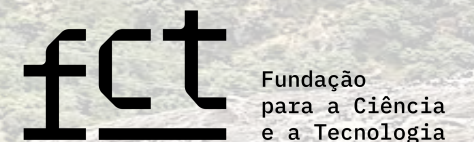




Lhenguabase como recurso lexicográfico do mirandês

Ana Salgado & Luís Trigo

FIESTA DE LAS LHÉNGUAS | 2025
20 de setembro



Agenda

Apresentação da Lhenguabase

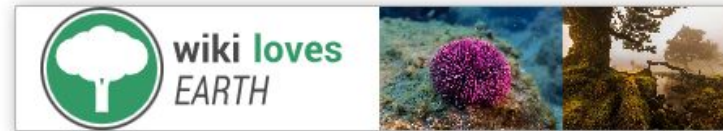
Dar a conhecer a Lhenguabase como recurso lexicográfico para o mirandês.

<https://lhenguabase.wikibase.cloud>

Tecnologias abertas e ecossistema global

O uso de tecnologias abertas como a Wikibase permite que o mirandês seja integrado num ecossistema global de dados multilingues, promovendo a interoperabilidade e a sustentabilidade.





Contents hide

- (Top)
- Life and learning
- Murray and the OED
- In literature and film
- Honorary degrees
- References
- Further reading
- External links

James Murray (le

Article Talk

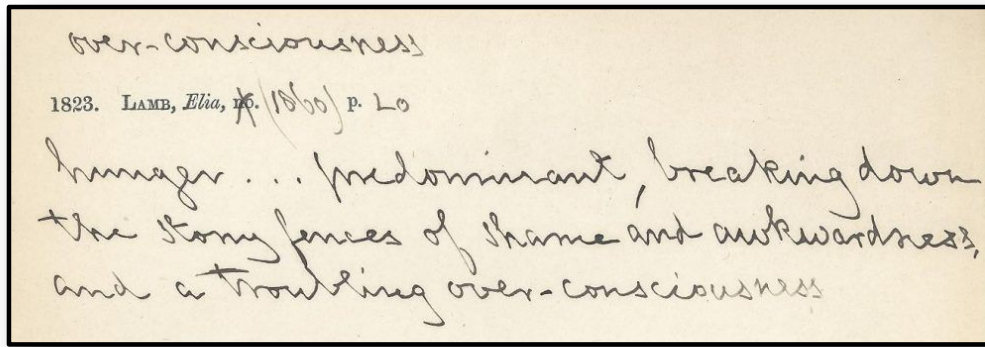
Read Edit View history Tools

Appearance hide

From Wikipedia, the free encyclopedia

For other people with the same name, see James Murray.

Sir James Augustus Henry Murray, FBA (/ˈmʌri/; 7 February 1837 – 26 July 1915) was a British lexicographer and philologist. He was the primary editor of the *Oxford English Dictionary* (OED) from 1879 until his death.^{[1][2]}

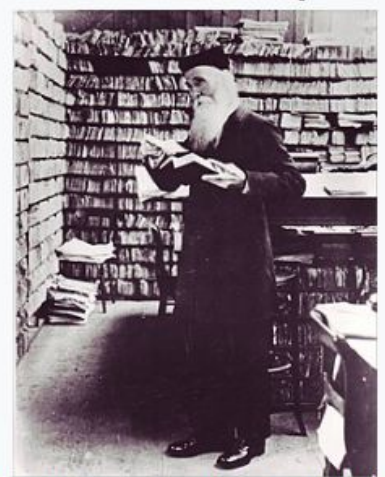


Henry' in order to distinguish himself from other James Murrays in the Hawick area.^[4] A precocious child with a voracious appetite for learning, he left school at fourteen because his parents were not able to afford to pay the

Contributors to the OED

The OED has always been a collaborative project with contributions from volunteer readers, proof-readers, consultants, and staff. It can be regarded as one of the first academic crowdsourcing projects of its kind. This was emphasized in each of the prefaces which accompanied the fascicles of the First Edition as they were issued.

Sir James Murray



Murray in the Scriptorium at Banbury Road, before 1910

Born James Murray
7 February 1837
Denholm, Roxburghshire, Scotland

- Text
- Small
 - Standard
 - Large
- Width
- Standard
 - Wide
- Color (beta)
- Automatic
 - Light
 - Dark

Porquê recursos digitais para o mirandês?

Escassez de recursos

Os materiais lexicográficos existentes são limitados e frequentemente desatualizados.

