

A lusofonia na Wikipédia em 150 tópicos

Cláudia Freitas

Linguatca/FCCN & PUC-Rio

maclaudia.freitas@gmail.com

Resumo

Este artigo descreve os tópicos usados no Páxico, a primeira avaliação conjunta em recolha de informação centrada em tópicos relacionados com a lusofonia, usando o material da Wikipédia em português. Depois de uma explicação sobre como os tópicos foram escolhidos e de questões associadas à sua escolha e à sua categorização posterior, os tópicos são apresentados por categoria. Comentamos também a forma de apresentação de certas informações na Wikipédia e, por fim, fazemos uma categorização geográfica e temporal dos tópicos.

Palavras chave

Páxico, Wikipédia

1 Apresentação

Este artigo descreve os tópicos usados no Páxico, a primeira avaliação conjunta em recolha de informação centrada em tópicos relacionados com a lusofonia, usando o material da Wikipédia em português.

Depois de uma explicação de como os tópicos foram escolhidos¹, e de várias questões associadas à sua escolha e categorização posterior, comentamos algumas opções, decorrentes da forma pela qual certas informações estão presentes na Wikipédia e das características da própria tarefa. Por fim, fazemos uma caracterização dos tópicos em termos de categorização por tema, por distribuição geográfica e por tempo.

2 Processo de elaboração e escolha dos tópicos

Já estamos habituados a usar buscadores como o Google para encontrar a informação que desejamos, embora nem sempre a tarefa seja fácil ou bem-sucedida. E em que contextos visitamos a Wikipédia? Para qual tipo de busca?

¹O grupo responsável pela elaboração dos tópicos foi constituído por mim, Paulo Rocha e Diana Santos.

De minha experiência pessoal, não vou à Wikipédia como quem vai ao Google, isto é, procurar coisas. Vejo a Wikipédia como um lugar para “aprofundar superficialmente” pontos que me interessam. De alguma maneira, já sei o que procuro, já sei por onde começar.

Por isso, a tarefa de elaborar perguntas / tópicos que envolvessem busca não-trivial na Wikipédia acabou sendo mais difícil do que o esperado. (Aliás, do que eu esperava). Que tipo de perguntas elaborar cujo processo de busca fosse “não-trivial”?

A ideia era formular questões suficientemente interessantes – para falantes lusófonos – e que, por outro lado, tivessem respostas pouco óbvias ou cuja busca fosse trabalhosa (isto é, aquelas que, preferencialmente, tivessem respostas “espalhadas” por diferentes páginas.) Assim, diferentemente do relatado em (Simões, Rocha e Fonseca, 2009) sobre os diferentes tipos de pergunta comuns em WebPapers (competição em que os participantes buscam, na Web, respostas a perguntas), evitamos perguntas com respostas únicas, bem como perguntas com muitas indireções, do tipo “Qual a capital da quarta província mais populosa do Canadá”.

Uma pergunta/tópico como [Músicos associados ao surgimento da Bossa-nova], por exemplo, seria uma pergunta relativamente fácil, pois rapidamente respondida com uma visita à página “Bossa Nova” da Wikipédia (ainda que, do ponto de vista dos sistemas, encontrar a resposta correta esteja longe de ser uma tarefa simples).

A estratégia esteve, portanto, em privilegiar perguntas consideradas difíceis/trabalhosas do ponto de vista dos humanos. Mesmo que essa não tenha sido a melhor opção, é defensável quando se pensa em uma tarefa o menos artificial possível.

Ainda na perspectiva de naturalizar a tarefa, outro aspecto relevante durante a elaboração das perguntas/tópicos foi conciliar perguntas cuja resposta não fosse facilmente obtida por uma busca no Google, o que tornaria artificial a necessidade de procura na Wikipédia.

Assim, embora subjacente a essa estratégia

esteja a mistura de duas tarefas - resposta automática a perguntas/recuperação de informação com a tarefa mais simples de recuperação de documentos - o que eu esperava era garantir que a busca de respostas - pelos participantes humanos ou sistemas - seria em grande medida dependente do conteúdo da Wikipédia.

Tentei, assim, criar perguntas que evitassem (ou minimizassem) uma busca geral na Internet, para posterior mapeamento, na Wikipédia, da página referente à resposta.

Naturalmente essa estratégia pode ser vista como complementar: a busca geral na Internet fornece documentos com a indicação da resposta, e o conteúdo da Wikipédia confirma/aprofunda os resultados obtidos. Mas, para os objetivos da tarefa proposta no Páxico, tentei criar perguntas cuja resposta não fosse tão óbvia para o Google.

Vale notar, entretanto, que nem todas as perguntas foram guiadas por essas estratégias - mas isso contou para cerca de 60% delas.

