

Um serviço público **transversal** para a **transformação digital**

daniel.gomes@fccn.pt

A transformação digital **mudou** o consumo de informação



Flagelo das sociedades digitais: **efemeridade de informação**

Informação centenária
acessível



Passado recente
inacessível (século XXI)



80%

das páginas da web **desaparecem**
ou mudam
passado apenas **1 ano.**

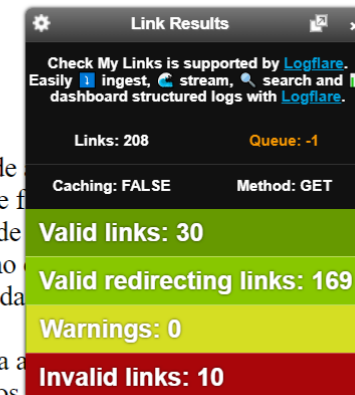
CVs incompletos

“O meu melhor trabalho **estava** num website que já não existe.”

Projectos

Investigação

- **QualWeb** (2008-2011): realizou avaliações de pessoas com deficiências em larga escala e de f
- **Tomba** (2005-2006): é o primeiro protótipo de web portuguesa. Foi responsável pelo desenho sistema de integração de dados provenientes da técnica do projecto.
- **XMLBase** (2002-2004): este projecto visou a implementação de sistemas de gestão de dados responsável pelo desenho e desenvolvimento da infra-estrutura de armazenamento e acesso a dados.
- **Tumba!** **503** **503** (2000-2006): é um motor de busca otimizado para a web portuguesa. Foi responsável pelo desenho, desenvolvimento e manutenção do sistema de recolha, armazenamento e acesso a dados web.
- Digital Deposit (12/2000-6/2001): estudou o desenho de um sistema de suporte ao arquivo selectivo de publicações *online*. Este projecto foi desenvolvido em colaboração com a **Biblioteca Nacional**. Foi responsável pelo sistema de colecção e gestão de publicações.
- Ariadne (12/2000-6/2001): visou a criação de uma infra-estrutura de publicação de notícias através da Internet. Este projecto foi desenvolvido em colaboração com o jornal **Publico**. Foi responsável pelo desenvolvimento de um Guia do Lazer personalizado móvel baseado na utilização de dispositivos móveis Palm Pilot e pela middleware e interface de utilização de um sistema de pesquisa temática sobre várias publicações online (**Newssearch**).



Habilitações académicas

- Doutoramento em Engenharia Informática pela **Universidade de Lisboa** aprovado com distinção e louvor (2003/2007).

“Apagão” de eventos recentes



This site can't be reached

Check if there is a typo in www.parquedasnacoes.pt.

If spelling is correct, [try running windows network Diagnostics](#).

DNS_PROBE_FINISHED_NXDOMAIN

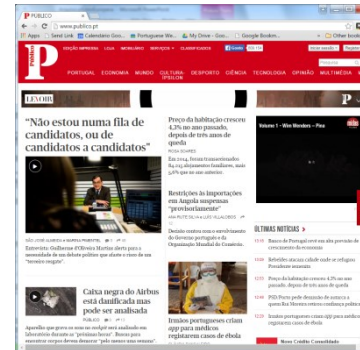
Reload

Paradigma da preservação de informação **transformou-se** com o digital

Preservar a Imprensa



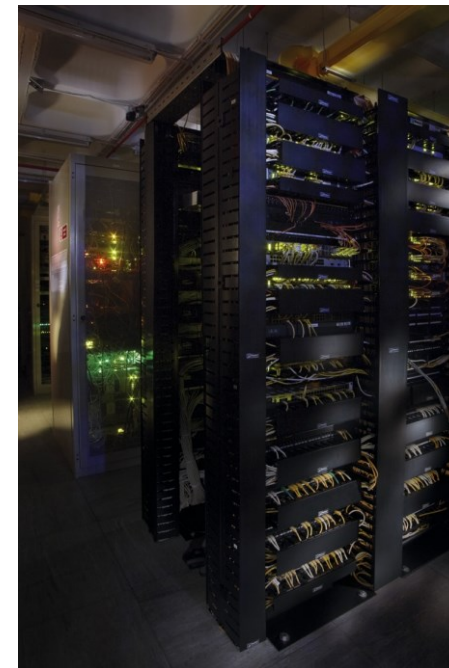
Preservar a Web



Arquivo.pt: serviço **cloud** para preservação digital da informação publicada online.

Dados são o Ouro do século XXI!

