

Making the most of a 100-year-old dictionary

Alberto Simões, Álvaro Iriarte



**Dicionário
Aberto**

<http://dicionario-aberto.net>

- * Tarefa 1: Criar utilizador no Dicionário-Aberto.
1. Aceder ao sítio do dicionário;
 2. Usar ligação “entrar/registar”;
 3. Preencher formulário “registar”;
 4. Visitar caixa do correio, e esperar pelo e-mail de registo;
 5. Confirmar registo no sítio do dicionário;
 6. Entrar no sítio do dicionário com o utilizador criado.

Parte I

História do Dicionário Aberto



**Dicionário
Aberto**

Motivação

- * Não existência de um “dicionário livre de língua”, da língua Portuguesa para PLN:
 - * Disponível para acesso local;
 - * Disponível em formato electrónico;
 - * Disponível em formato legível por computador;

Problemas

- * Inexistência de mão de obra para a criação de um dicionário de raiz;
- * Inexistência de editoras interessadas em disponibilizar versões livres dos seus dicionários;
- * Existência de dicionários com mais de 90 anos, no domínio público... em formato papel!

Desafio

- * Desafiar equipa portuguesa dos *Distributed Proofreaders* do Projecto Gutenberg para transcrever um destes dicionários;
- * Dicionário escolhido: Novo Diccionário da Língua Portuguesa de Cândido de Figueiredo (1913)
- * Razão: digitalizado (imagens) pela Biblioteca Nacional.

