and MetamorphoSys cannot be used for this purpose. |
| CVF | Content View Flag. Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use. |

Table 2. Simple Concept and Atom Attributes (File = MRSAT.RRF)

| Col. | Description |
|---|---|
| CUI | Unique identifier for concept (if METAUI is a relationship identifier, this will be CUI1 for that relationship) |
| LUI | Unique identifier for term (optional - present for atom attributes, but not for relationship attributes) |
| SUI | Unique identifier for string (optional - present for atom attributes, but not for relationship attributes) |
| METAUI | Metathesaurus atom identifier (will have a leading A) or Metathesaurus relationship identifier (will have a leading R) or blank if it is a concept attribute. |
| STYPE | The name of the column in MRCONSO.RRF or MRREL.RRF that contains the identifier to which the attribute is attached, e.g., SAUI, SCUI, SRUI, CODE, CUI, AUI. As vocabularies that were added to the Metathesaurus prior to the development of the RRF are updated and brought into complete alignment with the RRF, many attributes currently shown as linked to Metathesaurus AUIs will be linked to one of the source vocabulary identifiers. |
| CODE | Most useful source asserted identifier (if the source vocabulary contains more than one) or a Metathesaurus-generated source entry identifier (if the source vocabulary has none). Optional - present if METAUI is an AUI. |
| ATUI | Unique identifier for attribute |
| SATUI | Source asserted attribute identifier (optional - present if it exists) |
| ATN | Attribute name. Possible values appear in MRDOC.RRF and are described on the Attribute Names page. |
| SAB | Abbreviated source name (SAB). Maximum field length is 20 alphanumeric characters. Two source abbreviations are assigned:<br>• Root Source Abbreviation (RSAB) — short form, no version information, for example, AI/RHEUM, 1993, has an RSAB of "AIR"<br>• Versioned Source Abbreviation (VSAB) — includes version information, for example, AI/RHEUM, 1993, has an VSAB of "AIR93"<br><br>Official source names, RSABs, and VSABs are included on the Source Vocabularies page. |
| ATV | Attribute value described under specific attribute name on the Attributes Names page. A few attribute values exceed 1,000 characters. Many of the abbreviations used in attribute values are explained in MRDOC.RRF and included on the Abbreviations Used in Data Elements page. |
| SUPPRESS | Suppressible flag. Values = O, E, Y, or N. Reflects the suppressible status of the attribute; not yet in use. See also SUPPRESS in MRCONSO.RRF and MRDEF.RRFand MRREL.RRF. |
| CVF | Content View Flag. Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use. |

UMLS® Reference Manual

UMLS® Reference Manual

Table 3. Definitions (File = MRDEF.RRF)

| Col. | Description |
|---|---|
| CUI | Unique identifier for concept |
| AUI | Unique identifier for atom - variable length field, 8 or 9 characters |
| ATUI | Unique identifier for attribute |
| SATUI | Source asserted attribute identifier [optional-present if it exists] |
| SAB | Abbreviated source name (SAB) of the source of the definition<br>Maximum field length is 20 alphanumeric characters. Two source abbreviations are assigned:<br>• Root Source Abbreviation (RSAB) — short form, no version information, for example, AI/RHEUM, 1993, has an RSAB of "AIR"<br>• Versioned Source Abbreviation (VSAB) — includes version information, for example, AI/RHEUM, 1993, has an VSAB of "AIR93"<br><br>Official source names, RSABs, and VSABs are included on the Source Vocabularies page. |
| DEF | Definition |
| SUPPRESS | Suppressible flag. Values = O, E, Y, or N. Reflects the suppressible status of the attribute; not yet in use. See also SUPPRESS in MRCONSO.RRF and MRDEF.RRFand MRREL.RRF. |
| CVF | Content View Flag. Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use. |

