

# QuickStatements

**Éder Porto, Wiki Movimento Brasil**

Curso de Estruturação e Georreferenciamento de Dados Abertos sobre Saúde  
20 de agosto de 2024



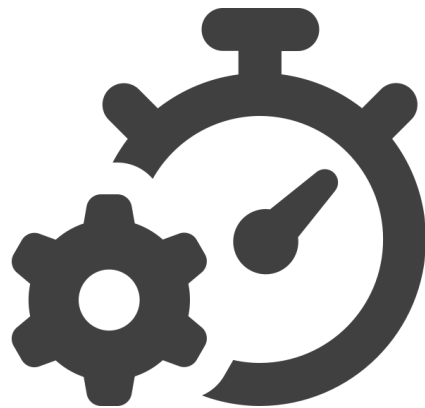
# Introdução

Otimizar o carregamento de grandes volumes de dados para o Wikidata implica em mais tempo disponível para se dedicar a outras tarefas.

Para otimização, utilizam-se ferramentas específicas que facilitam o esforço humano envolvido.

Uma dessas ferramentas é o **QuickStatements**, desenvolvida por Magnus Manske.

<https://quickstatements.toolforge.org>



# Comandos

A ferramenta interpreta comandos de texto wikidatificados e adiciona as informações ao Wikidata.

Estes comandos são escritos como triplas.  
Por exemplo:

**QID|PID|Valor**

**QID** é o número Wikidata do item.

Por exemplo: Q6041378 é o item do Instituto Moreira Salles

**PID** é o número Wikidata da propriedade.

Por exemplo: P112 é o valor para a propriedade "Fundador"

**Valor** pode ser um dentre diversos tipos de dados: QID, quantidades, texto, data, coordenadas geográficas são os mais comuns.

# Comandos

Em um comando, um valor pode ser acompanhado de qualificadores.

Estes qualificadores seguem as mesmas definições de **PID** e **Valor** citadas anteriormente e são escritos logo após o Valor da declaração:

**QID|PID1|Valor1|PID2|Valor2|PID3|Valor3|...**

Por exemplo:

**Q56285719|P195|Q71989864|P217|"002001MF005003"|...**

# Comandos

Cada comando deve ser seguido de uma referência. Propriedades para referências no Wikidata aceitam valores variados. As mais comuns são a P854 (Url de referência) e a P248 (afirmado em). Vêm sempre no final do comando e a letra "P" deve ser substituída por "S" de *source*. Assim:

**QID|PID1|Valor1|PID2|Valor2|SID1|Valor3|...**

Por exemplo:

**Q56285719|P195|Q71989864|P217|"002001MF005003"|S854|"https://acervos.ims.com.br/portals/#/detailpage/75863"**

# Tipos de dados

**Item (QID):**

Q12345

**Propriedade (PID):**

P12345

**Valores (Valor):**

**Item:**

Q12345

**Texto:**

"Texto"

**Data:**

+AAAA-MM-DDT00:00:00

**Coordenadas:**

@LAT/LON

**Quantidade:**

+NUM[MAX,MIN]UID

**Texto monolíngue:**

lang:"Texto"

LAT e LON são coordenadas geográficas em números decimais.

NUM, MAX e MIN são números reais.

MAX, MIN e UID são opcionais na maioria dos casos.

UID é o QID da unidade de medida, trocando-se "P" por "U".

lang é o código da língua.

# Comandos em massa

Para construir muitos comandos, utiliza-se em geral uma tabela, em que cada coluna representa um **único** valor para uma propriedade, para um qualificador ou para uma referência.

Lpt-br	Dpt-br	P31	P31	P7835
Largo da Sé na década de 1880	Largo da Sé na década de 1880 fotografado por Marc Ferrez	Q125191		75863





# Comandos em massa

Para construir muitos comandos, utiliza-se em geral uma tabela, em que cada coluna representa um **único** valor para uma propriedade, para um qualificador ou para uma referência.

P7835	QS			
75863	CREATE  LAST Lpt-br "Largo da Sé na década de 1880"  LAST Dpt-br "Largo da Sé na década de 1880 fotografado por Marc Ferrez"  LAST P31 Q125191  LAST P7835 "75863"			

# Comandos - Criação de entidades

```
CREATE||LAST|Lpt-br|"Largo da Sé na  
década de 1880"||LAST|Dpt-br|"Largo da Sé  
na década de 1880 fotografado por Marc  
Ferrez"||LAST|P31|Q125191||LAST|P7835|"758  
63"
```

- Uma linha

```
CREATE
```

```
LAST    Lpt-br    "Largo da Sé na década  
de 1880"
```

```
LAST    Dpt-br    "Largo da Sé na década  
de 1880 fotografado por Marc Ferrez"
```

```
LAST    P31    Q125191
```

```
LAST    P7835    "75863"
```

