Instalação, configuração e teste

Metamap DFB 2021

Tester: Perciliano

Test: nº 40

**Sumário**

[1. Instalação Metamap atual 3](#_Toc81826550)

[1.1 Instalar léxico 3](#_Toc81826551)

[1.2 Instalar Metamap 3](#_Toc81826552)

[1.3 Instalar UMSL 2021 3](#_Toc81826553)

[1.4 Metamorphosys 4](#_Toc81826554)

[2. Pré-requisitos antes da execução do test 5](#_Toc81826555)

[3. Dados do Test para criação de bases em Português do Brasil 6](#_Toc81826556)

[4. Passo a passo 6](#_Toc81826557)

[5. Referências 20](#_Toc81826558)

# **Instalação Metamap atual**

Nova instalação completa do Metamap feita em 30-08-2021.

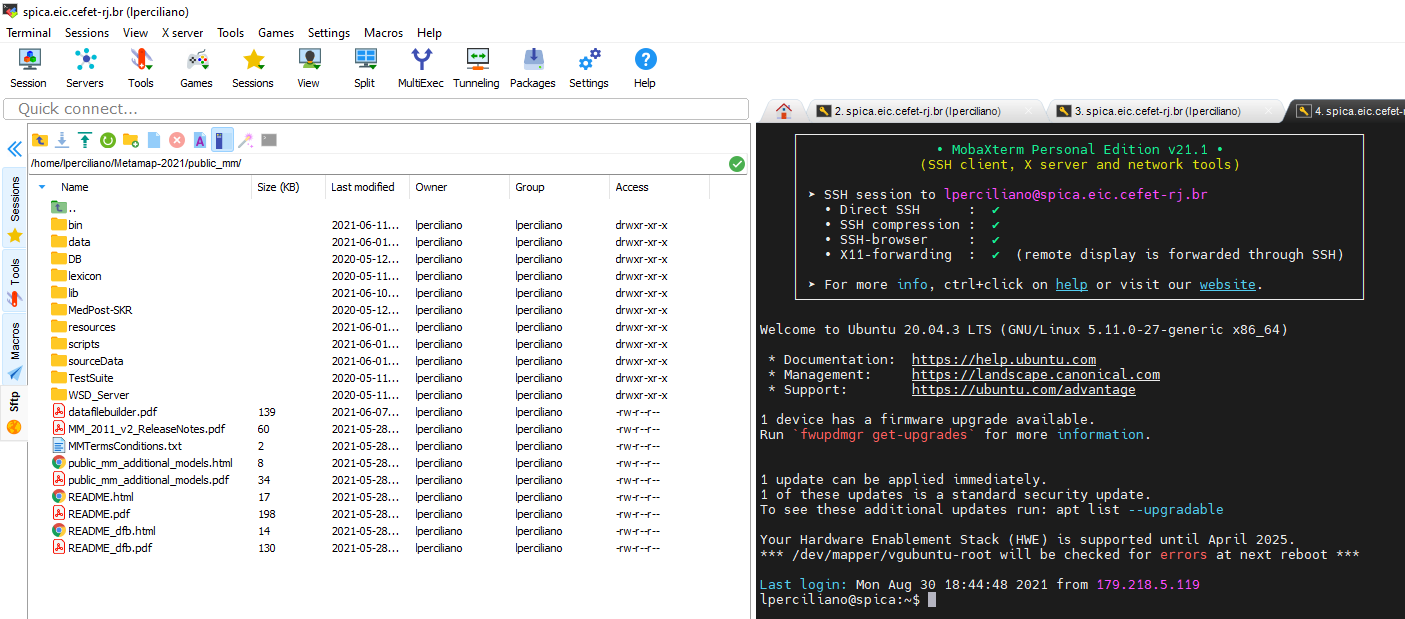
## Instalar léxico

tar -vzxf ../arquivos-metamap/lvg2016.tgz

## Instalar Metamap

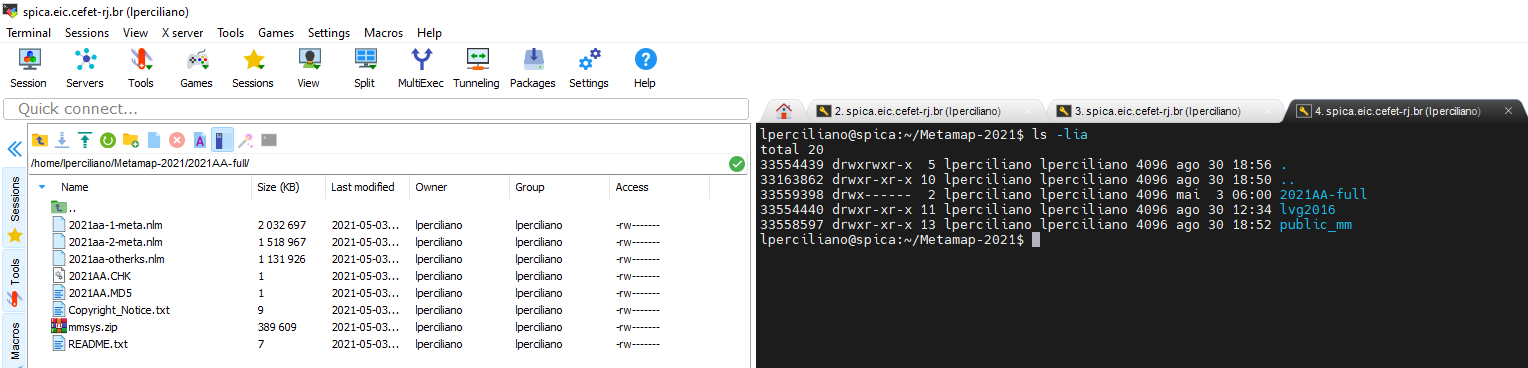
tar -jxvf ../arquivos-metamap/public\_mm\_linux\_main\_2020.tar.bz2

