

# *Scholia: criando visualizações a partir do Wikidata*

## *A experiência do NeuroMAT*

**Erika Guetti Suca**

**WIKI MOVIMENTO BRASIL e Universidad de São Paulo  
18 de Maio do 2023 São Paulo**

**E-MAIL: [egueti@usp.br](mailto:egueti@usp.br)  
USER: ErikaGuetti**

Grupo de usuários

**WMB**  
Wiki Movimento Brasil

# Erika Guetti Suca



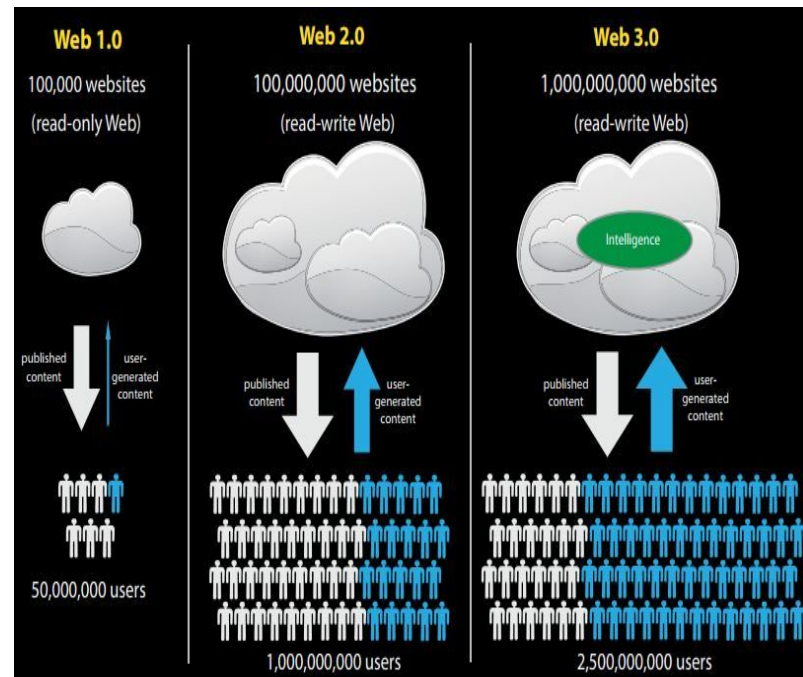
**Estudante de doutorado em ciências da computação da Universidade de São Paulo (USP) no Instituto de Matemática e Estatística (IME), atuando nas áreas de representação do conhecimento, privacidade de dados e web semântica. Membro do Wiki Movimento Brasil.**

Grupo de usuários

**WMB**  
Wiki Movimento Brasil

# A Web Semântica ou Web 3.0

- **Representação dos dados:** Os dados são enriquecidos com **contexto** e **links** significativos e **interpretáveis** por **máquinas** viabilizando que computadores encontrem e manipulem informações com maior precisão.
- **Recuperação dos dados:** as experiências semânticas permitem aos usuários encontrar **experiências** de conteúdo com **menos discordância**, enquanto uma máquina encontra, filtra, classifica, combina, organiza respostas mais.



**Evolução da internet**

# O que é Wikidata?

Banco de dados secundário, livre, colaborativo e multilíngue que coleta dados estruturados.



**Livre:** os dados no Wikidata estão publicados sob a licença Creative Commons Dedicção ao Domínio Público 1.0, que permite a sua reutilização nos mais diversos ambientes.



**Colaborativo:** dados adicionados e mantidos por editores, que definem as regras quanto à criação e ao gerenciamento do conteúdo. Robôs também inserem dados no Wikidata.