Além de tentar elaborar tópicos interessantes e com resposta trabalhosa, evitamos perguntas com uma única resposta, justamente para motivar a procura em diferentes páginas.² E, a fim de evitar respostas abrangentes/amplas demais, ou com informação subjetiva, fugimos de tópicos/perguntas que envolvessem julgamentos de valor, do tipo “os melhores...”.

3 Dificuldades

Assumindo que o Páxico só teria perguntas que pudessem ser respondidas por meio da consulta à Wikipédia, tentamos já fornecer respostas à medida que elaboramos as perguntas como garantia de que estávamos diante de uma pergunta válida. Nesse processo, fomos surpreendidos por páginas cuja ligação não apontava para o termo referido - ou por erro, ou por inexistência da página. Comento a seguir dois desses casos:

tópico 52 [Gêneros musicais que misturam samba e ritmos norte-americanos]. Em uma série

²De acordo com a Diana,

à medida que o tempo passava, os critérios de escolha e distribuição foram sendo mais pelo interesse pelo aumento do leque de assuntos e menos pela existência de respostas na wikipédia. Associado ao fato infeliz de todo o trabalho ter sido feito sob grande pressão de prazos, isso explica que em alguns casos os donos dos tópicos (eu, por exemplo) não foram capazes ou não tiveram tempo de achar respostas antes de o conjunto estar completo e divulgado aos participantes.

de páginas, existe referência à Pilantragem (que seria uma resposta adequada):

... e que foi chamado de pilantragem (uma mistura de samba e soul), movimento também idealizado e capitaneado por Carlos Imperial.

Wikipédia

A ligação de pilantragem, no entanto, leva à página *Turma da Pilantragem* (Wikipédia), que não é uma resposta adequada à pergunta:

A Turma da Pilantragem foi o nome de um grupo musical surgido no movimento cultural brasileiro denominado Pilantragem, em fins da década de 1960.

Turma da Pilantragem
Wikipédia

Ou seja, pilantragem, como gênero musical, não tem uma página específica, embora, na página do referido grupo musical existam dados históricos sobre o movimento.

tópico 62 [Praias de Portugal boas para a prática de surf]. Na página *Turismo em Portugal* (Wikipédia), temos que

Portugal é também um país onde se pratica, além de muitos outros desportos, surf. Entre os melhores spots estão o Guincho, Peniche, Ericeira, Carcavelos, São Pedro e São João do Estoril, Costa da Caparica e São Torpes.

Turismo em Portugal
Wikipédia

No entanto, a página/documento *São Torpes* (Wikipédia) não existe, e a ligação de São Torpes é para a página do *Parque Natural do Sudoeste Alentejano e Costa Vicentina* (Wikipédia). Na página do referido parque, por sua vez, não há qualquer menção a São Torpes. Assim, embora seja uma resposta correta, “São Torpes” não pode ser adicionada como página-resposta do Páxico (e nem aceitamos a página *Parque Natural do Sudoeste Alentejano e Costa Vicentina* (Wikipédia), pois não consideramos que seja uma resposta adequada para o tópico.)

Ainda no mesmo trecho, a ligação de *Guincho* (Wikipédia) nos direciona para a página *Forte do Guincho* (Wikipédia), que também não consideramos correta - embora exista a página *Praia do Guincho* (Wikipédia), que responde ao tópico. Para a discussão referente à adequação

de uma resposta, bem como de justificativas associadas às respostas, veja-se (Freitas et al., 2012).

Ou seja, tivemos dois tipos de problemas:

1. a ligação aponta para uma página incorreta, e a página correta ainda não existe na Wikipédia
2. a ligação aponta para uma página incorreta, e a página correta existe na Wikipédia

3.1 Redação dos tópicos

Como já mencionado, evitamos tópicos sem resposta – ou melhor, sem página de resposta. Tentamos evitar, também, tópicos com um número excessivo de respostas. Tópicos como [Orquestras filarmônicas com mais de cinquenta anos de vida],[Violoncelistas de língua portuguesa] e [Excetuando-se o português, que outras línguas faladas são faladas em países lusófonos], por exemplo, passaram respetivamente a **tópico 3** [Orquestras filarmônicas com mais de cinquenta anos de vida em países lusófonos] **tópico 4** [Mulheres violoncelistas de língua portuguesa] e a **tópico 25** [Que línguas bantas ou bantus são faladas em países lusófonos?]

Outro cuidado esteve em procurar um “português universal”, isto é, garantir que diferenças entre as variantes não levariam a dificuldades na compreensão dos tópicos. Em alguns casos, tratava-se apenas de uma variação no uso (futebolistas / jogadores de futebol), não sendo difícil perceber do que se tratava. Noutros casos, embora fosse possível inferir a intenção da pergunta, as alterações foram necessárias para um rápido entendimento comum. Por exemplo, [Ilhas e ilhéus desabitados de Cabo Verde] passou a “Ilhas e ilhotas”, uma vez que, no português falado no Brasil, *ilhéu* é principalmente aquele que mora em uma ilha, o que tornava o tópico sem sentido.