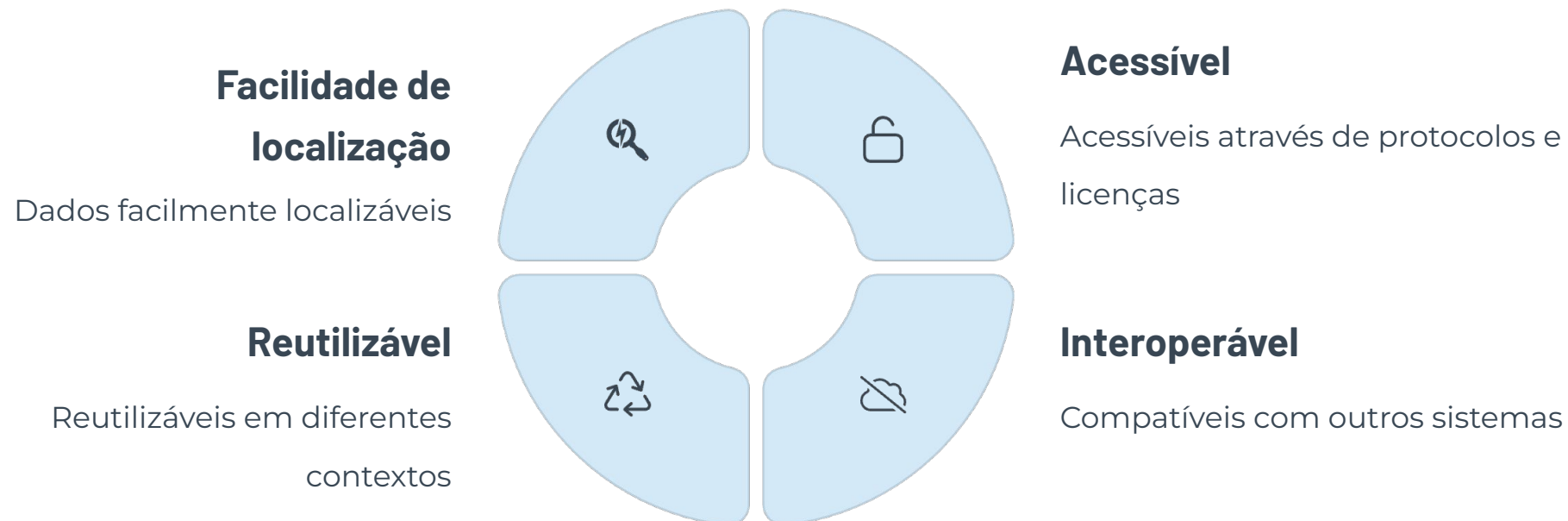
Ausência no digital

A ausência de recursos digitais acessíveis limita o ensino, a investigação e a documentação sistemática do mirandês.

Importância estratégica

Recursos digitais são essenciais para documentação, ensino e investigação, garantindo a preservação e transmissão da língua.

Princípios FAIR Data



❌ Dicionários em PDF ou papel **não** são FAIR - limitam a análise, acesso e reutilização dos dados.

Boas práticas em línguas minorizadas

Projetos internacionais demonstram o potencial dos recursos digitais para línguas minoritárias:

Galego: Termigal

Portal terminológico que oferece recursos lexicográficos especializados, promovendo o uso técnico e científico da língua galega.

Basco: Elhuyar Hiztegia

Dicionário digital multilíngue que serve de modelo para a digitalização de recursos lexicográficos em línguas minoritárias.

Quéchuwa: Qichwabase

Base de conhecimento da língua e comunidade quéchuwa resultante da colaboração de académicos e comunidade.

<https://qichwa.wikibase.cloud/>

Estas iniciativas demonstram que é possível criar recursos digitais sustentáveis, servindo de inspiração para o desenvolvimento do mirandês.

OS PRIMEIROS PASSOS...

U.PORTO
FLUP FACULDADE DE LETRAS
UNIVERSIDADE DO PORTO

#oficina | 16.janeiro.2025 | FLUP

Lexicografia

e sus•ten•ta•bi•li•da•de lin•guís•ti•ca

documentação do Mirandês com
recurso à Wikidata

Auditório Nobre
10h00 Abertura
10h10 *Componentes estruturais para um dicionário mirandês*
Ana Salgado _ Academia de Ciências de Lisboa | NOVA CLUNL
11h30 Pausa para café
11h45 *Estrutura dos lexemas na Wikidata: exemplos de outras línguas*
David Lindemann _ Universidade do País Basco | Wikimedia País Basco
13h00 Almoço

