

Análise e extração de informação de ficheiros PDF

Processamento de Linguagem Natural em
Engenharia Biomédica

2023/2024

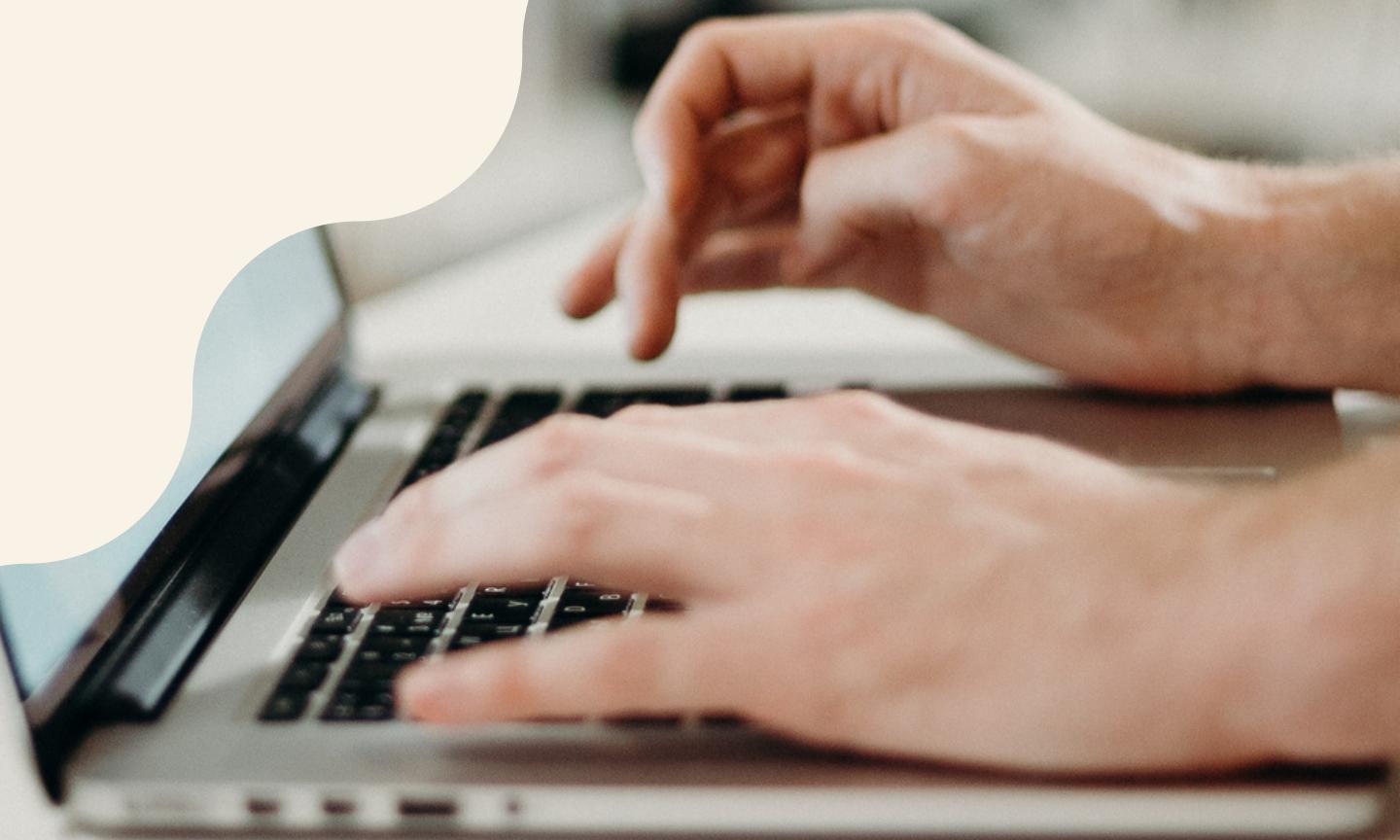
Documentos Analisados

01 - Glossário Ministério da Saúde

02 - Medicina

03 - Anatomia Geral

04 - Glossário de Termos Médicos e Populares



01 - Glossário Ministério da Saúde

MINISTÉRIO DA SAÚDE



Série F. Comunicação e Educação em Saúde

 Brasília – DF
2004

- Conversão do ficheiro PDF num ficheiro XML e abertura do mesmo;
- Limpeza do ficheiro XML;
- Arquivo das alterações efetuadas num ficheiro XML;
- Extração da informação;
- Arquivo da informação retida num ficheiro JSON.