tar -jxvf ../arquivos-metamap/public\_mm\_linux\_dfb\_2021.tar.bz2

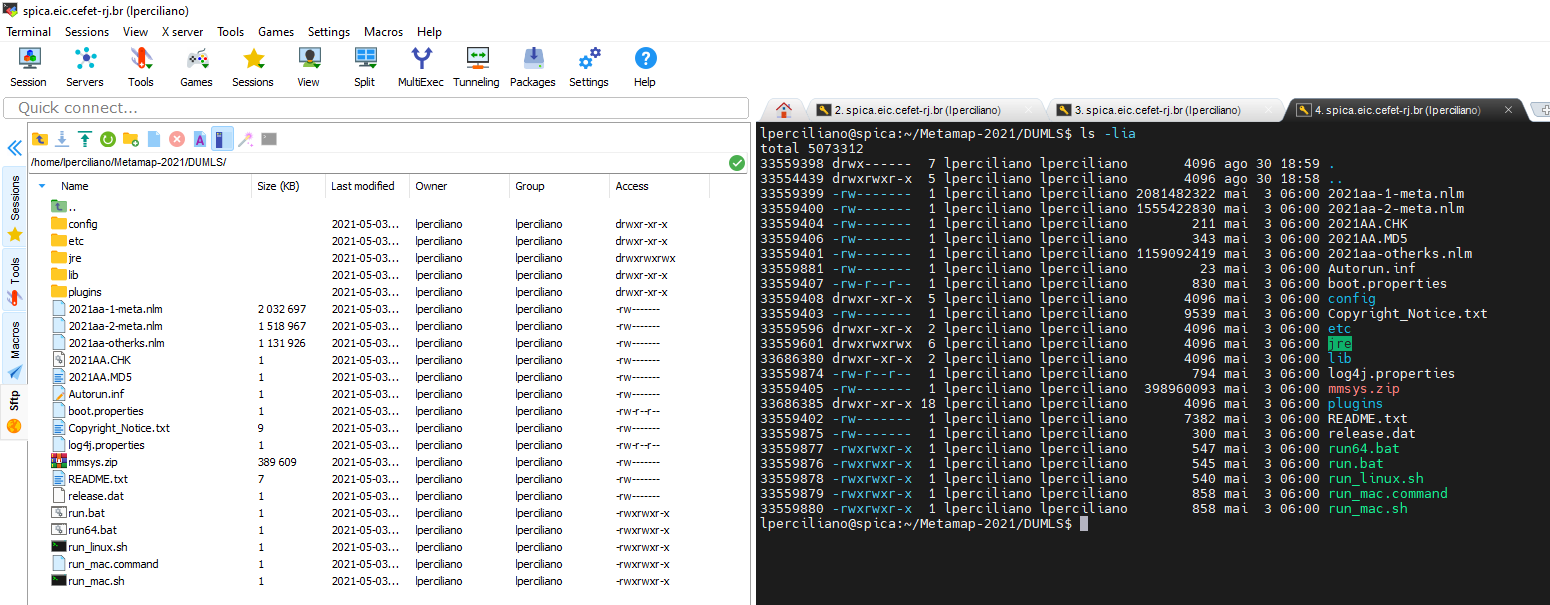


## Instalar UMSL 2021

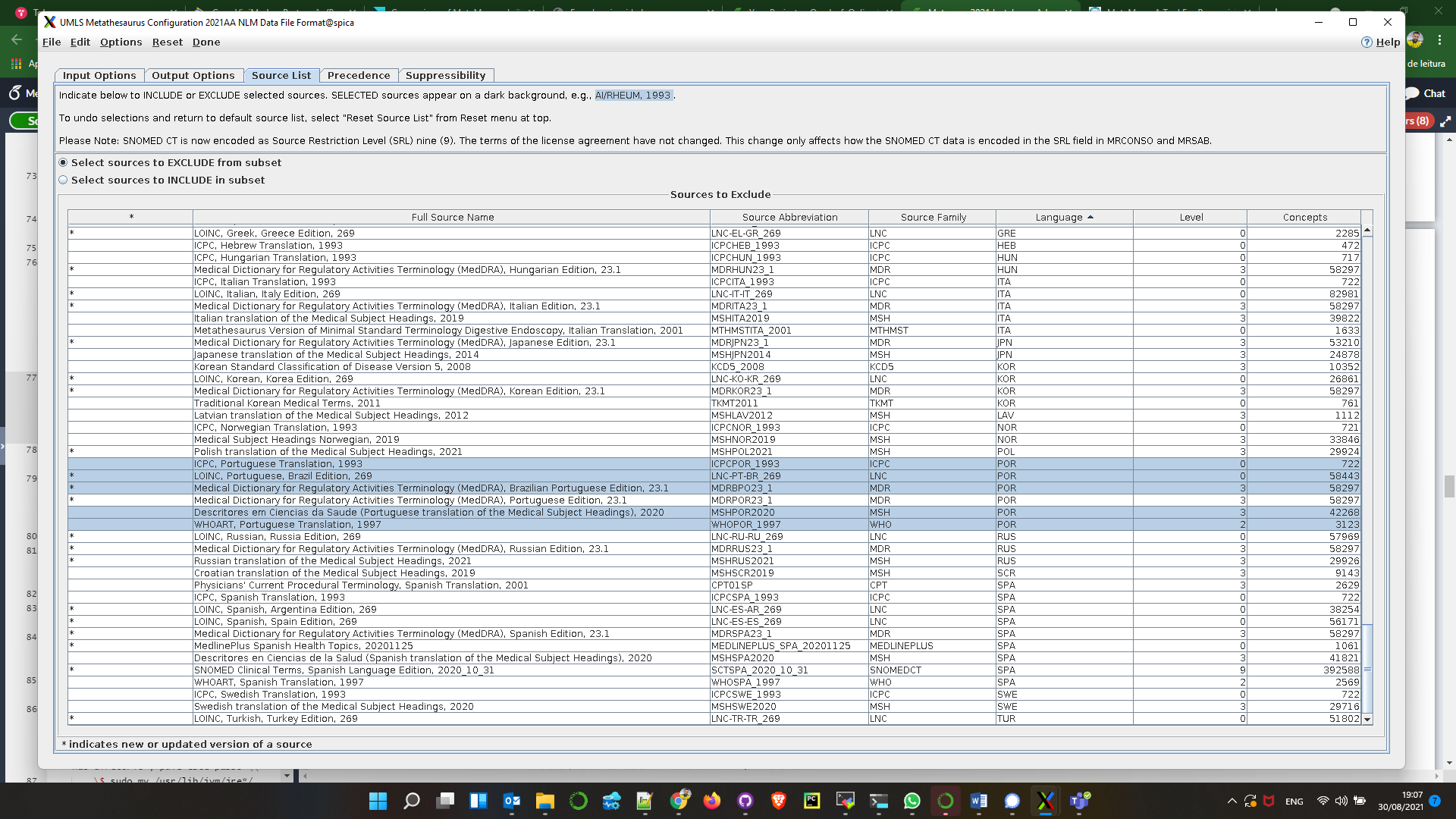
unzip ../arquivos-metamap/umls-2021AA-full.zip



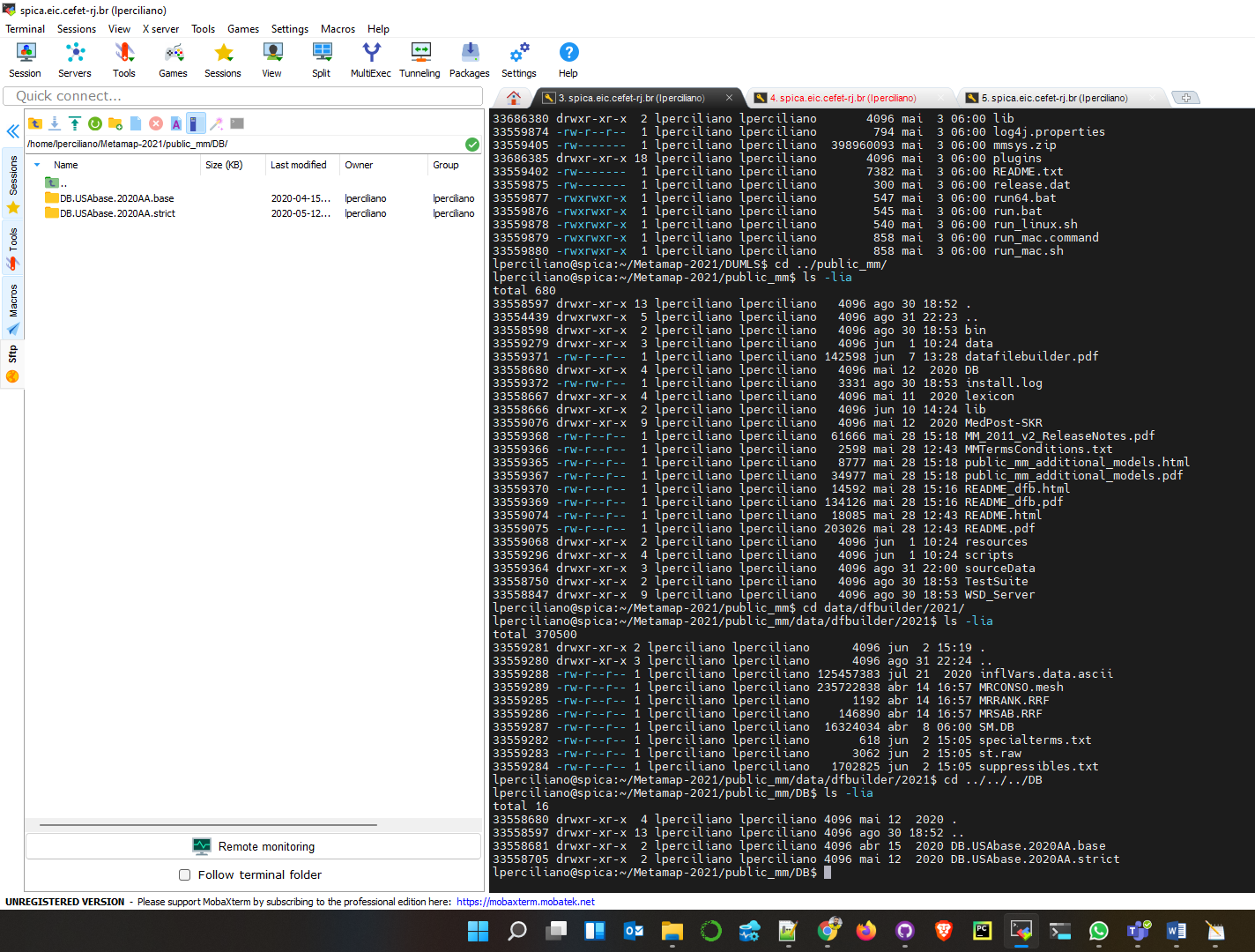
-



## Metamorphosys



Antes da instalação da base em português do Brasil, já existe as bases em US.



# **Pré-requisitos antes da execução do test**

A nova versão dfb 2021 do arquivo "filter\_mrconso" está diferente, não alterei nada por enquanto.

