

Instalação, configuração e teste

MetaMap DFB 2022

Tester: Perciliano

Date: 29/03/2023 20:27h a 30/03/2023 00:47h

Sumário

1	INSTALAÇÃO METAMAP	3
2	DADOS DO TEST PARA CRIAÇÃO DE BASES EM PORTUGUÊS DO BRASIL	3
3	SÓ EM CASO DE INSUCESSO EXECUTAR ESTE TÓPICO	3
4	PASSO A PASSO PARA CRIAR OS ARQUIVOS DE TRABALHO	3
5	EXECUTANDO OS 5 ARQUIVOS DO \01METAWORDINDEX	4
5.1	ARQUIVO ./01CREATEWORKFILES	4
5.2	ARQUIVO ./02SUPPRESS	5
5.3	ACESSAR O ./03FILTERPREP	6
5.4	ACESSAR O ./04FILTERSTRICT.	7
5.5	ACESSAR O ./05GENERATEMWIFILES	8
6	EXECUTANDO ARQUIVOS DE OUTROS DIRETÓRIOS	8
7	ARQUIVO LOADDATAFILES	9
8	DETALHES DOS ARQUIVOS DE TESTE	11
9	TESTE DO MODELO EM PT-BR	11
10	REFERÊNCIAS	11

1 Instalação MetaMap

Nova instalação completa do MetaMap feita em 19-03-2023, note Linux Perci. [1 - versão no github]

2 Passo a passo para criar os arquivos de trabalho

Para evitar ficar colocando dados no terminal, ajustar o ./bin/BuildDataFiles

Alterar as linhas abaixo para:

- 44 = ui_input="PTBRbase" (nome da base customizada)
- 54 = response="yes" (confirmar que a base acima será usada)
- 83 = DB_SELECT="" (confirmar que o DB padrão USABase será usado)
- 133 = LEX_SELECT="" (confirmar que será usado o léxico ???)
- 159 = RES_SELECT="" (confirmar dfb padrão)
- 172 = ui_input="2022" (qual o ano da base UMLS usada)
- 194 = ui_input="AB" (qual a release UMLS usada)

Para geração de modelos sem precisar nova instalação, esse arquivo as vezes não pode ser ajustado ou tem que ser ajustado em conformidade com as bases baixadas, logo o arquivo original e as linhas eram:

- 44 = read ui_input
- 54 = read response
- 83 = read DB_SELECT
- 133 =read LEX_SELECT
- 159 =read RES_SELECT
- 172 = read ui_input
- 194 = read ui_input

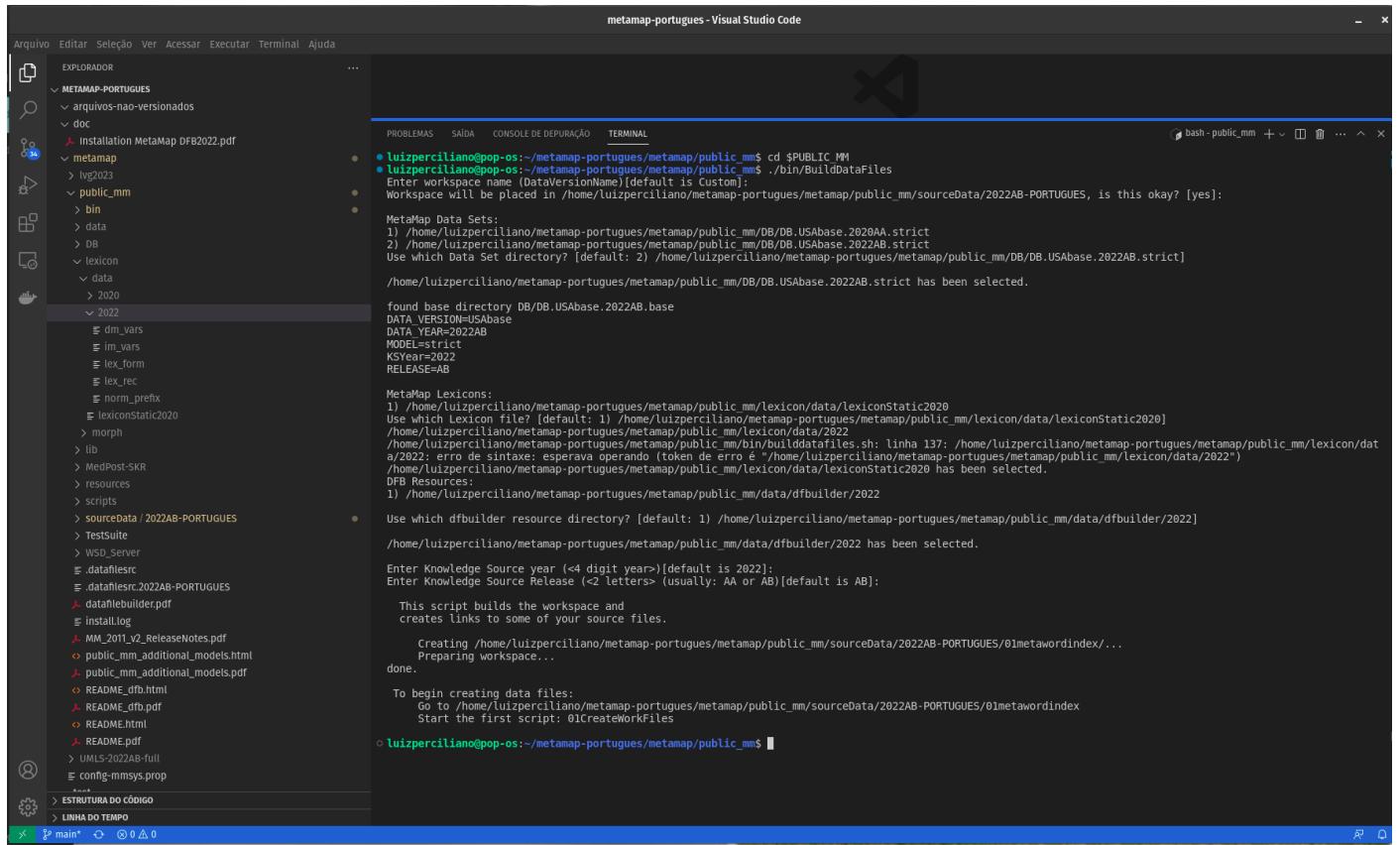
```
cd $PUBLIC_MM
./bin/BuildDataFiles
UMLS PORTUGUES

/home/luizperciliano/metamap-portugues/metamap/public_mm/DB/DB.USAbase
.2022AB.strict

/home/luizperciliano/metamap-portugues/metamap/public_mm/lexicon/data/
2022

...dfbuilder/2022

2022
AB
```



Iniciar configuração da [planilha de teste](#)

Problema no arquivo `./bin/BuildDataFiles`, linha 91 qd coloca a variável `UMLS_portugues`. <Mas criou os diretórios. Esse ocorreu no docker com o Malah.

3 Executando os 5 arquivos do \01metawordindex

```
cd $PUBLIC_MM/sourceData/2022AB-PORTUGUES/01metawordindex
nano 01CreateWorkFiles
```

No arquivo acima, escolher um dos ajustes abaixo:

- V1 – gawk em “POR” - Ajustar a [linha 45](#) de: `gawk -F"\t" $2 == "ENG" \ para: gawk -F"\t" $2 == "POR" \`

```
gawk -F'\t' '$2 == "POR" \
&& ! ( $12 == "NEU" \
&& ! ( $13 == "ACR" || $13 == "LATSY" || $13 == "PT" || $13 == "SY" ) ) \
{ print $0 }' MRCNSO.RRF > mrconso.ALLENG
```

```
gawk -F"\t" '$2 == "POR" \
&& ! ( $12 == "NEU" \
&& ! ( $13 == "ACR" || $13 == "LATSY" || $13 == "PT" || $13 == "SY" ) ) \
{ print $0 }' MRCNSO.RRF > mrconso.ALLENG
```

- V2 ~~grep E e '^[A-Za-zA-Z]+[0-9][0-9]*\|POR\|ENG\|' MRCNSO.RRF \~~

```

grep -E -e '^[A-Za-z\_]+'[0-9][0-9]*\|POR|ENG\|' MRCNSO.RRF \
| sort -t '\|' -k1,1 -k3,4 -k5,6 -k7,7r -k12,14 \
| sort -t '\|' -u -k1,1 -k3,4 -k5,6 -k12,14 \
> mrconso.ALLENG

```

- ~~V3 grep -E -e '^[A-Za-z_]+'[0-9][0-9]*\|POR|' MRCNSO.RRF \~~. Trocar da linha 45 a 48 pelo código abaixo.

```

grep -E -e '^[A-Za-z\_]+'[0-9][0-9]*\|POR|' MRCNSO.RRF \
| sort -t '\|' -k1,1 -k3,4 -k5,6 -k7,7r -k12,14 \
| sort -t '\|' -u -k1,1 -k3,4 -k5,6 -k12,14 \
> mrconso.ALLENG

```

```

36 # /bin/rm -f mrconso.* 
37 echo "creating and sorting mrconso.ALLENG"
38
39 # After investigations and e-mail discussions with Steve Emrick and Brian Carlsen
40 # on 05/30/2013 and 05/31/2013:
41 # We want to keep MRCNSO all rows in which
42 # (1) LAT == "ENG", and
43 # (2) not ( SAB == "NEU" and TTY not one of ( "ACR", "LATSY", "PT", or "SY" ) )
44
45 grep -E -e '^[A-Za-z\_]+'[0-9][0-9]*\|POR|' MRCNSO.RRF \
46 | sort -t '\|' -k1,1 -k3,4 -k5,6 -k7,7r -k12,14 \
47 | sort -t '\|' -u -k1,1 -k3,4 -k5,6 -k12,14 \
48 > mrconso.ALLENG
49
50 # chmod -w mrconso.ALLENG
51 echo "DONE"
52
53 # ensure that there are no non-ascii lines in mrconso
54 # echo "excluding all non-ASCII lines from mrconso.ALLENG"