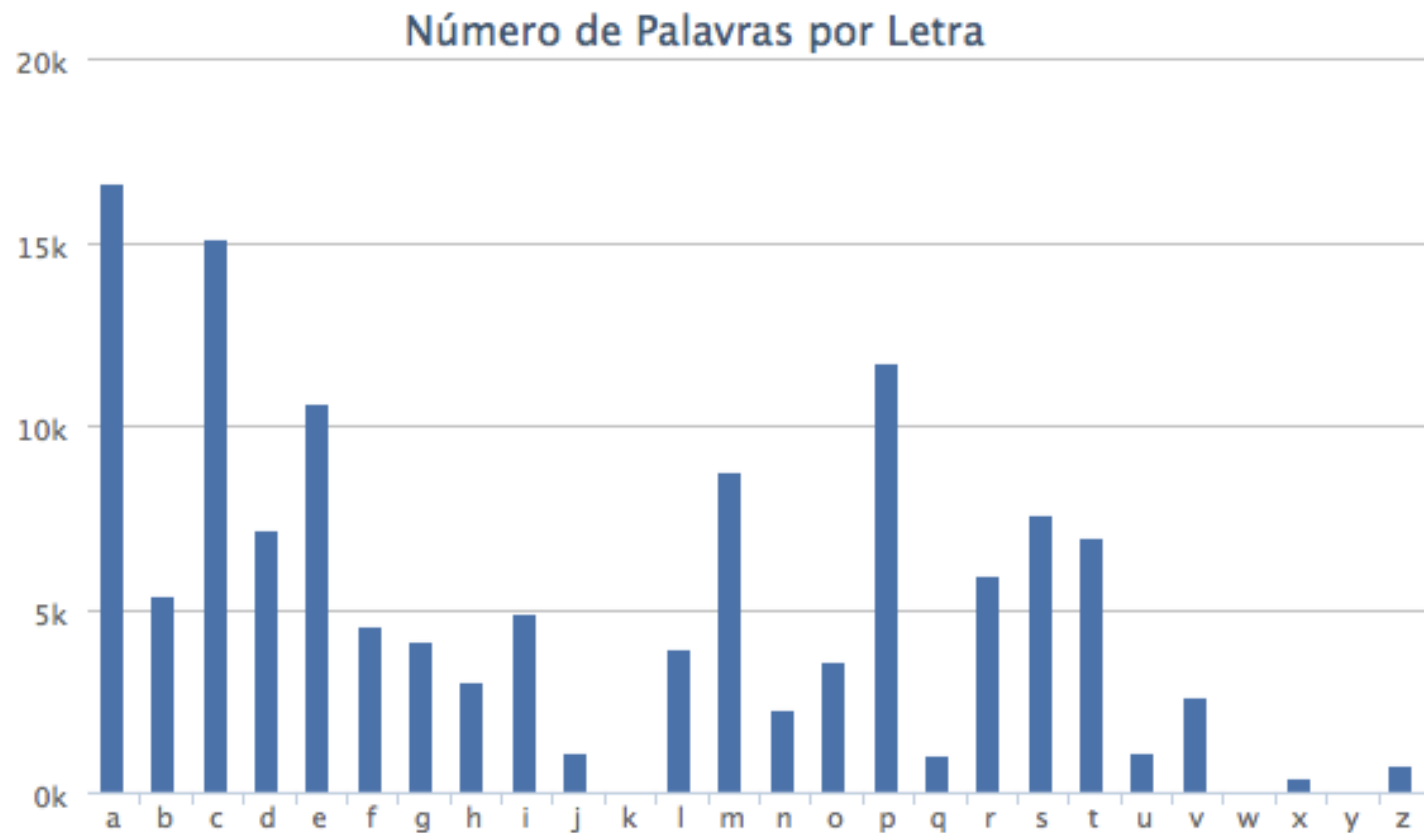
Transcrição

- * Processo demorado, por etapas, por voluntários;
 - * Obrigado, Manuela!
- * Revisto por mais voluntários;
 - * Obrigado, Rita!
- * Processo com mais de quatro anos!