4 Conteúdo e distribuição dos tópicos

Considerando o espectro assumidamente geral de “cultura lusófona”, e também que um tópico pode ser sobre mais de um tema ([Filmes sobre as relações entre Portugal e suas colônias], por exemplo, pode ser tanto sobre Cinema quanto História), os 150 tópicos distribuem-se da maneira indicada na tabela 1.

Esta tabela apresenta a negrito a distribuição dos tópicos pelo que chamamos grandes temas – categorias mais abrangentes. A maior parte dos 150 tópicos pertence ao super tema Letras,

Grande tema -tema	Tópicos	
	#	%
Letras	69	46,00
- história	50	33,33
- literatura	15	10,00
- linguística	6	4,00
- jornalismo	3	2,00
- filosofia	2	1,33
Artes	36	24,00
- música	19	12,67
- cinema	10	6,67
- televisão	4	2,67
- artes plásticas	2	1,33
- artes	2	1,33
Geografia	34	22,67
- geografia	26	17,33
- arquitetura/urbanismo	7	4,67
- demografia	4	2,67
- geologia	2	1,33
Cultura	27	18,00
- antropologia/folclore	12	8,00
- religião	7	4,67
- culinária	5	3,33
- cultura	3	2,00
- ensino	2	1,33
Política	19	12,67
Desporto/Esportes	18	12,00
Ciência	14	9,33
- saúde	4	2,67
- zoologia	3	2,00
- ciência	2	1,33
- botânica	2	1,33
- geologia	2	1,33
- matemática	1	0,67
Economia	6	4,00

Tabela 1: Distribuição dos tópicos por tema

decorrência principalmente da presença do tema História – que corresponde ao tipo de informação mais naturalmente associada ao conhecimento enciclopédico – nessa categoria. A segunda categoria mais frequente é Artes, que engloba tópicos relacionados a música e cinema, dentre outros. Logo em seguida aparecem o super tema Geografia e Cultura. Os super temas menos frequentes foram Política, Desporto/Esportes, Ciência e Economia.

Como se pode perceber, a totalidade de tópicos distribuídos é superior a 150, indicando que alguns tópicos foram classificados em mais de um grande tema. Mais especificamente, a maioria dos tópicos (56%) foi atribuído apenas um tema; a 39% dos tópicos foram atribuídos 2 temas e a 4,6% dos tópicos, 3 temas; não há tópicos com mais de 3 temas (veja-se a tabela 2).

# Temas atribuídos	Grandes temas		Temas	
	Total	%	Total	%
1	84	56,00	75	50,00
2	59	39,33	64	42,67
3	7	4,67	11	7,33

Tabela 2: Atribuição de temas e grandes temas por tópico

A tabela 1 mostra igualmente, abaixo da distribuição por grandes temas, a distribuição dos tópicos por temas, o que permite uma análise mais detalhada do conteúdo do Páigico. A partir desses valores observamos que, dentre os temas, o mais frequente foi História (50 tópicos), seguido de Geografia (26 tópicos), Música e Política (ambos com 19 tópicos) e Desporto/Esportes (18 tópicos). O menos frequente foi Matemática, com apenas um tópico. A tabela 2 indica também que metade dos tópicos é apenas sobre um tema; 42% (64 tópicos) dos tópicos envolve 2 temas e apenas 7% (11) dos tópicos foi classificado como 3 temas. Contrastando a distribuição por tópicos em termos de temas e super temas (consulte-se novamente a tabela 2), notamos que, quando tratamos das categorias mais abrangentes (super temas), há uma ligeira prevalência de tópicos com um tema (56% vs 50%). A atribuição de dois ou mais temas a um tópico foi decorrência não só do conteúdo misto do tópico propriamente, como no exemplo de filmes sobre um dado tema, mas também quando houve divergência, entre os membros da organização, sobre qual tema atribuir – e, nesse caso, todas as possibilidades foram consideradas. Por essa perspectiva, a maior frequência de tópicos com categorização única no âmbito dos super temas sugere a maior facilidade de concordância quando estamos diante de categorias mais amplas.

A tabela 5 lista todos os tópicos criados para o Páigico. Para cada tópico, informamos: identificador (ID), descrição, a sua classificação em grande tema, tema e local (consulte-se a tabela 4 para uma correspondência entre as abreviaturas e os países respetivos).