No diretório public\_mm/data/dfbuilder/202, executar o script abaixo:

* 1. cd ln -s ../2016/\* .k

No arquivo "/public\_mm/bin/builddatafiles.sh" trocar a linha 498 para:

* 1. echo "# NOTE: MODELOPTION is currently not supported in filter\_mrconso and mm\_variants" >> $WORKSPACE/dfbuilder.profile

ou ainda pode-se remover a linha ou ainda colocar # no início da linha

Linha 486 para (verificar se não é a mesma coisa acima ou algo a ser trocado):

* 1. echo "NOTE: MODELOPTION is currently not supported in filter\_mrconso and mm\_variants" >> $WORKSPACE/dfbuilder.profile

No arquivo public\_mm/bin/BuildDatafiles alterar respectivamente a 1 e 7 para:

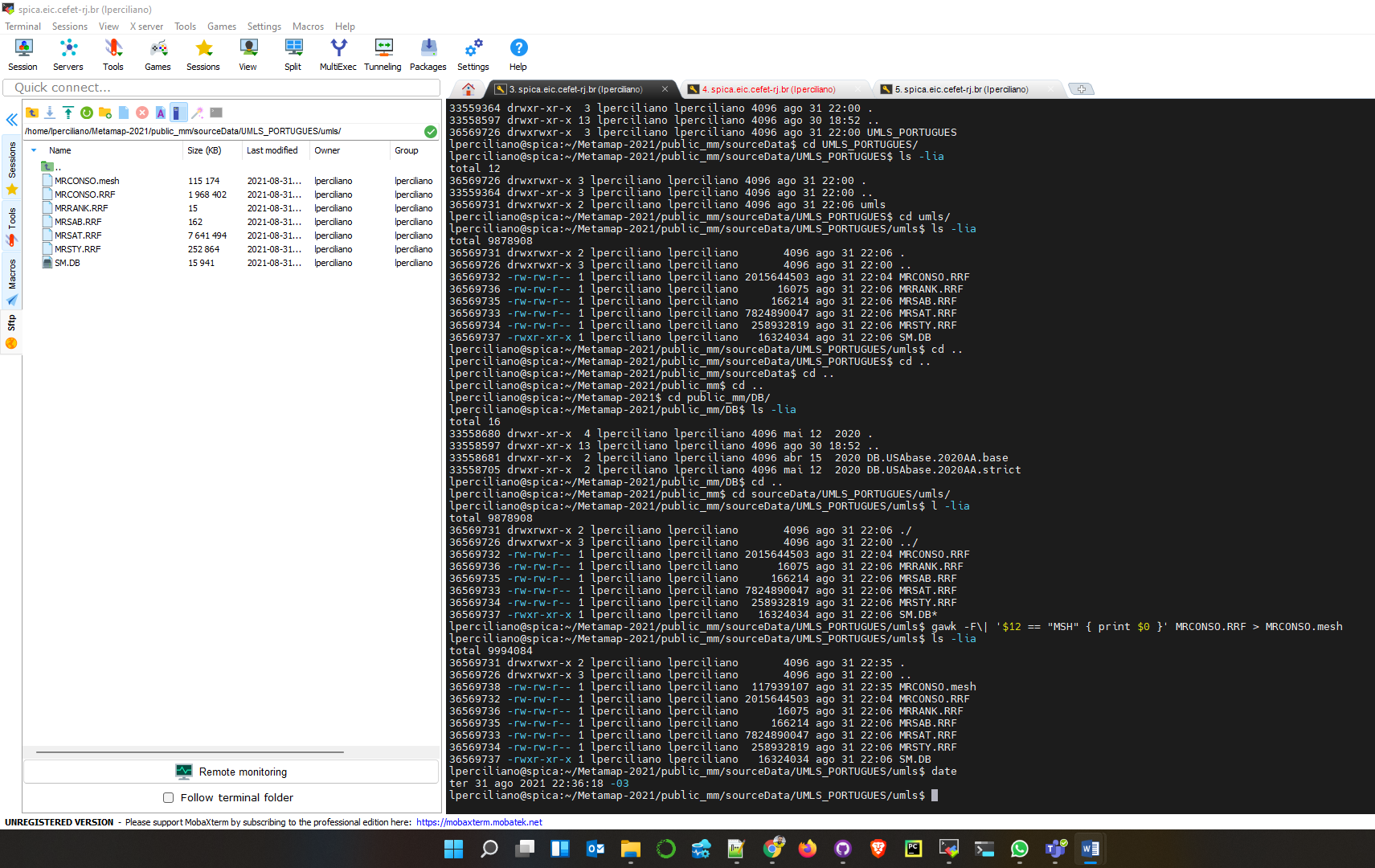
* 1. #!/bin/bash
  2. /bin/bash ${BASEDIR}/bin/builddatafiles.sh $\*

Verificar se a variável do LVG está criada

* 1. printenv | grep LVG\_2016

No diretório public\_mm/sourceData/UMLS\_PORTUGUES/umls gerar outro arquivo MRCONSO.mesh com o comando abaixo:

* 1. gawk -F\| '$12 == "MSH" { print $0 }' MRCONSO.RRF > MRCONSO.mesh



# **Dados do Test para criação de bases em Português do Brasil**

Teste realizado em 31-08-2021

Início em 22:56h até 01/09/2021 00:28h

Sucesso sim ( ) não(x)

OBS.: No teste 36 criou as duas bases.

# **Passo a passo**

Passos executados.

* 1. Criar arquivo de log para guardar os resultados gerados com as execuções

bash | tee log\_test40.txt

date

* 1. Verificar diretórios

cd Metamap-2021/

ls –lia

cd DUMLS/

ls –lia

cd ../public\_mm/

ls –lia

cd data/dfbuilder/2021/

ls –lia

cd ../../../DB

ls –lia

cd ../sourceData

ls –lia

cd UMLS\_PORTUGUES (caso tenha arquivos. Para nova instalação tem que apagar estes em comando mais abaixo)

ls –lia

cd umls

ls –lia

* 1. Excluir as bases geradas em português do diretório /Metamap-2021/public\_mm/DB/

ls –lia

rm -rf "DB.UMLS\_PORTUGUES.2020AA.base" (verificar se tem q excluir essa mesma ou se ela é sobreposta – no passo 36 de sucesso não foi feito isso)

* 1. Excluir diretórios do /public\_mm/sourceData/UMLS\_PORTUGUES/

ls –lia

rm -rf "01metawordindex" "02treecodes" "03variants" "04synonyms" "05abbrAcronyms" "dfbuilder.profile"

Iniciar instalação

* 1. cd Metamap-2021/public\_mm
  2. ./bin/skrmedpostctl start
  3. ./bin/BuildDataFiles
  4. UMLS\_PORTUGUES e digite enter

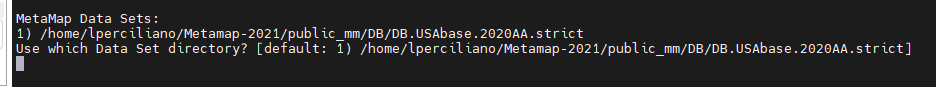
Na pergunta: Workspace will be placed in /home/lperciliano/Metamap-2021/public\_mm/sourceData/UMLS\_PORTUGUES, is this okay? digitar

* 1. yes e enter
  2. Na pergunta:

MetaMap Data Sets: 1) /home/lperciliano/Metamap-2021/public\_mm/DB/DB.USAbase.2020AA.strict

Use which Data Set directory? [default: 1) /home/lperciliano/Metamap-2021/public\_mm/DB/DB.USAbase.2020AA.strict] - usar o default:

dá enter ????? Aqui tem colocar o meu?

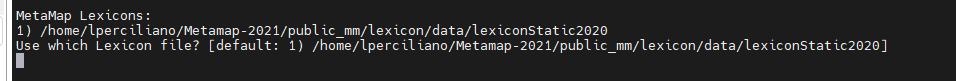


* 1. Na pergunta:

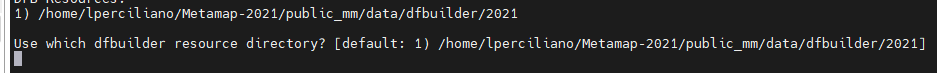
MetaMap Lexicons: 1) /home/lperciliano/Metamap-2021/public\_mm/lexicon/data/lexiconStatic2020

Use which Lexicon file? [default: 1) /home/lperciliano/Metamap-2021/public\_mm/lexicon/data/lexiconStatic2020]

use o default: digite enter



* 1. Na pergunta:



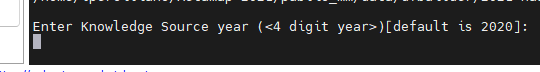
1) /home/lperciliano/Metamap-2021/public\_mm/data/dfbuilder/2021

Use which dfbuilder resource directory? [default: 1) /home/lperciliano/Metamap-2021/public\_mm/data/dfbuilder/2021]

Escolha dfbuilder dir default: digite enter (4)

????? Aqui tem colocar o meu?

* 1. Na pergunta:



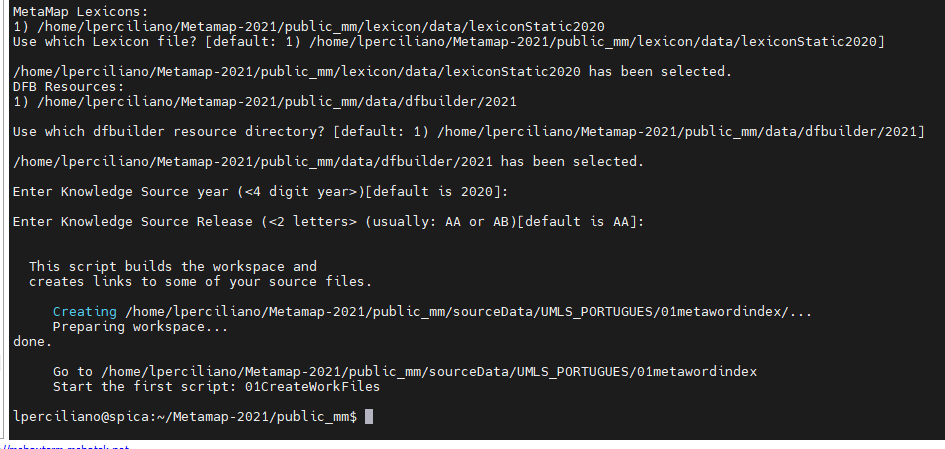
Enter Knowledge Source year (<4 digit year>)[default is 2020]:

???? 2020 ou 2021 escolha o ano do source, default: digite enter

* 1. Na pergunta:

Enter Knowledge Source Release (<2 letters> (usually: AA or AB)[default is AA]:

p/ default: digite enter



Ir no arquivo /UMLS\_PORTUGUES/dfbuilder.profile e trocar a primeira linha para

* 1. #!/bin/bash

Ir no diretório abaixo e deletar o arquivo "01CreateWorkFiles"

* 1. cd sourceData/UMLS\_PORTUGUES/01metawordindex/
  2. ls –lia
  3. rm -rf "01CreateWorkFiles"
  4. ls –lia

Copiar novo arquivo "01CreateWorkFiles" (este arquivo tem q estar anteriormente com permissao 777)

* 1. cp /home/lperciliano/arquivos-metamap/01CreateWorkFiles 01CreateWorkFiles
  2. ls –lia
  3. Alterar a primeira linha dos seguintes arquivos com: #!/bin/bash

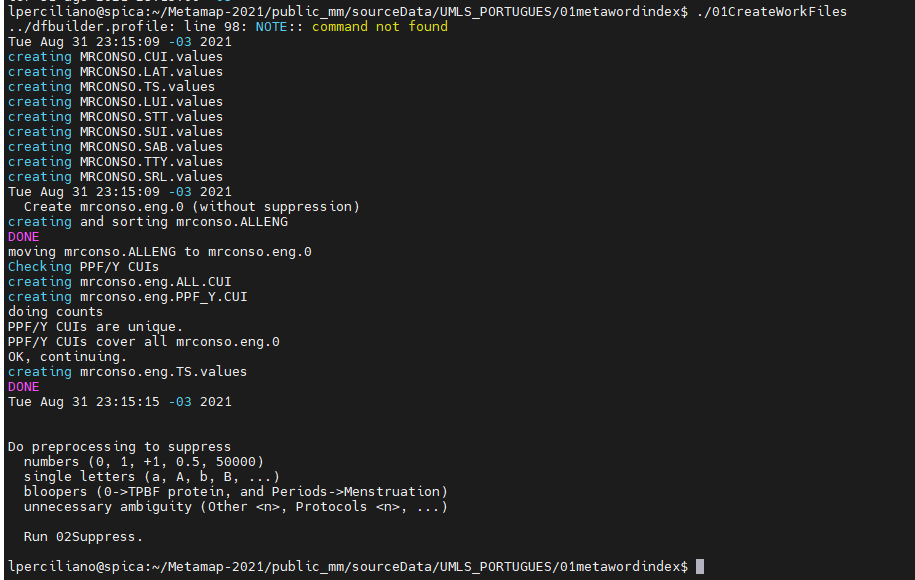
02Suppress

03FilterPrep

05GenerateMWIFiles

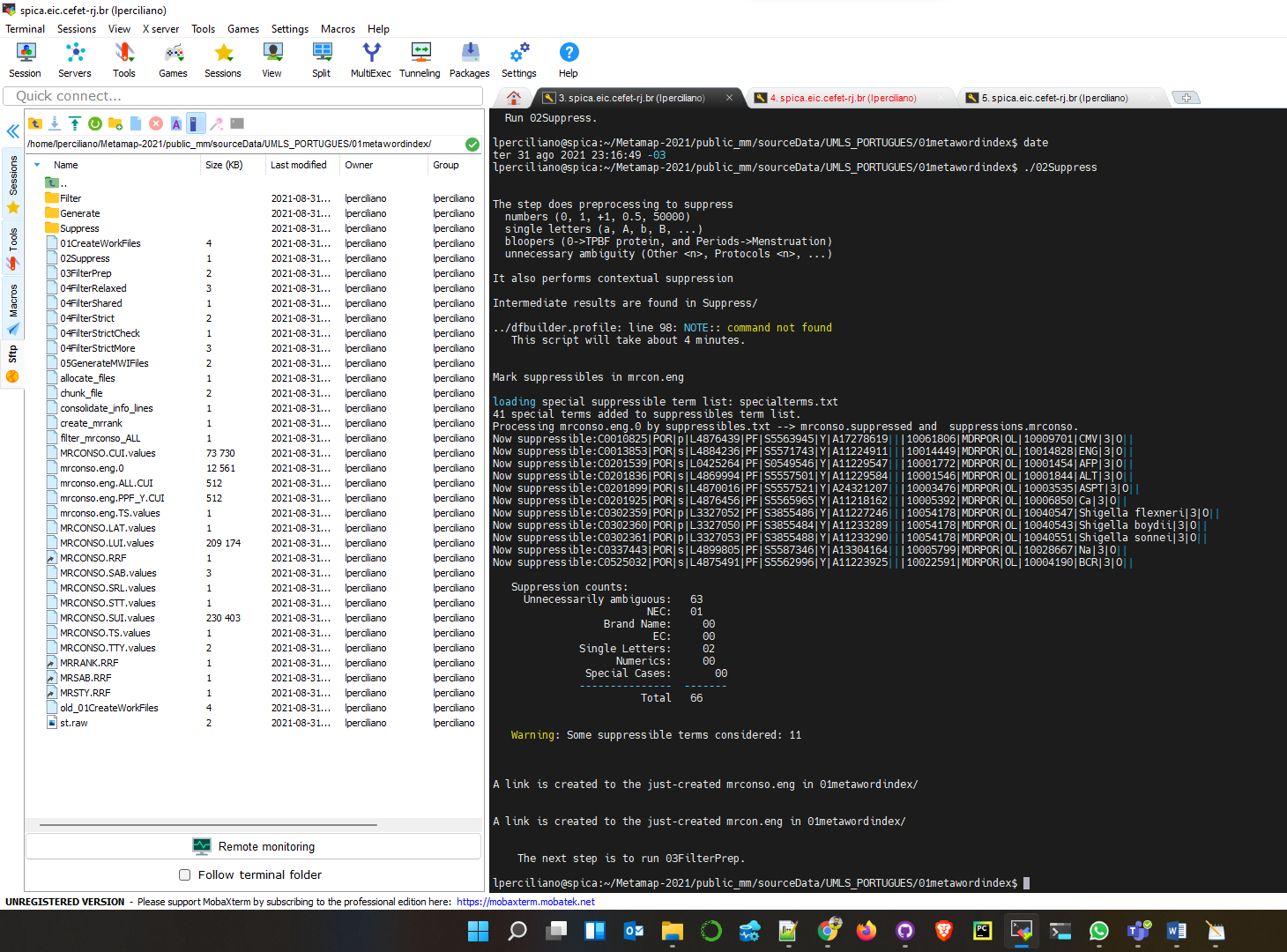
* 1. Executar o arquivo

./01CreateWorkFiles



* 1. Executar o arquivo

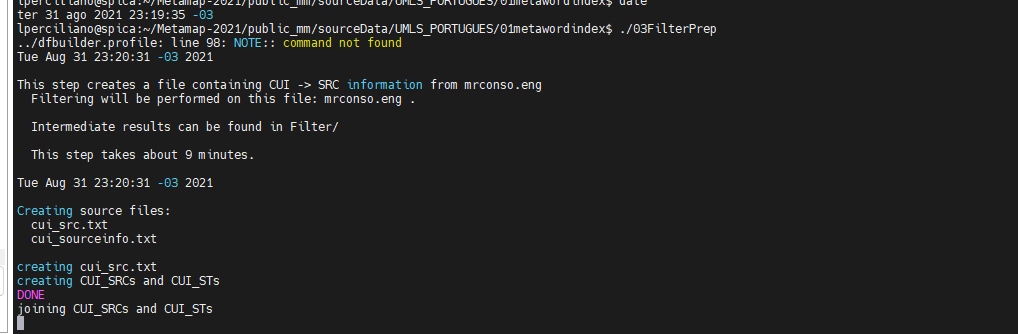
./02Suppress



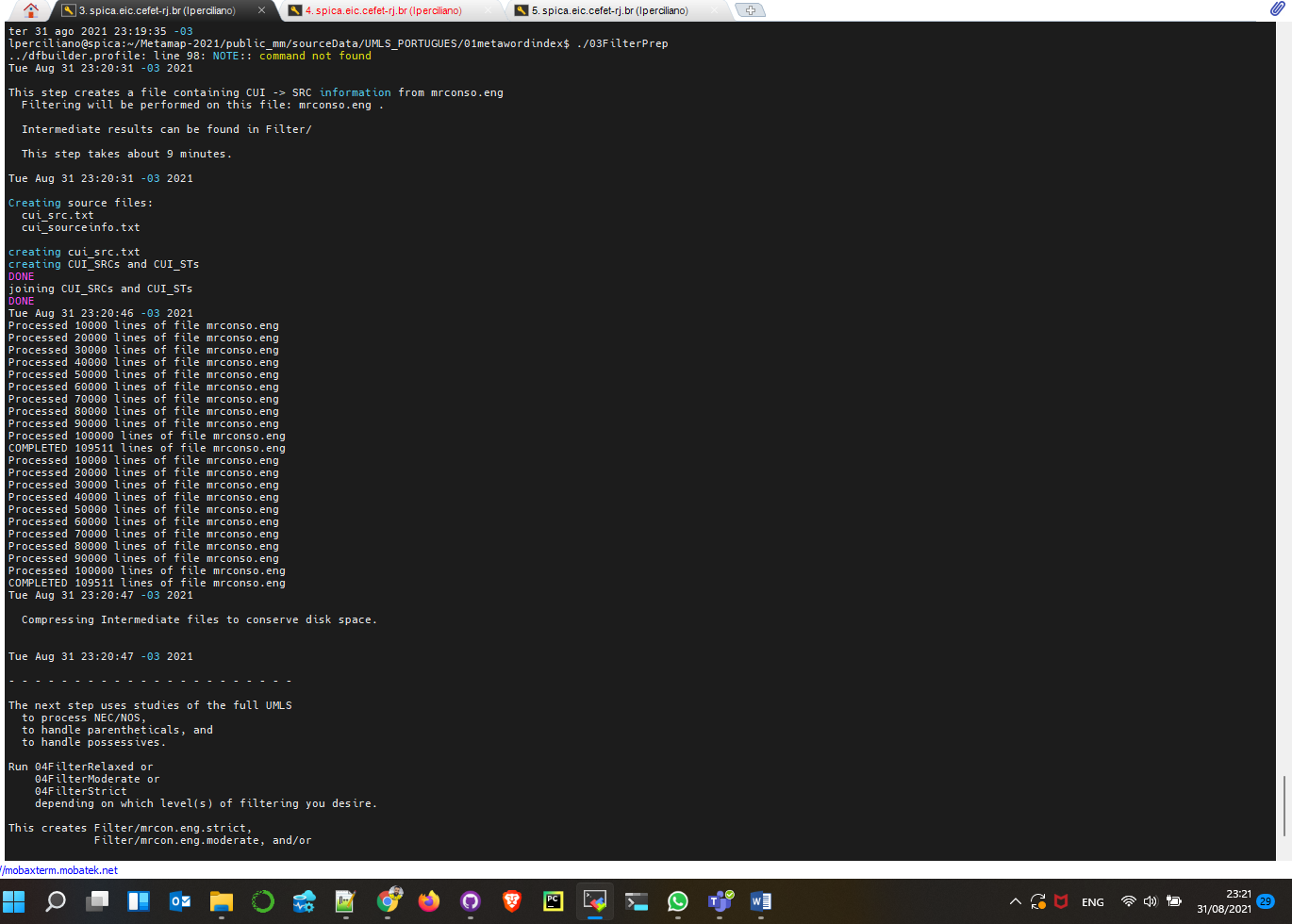
O resultado aqui segundo o Will deveria ser total de xxxx