UMLS® Reference Manual

UMLS® Reference Manual

Table 4. History (File = MRHIST.RRF)

| Col. | Description |
|---|---|
| CUI | Unique identifier for concept |
| SOURCEUI | Source asserted unique identifier |
| SAB | Abbreviated source name (SAB). Maximum field length is 20 alphanumeric characters. Two source abbreviations are assigned:<br>• Root Source Abbreviation (RSAB) — short form, no version information, for example, AI/RHEUM, 1993, has an RSAB of "AIR"<br>• Versioned Source Abbreviation (VSAB) — includes version information, for example, AI/RHEUM, 1993, has an VSAB of "AIR93"<br><br>Official source names, RSABs, and VSABs are included on the Source Vocabularies page. |
| SVER | Release date or version number of a source |
| CHANGETYPE | Source asserted code for type of change |
| CHANGEKEY | CONCEPTSTATUS (if history relates to a SNOMED CT concept) or DESCRIPTIONSTATUS (if history relates to a SNOMED CT atom) |
| CHANGEVAL | CONCEPTSTATUS value or DESCRIPTIONSTATUS value after the change took place. Note: The change may have affected something other than the status value. |
| REASON | Explanation of change if present |
| CVF | Content View Flag. Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use. |

Table 5. Related Concepts (File = MRREL.RRF)

| Col. | Description |
|---|---|
| CUI1 | Unique identifier of first concept |
| AUI1 | Unique identifier of first atom |
| STYPE1 | The name of the column in MRCONSO.RRF that contains the identifier used for the first concept or first atom in source of the relationship. |
| REL | Relationship of second concept or atom to first concept or atom |
| CUI2 | Unique identifier of second concept |
| AUI2 | Unique identifier of second atom |
| STYPE2 | The name of the column in MRCONSO.RRF that contains the identifier used for the second concept or second atom in the source of the relationship. |
| RELA | Additional (more specific) relationship label (optional) |
| RUI | Unique identifier of relationship |
| SRUI | Source asserted relationship identifier, if present |
| SAB | Abbreviated source name of the source of relationship. Maximum field length is 20 alphanumeric characters. Two source abbreviations are assigned:<br>• Root Source Abbreviation (RSAB) — short form, no version information, for example, AI/RHEUM, 1993, has an RSAB of "AIR"<br>• Versioned Source Abbreviation (VSAB) — includes version information, for example, AI/RHEUM, 1993, has an VSAB of "AIR93"<br><br>Official source names, RSABs, and VSABs are included on the Source Vocabularies page. |
| SL | Source of relationship labels |
| RG | Relationship group. Used to indicate that a set of relationships should be looked at in conjunction. |
| DIR | Source asserted directionality flag. Y indicates that this is the direction of the relationship in its source; N indicates that it is not; a blank indicates that it is not important or has not yet been determined. |
| SUPPRESS | Suppressible flag. Values = O, Y, E, or N. Reflects the suppressible status of the relationship; not yet in use. See also SUPPRESS in MRCONSO.RRF and MRDEF.RRF and MRREL.RRF. |
| CVF | Content View Flag. Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use. |

Table 6. Co-occurring Concepts (File = MRCOC.RRF)

| Col. | Description |
|---|---|
| CUI1 | Unique identifier of first concept |
| AUI1 | Unique identifier of first atom |
| CUI2 | Unique identifier of second concept or not present<br>Note: Where CUI2 is not present and COT is LQ (MeSh topical qualifier), the count of citations of CUI1 with no MeSH qualifiers is reported in COF. |
| AUI2 | Unique identifier of second atom |
| SAB | Abbreviation of the source of co-occurrence information |
| COT | Type of co-occurrence |
| COF | Frequency of co-occurrence, if applicable |
| COA | Attributes of co-occurrence, if applicable |
| CVF | Content View Flag. Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use. |