4.1 Entidades Mencionadas nos tópicos

Dos 150 tópicos, 96 continham entidades mencionadas em sua formulação. A tabela 3 apresenta a distribuição de entidades mencionadas (EM) por tópico. Para a identificação e classificação das EM, seguimos as diretrizes do Segundo HAREM (Mota e Santos, 2008): consideramos apenas as EM iniciadas por letras maiúsculas – exceção para as entidades temporais – e tomamos como base as categorias descritas em Carvalho et al. (2008), uma vez que se mostraram produtivas

para a classificação das entidades nesta tarefa. Ainda seguindo a maneira de classificação do Segundo HAREM, consideramos possível uma mesma EM receber mais de uma classificação. Assim, a EM Brasil no **tópico 34** [Viajantes ou exploradores que escreveram sobre o Brasil do século XVI] pode ser vista tanto como uma ORGANIZAÇÃO, como um LOCAL ou como PESSOA (os brasileiros), e portanto recebeu as 3 classificações.

local	59
tempo	24
organização	21
peessoa	10
acontecimento	8
abstração	5
coisa	3

Tabela 3: EM no Páigico

4.2 Categorização geográfica e temporal

Com relação à localização geográfica, os 150 tópicos do Páigico estão distribuídos como indica a tabela 4:

Lugar (abreviatura)	Tópicos	
	#	%
Brasil (br)	50	33,33
Lusofonia (lus)	44	29,33
Portugal (pt)	16	10,67
Moçambique (mo)	11	7,33
Angola (ao)	8	5,33
Geral (ger)	7	4,67
Cabo Verde (cv)	6	4,00
Macao (mo)	4	2,67
Timor (ti)	3	2,00
Guiné Bissau (gw)	2	1,33
São Tomé e Príncipe (st)	2	1,33

Tabela 4: Lugares mencionados no Páigico

A categoria “lusofonia geral” corresponde a tópicos em que não há especificação de região geográfica; e os tópicos “gerais” não mencionam explicitamente elementos da cultura lusófona, mas integram o Páigico por conterem, em sua resposta, elementos da cultura de países lusófonos.

Com relação à localização temporal, os 150 tópicos estão assim distribuídos: a partir do século XX em diante, existem 43, até o século XIX inclusive, 13, enquanto os outros não envolvem localização temporal explícita, quer por corresponderem a perguntas que não envolvem tempo (1), ou perguntas cujas respostas podem estar distribuídas por diferentes períodos (2):

1. **tópico 123** [Rios de Angola com mais de 500 quilómetros de comprimento]
2. **tópico 103** [Movimentos culturais surgidos no nordeste do Brasil]

4.3 E os tópicos de exemplo?

Rodrigues, Gonçalo Oliveira e Gomes (2012) mencionam que os tópicos do Págico eram muito diferentes dos tópicos de exemplo. Não tínhamos essa impressão, mas fomos aplicar a análise acima a esses 11 tópicos:

Comparativamente aos 11 tópicos apresentados como exemplo, os tópicos do Págico têm uma distribuição bastante parecida em termos de localização geográfica. No critério conteúdo, como era de se esperar, há uma distribuição mais homogênea dos temas por tópico. Os supertemas mais frequentes foram Artes, Ciência, Cultura, Desporto/Esportes, todos com 2 tópicos cada, e Economia, Geografia, Letras e Política, com 1 tópico cada. Quanto à distribuição geográfica, dos 11 tópicos, 5 (45%) estão relacionados ao Brasil; 3 relacionados a Portugal; 2 à Lusofonia geral e 1 é um "tópico geral". Dos 11 tópicos, apenas 6 (cerca de 55%) contêm entidades mencionadas, em uma distribuição parecida à do Págico (64% dos tópicos contêm EM). Nos exemplos, a distribuição das categorias de EM se deu da seguinte maneira (lembrando sempre que uma EM pode receber mais de uma categoria): ORG e LOCAL foram as mais frequentes, com 2 ocorrências cada, e em seguida PESSOA, TEMPO e OUTRO, com 1 ocorrência cada. Por fim, considerando a dimensão temporal, assim como no Págico, os tópicos exemplo tratam, na imensa maioria (7 tópicos), de questões relativas ao século XX.

5 Considerações finais

Neste breve artigo apresentamos as motivações e decisões subjacentes à criação dos tópicos do Págico. Caracterizamos detalhadamente o conjunto de tópicos quanto ao tema, distribuição geográfica e temporal, em um instantâneo que

tenta capturar, em 150 perguntas, um pouco da cultura dos países falantes da língua portuguesa.

Agradecimentos

O trabalho aqui descrito enquadra-se no âmbito da Linguateca, co-financiada desde o seu início pelo Governo Português, pela União Europeia (FEDER e FSE), sob o contrato POSC/339/1.3/C/NAC, pela UMIC e pela FCCN, e em 2011 pela Fundação da Ciência e da Tecnologia (FCT) e pela Fundação para a Computação Científica Nacional (FCCN).

