

MetaMap

Perciliano

Created in March, 18, 2023

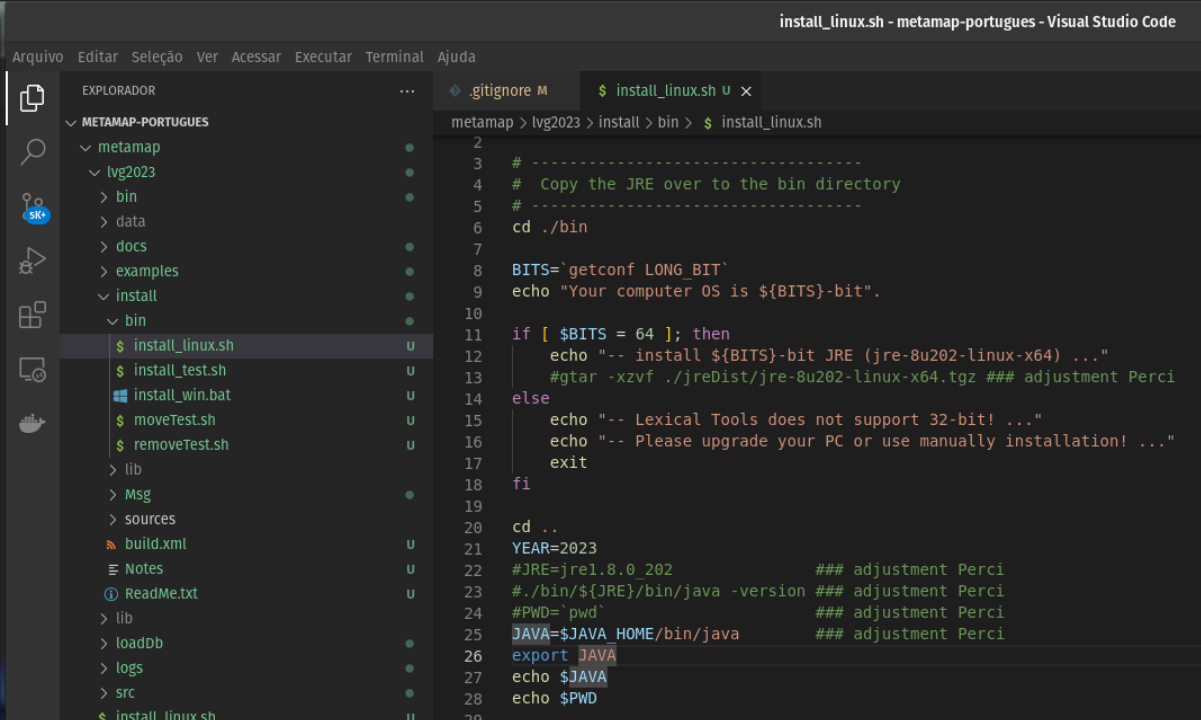
Update in March, 19, 2023

Infra	3
Install LVG2023	3
Install MetaMap + DFB	4
UMLS e Metamorphosys	4
Gerando arquivos de dados	5
./01CreateWorkFiles	5
./02Suppress	5
./03FilterPrep	6
./04FilterStrict	8

