

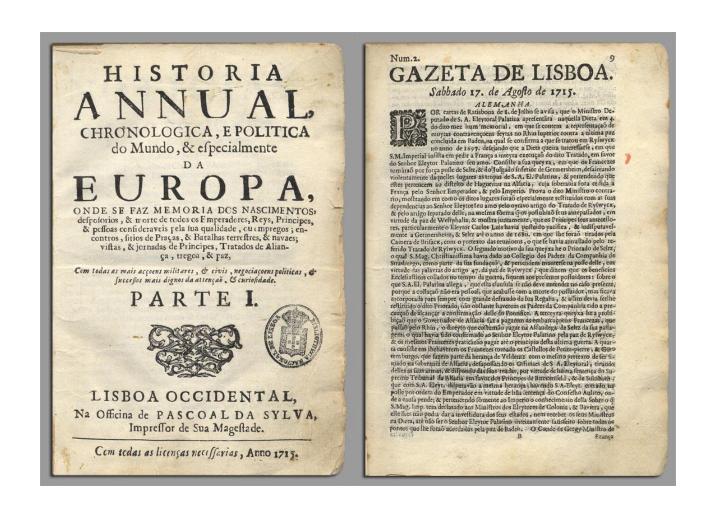
Arquivo.pt: catálogo de serviços para preservação digital

Daniel.Gomes@fccn.pt

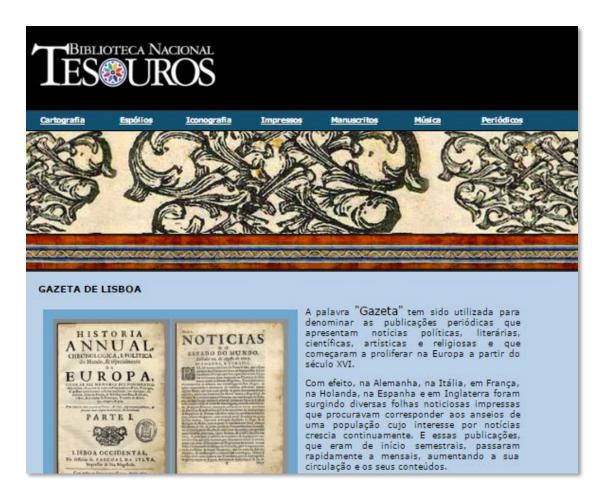




"Gazeta de Lisboa" foi 1º jornal impresso português, iniciado em 1715

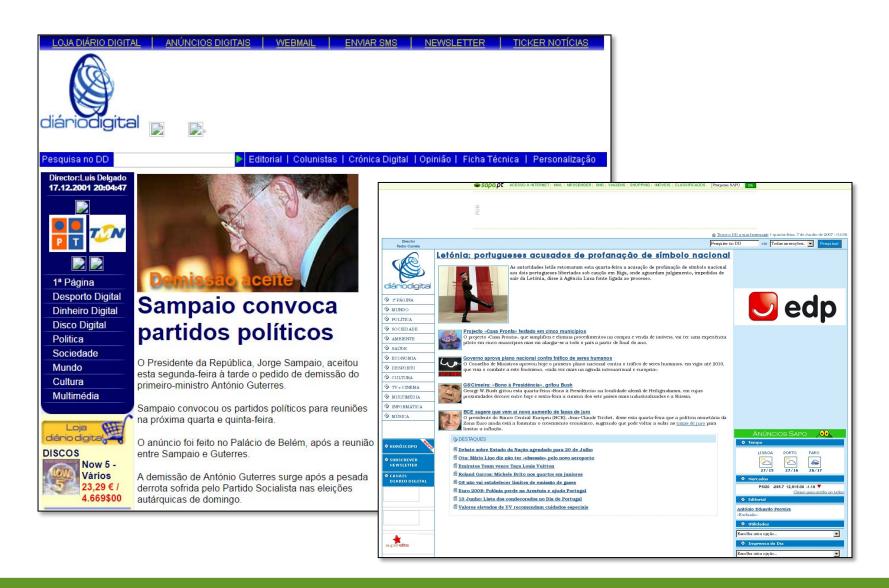


"Gazeta de Lisboa" foi suspenso em 1762.