* 1. Acessar o ./03FilterPrep e apagar da linha 101 a 111

./03FilterPrep



--



* 1. Executar o arquivo

./04FilterStrict

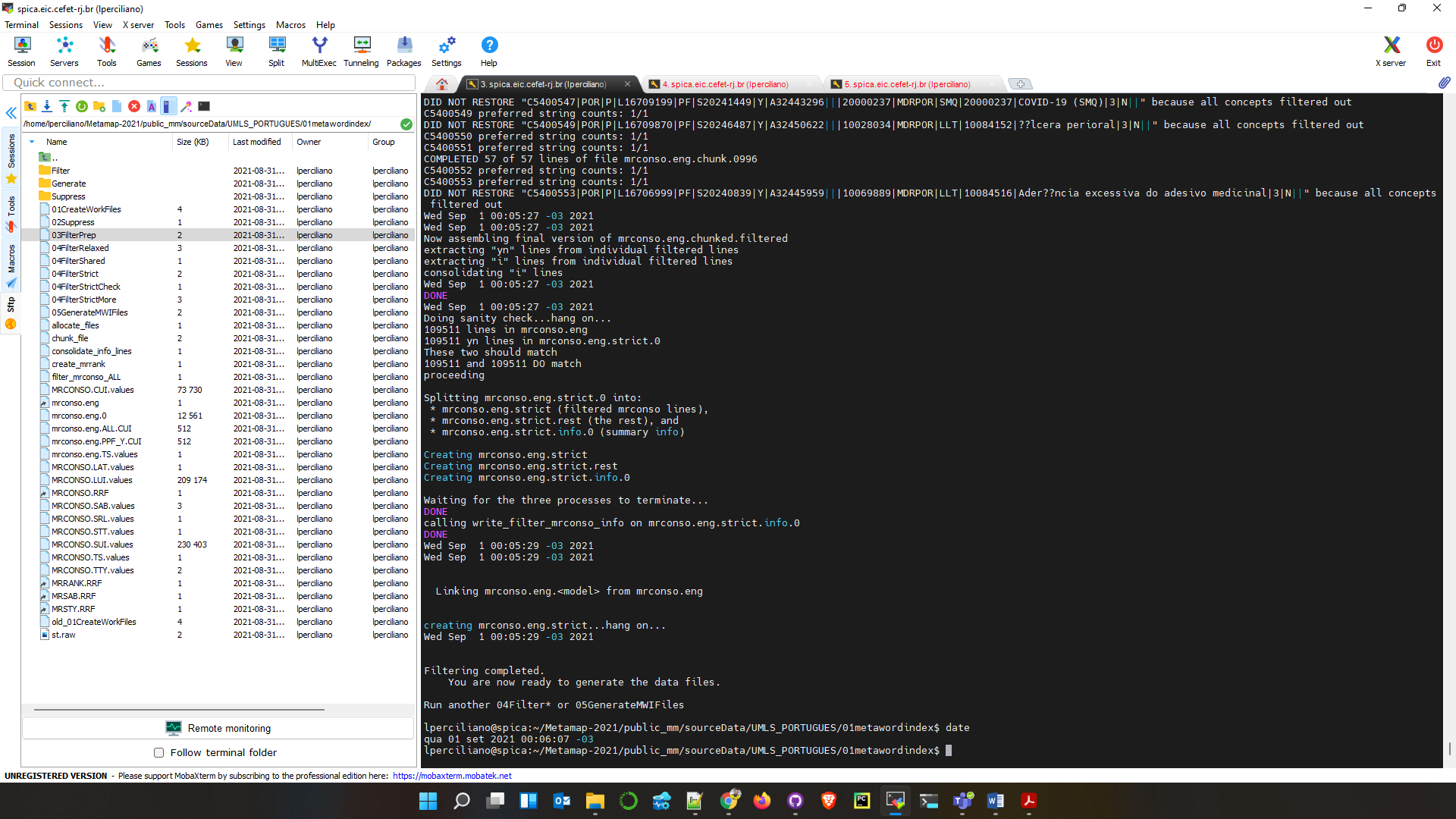
yes

enter

2

enter

Este demora demora +- 0:30h

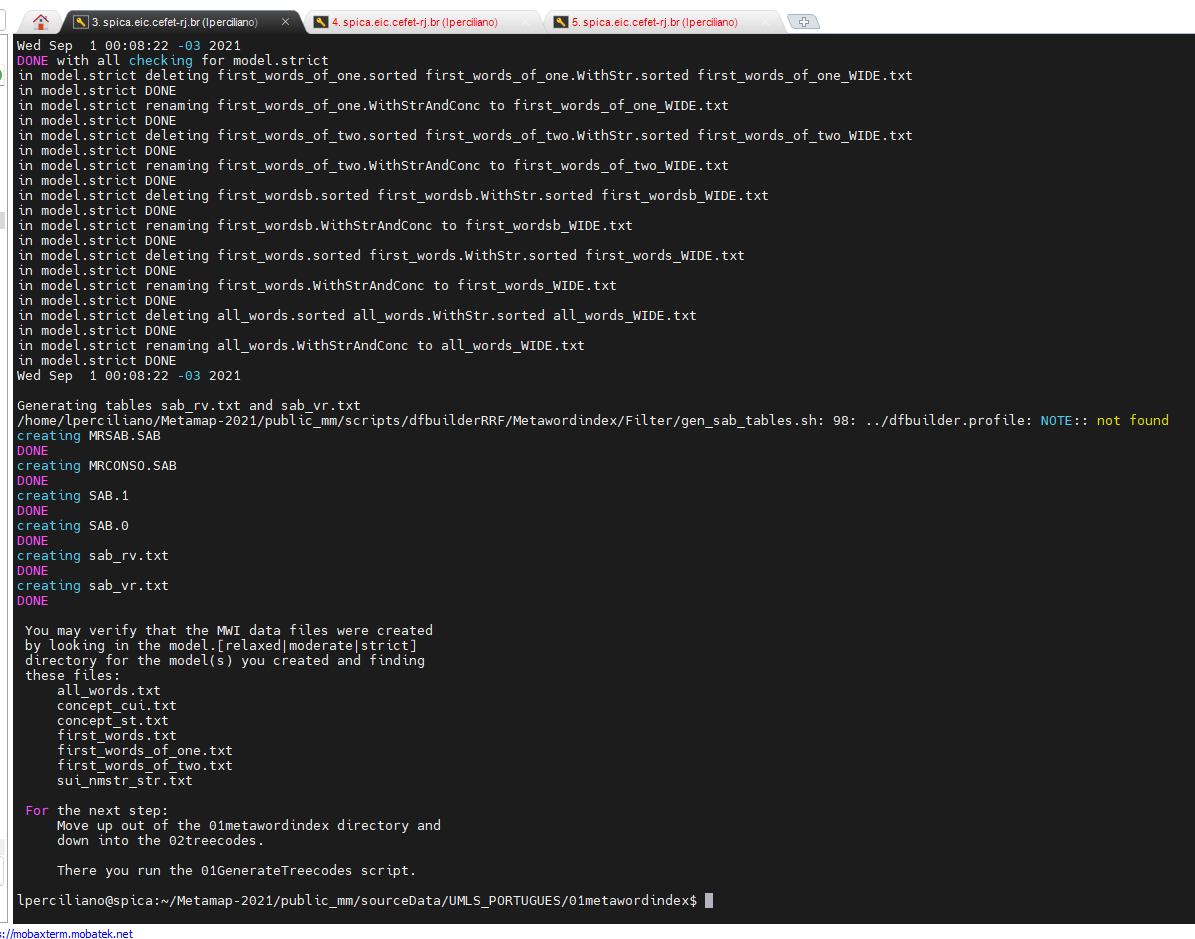


Executar o arquivo

* 1. cd sourceData/UMLS\_PORTUGUES/01metawordindex/
  2. ls –lia
  3. ./05GenerateMWIFiles (demora uns 2 minutos)

gera alguns erros??? verificar???