```

```

$ 01CreateWorkFiles
metamap > public_mm > sourceData > 202AB-PORTUGUES > 01metawordindex > $ 01CreateWorkFiles
17 # create links to the necessary MR+ files
18 # echo 'Linking to Metathesaurus files...
19 # ln -s ../../umls/MRCONSO.RRF
20 # ln -s ../../umls/MRSAT.RRF
21 # ln -s ../../umls/MRSTY.RRF
22 # echo 'Finished creating links.'
23
24 # collect frequency data about each of the fields
25 /bin/rm -rf MRCONSO.*.values

luisperciliano@pop-os:~/metamap-portugues/metamap/public_mm$ cd $PUBLIC_MM/sourceData/UMLS-PORTUGUES/01metawordindex
bash: cd: /home/luisperciliano/metamap-portugues/metamap/public_mm/sourceData/UMLS-PORTUGUES/01metawordindex: Arquivo ou diretório inexistente
luisperciliano@pop-os:~/metamap-portugues/metamap/public_mm$ cd $PUBLIC_MM/sourceData/UMLS-PORTUGUES/01metawordindex\c
ambin 01CreateWorkFiles
luisperciliano@pop-os:~/metamap-portugues/metamap/public_mm$ cd $PUBLIC_MM/sourceData/202AB-PORTUGUES/01metawordindex\dex
luisperciliano@pop-os:~/metamap-portugues/metamap/public_mm/sourceData/202AB-PORTUGUES/01metawordindex$ cd $PUBLIC_MM/sourceData/202AB-PORTUGUES/01metawordindex$ ./01CreateWorkFiles
Wed Mar 29 22:24:10 -03 2023
creating MRCONSO.CUI.values
creating MRCONSO.LAT.values
creating MRCONSO.TS.values
creating MRCONSO.UUI.values
creating MRCONSO.SUI.values
creating MRCONSO.SAB.values
creating MRCONSO.TTY.values
creating MRCONSO.SRL.values
Wed Mar 29 22:24:10 -03 2023
    Create mrconso.eng.0 (without suppression)
    creating and sorting mrconso.ALLENG
DONE
moving mrconso.ALLENG to mrconso.eng.0
Checking PPF/Y CUIs
creating mrconso.eng.ALL_CUI
creating mrconso.eng.PPF_Y.CUI
doing counts
PPF/Y CUIs are unique.
PPF/Y CUIs cover all mrconso.eng.0
OK, continuing.
creating mrconso.eng.TS.values
DONE
Wed Mar 29 22:24:11 -03 2023
Do preprocessing to suppress
numbers (0, 1, +1, 0.5, 50000)
single letters (a, A, b, B, ...)
bloopers (0->TPB protein, and Periods->Menstruation)
unnecessary ambiguity (Other <n>, Protocols <n>, ...)
Run 02Suppress.

luisperciliano@pop-os:~/metamap-portugues/metamap/public_mm/sourceData/202AB-PORTUGUES/01metawordindex$ 

```

3.1 Arquivo ./02Suppress

```
./02Suppress
```

Demorou menos de 1 minuto e não 4 como esperado. Certifique-se de na pasta Suppress, os arquivos mrconso.eng.0 e mrconso.suppressed correspondam ao MRCONSO.RRF. Deve conter registros com o campo LAT (idioma) com strings “POR” e não “ENG” (a menos que você esteja incluindo registros em inglês).

O resultado aqui segundo o Will deveria ser total de xxxx

Arquivo Editar Seleção Ver Acessar Executar Terminal Ajuda

EXPLORADOR

- METAMAP-PORTUGUES
 - arquivos-nao-versionados
 - d c
 - Installation MetaMap DFB2022.pdf
 - metamap
 - lgv2023
 - public_mm
 - bin
 - data
 - DB
 - lexicon
 - lib
 - MedPost-SKR
 - resources
 - scripts
 - sourceData / 2022AB-PORTUGUES
 - 01metawordindex
 - Filter
 - Generate
 - Suppress
 - 01CreateWorkFiles
 - 02suppress
 - 03FilterPrep
 - 04FilterRelaxed
 - 04FilterShared
 - 04FilterStrict
 - 04FilterStrictCheck
 - 04FilterStrictMore
 - 05GenerateMRFfiles
 - allocate_lines
 - chunk_nle
 - consolidate_info_lines
 - create_mrrank
 - filter_mrconsom_ALL
 - MRCONSO.CUI.values
 - mrconsom.ENG
 - mrconsom.ENG.0
 - mrconsom.ENG.ALL.CUI
 - mrconsom.ENG.PPR_Y.CUI
 - mrconsom.ENG.TS.values
 - MRCONSO.LAT.values
- ESTRUTURA DO CÓDIGO
- LINHA DO TEMPO

\$ 01CreateWorkFiles ×

```
metamap > public_mm > sourceData > 2022AB-PORTUGUES > 01metawordindex > $ 01CreateWorkFiles
17 # create links to the necessary MR* files
PROBLEMAS SAIDA CONSOLE DE DEPURA O TERMINAL
INTERMEDIATE results are found in Suppress/
This script will take about 4 minutes.

Mark suppressibles in mrcon.eng
loading special suppressible term list: specialterms.txt
41 special terms added to suppressibles term list.
Processing mrconso.eng.0 by suppressibles.txt -> mrconso.suppressed and suppressions.mrconso.
Now suppressible:C0010825 [POR|s]L4876439 [PF|S5563945 |Y|A32237216 |||10061806|MDRBPO|OL|10009701|CMV|3|0|||
Now suppressible:C0013853 [POR|s]L4884236 [PF|S5571743 |Y|A32125737 |||10014449|MDRBPO|OL|10014828|ENG|3|0|||
Now suppressible:C0026916 [POR|s]L16570555 [PF|S200952420 |Y|A32176499 |||10070407|MDRBPO|OL|10025372|MAC|3|0|||
Now suppressible:C0030519 [POR|s]L16614619 [PF|S200958711 |Y|A32215114 |||10003568|MDRBPO|OL|10034114|PA|3|0|||
Now suppressible:C0031304 [POR|s]L16614619 [PF|S200958711 |Y|A32215114 |||10003568|MDRBPO|OL|10034114|PA|3|0|||
Now suppressible:C0031306 [POR|s]L165717354 [PF|S20098701 |Y|A32129839 |||10042602|MDRBPO|OL|10033350|PA|3|0|||
Now suppressible:C0201539 [POR|s]L0425264 [PF|S0549946 |Y|A32137947 |||10001772|MDRBPO|OL|10001454|API|3|0|||
Now suppressible:C0201836 [POR|s]L4869994 [PF|S5557591 |Y|A32125699 |||10001546|MDRBPO|OL|10001844|ALT|3|0|||
Now suppressible:C0201925 [POR|s]L4876456 [PF|S5556595 |Y|A32144137 |||10005392|MDRBPO|OL|10006656|Ca|3|0|||
Now suppressible:C0382358 [POR|p]L3227051 [PF|S3855485 |Y|A32144231 |||10054178|MDRBPO|OL|10040545|Shigella dysenteriae|3|0|||
Now suppressible:C0382359 [POR|p]L3227052 [PF|S3855486 |Y|A32199034 |||10054178|MDRBPO|OL|10040547|Shigella flexneri|3|0|||
Now suppressible:C0382360 [POR|p]L3227053 [PF|S3855488 |Y|A32249528 |||10054178|MDRBPO|OL|10040543|Shigella boydii|3|0|||
Now suppressible:C0382361 [POR|p]L3227053 [PF|S3855488 |Y|A32249528 |||10054178|MDRBPO|OL|10040551|Shigella sonnei|3|0|||
Now suppressible:C0486222 [POR|s]L12216542 [PF|S15404627 |Y|A30161111 |1064-5||LNC-PT-BR|00SN|1064-5|I|NOS Ag RBC BPU QL|0|0|0|||
Now suppressible:C0486223 [POR|s]L2411093 [PF|S15404628 |Y|A30182778 |1065-2||LNC-PT-BR|00SN|1065-2|I|NOS Ag RBC Donr QL|0|0|0|||
Now suppressible:C0525032 [POR|s]L4875491 [PF|S5562996 |Y|A32172414 |||10022591|MDRBPO|OL|10004190|BCR|3|0|||
Now suppressible:C1956346 [POR|s]L16567243 [PF|S20066200 |Y|A32207000 |||10011078|MDRBPO|OL|10006896|CAD|3|0|||
Suppression counts:
  Unnecessarily ambiguous: 261
    NEC: 01
    Brand Name: 00
    C: 00
    Single Letters: 02
    Numerics: 00
    Special Cases: 01
    Total: 265

Warning: Some suppressible terms considered: 18

A link is created to the just-created mrconso.eng in 01metawordindex/
A link is created to the just-created mrcon.eng in 01metawordindex/
The next step is to run 03FilterPrep.
```

luisperciliano@pop-os:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/01metawordindex\$

Lin 45, Col 24 Espa os: 3 UTF-8 LF Shell Script A? D

3.2 Acessar o ./03FilterPrep

Versões do arquivo a usar:

03FilterPrep_V1 – sem ajuste

03FilterPrep_V2 – sem as linhas da 101 a 111

```
# gzip mrso.lbls.d

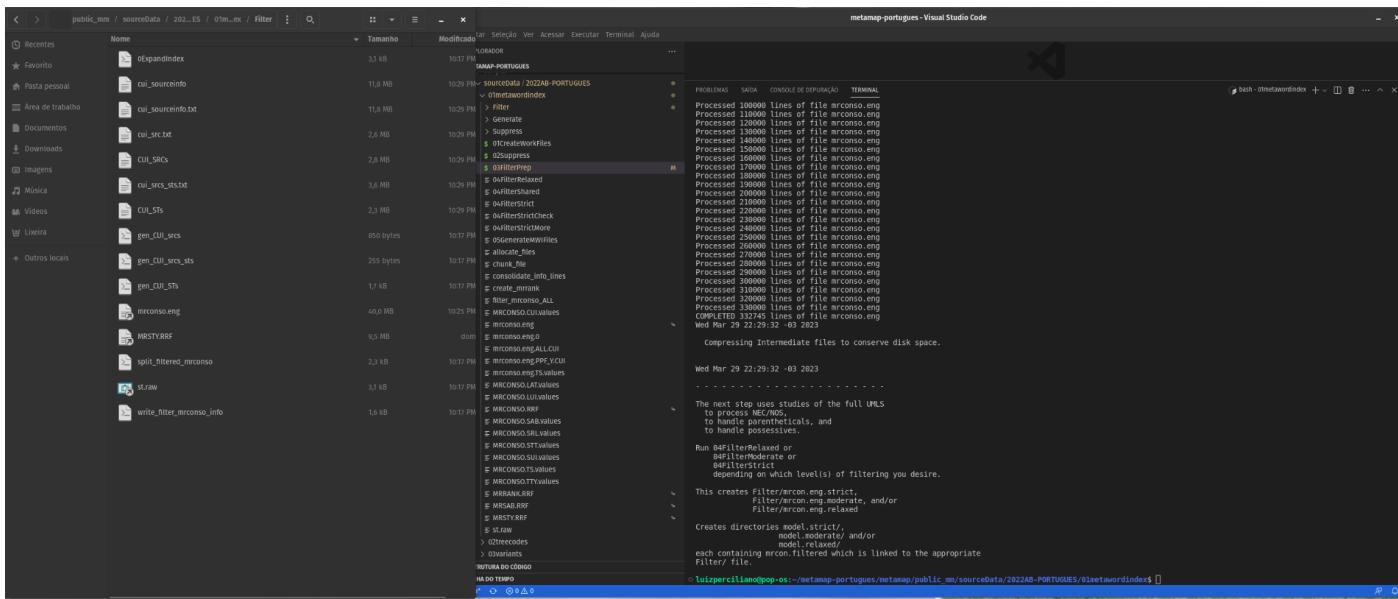
case $OSTYPE in
  MINGW*)
    cp Filter/mrconso.eng mrconso.eng
    ;;
  CYGWIN*)
    cp Filter/mrconso.eng mrconso.eng
    ;;
  *)
    ln -s Filter/mrconso.eng mrconso.eng
    ;;
esac

date

cd ../
```

```
nano 03FilterPrep  
./03FilterPrep
```

OBS.: Na v2, não gera o erro de criar o link, porém, diz que esse processo leva 9 minutos, contudo não levou nem 1 minuto.