UMLS® Reference Manual

UMLS® Reference Manual

Table 7. Computable Hierarchies (File = MRHIER.RRF)

| Col. | Description |
|---|---|
| CUI | Unique identifier of concept |
| AUI | Unique identifier of atom - variable length field, 8 or 9 characters |
| CXN | Context number (e.g., 1, 2, 3) |
| PAUI | Unique identifier of atom's immediate parent within this context |
| SAB | Abbreviated source name (SAB) of the source of atom (and therefore of hierarchical context). Maximum field length is 20 alphanumeric characters. Two source abbreviations are assigned:<br>• Root Source Abbreviation (RSAB) — short form, no version information, for example, AI/RHEUM, 1993, has an RSAB of "AIR"<br>• Versioned Source Abbreviation (VSAB) — includes version information, for example, AI/RHEUM, 1993, has an VSAB of "AIR93"<br><br>Official source names, RSABs, and VSABs are included on the Source Vocabularies page. |
| RELA | Relationship of atom to its immediate parent |
| PTR | Path to the top or root of the hierarchical context from this atom, represented as a list of AUIs, separated by periods (.) The first one in the list is top of the hierarchy; the last one in the list is the immediate parent of the atom, which also appears as the value of PAUI. |
| HCD | Source asserted hierarchical number or code for this atom in this context; this field is only populated when it is different from the code (unique identifier or code for the string in that source). |
| CVF | Content View Flag. Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use. |

Table 8. Contexts (File = MRCXT.RRF)

| Col. | Description |
|---|---|
| CUI | Unique identifier of concept |
| SUI | Unique identifier of string used in this context |
| AUI | Unique identifier of atom that has this context (variable length field, 8 or 9 characters) |
| SAB | Abbreviated source name (SAB). Maximum field length is 20 alphanumeric characters. Two source abbreviations are assigned:<br>• Root Source Abbreviation (RSAB) — short form, no version information, for example, AI/RHEUM, 1993, has an RSAB of "AIR"<br>• Versioned Source Abbreviation (VSAB) — includes version information, for example, AI/RHEUM, 1993, has an VSAB of "AIR93"<br><br>Official source names, RSABs, and VSABs are included on the Source Vocabularies page. |
| CODE | Unique identifier or code for string in that source |
| CXN | The context number (if the atom has multiple contexts) |
| CXL | Context member label, i.e., ANC for ancestor of this atom, CCP for the atom itself, SIB for sibling of this atom, CHD for child of this atom |
| RNK | For rows with a CXL value of ANC, the rank of the ancestors (e.g., a value of 1 denotes the most remote ancestor in the hierarchy) |
| CXS | String or concept name for context member |
| CUI2 | Concept identifier of context member (may be empty if context member is not yet in the Metathesaurus) |
| AUI2 | Atom identifier of context member |
| HCD | Source hierarchical number or code of context member (if present) |
| RELA | Additional relationship label providing further categorization of the CXL, if applicable and known. Valid values listed on the Abbreviations Used in Data Elements page. |
| XC | A plus (+) sign indicates that the CUI2 for this row has children in this context. If this field is empty, the CUI2 does not have children in this context. |
| CVF | Content View Flag. Bit field used to flag rows included in Content View. |

Table 9. Mappings (File = MRMAP.RRF)