- N linhas

# Comandos - Edição de entidades

```
Q56285719|Lpt-br|"Largo da Sé na década de 1880"||Q56285719|Dpt-br|"Largo da Sé na década de 1880 fotografado por Marc Ferrez"||Q56285719|P31|Q125191||Q56285719|P7835|"75863"
```

- Sem “CREATE”
- LAST é substituído pelo QID da entidade

```
Q56285719    Lpt-br    "Largo da Sé na década de 1880"
```

```
Q56285719    Dpt-br    "Largo da Sé na década de 1880 fotografado por Marc Ferrez"
```

```
Q56285719    P31    Q125191
```

```
Q56285719    P7835    "75863"
```

# Comandos - Remoção de declarações

-Q56285719|P625|@-23.54965,-46.6335||  
Q56285719|P1259|@-23.54965,-46.6335

-Q56285719	P625	@-23.54965,-46.63354
Q56285719	P1259	@-23.54965,-46.63354

- Remoção total da declaração  
(valor, qualificadores e referências)

# Comandos de tabelas

- Função Concatenar
  - =CONCATENAR(<texto\_1>;<texto\_2>;...;<texto\_n>)
  - Se <texto\_i> for:
    - um texto digitado diretamente na fórmula: deve estar acompanhado de aspas duplas;
    - um ponteiro para uma célula de Valor = Texto da tabela: as aspas devem ser concatenadas à parte;
    - um ponteiro para uma célula da tabela: não são necessárias aspas.
  - Exemplo:
    - =CONCATENAR("||**LAST**||";**\$D\$1**;"|";"""";D2;"""")

# Comandos de tabelas

- Função Se

- =SE(teste\_lógico; valor\_se\_verdadeiro; valor\_se\_falso)

- Exemplo:

- SE(B2<>"";CONCATENAR("||LAST|";\$B\$1;"|";"""";B2;"""");""");



Se a célula B2 for diferente (<>) de vazio ("")

Devolva a concatenação (CONCATENAR("||LAST|";\$B\$1;"|";"""";B2;""""))

Caso contrário, devolva vazio ("")

# Comandos concatenados

- Os comandos podem ser concatenados em cascata para maior legibilidade.

```
=CONCATENAR(  
    "CREATE";  
    SE(A2<>"";CONCATENAR(" | | LAST |";$A$1;" |";"""";A2;"""");""");  
    SE(B2<>"";CONCATENAR(" | | LAST |";$B$1;" |";"""";B2;"""");""");  
    SE(C2<>"";CONCATENAR(" | | LAST |";$C$1;" |";C2);""");  
    SE(D2<>"";CONCATENAR(" | | LAST |";$D$1;" |";D2;SE(R2<>"";CONCA  
TENAR(" | | LAST |";$R$1;" |";"""";R2;"""");"""););""")  
)
```

# Comandos concatenados

- Os comandos podem ser concatenados em cascata para maior legibilidade.

```
=CONCATENAR(  
    SE(B2<>"";CONCATENAR("|";A2;"|";$B$1;"|";"""";B2;"""");""");  
    SE(C2<>"";CONCATENAR("|";A2;"|";$C$1;"|";C2);""");  
    SE(D2<>"";CONCATENAR("|";A2;"|";$D$1;"|";D2;SE(R2<>"";CONCA  
TENAR("|";LAST;"|";$R$1;"|";"""";R2;"""");""));""")  
)
```



# Comandos em massa

O formato de comandos apresentado é o modo mais legível para humanos, mas principalmente para metadados muito bem estruturados um formato que evita toda essa concatenação é o CSV.

Nem sempre o QuickStatements interpreta o CSV de forma correta, principalmente quando há valores vazios entre as colunas.

Para isso, criamos uma ferramenta chamada **CSV2QS** para converter entre formatos.

<https://csv2qs.toolforge.org>



CSV2QS

# Quick Statements - Boas práticas

- Edições muito grandes ou potencialmente controversas devem passar pelo crivo da comunidade;
- Todas as declarações (exceto talvez para identificadores) devem ter referências. Rótulos (Lxx e Axx), descrições (Dxx) e ligações internas para projetos Wiki (Sxxwiki) são propriedades especiais e não têm espaço para referências;
- Editores utilizando a ferramenta são responsáveis por arrumar ou desfazer suas edições equivocadas;
  - As edições podem ser acompanhadas através da ferramenta EditGroups (<https://editgroups.toolforge.org>).
- Evite criar itens duplicados.
  
- Mais informações sobre a ferramenta em [\[\[d:Help:QuickStatements\]\]](#)

Éder Porto

Ederporto | EPorto (WMB)  
Wiki Movimento Brasil

[eder.porto@wmnobrasil.org](mailto:eder.porto@wmnobrasil.org)