300 anos depois as suas publicações continuam acessíveis

"Diário Digital" foi o 1º jornal português **exclusivamente online**, iniciado em **1999**



"Diário Digital" desapareceu em 2017



Apenas após algumas décadas, como acedemos às publicações do "Diário Digital"?



Página não encontrada.

Era Digital

Abundância de informação vs.

Escassez de Memória

São também necessários Arquivos para a Web!

Preservar a Imprensa



Preservar a Web

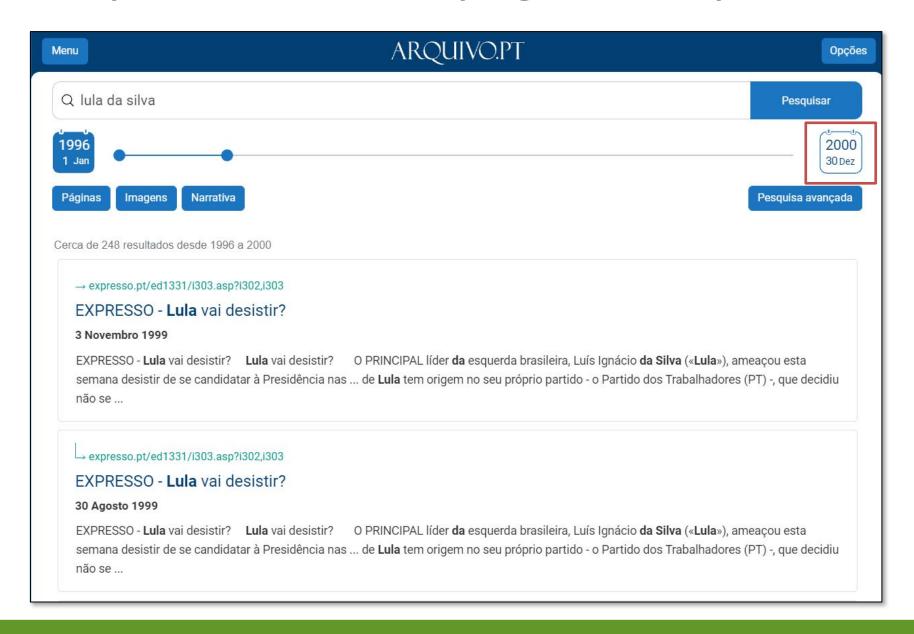






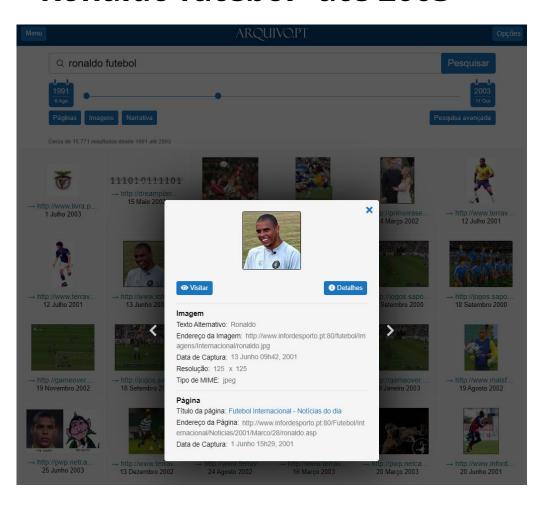