Infra

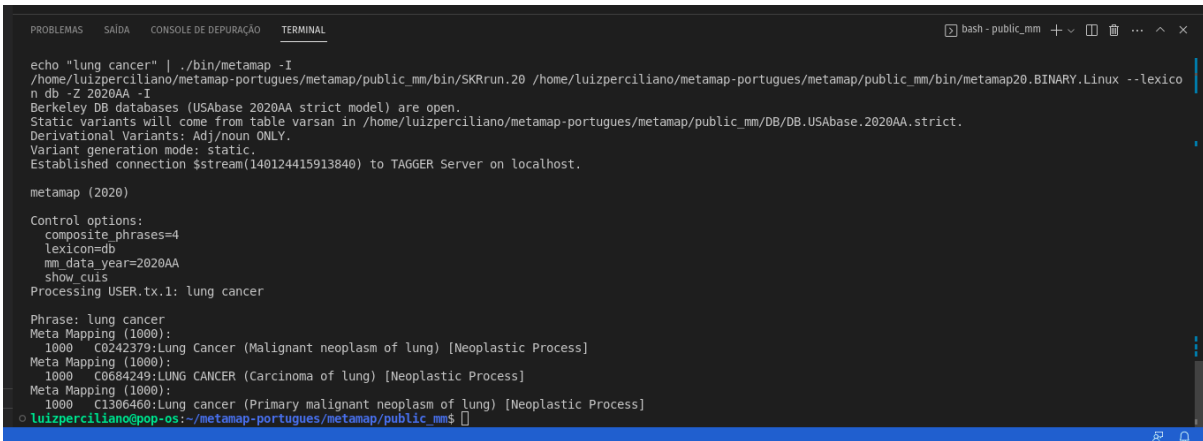
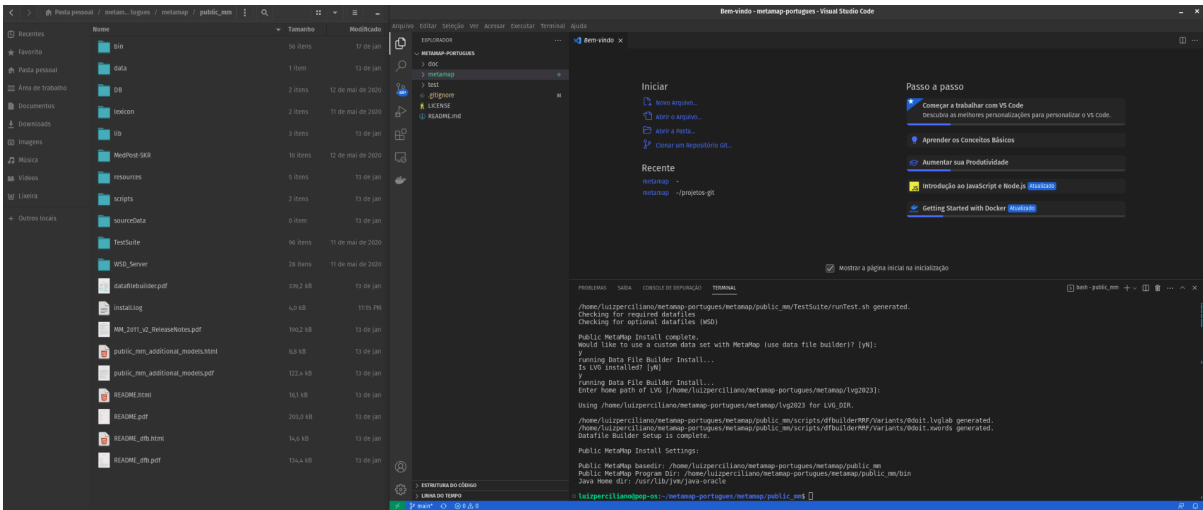
```
NAME="Pop!_OS"
VERSION="22.04 LTS"
ID=pop
ID_LIKE="ubuntu debian"
PRETTY_NAME="Pop!_OS 22.04 LTS"
VERSION_ID="22.04"
HOME_URL="https://pop.system76.com"
SUPPORT_URL="https://support.system76.com"
BUG_REPORT_URL="https://github.com/pop-os/pop/issues"
PRIVACY_POLICY_URL="https://system76.com/privacy"
VERSION_CODENAME=jammy
UBUNTU_CODENAME=jammy
LOGO=distributor-logo-pop-os
```