Abertura e limpeza do ficheiro XML

```
import re
import json

ficheiro = open('glossario_ministerio_saude.xml','r', encoding="utf-8")
texto = ficheiro.read()

#Limpeza de texto e marcação
texto = re.sub(r'<page number="107" position="absolute" top="0" left="0" height="1263" width="892">\n[\d\d]+',r'',texto) # Remove parte final
texto = re.sub(r'<page number="15" (.*\n*){11}', r'', texto) # Remover parte inicial
texto = re.sub(r"</?page.*>", "", texto) # Remover <page> e </page>
texto = re.sub (r"</?text.*?>", "", texto) # remover <text> e </text>
texto = re.sub(r"</?fontspec.*?>", "", texto) # Remover <fontspec> e </fontspec>
texto = re.sub(r'</?pdf2xml.*?>', r'', texto)
texto = re.sub(r'<image.+?>', r'',texto)
texto = re.sub(r'<[?]xml(.*)>', "", texto) # Remover <xml> e </xml>
texto = re.sub(r'<![?DOCTYPE(.*)?>', "", texto) # Remover <DOCTYPE> e </DOCTYPE>
texto = re.sub(r'</?pdf(.*)>', "", texto) # Remover <pdf> e </pdf>
texto = re.sub(r'<b>\d{2}[15|3|5|7|9]</b>\n<b>.*</b>\n', r'', texto) # Retirar número de páginas ímpares
texto = re.sub(r'<b>.*</b>\n<b>\d{2}[16|2|4|6|8]</b>\n', '', texto) # Retirar número de páginas pares

texto = re.sub(r"<text(.*)>\s*(\*)(.*)</text>", "", texto) # Remover frases inicializadas com *, ** e ***
texto = re.sub(r"\n(\s*)\n", "\n", texto) # Remover espaços
texto = re.sub(r"[A-Z]\n\d{1,3}", "", texto) # Retira letras
texto = re.sub(r"\d{1,3}\n", "", texto) # Retira números
texto = re.sub(r"<b> </b>", "", texto) # Remove os <b> e </b>

texto = re.sub(r'<text .+?>','',texto) # retira <text top="605" left="176" width="289" height="16" font="14"> de todo o texto
texto = re.sub(r'<position="absolute" top="0" left="0" height="918" width="663">[\d\d]+',r'',texto)
texto = re.sub(r'</text>', r'', texto) # Remove os </text>
texto = re.sub(r'</b>\n<b>','',texto) # Colocar os títulos todos na mesma linha
texto = re.sub(r'</i>\n<i>','',texto) # Colocar as categorias todas na mesma linha

texto = re.sub(r'<i>([^\Category])).*</i>',r'\1\2', texto)
texto = re.sub(r'<i>Category:?\s?</i>\n(.*)\s\n\s', r"<i>Category: \1\2</i>\n", texto)
texto = re.sub(r'<i>Category:?\s?</i>\n(.*)\n', r"<i>Category: \1</i>\n", texto)
texto = re.sub(r'<em </i>\nSaúde", r" em Saúde</i>\n", texto)
texto = re.sub(r'</i>\n', r"<!--", texto)
texto = re.sub(r'&lt;\n&gt;', r"@\n&lt;b&gt;", texto)
texto = re.sub(r'\\n', "", texto)
texto = re.sub(r'\\n@', "", texto)</pre>
```