Sala de computadores 206
14h00 *Mãos na massa com dados do mirandês*
16h20 *Recolha fonológica com Lingua Libre*
André Barbosa _ Wikimedia Portugal
Carlos Silva _ Wikimedia Portugal | CLUP
Ana Afonso _ Associação de Lhéngua i Cultura Mirandesa
17h00 Fecho


+ info inscrições




Organização



Apoio



II PORTO MEETING




Inscrição

Encontro Wikimedia de Línguas Minoritárias


<https://www.portomeeting.wikimedia.pt>

- 21 MAI** Oficina: Línguas minoritárias e Tecnologias de Processamento da Linguagem
FLUP | online
- 22 MAI** Dos falantes aos dados estruturados e abertos
FLUP | online
- 23 MAI** Ferramentas Wiki na preservação e revitalização das línguas
FLUP | online
- 24 MAI** Onde moram as línguas? As comunidades ibéricas em foco
Miranda do Douro
- 25 MAI** Memórias, lugares i futuro de la lhéngua mirandesa
Miranda do Douro

Organização



Apoio



O que é a Lhenguabase?

Objetivos principais

- Reunir dados lexicais do mirandês de forma sistemática
- Promover colaboração entre investigadores e falantes nativos
- Garantir estruturação e interoperabilidade dos dados
- Facilitar a reutilização em diferentes contextos
- Apoiar o processo de ensino-aprendizagem do mirandês
- Incentivar o uso ativo da língua



A Lhenguabase nasceu de uma oficina sobre Lexicografia e Sustentabilidade Linguística realizada na Faculdade de Letras da Universidade do Porto, demonstrando o compromisso académico com a preservação das línguas minorizadas.

Características da Lhenguabase



Repositório lexicográfico

Funciona como repositório central de dados lexicográficos do mirandês, organizados de forma estruturada e acessível.



Alojamento em wikibase.cloud

Alojado em wikibase.cloud, beneficia de ligação direta à web semântica e aos padrões internacionais de dados abertos.



Estrutura colaborativa

Estrutura aberta, colaborativa e multilingue que permite contribuições de diferentes utilizadores e comunidades.



Inspiração internacional

Inspirado por projetos similares.

Estrutura de uma entrada

Exemplo: "xaboneteira"

- 1 Forma**
Representação gráfica da palavra tal como aparece na norma ortográfica do mirandês
- 2 Definição**
Descrição clara e precisa do significado da entrada lexical
- 3 Variante**
Formas alternativas regionais ou históricas da mesma unidade lexical
- 4 Equivalentes**
Correspondências em português, espanhol, leonês e outras línguas relacionadas
- 5 Fonte**
Referência bibliográfica ou documental que atesta o uso da forma
- 6 Observações**
Informações adicionais sobre uso, registo ou características especiais

Modelo de dados e interoperabilidade

Modelo adaptado

Estrutura de dados especificamente adaptada à realidade sociolinguística do mirandês



Classes e propriedades

Sistema de classes e propriedades personalizadas para representar fenómenos linguísticos específicos

Ligações externas

Conectividade com sistemas externos como Wikidata, garantindo interoperabilidade global

Interface e usabilidade



Visita à página

Interface web intuitiva e responsiva, acessível a partir de qualquer dispositivo com ligação à Internet



Navegação simples

Sistema de navegação claro que permite exploração fácil das entradas lexicais, com funcionalidades de pesquisa avançada