3.3 Acessar o ./04FilterStrict.

Se estiver usando 04FilterStrict, certifique-se de que o tagger SKR / MedPOST esteja rodando. Acho que este serviço não fez diferença, executei sem ela e foi de boa. Script 04 alterado, para não precisar imputar

```
cd $PUBLIC_MM/  
./bin/skrmedpostctl start  
cd $PUBLIC_MM/sourceData/2022AB-PORTUGUES/01metawordindex  
./04FilterStrict  
[yes/no]  
Enter (9 processadores)
```

Esta execução demora +- 0:30h. Porém demorou bem menos, aproximadamente **07 minutos** e gerou varios erros, contudo, rodou com x processadores.

Iniciou +- 22:33 até 00:15

Erros de português

EXPLORADOR

METAMAP-PORTUGUES

- sourceData / 202AB-PORTUGUES
- 01metawordindex
- > Filter
- > Generate
- > Suppress
- \$ 01CreateWorkFiles
- \$ 02Suppress
- \$ 03Filterrep
- M 04FilterRelaxed
- 04Filtershared
- 04FilterStrict
- 04FilterStrictCheck
- 04FilterStrictMore
- 05GenerateMWIFiles
- allocate_nies
- chunk_nie
- consolidate_info_lines
- create_mrank
- filter_mrconso_ALL
- MRCONSO.CUIvalues
- mrconso.ENG
- mrconso.ENG_0
- mrconso.ALL.CUI
- mrconso.ENG_PPF_YCUI
- mrconso.TSvalues
- MRCONSO.LATvalues
- MRCONSO.LUIvalues
- MRCONSO.RRF
- MRCONSO.SAB.values
- MRCONSO.SRLvalues
- MRCONSO.STTvalues
- MRCONSO.USvalues
- MRCONSO.TSvalues
- MRCONSO.TTvalues
- MRRANK.RRF
- MRSAB.BRF
- MRSTY.RRF
- SLRaw
- > 02treecodes
- > 03variants
- > EXTRUTOR DO CÓDIGO
- > LINHA DO TEMPO

PROBLEMAS SAÍDA CONSOLE DE DEPURAÇÃO TERMINAL

Restored "C2706754|POR|P|L12584536|[PF|S15228925|Y|A25444875||54193-8||LNC-PT-BR|LN|54193-8|Cefotaxima+Sulbactam:#N/A:Pt:Isolado:OrdQn:Difuso em agar (disco)|0|"

" previously deleted because of syntax

C2706755 preferred string counts: 1/1

Restored "C2706755|POR|P|L12810199|[PF|S15476899|Y|A25341035||54194-6||LNC-PT-BR|LN|54194-6|Mezlocilina+Sulbactam:#N/A:Pt:Isolado:OrdQn|0|N||" previously deleted because of syntax

C2706756 preferred string counts: 1/1

Restored "C2706756|POR|P|L12428927|[PF|S15476897|Y|A25653806||54195-3||LNC-PT-BR|LN|54195-3|Mezlocilina+Sulbactam:#N/A:Pt:Isolado:OrdQn:Concentrao inibitria mnima|0|N||" previously deleted because of syntax

C2706757 preferred string counts: 1/1

Restored "C2706757|POR|P|L12810198|[PF|S15476898|Y|A25548916||54196-1||LNC-PT-BR|LN|54196-1|Mezlocilina+Sulbactam:#N/A:Pt:Isolado:OrdQn:Difuso em agar (disco)|0|"

" previously deleted because of syntax

C2706758 preferred string counts: 1/1

Restored "C2706758|POR|P|L12428950|[PF|S15548014|Y|A24924900||54197-9||Piperacilina+Sulbactam:#N/A:Pt:Isolado:OrdQn|0|N||" previously deleted because of syntax

C2706759 preferred string counts: 1/1

Restored "C2706759|POR|P|L12334423|[PF|S15548011|Y|A25132558||54198-7||LNC-PT-BR|LN|54198-7|Piperacilina+Sulbactam:#N/A:Pt:Isolado:OrdQn:Concentrao inibitria mnima|0|N||" previously deleted because of syntax

C2706760 preferred string counts: 1/1

Restored "C2706760|POR|P|L20444949|[PF|S15548012|Y|A25028909||54199-5||LNC-PT-BR|LN|54199-5|Piperacilina+Sulbactam:#N/A:Pt:Isolado:OrdQn:Difuso em agar (disco)|0|"

" previously deleted because of syntax

C2706761 preferred string counts: 1/1

Restored "C2706761|POR|P|L12503639|[PF|S15365228|Y|A25444878||54200-1||LNC-PT-BR|LN|54200-1|Griseofulvina:#N/A:Pt:Isolado:OrdQn|0|N||" previously deleted because of syntax

C2706762 preferred string counts: 1/1

Restored "C2706762|POR|P|L12466224|[PF|S15365226|Y|A25548920||54201-9||LNC-PT-BR|LN|54201-9|Griseofulvina:#N/A:Pt:Isolado:OrdQn:Concentrao inibitria mnima|0|N||"

" previously deleted because of syntax

C2706763 preferred string counts: 1/1

Restored "C2706763|POR|P|L12698858|[PF|S15365227|Y|A25236734||54202-7||LNC-PT-BR|LN|54202-7|Griseofulvina:#N/A:Pt:Isolado:OrdQn:Difuso em agar (disco)|0|N||" previously deleted because of syntax

C2706764 preferred string counts: 1/1

Restored "C2706764|POR|P|L12191008|[PF|S15228611|Y|A25132561||54203-5||LNC-PT-BR|LN|54203-5|Cefmenoxime:#N/A:Pt:Isolado:OrdQn:Concentrao inibitria mnima|0|N||" previously deleted because of syntax

C2706765 preferred string counts: 1/1

Restored "C2706765|POR|P|L12831122|[PF|S15644636|Y|A25653807||54204-3||LNC-PT-BR|LN|54204-3|Tiacetazona:#N/A:Pt:Isolado:OrdQn:Concentrao inibitria mnima|0|N||" previously deleted because of syntax

C2706766 preferred string counts: 1/1

Restored "C2706766|POR|P|L12318909|[PF|S15445900|Y|A24924904||54205-0||Legionella pneumophila 7+8+9+10+11+12+13+14 Ac:#N/A:Pt:Soro:On|0|N||" previously deleted because of syntax

C2706767 preferred string counts: 1/1

Restored "C2706767|POR|P|L12785596|[PF|S15279581|Y|A24924906||54206-8||LNC-PT-BR|LN|54206-8|Cytomegalovrus DNA:#N/A:Pt:Sor:Plas:Qn:Sonda com limite de detecco da amplificaco de alvo|0|N||" previously deleted because of syntax

C2706768 preferred string counts: 1/1

Restored "C2706768|POR|P|L1243005|[PF|S15638486|Y|A24924908||54207-6||LNC-PT-BR|LN|54207-6|Tempo de Tromboplastina parcial ativada.fator de substituio^imediamente aps adio de plasma deficiente de fator V:Tempo:Pt:PPPO:Qn:Coagulao|0|N||" previously deleted because of syntax

C2706769 preferred string counts: 1/1

Restored "C2706769|POR|P|L12927747|[PF|S15638485|Y|A24924910||54208-4||LNC-PT-BR|LN|54208-4|Tempo de Tromboplastina parcial ativada.fator de substituio^imediamente aps adio de plasma deficiente de fator V+A:PC:Tempo:Pt:PPPO:Qn:Coagulao|0|N||" previously deleted because of syntax

C2706770 preferred string counts: 1/1

Restored "C2706770|POR|P|L12638668|[PF|S15383623|Y|A25028910||54209-2||LNC-PT-BR|LN|54209-2|Hemoglobina C-Harlem/Hemoglobina total:#N/A:Pt:SgTotal:On|0|N||" previously deleted because of syntax

C2706772 preferred string counts: 1/1

Restored "C2706772|POR|P|L2664046|[PF|S15385186|Y|A25444886||54210-0||Hepatite B virus core basal promotor de mutao:#N/A:Pt:Soro:Nom:Sonda com limite de detecco da amplificaco de alvo|0|N||" previously deleted because of syntax

C2706773 preferred string counts: 1/1

EXPLORADOR

METAMAP-PORTUGUES

sourceData / 2022B-PORTUGUES

01metawordindex

> Filter

> Generate

> Suppress

\$ 01createWorkFiles

\$ 02Suppress

\$ 03filterprep M

04filterRelaxed

04filterShared

04filterStrict

04filterStrictCheck

04filterStrictMore

05generateWifiles

allocate_files

chunk_file

consolidate_info_lines

create_mrmark

filter_mirconso_ALL

MRCONSO.CUIvalues

mrconso.eng

mrconso.eng.0

mrconso.eng.ALLCUI

mrconso.eng.PPF_YCU

mrconso.eng.TSvalues

MRCONSO.LATvalues

MRCONSO.UVvalues

MRCONSO.RRF

MRCONSO.SABvalues

MRCONSO.SRLvalues

MRCONSO.STTvalues

MRCONSO.SUvalues

MRCONSO.TSvalues

MRCONSO.TTvalues

MRKRRN.RRF

MRSAB.RRF

MRSTY.RRF

s.raw

> 02treecodes

> 03variants

ESTRUTURA DO CÓDIGO

LINHA DO TEMPO

PROBLEMAS SAÍDA CONSOLE DE DEPURAÇÃO TERMINAL

Restored "C2708338|POR|P|L12874807|PF|S15216412|Y|A25132728||55142-4||LNC-PT-BR|LN|55142-4|Cache valley virus RNA:#N/A:Pt:XXX:Ord:Sonda com limite de deteção da amplificação de alvo|0[N]|" previously deleted because of syntax
C2708339 preferred string counts: 1/1

Restored "C2708339|POR|P|L12875173|PF|S15219513|Y|A25132730||55143-2||LNC-PT-BR|LN|55143-2|California virus sorogrupo RNA:#N/A:Pt:XXX:Ord:Sonda com limite de deteção da amplificação de alvo|0[N]|" previously deleted because of syntax
C2708340 preferred string counts: 1/1

Restored "C2708340|POR|P|L12333080|PF|S15558490|Y|A25431240||55144-0||LNC-PT-BR|LN|55144-0|Powassan virus RNA:#N/A:Pt:XXX:Ord:Sonda com limite de deteção da amplificação de alvo|0[N]|" previously deleted because of syntax
C2708341 preferred string counts: 1/1