Extração de informação

```
file = open('glossario_atualizado.xml', 'w', encoding = 'utf-8')
file.write(texto)
file.close
```

```
#Extração de Informação

dicionario = {}
#ver termos que n tem descrições

all_titles = re.findall(r'<b>(.+)</b>', texto) # Extrai o título
all_descriptions = re.findall(r'@([^\@<]+)@', texto) # Extrai tudo o que está a frente do arroba
all_categories = re.findall(r'<i>.+\\n.+', texto) # Extrai Categoria: categoria
```

```
titulos = []
descricaoes = []
categorias = []

for titulo in all_titles:
    titulos.append(titulo)

for categoria in all_categories:
    categorias.append(categoria)

for descricao in all_descriptions:
    descricaoes.append(descricao)

# Verificar o comprimento das listas e usando o menor comprimento para iterar
lista = []
comprimento_minimo = min(len(titulos), len(descricaoes), len(categorias))

for i in range(comprimento_minimo):
    tuplo = (titulos[i], categorias[i] if i < len(categorias) else None, descricaoes[i])
    lista.append(tuplo)

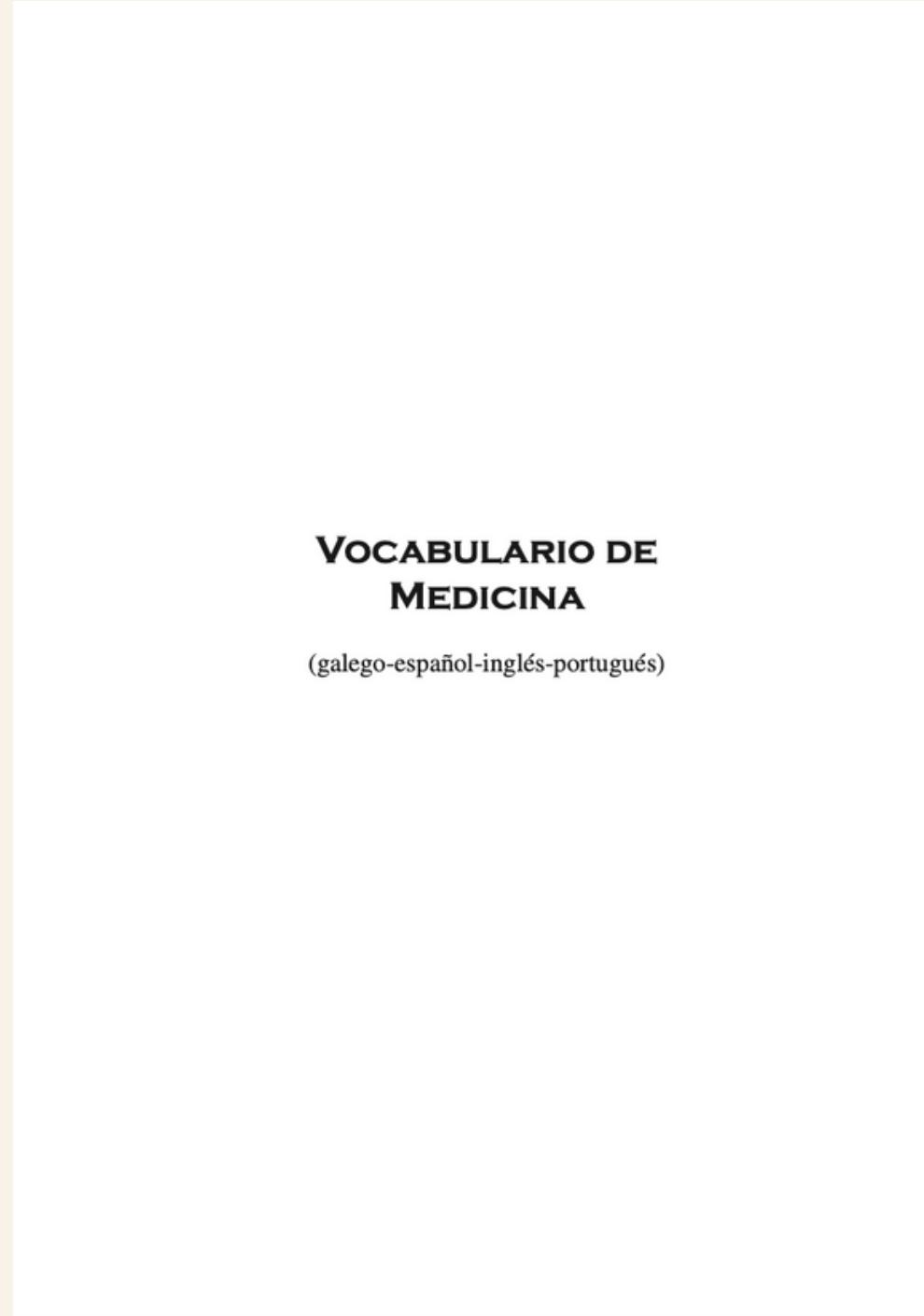
output = open("glossario.json", "w", encoding="utf-8")
json.dump(lista, output, ensure_ascii=False, indent=4)

output.close()
ficheiro.close()
```

Resultado do ficheiro JSON

```
[  
    "Abuso incestuoso",  
    "<i>Categoria: Acidentes e Violência</i>\n@Consiste no abuso sexual envolvendo pais ou ",  
    "Consiste no abuso sexual envolvendo pais ou \noutro parente próximo, os quais se encon-\ntram em uma posição de maior poder em re-\nlação à vítima."  
,  
    [  
        "Abuso sexual na adolescência",  
        "<i>Categoria: Acidentes e Violência</i>\n@É todo ato ou jogo sexual, relação heterosse-",  
        "É todo ato ou jogo sexual, relação heterosse-\nual ou homossexual, cujo agressor está em \nestágio de desenvolvimento psicossexual mais \nadiantado que a cr"  
,  
        [  
            "Abuso sexual na infância",  
            "<i>Categoria: Atenção à Saúde</i>\n@Modelo de intervenção centrado no indiví-",  
            "Modelo de intervenção centrado no indiví-\nduo no qual permite a relação entre a epide-\nmiologia e a dimensão sociocultural do tra-\nbalho de prevenção."  
,
```

02 - Medicina



- Conversão do ficheiro PDF num ficheiro XML e abertura do mesmo;
- Limpeza do ficheiro XML;
- Arquivo das alterações efetuadas num ficheiro XML;
- Extração da informação;
- Arquivo da informação retida num ficheiro JSON.