Edição colaborativa

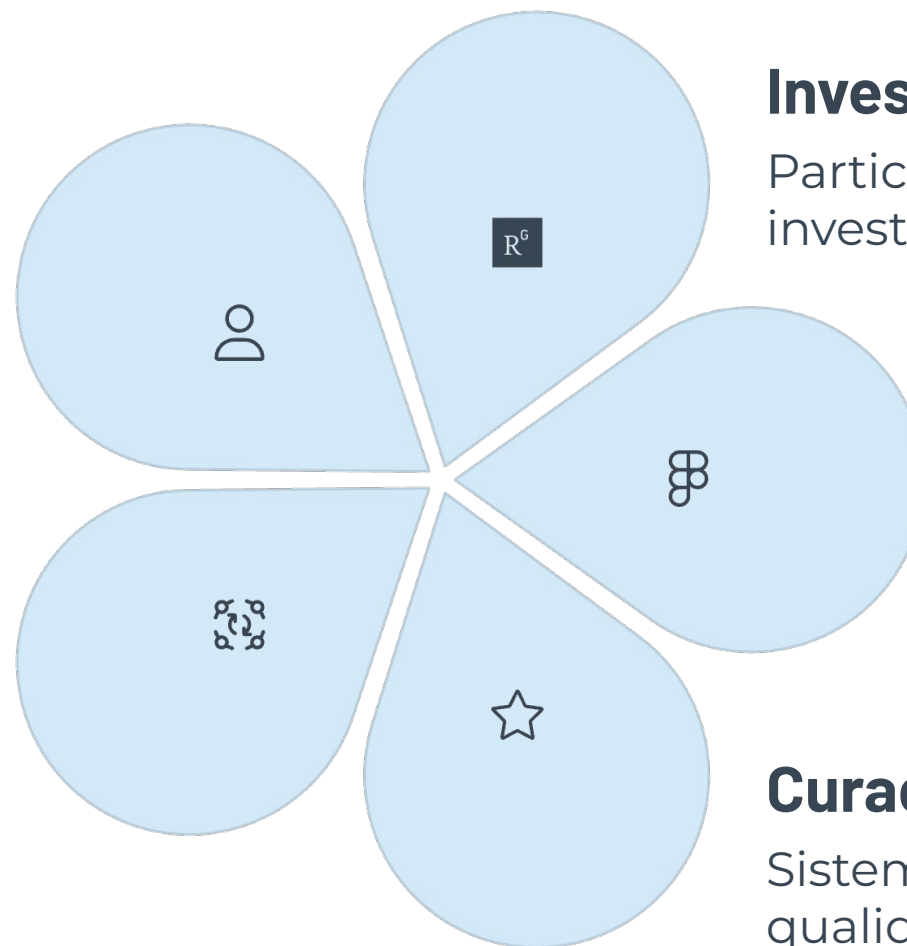
Ferramentas de edição acessíveis que permitem contribuições da comunidade com supervisão especializada

<https://lhenguabase.wikibase.cloud>

Dimensão colaborativa

Falantes nativos
Contribuições valiosas de falantes que conhecem a língua

Comunidade ativa
Rede de colaboradores comprometidos com a língua mirandesa



Investigadores

Participação ativa de linguistas e investigadores especializados

Interface acessível

Ferramentas técnicas acessíveis a utilizadores sem formação informática

Curadoria rigorosa

Sistema de revisão que mantém a qualidade científica dos dados

Funcionalidades de pesquisa avançada

Entradas completas

Pesquisa de entradas com os respectivos sentidos e glosas explicativas, permitindo uma compreensão completa do significado

Exemplos de uso contextualizado

Consulta de entradas acompanhadas de exemplos práticos de uso, facilitando a aprendizagem e aplicação correta

Classes ontológicas

Exploração de classes ontológicas e suas instâncias, permitindo análises linguísticas mais sofisticadas

Alinhamentos semânticos

Alinhamentos com a Wikidata e outras bases de dados semânticas com identificadores universais

lexical_entry	wikidata	posLabel	lemma	sense	gloss_mwl	gloss_pt
Q <https://lhenguabase.wikibase.cloud/entity/L190>	Q wd:L1216038	substantivo	nieto	<https://lhenguabase.wikibase.cloud/entity/L190-S1>		neto
Q <https://lhenguabase.wikibase.cloud/entity/L352>	Q wd:L1126204	substantivo	lhobo	<https://lhenguabase.wikibase.cloud/entity/L352-S1>		lobo
Q <https://lhenguabase.wikibase.cloud/entity/L2>	Q wd:L1408606	substantivo	xaboneteira	<https://lhenguabase.wikibase.cloud/entity/L2-S1>		saboneteira, objeto para colocar sabonetes
Q <https://lhenguabase.wikibase.cloud/entity/L3>	Q wd:L1408607	adjective	passibo	<https://lhenguabase.wikibase.cloud/entity/L3-S1>		passivo
Q <https://lhenguabase.wikibase.cloud/entity/L4>	Q wd:L1408611	substantivo	passibo	<https://lhenguabase.wikibase.cloud/entity/L4-S1>		passivo
Q <https://lhenguabase.wikibase.cloud/entity/L5>	Q wd:L1408613	substantivo	sábado	<https://lhenguabase.wikibase.cloud/entity/L5-S1>		sábado
Q <https://lhenguabase.wikibase.cloud/entity/L37>	Q wd:L1216004	adjective	burmeilho	<https://lhenguabase.wikibase.cloud/entity/L37-S1>		vermelho
Q <https://lhenguabase.wikibase.cloud/entity/L212>	Q wd:L1216000	adjective	branco	<https://lhenguabase.wikibase.cloud/entity/L212-S1>		branco
Q <https://lhenguabase.wikibase.cloud/entity/L265>	Q wd:L1216007	adjective	berde	<https://lhenguabase.wikibase.cloud/entity/L265-S1>		verde
Q <https://lhenguabase.wikibase.cloud/entity/L293>	Q wd:L1216021	substantivo	tiu	<https://lhenguabase.wikibase.cloud/entity/L293-S1>		senhor
Q <https://lhenguabase.wikibase.cloud/entity/L302>	Q wd:L1216041	substantivo	padrino	<https://lhenguabase.wikibase.cloud/entity/L302-S1>		padrinho