**Multilíngue:** A edição, o consumo, a navegação e a reutilização dos dados são completamente multilíngues.



**Banco de dados secundário:** o Wikidata armazena declarações e suas fontes e conexões a outros bancos de dados.



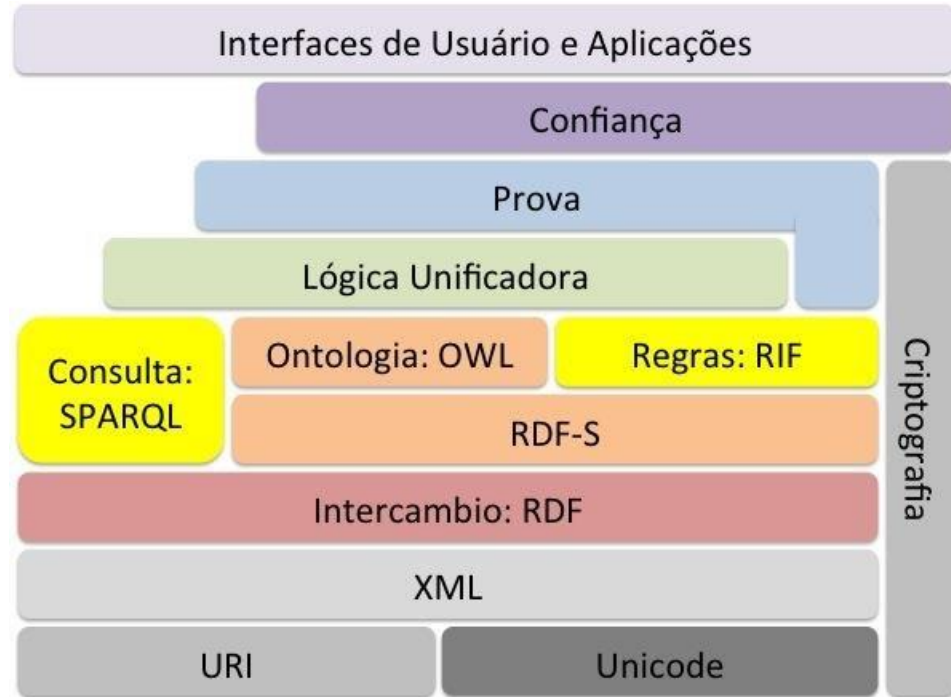
**Suporte:** o Wikidata auxilia a Wikipédia com dados que são mais fáceis de manter em espaços direcionados, reduzindo o trabalho de edição e melhorando a qualidade da informação.



**Dados estruturados:** possui alto grau de organização estruturada e permite o fácil reuso dos dados pelos projetos Wikimedia ou por terceiros.

# Estrutura da Web Semântica

Principais tecnologias, protocolos, padrões de formato que fundamentam a proposta da Web Semântica.



# O que é Scholia?

- Scholia é um projeto que ajuda a apresenta **informações bibliográficas e perfis acadêmicos de autores e instituições** usando Wikidata, o banco de dados com curadoria da comunidade que suporta a Wikipedia e todos os outros projetos da Wikimedia.
- O projeto usa vários **formatos** de exibição disponíveis, **predefinidos**, como listas de publicações para pesquisadores e organizações individuais, publicações por ano, cronogramas de empregos, bem como redes de co-autores e tópicos e gráficos de citações.

# Perfil do NeuroMAT visualizado no Scholia

- <https://scholia.toolforge.org/organization/Q18477654>

**SCHOLIA** Author Work Organization Location Event Project Award Topic Tools Help Search...  
organization / Q18477654 Improve data