Pesquise textos de páginas do passado

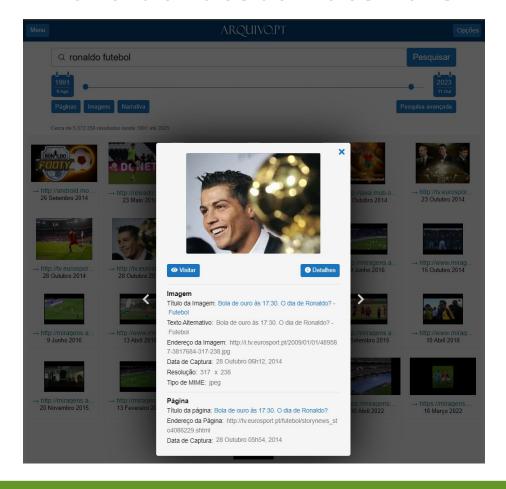


Pesquise imagens do passado

"Ronaldo futebol" até 2003



"Ronaldo futebol" até 2023



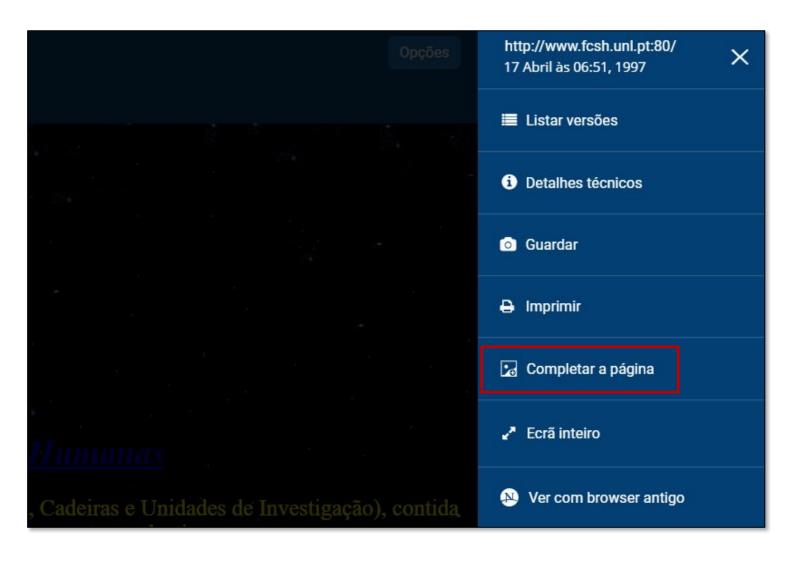
Pesquise o **histórico** de um endereço



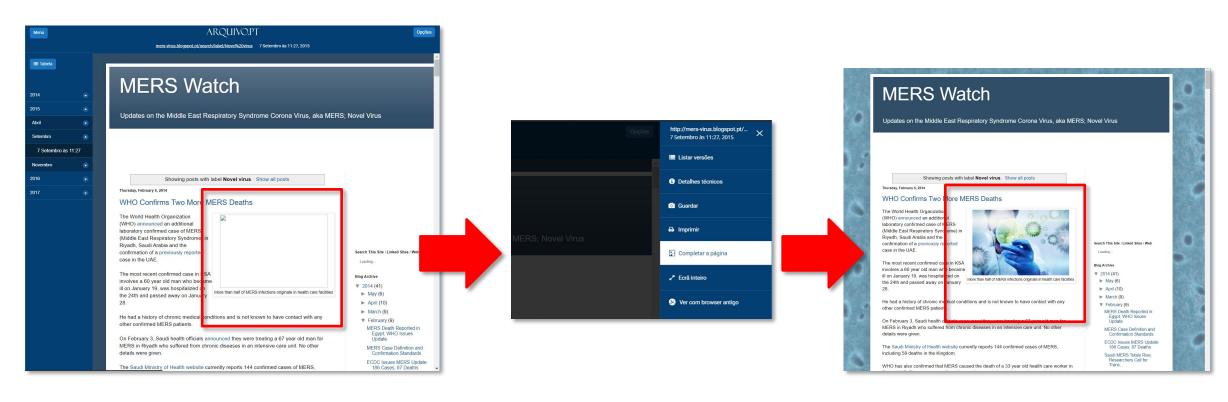
Reprodução de uma página arquivada



Opções de reprodução: Completar a página



"Completar página"



Ficheiros obtidos são **integrados** no acervo do Arquivo.pt: Curadoria colaborativa!

Pesquisar noutros arquivos (ex. Internet Archive)



