

# Wikidata: Uma Fonte Aberta para a Gestão e Interoperabilidade da Informação

**André Barbosa**

*AWikimedia Portugal / Universidade do Porto*

Wikidata	<p>O Wikidata é uma base de conhecimento livre, colaborativa e estruturada que se destaca como um repositório central de dados interligados, sustentando iniciativas open data e ciência aberta. Criado pela Wikimedia Foundation para dar suporte aos projetos Wikimedia, a sua evolução levou-o ao ponto onde se assume como elemento central das redes de dados interligados, permitindo o armazenamento e a reutilização de dados estruturados para diversas aplicações, desde bibliotecas e catálogos digitais até inteligência artificial. No contexto da ciência da informação, o Wikidata apresenta-se como um recurso inovador para bibliotecas, arquivos e museus, promovendo a interoperabilidade e a acessibilidade dos dados conforme as melhores práticas. A sua arquitetura permite a ligação entre diferentes bases de dados, facilitando a normalização da informação e enriquecendo catálogos bibliográficos com identificadores únicos e referências cruzadas. Nesta comunicação, exploramos o papel do Wikidata como fonte aberta na gestão e disseminação do conhecimento. Destacamos as suas vantagens, como a atualização colaborativa e a flexibilidade na estruturação de dados, assim como desafios relacionados com a curadoria e a qualidade da informação, usando quatro casos de estudo de diferentes tipologias.</p>
Wikimedia	
Dados abertos	
Interoperabilidade	
Gestão da informação	

# Wikidata: An Open Source for Information Management and Interoperability

Wikidata	<p>Wikidata is a free, collaborative and structured knowledge base that stands out as a central repository of linked data, supporting open data and open science initiatives. Created by the Wikimedia Foundation to support Wikimedia projects, its evolution has taken it to the point where it is the central element of networks of linked data, allowing the storage and reuse of structured data for various applications, from libraries and digital catalogues to artificial intelligence. In the context of information science, Wikidata is an innovative resource for libraries, archives and museums, promoting interoperability and data accessibility in line with best practices. Its architecture makes it possible to link diferente databases, facilitating the standardisation of information and enriching bibliographic catalogues with unique identifiers and cross-references. In this communication, we explore the role of Wikidata as an open source for managing and disseminating knowledge. We highlight its advantages, such as collaborative updating and flexibility in structuring data, as well as challenges related to curation and information quality, using four case studies of different typologies.</p>
Wikimedia	
Open data	
Interoperability	
Information management	

## INTRODUÇÃO

O Wikidata é uma base de conhecimento livre, colaborativa, estruturada e multilíngue, parte integrante do ecossistema de projetos colaborativos e abertos da Fundação Wikimedia. Foi criado em 2012, com o objetivo de fornecer dados estruturados aos projetos wikimedia, centralizando informações como ligações inter-línguas e dados de infocaixas. Isso permitiu reduzir a complexidade e o esforço da gestão da informação, e o aumento da consistência da informação nos vários projetos Wikimedia (Roth, 2012).

Partilha dos princípios da wikipédia, nomeadamente a disponibilização do seu conteúdo sob licença Creative Commons CC0 (Wikidata:Introduction - Wikidata, sem data), que permite a reutilização dos dados em diferentes cenários, permitindo a cópia, modificação, análise, distribuição e visualização dos mesmos, e tornando o projeto num Open Educational Resource (OER). Visa permitir que os dados sejam utilizados tanto na Wikipédia como em aplicações externas (Ford & Iliadis, 2023). Para isso, os dados podem ser exportados mediante serviços Web em vários formatos, incluindo JavaScript Object Notation, ou JSON, e Resource Description Framework, ou RDF (Erxleben et al., 2014; Vrandečić & Krötzsch, 2014).

A conjunção do espírito colaborativo do movimento wikimédia com uma API forte e robusta faz com que sejam desenvolvidas inúmeras ferramentas de apoio à edição em massa e visualização da informação. A par da API e de outras formas de obtenção dos dados, como dumps ou o sistema de exportação de informação, a Fundação Wikimedia disponibiliza o Wikimedia Cloud Services, e, particularmente o Wikimedia Toolforge, um serviço que permite, entre outras coisas, hospedar ferramentas de software para o movimento Wikimedia sem custo (Help:Cloud

Services introduction, sem data). É neste ecossistema que são disponibilizadas diversas ferramentas de edição e visualização de dados, criadas e mantidas essencialmente por uma comunidade de desenvolvedores, na sua maioria voluntários.