## NeuroMat (Q18477654)

The Research, Innovation, and Dissemination Center for Neuromathematics is a Brazilian research center established in 2013 at the University of São Paulo that is dedicated to integrating mathematical modeling and theoretical neuroscience. Among the core missions of NeuroMat are the creation of a new mathematical system to understanding neural data and the development of neuroscientific open-source computational tools, keeping an active role under the context of open knowledge, open science and scientific dissemination. The research center is headed by Antonio Galves, from USP's Institute of Mathematics and Statistics, and is funded by the São Paulo Research Foundation (FAPESP). As of 2019, the co-principal investigators are Oswaldo Baffa Filho (USP), Pablo A. Ferrari (USP/UBA), Fernando da Paixão (UNICAMP), Antonio Carlos Roque (USP), Jorge Stolfi (UNICAMP), and Cláudia D. Vargas (UFRJ). Ernst W. Hamburger (USP) was the former director of scientific dissemination. NeuroMat's International Advisory Board consists of David R. Brillinger, Leonardo G. Cohen (NIH), Markus Diesmann (Jülich), Francesco Guerra, Wojciech Szpankowski (Purdue). ([Read more on English Wikipedia](#)) <http://neuromat.numec.prp.usp.br/>

### Table of Contents

- [Employees and affiliated](#)
- [Co-author graph](#)
- [Advisor graph](#)
- [Topics that employees and affiliates have published on](#)
- [Recent publications](#)
- [Uses](#)
- [Page production](#)
- [Citations](#)
  - [Recent citations](#)



# Perfil do pesquisador principal do NeuroMAT

- <https://scholia.toolforge.org/author/Q17489997>

author / Q17489997

Improve data

## Antonio Galves (Q17489997)

Jefferson Antonio Galves, known simply as Antonio Galves, is a Brazilian mathematician, professor of the Institute of Mathematics and Statistics of the University of São Paulo and member of the Brazilian Academy of Sciences. His field of studies is related to statistician issues models, in particular models that have stochasticity and variable range of memory. In 2007 he won the National Order of Scientific Merit. Professor Galves is also the leader of NeuroMat, a research center established in 2013 at the University of São Paulo that is dedicated to integrating mathematical modeling and theoretical neuroscience.

[\(Read more on English Wikipedia\)](#)

<https://orcid.org/0000-0001-8757-715X>

### Table of Contents

- List of publications
  - Number of publications per year
  - Number of pages per year
- Topics
  - Topic scores
  - Topics of authored works
  - Topics-works matrix
- Use
- Venue statistics
- Review statistics
- Co-author graph
- Co-author map
- Other locations
- Timeline
- Academic tree
- Citation statistics
  - Most cited works
  - Citations by year
  - Citing authors
- Associated images
- Events





[https://www.wikidata.org/wiki/Wikidata:Main\\_Page](https://www.wikidata.org/wiki/Wikidata:Main_Page)

← → ↻ [https://www.wikidata.org/wiki/Wikidata:Main\\_Page](https://www.wikidata.org/wiki/Wikidata:Main_Page)

🇵🇹 português ErikaGuetti Discussão [Preferências](#) [Beta](#) [Páginas vigiadas](#) [Contribuições](#) [Sair](#)

Página principal [Discussão](#) [Ler](#) [Ver fonte](#) [Ver histórico](#) [Mais](#)

**WIKIDATA**

[Página principal](#)  
[Portal comunitário](#)  
[Esplanada](#)  
[Criar um objeto novo](#)  
[Mudanças recentes](#)  
[Página aleatória](#)  
[Serviço de consulta](#)  
[Nas redondezas](#)  
[Ajuda](#)  
[Donativos](#)

[Dados lexicográficos](#)  
[Criar um lexema novo](#)  
[Mudanças recentes](#)  
[Lexema aleatório](#)

[Ferramentas](#)  
[Páginas afluentes](#)  
[Alterações relacionadas](#)  
[Páginas especiais](#)  
[Hierarização](#)

**Ajude a escolher a melhor imagem do ano. Consulte as finalistas e vote na sua preferida.**

[Ajuda-nos com as traduções!]

**Boas-vindas à Wikidata**

a base de conhecimentos livre com 103 185 182 itens de dados que todos podem editar.