Agradeço ao resto da organização do Págico, sem a qual o mesmo não teria sido possível, e a todos os participantes, cujas respostas ajudaram a iluminar os tópicos e a esclarecer pontos pouco claros.

Agradeço também ao Fernando Perdigão e à Belinda Maia, pela revisão e comentários que ajudaram a tornar o artigo mais esclarecedor.

Referências

- Carvalho, Paula, Hugo Gonçalo Oliveira, Diana Santos, Cláudia Freitas, e Cristina Mota. 2008. Segundo HAREM: Modelo geral, novidades e avaliação. Em Cristina Mota e Diana Santos, editores, *Desafios na avaliação conjunta do reconhecimento de entidades mencionadas*. Linguateca, pp. 11–31, 31 de Dezembro, 2008.
- Freitas, Cláudia, Paulo Rocha, Cristina Mota, Luís Costa, e Diana Santos. 2012. O que é uma resposta? Notas de uns avaliadores estafados. *Linguamática*, 4(1), Abril, 2012. Neste volume.
- Mota, Cristina e Diana Santos, editores. 2008. *Desafios na avaliação conjunta do reconhecimento de entidades mencionadas: O Segundo HAREM*. Linguateca.
- Rodrigues, Ricardo, Hugo Gonçalo Oliveira, e Paulo Gomes. 2012. Uma abordagem ao Págico baseada no processamento e análise de sintagmas dos tópicos. *Linguamática*, 4(1), Abril, 2012. Neste volume.
- Simões, Alberto, Paulo Rocha, e Rúben Fonseca. 2009. Webpaper — más perguntas e boas respostas: a arte de interrogar. Em Paulo Dias, António José Osório, e Altina Ramos, editores, *O digital e o currículo*. Centro de Competência da Universidade do Minho, pp. 227–238, Maio, 2009.

Tabela 5: Lista de tópicos do Páxico.

ID	Descrição	Grande tema	Tema	Lugar
1	Filmes sobre a ditadura ou sobre o golpe militar no Brasil	Artes, Letras	cinema, história	br
2	Telenovelas brasileiras do gênero realismo fantástico.	Artes	televisão	br
3	Orquestras filarmônicas com mais de cinquenta anos de vida em países lusófonos	Artes	música	lus
4	Mulheres violoncelistas de língua portuguesa	Artes	música	lus
5	Flautistas que se naturalizaram brasileiros ou portugueses.	Artes	música	lus
6	Instrumentistas famosos, de países de língua portuguesa, que são conhecidos por tocarem bem mais do que um instrumento.	Artes	música	lus
7	Guitarristas portugueses que também foram compositores.	Artes	música	pt
8	Telenovelas brasileiras passadas no tempo da escravatura no Brasil	Artes, Letras	televisão, história	br
9	Comidas de santo (comidas rituais do Candomblé ou Umbanda) que também fazem parte da culinária brasileira.	Cultura	culinária, religião, antro./fol.	br
10	Pratos brasileiros de origem ou influência indígena.	Cultura, Letras	culinária, história	br
11	Filmes sobre o cangaço.	Artes, Letras	cinema, história	br
12	Políticos de movimentos de libertação lusófonos que foram recebidos pelo Papa	Letras, Política	política, história	lus
13	Dinossauros carnívoros que habitaram o Brasil.	Ciência	zoologia	br
14	Filmes baseados em histórias de Guimarães Rosa.	Artes, Letras	cinema, literatura	br
15	Filmes brasileiros sobre futebol.	Artes, Esportes	cinema, esportes	br
16	Membros da igreja associados à Teologia da Libertação.	Cultura, Letras, Política	filosofia, religião, política	ger
17	Documentários sobre políticos brasileiros.	Artes, Letras, Política	cinema, política, história	br
18	Discos brasileiros considerados marcantes na história da música brasileira	Artes	música	br
19	Tribos indígenas que vivem na Amazônia.	Geografia	demografia, geografia	br
20	Que etnias africanas entraram no Brasil no período da escravatura?	Geografia, Letras	história, demografia	br
21	Lendas e figuras folclóricas de origem portuguesa populares no Brasil	Cultura, Letras	antro./fol., história	lus
22	Filmes sobre o Brasil colonial	Artes, Letras	cinema, história	br
23	Filmes sobre as relações entre Portugal e suas colônias	Artes, Letras	cinema, história	lus
24	Filmes que abordem movimentos históricos como rebeliões, revoltas, batalhas, levantes ou levantamentos populares entre a população (ou parte dela) e o governo vigente em países lusófonos	Artes, Letras, Política	cinema, história, política	lus
25	Que línguas bantas ou bantus são faladas em países lusófonos?	Letras	linguística	lus
26	Artistas brasileiros cujas obras integram o acervo do MoMA ou que já expuseram lá.	Artes	artes plásticas	br
27	Doenças letais comuns em países lusófonos transmitidas por mosquitos	Ciência	saúde	lus
28	Festas populares brasileiras de origem portuguesa	Cultura	antro./fol.	lus
29	Escritores lusófonos que se filiaram a partidos políticos	Letras, Política	literatura, política	lus

Continua na página seguinte...