Install LVG2023

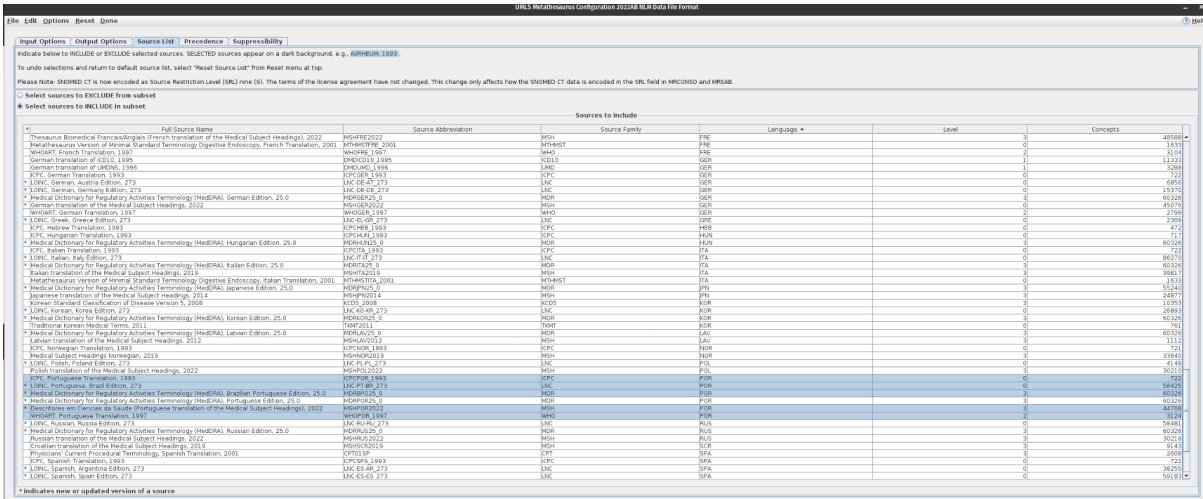


```
install_linux.sh - metaportugues - Visual Studio Code
Arquivo  Editar  Seleção  Ver  Acessar  Executar  Terminal  Ajuda
EXPLORADOR
METAMAP-PORTUGUES
  metaportugues
    lv2023
      bin
      data
      docs
      examples
      install
        bin
          install_linux.sh
          install_test.sh
          install_win.bat
          moveTest.sh
          removeTest.sh
      lib
      Msg
      sources
      build.xml
      Notes
      ReadMe.txt
      lib
      loadDb
      logs
      src
      install_linux.sh
.gitignore M  $ install_linux.sh U x
metaportugues > lv2023 > install > bin > $ install_linux.sh
1
2
3 # -----
4 # Copy the JRE over to the bin directory
5 # -----
6 cd ./bin
7
8 BITS=`getconf LONG_BIT`
9 echo "Your computer OS is ${BITS}-bit".
10
11 if [ $BITS = 64 ]; then
12     echo "-- install ${BITS}-bit JRE (jre-8u202-linux-x64) ..."
13     #gtar -xvzf ./jreDist/jre-8u202-linux-x64.tgz ### adjustment Perci
14 else
15     echo "-- Lexical Tools does not support 32-bit! ..."
16     echo "-- Please upgrade your PC or use manually installation! ..."
17     exit
18 fi
19
20 cd ..
21 YEAR=2023
22 #JRE=jre1.8.0_202 ### adjustment Perci
23 #./bin/${JRE}/bin/java -version ### adjustment Perci
24 #PWD=`pwd` ### adjustment Perci
25 JAVA=$JAVA_HOME/bin/java ### adjustment Perci
26 export JAVA
27 echo $JAVA
28 echo $PWD
29
```

