

The image features a series of vertical bars in red, green, and blue, arranged in two rows. The top row consists of nine bars of varying heights and widths, while the bottom row consists of eight bars. The bars are positioned around the central text.

WIKIDATA LAB IX

Structured narratives on Wikipedia



Event schedule:

- The concept of structured narratives
- The Mbabel tool: operation and use
- Understanding the code: creating templates
- Proposed activities



What is structured narrative?

The development of verbal texts, understandable by humans, automated from predetermined arrangements that process information from structured databases.



What is structured narrative?

Experimenting with structured narratives occurs in Natural Language Generation (NLG) software, especially in automated journalism for sports, finance, weather forecasting, entertainment and crime coverage.



NLG softwares

Softwares write basic textual structures on certain subjects based on databases:

- Automated Insights ▪ AX Semantics
- Heliograf ▪ Narrative Science
- Quakebot ▪ Syllabs
- Associated Press ▪ Aexea
- The Washington Post ▪ Forbes
- Los Angeles Times ▪ Le Monde

What is structured narrative?

The development of verbal texts, understandable by humans, automated from predetermined arrangements that process information from structured databases.

Wikidata

Mbabel

Articles



Structured drafts

- Structured drafts are generated by the Mbabel tool, following the concept of structured narratives.
- Integration of Wikidata and Wikipedia.
- Creation of pre-determined, generic and editable entries.



Structured drafts

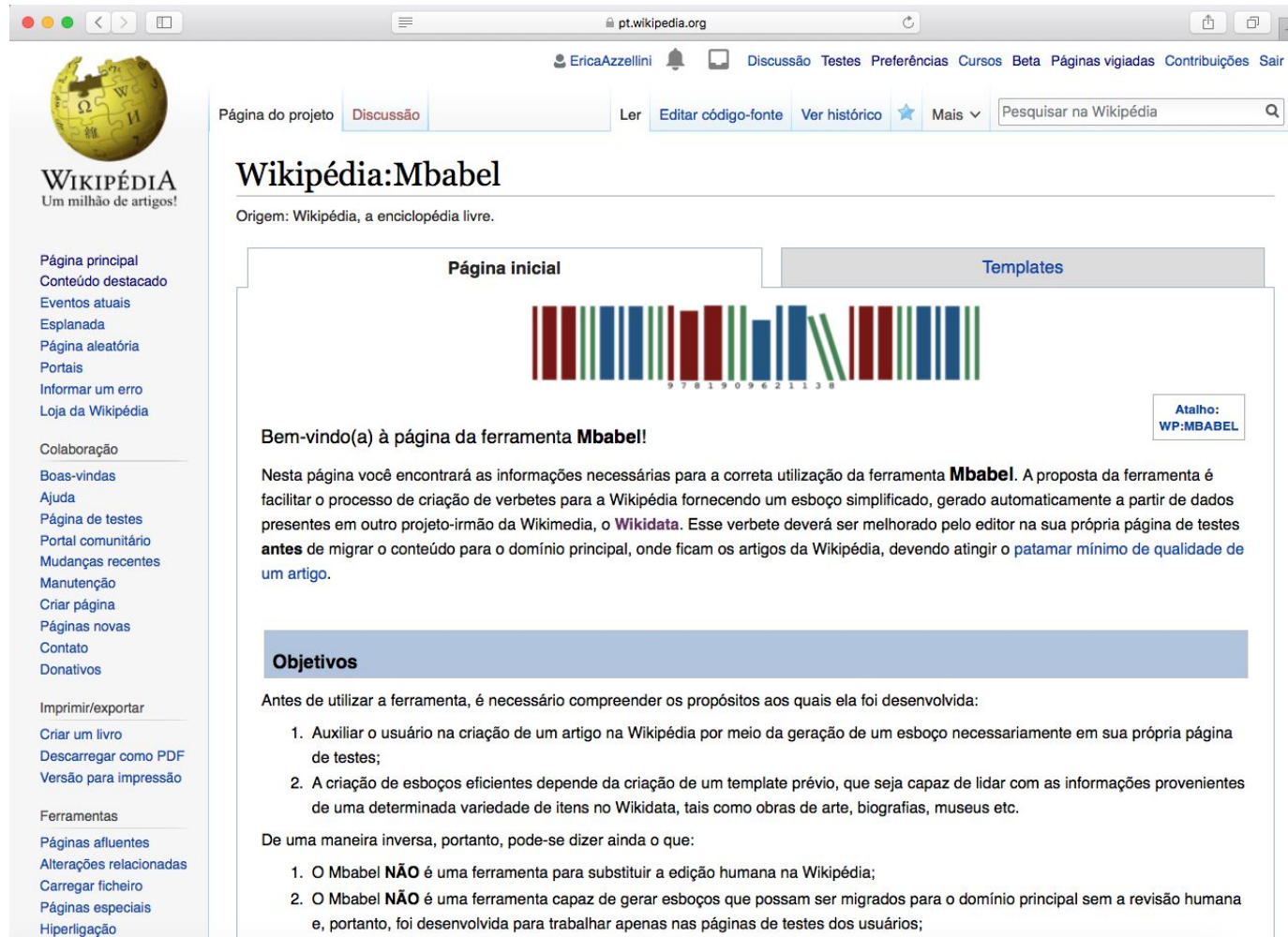
- Geração de rascunhos de verbetes de forma:

Pre-determined: textual frame is pre-made

Generic: covers different qids within a thematic

Editable: manipulable by the editor on the sandbox

Mbabel tool



The screenshot shows the Portuguese Wikipedia page for the Mbabel tool. The page title is "Wikipédia:Mbabel". The main content area features a decorative graphic of vertical bars in red, blue, and green, with numbers 9 7 8 1 9 0 9 6 2 1 1 3 8 below it. Below the graphic, there is a welcome message: "Bem-vindo(a) à página da ferramenta **Mbabel**!". This is followed by a paragraph explaining the tool's purpose: "Nesta página você encontrará as informações necessárias para a correta utilização da ferramenta **Mbabel**. A proposta da ferramenta é facilitar o processo de criação de verbetes para a Wikipédia fornecendo um esboço simplificado, gerado automaticamente a partir de dados presentes em outro projeto-irmão da Wikimedia, o **Wikidata**. Esse verbete deverá ser melhorado pelo editor na sua própria página de testes **antes** de migrar o conteúdo para o domínio principal, onde ficam os artigos da Wikipédia, devendo atingir o **patamar mínimo de qualidade de um artigo**." Below this is a section titled "Objetivos" (Objectives) with two numbered points: "1. Auxiliar o usuário na criação de um artigo na Wikipédia por meio da geração de um esboço necessariamente em sua própria página de testes;" and "2. A criação de esboços eficientes depende da criação de um template prévio, que seja capaz de lidar com as informações provenientes de uma determinada variedade de itens no Wikidata, tais como obras de arte, biografias, museus etc." At the bottom, there is another paragraph: "De uma maneira inversa, portanto, pode-se dizer ainda o que:" followed by two numbered points: "1. O Mbabel **NÃO** é uma ferramenta para substituir a edição humana na Wikipédia;" and "2. O Mbabel **NÃO** é uma ferramenta capaz de gerar esboços que possam ser migrados para o domínio principal sem a revisão humana e, portanto, foi desenvolvida para trabalhar apenas nas páginas de testes dos usuários;". The page also includes a sidebar with navigation links and a top navigation bar with user options and a search box.

Access:

[Wikipédia:Mbabel](https://pt.wikipedia.org/wiki/Wikip%C3%A9dia:Mbabel)

[WP:MBABEL](https://pt.wikipedia.org/wiki/ WP:MBABEL)



Mbabel tool

- Tool generates drafts of automated articles from structured information on Wikidata
- Adaptation and extension of the artwork's Mbabel template created by [User:Pharos](#), from [GLAM The Met](#)
- Development of four different templates and landing page with button for automatic draft generation



What is it for?

- Assist in the first steps of preparing an article on Wikipedia
- Complement features for editing on Wikipedia
- Encourage the integration between Wikidata and Wikipedia
from the manipulation of computational structures in
Wikimedia
- Encourage improvements in data structuring
- Experimentation with linguistic resources



What it is **not** for?