Restored "C2708341|POR|P|L12923204|PF|S16062422|Y|A25549090||55145-7|Saint Louis encefalite virus RNA:#N/A:Pt:XXX:Ord:Sonda com limite de deteção da amplificação de alvo|0[N]|" previously deleted because of syntax
C2708342 preferred string counts: 1/1

Restored "C2708342|POR|P|L12695221|PF|S15325499|Y|A25445042||55146-5||LNC-PT-BR|LN|55146-5|Familiar longo QTS syndrome analise de mutação do gene:Ident:Pt:Sg/Tecido or:Nar:Genética molecular|0[N]|" previously deleted because of syntax
C2708344 preferred string counts: 1/1

Restored "C2708344|POR|P|L12575188|PF|S15173189|Y|A25029085||55147-3||LNC-PT-BR|LN|55147-3|BCR-ABL1 b2a2+b3a2 transcripto de fuso/AcL 1 transcripto:#N/A:Pt:SgTotal:On:Genética molecular|0[N]|" previously deleted because of syntax
C2708346 preferred string counts: 1/1

DID NOT RESTORE "C2708346|POR|P|L12869535|PF|S15173188|Y|A24925086||55148-1||LNC-PT-BR|LN|55148-1|BCR-ABL1 b2a2+b3a2 transcripto de fuso/AcL 1 transcripto:#N/A:Pt:Sg/Tecido:Ord:genética molecular|0[N]|" because all concepts filtered out
C2708347 preferred string counts: 1/1

Restored "C2708347|POR|P|L12284938|PF|S15173205|Y|A25236928||55149-9||LNC-PT-BR|LN|55149-9|BCR-ABL1 e1a2 transcripto de fuso/AcL 1 transcripto:#N/A:Pt:Sg/Tecido:Ord:genética molecular|0[N]|" previously deleted because of syntax
C2708348 preferred string counts: 1/1

Restored "C2708348|POR|P|L12286666|PF|S15155213|Y|A24925087||55150-7||LNC-PT-BR|LN|55150-7|Anaplasma phagocytophilum Ac.IgG & IgM painel:#N/A:Pt:Soro:Ord|0[N]|" previously deleted because of syntax
C2708350 preferred string counts: 1/1

Restored "C2708350|POR|P|L12573653|PF|S15142213|Y|A25132736||55151-5||LNC-PT-BR|LN|55151-5|Aldosterona & Renina Painel:#N/A:Pt:Sor:Plas:On|0[N]|" previously deleted because of syntax
C2708352 preferred string counts: 1/1

Restored "C2708352|POR|P|L12574806|PF|S15151353|Y|A25445044||55152-3||LNC-PT-BR|LN|55152-3|Amiodarona & Desetilamiodarona painel:#N/A:Pt:Sor:Plas:On|0[N]|" previously deleted because of syntax
C2708354 preferred string counts: 1/1

Restored "C2708354|POR|P|L12384079|PF|S15188403|Y|A24925090||55153-1||LNC-PT-BR|LN|55153-1|Biotinidase painel:#N/A:Pt:Sor:Plas:On|0[N]|" previously deleted because of syntax
C2708355 preferred string counts: 1/1

Restored "C2708355|POR|P|L12887717|PF|S151319413|Y|A25445046||55154-9||LNC-PT-BR|LN|55154-9|Etambutol+Rifampina:#N/A:Pt:Isolado:Ord:Qn:Método de crescimento lento das micobactérias|0[N]|" previously deleted because of syntax
C2708357 preferred string counts: 1/1

Restored "C2708357|POR|P|L12287469|PF|S15333456|Y|A25029088||55155-6||LNC-PT-BR|LN|55155-6|Flunitrazepam e 7-Aminoflunitrazepam painel:#N/A:Pt:Urina:On|0[N]|" previously deleted because of syntax
C2708359 preferred string counts: 1/1

Restored "C2708359|POR|P|L12987771|PF|S15479531|Y|A25549093||55156-4||LNC-PT-BR|LN|55156-4|Minociclinina:#N/A:Pt:Isolado:Ord:Qn:Método de crescimento lento das micobactérias|0[N]|" previously deleted because of syntax
C2708360 preferred string counts: 1/1

Restored "C2708360|POR|P|L12291595|PF|S15585156|Y|A25341246||55157-2||LNC-PT-BR|LN|55157-2|Risperidona & 9-Hidroxrisperidona painel:#N/A:Pt:Sor:Plas:On|0[N]|" previously deleted because of syntax
C2708363 preferred string counts: 1/1

Restored "C2708363|POR|P|L12290960|PF|S15645257|Y|A25445047||55158-0||LNC-PT-BR|LN|55158-0|Tigeciclina:#N/A:Pt:Isolado:Ord:Qn:Método de crescimento lento das micobactérias|0[N]|" previously deleted because of syntax
C2708363 preferred string counts: 1/1

Restored "C2708363|POR|P|L12444047|PF|S15668248|Y|A25029091||55159-8||LNC-PT-BR|LN|55159-8|Ursodeoxicólico:#N/A:Pt:Soro:Qn|0[N]|" previously deleted because of syntax
C2708364 preferred string counts: 1/1

EXPLORADOR

METAMAP-PORTUGUES

sourceData / 2022B-PORTUGUES

01metawordindex

> Filter

> Generate

> Suppress

\$ 01CreateWorkFiles

02Suppress

\$ 03Filterrep M

04FilterRelaxed

04FilterShared

04Filterstrict

04FilterstrictCheck

04FilterstrictMore

05GenerateMVIfiles

allocate_nies

chunk_file

consolidate_info_lines

create_mrrank

filter_mrconsol_ALL

MRCONSO.CUIvalues

MRCONSO.LATvalues

MRCONSO.LLUIvalues

MRCONSO.RRF

MRCONSO.SABvalues

MRCONSO.SRLvalues

MRCONSO.STTvalues

MRCONSO.USvalues

MRCONSO.Tvalues

MRCONSO.TTvalues

MRRANK.RRF

MRSAB.RRF

MRSTY.RRF

S.lRaw

> 02treecodes

> 03variants

ESTRUTURA DO CÓDIGO

LINHA DO TEMPO

PROBLEMAS SAÍDA CONSOLE DE DEPURAÇÃO TERMINAL

C2708855 preferred string counts: 1/1
Restored "C2708855|POR|P|L12319716|PF|S15451862|Y|A24925235|55598-7||LNC-PT-BR|LN|55598-7|Leuccitos.desintegrados/100 leuccitos:#N/A:Pt:LiqOrg:Qn:Contagem manual[0|N]" previously deleted because of syntax
C2708857 preferred string counts: 1/1
DID NOT RESTORE "C2708857|POR|P|L12807009|PF|S15451861|Y|A25029267||55599-5||LNC-PT-BR|LN|55599-5|Leuccitos.desintegrados/100 leuccitos:#N/A:Pt:LCR:Qn:Contagem manual[0|N]" because all concepts filtered out
C2708858 preferred string counts: 1/1
Restored "C2708858|POR|P|L12630220|PF|S15514840|Y|A25654166||55600-1||LNC-PT-BR|LN|55600-1|Observao microscpica:Ident:Pt:LiqPerit/Ascit:Nom:Colorao citolgica[0|N]" previously deleted because of syntax
C2708859 preferred string counts: 1/1
Restored "C2708859|POR|P|L12630219|PF|S15514838|Y|A25654167||55601-9||LNC-PT-BR|LN|55601-9|Observao microscpica:Ident:Pt:LiqPeric:Nom:Colorao citolgica[0|N]" previously deleted because of syntax
C2708860 preferred string counts: 1/1
Restored "C2708860|POR|P|L12771751|PF|S15514843|Y|A25654168||55602-7||LNC-PT-BR|LN|55602-7|Observao microscpica:Ident:Pt:LiqPleur:Nom:Colorao citolgica[0|N]" previously deleted because of syntax
C2708861 preferred string counts: 1/1
Restored "C2708861|POR|P|L12320215|PF|S15514846|Y|A25341408||55603-5||LNC-PT-BR|LN|55603-5|Observao microscpica:Ident:Pt:LiqSinov:Nom:Colorao citolgica[0|N]" previously deleted because of syntax
C2708862 preferred string counts: 1/1
Restored "C2708862|POR|P|L12427752|PF|S15514810|Y|A25445247||55604-3||LNC-PT-BR|LN|55604-3|Observao microscpica:Ident:Pt:Endomet:Nom:Colorao citolgica[0|N]" previously deleted because of syntax
C2708863 preferred string counts: 1/1
Restored "C2708863|POR|P|L12525892|PF|S15537917|Y|A25132917||55605-0||LNC-PT-BR|LN|55605-0|Penas de periquito Ac.IgE:#N/A:Pt:Soro:Ord[0|N]" previously deleted because of syntax
C2708864 preferred string counts: 1/1
Restored "C2708864|POR|P|L12427762|PF|S15537906|Y|A25132918||55606-8||LNC-PT-BR|LN|55606-8|Penas de papagaio Ac.IgE:#N/A:Pt:Soro:Ord[0|N]" previously deleted because of syntax
C2708865 preferred string counts: 1/1
Restored "C2708865|POR|P|L12427819|PF|S15655291|Y|A25445249||55607-6||LNC-PT-BR|LN|55607-6|Triglicerides/Colesterol.HDL:#N/A:Pt:Sor/Plas:Qn[0|N]" previously deleted because of syntax
C2708866 preferred string counts: 1/1
Restored "C2708866|POR|P|L12313215|PF|S15537919|Y|A25237103||55608-4||LNC-PT-BR|LN|55608-4|Penas de per Ac.IgE:#N/A:Pt:Soro:Ord[0|N]" previously deleted because of syntax
C2708867 preferred string counts: 1/1
Restored "C2708867|POR|P|L12931534|PF|S15668599|Y|A25132921||55609-2||LNC-PT-BR|LN|55609-2|Uria:#N/A:XXX:Urina:Qn[0|N]" previously deleted because of syntax
C2708868 preferred string counts: 1/1
Restored "C2708868|POR|P|L12695181|PF|S15333556|Y|A25549289||55610-0||LNC-PT-BR|LN|55610-0|Folitropina^ amostra aps administrao de XXX:#N/A:Pt:Sor/Plas:Qn[0|N]" previously deleted because of syntax
C2708869 preferred string counts: 1/1
Restored "C2708869|POR|P|L12305280|PF|S15335537|Y|A25445253||55611-8||LNC-PT-BR|LN|55611-8|Folitropina^ amostra aps administrao de XXX:#N/A:Pt:Sor/Plas:Qn[0|N]" previously deleted because of syntax
C2708870 preferred string counts: 1/1
Restored "C2708870|POR|P|L12305284|PF|S15335631|Y|A25341411||55612-6||LNC-PT-BR|LN|55612-6|Folitropina^ amostra aps administrao de XXX:#N/A:Pt:Sor/Plas:Qn[0|N]" previously deleted because of syntax
C2708871 preferred string counts: 1/1
Restored "C2708871|POR|P|L12876238|PF|S15227912|Y|A25132932||55628-2||LNC-PT-BR|LN|55628-2|Cefadroxil:#N/A:Pt:Isolado:OrdOn:Fitas gradiente[0|N]" previously deleted because of syntax
C2708872 preferred string counts: 1/1
Restored "C2708872|POR|P|L12779803|PF|S15227996|Y|A25654178||55629-0||LNC-PT-BR|LN|55629-0|Cefaloglicina:#N/A:Pt:Isolado:OrdOn:Fitas gradiente[0|N]" previously deleted because of syntax
C2708873 preferred string counts: 1/1
Restored "C2708873|POR|P|L12839812|PF|S15228083|Y|A25029276||55630-8||LNC-PT-BR|LN|55630-8|Cefalotina+Sulbactam:#N/A:Pt:Isolado:OrdOn:Concentrao inibitria minima[0|N]" previously deleted because of syntax
C2708875 preferred string counts: 1/1