Install MetaMap + DFB



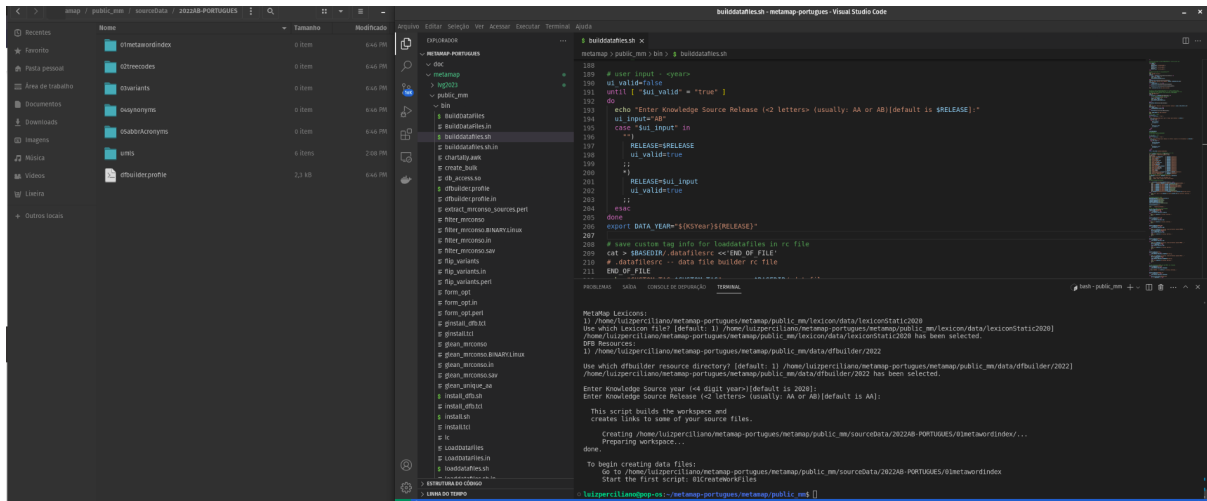
UMLS e Metamorphosys



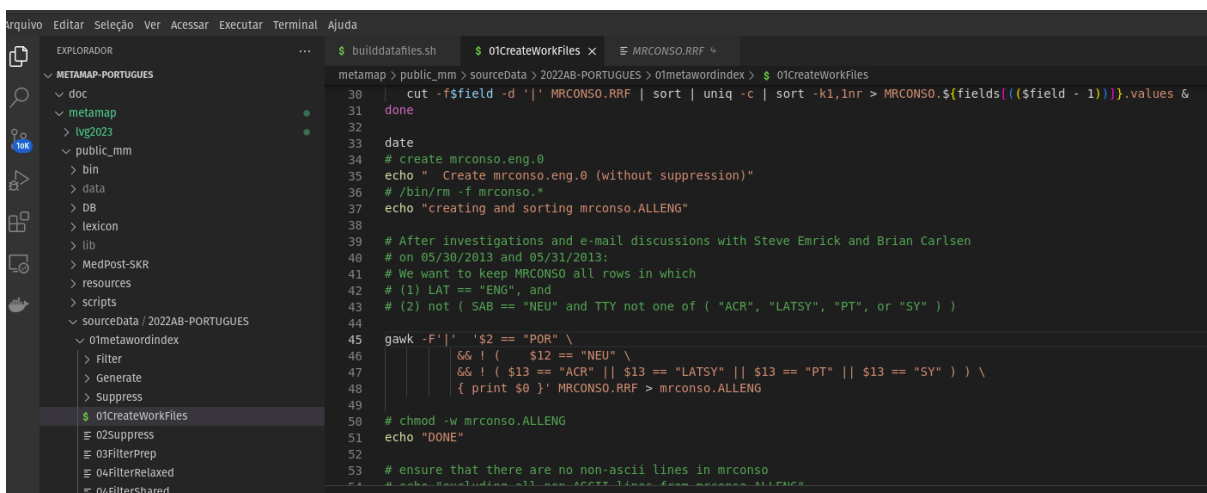
In the document DFB, the step 10, page 5 (datafilebuilder.pdf) below, was ignored, it's wasn't find:

In the "Eliminate Extended Unicode Characters" box, select option: "Remove records containing extended UTF-8 characters".

Gerando arquivos de dados



./01CreateWorkFiles



./02Suppress

MRCONSO.STT.values - metaportugues - Visual Studio Code

Arquivo Editar Seleção Ver Acessar Executar Terminal Ajuda

EXPLORADOR ... \$ bulddatfiles.sh \$ 01createworkfiles MRCONSO.STT.values x

metaport > public_mm > sourcedata > 2022AB-PORTUGUES > 01metawordindex > MRCONSO.STT.values