| Col. | Description |
|---|---|
| MAPSETCUI | Unique identifier for the UMLS concept which represents the whole map set. |
| MAPSETSAB | Source abbreviation (SAB) for the provider of the map set. |
| MAPSUBSETID | Map subset identifier used to identify a subset of related mappings within a map set. This is used for cases where the FROMEXPR may have more than one potential mapping (optional). |
| MAPRANK | Order in which mappings in a subset should be applied. Used only where MAPSUBSETID is used. (optional) |
| MAPID | Unique identifier for this individual mapping. Primary key of this table to identify a particular row. |
| MAPSID | Source asserted identifier for this mapping (optional). |
| FROMID | Identifier for the entity being mapped from. This is an internal UMLS identifier used to point to an external entity in a source vocabulary (represented by the FROMEXPR). When the source provides such an identifier, it is reused here. Otherwise, it is generated by NLM. The FROMID is only unique within a map set. It is not a pointer to UMLS entities like atoms or concepts. There is a one-to-one correlation between FROMID and a unique set of values in FROMSID, FROMEXPR, FROMTYPE, FROMRULE, and FROMRES within a map set. |
| FROMSID | Source asserted identifier for the entity being mapped from (optional). |
| FROMEXPR | Entity being mapped from - can be a single code/identifier /concept name or a complex expression involving multiple codes/ identifiers/concept names, Boolean operators and/or punctuation |
| FROMTYPE | Type of entity being mapped from. |
| FROMRULE | Machine processable rule applicable to the entity being mapped from (optional) |
| FROMRES | Restriction applicable to the entity being mapped from (optional). |
| REL | Relationship of the entity being mapped from to the entity being mapped to. |
| RELA | Additional relationship label (optional). |
| TOID | Identifier for the entity being mapped to. This is an internal identifier used to point to an external entity in a source vocabulary (represented by the TOEXPR). When the source provides such an identifier, it is reused here. Otherwise, it is generated by NLM. The TOID is only unique within a map set. It is not a pointer to UMLS entities like atoms or concepts. There is a one-to-one correlation between TOID and a unique set of values in TOSID, TOEXPR, TOTYPE, TORULE, TORES within a map set. |
| TOSID | Source asserted identifier for the entity being mapped to (optional). |
| TOEXPR | Entity being mapped to - can be a single code/identifier/concept name or a complex expression involving multiple codes/ identifiers/concept names, Boolean operators and/or punctuation. |
| TOTYPE | Type of entity being mapped to. |
| TORULE | Machine processable rule applicable to the entity being mapped to (optional). |
| TORES | Restriction applicable to the entity being mapped to (optional). |
| MAPRULE | Machine processable rule applicable to this mapping (optional). |
| MAPRES | Restriction applicable to this mapping (optional). |
| MAPTYPE | Type of mapping (optional). |
| MAPATN | The name of the attribute associated with this mapping [not yet in use] |
| MAPATV | The value of the attribute associated with this mapping [not yet in use] |
| CVF | The Content View Flag is a bit field used to indicate membership in a content view. |

UMLS® Reference Manual

UMLS® Reference Manual

Table 10. Simple Mappings (File = MRSMAP.RRF)

| Col. | Description |
|---|---|
| MAPSETCUI | Unique identifier for the UMLS concept which represents the whole map set. |
| MAPSETSAB | Source abbreviation for the map set. |
| MAPID | Unique identifier for this individual mapping. Primary key of this table to identify a particular row. |
| MAPSID | Source asserted identifier for this mapping (optional). |
| FROMEXPR | Entity being mapped from - can be a single code/identifier/concept name or a complex expression involving multiple codes/ identifiers/concept names, Boolean operators and/or punctuation. |
| FROMTYPE | Type of entity being mapped from. |
| REL | Relationship of the entity being mapped from to the entity being mapped to. |
| RELA | Additional relationship label (optional). |
| TOEXPR | Entity being mapped to - can be a single code/identifier /concept name or a complex expression involving multiple codes/ identifiers/concept names, Boolean operators and/or punctuation. |
| TOTYPE | Type of entity being mapped to. |
| CVF | The Content View Flag is a bit field used to indicate membership in a content view. |

Table 11. Source Information (File = MRSAB.RRF)