<u>deutsche-guggenheim-berlin.de</u>, 2001 @Arquivo.pt



<u>deutsche-guggenheim-berlin.de</u>, 2001 @Archive.org



1 único arquivo da web não chega





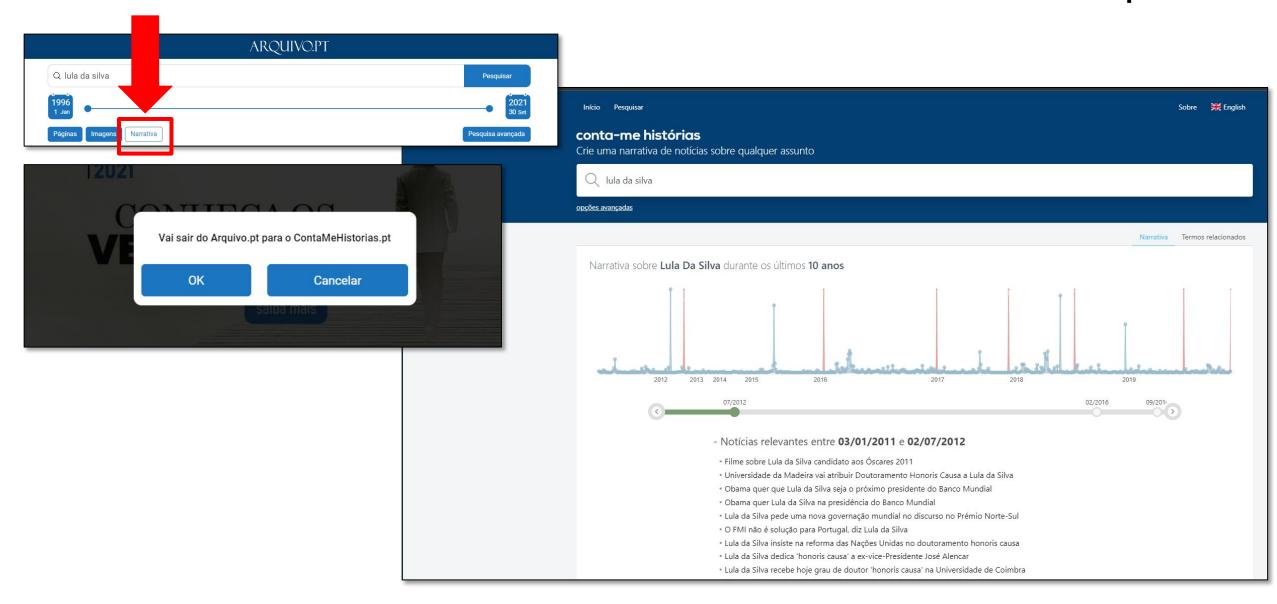
Temporarily Offline

Internet Archive services are temporarily offline.

Please check our official accounts, including Twitter/X, Bluesky or Mastodon for the latest information.

We apologize for the inconvenience.

Narrativas automáticas com ContaMeHistorias.pt





Exemplos de conteúdo web brasileiro preservado

Eleições brasileiras em 1998, descritas no jornal português Expresso



Lula vai desistir?



O PRINCIPAL líder da esquerda brasileira, Luís Ignácio da Silva («Lula»), ameaçou esta semana desistir de se candidatar à Presidência nas eleições de Outubro, deixando o campo livre para a reeleição de Fernando Henrique Cardoso.



O desânimo de Lula tem origem no seu próprio partido - o Partido dos Trabalhadores (PT) -, que decidiu não se aliar ao Partido Democrático Trabalhista (PDT) no Rio de Janeiro, inviabilizando a candidatura da dupla Lula-Leonel Brizola (líder do PDT), o primeiro à chefia do Estado e o segundo à vice-presidência.



O fracasso deve-se à vitória dos radicais do PT, que conseguiram impor a candidatura do ex-líder estudantil Vladimir Palmeira ao Governo do Rio, derrotando os moderados, que se tinham comprometido a apoiar o candidato de Brizola, para garantirem que este último concorreria como «vice» de Lula







Notícias de alguns jornais brasileiros Folha de São Paulo, 18/2/2008



Estadão, 21/5/2009



Brasil na final

Jade de graça

Justica manda PF

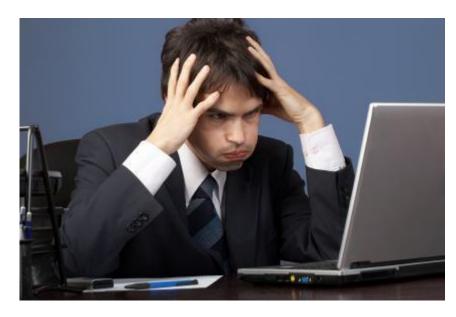
Líder opositor é



Arquivo.pt

Casos de uso

Gestão: documentação/auditoria de projeto concluído



"Estava tudo no site"