- Does not replace human edition
- Must not be published to primary domain without proper user review
- Not competing with other editing tools



Structured drafts: characteristics

- The structured drafts are Mbabel's templates
- Each template includes Wikidata infobox, introduction, suggestions of sections for the entry and references with identifiers



Structured drafts: gender

- The sentences try as far as possible to avoid gender markings typical of the Portuguese language, in order to sound natural to the editor

Structured drafts: gender

O filme foi dirigido por **P57** (Sophia Coppola/Steven Spielberg)



GENDER NEUTRAL
(PASSIVE VOICE)

P57 é o diretor do filme

MASCULINE

Sophia Coppola é o diretor do filme



Steven Spielberg é o diretor do filme

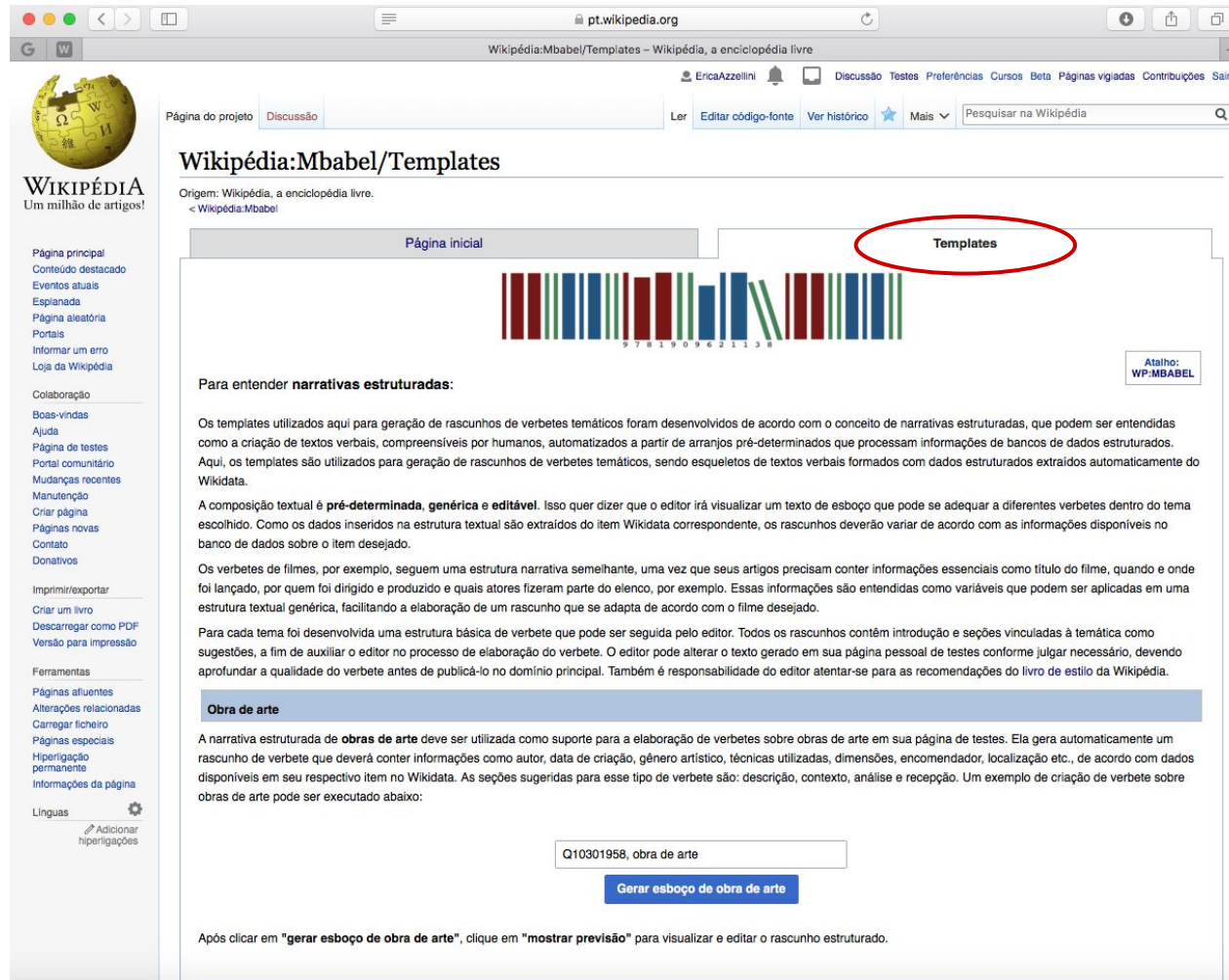




Structured drafts: properties

- The availability of sentences in the structured draft depends on the availability of properties registered in the corresponding item in Wikidata
- The more properties, the more automatic sentences in the draft!

Structured drafts



The screenshot shows the Portuguese Wikipedia page 'Wikipédia:Mbabel/Templates'. The page title is 'Wikipédia:Mbabel/Templates' and the origin is 'Wikipédia, a enciclopédia livre'. The page content is organized into a structured draft format. A navigation bar at the top includes 'Página inicial' and 'Templates', with 