Abertura e limpeza do ficheiro XML

```
import re
import json
ficheiro = open('medicina.xml','r')
texto = ficheiro.read()

#Limpar o texto
texto = re.sub(r'^[\d\D]+<text\stop="128"\sleft="121"\swidth="3"\sheight="14"\sfont="17">\s</text>', r'', texto) # Remover parte inicial
texto = re.sub(r'<page number="544" position="absolute" top="0" left="0" height="918" width="663">\n[\d\D]+', r'', texto) # Remover parte final
texto = re.sub(r"</?page.*>", "", texto) # Remover <page> e </page>
texto = re.sub(r"</?text.*?>", "", texto) # remover <text> e </text>
texto = re.sub(r"</?fontspec.*?>", "", texto) # Remover <fontspec> e </fontspec>
texto = re.sub(r'</?pdf2xml.*?>', r'', texto)
texto = re.sub(r"<[?]xml(.*)>", "", texto) # Remover <xml> e </xml>
texto = re.sub(r"<[!]DOCTYPE(.*)>", "", texto) # Remover <DOCTYPE> e </DOCTYPE>
texto = re.sub(r"</?pdf(.*)>", "", texto) # Remover <pdf> e </pdf>
texto = re.sub(r"\n(\s*)\n", "", texto) # Remover espaços
texto = re.sub(r'[ ]{2,}', r'', texto) # Remover espaços nos meios das palavras
texto = re.sub(r'V\nvocabulario\n\d+', r'', texto) # Apagar informações sobre o cabeçalho
texto = re.sub(r'<text[^>]*><b>[A-Z]</b>\s*</text>\n', '', texto) # Remover letras isoladas s/ espaços
```

```
texto = re.sub(r"</b>+", r"</b>\n", texto) # Colocar as categorias em linhas diferentes
texto = re.sub(r'<b></b>', r'', texto)
texto = re.sub(r'<b> </b>', r'', texto)
texto = re.sub(r'<i> </i>', r'', texto)

#Marcação do texto
texto = re.sub(r'(SIN\.-)', r'\n\1@', texto) #Sinónimos
texto = re.sub(r'(VAR\.-)', r'\n\1$', texto) #Variantes
texto = re.sub(r'(Nota\.-)', r'\n\1£', texto) #Notas

#Traduções
texto = re.sub(r'(\bes\b)', r'\n\1#', texto) #Espanhol
texto = re.sub(r'(\ben\b)', r'\n\1$', texto) #Inglês
texto = re.sub(r'(\bla\b)', r'\n\1%', texto) #Latim
texto = re.sub(r'(\bpt\b)', r'\n\1&', texto) #Português
```

Extração da informação

```
file = open('med_atualizado.xml', 'w', encoding = 'utf-8')
file.write(texto)
file.close
```

```
portugues = []
espanhol = []
ingles = []
latim = []
portugues = []

all_ES = re.findall(r'\#\s*<i>(\w+)</i>', texto)
for palavra in all_ES:
    espanhol.append(palavra)

all_IN = re.findall(r'\$\s*<i>(\w+)</i>', texto)
for palavra in all_IN:
    ingles.append(palavra)

all_LA = re.findall(r'\%\s*<i>(\w+)</i>', texto)
for palavra in all_LA:
    latim.append(palavra)

all_PT = re.findall(r'\&\s*<i>(\w+)</i>', texto)
for palavra in all_PT:
    portugues.append(palavra)
```

```
dicE = {'Termos': espanhol}
dicI = {'Termos': ingles}
dicL = {'Termos': latim}
dicP = {'Termos': portugues}

dicionario = {}
dicionario['Espanhol'] = dicE
dicionario['Inglês'] = dicI
dicionario['Latim'] = dicL
dicionario['Português'] = dicP
```

```
output = open("medicina.json", "w", encoding="utf-8")
json.dump(dicionario, output, ensure_ascii=False, indent=4)

output.close()

ficheiro.close()
```