ls -lia



Acessar o diretório

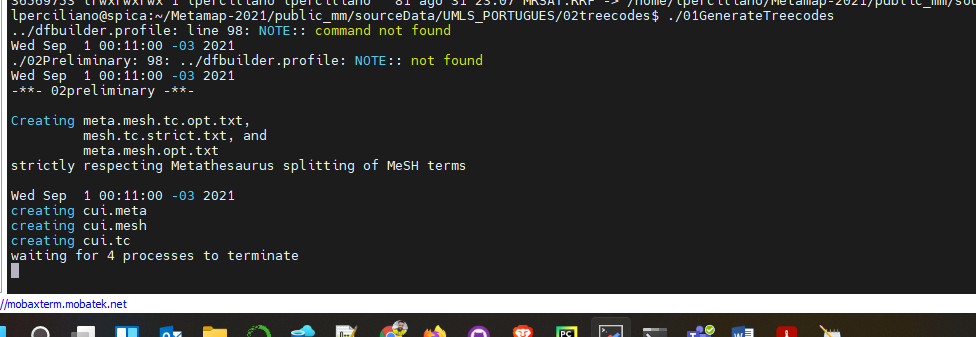
* 1. cd ../02treecodes/

Alterar a primeira linha do 01GenerateTreecodes para: #!/bin/bash

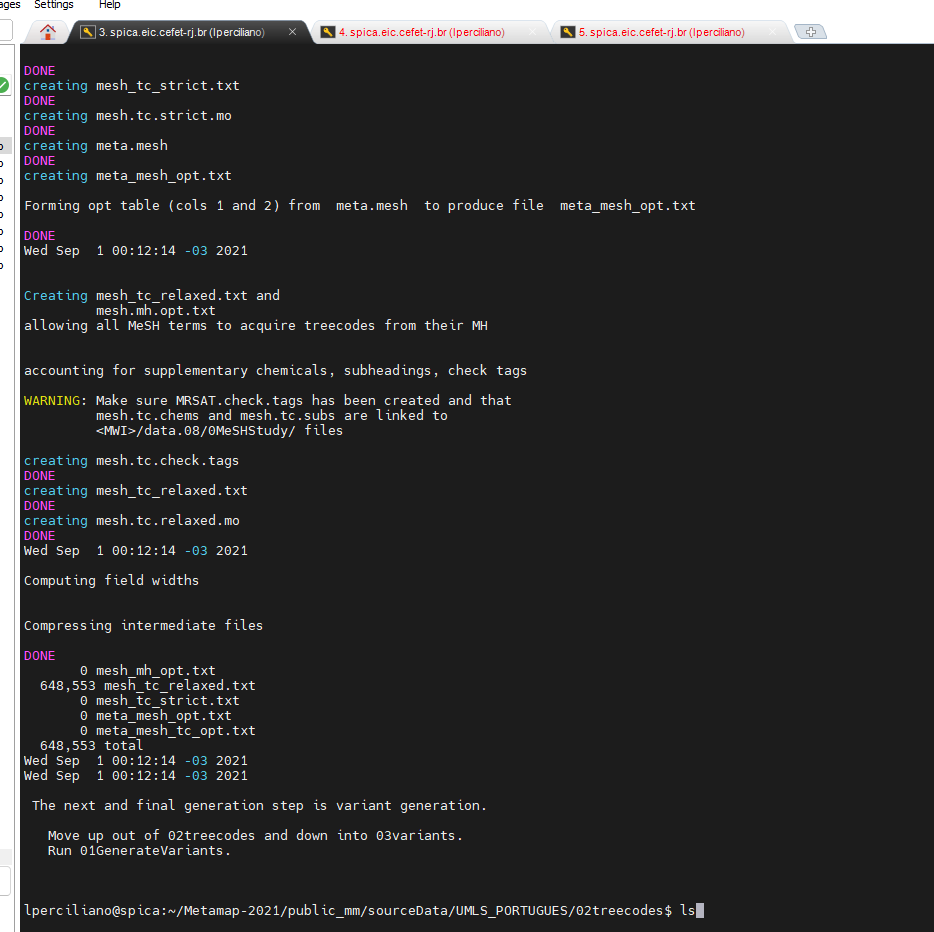
* 1. ./01GenerateTreecodes

gera alguns erros

ls –lia



-



* 1. Acessar

cd ../03variants/

Alterar a primeira linha do arquivo ./01GenerateVariants para: #!/bin/bash

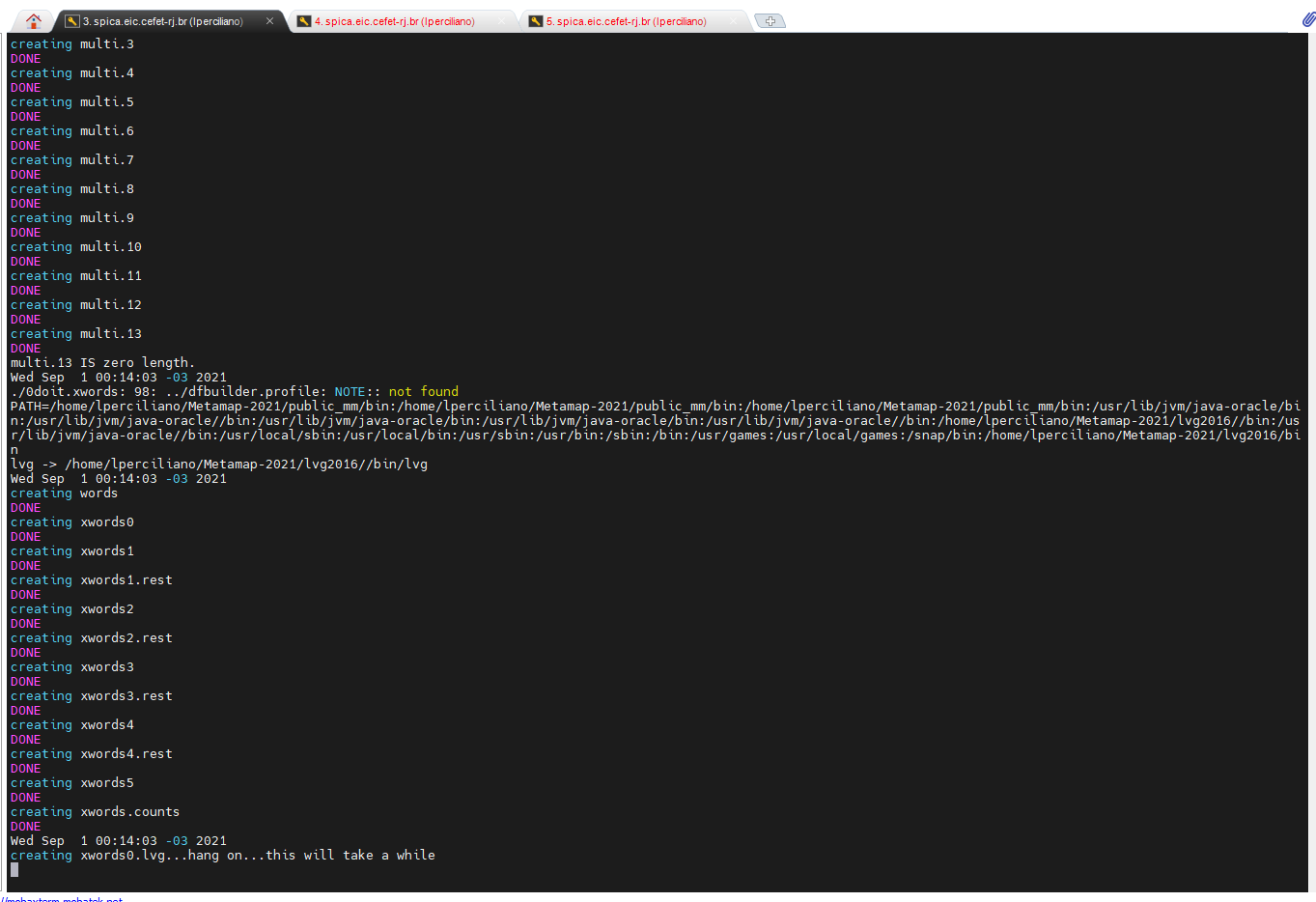
ls -lia

./01GenerateVariants

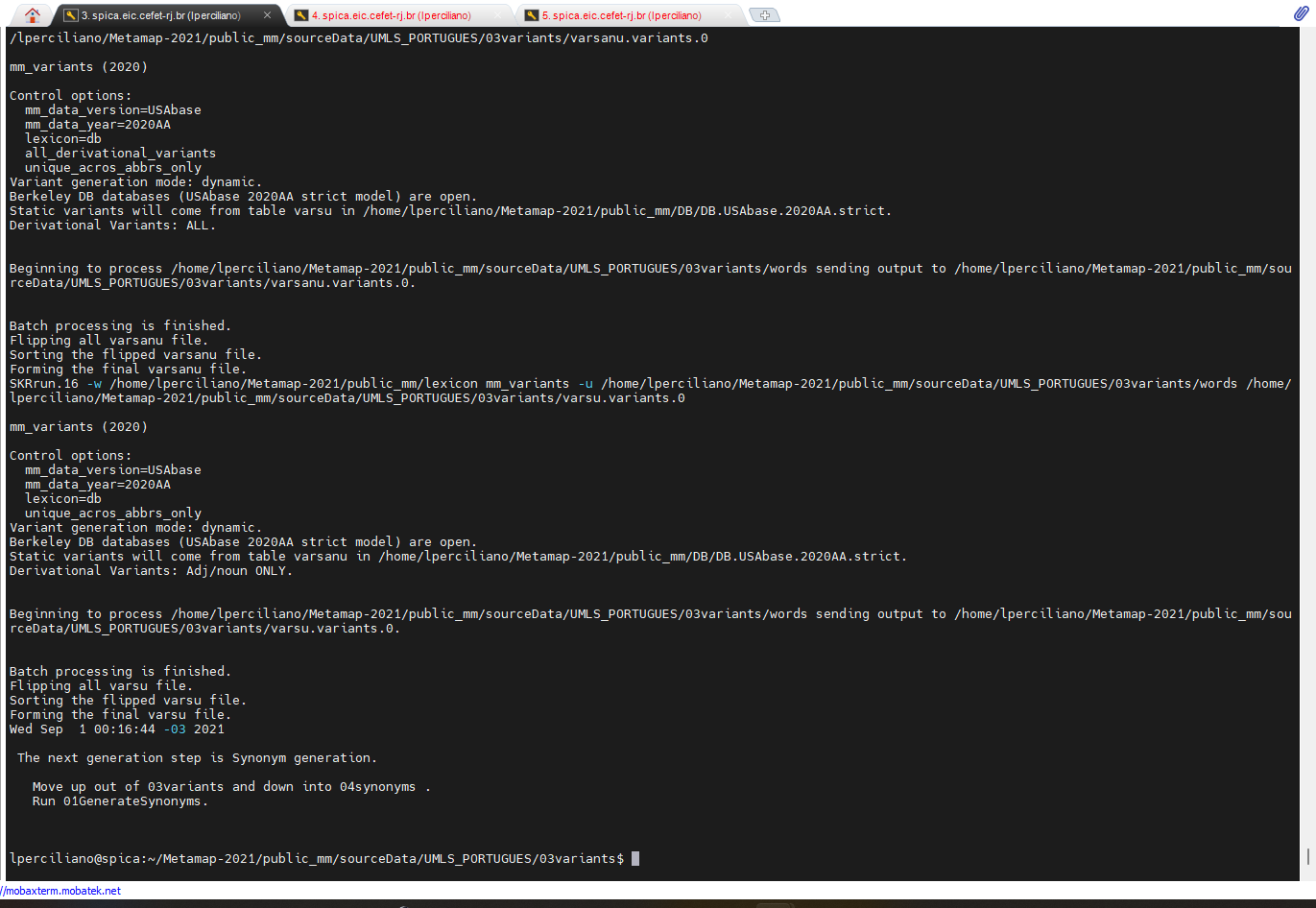
demora uns 2 minutos

gera alguns erros + cria vários arquivos

ls –lia



-



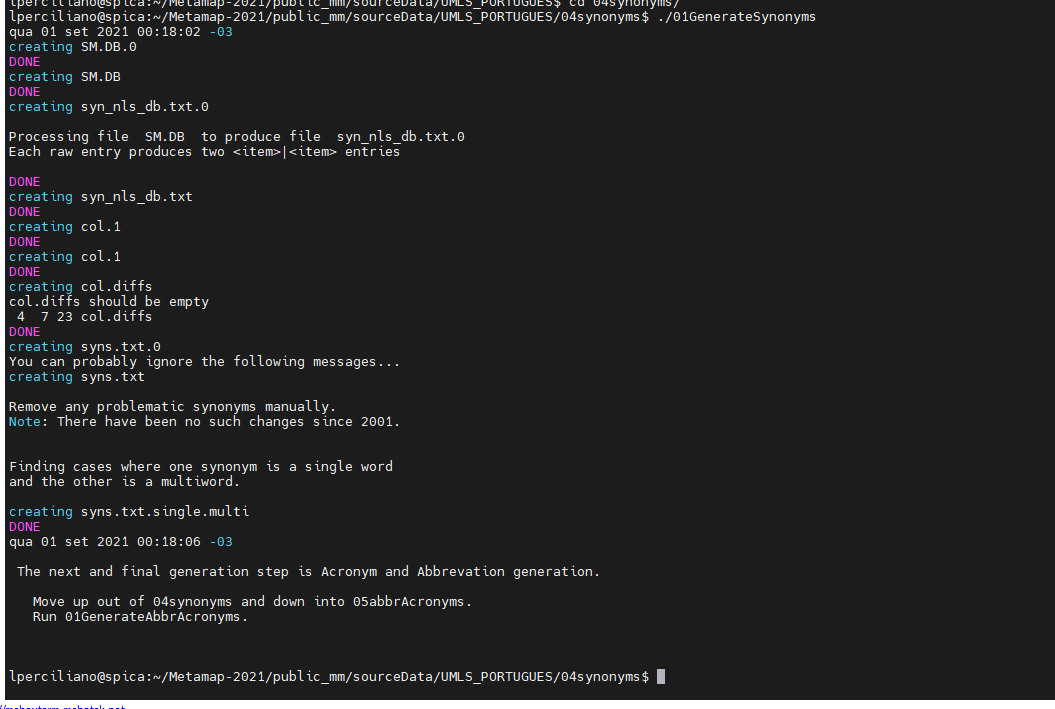