EXPLORADOR

✓ METAMAP-PORTUGUES

✓ sourceData / 2022AB-PORTUGUES

✓ 01metawordindex

✓ > Filter

✓ > Generate

✓ > Suppress

✓ \$ 01CreateWorkFiles

✓ \$ 02Suppress

✓ \$ 03filterprep M

✓ E 04filterrelated

✓ E 04filtershared

✓ E 04filterstrict

✓ E 04filterstrictcheck

✓ E 04filterstrictmore

✓ E 05GenerateMWIFiles

✓ E allocate_files

✓ E chunk_nle

✓ E consolidate_info_lines

✓ E create_mrank

✓ E filter_mrconso_ALL

✓ E MRCONSO.CUIvalues

✓ E mrconso.eng

✓ E mrconso.eng.0

✓ E mrconso.eng.ALLCUI

✓ E mrconso.eng.PPF.YCU

✓ E mrconso.eng.TsValues

✓ E MRCONSOLAT.values

✓ E MRCONSOLU.Values

✓ E MRCONSO.RRF

✓ E MRCONSO.SAB.Values

✓ E MRCONSO.SRLvalues

✓ E MRCONSO.STValues

✓ E MRCONSO.UVvalues

✓ E MRCONSO.TSValues

✓ E MRCONSO.TTValues

✓ E MRCONSO.VTValues

✓ E MRANKRRF

✓ E MRSAB.BRF

✓ E MRSITY.RRF

✓ E SRAW

> 02treecodes

> 03variants

✓ EXTRUSÃO DO CÓDIGO

> LINHA DO TEMPO

PROBLEMAS SAÍDA CONSOLE DE DEPURAÇÃO TERMINAL

DID NOT RESTORE "C2923437|POR|P|L22280933|PF|S15141794|Y|A25238179||59102-4||LNC-PT-BR|LN|59102-4|Alcance do fluxo:Comprimento:Pt:Sistema Circulatrio.XXX:Qn:US|0[N]|" because all concepts filtered out

C2923439 preferred string counts: 1/1

DID NOT RESTORE "C2923439|POR|P|L2203499|PF|S15301962|Y|A25134024||59103-2||LNC-PT-BR|LN|59103-2|Durao A-C:Tempo:Pt:Corao.vlvula.XXX:Qn:US|0[N]|" because all concepts filtered out

C2923400 preferred string counts: 1/1

DID NOT RESTORE "C2923400|POR|P|L2620555|PF|S15517550|Y|A25446287||59104-0||LNC-PT-BR|LN|59104-0|Onda E.mimo:Onda A.mximo:Pt:Corao.vlvula.XXX:Qn:US|0[N]|" because all concepts filtered out

C2923411 preferred string counts: 1/1

DID NOT RESTORE "C2923411|POR|P|L24258188|PF|S15302056|Y|A25238180||59105-7||LNC-PT-BR|LN|59105-7|Durao de onda A:Tempo:Pt:Sistema Circulatrio.XXX:Qn:US|0[N]|" because all concepts filtered out

C2923443 preferred string counts: 1/1

DID NOT RESTORE "C2923443|POR|P|L2412691|PF|S15413208|Y|A25134026||59106-5||LNC-PT-BR|LN|59106-5|Inclinao.mxima^em estenose:Pres:Pt:Sistema Circulatrio.XXX:Qn:US|0[N]|" because all concepts filtered out

C2923444 preferred string counts: 1/1

DID NOT RESTORE "C2923444|POR|P|L2347348|PF|S15673454|Y|A25342459||59107-3||LNC-PT-BR|LN|59107-3|Velocidade.mxima^em estenose:Pt:Sistema Circulatrio.XXX:Qn:US|0[N]|" because all concepts filtered out

C2923446 preferred string counts: 1/1

DID NOT RESTORE "C2923446|POR|P|L2301142|PF|S15301976|Y|A25503073||59108-1||LNC-PT-BR|LN|59108-1|Durao da medida:Tempo:Pt:XXX:Qn:US|0[N]|" because all concepts filtered out

C2923448 preferred string counts: 1/1

DID NOT RESTORE "C2923448|POR|P|L2436506|PF|S15680632|Y|A25446288||59109-9||LNC-PT-BR|LN|59109-9|Separao leaflet:Comprimento:Pt:Corao.vlvula.XXX:Qn:US|0[N]|" because all concepts filtered out

C2923450 preferred string counts: 1/1

DID NOT RESTORE "C2923450|POR|P|L2303079|PF|S15317670|Y|A25655203||59110-7||LNC-PT-BR|LN|59110-7|Espessura leaflet:Comprimento:Pt:Corao.vlvula.XXX:Qn:US|0[N]|" because all concepts filtered out

C2923452 preferred string counts: 1/1

DID NOT RESTORE "C2923452|POR|P|L2620552|PF|S15517534|Y|A24926326||59111-5||LNC-PT-BR|LN|59111-5|Onda E fluxo sangueno/ anis de onda E:Pt:Corao.vlvula.XXX:Qn:US|0[N]|" because all concepts filtered out

C2923454 preferred string counts: 1/1

DID NOT RESTORE "C2923454|POR|P|L16770326|PF|S20308041|Y|A32591001||59112-3||LNC-PT-BR|LN|59112-3|Durao de onda A.veia pulmonar/durao de onda A.vlvula mitral:Pt:Sistema Circulatrio.XXX:Qn:US|0[N]|" because all concepts filtered out

C2923456 preferred string counts: 1/1

DID NOT RESTORE "C2923456|POR|P|L2927789|PF|S15638812|Y|A25134027||59113-1||LNC-PT-BR|LN|59113-1|Tempo-velocidade integral.veia pulmonar onda A/Tempo-velocidade integral.vlvula mitral onda A:Pt:Veia Pulmonar:On:US|0[N]|" because all concepts filtered out

C2923458 preferred string counts: 1/1

DID NOT RESTORE "C2923458|POR|P|L2203519|PF|S15362054|Y|A25550376||59114-9||LNC-PT-BR|LN|59114-9|Durao de onda A.veia pulmonar - durao de onda A.vlvula mitral ::Pt:Veia Pulmonar:On:US|0[N]|" because all concepts filtered out"

C2923460 preferred string counts: 1/1

DID NOT RESTORE "C2923460|POR|P|L2430966|PF|S15563875|Y|A25446289||59115-6||LNC-PT-BR|LN|59115-6|Propagao do fluxo::Pt:Corao.vlvula.XXX:Qn:US|0[N]|" because all concepts filtered out

C2923462 preferred string counts: 1/1

DID NOT RESTORE "C2923462|POR|P|L2620550|PF|S15520860|PF|S15299334|Y|A25030344||59116-4||LNC-PT-BR|LN|59116-4|Dimetro.juno articulo sinotubular/Dimetro.raiz da aorta::Pt:Aorta:US|0[N]|" because all concepts filtered out

C2923464 preferred string counts: 1/1

DID NOT RESTORE "C2923464|POR|P|L2684289|PF|S15247792|Y|A25238180||59117-2||LNC-PT-BR|LN|59117-2|Circumferential fiber shortening.mean::Pt:Corao.ventrculo.esquerdo:US|0[N]|" because all concepts filtered out

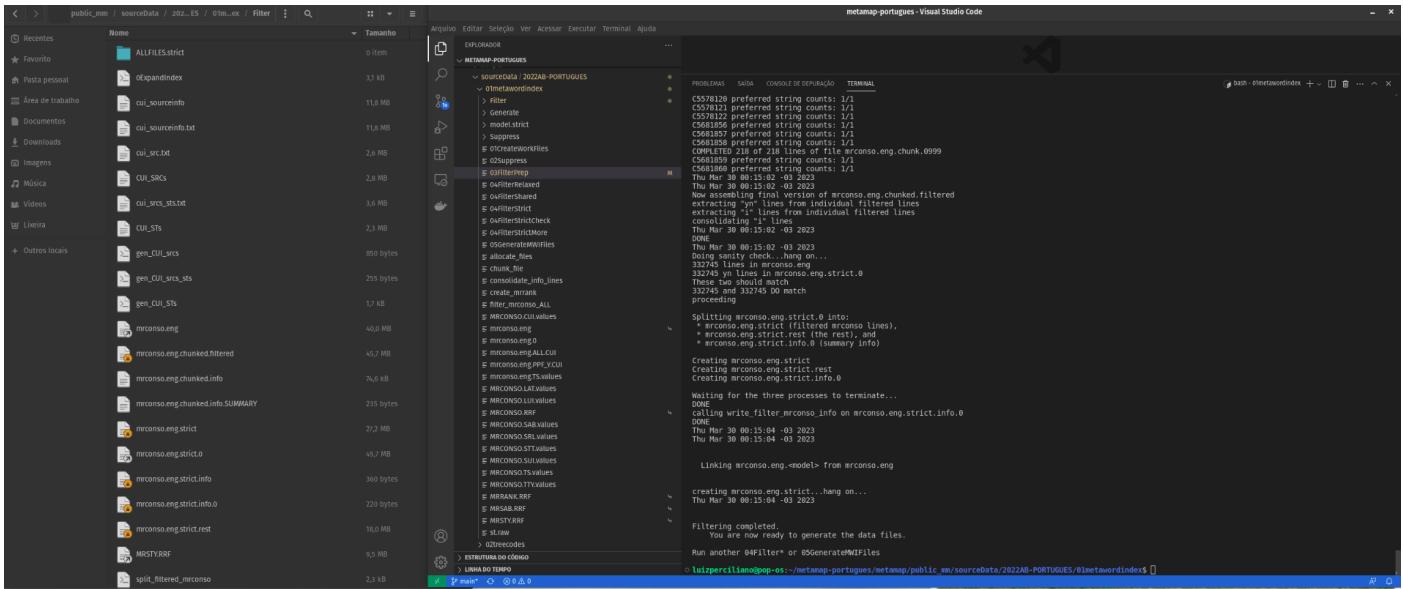
C2923466 preferred string counts: 1/1

DID NOT RESTORE "C2923466|POR|P|L2878747|PF|S15247793|Y|A25238182||59118-0||LNC-PT-BR|LN|59118-0|Circumferential fiber shortening.mean^~corrected for heart rate::Pt:Corao.ventrculo.esquerdo:US|0[N]|" because all concepts filtered out