Resultado do ficheiro JSON

```
{  
    "Espanhol": {  
        "Termos": []  
        "ala",  
        "abdomen",  
        "ablación",  
        "abordaje",  
        "absceso",  
        "abulia",  
        "acalasia",  
        "acantolisis",  
        "ácaro",  
        "aceptación",  
        "acetábulo",  
        "acetilcolina",  
        "acidemia",  
        "acidificación",  
        "acidificante",  
        "acidosis",  
        "aciduria",  
        "acinesia",  
        "acino",  
        "....."  
    }  
}
```

```
{  
    "Inglês": {  
        "Termos": []  
        "ginecomastia",  
        "gínglimo",  
        "jeringa",  
        "sordera"  
    },  
    "Espanhol": {  
        "Termos": []  
        "wing",  
        "abdomen",  
        "abduction",  
        "ablation",  
        "approach",  
        "abscess",  
        "aboulia",  
        "achalasia",  
        "acantholysis",  
        "acarus",  
        "acceptance",  
        "acetabulum",  
        "acetylcholine",  
        "acidemia",  
        "acidification",  
        "acidifier",  
        "glutamate",  
        "acidosis"  
    }  
}
```

03 - Anatomia Geral

12 Anatomia Geral

1	ANATOMIA GERAL
2	PARTES DO CORPO HUMANO
3	Cabeça.
4	Fronte. Parte anterior da cabeça. A
5	Occipital. Parte posterior da cabeça (nuca). B
6	Tempora. A
7	Orelha.
8	Face. A
9	Olho.
10	Bochecha. C
11	Nariz.
12	Boca.
13	Mento.
14	Pescoço. Seu limite superior passa por uma linha ao longo da margem inferior da mandíbula, processo mastóide, linha nucal superior, até a protuberância occipital externa; seu limite inferior estende-se da margem superior do manúbrio do esterno, ao longo da clavícula, ao acrônio e à espinha da escápula, até o processo espinhosso de C VII. A, B
15	Tronco. A, B, C
16	Tórax. Parte do tronco, entre o pescoço e o abdome. Sua estrutura básica é a caixa torácica. Seu limite inferior é a abertura torácica inferior e o diafragma. A
17	Peito. C
18	Abdome. Parte do tronco entre o tórax, a margem superior do sacro, o ligamento inguinal e a sínfise pública. A, C
19	Pelve. Parte do tronco entre o abdome e o soalho da pelve. A pelve maior e a pelve menor são separadas pela linha terminal.
20	Dorso. Parte posterior do tronco. B
21	Membro superior. Constituído pelo cíngulo do membro superior e pela extremidade livre.
22	Cíngulo do membro superior. Sua estrutura óssea básica é formada pela escápula e pela clavícula. A, B
23	Axila. Cavidade axilar. Espaço de união entre o membro superior e a parede lateral do tórax. C
24	Braço. A
25	Cotovelo.
26	Antebraço. A
27	Mão.
28	Carpo. A
29	Metacarro. A
30	Palma. A
31	Dorso da mão.
32	Dedos da mão.
33	Membro inferior. Constituído pelo cíngulo do membro inferior e pela extremidade livre.

- Conversão do ficheiro PDF num ficheiro XML e abertura do mesmo;
- Limpeza do ficheiro XML;
- Arquivo das alterações efetuadas num ficheiro XML;
- Extração da informação;
- Arquivo da informação retida num ficheiro JSON.

Abertura e limpeza do ficheiro XML

```
import re
import json

file = open("C:\\\\Users\\\\liama\\\\OneDrive\\\\Ambiente de Trabalho\\\\anatomia geral\\\\Anatomia_GeralORIGINAL.xml", "r", encoding="utf-8")
text = file.read()

def clean(text):
    text = re.sub(r"\s+", " ", text) # substitui todos os espaços e \n por apenas 1 espaço
    return text.strip() # Remove o \n no início e fim da frase
```

```
text = re.sub(r"</?page.*>", "", text) # Remover <page> e </page>
text = re.sub(r"</?image.*?>", "", text) # Remover <image> e </image>
text = re.sub(r"<![!]DOCTYPE(.*)?>", "", text) # Remover <DOCTYPE> e </DOCTYPE>
text = re.sub(r"</?pdf(.*)?>", "", text) # Remover <pdf> e </pdf>
text = re.sub(r"</?fontspec.*?>", "", text) # Remover <fontspec> e </fontspec>
text = re.sub(r"<[?]xml(.*)?>", "", text) # Remover <xml> e </xml>
text = re.sub(r"<i><b>(.*)</b></i>", "", text) # Remover <i><b> e </b></i>
text = re.sub(r"</?text.*?>", "", text) # Remover <text> e </text>
```

```
text = re.sub(r"</?page.*>", "", text) # Remover <page> e </page>
text = re.sub(r"</?image.*?>", "", text) # Remover <image> e </image>
text = re.sub(r"<![!]DOCTYPE(.*)?>", "", text) # Remover <DOCTYPE> e </DOCTYPE>
text = re.sub(r"</?pdf(.*)?>", "", text) # Remover <pdf> e </pdf>
text = re.sub(r"</?fontspec.*?>", "", text) # Remover <fontspec> e </fontspec>
text = re.sub(r"<[?]xml(.*)?>", "", text) # Remover <xml> e </xml>
text = re.sub(r"<i><b>(.*)</b></i>", "", text) # Remover <i><b> e </b></i>
text = re.sub(r"</?text.*?>", "", text) # Remover <text> e </text>
```