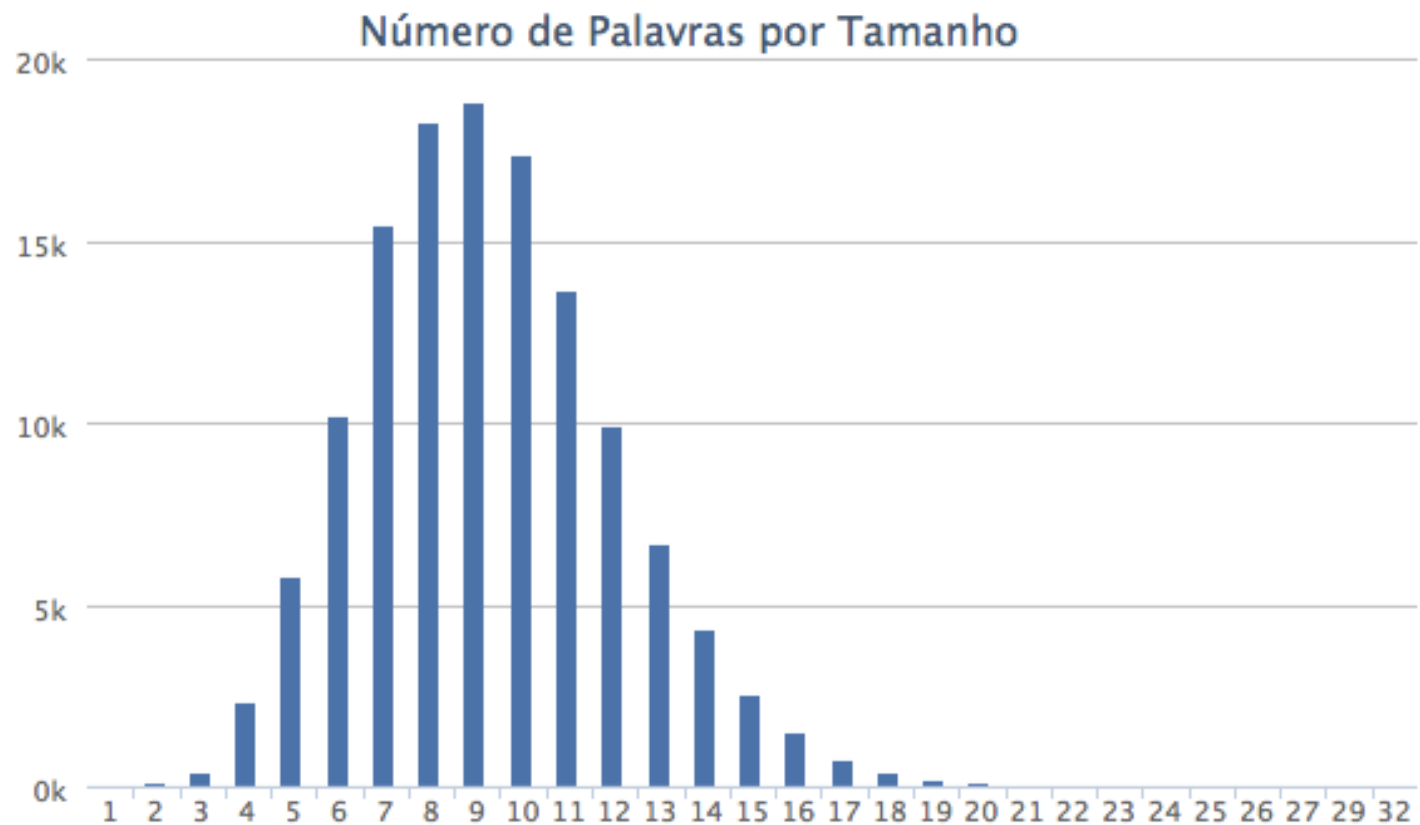
Incorporação

- * Transcrição num dialecto tipo Wiki;
- * Adição diária de 100/200 palavras, desde início de 2007;
 - * Sítio on-line a partir de Junho de 2007!
 - * Palavras adicionadas disponíveis no dia seguinte!
- * Incorporação terminada a 2 de Março de 2010!

Estatísticas



Estatísticas



Estruturação

- * Formato usado simples para ser aprendido e usado pelos voluntários;
- * Formato demasiado vago e ambíguo;
- * Transformação num formato XML rico;
- * Subconjunto do TEI para dicionários;
- * De forma completamente automática!!

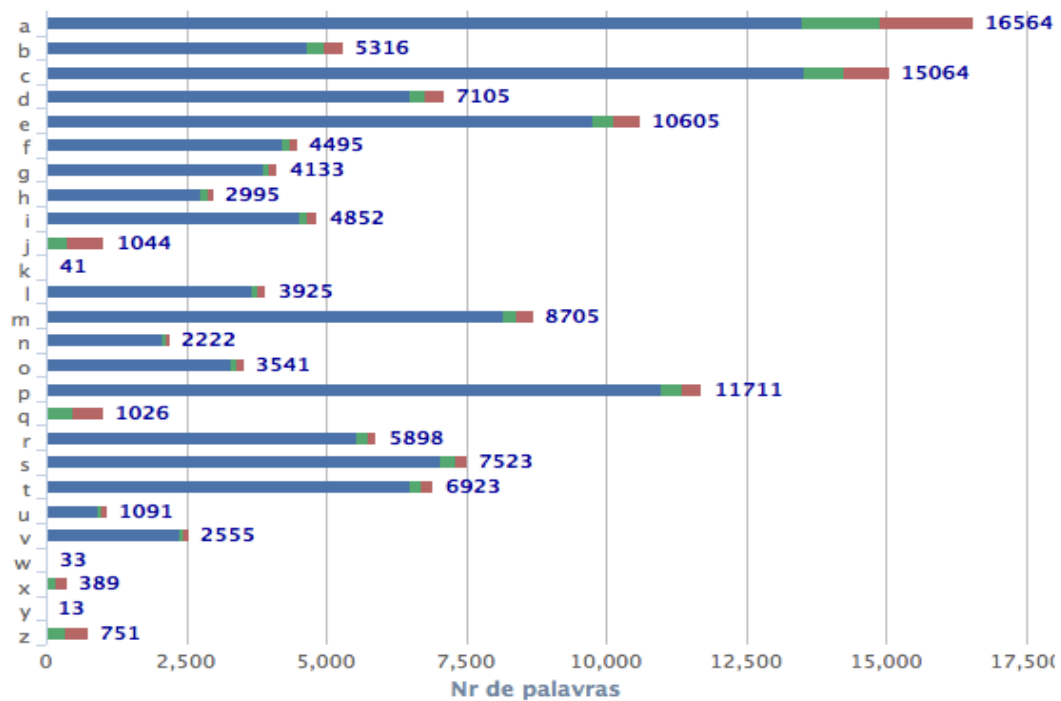
Modernização

- * Dicionário com grafia de 1913!
- * Modernização da grafia de forma semi-automática:
 1. Construção de regras de conversão;
 2. Validação manual!
 3. Voluntários responsáveis, precisam-se.!

Modernização

Estado do processo de Modernização

apagadas aprovadas por moderar



Outras Funcionalidades

- * Disponível em PDF e Base de Dados SQL;
- * Disponível como motor de pesquisa para Firefox;
- * Disponível em ePub;
- * Disponível em StarDict;
- * Disponível em API RESTless:
 - * Aplicação para iOS disponível;
 - * Aplicação para Android e Win8 em desenvolvimento.

Parte II

Interface Geral



**Dicionário
Aberto**

Pesquisa Simples

Cavalgadura

f.

Besta cavalgar, muar, ou asinina, que se pode cavalgar.

Fig.