Linguateca R

Estrutura Equipa

Apresentação Acesso a recurso

- Acesso a recursos
- CETEMPúblico
 CETENFolha
- CHAVE
 COMPARA
- Corpógrafo
- CorTrad
- Esfinge
 Floresta Sintá(c)tica
- GeoNetPT
 DADET
- PANTERA
- PoNTE
 ReEscreve
- REPENTINO
- Repositório
 WebJspell
- WPT

Avaliação conjunta
Catálogo de recursos
Catálogo de ferramentas
Catálogo de atores
Catálogo de publicações
Informação interessante
Perguntas já respondidas
Gramateca

English

Repositório de recursos, publicações e programas

Linguateca

Neste espaço providenciamos uma "estante electrónica" para todos quantos quiserem ter os seus trabalhos relacionados com o processamento computacional da lingua portuguesa acessiveis directamente da rede.

Basta enviarem-nos a versão electrónica, indicando a referência que queiram que apareça no nosso <u>catálogo de publicações</u>, ou, no caso de programas ou outros recursos, enviarem-nos o respectivo programa e documentação básica.

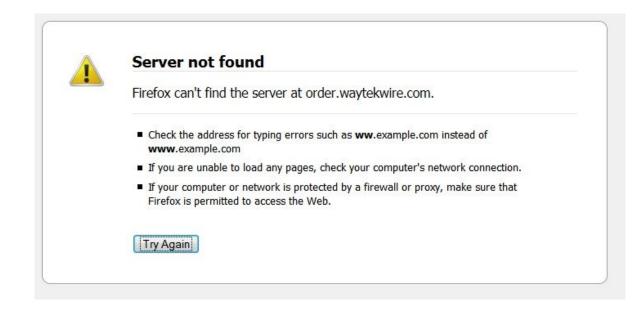
De momento, temos os seguinte recursos importantes (por ordem de disponibilização), em que indicamos também, com gratidão, o responsável pela sua cedência, e a data a partir da qual ficaram disponíveis no Repositório

- Regras de expressões bilingues inglês-português. Logos Corporation (Anabela Barreiro, Janeiro de 2001)
- Base de Dados para Sintetizador de Fala em Português Europeu, FEUP/IPB-BD (João Paulo Teixeira, Abril de 2001)
- <u>Dicionário Portuguez-Latino de Pedro José da Fonseca</u> (Digitalizado por Ricardo Ueda Karpischek, Julho de 2003)
- Colecção FolhaRIcol (Vera Lúcia Strube de Lima, Agosto de 2003)
- Corpus TeMário (Lucia H. Machado Rino, Outubro de 2003)
- Colecção WBR-99 (Berthier Ribeiro-Neto, Nivio Ziviani, Dezembro de 2003)
- Port4NooJ: recursos linguísticos para o português no formato NooJ (Anabela Barreiro, Novembro de 2007)
- Léxico-gramática dos adjectivos intransitivos humanos do português europeu (Paula Carvalho, Dezembro de 2008)
- Erros ortográficos em português de Portugal e sua correcção (José Carlos Medeiros, Agosto de 2000)
- Memórias de tradução de português de Portugal para inglês britânico (Março de 2010)
- Correcção ortográfica de português (Abril de 2010)
- ReLi, corno de resenhas de livros (Cláudia Freitas, Maio de 2013)

Portfólio/CV

"O meu melhor trabalho foi um site que já não existe."

60% dos sites desaparecem após apenas 2 anos.



Documentação acerca de um evento



EMPRESA

PRODUTOS

SERVIÇOS



MERCADOS

0.8

MÉDIA

CONTACTOS

VII Congresso Nacional de Arquivologia – CNA Brasil



Realizar um congresso, grandioso e importante como este para os profissionais arquivistas, é um trabalho árduo, mas também prazeroso, pois é ele um marco para o avanço da teoria arquivística e de suas tecnologias para a comunidade brasileira. Comunidade esta, cada vez mais, exigente e consciente da importância da gestão documental e informacional, considerando não somente a atividade fim da arquivística, mas ainda, sob um olhar na sustentabilidade e nas inovações que contribuem para o desenvolvimento e uma melhor aplicabilidade da gestão documental nas empresas e demais espaços de atuação do profissional arquivista.

espaços de atuação do profissional arquivista.

due contribuent para o descrivoramiento e una memor apricabilidade da Sestao documenta nas empresas e demais



Arquivo.pt

O que fazer para que um site seja preservado

Dissemine e siga boas práticas na construção de sites

sobre.arquivo.pt > Recomendações para a publicação na Web de informação preservável

Recomendações para a publicação na Web de informação preservável

Última atualização em 25 de Setembro de 2017 às 10:48

Apresenta recomendações para que as páginas de um sítio Web possam ser corretamente arquivadas e acedidas ao longo dos anos. É impossível preservar a web portuguesa sem a colaboração dos autores.