-

* 1. Acessar

$ cd ../04synonyms/

Alterar a primeira linha do arquivo ./01GenerateSynonyms com: #!/bin/bash

./01GenerateSynonyms

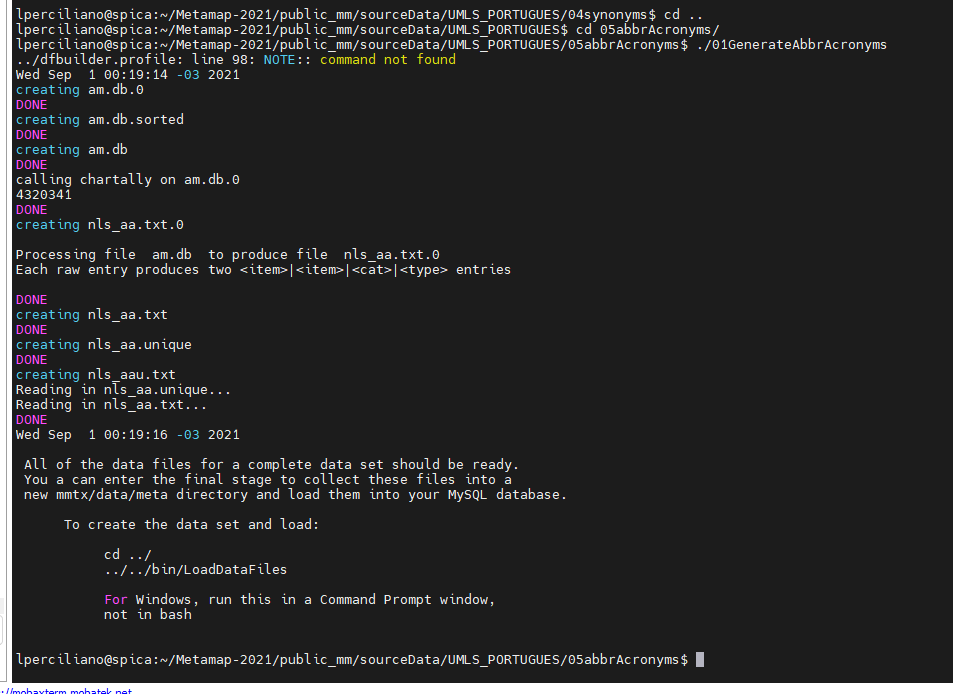


* 1. Acessar

cd ../05abbrAcronyms/

Alterar a primeira linha do arquivo ./01GenerateAbbrAcronyms para: #!/bin/bash

./01GenerateAbbrAcronyms



* 1. Voltar para o diretório public\_mm e executar

cd ../../../

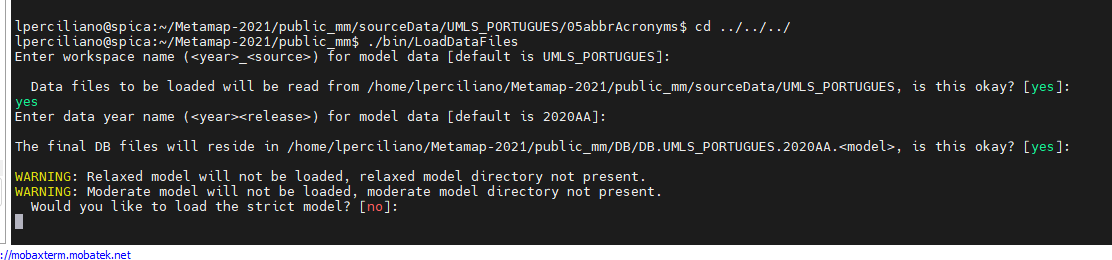
./bin/LoadDataFiles

enter

yes

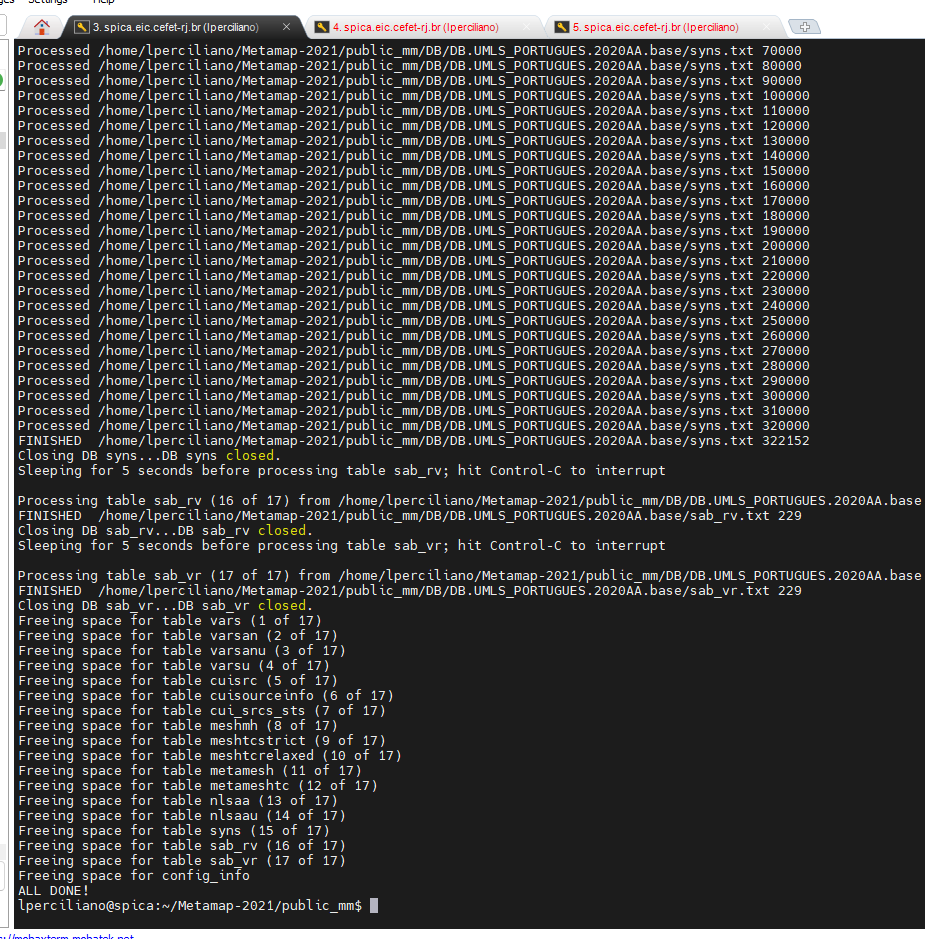
?????? enter (default 2020AA) ????? o meu é 2021AA

Yes



enter (no)