Tabela 5 – continuação da página anterior

ID	Descrição	Grande tema	Tema	Lugar
30	Batalhas entre brancos e índios brasileiros que ocorreram até o século XVIII	Letras	história	br
31	Grupos indígenas que habitavam o litoral do Brasil quando chegaram os europeus.	Geografia, Letras	história, de- mografia, ge- ografia	br
32	Tribos indígenas brasileiras que praticavam canibalismo em rituais religiosos.	Cultura, Le- tras	história, re- ligião	br
34	Viajantes ou exploradores que escreveram sobre o Brasil do século XVI	Geografia, Letras	história, geo- grafia, litera- tura	br
35	Que autores não lusófonos escreveram sobre o Brasil nos séculos XVIII e XIX?	Letras	história, lite- ratura	br
36	Escolas de samba fundadas ou sediadas em morros cariocas.	Artes, Cul- tura, Letras	música, história, antro./fol.	br
37	Ritmos lusófonos que utilizam zabumba, triângulo e sanfona.	Artes	música	lus
38	Esportes integrados às Olimpíadas nos últimos 30 anos	Esportes	esportes	ger
39	Modalidades esportivas em que países lusófonos já ganharam medalha de ouro nos Jogos Olímpicos.	Esportes	esportes	lus
40	Instrumentos musicais de origem africana comuns no Brasil	Artes, Letras	música, história	br
41	Congressos ou conferências que têm por tema as relações culturais e/ou sociais entre África e demais países lusófonos	Ciência, Cul- tura, Política	ciência, política, cultura	lus
42	Plantas com as quais índios brasileiros pintam seus corpos	Ciência, Cul- tura	botânica, an- tro./fol.	br
43	Produtos agrícolas com os quais se pode produzir combustível em escala comercial	Ciência, Economia	ciência, eco- nomia	ger
44	Lendas ou personagens folclóricas de origem indígena conhecidas no Brasil	Cultura	antro./fol.	br
45	Cantores vaiados nos festivais de música brasileira na década de 60.	Artes, Letras	música, história	br
46	Séries ou minisséries de televisão passadas na época da independência do Brasil	Artes, Letras	televisão, história	br
47	Filósofos franceses do século XX que deram cursos ou conferências no Brasil.	Letras	filosofia	br
48	Jornais portugueses que existiam no tempo da implantação da república	Letras	história, jor- nalismo	pt
49	Séries ou minisséries brasileiras baseadas em romances portugueses.	Artes, Letras	televisão, li- teratura	lus
50	Jornais que circularam no Rio de Janeiro entre 1910 e 1960.	Letras	história, jor- nalismo	br
51	Além do samba, que outros gêneros musicais são populares no carnaval brasileiro?	Artes, Cul- tura	música, an- tro./fol.	br
52	Gêneros musicais que misturam samba e ritmos norte-americanos.	Artes	música	ger
53	Parques do Rio de Janeiro que têm cachoeiras	Geografia	geografia	br
54	Igrejas do Rio de Janeiro construídas por irmandades ou confrarias de negros.	Geografia, Letras	arq./urb., história	br
55	Escritores estrangeiros que visitaram Portugal no século XIX e que publicaram descrições das suas viagens	Letras	literatura, história	pt
56	Jogadores de futebol brasileiros que foram jogar no exterior quando tinham menos de 21 anos de idade.	Esportes	esportes	br
57	Jogadores lusófonos que já foram campeões mundiais por clubes europeus.	Esportes	esportes	lus
58	Países que venceram a Copa do Mundo em uma disputa de pênaltis	Esportes	esportes	ger
59	Jogadores de basquetebol brasileiros que jogam ou jogaram em campeonatos da NBA	Esportes	esportes	br

Continua na página seguinte...