Pessoa grosseira e estúpida.

(De cavalgar)

- **sinónimo de:** encavalgadura, montaria

Revisão 2



Sugestões “léxicas”

Montaria²

f.



Provisão de cavalos para o exército; remonta.

Bras

O mesmo que cavalgada.

(De montar)

- **sinónimo de:** cavalgada

Revisão 2  



 Será que queria dizer: » contaria » monaria » montearia » monteria » pontaria » tontaria

Folhear o Dicionário

A ▲ ▲

B competente

C competentemente

D competição

E competidor

F competir

G compilação

H compilador

I compilar

J compilatório

K compita

L compitais

M cômputo


N complacência

← complacência complacentemente →

Complacente

adj.
Que tem complacência.
Em que há complacência.
(Lat. *complacens*)

- **antónimo de:** *incomplacente*
- **sinónimo de:** *comprazente*

Revisão 2 ⓘ 

Palavra aleatória

Enologia

f.

Tratado acerca dos vinhos e da sua preparação.

(Do gr. *oinos* + *logos*)

- **relacionado com:** [enológico](#)

Revisão 2 ⇅



🔍 Será que queria dizer: » [cenologia](#) » [etnologia](#) » [etologia](#) » [neologia](#)

Palavras “Favoritas”

Entradas Favoritas

- alforreca¹
- x. p. t. o.¹

Histórico Ortográfico

Alforreca

f.

Molusco, do feitio de umbrella e de tecidos semi-transparentes.

(Do ár. *al-horreque*)

- **co-hipónimo de:** [ecopleura](#)
- **hiperónimo de:** [pelágia](#), [vinagreira](#)
- **hipónimo de:** [acalefos](#), [acalephos](#)
- **sinónimo de:** [água-viva](#)
- **termo mais abrangente que:** [pelágia](#), [vinagreira](#)
- **termo mais específico que:** [acalefos](#), [acalephos](#)

Revisão 1 ▾



🔍 Será que queria dizer: » [alfarreca](#)

Parte III

Pesquisa Avançada



**Dicionário
Aberto**

Pesquisa Avançada

- * Pesquisa por afixos;
- * Ocorrências nas definições;
- * Relações léxico-conceituais entre termos introduzidos;
- * Resultados ordenados por relevância.

Pesquisa por “prefixo”

in...¹ – *pref.* ; (designativo de *privação, negação, lugar*)(Lat. *in*)

in-fólio¹ – *adj.* ; Diz-se de um livro ou de um formato, em que cada folha de impressão é ap...

in-quarto¹ – *m.* ; Volume, cujas folhas foram impressas a oito páginas cada uma.Formato de li...

inabalável¹ – *adj.* ; Que não pode ser abalado. *Fíg.*; Fixo; constante: *fé inabalável*.Inquebra...

inabalavelmente¹ – *adv.* ; De modo inabalável.

inabdicável¹ – *adj.* ; Que se não pode abdicar.(De *in...* + *abdicável*)

inábil¹ – *adj.* ; Não hábil; que não tem aptidão ou competência; incapaz.(Lat. *inhabilis*)

inabilidade¹ – *f.* ; Qualidade de inábil; falta de habilidade.(De *in...* + *habilidade*)

inabilitação¹ – *f.* ; Acto ou efeito de inabilitar.

inabilitar¹ – *v. t.* ; Tornar inábil.Impedir.Tirar a faculdade ou certos meios a.(De *in...* + *h...*)

inabilmente¹ – *adv.* ; De modo inábil.

inabitado¹ – *adj.* ; Não habitado.(Lat. *inhabitatus*)

inabitável¹ – *adj.* ; Que não pode ser habitado.(Lat. *inhabitabilis*)

inabitual¹ – *adj.* ; Que não é habitual.(De *in...* + *habitual*)

inabordável¹ – *adj.* ; Que não é abordable.(De *in...* + *abordable*)

inacabável¹ – *adj.* ; Que se não pode terminar; interminável, infindo.(De *in...* + *acabar*)

inacção¹ – *f.* ; Falta de acção.Inércia; indecisão; froxidão de carácter.(De *in...* + *acção*)

inaceitável¹ – *adj.* ; Que não é aceitável.Que se não pode aceitar ou admitir; inadmissível; in...

inacessibilidade¹ – *f.* ; Qualidade do que é inacessível.

inacessível¹ – *adj.* ; Que não é acessível; a que se não pode chegar: *uma felicidade inacessível...*