'Templates' circled in red. Below the navigation bar is a decorative graphic of vertical bars. The main content area is titled 'Para entender narrativas estruturadas:' and contains several paragraphs explaining the purpose and usage of the templates. A section titled 'Obra de arte' is highlighted in blue. At the bottom, there is a search box containing 'Q10301958, obra de arte' and a button labeled 'Gerar esboço de obra de arte'. The page also features a sidebar on the left with various navigation options and a top navigation bar with links like 'Discussão', 'Testes', and 'Preferências'.

On tab Templates:

- Books
- Movies
- Artwork
- Museums

Mbabel

- It should be used preferably with items from Wikidata that do not have entry on Wikipedia
- Use the button to generate the draft on your sandbox

Qid, obra de arte

Qid, museu

Qid, filme

Qid, livro



Q220741, filme

Gerar esboço de filme

Mbabel

Copy qid

WIKIDATA

Página [Discussão](#)

Pesquisar em Wikidata

Cidade de Deus (Q220741)

filme de Fernando Meirelles e Kátia Lund (2002) [editar](#)

- Em mais idiomas [Configurar](#)

Idioma	Rótulo	Descrição	Também conhecido como
português do Brasil	Cidade de Deus	filme de Fernando Meirelles e Kátia Lund (2002)	
inglês	City of God	2002 film by Fernando Meirelles, Kátia Lund	
português	Cidade de Deus	filme de 2002	
alemão	City of God	Film von Fernando Meirelles und Kátia Lund (2002)	