[Introdução](#) • [Chat do Projeto](#) • [Portal da Comunidade](#) • [Ajuda](#)

**Quer ajudar a traduzir? Traduza as mensagens que faltam.**

**Bem-vindo(a)!** **Saiba mais sobre dados**

<https://query.wikidata.org/>

The screenshot shows the Wikidata Query Service interface. At the top, there is a navigation bar with the Wikidata logo, the text "Wikidata Query Service", and buttons for "Examples", "Help", "More tools", and "Query Builder". The language is set to "English".

The main area contains a SPARQL query:

```
1 #defaultView:BarChart
2 SELECT DISTINCT ?Year (COUNT(?NeuroMat) AS ?Number_Keywords) ?keyLabel WHERE {
3   ?NeuroMat wdt:P361 wd:Q18477654.
4   ?NeuroMat wdt:P577 ?dataTemp.
5   BIND(STR(YEAR(?dataTemp)) AS ?Year)
6   ?NeuroMat wdt:P921 ?key.
7   SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en".}
8 }
9 GROUP BY ?key ?keyLabel ?Year
```



At the bottom, there is a status bar showing "173 results in 476 ms" and buttons for "Code", "Download", and "Link". The view is currently set to "Bar chart".

# Exemplos de itens criados na Wikidata do NeuroMAT

Pesquisador Antonio Galves

<https://www.wikidata.org/wiki/Q17489997>

## Identificadores

identificador VIAF  12549746  editar



↳ 1 referência

+ adicionar valor

ISNI  0000 0000 3259 0069  editar



↳ 1 referência

+ adicionar valor

identificador BnF  145720057  editar

↳ 1 referência

+ adicionar valor

identificador GND  1146394993  editar

↳ 1 referência

+ adicionar valor

identificador J9U da Biblioteca Nacional de Israel  987007427170305171  editar

↳ 1 referência

+ adicionar valor

# Exemplos de itens criados na Wikidata do NeuroMAT

Artigo do NeuroMAT com informações acadêmicas pertinentes

Infinite Systems of Interacting Chains with Memory of Variable Length—A Stochastic Model for Biological Neural Nets

<https://www.wikidata.org/wiki/Q56592766>

parte de	<div><div><span>NeuroMat</span></div><div><span>financiador</span> <span>Fundação de Amparo à Pesquisa do Estado de São Paulo</span></div><div><span>referido como</span> <span>FAPESP 2013/07699-0</span></div><div><span>► 1 referência</span></div></div> <div><span>+ adicionar valor</span></div>
título	<div><div><span>Infinite Systems of Interacting Chains with Memory of Variable Length—A Stochastic Model for Biological Neural Nets (inglês)</span></div><div><span>► 1 referência</span></div></div> <div><span>+ adicionar valor</span></div>
tema(s) principal(is)	<div><div><div><span>rede neural</span></div><div><span>referido como</span> <span>Biological neural nets</span></div><div><span>▼ 0 referência</span></div><div><span>+ adicionar referência</span></div></div><div><div><span>Sistema de partículas em interação</span></div><div><span>referido como</span> <span>Interacting particle system</span></div><div><span>▼ 0 referência</span></div><div><span>+ adicionar referência</span></div></div><div><div><span>Hawkes Process <sup>inglês</sup></span></div><div><span>referido como</span> <span>Hawkes process</span></div><div><span>▼ 0 referência</span></div><div><span>+ adicionar referência</span></div></div></div>

## Item de dados

Um item é qualquer tópico, pode ser um artigo da Wikipédia, um objeto, uma pessoa, um lugar, um conceito, etc.

Cada item tem sua própria página — onde todos os dados sobre ele são armazenados — e um **identificador único**, composto pela letra Q e um ou mais números.

Seu identificador único é usado para auxiliar pessoas e máquinas a diferenciar itens com nomes idênticos ou similares.