Se por um lado temos uma comunidade defensora do acesso livre ao conhecimento, que desenvolve e mantém um conjunto de projetos de livre acesso, estranhamente a adesão por parte da comunidade GLAM (Galleries, Libraries, Archives and Museums) têm sido tímida, embora nos últimos anos vários projetos apontem para uma inversão dessa tendência (Correia, 2023).

### 1. Wikidata

O Wikidata utiliza o mediawiki, o software por detrás da Wikipédia, com uma extensão para a gestão de dados estruturados, o Wikibase. Este permite que a informação seja descrita num sistema de tripla: Item -> Propriedade -> Valor. Um item é o objeto a ser descrito, podendo ser uma pessoa, um local, um acontecimento, etc.; uma propriedade será o fragmento de informação relevante ao item, por exemplo, a altura do item, ou as coordenadas onde se encontra; e o Valor corresponderá a outro item ou a um valor literal. Neste sistema de tripla cada item têm um identificador único que começa com a letra Q, e as propriedades seguem o mesmo sistema, começando com a letra P. O sistema permite outro tipo de entidades, como "Labels", "Values", "Descriptions", "Aliases", "Rank" ou "Qualifiers" (Evenstein Sigalov & Nachmias, 2023).

Assim, para modelizar "A capital de Portugal é Lisboa", no wikidata temos o seguinte:

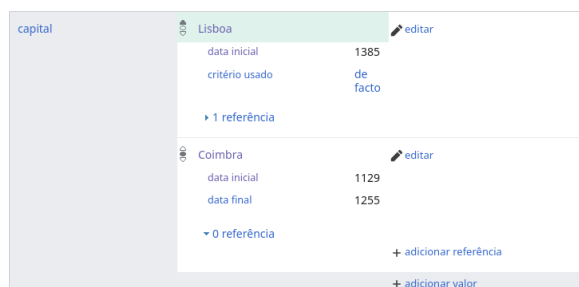
Q45 -> P36 -> Q597

a que corresponde

Portugal -> Capital -> Lisboa

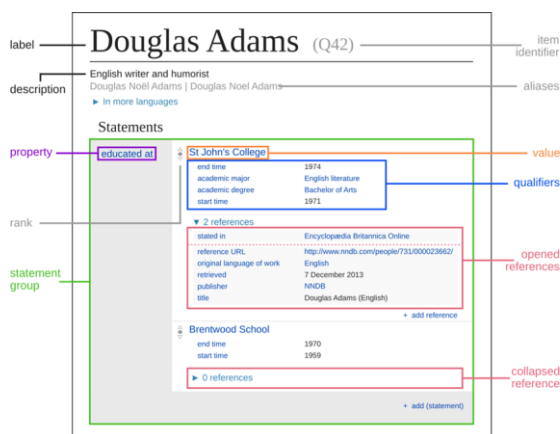
Mas Lisboa não foi sempre a capital de Portugal, e isso também é modelável usando o wikibase. Assim, ao valor Q597 estará associado um qualitativo e o valor correspondente, no caso específico o P580, ou seja, data inicial, e o ano. Ao valor Q597 estão ainda associados mais dois

qualificadores: um para descrever o critério usado para definir a capital e outro qualificador para registar a referência usada.



**Figura 1** - Representação gráfica de como a propriedade Capital está modelizada dentro do item de Portugal (Q45)

Na figura seguinte podemos visualizar com detalhe alguns destes componentes, presentes no item Wikidata de Douglas Adams.



**Figura 2** - 1. Representação gráfica modelo de dados no Wikidata com um grupo de afirmações e referências abertas. (Kritschmar, 2016)

Uma parte importante do modelo de dados do Wikidata é a secção dos identificadores externos, como bases de dados de referência, identificadores das bibliotecas nacionais, identificadores de instituições culturais, entre inúmeras outras bases de dados de referência (Bianchini et al., 2021; Evenstein Sigalov & Nachmias, 2023). Este é um fator que acrescenta uma importância extra ao wikidata (Veen, 2019). Não se trata apenas de modelizar dados, mas também de ligar várias bases de dados num ponto central, e o Wikidata está a tornar-se no elemento centralizador da informação.

Além de uma API forte e robusta, herdada do mediawiki, o wikidata permite a realização de

exportação RDF, e queries (Bianchini et al., 2021; Erxleben et al., 2014). O Wikidata Query Service (WDQS) disponível em <https://query.wikidata.org/>, permite fazer queries usando SPARQL, uma linguagem de consulta semântica capaz de recuperar e manipular dados armazenados em formato RDF (Reis & Correia, 2019).