Pesquisa por “sufixo”

acalefologia¹ — *f.* ; Parte da Zoologia, que trata dos acalephos.

acalephologia¹ — *f.* ; Parte da Zoologia, que trata dos acalephos.

acirologia¹ — *f.* ; Improriedade de expressão.(Gr. *akurologia*)

acologia¹ — *f. Med.*; O mesmo que *acognosia*.(Do gr. *akos* + *logos*)

acribologia¹ — *f.* ; Rigor e precisão no estilo.(Gr. *akribologia*)

acrologia¹ — *f.* ; Investigação do absoluto, dos primeiros princípios.(Gr. *akron* + *logos*)

actinologia¹ — *f.* ; Descrição dos animais radiados.(Do gr. *aktin* + *logos*)

acyrologia¹ — *f.* ; Improriedade de expressão.(Gr. *akurologia*)

adenologia¹ — *f.* ; Parte da Anatomia, que trata de glândulas.(Do gr. *aden* + *logos*)

aerologia¹ — *f.* ; Tratado do ar.(De *aerólogo*)

agmatologia¹ — *f. Med.*; Tratado das fracturas.(Do gr. *agma* + *logos*)

agrologia¹ — *f.* ; Ciência, que tem por objecto o conhecimento dos terrenos, nas suas relação...

agrostologia¹ — *f.* ; O mesmo que *agrostographia*.

alethologia¹ — *f.* ; Tratado ou discurso acerca da verdade.(Do gr. *aletheia* + *logos*)

aletologia¹ — *f.* ; Tratado ou discurso acerca da verdade.(Do gr. *aletheia* + *logos*)

algologia¹ — *f.* ; Tratado ou estudo das algas.(Cp. *algólogo*)

allotriologia¹ — *f.* ; Aplicação de doutrinas, estranhas ao assunto ocorrente.(Do gr. *allotrios* + ...

alogia¹ — *f.* ; Absurdo, disparate.(Do gr. *a* priv. e *logos*)

alotriologia¹ — *f.* ; Aplicação de doutrinas, estranhas ao assunto ocorrente.(Do gr. *allotrios* + ...

ampelologia¹ — *f.* ; Conjunto dos princípios ou teorias respeitantes à cultura e tratamento da ...

Pesquisa de “infixos”

cairiris¹ — *m. pl.* ; Numerosa tribo de Índios do Brasil, que viviam na cordilheira Borborem...

cariris¹ — *m. pl.* ; Selvagens, que habitaram na região de Paraíba, no Brasil.

ebridente¹ — *adj.* ; Que se ri, embriagado. Cf. Castilho, *Fastos*, I, 47.(De *ébrio* + *ridente*)

eribá¹ — *m.* ; Árvore leguminosa do Brasil.

iribá¹ — *m. Bras*; Árvore do sertão.

maririçó¹ — *m.* ; Planta irídea do Brasil, purgativa.

muirapirica¹ — *f.* ; Árvore brasileira, empregada em construções.

piririca¹ — *adj. Bras*; Áspero como a lixa.(Do tupi *piriri*)

piriricar¹ — *v. i. Bras*; Produzir um ligeiro estremecimento na água.(Cp. *pererecar*)

tariqui¹ — *m. Bras*; Planta medicinal.

tatapirica¹ — *f.* ; Planta terebintácea do Brasil.

tiririca¹ — *f.* ; Arbusto cyperáceo do Brasil.Nome de outras plantas.

...

Pesquisa Reversa

base de dados conceptuais / dicionário onomasiológico / dicionário de produção ou codificador

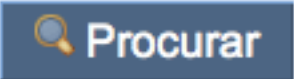
(mais do que um thesaurus, que ordena palavras por sinónimos, hiperónimos ou outras relações léxico-conceptuais)

Pesquisa Reversa

dicionário codificador:

endurecer + metal \Rightarrow temperar

de a a zythógala (128 521 entradas)

endurecer metal 

Prefixo Infixo Sufixo Pesquisa Reversa Pesquisa Ontológica

temperar¹ — *v. t.* ; Misturar proporcionalmente. Adubar. Moderar o gosto ou sabor de. Preparar....

...

Pesquisa Ontológica

* Criação de uma ontologia de forma automática usando padrões (Hearst 1992)

- **SYN**: o mesmo (ou melhor)? que ...
- **ANT**: que não é ...
- **HIPO**: espécie de ...
- **HIPER**: que tem por tipo ...
- **MERO**: cada uma das partes que formam ...
- **HOLO**: composto por ...