Identificador de item (QID)

Conceição Evaristo (Q16934999)

escritor brasileiro [editar](#)

▼ Em mais idiomas  
configurar

Linguagem	Rótulo	Descrição	Também conhecido como
Inglês	Conceição Evaristo	escritor brasileiro	
Português	Conceição Evaristo	escritora brasileira	
Portugues do Brasil	Conceição Evaristo	Linguista e escritora brasileira	Maria da Conceição Evaristo de...
inglês americano	Nenhum rótulo definido	Nenhuma descrição definida	

[Todos os idiomas inseridos](#)

# Wikidata

## Primeiras informações de um item

**Rótulo:** é o título do item. Cada item tem um rótulo por idioma humano. Rótulos não precisam ser únicos.

**Descrição:** uma frase descritiva específica do item. Ela contextualiza o rótulo evitando ambiguidades para itens com nomes idênticos ou similares. Cada item tem uma descrição por idioma humano.

**Também conhecida como:** nome alternativo pelo qual o item é conhecido. Pode haver mais de um.

The image shows a Wikidata item page for 'Conceição Evaristo' (Q16939999). The page is annotated with three boxes and lines:

- An orange box labeled 'Rótulo' points to the name 'Conceição Evaristo'.
- A pink box labeled 'Descrição' points to the description 'escritor brasileiro'.
- An orange box labeled '“Também conhecida como”' points to the 'Também conhecida como' section, which contains the text 'Maria da Conceição Evaristo de...'.

The main content of the page includes:

- Item name: **Conceição Evaristo** (Q16939999)
- Description: **escritor brasileiro**
- Section: **Em mais idiomas** (configurar)
- Table with columns: **Língua**, **Rótulo**, **Descrição**

Língua	Rótulo	Descrição
Inglês	Conceição Evaristo	escritor brasileiro
Português	Conceição Evaristo	escritora brasileira
Português do Brasil	Conceição Evaristo	Linguista e escritora brasileira
inglês americano	Nenhum rótulo definido	Nenhuma descrição definida

Todos os idiomas inseridos

[edit](#)

## Declarações

Toda a página de um item possui uma seção de "Declarações", que inclui informações — palavras, números, até mesmo arquivos de imagem — sobre o item de dados. Isso pode parecer complicado mas é razoavelmente simples. É importante que as informações inseridas nas declarações sejam verificáveis por meio de fontes.

Declarações		
instância de	<span>ser humano</span> ↳ 2 referências	 editar <a href="#">+ adicionar valor</a>
parte de	<span>Mead &amp; Bateson</span> <i>inglês</i> ↳ 0 referência	 editar <a href="#">+ adicionar referência</a> <a href="#">+ adicionar valor</a>
local de nascimento	<span>Filadélfia</span> ↳ 1 referência	 editar <a href="#">+ adicionar valor</a>
data de morte	<span>15 novembro 1978</span> ↳ 10 referências	 editar <a href="#">+ adicionar valor</a>

# Wikidata

## Estrutura de uma declaração

**Propriedade:** é uma categoria de dados. Da mesma forma que os itens de dados, as propriedades são formadas por números antecedidos por uma letra. Identificamos uma propriedade pela letra P. é uma e possui um nome único e distinto.

**Valor:** descreve o dado de determinada propriedade. O valor pode ser um outro item de dados ou um valor numérico. Podem ser atribuídos vários valores a uma propriedade e cada valor pode ter múltiplas fontes. O tipo de valor permitido depende da propriedade que se deseja editar. Por exemplo, propriedades baseadas em data só permitem a inserção de dados em formato de data.

alma mater	<span>edit</span> <span>+ adicionar referência</span>
Barnard College	<span>edit</span> <span>+ adicionar referência</span>
Universidade Columbia	<span>edit</span> <span>+ adicionar referência</span>
data final	1929
título acadêmico	Doutorado em filosofia



# Wikidata e suas possibilidades

## 01 **É uma alternativa para a preservação de informação que não se encaixa nos critérios de notoriedade da Wikipédia**

Na Wikipédia, os critérios de notoriedade ainda impedem que certas pessoas e conteúdos que não tenham ampla cobertura de fontes secundárias permaneçam na Wikipédia. Nestes casos, o Wikidata pode ser uma alternativa, sendo o primeiro local onde você poderá guardar as informações do tema ou biografia desejada.