1 332776 PF
2

EXPLORE METAMAP-PORTUGUES
MedPost-SKR
resources
scripts
sourcedata / 2022AB-PORTUGUES
01metawordindex
Filter
Generate
Suppress
01createworkfiles
02suppress
03filterprep
04filterrelaxed
04filtershared
04filterstrict
04filterstrictcheck
04filterstrictmore
05generatemainfiles
allocate_files
chunk_file
consolidate_info_lines
create_mrrank
filter_mrconso_all
MRCONSO.CUI.values
mrconso.eng
mrconso.eng 0
mrconso.eng ALL.CUI
mrconso.eng PPF_Y.CUI
mrconso.eng TS.values
MRCONSO.LAT.values
MRCONSO.LUI.values
MRCONSO.RRF
MRCONSO.SAB.values
MRCONSO.SRL.values
MRCONSO.STT.values
MRCONSO.SUI.values
MRCONSO.TS.values
MRCONSO.TTY.values
MRRANK.RRF
MRSAB.RRF
MRSTY.RRF
et cru
ESTRUTURA DO CÓDIGO
LINHA DO TEMPO

PROBLEMAS SAÍDA CONSOLE DE DEBURAÇÃO TERMINAL bash - 01metawordindex +

```
Now suppressible:C0833036|POR|s|L16571354|PF|S20098781|Y|A32129839|||10042602|MDRBP|OL|10033350|PAC|3|0||  
Now suppressible:C0201539|POR|s|L0425264|PF|S0549546|Y|A32137947|||10001772|MDRBP|OL|10001454|AFP|3|0||  
Now suppressible:C0201836|POR|s|L4869994|PF|S5557501|Y|A32125699|||10001546|MDRBP|OL|10001844|ALT|3|0||  
Now suppressible:C0201925|POR|s|L4876456|PF|S5555085|Y|A32144137|||10005392|MDRBP|OL|10006850|Ca|3|0||  
Now suppressible:C0302358|POR|p|L3327051|PF|S3855485|Y|A32144231|||10054178|MDRBP|OL|10040545|Shigella dysenteriae|3|0||  
Now suppressible:C0302359|POR|p|L3327052|PF|S3855486|Y|A32199034|||10054178|MDRBP|OL|10040547|Shigella flexneri|3|0||  
Now suppressible:C0302360|POR|p|L3327050|PF|S3855484|Y|A32198860|||10054178|MDRBP|OL|10040543|Shigella boydii|3|0||  
Now suppressible:C0302361|POR|p|L3327053|PF|S3855488|Y|A32249520|||10054178|MDRBP|OL|10040551|Shigella sonnei|3|0||  
Now suppressible:C0337443|POR|s|L4809805|PF|S5507346|Y|A32198984|||10005799|MDRBP|OL|10028667|Na|3|0||  
Now suppressible:C0486222|POR|s|L12216542|PF|S15404627|Y|A30101111|||1064-5||LNC-PT-BR|OOSN|1064-5|I NOS Ag RBC BPU Q|0|0||  
Now suppressible:C0486223|POR|s|L12411083|PF|S15404628|Y|A30182778|||1065-2||LNC-PT-BR|OOSN|1065-2|I NOS Ag RBC Donr Q|0|0||  
Now suppressible:C0525032|POR|s|L4875491|PF|S5562996|Y|A32172414|||10022591|MDRBP|OL|10004190|BCR|3|0||  
Now suppressible:C1956346|POR|s|L10567245|PF|S20066200|Y|A32207000|||10011070|MDRBP|OL|10006896|CAD|3|0||
```

Suppression counts:
Unnecessarily ambiguous: 261
NEC: 01
Brand Name: 00
EC: 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.