## 2. Casos de estudo

### Universidade de Edimburgo, Escócia: O projeto de caça às bruxas

Em 2003, investigadores da Universidade de Edimburgo (UoE) reuniram e sistematizaram documentação histórica no âmbito do projeto

The Survey of Scottish Witchcraft (1563–1736), uma base de dados digital que documenta cerca de 4.000 indivíduos acusados de bruxaria na Escócia. Quase 15 anos mais tarde, a base de dados estava inacessível ao público e já não era mantida, mas continha dados de alta qualidade provenientes de fontes fiáveis, selecionados por investigadores, sendo integrada no Wikidata (Evenstein Sigalov & Nachmias, 2023). Essa integração enriqueceu os dados originais com os já presentes no wikidata.

Após a integração, foram feitas queries aos dados para disponibilizar a informação em mapas, e integradas num novo sítio web. Hoje o projeto é largamente reconhecido, e um exemplo de como dados abandonados podem ganhar uma nova e inesperada vida (witches, n.d).

The Metropolitan Museum of Art, Estados Unidos: The Depiction Game

Lançado em 2019, o Wikidata game “Depiction”, permite que voluntários adicionem mais informação à base de dados, sobre nomeadamente o que é retratado no quadro ou imagem. Para este projeto, foi usada IA para analisar as imagens e sugerir aos participantes para confirmar essas informações (Evenstein Sigalov & Nachmias, 2023). Isto permite enriquecer a base de dados da instituição, ao mesmo tempo que envolve os participantes na descoberta das suas coleções.

## CONCLUSÃO

O wikidata apresenta-se cada vez mais como o ponto central no mundo Linked Open Data, funcionando com elo entre diversas bases de dados, de diferentes tipos e categorias (Godby et al., 2024; Veen, 2019). Essas ligações permitem o enriquecimento das bases de dados locais, e, em alguns casos, numa lógica de informação federada.

O ecossistema de ferramentas de visualização, assim como a API e o Wikidata Query System facilitam a extração e visualização da informação (Nielsen et al., 2017; Reis & Correia,

## GNB Alemão

Em 2019, a Biblioteca Nacional da Alemanha, em conjunto com a Wikimedia Alemanha, avaliou a adequação da Wikibase como software para a manutenção do Ficheiro de Autoridade Integrado. Desde 2020 que é usado o Wikibase para a gestão do Ficheiro de Autoridade Integrado, e as duas entidades continuam a trabalhar em conjunto para uma maior integração do Wikibase no seio da Biblioteca Nacional da Alemanha. (Fischer & Nationalbibliothek, 2024).

## SCHOLIA

O scholia é uma aplicação web que cria perfis académicos visuais para tópicos, pessoas, organizações, elementos, etc. utilizando informações bibliográficas e outras informações constantes no Wikidata. Utiliza o WDQS para obter a informação do wikidata, não sendo necessária qualquer outra intervenção humana além da adição dos dados (Nielsen et al., 2017; Reis & Correia, 2019).

Em 2019, investigadores da NOVA SBE desenvolveram um projeto-piloto de importação de dados de investigadores da instituição para o Wikidata, para posterior consulta e visualização no Scholia (Reis & Correia, 2019). O perfil pode ser visualizado em

<https://scholia.toolforge.org/organization/Q392754>

2019; Tharani, 2021), podendo ser um fator decisivo em instituições com poucos recursos para plataformas de gestão da informação. O envolvimento comunitário pode alavancar novos projetos, como no *The Survey of Scottish Witchcraft (1563–1736)*.

## Ligações web úteis

<https://www.wikidata.org>

<https://outreach.wikimedia.org/wiki/GLAM>

<https://wikidatadays.wikimedia.pt/>

<https://witches.is.ed.ac.uk/>

<https://wikidata->

[game.toolforge.org/distributed/#game=49&op\\_t=%7B%22type%22%3A%22dog%22%7D](https://game.toolforge.org/distributed/#game=49&op_t=%7B%22type%22%3A%22dog%22%7D)  
<https://wikidata-game.toolforge.org/distributed/#>  
<https://wiki.dnb.de/pages/viewpage.action?pageId=167019461>  
<https://scholia.toolforge.org/>