lexical_entry	wikidata	posLabel	lemma	usage_example
Q <https://lhenguabase.wikibase.cloud/entity/L584>	Q wd:L1408692	substantivo	çcremaçon	defende ua çcremaçon positiba cun políticas para sigurar las pessonas.
Q <https://lhenguabase.wikibase.cloud/entity/L663>	Q wd:L1408704	adjective	zgraciado	Rompe l zgraciado die 8 de Maio de 1762.
Q <https://lhenguabase.wikibase.cloud/entity/L673>	Q wd:L1408706	substantivo	murclegano	l assi, un murclegano serie, nó un "rato ciego", mas un "rato ceguito"!
Q <https://lhenguabase.wikibase.cloud/entity/L770>	Q wd:L1408724	adjective	cúrtio	Eiduarte fui un rei de Pertual,cun un reinado cúrtio mais memorable

Exemplos de queries SPARQL



As queries SPARQL permitem extrações precisas de dados para investigação, ensino e desenvolvimento de aplicações linguísticas.

Perspetivas futuras e sustentabilidade



Metas imediatas

Expansão contínua dos dados lexicais e melhoria da cobertura vocabular



Desenvolvimento multilingue

Adição de traduções e equivalentes em múltiplas línguas ibero-românicas



Interconexões

Estabelecimento de ligações com outros projetos lexicográficos



Digitalização de recursos

Conversão de recursos lexicais existentes (PDFs, fichas) para formatos estruturados e pesquisáveis



Parcerias estratégicas

Exploração de colaborações com projetos de humanidades digitais

Conclusão

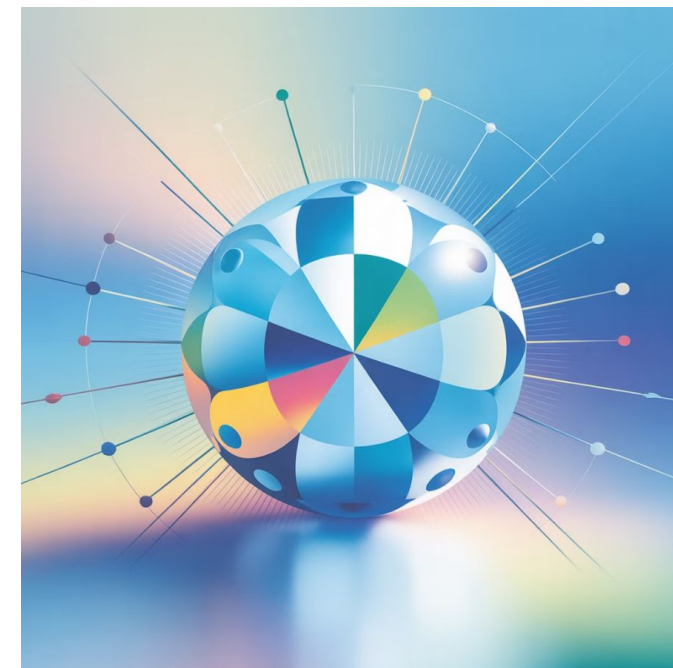
A Lhenguabase representa um passo essencial para a valorização e sustentabilidade do mirandês.

Primeira plataforma lexicográfica digital especificamente dedicada ao mirandês, utilizando tecnologias de ponta da web semântica.

Contribuição significativa para a documentação e preservação digital de uma língua minorizada.

Convite à participação

A Lhenguabase só cumprirá plenamente a sua missão com a participação ativa da comunidade académica, dos falantes nativos e dos entusiastas da língua mirandesa. Convidamos todos a explorar a plataforma, contribuir com dados e colaborar neste projeto de preservação linguística.



Ana Salgado

acsalgado@letras.up.pt

Luís Trigo

ltrigo@letras.up.pt

* Um reconhecimento especial vai para a comunidade de falantes de mirandês, cuja colaboração é essencial para o êxito desta iniciativa.

Visite a Lhenguabase em: <https://lhenguabase.wikibase.cloud>

Este projeto representa o compromisso das instituições portuguesas com a comunidade mirandesa e o seu património cultural imaterial.