(+- 3 minutos)



* 1. Verificar se está funcionando.

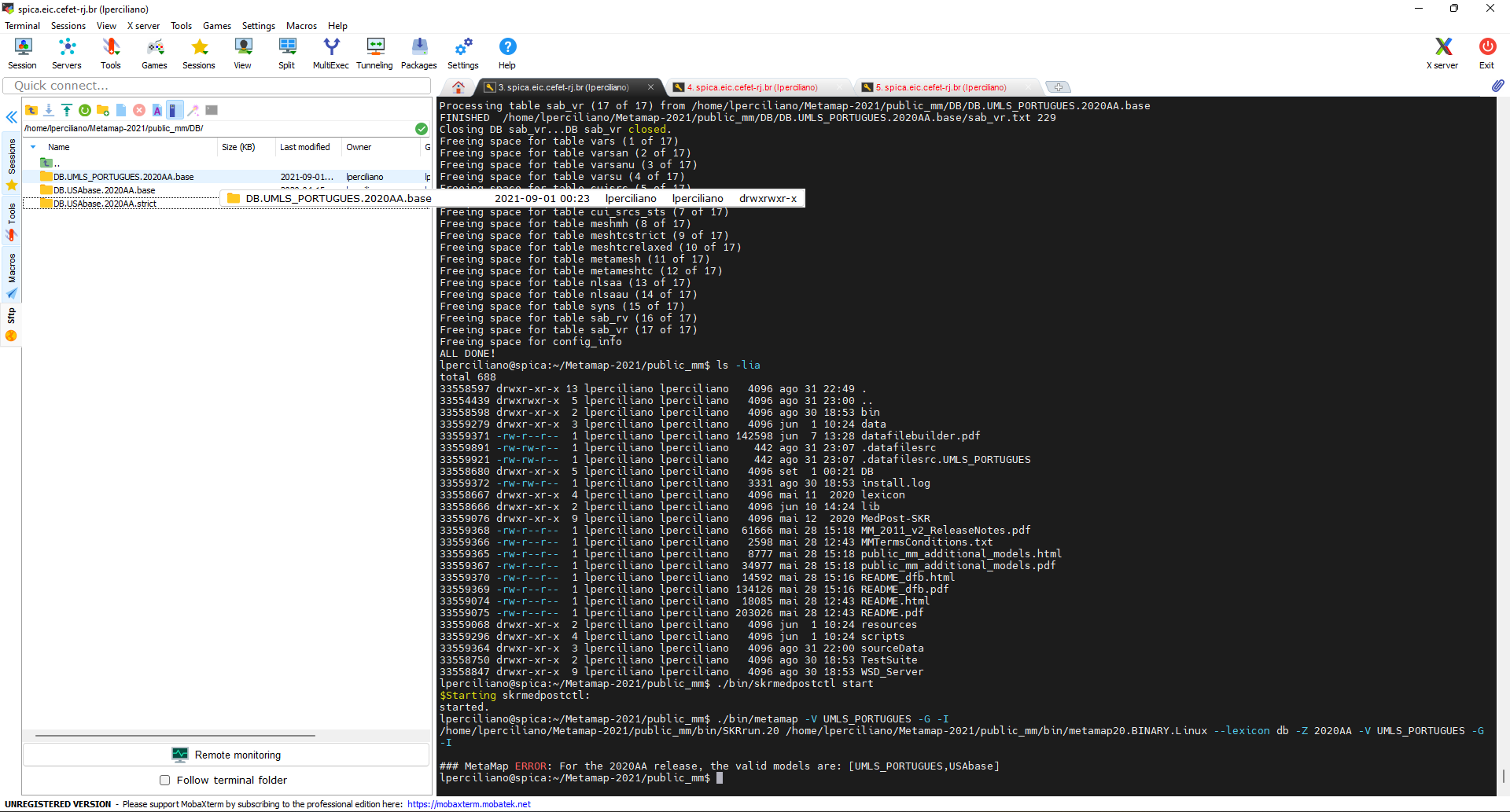
Deve gerar 2 diretórios com base em português.

cd Metamap-2020/public\_mm

./bin/skrmedpostctl start

./bin/metamap -V UMLS\_PORTUGUES -G -I

Testar os termos dor, fígado, bexiga, urina



# **Referências**

Instalação Metamap 2021 conforme documento <https://www.overleaf.com/project/603c54d389ae449739696960> com passos executados até (4.1.1. Executar Scripts)

Log do test <https://github.com/LuizPerciliano/Evento-Adverso-com-Metamap-e-Twitter-com-Portugues-Brasil/blob/main/tests/test-40/log_test40.txt>