## REFERÊNCIAS

- Bianchini, C., Bargioni, S., & Pellizzari di San Girolamo, C. C. (2021). Beyond VIAF. *Information Technology and Libraries*, 40(2). <https://doi.org/10.6017/ital.v40i2.12959>
- Correia, M. M. (2023). Wikidata: Uma central para autoridades ligadas. *Actas do Congresso Nacional de Bibliotecários, Arquivistas e Documentalistas*, 14, Artigo 14. <https://doi.org/10.48798/congressobad.3005>
- Erxleben, F., Günther, M., Krötzsch, M., Mendez, J., & Vrandečić, D. (2014). Introducing Wikidata to the Linked Data Web. Em P. Mika, T. Tudorache, A. Bernstein, C. Welty, C. Knoblock, D. Vrandečić, P. Groth, N. Noy, K. Janowicz, & C. Goble (Eds.), *The Semantic Web – ISWC 2014* (pp. 50–65). Springer International Publishing. [https://doi.org/10.1007/978-3-319-11964-9\\_4](https://doi.org/10.1007/978-3-319-11964-9_4)
- Evenstein Sigalov, S., & Nachmias, R. (2023). Investigating the potential of the semantic web for education: Exploring Wikidata as a learning platform. *Education and Information Technologies*, 28(10), 12565–12614. <https://doi.org/10.1007/s10639-023-11664-1>
- Fischer, B. K., & Nationalbibliothek, D. (2024, novembro 13). Why we use Wikibase at the German National Library. <https://doi.org/10.5281/zenodo.14144623>
- Ford, H., & Iliadis, A. (2023). Wikidata as Semantic Infrastructure: Knowledge Representation, Data Labor, and Truth in a More-Than-Technical Project. *Social Media + Society*, 9(3), 20563051231195552. <https://doi.org/10.1177/20563051231195552>
- Godby, J., Smith-Yoshimura, K., Washburn, B., Davis, K. K., Detling, K., Eslao, C. F., Folsom, S., Li, X., McGee, M., Miller, K., Moody, H., Thomas, C., & Tomren, H. (2024, agosto 22). Creating Library Linked Data with Wikibase: Lessons Learned from Project Passage. OCLC. <https://www.oclc.org/research/publications/2019/oclcresearch-creating-library-linked-data-with-wikibase-project-passage.html>
- Help:Cloud Services introduction—Wikitech. (sem data). Obtido 8 de abril de 2025, de [https://wikitech.wikimedia.org/wiki/Help:Cloud\\_Services\\_introduction](https://wikitech.wikimedia.org/wiki/Help:Cloud_Services_introduction)
- Kritschmar, C. (2016). English: Graphic representing the datamodel in Wikidata with a statement group and opened references [Gráfico]. Obra do próprio. [https://commons.wikimedia.org/wiki/File:Datamodel\\_in\\_Wikidata.svg](https://commons.wikimedia.org/wiki/File:Datamodel_in_Wikidata.svg)
- Nielsen, F. Å., Mietchen, D., & Willighagen, E. (2017). Scholia and scientometrics with Wikidata (No. arXiv:1703.04222). arXiv. <https://doi.org/10.48550/arXiv.1703.04222>
- Reis, A. C. M., & Correia, M. M. (2019). Q392754: Criação de perfis académicos com Wikidata e Scholia – projecto-piloto. *PRISMA.COM*, 40, Artigo 40.
- Roth, M. (2012, março 30). The Wikipedia data revolution. Diff. <https://diff.wikimedia.org/2012/03/30/the-wikipedia-data-revolution/>
- Tharani, K. (2021). Much more than a mere technology: A systematic review of Wikidata in

libraries. The Journal of Academic Librarianship, 47(2), 102326. <https://doi.org/10.1016/j.acalib.2021.102326>

Veen, T. van. (2019). Wikidata: From “an” Identifier to “the” Identifier. Information Technology and Libraries, 38(2), Artigo 2. <https://doi.org/10.6017/ital.v38i2.10886>

Vrandečić, D., & Krötzsch, M. (2014). Wikidata: A free collaborative knowledgebase. Commun. ACM, 57(10), 78–85.

<https://doi.org/10.1145/2629489>

Wikidata:Introduction—Wikidata. (sem data). Obtido 6 de abril de 2025, de <https://www.wikidata.org/wiki/Wikidata:Introduction>

Witches. (sem data). Obtido 19 de abril de 2025, de <https://witches.is.ed.ac.uk/about>

Barbosa, André (2025). "Wikidata: Uma Fonte Aberta para a Gestão e Interoperabilidade da Informação". Cadernos BAD, n. 1-2. <https://doi.org/10.48798/cadernosbad.3092>

**Acesso e licença**

Artigo em acesso aberto distribuído nos termos da licença Creative Commons Atribuição 4.0 Internacional (CC-by 4.0).

**Revisão por Pares**

Esta revista usa um sistema de revisão duplamente cega por pares assegurada pelo conselho científico da Cadernos BAD.

**Financiamento, apoio e patrocínios**

**Conflitos de Interesse**

O autor declara a inexistência de conflitos de interesse na realização do presente trabalho.

**Confidencialidade dos Dados**

O autor declara ter seguido os protocolos de RGPD.

**Recebido**

23/02/2025

**Aceite**

17/03/2025

**Publicado**

27/07/2025