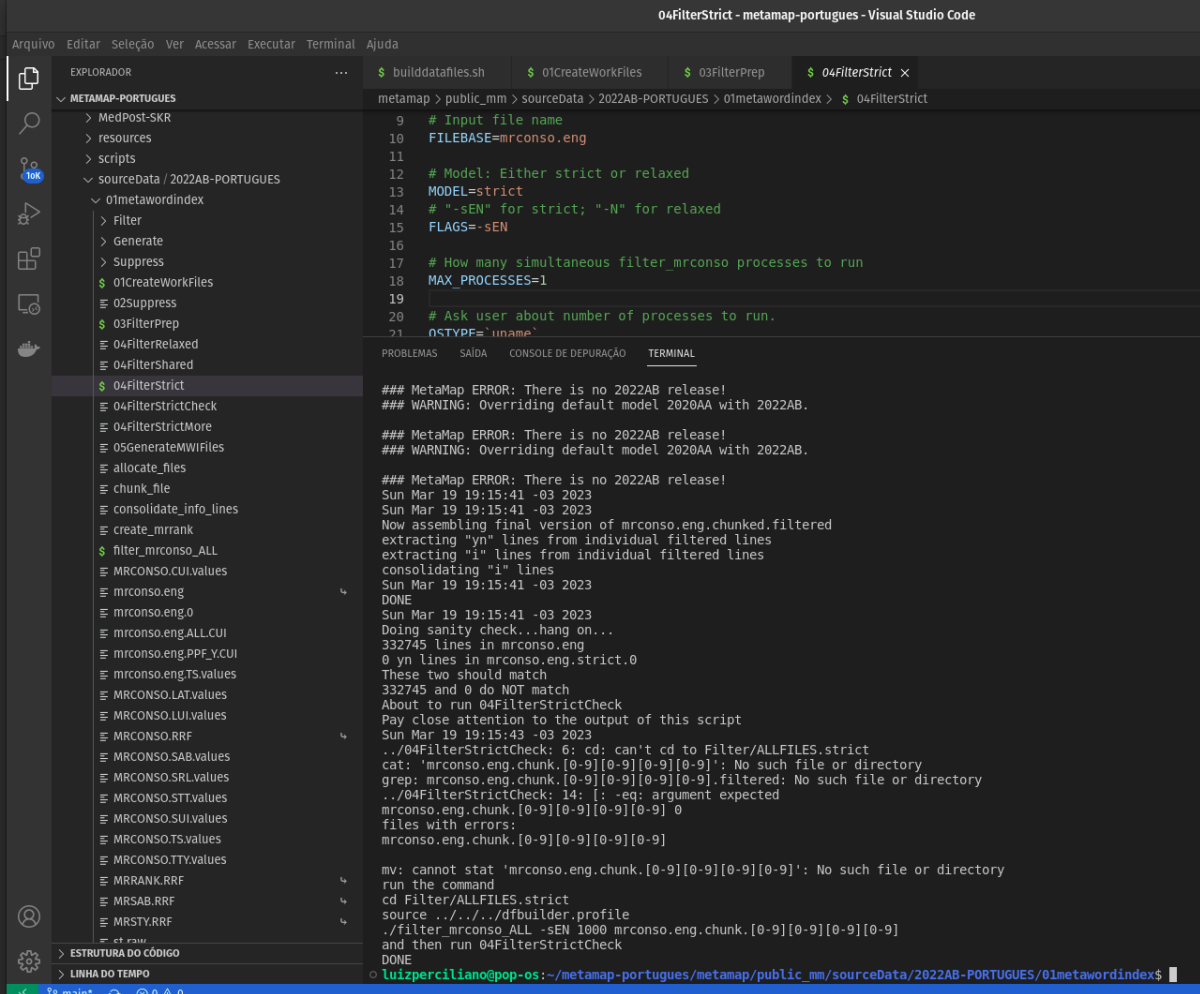
luizperciliano@pop-os:~/metaportugues/metaportugues/public_mm/sourcedata/2022AB-PORTUGUES/01metawordindex\$

./03FilterPrep

```
88  echo ' '
89
90
91  # gzip intermediate files
92
93  # gzip mrcon.eng.labels
94  # gzip mrcon.eng.lbld
95  # gzip mrcon.eng.lbld.s
96  # gzip mrcon.eng.n
97  # gzip mrso.labels
98  # gzip mrso.lbld
99  # gzip mrso.lbld.s
100
101  case $OSTYPE in
102  *MINGW*)
103      cp Filter/mrconso.eng mrconso.eng
104      ;;
105  *CYGWIN*)
106      cp Filter/mrconso.eng mrconso.eng
107      ;;
108  *)
109      ln -s Filter/mrconso.eng mrconso.eng
110      ;;
111  esac
112
113  date
114
115  cd ../
116
117  echo ' '
```