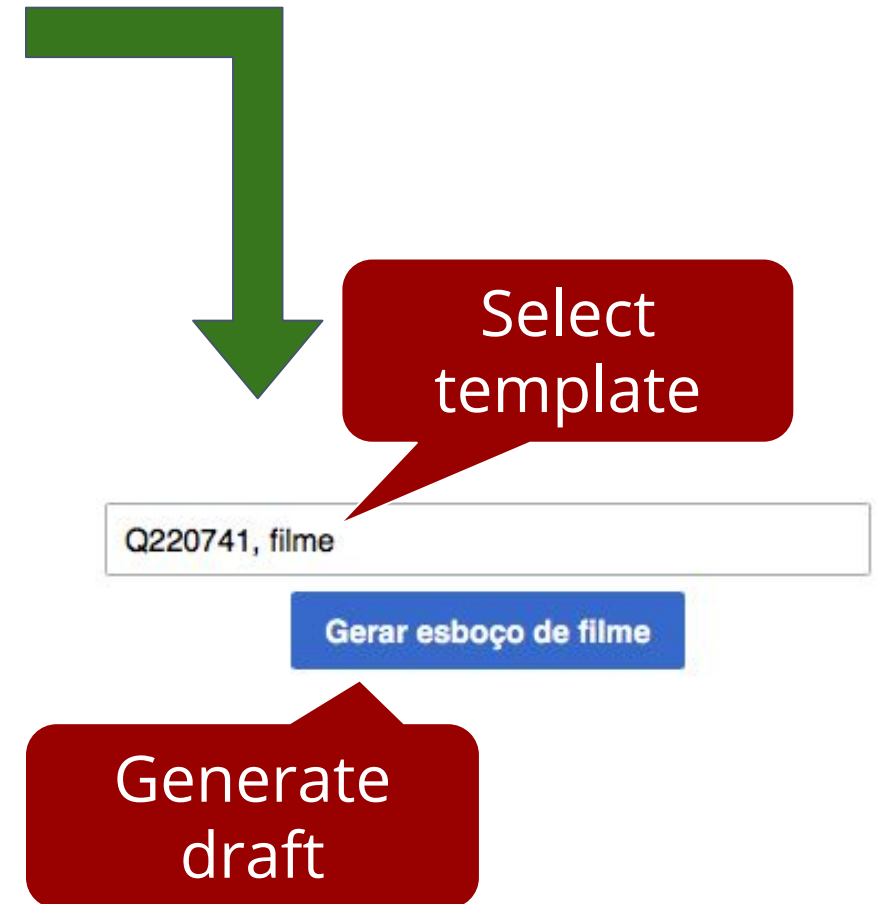
Todos os idiomas introduzidos

Declarações

instância de [filme](#) [editar](#)

[1 referência](#)

[+ adicionar valor](#)



Mbabel

EricaAzzellini

Página de usuário(a) | Discussão

Discussão Testes Preferências Cursos Beta Páginas vigiadas Contribuições Sair

criar Criar código-fonte

Pesquisar na Wikipédia

A criar Usuário(a):EricaAzzellini/Testes/Q220741, filme

Clique em **Mostrar previsão** para visualizar como o rascunho será salvo na sua página de testes.

Atenção! Esta é uma subpágina do(a) usuário(a) EricaAzzellini!

- **Não** deve usar esta página para anúncios, propaganda, blogues, conteúdo de artigos que tenham sido eliminados, ou de qualquer tipo de rede social.
- Pode usá-la para descrever algo sobre si de forma resumida e outras informações relacionadas à Wikipédia.
- Lembre-se de que a Wikipédia não é um servidor de hospedagem gratuita.
- Se colocar muita informação que não seja relevante para o trabalho colaborativo na Wikipédia, a página poderá ser eliminada.
- Suas informações pessoais não devem aparecer nas páginas do domínio principal, que estão reservadas para conteúdo enciclopédico.
- Lembre-se de que sua página de usuário(a) é uma página da Wikipédia sobre você, mas **não lhe pertence**. Portanto, edite-a com responsabilidade.
- Novo(a) na Wikipédia? Seja bem-vindo(a) e tire suas dúvidas com os usuários experientes.
- Caso esta página não seja sua, prefira contatar o(a) usuário(a) EricaAzzellini antes de editar.

N I Edição avançada Caracteres especiais Ajuda

```
{{subst:#switch:{{subst:#titleparts:{{subst:#invoke:String|replace|source={{subst:SUBPAGENAME}}, |count=1}}|2|2}}
obra de arte={{subst:Mbabel/Obra de arte/2|qid={{subst:#titleparts:{{subst:#invoke:String|replace|source={{subst:SUBPAGENAME}}, |count=1}}|1}}}}
museu={{subst:Mbabel/Museu/2|qid={{subst:#titleparts:{{subst:#invoke:String|replace|source={{subst:SUBPAGENAME}}, |count=1}}|1}}}}
livro={{subst:Mbabel/Livro/2|qid={{subst:#titleparts:{{subst:#invoke:String|replace|source={{subst:SUBPAGENAME}}, |count=1}}|1}}}}
filme={{subst:Mbabel/Filme/2|qid={{subst:#titleparts:{{subst:#invoke:String|replace|source={{subst:SUBPAGENAME}}, |count=1}}|1}}}}
Erro ao selecionar a predefinição
}}
```

Inserir Assinar em páginas de discussão: Citar suas fontes:

Resumo:

Vigiar esta página

Ao salvar, você concorda irrevogavelmente em liberar as suas contribuições sob as licenças Creative Commons Attribution/Share-Alike License 3.0 e GFDL. Você concorda em ser creditado por reutilizadores, no mínimo, por meio de uma hiperligação ou URL para a página de usuário(a) contribuindo. Veja as condições de uso para detalhes.