C2923467 preferred string counts: 1/1

DID NOT RESTORE "C2923467|POR|P|L2525511|PF|S15540987|Y|A25134028||59119-8||LNC-PT-BR|LN|59119-8|Periodo enchimento.distole:Tempo:Pt:Corao.ventrculo.XXX:Qn:US|0[N]|" because all concepts filtered out

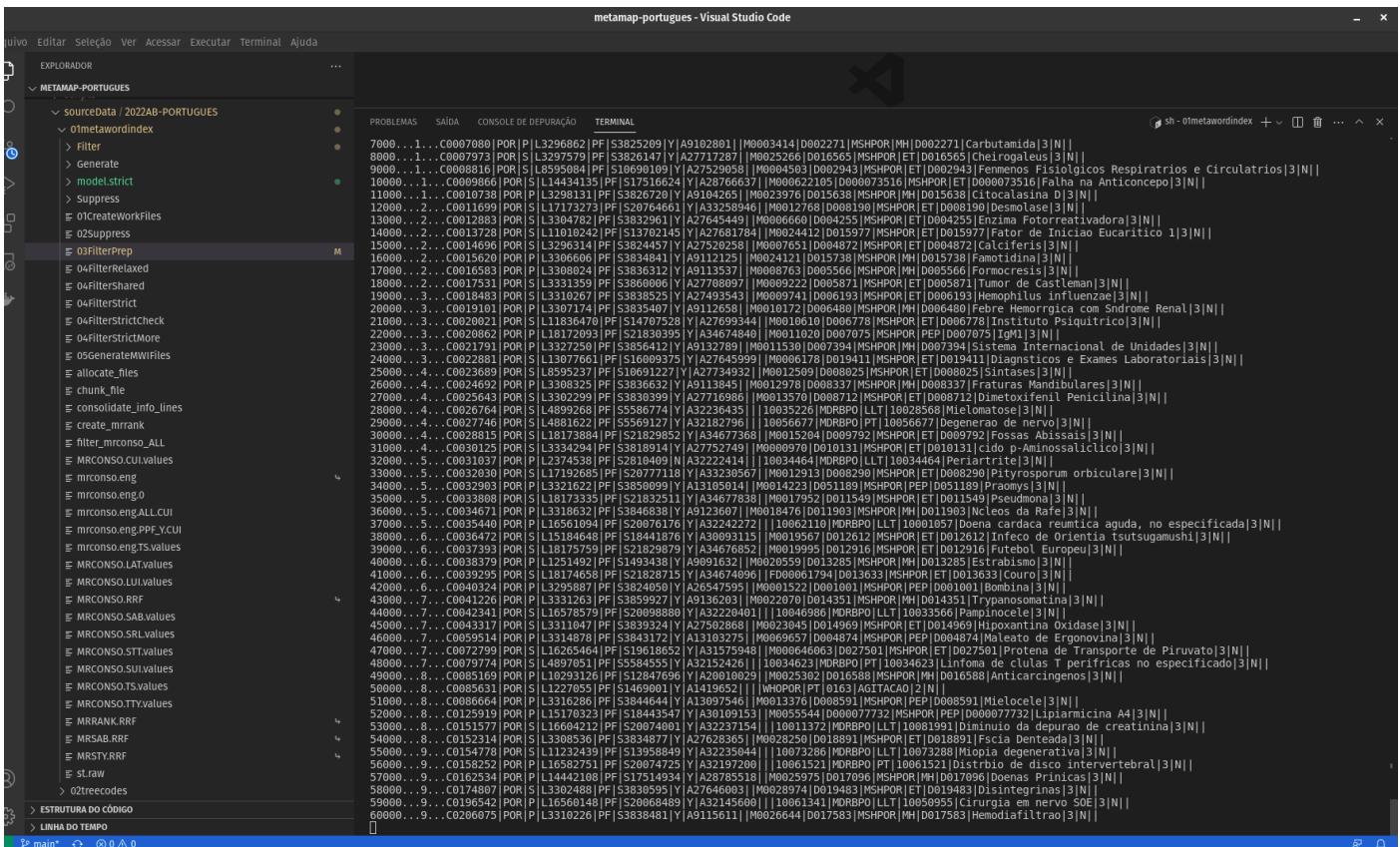
C2923468 preferred string counts: 1/1



3.4 Acessar o ./05GenerateMWIFiles

./05GenerateMWIFiles

Inicio em 30/03/2023 - 00:20h a 00:21



metamap-portugues - Visual Studio Code

Arquivo Editar Seleção Ver Acessar Executar Terminal Ajuda

EXPLORADOR

METAMAP-PORTUGUES

sourceData / 2022AB-PORTUGUES

01metawordindex

- > Filter
- > Generate
- > model.strict
- > SAB
- > Suppress
- 01CreateWorkfiles
- 02Suppress
- 03FilterPrep
- 04FilterRelaxed
- 04FilterShared
- 04FilterStrict
- 04FilterStrictCheck
- 04FilterStrictMore
- 05GenerateMWIFiles
- allocate_files
- chunk_file
- consolidate_info_lines
- create_mrank
- filter_mrcons0_ALL
- MRCONSO.CUI.values
- mrcons0.eng
- mrcons0.eng.0
- mrcons0.eng.ALL.CUI
- mrcons0.eng.PPF-Y.CUI
- mrcons0.eng.TS.values
- MRCONSO.LAT.values
- MRCONSO.LUI.values
- MRCONSO.RRF
- MRCONSO.SAB.values
- MRCONSO.SRLvalues
- MRCONSO.STTvalues
- MRCONSO.SUI.values
- MRCONSO.TS.values
- MRCONSO.TTY.values
- MRRANK.RRF
- MRSAB.RRF
- MRTSTY.RRF
- srt.raw

ESTRUTURA DO CÓDIGO

LINHA DE TEMPO

PROBLEMAS SAÍDA CONSOLE DE DEPURAÇÃO TERMINAL

```
in model.strict deleting first_words_of_one.sorted first_words_of_one.WithStr.sorted first_words_of_one_WIDE.txt
in model.strict DONE
in model.strict renaming first_words_of_one.WithStrAndConc to first_words_of_one_WIDE.txt
in model.strict DONE
in model.strict deleting first_words_of_two.sorted first_words_of_two.WithStr.sorted first_words_of_two_WIDE.txt
in model.strict DONE
in model.strict renaming first_words_of_two.WithStrAndConc to first_words_of_two_WIDE.txt
in model.strict DONE
in model.strict deleting first_wordsb.sorted first_wordsb.WithStr.sorted first_wordsb_WIDE.txt
in model.strict DONE
in model.strict renaming first_wordsb.WithStrAndConc to first_wordsb_WIDE.txt
in model.strict DONE
in model.strict DONE
in model.strict renaming first_wordsb.WithStrAndConc to first_wordsb_WIDE.txt
in model.strict DONE
in model.strict renaming first_wordsb.WithStrAndConc to first_wordsb_WIDE.txt
in model.strict DONE
in model.strict DONE
in model.strict renaming all_words.sorted all_words.WithStr.sorted all_words_WIDE.txt
in model.strict DONE
in model.strict renaming all_words.WithStrAndConc to all_words_WIDE.txt
in model.strict DONE
Thu Mar 30 00:21:15 -03 2023

Generating tables sab_rv.txt and sab_vr.txt
creating MRSAF_SAB
DONE
creating MRCONSO.SAB
DONE
creating SAB_1
DONE
creating SAB_0
DONE
creating sab_rv.txt
DONE
creating sab_vr.txt
DONE

You may verify that the MWI data files were created by looking in the model.[relaxed|moderate|strict] directory for the model(s) you created and finding these files:
    all_words.txt
    concept_cui.txt
    concept_st.txt
    first_words.txt
    first_words_of_one.txt
    first_words_of_two.txt
    sui_nmstr_str.txt

For the next step:
Move up out of the 01metawordindex directory and down into the 02treecodes.

There you run the 01GenerateTreecodes script.
```

luizperciliano@pop-os:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/01metawordindex\$

4 Executando arquivos de outros diretórios

Última atualização: 21/02/2023 23:40

13 de 21

```

cd $PUBLIC_MM/sourceData/2022AB-PORTUGUES/02
./01GenerateTreecodes

cd $PUBLIC_MM/sourceData/2022AB-PORTUGUES/03variants
./01GenerateVariants

cd $PUBLIC_MM/sourceData/2022AB-PORTUGUES/04synonyms
./01GenerateSynonyms

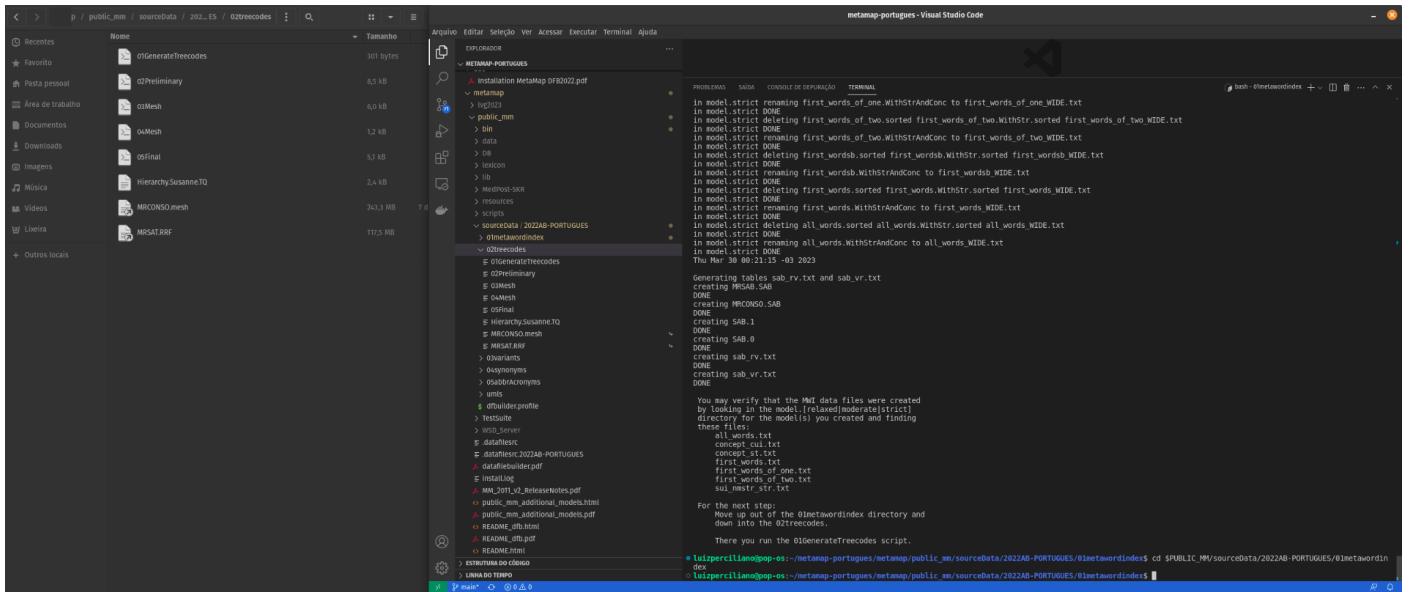
cd $PUBLIC_MM/sourceData/2022AB-PORTUGUES/05abbrAcronyms
./01GenerateAbbrAcronyms

```

Este script leva cerca de 3 minutos e gera o arquivo treecodes.txt. (não está gerando este)

- O diretório 02treecodes inicialmente tinha 8 itens com total de 23,7kb, após executar o ./02Suppress, o diretório ficou com 95 itens, totalizando 158,7MB.
- O diretório 03variants inicialmente tinha 7 itens com total de 16,2kb, após executar o ./01GenerateVariants, o diretório ficou com 79 itens, totalizando 16,4kb (o arquivo mrconso.filtered ao tentar abrir deu quebrado).
- O diretório 04synonyms inicialmente tinha 6 itens com total de 3,2kb, após executar o ./01GenerateSynonyms, o diretório ficou com 17 itens, totalizando 103,2MB (é gerado sinônimos, porém todos em inglês, visto no arquivo syn_nls_db.txt.0).
- O diretório 05abbrAcronyms inicialmente tinha 5 itens com total de 9,0kb, após executar o ./01GenerateAbbrAcronyms, o diretório ficou com 12 itens, totalizando 9kb (**Arquivos gerados com conteúdo em inglês**).