Deleted lines 101 until 111

./04FilterStrict



The screenshot shows the Visual Studio Code interface with the file explorer on the left, the script editor in the center, and the terminal at the bottom. The file explorer shows the project structure for '04FilterStrict'. The script editor displays the '04FilterStrict' script, which includes comments and variables for file names, models, and flags. The terminal shows the execution of the script, including error messages from MetaMap and the output of the '04FilterStrictCheck' command.

```
04FilterStrict - metaportugues - Visual Studio Code
Arquivo Editar Seleção Ver Acessar Executar Terminal Ajuda
EXPLORADOR
METAMAP-PORTUGUES
  MedPost-SKR
  resources
  scripts
  sourceData / 2022AB-PORTUGUES
    01metawordindex
      Filter
      Generate
      Suppress
    01CreateWorkFiles
    02Suppress
    03FilterPrep
    04FilterRelaxed
    04FilterShared
    04FilterStrict
    04FilterStrictCheck
    04FilterStrictMore
    05GenerateMWFiles
    allocate_files
    chunk_file
    consolidate_info_lines
    create_mrrank
    filter_mrconso_ALL
    MRCONSO.CUI.values
    mrconso.eng
    mrconso.eng.0
    mrconso.eng.ALL.CUI
    mrconso.eng.PPF.Y.CUI
    mrconso.eng.TS.values
    MRCONSO.LAT.values
    MRCONSO.LUI.values
    MRCONSO.RRF
    MRCONSO.SAB.values
    MRCONSO.SRL.values
    MRCONSO.STT.values
    MRCONSO.SUI.values
    MRCONSO.TS.values
    MRCONSO.TTY.values
    MRRANK.RRF
    MRSAB.RRF
    MRSTY.RRF
    et raw
  ESTRUTURA DO CÓDIGO
  LINHA DO TEMPO
$ builddatafiles.sh $ 01CreateWorkFiles $ 03FilterPrep $ 04FilterStrict x
metamap > public_mm > sourceData > 2022AB-PORTUGUES > 01metawordindex > $ 04FilterStrict
9 # Input file name
10 FILEBASE=mrconso.eng
11
12 # Model: Either strict or relaxed
13 MODEL=strict
14 # "-sEN" for strict; "-N" for relaxed
15 FLAGS=-sEN
16
17 # How many simultaneous filter_mrconso processes to run
18 MAX_PROCESSES=1
19
20 # Ask user about number of processes to run.
21 NSTYPE='uname'
PROBLEMAS SAÍDA CONSOLE DE DEBURAÇÃO TERMINAL
### MetaMap ERROR: There is no 2022AB release!
### WARNING: Overriding default model 2020AA with 2022AB.
### MetaMap ERROR: There is no 2022AB release!
### WARNING: Overriding default model 2020AA with 2022AB.
### MetaMap ERROR: There is no 2022AB release!
Sun Mar 19 19:15:41 -03 2023
Sun Mar 19 19:15:41 -03 2023
Now assembling final version of mrconso.eng.chunked.filtered
extracting "yn" lines from individual filtered lines
extracting "i" lines from individual filtered lines
consolidating "i" lines
Sun Mar 19 19:15:41 -03 2023
DONE
Sun Mar 19 19:15:41 -03 2023
Doing sanity check...hang on...
332745 lines in mrconso.eng
0 yn lines in mrconso.eng.strict.0
These two should match
332745 and 0 do NOT match
About to run 04FilterStrictCheck
Pay close attention to the output of this script
Sun Mar 19 19:15:43 -03 2023
../04FilterStrictCheck: 6: cd: can't cd to Filter/ALLFILES.strict
cat: 'mrconso.eng.chunk.[0-9][0-9][0-9][0-9]': No such file or directory
grep: mrconso.eng.chunk.[0-9][0-9][0-9][0-9].filtered: No such file or directory
../04FilterStrictCheck: 14: [: -eq: argument expected
mrconso.eng.chunk.[0-9][0-9][0-9][0-9] 0
files with errors:
mrconso.eng.chunk.[0-9][0-9][0-9][0-9]
mv: cannot stat 'mrconso.eng.chunk.[0-9][0-9][0-9][0-9]': No such file or directory
run the command
cd Filter/ALLFILES.strict
source ../../../../dfbuilder.profile
./filter_mrconso_ALL -sEN 1000 mrconso.eng.chunk.[0-9][0-9][0-9][0-9]
and then run 04FilterStrictCheck
DONE
luizperciliano@pop-os:~/metamap-portugues/metamap/public_mm/sourceData/2022AB-PORTUGUES/01metawordindex$
```