Extração de informação

```
# Pesquisar a expressão
list = re.findall(r"<b>\s(.*)\.</b>(.*)\s+@", text)
nova_lista = [(designacao, re.sub(r"@<b>|</b>"," ", descricao)) for designacao, descricao in list] #remover @<b> de alguns termos que originava erros

# Conversão
new_list = []
for designacao, descricao in nova_lista:
    new_designacao = designacao.lower()
    new_descricao = clean(descricao)
    new_list.append([new_designacao, new_descricao])

list = new_list
```

```
dicionario = dict(list)

output = open("C:\\\\Users\\\\liama\\\\OneDrive\\\\Ambiente de Trabalho\\\\anatomia geral\\\\anatomia geral.json", "w", encoding="utf-8")
json.dump(dicionario, output, ensure_ascii=False, indent=4)

output.close()

file.close()
```

Resultado do ficheiro JSON

```
{  
    "fronte": "Testa",  
    "occipital": "Pertencente ao occípicio (nuca); situado em direção ao occípicio (nuca).",  
    "pescoço": "Seu limite superior passa por uma linha ao longo da margem inferior da mandíbula, processo mastóide, linha nucal superior, até a protuberância occipital externa; seu limite inferior estende-se ao nível da articulação entre o occipital e o atlas.",  
    "tórax": "Parte do tronco, entre o pescoço e o abdome. Sua estrutura básica é a caixa torácica. Seu limite inferior é a abertura torácica inferior e o diafragma.",  
    "abdome": "Parte do tronco entre o tórax, a margem superior do sacro, o ligamento inguinal e a sinfise púbica.",  
    "pelve": "Parte do tronco entre o abdome e o seio da pelve. A pelve maior e a pelve menor são separadas pela linha terminal.",  
    "dorso": "Parte posterior do tronco.",  
    "membro superior": "Constituído pelo cíngulo do membro superior e pela extremidade livre.",  
    "cíngulo do membro superior": "Sua estrutura óssea básica é formada pela escápula e pela clavícula.",  
    "axila": "Cavidade axilar. Espaço de união entre o membro superior e a parede lateral do tórax.",  
    "membro inferior": "Constituído pelo cíngulo do membro inferior e pela extremidade livre.",  
    "cíngulo do membro inferior": "Sua estrutura óssea básica é formada pelo osso do quadril.",  
    "quadril": "Região de união da pelve com a extremidade livre do membro inferior.",  
    "região poplitea": "Fossa superficial do joelho.",  
    "sura": "Panturrilha.",  
    "cavidade abdominopélvica": "Cavidade conjunta do abdome e da pelve.",  
    "mediano": "Situado no plano mediano.",  
    "coronal": "Situado no plano da sutura coronal.",  
    "sagital": "Situado no plano da sutura sagital, ou situado em direção dorsal-ventral.",  
    "intermédio": "Situado entre duas estruturas.",  
    "medial": "Situado próximo do plano mediano.",  
    "lateral": "Situado afastado do plano mediano.",  
    "anterior": "Situado à frente de.",  
    "posterior": "Situado atrás de.",  
    "ventral": "Situado em direção ao ventre.",  
    "dorsal": "Situado em direção ao dorso.",  
    "frontal": "Pertencente à frente (testa); situado paralelamente à fronte (testa).",  
    "superior": "Situado em direção à cabeça.",  
    "inferior": "Situado em direção à cauda (cóccix).",  
    "cranial": "Pertencente ao crânio; situado em direção do crânio.",  
    "caudal": "Situado em direção à cauda.",  
    "rostral": "Situado em direção ao rostro do corpo caloso.",  
    "apical": "Pertencente ao ápice; em direção ao ápice."}
```

Glossário de Termos Médicos Técnicos e Populares (em português de Portugal)

Fonte: <http://users.ugent.be/~rvdstich/eugloss/welcome.html>

Multilingual Glossary of Technical and Popular Medical Terms in Nine European Languages

Observação: This project was commissioned by The European Commission (<http://europa.eu.int/comm/index.htm>) and executed by Heymans Institute of Pharmacology (<http://www.heymans.ugent.be/>) and Mercator School, Department of Applied Linguistics.