No final de tudo, o diretório UMLS-PORTUGUES fica com um total de 2.294 itens, totalizando 879,6 MB.



Screenshot of Visual Studio Code showing the command line interface for generating treecodes. The terminal output shows the process of creating mesh files, generating check tags, and compressing intermediate files. The command run is `luisperciliano@pop-os:~/metmap-portugues/metmap/public_mm/sourceData/2022AB-PORTUGUES/63treecodes\$`.

```

PROBLEMAS SAÍDA CONSOLE DE DESENVOLVIMENTO TERMINAL
[...]
DONE
Creating mesh_tc.strict.mo
DONE
Creating meta.mesh
DONE
creating meta_mesh.opt.txt
Forcing opt table (cols 1 and 2) from meta.mesh to produce file meta_mesh.opt.txt
DONE
Thu Mar 30 00:25:12 -03 2023

Creating mesh_tc.relaxed.txt and
mesh_mm.opt.txt
allowing all Mesh terms to acquire treecodes from their MN
accounting for supplementary chemicals, subheadings, check tags
WARNING: Make sure MSAT.check.tags has been created and that
mesh_tc.chems and mesh_tc.subs are linked to
<NW>/data/08/MSHStudy/ files
creating mesh_tc.check.tags
DONE
Creating mesh_tc.relaxed.mo
DONE
Thu Mar 30 00:25:12 -03 2023

Computing field widths
Compressing intermediate files
DONE
0 mesh_mm.opt.txt
0 mesh_tc_relaxed.txt
0 mesh_tc.strict.txt
0 mesh_mm_relaxed.txt
0 meta_mesh_tc.opt.txt
58 total
Thu Mar 30 00:25:12 -03 2023
Thu Mar 30 00:25:12 -03 2023

The next and final generation step is variant generation.
Move up out of 63treecodes and down into 03variants.
Run 01GenerateVariants.

luisperciliano@pop-os:~/metmap-portugues/metmap/public_mm/sourceData/2022AB-PORTUGUES/63treecodes$ 

```

Screenshot of Visual Studio Code showing the command line interface for generating variants. The terminal output shows the process of creating mesh files, generating check tags, and compressing intermediate files. The command run is `luisperciliano@pop-os:~/metmap-portugues/metmap/public_mm/sourceData/2022AB-PORTUGUES/03variants\$`.

```

PROBLEMAS SAÍDA CONSOLE DE DESENVOLVIMENTO TERMINAL
[...]
creating mesh_tc.strict.mo
DONE
Creating meta.mesh
DONE
creating meta_mesh.opt.txt
Forcing opt table (cols 1 and 2) from meta.mesh to produce file meta_mesh.opt.txt
DONE
Thu Mar 30 00:25:12 -03 2023

Creating mesh_tc.relaxed.txt and
mesh_mm.opt.txt
allowing all Mesh terms to acquire treecodes from their MN
accounting for supplementary chemicals, subheadings, check tags
WARNING: Make sure MSAT.check.tags has been created and that
mesh_tc.chems and mesh_tc.subs are linked to
<NW>/data/08/MSHStudy/ files
creating mesh_tc.check.tags
DONE
Creating mesh_tc.relaxed.mo
DONE
Thu Mar 30 00:25:12 -03 2023

Computing field widths
Compressing intermediate files
DONE
0 mesh_mm.opt.txt
58 mesh_tc_relaxed.txt
0 mesh_tc.strict.txt
0 mesh_mm_relaxed.txt
0 meta_mesh_tc.opt.txt
59 total
Thu Mar 30 00:25:12 -03 2023
Thu Mar 30 00:25:12 -03 2023

The next and final generation step is variant generation.
Move up out of 03treecodes and down into 03variants.
Run 01GenerateVariants.

luisperciliano@pop-os:~/metmap-portugues/metmap/public_mm/sourceData/2022AB-PORTUGUES/03variants$ 

```

Screenshot of Visual Studio Code showing the command line interface for generating variants. The terminal output shows the process of creating mesh files, generating check tags, and compressing intermediate files. The command run is `luisperciliano@pop-os:~/metmap-portugues/metmap/public_mm/sourceData/2022AB-PORTUGUES/03variants\$`.

```

PROBLEMAS SAÍDA CONSOLE DE DESENVOLVIMENTO TERMINAL
[...]
SKRrun_16 -w /home/luisperciliano/metmap-portugues/metmap/public_mm/lexicon_mm_variants -Du /home/luisperciliano/metmap-portugues/metmap/public_mm/sourceData/2022AB-PORTUGUES/03variants/versmu.variants.o
mm_variants (2020)
control options:
    mm_data_version=5
    mm_data_year=2022AB
    lexicon_db
        all_across_abbrs
        unique_across_abbrs_only
        Variant generation mode: dynamic.
        ## WARNING: Overriding default model 2020AA with 2022AB.

#####
##### ERROR!!!
##### Predicate exec:unit_db/1 is UNKNOWN in module db_access
#####
Flipping all versmu file.
Sorting the flipped versmu file.
SKRrun_16 -w /home/luisperciliano/metmap-portugues/metmap/public_mm/lexicon_mm_variants -u /home/luisperciliano/metmap-portugues/metmap/public_mm/sourceData/2022AB-PORTUGUES/03variants/versmu.variants.o
mm_variants (2020)
Control options:
    mm_data_version=5
    mm_data_year=2022AB
    lexicon_db
        unique_across_abbrs_only
        Variant generation mode: dynamic.
        ## WARNING: Overriding default model 2020AA with 2022AB.

#####
##### ERROR!!!
##### Predicate exec:unit_db/1 is UNKNOWN in module db_access
#####
Flipping all versmu file.
Sorting the flipped versmu file.
From the last versmu file.
Thu Mar 30 00:28:32 -03 2023

The next generation step is Synonym generation.
Move up out of 03variants and down into 04synonyms.
Run 01GenerateSynonyms.

luisperciliano@pop-os:~/metmap-portugues/metmap/public_mm/sourceData/2022AB-PORTUGUES/03variants$ 

```

Última atualização: 21/02/2023 23:40

metamap-portugues - Visual Studio Code

PROBLEMAS SAÍDA CONSOLE DE DEPURAÇÃO TERMINAL

```
luisperciliano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/04synonyms$ cd $PUBLIC_MM/sourceData/2022AB-PORTUGUES/04synonyms
luisperciliano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/04synonyms$ ./04GenerateSynonyms
```

metamap-portugues - Visual Studio Code

PROBLEMAS SAÍDA CONSOLE DE DEPURAÇÃO TERMINAL

```
luisperciliano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/04synonyms$ cd $PUBLIC_MM/sourceData/2022AB-PORTUGUES/04synonyms
luisperciliano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/04synonyms$ ./04GenerateSynonyms
```

metamap-portugues - Visual Studio Code

PROBLEMAS SAÍDA CONSOLE DE DEPURAÇÃO TERMINAL

```
luisperciliano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/04synonyms$ cd $PUBLIC_MM/sourceData/2022AB-PORTUGUES/04synonyms
luisperciliano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/04synonyms$ ./04GenerateSynonyms
```

Última atualização: 21/02/2023 23:40

metamap-portugues - Visual Studio Code

PROBLEMAS SAÍDA CONSOLE DE DEPURAÇÃO TERMINAL

```
luisperciliano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/04synonyms$ cd $PUBLIC_MM/sourceData/2022AB-PORTUGUES/04synonyms
luisperciliano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/04synonyms$ ./04GenerateSynonyms
```

metamap-portugues - Visual Studio Code

PROBLEMAS SAÍDA CONSOLE DE DEPURAÇÃO TERMINAL

```
luisperciliano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/04synonyms$ cd $PUBLIC_MM/sourceData/2022AB-PORTUGUES/04synonyms
luisperciliano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/04synonyms$ ./04GenerateSynonyms
```

metamap-portugues - Visual Studio Code

PROBLEMAS SAÍDA CONSOLE DE DEPURAÇÃO TERMINAL

```
luisperciliano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/04synonyms$ cd $PUBLIC_MM/sourceData/2022AB-PORTUGUES/04synonyms
luisperciliano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/04synonyms$ ./04GenerateSynonyms
```

```

luisperciliiano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/04synonyms$ cd $PUBLIC_MM/sourceData/2022AB-PORTUGUES/04synonyms
luisperciliiano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/04synonyms$ ./04GenerateSynonyms
Thu Mar 30 2023 00:30:32 -03
creating SM.0B
DONE
creating SM.0B
DONE
creating syn.nls.db.txt.0
Processing file SM.0B to produce file syn.nls.db.txt.0
Each raw entry produces two <item><cat><type> entries
DONE
creating syn.nls.db.txt
DONE
creating col.1
DONE
creating col.1
DONE
creating col.diff
col.diff is empty
4 7 23 col.diff
DONE
creating syns.txt.0
You can probably ignore the following messages...
creating syns.txt
Remove any problematic synonyms manually.
Note: there have been no such changes since 2001.

The next generation step is Synonym generation.
Move up out of 04synonyms and down into 05abbrAcronyms.
Run ./05GenerateAbbrAcronyms.

luisperciliiano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/05abbrAcronyms$ cd $PUBLIC_MM/sourceData/2022AB-PORTUGUES/05abbrAcronyms
luisperciliiano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/05abbrAcronyms$ ./05GenerateAbbrAcronyms