MetaMap ERROR: There is no 2022AB release!

Sun Mar 19 19:15:41 -03 2023

Sun Mar 19 19:15:41 -03 2023

Now assembling final version of mrconso.eng.chunked.filtered

extracting "yn" lines from individual filtered lines

extracting "i" lines from individual filtered lines

consolidating "i" lines

Sun Mar 19 19:15:41 -03 2023

DONE

Sun Mar 19 19:15:41 -03 2023

Doing sanity check...hang on...

332745 lines in mrconso.eng

0 yn lines in mrconso.eng.strict.0

These two should match

332745 and 0 do NOT match

About to run 04FilterStrictCheck

Pay close attention to the output of this script

Sun Mar 19 19:15:43 -03 2023


```
../04FilterStrictCheck: 6: cd: can't cd to Filter/ALLFILES.strict
cat: 'mrconso.eng.chunk.[0-9][0-9][0-9][0-9]': No such file or directory
grep: mrconso.eng.chunk.[0-9][0-9][0-9][0-9].filtered: No such file or directory
../04FilterStrictCheck: 14: [: -eq: argument expected
mrconso.eng.chunk.[0-9][0-9][0-9][0-9] 0
files with errors:
mrconso.eng.chunk.[0-9][0-9][0-9][0-9]
```

```
mv: cannot stat 'mrconso.eng.chunk.[0-9][0-9][0-9][0-9]': No such file or directory
run the command
cd Filter/ALLFILES.strict
source ../../../../dfbuilder.profile
./filter_mrconso_ALL -sEN 1000 mrconso.eng.chunk.[0-9][0-9][0-9][0-9]
and then run 04FilterStrictCheck
DONE
```

Knowing the mistake, I continued the steps anyway.

Traduzido:

```
### MetaMap ERROR: Não há versão 2022AB!
Agora montando a versão final de mrconso.eng.chunked.filtered
extraindo linhas "yn" de linhas filtradas individuais
extraindo linhas "i" de linhas filtradas individuais
consolidando linhas "i"
Dom 19 de março 19:15:41 -03 2023
FEITO
Dom 19 de março 19:15:41 -03 2023
Fazendo um teste de sanidade... espere...
332745 linhas em mrconso.eng
0 yn linhas em mrconso.eng.strict.0
Esses dois devem corresponder
332745 e 0 NÃO correspondem
Prestes a executar 04FilterStrictCheck
Preste muita atenção à saída deste script
Dom 19 de março 19:15:43 -03 2023
../04FilterStrictCheck: 6: cd: não é possível fazer cd para Filter/ALLFILES.strict
cat: 'mrconso.eng.chunk.[0-9][0-9][0-9][0-9]': Arquivo ou diretório inexistente
grep: mrconso.eng.chunk.[0-9][0-9][0-9][0-9].filtered: Arquivo ou diretório inexistente
../04FilterStrictCheck: 14: [: -eq: argumento esperado
mrconso.eng.chunk.[0-9][0-9][0-9][0-9] 0
arquivos com erros:
mrconso.eng.chunk.[0-9][0-9][0-9][0-9]

mv: não é possível declarar 'mrconso.eng.chunk.[0-9][0-9][0-9][0-9]': Arquivo ou diretório
inexistente
execute o comando
cd Filter/ALLFILES.strict
```

fonte ../../dfbuilder.profile

./filter_mrconso_ALL -sEN 1000 mrconso.eng.chunk.[0-9][0-9][0-9][0-9]

e então execute 04FilterStrictCheck

FEITO