Tabela 5 – continuação da página anterior

ID	Descrição	Grande tema	Tema	Lugar
60	Jogadores de basquetebol brasileiros com mais de 2, 10 metros	Esportes	esportes	br
61	Movimentos culturais em países lusófonos que se refletiram nas artes plásticas e na música	Artes	artes plásticas, música	lus
62	Praias de Portugal boas para a prática de surf	Esportes, Geografia	esportes, geografia	pt
63	Estudiosos da música indígena brasileira	Artes, Cultura	música, antro./fol.	br
64	Doces brasileiros que têm origem nos doces portugueses	Cultura, Letras	culinária, história	lus
66	Fotógrafos de cinema ou diretores de fotografia brasileiros que já dirigiram ou realizaram filmes	Artes	cinema	br
67	Doenças que acometeram a população indígena nas Américas	Ciência, Letras	história, saúde	ger
68	Bandas brasileiras de punk formadas até 1980 em São Paulo.	Artes, Letras	música, história	br
69	Complexos esportivos construídos para os Jogos Pan Americanos do Rio de Janeiro (2007) que nunca receberam quaisquer outros eventos esportivos.	Esportes, Geografia	esportes, arq./urb.	br
70	Centros culturais e faculdades do Rio de Janeiro sediados em prédios históricos	Geografia	arq./urb.	br
71	Doenças presentes no Brasil no século XVII	Ciência, Letras	história, saúde	br
72	Autores lusófonos que escrevem literatura fantástica e que tenham pelo menos um livro publicado	Letras	literatura	lus
74	Nomes ligados à luta contra o racismo no século XX no Brasil	Letras, Política	política, história	br
75	Organizações ou grupos armados que lutaram contra o regime militar no Brasil	Letras, Política	política, história	br
76	Ritmos brasileiros de origem portuguesa	Artes, Letras	música, história	lus
77	Médicos portugueses do século XVIII que viveram a maior parte da sua vida no estrangeiro	Ciência, Letras	história, saúde	pt
78	Escritoras de língua portuguesa que tenham publicado livros para crianças entre 1850 e 1940	Letras	literatura, história	lus
79	Povos indígenas brasileiros considerados extintos.	Geografia, Letras	história, demografia	br
80	Línguas faladas em Timor Leste	Letras	linguística	ti
81	Igrejas em Macau	Cultura, Geografia	religião, arq./urb.	mo
82	Políticos timorenses que participaram na luta armada contra a Indonésia	Letras, Política	política, história	ti
83	Que equipes da primeira divisão do futebol brasileiro desceram para a segunda divisão e nunca mais conseguiram voltar?	Esportes	esportes	br
84	Escritores lusófonos traduzidos para 5 ou mais idiomas	Letras	literatura	lus
85	Destinos turísticos do Brasil cuja temperatura no Inverno pode ser negativa	Geografia	geografia	br
86	Compositoras brasileiras de samba	Artes	música	br
87	Sambistas negros que abordam o racismo em suas letras	Artes, Letras, Política	música, literatura, política	ger
88	Cidades portuguesas que têm festivais medievais	Cultura, Geografia	geografia, antro./fol.	pt
89	Excetuando o português, para que outras línguas consideradas oficiais em países lusófonos existe uma versão da Wikipédia?	Letras	linguística	lus
90	Filmes brasileiros premiados na categoria Montagem.	Artes	cinema	br
91	Estados fronteiriços de Moçambique	Geografia	geografia	mz

Continua na página seguinte...