A

a milionésima parte de um grama (pop) , **micrograma**

à volta da boca (pop) , **perioral**

à volta da órbita (pop) , **periorbital**

à volta dos vasos sanguíneos (pop) , **perivascular**

abaixamento, abatimento, prostração (pop) , **depressão**

acesso , acesso, tumor (pop)

acesso, tumor (pop) , **acesso**

acesso; acumulação de pus (pop) , **empíema**

abdómen , barriga, ventre (pop)

abdominal , ventral (pop)

aberrante , anormal (pop)

abertura; orifício (pop) , **perfuração**

ablação (pop) , **extração**

ablação dos órgãos性uais, capaçao, eviraçao, emasculação (pop) , **castração**

abocadamento (pop) , **anastomose**

abortamento, desmancho (pop) , **aborto**

aborto , abortamento, desmancho (pop)

- Conversão do ficheiro PDF num ficheiro XML e abertura do mesmo;
- Limpeza do ficheiro XML;
- Arquivo das alterações efetuadas num ficheiro XML;
- Extração da informação;
- Arquivo da informação retida num ficheiro JSON.

04 - Glossário de Termos Médicos e Populares

Script para limpeza e extração de conteúdos

```
1 import re
2 import json
3
4 ficheirotxt = open("Glossario_de_termos_Medicos_Tecnicos_Populares.xml", encoding="utf-8")
5 text = ficheirotxt.read()
6
7 text = re.sub(r"</?fontspec.*?", "", text)
8 text = re.sub(r"</?page.*?", "", text)
9 text = re.sub(r"</?text.*?", "", text)
```

```
12 text = re.sub(r"<b>[A-Z]</b>", "", text) #termos que correspondem a apenas uma maiuscula sao eliminados, especie de blacklist
13 #maiusculas apenas aparecem nesses casos especificos
14
15 text = re.sub (r"<i>", "", text)
16 text = re.sub (r"</i>", "", text)
17 termos = re.findall(r"<b>(.*)</b>", text) #recolha dos termos
```

```
20 text2 = re.sub(r"<b>(.*)</b>", "", text) #remocão dos termos na variavel text2
21 text2= re.sub(r"\n+", "", text2) #remocão dos paragrafos
22 text2 = re.sub(r"\s,", "", text2) #remover espacos seguidos de virgula, presente apos o (pop)
23
24
25
26 text2 = re.sub(r"\(pop\)", "@@", text2) #substituir o (pop) por @@ para marcação
27 #text2 = re.sub(r"@@ @@", "@@", text2)
```

Script para limpeza e extração de conteúdos

```
30     resultado=open("glossario_termos_populares.txt", "w", encoding='UTF-8')
31     resultado.write(text2)
32     resultado.close()
```

```
35     designacoes1 = re.findall(r"@(?:\s|,|X)(.*?)", text2) #encontra todas as que começam em
36     #@ seguido por um espaço em branco, vírgula ou "X", e terminam com "@".
37     # As partes entre "@" e "@" são capturadas pelo grupo (.*) e retornadas como uma lista.
38     #(?: ) identifica o conteúdo como necessário mas não o captura
39     designacoes2= re.findall(r"(^.*?)@", text2) #encontra o que começa no início de uma linha e termina com @
40
41     designacoes= designacoes2 + designacoes1 #conjunto total de designacoes
42
43     lista_de_listas = [[x, y] for x, y in zip(termos, designacoes)]
44
45     dicti={}
46     dicti = dict(lista_de_listas)
```

```
48     output = open ("json_termos_populares.json", "w", encoding="utf-8")
49     json.dump(dicti, output, ensure_ascii=False, indent=4)
50     output.close()
```