| Field | Full Name | Description |
|---|---|---|
| VCUI | CUI | CUI of the versioned SRC concept for a source |
| RCUI | Root CUI | CUI of the root SRC concept for a source |
| VSAB | Versioned Source Abbreviation | The versioned source abbreviation for a source, e.g., MSH2003_2002_10_24 |
| RSAB | Root Source Abbreviation | The root source abbreviation for a source e.g., MSH |
| SON | Official Name | The official name for a source |
| SF | Source Family | The source family for a source |
| SVER | Version | The source version, e.g., 2001 |
| VSTART | Meta Start Date | The date a source became active, e.g., 2001_04_03 |
| VEND | Meta End Date | The date a source ceased to be active, e.g., 2001_05_10 |
| IMETA | Meta Insert Version | The version of the Metathesaurus in which a source first appeared, e.g., 2001AB |
| RMETA | Meta Remove Version | The version of the Metathesaurus in which the source last appeared, e.g., 2001AC |
| SLC | Source License Contact | The source license contact information |
| SCC | Source Content Contact | The source content contact information |
| SRL | Source Restriction Level | 0, 1, 2, 3, 4 - explained in the License Agreement |
| TFR | Term Frequency | The number of terms for this source in MRCONSO.RRF, e.g., 12343 |
| CFR | CUI Frequency | The number of CUIs associated with this source, e.g., 10234 |
| CXTY | Context Type | The type of contexts for this source. Values are FULL, FULL-MULTIPLE, FULL-NOSIB, FULL-NOSIB-MULTIPLE, FULL-MULTIPLE-NOSIB-RELA, null. |
| TTYL | Term Type List | Term type list from source, e.g., MH, EN, PM, TQ |
| ATNL | Attribute Name List | The attribute name list (from MRSAT.RRF), e.g., MUI, RN, TH |
| LAT | Language | The language of the terms in the source |
| CENC | Character Encoding | All UMLS content is provided in Unicode, encoded in UTF-8.<br>MetamorphoSys will allow exclusion of extended characters with some loss of information.<br>Transliteration to other character encodings is possible but not supported buy NLM; for further information, see http://www.unicode.org |
| CURVER | Current Version | A Y or N flag indicating whether or not this row corresponds to the current version of the named source |
| SABIN | Source in Subset | A Y or N flag indicating whether or not this row is represented in the current MetamorphoSys subset. Initially always Y where CURVER is Y, but later is recomputed by MetamorphoSys. |
| SSN | Source Short Name | The short name of a source as used by the NLM Knowledge Source Server |
| SCIT | Source Citation | Citation information for a source. This is intended to replace the SOS attributes in the SRC concepts. |

UMLS® Reference Manual

UMLS® Reference Manual

Table 12. Concept Name Ranking (File = MRRANK.RRF)

| Col. | Description |
|---|---|
| RANK | Numeric order of precedence, higher value wins |
| SAB | Abbreviated source name (SAB) for source vocabulary. Maximum field length is 20 alphanumeric characters. Two source abbreviations are assigned:<br>• Root Source Abbreviation (RSAB) — short form, no version information, for example, AI/RHEUM, 1993, has an RSAB of "AIR"<br>• Versioned Source Abbreviation (VSAB) — includes version information, for example, AI/RHEUM, 1993, has an VSAB of "AIR93"<br><br>Official source names, RSABs, and VSABs are included on the Source Vocabularies page. |
| TTY | Abbreviation for term type in source vocabulary, for example PN (Metathesaurus Preferred Name) or CD (Clinical Drug). Possible values are listed in Abbreviations Used in Data Elements page. |
| SUPPRESS | NLM-recommended Source and Term Type (SAB/TTY) Suppressiblity. Values = Y or N. Indicates the suppressible status of all atoms (names) with this Source and Term Type (SAB/TTY). Note that changes made in MetamorphoSys at the Suppressible tab are recorded in your configuration file. Status E does not occur here, as it is assigned only to individual cases such as the names (atoms) in MRCONSO.RRF. See also SUPPRESS in MRCONSO.RRF, MRDEF.RRF, and MRREL.RRF. |

Table 13. Retired CUI Mapping (File = MRCUI.RRF)

| Col. | Description |
|---|---|
| CUI1 | Unique identifier for first concept - Retired CUI - was present in some prior release, but is currently missing |
| VER | The last release version in which CUI1 was a valid CUI |
| REL | Relationship |
| RELA | Relationship attribute |
| MAPREASON | Reason for mapping |
| CUI2 | Unique identifier for second concept - the current CUI that CUI1 most closely maps to |
| MAPIN | Is this map in current subset? Values of Y, N, or null. MetamorphoSys generates the Y or N to indicate whether the CUI2 concept is or is not present in the subset. The null value is for rows where the CUI1 was not present to begin with (i.e., REL=DEL). |

Table 14. AUI Movements (File = MRAUI.RRF)

| Col. | Description |
|---|---|
| AUI1 | Atom unique identifier |
| CUI1 | Concept unique identifier |
| VER | Version in which this change to the AUI first occurred |
| REL | Relationship |
| RELA | Relationship attribute |
| MAPREASON | Reason for mapping |
| AUI2 | Unique identifier for second atom |
| CUI2 | Unique identifier for second concept - the current CUI that CUI1 most closely maps to |
| MAPIN | Mapping in current subset: is AUI2 in current subset? Values of Y, N, or null. |