## 02 **Oferece mapeamento de dados por meio da ferramenta Wikidata Query Service**

A ferramenta Wikidata Query Service realiza buscas nos dados estruturados do Wikidata, tornando possível visualizar de forma rápida e simples um conjunto de itens por meio de características em comum. Diante disso, você pode formar listagens e agrupamentos úteis em diversas situações, como as listas utilizadas neste evento.

## Possibilidade de trabalhar informações mais sensíveis:

Itens no Wikidata descrevem dados

“brutos” - como data de nascimento, ocupação, sexo e nacionalidade - sobre pessoas, conceitos e etc. Porém, dados também possuem vieses.

Idioma	Rótulo	Descrição	Também conhecido como
português do Brasil	grupo étnico	Autodeclaração de cor, raça ou etnia (segundo o IBGE)	cor ou raça
inglês	ethnic group	subject's ethnicity (consensus is that a VERY high standard of proof is needed for this field to be used. In general this means 1) the subject claims it themselves, or 2) it is widely agreed on by scholars, or 3) is fictional and portrayed as such)	ethnicity culture people (cultural) nationality race ethnic or cultural origin

Os requisitos para inserção da propriedade grupo étnico em um item no Wikidata são mais elevados do que o normal, pois a comunidade de editores considera dados desse tipo como controversos.

Porém, quando não há no item de uma pessoa qualquer informação que a identifique como negra, pesquisas no Wikidata com esse recorte são dificultadas. Se desejássemos fazer uma busca por todas as escritoras negras brasileiras, por exemplo, teríamos apenas 4 resultados.

Em suma, considerar dados étnicos como controversos leva à invisibilização de pessoas negras.

# Critérios de notoriedade do Wikidata

- ✔ Um item é considerado notório se puder **existir enquanto um verbete** da Wikipédia ou de qualquer outro projeto Wikimedia.
- ✔ Um item é considerado notório se referir a uma **entidade conceitual**, como feminismo, claramente identificáveis e que possa ser descrita por meio de referências reputadas.
- ✔ Um item é considerado notório se referir a uma **entidade material**, como lápis, claramente identificável e que possa ser descrita por meio de referências reputadas.
- ✔ Um item é considerado notório quando é enriquecido com informações.

# Bibliografia

- Scholia, Scientometrics and Wikidata. Nielsen, F.Å., Mietchen, D., Willighagen, E. 2017.
- Investigating the potential of the semantic web for education: Exploring Wikidata as a learning platform. Evenstein Sigalov, S., Nachmias, R. 2023.
- Perfil no Scholia do NeuroMAT
- <https://www.wikidata.org/wiki/Wikidata:Scholia>
- Vitrine NeuroMAT - USP
- <https://vitrine.numec.prp.usp.br/>
- <https://github.com/eguetti/VitrineNeuroMAT>
- Apresentação do criador do Scholia
- [https://commons.wikimedia.org/wiki/File:WikidataCon\\_2017\\_-\\_Scholia.web](https://commons.wikimedia.org/wiki/File:WikidataCon_2017_-_Scholia.web)
-

# Exemplos SPARQL usando dados do NeuroMAT

Exemplos não criados usando Scholia

- Query QUALIS das publicações do NeuroMAT
- <https://w.wiki/6GCCG>
- O grafo de co autores por nacionalidade da afiliação é gerado por esta consulta (Muito legal de ver!)
- <https://w.wiki/6UTU>
- Item do professor Antonio Galves
- <https://www.wikidata.org/wiki/Q17489997>
- Artigo do NeuroMAT com informações acadêmicas
- <https://www.wikidata.org/wiki/Q56592766>
- Infinite Systems of Interacting Chains with Memory of Variable Length—A Stochastic Model for Biological Neural Nets
-

**Obrigada :)**