Última atualização em 2015-08-03.

O funcionamento dos arquivos da web é semelhante ao dos motores de busca como o Google. O respeito pelas recomendações seguintes contribuem para o correto funcionamento dos arquivos e dos motores de busca sobre a web.

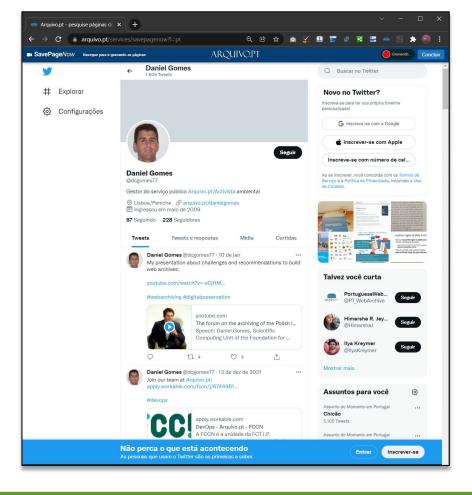
As recomendações apresentadas em seguida são classificadas em *fundamentais* e *aconselháveis*. O desrespeito por uma recomendação fundamental impossibilita o arquivo da informação de um sítio web.

arquivo.pt/recomenda

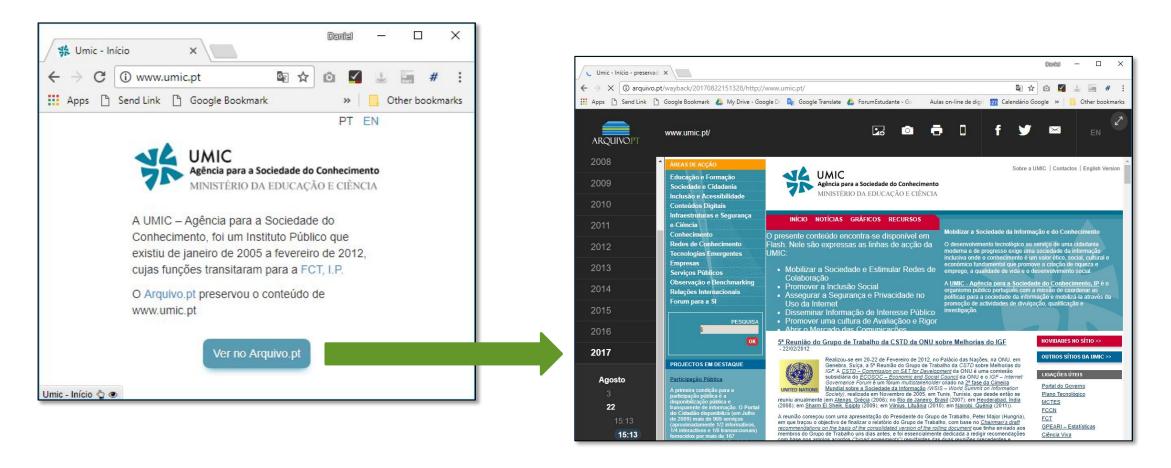
Arquive página agora!



arquivo.pt/savepagenow



Não mate o seu site histórico! Preserve-o no Memorial do Arquivo.pt: <u>arquivo.pt/memorial</u>



Antes que o seu site "abandonado" seja atacado, envie-o para o Memorial e evite problemas de segurança, custos e perdas de informação!



