



Trabalho Prático

Extração e Estruturação de Termos Médicos

Beatriz Amorim, pg56112

Carolina Santos, pg56116

Catarina Nunes, pg56117



Objetivos

- Extrair informação relevante de vários documentos PDF;
- Juntar estruturas diferentes num único formato;
- Evitar perda de informação.

Documentos Utilizados

- Dicionário Multilingue da COVID-19;
- Glossário de Neologismos da Saúde Humana
- Glossário Temático - Monitoramento e Avaliação
- Glossário de Termos Médicos e Populares

Dicionário Multilingue da COVID-19:

```
{
  "Conceito": {
    "Categoria Lexical": "",
    "Traduções": { 'oc': [], 'eu': [], 'gl': [], 'es': [], 'en':
      ↪ [], 'fr': [], 'pt': [], 'pt_PT': [], 'pt_BR': [], 'nl':
      ↪ [], 'ar': [] },
    "Categoria": "",
    "Descrição": "",
    "Notas": [],
    "Sigla": [],
    "Símbolo": "",
    "Entrada Principal": "",
    "Sinónimo": "",
    "Sinónimo Complementar": [],
    "Denominação Comercial": "",
    "Nome Científico": "",
    "Número CAS": []
  }
}
```

Glossário de Neologismos da Saúde Humana:

```
{
  conceito : {
    "categoria_lexical": "",
    "traducoes": {
      "en": "",
      "es": ""
    },
    "descricao": "",
    "sigla": "",
    "inf_encicl": "",
    "citacoes": [""]
  }, ...
}
```

Glossário Temático - Monitoramento e Avaliação:

```
{  
  "Nome do Conceito": {  
    "Traduções": {"es": "tradução", "en": "translation"},  
    "Descrição": "Texto descritivo.",  
    "Gênero": "masc",  
    "Número": "pl",  
    "Sinónimos": ["sinónimo1", "sinónimo2"],  
    "Notas": ["Nota 1", "Nota 2"],  
    "Sigla": "sigla",  
    "Remissivas": ["relacionado1", "relacionado2"]  
  }  
}
```

Glossário de Termos Médicos e Populares:

```
{  
  "Conceito": {  
    "Termo Popular": [],  
  }  
}
```

Ficheiro JSON Global:

Total de termos no dicionário multilingue: 739

Total de termos no glossário de neologismos: 306

Total de termos no glossário de termos médicos populares: 1822

Total de termos no glossário de monitoramento e avaliação: 222

soma = 3089

Total de termos depois da concatenação (não há chaves repetidas): 3085

Termos em comum entre glossários:

- placebo: dicionario_multilingue, termos_populares
- antiviral: dicionario_multilingue, termos_populares
- broncodilatação: glossario_neologismos, termos_populares
- radiografia: dicionario_multilingue, termos_populares

Ficheiro JSON Global:

Exemplo de fusão de valores de um termo no JSON final:
(com dados do glossário de neologismos e termos populares)

```
"broncodilatação": {
  "descricao": "Dilatação dos brônquios.",
  "sigla": "",
  "inf_encicl": "",
  "traducoes": {
    "en": "bronchodilatation",
    "es": "bronco dilatación"
  },
  "termo_popular": [
    "dilação dos brônquios"
  ],
  "categoria_lexical": "s.f.",
  "citacoes": [
    "...Assim, o bloqueio desses receptores pelos princípios ativos do guaco provoca a broncodilatação e a diminuição da secreção brônquica..."
  ]
}
```