Pesquisa Ontológica

- * Completação da Ontologia por regras matemáticas

- $a \text{ SYN } b \Rightarrow b \text{ SYN } a$
- $a \text{ HIPO } b \wedge c \text{ HIPO } b \Rightarrow a \text{ COHIPO } c$
- $a \text{ HIPO } b \wedge b \text{ HIPO } c \Rightarrow a \text{ HIPO } c$

- * Permite que se use informação incompleta:

- * **Verbete A** indica sinonímia com B.
- * **Verbete B** não faz qualquer referência a A...

Estrutura Ontológica

Pirilampo

m.

Género de insectos coleópteros pentâmeros, que emitem uma luz fosforescente; vagalume.

(Gr. *purilampis*)

- **hiperónimo de:** [deríbia](#)
- **sinónimo de:** [avincu](#), [abrecu](#), [arancu](#), [caga-lume](#), [luzecuco](#), [luze-luze](#), [luzincu](#), [noctiluz](#), [oirincu](#), [ourincu](#), [pirífora](#), [pyríphora](#), [vagalume](#)
- **termo mais abrangente que:** [deríbia](#)


Revisão 2 ↕



🔍 Será que queria dizer: » [pyrilampo](#)

Pesquisa Ontológica

veado animal

 Procurar

Prefixo Infixo Sufixo Pesquisa Reversa Pesquisa Ontológica

cervo¹ — *m.* ; O mesmo que *veado*¹.(Lat. *cervus*)

alimal¹ — *m. Pop.*; O mesmo que *animal*¹.

mandiocaba¹ — *f.* ; Espécie de mandioca.

san-pedro-branco¹ — *m. Bras*; Espécie de mandioca, de talo branco e grandes raízes.

san-pedro-vermelho¹ — *m. Bras*; Espécie de mandioca, de talo cor de rosa.

gamo¹ — *m.* ; Espécie de veado, que tem achatada a parte superior dos galhos e comprida ...

ôlho-roxo¹ — *m. Bras*; Espécie de mandioca, de raiz comprida.

çuçupara¹ — *f. Bras*; Espécie de veado.

zorlito¹ — *m.* ; Espécie de veado, (*cervus capreolus*), que é o corço vulgar.

mata-negro¹ — *f.* ; Espécie de mandioca.

mónica¹ — *f. Açor*; O mesmo que *nêspora*. *Bras*; Espécie de mandioca.

perdido¹ — *adj.* ; Disperso.Naufragado.Extraviado.Que gosta muito: *é perdido por morangos.M...*

san-pedro-molle¹ — *m. Bras*; Espécie de mandioca, de talo muito sucoso.

nuri¹ — *m. Bras*; Espécie de mandioca

Descarregar

TEI

TXT

Parte IV

Exercícios



**Dicionário
Aberto**

Procura por Afixos

* Palavras Cruzadas:

_____ F E U (Vértebra; 10 letras)

E P I _____ (Vértebra; 10 letras)

Pesquisa por Afixos

- * Dicionário de Rimas (gráficas)

Que palavras rimam com “camafeu”?

Pesquisa por Afixos

* Estudo de Morfologia (produtividade de afixos):

Todos os adjectivos em **-vel** (como *amável*) podem formar advérbios em **-velmente** (*amavelmente*)?

Pesquisa por Afixos

* Estudo de Morfologia (produtividade de afixos):

Um professor de Língua Portuguesa disse aos alunos que o sufixo "-ería" em Espanhol, é sempre "-aria" em Português.

cafetería_{ES} = cafeteria_{PT}

É verdade?

Pesquisa Reversa

- * Dicionários Ideológicos/Onomasiológicos:
 - * O que acontece à “água” com o “frio”?
 - * Quem é o “médico” dos “olhos”?
- * Que palavras derivam do prefixo grego “orthos”?

Pesquisa Ontológica

- * Procurar “sargaço”.
- * O que difere da pesquisa reversa?

Pesquisa Ontológica

- * Procurar “pirilampo” na pesquisa reversa, e na pesquisa ontológica.
- * Quais são as diferenças?

Pesquisa Ontológica

- * Que variedades de oliveiras (ou “espécie de oliveiras”) há?
 - * Uso da pesquisa reversa
 - * Uso da pesquisa ontológica

Making the most of a 100-year-old dictionary

Alberto Simões, Álvaro Iriarte



**Dicionário
Aberto**