Colaboração com Brasil

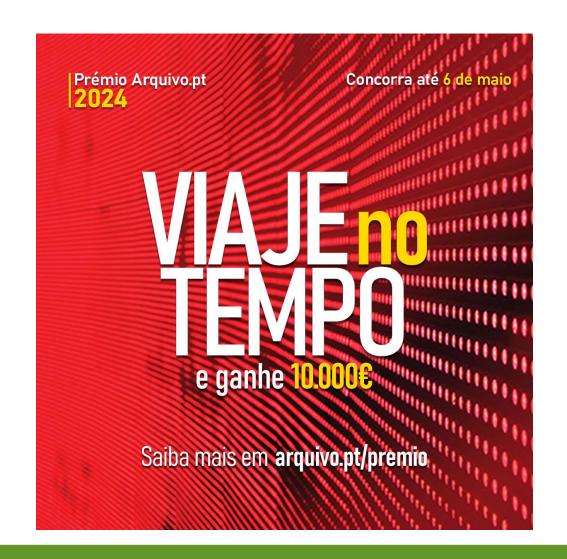


Origem dos utilizadores do Arquivo.pt

Country	Users % Users	
1. Portugal	72,493 44.9	6%
2. United States	23,006 14.27%	
3. Brazil	6,356 3.94%	
4. 📺 United Kingdom	5,394 3.35%	
5. 🇉 Japan	4,231 2.62%	
6. Germany	3,460 2.15%	
7. Maria Canada	2,667 1.65%	
8. India	2,494 1.55%	
9. Russia	2,414 1.50%	
10. Spain	2,340 1.45%	

Precisamos de ajuda na **divulgação** para que os brasileiros também tirem partido do Arquivo.pt!

Divulgue o Prémio **anual** Arquivo.pt!





arquivo.pt/premios



Para saber mais acerca de arquivos da web

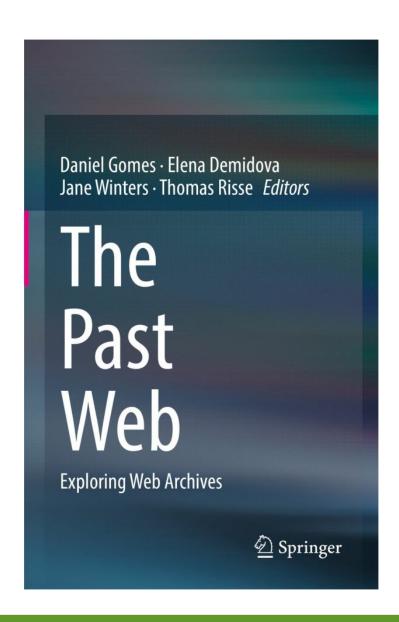
Documentação **para** formação: arquivo.pt/forma

- Uma nova ferramenta para pesquisar o passado
- Bem publicar, para bem preservar
- Acesso e processamento automático de informação preservada da Web
- Arquivar a web: faça-você-mesmo!



WARC file format - ISO 28500:2017

Livro acerca de memória e preservação digital



- Versão "pre-print" em acesso aberto
- arquivo.pt/livro



Catálogo de serviços do Arquivo.pt

Última atualização em 19 de Setembro de 2023 às 10:20

Públicos

- Pesquisa e acesso sobre a informação preservada da web
- Interfaces para Programação de Aplicações (APIs)
- · Sugira websites para serem preservados
- · SavePageNow: arquive imediatamente páginas da web
- Integração de coleções históricas de dados da web
- Formação acerca de preservação da web
- Dados abertos que listam informação arquivada da web relativa a vários temas
- CitationSaver: extrai os links contidos em documentos e arquiva as correspondentes páginas
- Arquivo404: mostra páginas preservadas em vez de páginas não encontradas

Para entidades parceiras

- Memorial: preserve a informação do seu antigo website antes de o desativar
- Arquivo a-pedido de websites em alta-qualidade
- Criação online de coleções e exposições temáticas
- Exposições itinerantes de cartazes que exibem páginas históricas da web

Todos os serviços em arquivo.pt/catalogo

Documentos publicados em-linha **também** são Património!

A informação publicada online é **valiosa**.

Preservar a memória digital é responsabilidade de todos.

Inscreva-se na lista de email para não perder as **informações mais importantes**

Formações, coleções colaborativas, prémios, recrutamento, novos projectos e serviços <u>arquivo.pt/inscrever</u>

Ajude a divulgar o Arquivo.pt para que seja útil a mais pessoas!

Inscreva-se para poder receber e divulgar novidades acerca do Arquivo.pt (ex. novas coleções, projetos, colaborações, bolsas).

Enviaremos apenas as informações mais importantes e pode desinscrever-se quando quiser.

Nome	Email	Instituição
	_	