Publicar página **Mostrar previsão** **Mostrar alterações** Cancelar

Click on "show preview"

DO NOT publish the page yet

DO NOT edit the code

Mbabel

Página de usuário(a)

[Discussão](#)

[Criar](#) [Criar código-fonte](#)

A criar Usuário(a):EricaAzzellini/Testes/Q220741, filme

Atenção! Esta é uma subpágina do(a) usuário(a) EricaAzzellini!



- **Não** deve usar esta página para [anúncios](#), [propaganda](#), [blogs](#), conteúdo de artigos que tenham sido [eliminados](#), ou de qualquer tipo de [rede social](#).
- Pode usá-la para descrever algo sobre si de forma resumida e outras informações relacionadas à Wikipédia.
- Lembre-se de que a [Wikipédia não é um servidor de hospedagem gratuita](#).
- Se colocar muita informação que não seja relevante para o trabalho colaborativo na Wikipédia, a página [poderá ser eliminada](#).
- Suas informações pessoais **não** devem aparecer nas páginas do [domínio principal](#), que estão reservadas para conteúdo enciclopédico.
- Lembre-se de que sua página de usuário(a) é uma página da Wikipédia *sobre* você, mas **não lhe pertence**. Portanto, edite-a com responsabilidade.
- Novo(a) na Wikipédia? Seja [bem-vindo\(a\)](#) e [tire suas dúvidas](#) com os usuários experientes.
- Caso esta página não seja sua, prefira [contatar](#) o(a) usuário(a) EricaAzzellini antes de editar.

Previsão da edição

Lembre-se de que esta é apenas uma antevisão do resultado. As modificações ainda não foram gravadas! → Ir para a área de edição

Cidade de Deus é um filme do gênero filme dramático, filme policial, História coming-of-age. Dirigido por Fernando Meirelles, Kátia Lund. O filme é baseado na obra Cidade de Deus, de Paulo Lins. Lançado em 18 de maio de 2002, tendo como idioma original o português, é um filme de Brasil, França. ^{[1][2][3][4]}

Índice [\[esconder\]](#)

- 1 [Sinopse](#)
- 2 [Elenco](#)
- 3 [Produção](#)
- 4 [Prêmios e recepção](#)
- 5 [Referências](#)

Sinopse

Baseado em [Cidade de Deus](#), de Paulo Lins. É seguido pelo filme [Cidade dos Homens](#). O filme falha no Teste de Bechdel, parâmetro utilizado como indicativo de preconceito de gênero em produções culturais.

Elenco

O elenco do filme é composto por Alexandre Rodrigues, Alice Braga, Leandro Firmino, Phellipe Haagensen, Douglas Silva, Jonathan Haagensen, Matheus Nachtergaele, Seu Jorge, Graziella Moretto, Roberta Rodrigues, Paulo Lins, Thiago Martins, Micael Borges, Dani Ornellas, Marcello Melo, Edward Boggis, Mary Sheyla, Daniel Zettel, Darlan Cunha, Gero Camilo, Babu Santana, Charles Paraventi, Olívia Araújo, Sérgio Chapelin, Michel Gomes, Arlindo Lopes.

Cidade de Deus

País	Brasil, França
Lançamento	18 de maio de 2002, 8 de maio de 2003
Diretor(a)	Fernando Meirelles Kátia Lund
	Elenco [Expandir]
Gênero	História coming-of-age filme policial filme dramático
Inspiração	Cidade de Deus
Duração	128 minutos

[\[edite no Wikidata\]](#)

The structured draft is moved to the user's sandbox and can be edited

Mbabel

You can also use Mbabel directly on your sandbox without using the landing page button as follows:

```
{{Mbabel | Qid | filme}}
{{Mbabel | Qid | livro}}
{{Mbabel | Qid | museu}}
{{Mbabel | Qid | obra de arte}}
```