```

```

luisperciliiano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/05abbrAcronyms$ cd $PUBLIC_MM/sourceData/2022AB-PORTUGUES/05abbrAcronyms
luisperciliiano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/05abbrAcronyms$ ./05GenerateAbbrAcronyms
Thu Mar 30 00:30:37 -03 2023
creating syns.txt.singlemulti
DONE
qui 30 mar 2023 00:30:37 -03
The next and final generation step is Acronym and Abbreviation generation.
Move up out of 05abbrAcronyms and down into 06abbrAcronyms.
Run ./06GenerateAbbrAcronyms.

luisperciliiano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/06abbrAcronyms$ cd $PUBLIC_MM/sourceData/2022AB-PORTUGUES/06abbrAcronyms
luisperciliiano@pop-es:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/06abbrAcronyms$ ./06GenerateAbbrAcronyms
Thu Mar 30 00:31:53 -03 2023
Can't open /home/luisperciliano/metamap-portugues/metamap/public_mm/lexicon/data/lexiconStatic2022: No such file or directory at ./extract_abbr.perl line 9.
creating am.db.sorted
DONE
creating am.db
DONE
calling charatty on am.db.0
x
DONE
creating nls_aa.txt.0
Processing file am.db to produce file nls_aa.txt.0
Each raw entry produces two <item><cat><type> entries
DONE
creating nls_aa.txt
DONE
creating nls_aa.unique
DONE
creating nls_aa.unique
DONE
creating nls_aa.unique
Reading in nls_aa.unique...
Reading in nls_aa.unique...
DONE
Thu Mar 30 00:31:53 -03 2023
All of the data files for a complete data set should be ready.
You can enter the final stage to collect these files into a new mmf/data/meta directory and load them into your MySQL database.

To create the data set and load:
cd ../
.../bin/LoadDataFiles
For Windows, run this in a Command Prompt window,
```

5 Arquivo LoadDataFiles

Sem ajustes neste arquivo. Executar a partir de /public_mm

Cria a versão final dos arquivos de dados e os carrega em um banco de dados Berkeley DB usado pelo MetaMap

```

cd $PUBLIC_MM
./bin/LoadDataFiles
Enter (4x)

```

Pergunta 1

```
luizperciliano@pop-os:~/metamap-umls2022ab/public_mm/sourceData/UMLS-PORTUGUES/05abbrAcronyms$ cd $PUBLIC_MM  
luizperciliano@pop-os:~/metamap-umls2022ab/public_mm$ ./bin/LoadDataFiles  
Enter workspace name (<year>_<source>) for model data [default is UMLS-PORTUGUES]:
```

Pergunta 2

```
Enter workspace name (<year>_<source>) for model data [default is UMLS-PORTUGUES]: 2022ab / 2  
Data files to be loaded will be read from /home/luiperciliano/metamap-umls2022ab/public_mm/sourceData/UMLS-PORTUGUES, is this okay? [yes]:  
| no script file | length: 3,201 lines: 109 Ln: 9 Col: 18 Pos: 171
```

Pergunta 3

```
luizperciliano@pop-os:~/metamap-umls2022ab/public_mm$ ./bin/LoadDataFiles
Enter workspace name (<year> <source>) for model data [default is UMLS-PORTUGUES]:
Data files to be loaded will be read from /home/luizperciliano/metamap-umls2022ab/public_mm/sourceData/UMLS-PORTUGUES, is this okay? [yes]: y
Enter data year name (<year><release>) for model data [default is 2022AB]: 2022AB
```

Pergunta 4

```
luizperciliano@pop-os:~/metamap-umls2022ab/public_mm$ ./bin/LoadDataFiles
Enter workspace name (<year>_<source>) for model data [default is UMLS-PORTUGUES]:
  Data files to be loaded will be read from /home/luizperciliano/metamap-umls2022ab/public_mm/sourceData/UMLS-PORTUGUES, is this okay? [yes]: yes
  Use default processors by default
Enter data year name (<year><release>) for model data [default is 2022AB]: PROCESSORS / 2
The final DB files will reside in /home/luizperciliano/metamap-umls2022ab/public_mm/DB/DB.UMLS-PORTUGUES.2022AB.<model>, is this okay? [yes]: yes
length: 3201 lines: 128 | Ln: 9 Col: 10 Pos: 177
```

Ajustado perguntas para by passar da seguinte forma.

Tem que rodar os 17 passos e gerar duas bases em português. (+- 3 minutos, foi menos)

O novo conjunto de dados será depositado na diretoria \$BASEDIR / DB /DB.09_custom.base e \$BASEDIR / DB / DB.09_custom.strict ou \$BASEDIR / DB / DB.09_custom.relaxed. Um conjunto de dados completo contém pelo menos dois subdiretórios: O subdiretório base contém arquivos compartilhados pelos modelos restritos e relaxados, e os diretórios restritos e relaxados contêm todos os modelos específicos.

```
$ loaddatabases.sh
metamap / public_mm > bin > $ loaddatabases.sh
23
24 export BASEDIR
25 export PATH
26
27 # Load environment log if user set variable previously
28 # Load environment log if user set variable previously
29 # Load sr file if present
30 If [ -r $BASEDIR/bin/datasrcsr ]; then
31   . $BASEDIR/bin/datasrcsr
32 fi
33
34 If [ $CUSTOM_TAG == "unknown" ]; then
35   CUSTOM_TAG=01_custom
36 fi
37
38 If [ -z "$LEX_YEAR" ]; then
39   LEX_YEAR=2016
40 fi
41
42 If [ -z "$DATA_YEAR" ]; then
43   DATA_YEAR=2016A
44 fi
45
46 # ask user for custom tag for dataset.
47 # user input ->year->model>
48 ui validate
49 until [ $ui.valid == "true" ]
50 do
51   echo "Enter workspace name (<year>-<source>) for model data [default is $CUSTOM_TAG]:"
52   read ui.input
53   case "$ui.input" in
54     *)
55       CUSTOM_TAG=$ui.input
56   esac
57 done
58
59 Data files to be loaded will be read from /home/uiluzperciliiano/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES, is this okay? [yes]
60 Enter data year name (<year>-release) for model data [default is 2022AB]:
61
62 The final DB files will result in /home/uiluzperciliiano/metamap-portugues/metamap/public_mm/DB/DB.2022AB-PORTUGUES.2022AB.<model>, is this okay? [yes]
63 WARNING: Relaxed model will not be loaded, relaxed model directory not present.
64 WARNING: Moderate model will not be loaded, moderate model directory not present.
65 Would you like to load the strict model [no]?
66
67 LINHA DO TÉRMINO
```

The screenshot shows a Windows desktop environment. In the foreground, a Visual Studio Code window is open, displaying a terminal session. The terminal window has the title "loaddatafiles.sh - metanmap-portugues - Visual Studio Code". It contains a script for loading data files, with syntax highlighting for shell commands and variables. The script includes sections for setting up environment variables, handling custom tags, and reading user input for workspace names and model tags. Below the terminal, there are tabs for "PROBLEMAS", "SAÍDA", "CONSOLE DE DEPURAÇÃO", and "TERMINAL". The background shows a file explorer window titled "Nome" (Name) with a tree view of files and folders, including "DB.2022AB-PORTUGUES.2022AB.base" and "DB.USAbase.2020AA.strict". The desktop also features icons for Recentes (Recent), Favorito (Favorites), Pasta pessoal (Personal Folder), Área de trabalho (Workplace), Documentos (Documents), Downloads, Imagens (Images), Música (Music), Vídeos (Videos), Lixeira (Recycle Bin), and Outros locais (Other locations). The taskbar at the bottom shows several pinned application icons.

6 Detalhes dos arquivos de teste

Colocar o resultado do arquivo de configuração da [planilha de teste](#).

7 Teste do modelo em PT-BR

Verificar se está funcionando

```
cd metamap-2022/public_mm/  
./bin/skrmedpostctl start  
./bin/metamap V 2022AB PORTUGUES G I  
  
./bin/metamap
```

Arquivo Editar Seleção Ver Acessar Executar Terminal Ajuda

EXPLORADOR METAMAP-PORTUGUES

- Installation MetaMap DFB2022.pdf
- metamap
- lvg2023
- public_mm
- bín
- data
- DB
- DB.2022AB-PORTUGUES.2022AB.base
- DB.2022AB-PORTUGUES.2022AB.strict
- DB.USAbase.2020AA.base
- DB.USAbase.2020AA.strict
- DB.USAbase.2022AB.base
- DB.USAbase.2022AB.strict
- lexicon
- lib
- MedPost-SKR
- resources
- scripts
- sourcedata / 2022AB-PORTUGUES
- 01metawordindex
- 02treecodes
- 03variants
- 04synonyms
- 05abacronyms
- umls
- dfbuilder.profile
- TestSuite
- WS_D_Server
- datafilesrc
- datafilesrc.2022AB-PORTUGUES
- datafilebuilder.pdf
- install.log
- MM_2011_V2_ReleaseNotes.pdf
- public_mm_additional_models.html
- public_mm_additional_models.pdf
- README_dfb.html
- README_dfb.pdf
- README.html
- README.pdf
- UMLS-2022AB-full
- ESTRUTURA DO CÓDIGO
- LINHA DO TEMPO

\$ loaddatafiles.sh x
23
24 export BASEDIR
25 export PATH
26
luizperciliano@pop-os:/home/luzperciliano/metamap-portugues/metamap/public_mm/bin\$./bin/metamap
/home/luzperciliano/metamap-portugues/metamap/public_mm/bin/SKRun.20 /home/luzperciliano/metamap-portugues/metamap/public_mm/bin/metamap20.BINARY.Linux --le
xicon db -Z 2020AA
Berkeley DB databases (USAbase 2020AA strict model) are open.
Static variants will come from table varsan in /home/luzperciliano/metamap-portugues/metamap/public_mm/DB/DB.USAbase.2020AA.strict.
Derivational Variants: Adj/noun ONLY.
Variant generation mode: static.
Established connection \$stream(139912361457008) to TAGGER Server on localhost.
metamap (2020)
Control options:
composite_phrases=4
lexicon=db
mm_data_year=2020AA
|: cancer
|:
Processing USER.tx.1: cancer
|:
Phrase: cancer
Meta Mapping (1000):
1000 CANCER [Malignant Neoplasms] [Neoplastic Process]
Meta Mapping (1000):
1000 Cancer (Cancer Genus) [Eukaryote]
Meta Mapping (1000):
1000 Cancer (Primary malignant neoplasm) [Neoplastic Process]
|: pulmão
|:
Processing USER.tx.1: pulmão
Phrase: pulmão
|: canela
|:
Processing USER.tx.1: canela
Phrase: canela
Meta Mapping (1000):
1000 Canela [Intellectual Product]
|: pulmao
|:
Processing USER.tx.1: pulmao
Phrase: pulmao
|: glicose
|:
Processing USER.tx.1: glicose
Phrase: glicose
|: Tumor
|:

Testar os termos dor, fígado, pulmão, bexiga, urina

8 Referências

- [1] <https://github.com/LuizPerciliano/metamap-portugues>
[2]