Resultado do ficheiro txt

a milionésima parte de um grama @@ à volta da boca @@ à volta da órbita @@ à volta dos vasos sanguíneos @@ abaixamento, abatimento, prostração @@ abcesso, tumor @@ abcesso, tumor @@ abcesso; acumulação de pus @@ barriga, ventre @@ ventral @@ anormal @@ abertura; orifício @@ ablação @@ ablação dos órgãos sexuais, capaço, eviração, emasculação @@ abocamento @@ abortamento, desmancho @@ abortamento, desmancho @@ repentina, brusco @@ absorvimento, absorvência @@ absorção de água e de solutos por células vivas @@ absorvimento, absorvência @@ jejum @@ abstracção, suposição @@ acariáse sarcóptica, sarna @@ incapacidade em permanecer sentado @@ acção de certos corpos sobre outros @@ acção de enferrujar @@ acção de inalar, extracção de líquidos ou gases @@ acção de urinar, urinação @@ aceleração da frequência cardíaca @@ acessório, anexo, indesejável @@ acidez, azedume @@ por acaso, sem importância @@ acidez, azedume @@ ácido animado @@ alteração do equilíbrio ácido básico do sangue e líquidos teciduais @@ ausência de movimento, acinese @@ espinha @@ adaptação @@ acomodação @@ acompanhante @@ acostumado @@ acrescentar um solvente a um medicamento em pó; tipo de regeneração de uma superfície lesionada; retorno à forma líquida do soro ou plasma sanguíneo previamente dessecado@@ cor azulada das extremidades (mãos-pés) @@ hormônio adreno-corticotrófico, corticotrofina @@ activação de uma droga por outra @@ activação de uma droga por outra @@ activado

Resultado do ficheiro JSON

```
1 {  
2     "micrograma": " a milionésima parte de um grama ",  
3     "perioral": " à volta da boca ",  
4     "periorbital": " à volta da órbita ",  
5     "perivasicular": " à volta dos vasos sanguíneos ",  
6     "depressão": " abaixamento, abatimento, prostração ",  
7     "abcesso": " abcesso, tumor ",  
8     "empiema": " abcesso; acumulação de pus ",  
9     "abdómen": " barriga, ventre ",  
10    "abdominal": " ventral ",  
11    "aberrante": " anormal ",  
12    "perfuração": " abertura; orifício ",  
13    "extracção": " ablação ",  
14    "castração": " ablação dos órgãos sexuais, capaço, eviração, emasculação ",  
15    "anastomose": " comunicação natural ou artificial entre dois vasos ou nervos ",  
16    "aborto": " abortamento, desmancho ",  
17    "abrupto": " repentina, brusco ",  
18    "absorção": " absorvimento, absorvência ",  
19    "ressorção": " absorção de água e de solutos por células vivas ",  
20    "abstinênci": " jejum ",  
21    "acidez": " acidez, azedume ",  
22    "acessório": " acessório, anexo, indesejável ",  
23    "adreno-corticotrófico": " hormônio adreno-corticotrófico, corticotrofina ",  
24    "adaptação": " adaptação ",  
25    "acomodação": " acomodação ",  
26    "acostumado": " acostumado ",  
27    "acção": " acção de enferrujar, acção de inalar, acção de urinar, aceleração da frequência cardíaca ",  
28    "acção de enferrujar": " acção de enferrujar ",  
29    "acção de inalar": " acção de inalar ",  
30    "acção de urinar": " acção de urinar ",  
31    "aceleração da frequência cardíaca": " aceleração da frequência cardíaca ",  
32    "acidez, azedume": " acidez, azedume ",  
33    "acido animado": " ácido animado ",  
34    "ausência de movimento": " ausência de movimento ",  
35    "acinese": " acinese ",  
36    "espinha": " espinha ",  
37    "adaptação": " adaptação ",  
38    "acomodação": " acomodação ",  
39    "acompanhante": " acompanhante ",  
40    "acostumado": " acostumado ",  
41    "acrescentar": " acrescentar ",  
42    "um solvente": " um solvente ",  
43    "um medicamento": " um medicamento ",  
44    "em pó": " em pó ",  
45    "tipo de regeneração": " tipo de regeneração ",  
46    "uma superfície lesionada": " uma superfície lesionada ",  
47    "retorno": " retorno ",  
48    "à forma líquida": " à forma líquida ",  
49    "do soro": " do soro ",  
50    "ou plasma": " ou plasma ",  
51    "sanguíneo": " sanguíneo ",  
52    "previamente dessecado": " previamente dessecado ",  
53    "cor azulada": " cor azulada ",  
54    "das extremidades": " das extremidades ",  
55    "(mãos-pés)": " (mãos-pés) ",  
56    "hormônio adreno-corticotrófico": " hormônio adreno-corticotrófico ",  
57    "corticotrofina": " corticotrofina ",  
58    "activação de uma droga": " activação de uma droga ",  
59    "por outra": " por outra ",  
60    "activação de uma droga": " activação de uma droga ",  
61    "por outra": " por outra ",  
62    "activado": " activado "
```