- Structured draft is generated when there is still no entry for the item on Wikipedia
- The user is redirected to the article if the item already has an entry



Understanding the code

- The code is based on a `#invoke`, which invokes a Lua module
- A Lua module is a set of functions in the Lua programming language
- `#invoke` is used as follows: `{{#invoke:<module name>|<function name>|<parameters>}}`

Understanding the code

Module

Function name

```
{{#invoke:WikidataB |getValue ou getPreferredValue |<PropertyID>  
|name=<fieldname> |fetchwikidata=<list of fields to fetch values from  
Wikidata> |onlysourced=<yes/no> |noicon=<yes/no> |<local parameter>}}
```

Parameters

Understanding the code

- The code has many pre-determined parts:

```
{{#invoke:Wikidata|B |getValue ou getPreferredValue |<PropertyID>  
|name=<fieldname> |fetchwikidata=ALL |onlysourced=no/yes |noicon=yes  
|qid={{q|id}}}}
```

Wikidata ID

List of fields to search
for values in Wikidata
(all)

Select only
information with
references or not

Delete the
"Edit in
Wikidata"
button

Understanding the code

```
{{subst:<noinclude  
/>#invoke:WikidataB|getPreferredValue|P31|fetchwikidata=ALL|noicon=sim|  
onlysourced=não|qid={{q|id}}|name=instanciade}}
```

- We use the subst: to add the values taken from Wikidata only once
- We use <noinclude /> in the template code so that the code does not run on the template page

Understanding the code

- Juntamente com o código do `#invoke`, usamos o código `#if`, para que sejam adicionadas no texto apenas informações existentes no Wikidata
- O código `#if` é usado da seguinte forma:

`{{#if:<expressão invoke>|<valor obtido no invoke + texto puro>}}`

`{{subst:<noinclude />#if:<código invoke>| é um .<código invoke>}}`

Examples

```
{{subst:<noinclude />#if:{{subst:<noinclude />#invoke:WikidataB|getLabel|qid={{q|}}|name=label}}|{{subst:<noinclude />#invoke:WikidataB|getLabel|qid={{q|}}}}}}
```

```
{{subst:<noinclude />#if:{{subst:<noinclude />#invoke:WikidataB|getPreferredValue|P31|fetchwikidata=ALL|noicon=sim|onlysourced=não|qid={{q|}}|name=instanciade}}|&nbsp;is a&nbsp;{{subst:<noinclude />#invoke:WikidataB|getPreferredValue|P31|fetchwikidata=ALL|noicon=sim|onlysourced=não|qid={{q|}}}}.}}
```

Don Quixote is a **book**.

Examples

```
{{subst:<noinclude />#if:{{subst:<noinclude />#invoke:WikidataB|getPreferredValue|P50|fetchwikidata=ALL|noicon=sim|onlysourced=não|qid={{qid}}|name=autor}}|The book's authorship is&nbsp;{{subst:<noinclude />#invoke:WikidataB|getPreferredValue|P50|fetchwikidata=ALL|noicon=sim|onlysourced=não|qid={{qid}}}}.}}
```

The book's authorship is **Miguel de Cervantes**.

Examples

```
{{subst:<noinclude />#if:{{subst:<noinclude />#invoke:WikidataB|getPreferredValue|P57 1|fetchwikidata=ALL|noicon =sim|onlysourced=não|qid={{qid}}|name=datafundacao}}|The museum was founded in&nbsp;{{subst:<noinclude />#invoke:WikidataB|getPreferredValue|P57 1|fetchwikidata=ALL|noicon =sim|onlysourced=não|qid={{qid}}}}.}}
```

The museum was founded in **1947**.