Tabela 5 – continuação da página anterior

ID	Descrição	Grande tema	Tema	Lugar
92	Cidades que fizeram parte do domínio português na Índia	Geografia, Letras	história, geo- grafia	pt
93	Estilos musicais cabo-verdianos	Artes	música	cv
94	Parques nacionais de Moçambique	Geografia	geografia	mz
95	Partidos políticos de São Tomé e Príncipe	Política	política	st
96	Montanhas mais altas de cada país lusófono	Geografia	geografia	lus
97	Escritores cabo-verdianos com obra publicada em crioulo	Letras	literatura, linguística	cv
98	Cidades dos Estados Unidos que tiveram forte imigração portuguesa	Geografia, Letras	história, geo- grafia	pt
99	Matemáticos de língua portuguesa que estudaram ou trabalharam em Itália	Ciência	matemática	lus
100	Ilhas de Moçambique	Geografia	geografia	mz
101	Cidades e vilas em países não-lusófonos que se situem junto à fronteira de um país lusófono	Geografia	geografia	lus
102	Ordens religiosas que vieram para o Brasil na época colonial.	Cultura, Le- tras	história, re- ligião	br
103	Movimentos culturais surgidos no nordeste do Brasil.	Cultura, Ge- ografia	história, geo- grafia	br
104	Pesquisadores do folclore brasileiro	Cultura	antro./fol.	br
105	Empresas de mineração angolanas	Economia	economia	ao
106	Vice-reis da Índia Portuguesa	Letras	história	pt
107	Dioceses católicas de Moçambique	Cultura	religião	mz
108	Jogadores de futebol nascidos em Cabo Verde que representaram a seleção portuguesa	Esportes	esportes	cv, pt
109	Candidatos a alguma das eleições presidenciais na Guiné-Bissau	Política	política	gw
110	Políticos da África lusófona que estudaram na União Soviética	Política	política	lus
111	Padres católicos que estão ou estiveram ativos em Timor	Cultura	religião	ti
112	Capitais das províncias de Angola	Geografia	geografia	ao
113	Ilhas e ilhotas de Cabo Verde que não são habitadas	Geografia	geografia	cv
114	Antigas moedas dos países lusófonos e suas colónias	Economia, Letras	economia, história	lus
115	Clubes que venceram o Girabola	Esportes	esportes	pt
116	Escritores lusófonos que passaram temporadas na prisão	Letras	literatura	lus
117	Produtos utilizados pelos portugueses no comércio de escravos com a África	Economia, Letras	história, eco- nomia	pt
118	Escritores moçambicanos que receberam o Prémio Camões	Letras	literatura	mz
119	Pratos típicos da gastronomia de Cabo Verde	Cultura	culinária	cv
120	Cervejas consumidas em Angola	Cultura	culinária	ao
121	Frutos de Angola	Ciência, Ge- ografia	geografia, botânica	ao
122	Políticos lusófonos do século XX assassinados	Letras, Política	política, história	lus
123	Rios de Angola com mais de 500 quilómetros de comprimento	Geografia	geografia	ao
124	Cabo-verdianos que participaram na guerra colonial na Guiné	Letras, Política	história, política	gb, cv
125	Fortalezas e feitorias portuguesas na costa africana	Geografia, Letras	história, geografia, arq./urb.	pt
126	Atletas da África lusófona que tenham competido nos Jogos Olímpicos	Esportes	esportes	lus
127	Mamíferos herbívoros existentes em Moçambique	Ciência, Ge- ografia	geografia, zoologia	mz
128	Escritores portugueses que tenham vivido em Macau	Letras	literatura	mo, pt
129	Antigos alunos da Universidade Eduardo Mondlane e da sua antecessora, a Universidade de Lourenço Marques	Cultura	ensino	mz

Continua na página seguinte...

Tabela 5 – continuação da página anterior

ID	Descrição	Grande tema	Tema	Lugar
130	Acordos, tratados e outros protocolos assinados entre as fações da Guerra Civil angolana	Letras, Política	história, política	ao
131	Quem descobriu São Tomé e Príncipe?	Letras	história	st
132	Deputados da FRELIMO	Política	política	mz
133	Futebolistas do Petro de Luanda	Esportes	esportes	ao
134	Personagens do século XIX ligadas à luta anti-colonial em Moçambique	Letras, Política	história, política	mz
135	Aves de Angola	Ciência, Ge- ografia	geografia, zoologia	ao
136	Provas desportivas internacionais com participação da seleção moçambicana	Esportes	esportes	mz
137	Eventos onde Maria de Lurdes Mutola foi medalha de ouro	Esportes	esportes	mz
138	Jornais, revistas e outras publicações periódicas de Macau	Letras	jornalismo	mo
139	Infra-estruturas de transportes (aeroportos, estações, rodoviárias e ferroviárias, pontes, etc.) existentes em Macau	Geografia	arq./urb.	mo
140	Cidades lusófonas conhecidas pelo seu Carnaval	Cultura	antro./fol.	lus
141	Cidades das antigas colónias portuguesas que têm ou tiveram a designação de Nova	Geografia, Letras	geografia, linguística	lus
142	Diplomatas lusófonos que trabalharam no Japão	Letras, Política	história, política	lus
143	Cidades do império português ocupadas pelos holandeses no período filipino	Letras	história	pt
144	Locais referidos n Os Lusíadas	Letras	literatura	pt
145	Minas de países lusófonos em atividade durante mais de 30 anos consecutivos, ainda em funcionamento ou não	Ciência, Economia, Geografia	geografia, geologia, economia	lus
146	Vulcões em território português e brasileiro	Ciência, Ge- ografia	geografia, geologia	lus
147	Museus em capitais de países lusófonos	Artes, Cul- tura	cultura, ar- tes	lus
148	Primeiras universidades de cada país lusófono	Cultura, Le- tras	história, en- sino	lus
149	Arquitetos de países lusófonos com obras em países estrangeiros na América do Norte e na Europa	Geografia	arq./urb.	lus
150	Empresários lusófonos com uma fortuna considerável	Economia	economia	lus
151	Cidades em países não lusófonos com mais de 500 mil habitantes ou com mais de 10% da população falando português	Geografia, Letras	geografia, linguística	lus
152	Pintores estrangeiros com uma ligação forte a Portugal ou ao Brasil	Artes	artes	lus
153	Toureiros a cavalo de países lusófonos com carreira internacional	Cultura, Es- portes	antro./fol., esportes	lus