Code of Identifiers

- Much like the code we have already seen
- The difference is that we use `#tag:ref` because we are dealing with a reference and we use the "Citar web" template

- The code `#tag:ref` is used as follows:

```
{{#tag:ref|<Citar Web template>}}
```

- The "Citar web" template is used as follows:

```
{{citar web|<url>|<título>|<publicação>|<acessodata>|<name>}}
```

Example



Encyclopædia
Universalis

```
{{subst:<noinclude />#if:{{subst:<noinclude  
</>#invoke:WikidataB|getPreferredValue|P3219|qid={{qid}}|fetchwikidata  
=ALL|onlysourced=no|noicon=sim}}|{{subst:<noinclude />#tag:ref|{{citar  
web|url=https://www.universalis.fr/encyclopedie/{{subst:<noinclude  
</>#invoke:WikidataB|getPreferredValue|P3219|qid={{qid}}|fetchwikidata  
=ALL|onlysourced=no|noicon=sim|maxvals=1}}|título={{subst:<noinclude  
</>#invoke:WikidataB|getLabel|{{qid}}}}|publicado=[[Encyclopædia  
Universalis|Encyclopædia Universalis]]|acessodata={{subst:<noinclude  
</>CURRENTDAY}} de {{subst:<noinclude />CURRENTMONTHNAME}} de  
{{subst:<noinclude />CURRENTYEAR}}}}|name="universalis" }}}}}
```


Towards the future: Lexicographical data

Wikidata:Lexicographical data - Wikidata

Project page Discussion Read Edit View history More Search Wikidata

Wikidata:Lexicographical data

(Redirected from Wikidata:LD)

Translate this page

Other languages: العربية • مصرى • asturianu • български • català • čeština • dansk • Deutsch • English • Esperanto • español • euskara • فارسی • français • galego • ગુજરાતી • magyar • باسا اندونيسيا • Bahasa Indonesia • italiano • 日本語 • 한국어 • македонски • norsk bokmål • नेपाली • Nederlands • polski • português do Brasil • русский • тоҷикӣ • ไทย • Türkçe • українська • 中文

Overview Documentation Development Tools Support for Wiktionary How to help Discussion

Wikidata:Lexicographical data

Welcome on the project page for lexicographical data!

What is lexicographical data?

Since the start of Wikidata in 2012, the multilingual knowledge base was mainly focused on *concepts*: Q-items are related to a thing or an idea, not to the *word* describing it. Since 2018, Wikidata has also stored a new type of data: words, phrases and sentences, in many languages, described in many languages. This information is stored in new types of entities, called Lexemes (L), Forms (F) and Senses (S). You can learn more about the data model on the [documentation page](#).

The structured description of the words will be directly connected to the concepts. It will allow editors to describe precisely all words in all languages, and will be reusable, just like the whole content of Wikidata, by multiple **tools** and **queries**—everything that the community creates to play with words. Lexicographical data can be reused on the Wikimedia projects, and can provide **support for Wiktionary**.

Timeline

- 2012: first discussions about including lexicographical data into Wikidata
- 2013–2016: many discussions with editors and developers, leading to several versions of the development plan
- 2016: start of the development
- 2017: continuing the development of the structure (Wikibase/Lexeme), development of several tools for Wiktionary (SiteLinks)
- May 23rd, 2018: deployment of the first version of lexicographical data (without some of the features)
- 2018–2019: iteration of the project, deployment of new features (Senses, RDF..)

Useful links

- Create a new Lexeme
- First created lexeme on Wikidata
- Frequently asked questions
- Documentation
- Data Model
- Existing tools
- Ideas on tools based on lexicographical data
- Support for Wiktionary
- Development plan
- How to help?
- Property proposals

Categories: Wikidata:Lexicographical data | Wikidata-en | Wikidata:Wiktionary

Access:

[Wikidata:Lexicographical data](#)

[WD:LD](#)



Towards the future: Lexicographical data

- Knowledge on Wikidata was mainly stored according to concepts, not words
- Experiments with lexicographical data on Wikidata
- Storing words, phrases and sentences in multiple languages: lexemes, forms and senses



Towards the future: Lexicographical data

- Automated text generation
- Automated translation
- Dictionary applications
- Language learning tools
- Linguistic research
- Text analysis

Activities

1. Data structuring

Improve items on Wikidata, since structured drafts depend directly on structuring data on Wikidata

2. Using Mbabel

Use the tool to generate structured drafts of the improved items

3. Creation of new templates

Development of two new Mbabel templates:
Earthquakes and Newspapers