- 19 000 milhões de ficheiros desde 1990s
- 46 milhões de websites
- 1 PetaByte
- 408 TB de dados históricos servidos em 2023



Datacenter do Arquivo.pt

Investimento preservado no Arquivo.pt

Resources » Articles » How Much Does It Cost to Build a Website? (2023 Data)

How Much Does It Cost to Build a Website? (2023 Data)

 The Upwork Team
Apr 21, 2021 | 10 Min Read

€5 600

€5 600 x 46M websites = €**257 600 milhões**

PIB Portugal 2022 = €**242 300 milhões**

Pesquisar **histórico** de uma página

ARQUIVO.PT

idc.pt

Pesquisar

1991

6 Ago

2023

2 Nov

Páginas

Imagens

Narrativa

Pesquisa avançada

Cerca de 538 resultados desde 1991 até 2023

Tabela

Lista

2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
21 Jul	18 Jan	6 Fev	29 Abr	6 Fev	10 Jan	5 Fev	12 Mar	23 Jun	5 Jun	22 Jan	21 Jan	5 Nov	26 Set	16 Abr	7 Fev	19 Fev	17 Jul	1 Jan	3 Jan	1 Jan	21 Jan
22 Set	28 Mar	14 Fev	5 Jun	8 Fev	14 Fev	25 Jun	21 Out	25 Set	4 Ago	20 Mai	22 Jan	7 Nov	30 Set	23 Abr	10 Fev	22 Fev	20 Jul	2 Jan	4 Jan	2 Jan	21 Jan
29 Nov	5 Jun	16 Fev	6 Jun	9 Fev	19 Fev	26 Jun		18 Dez		21 Mai				19 Nov	14 Mar	3 Ago	25 Jul	3 Jan	5 Jan	12 Jan	3 Fev
	21 Ago	16 Fev	9 Jun	4 Mar	18 Jun	19 Jul				1 Jul					17 Mai	6 Ago	26 Jul	4 Jan	6 Jan	14 Jan	21 Fev
	24 Set	30 Mar	30 Jul		8 Jul	11 Ago				2 Jul							27 Jul	12 Jan	7 Jan	15 Jan	15 Abr
	27 Set	10 Abr	28 Ago		31 Ago	14 Ago											28 Jul	13 Jan	8 Jan	3 Abr	15 Abr
	27 Nov	23 Abr	12 Set		6 Set	11 Set											29 Jul	14 Jan	9 Jan	6 Abr	4 Mai
	29 Nov	19 Jun	26 Set		6 Set	5 Out											30 Jul	15 Jan	10 Jan	7 Abr	26 Mai
		24 Jun	26 Set		8 Set	17 Dez											31 Jul	16 Jan	11 Jan	13 Mai	29 Jul
		25 Jul	13 Dez														1 Ago	17 Jan	13 Jan	13 Mai	29 Jul

Navegar pelo passado de um website

idc.pt, 2002

Menu

ARQUIVO.PT

idc.pt:80 5 Junho às 22h02, 2002

Opções

Tabela

2001

2002

Janeiro

Março

Junho

5 Junho 22h02, 2002

Agosto

Setembro

Novembro

2003

2004

2005

2006

2007

2008

2009

Sobre a IDC

Relatórios Locais

Relatórios Internacionais

Cadernos IDC

Conferências

Contacte-nos

A IDC no Mundo

Novas Notícias

IDCNet Login

NOVAS NOTÍCIAS

A IDC prevê que a taxa de crescimento média anual composta dos gastos no sector industrial seja, entre 2001 e 2005, ligeiramente inferior à média de todos os sectores, cifrando-se nos 8,5%. ▶ . . .

NOVAS PESQUISAS

ERM - Enterprise Resource Management: Definições, Mercado e Tendências ▶ . . .

Software de

Empresariais: Definições, Mercado e Tendências ▶ . . .

DIRECTIONS '2002

15 e 16 de Outubro de 2002
Hotel Mendon - Lisboa

IDC
Analyze the Future

A liderança das TMTs na retoma da economia: mudanças à vista

Para informações e inscrições: portugal@idc.com

idc.pt/press.asp#noticia14, 2002

Tabela

2001

2002

Fevereiro

Agosto

2 Agosto 03h37, 2002

2003

A IDC prevê que a taxa de crescimento média anual composta dos gastos no sector industrial seja, entre 2001 e 2005, ligeiramente inferior à média de todos os sectores, cifrando-se nos 8,5%.

Esta foi uma das conclusões do estudo "Indústria, Mercado e Tendências de Investimento em TI: 2001-2005".

O sector industrial apresenta um peso considerável no tecido empresarial português e no total dos gastos em Tecnologias de Informação em Portugal. Com quase 25% do total do volume de negócios gerados pelos sectores económicos e mais de 30% dos trabalhadores, estimamos que o sector industrial tenha representado, em 2001, um pouco mais de 1/5 dos gastos totais em Tecnologias de Informação.

Opções avançadas de reprodução (ex. *Completar página*)



Ficheiros em falta são obtidos automaticamente a partir dos websites originais ou de outros arquivos da web.

Pesquisa de páginas com o Arquivo.pt: um “Google” para o **passado**!



bbc.com
https://www.bbc.com › articles

'às vezes eu simplesmente paro e choro atrás da câmara', diz ...

2 days ago — As memórias da **guerra** podem surgir, mas pelo menos temos segurança física, assim como as pessoas que amamos. A Faixa de **Gaza** é um lugar muito ...

expresso.pt
https://expresso.pt › eixo-do-mal

Cerco a Gaza, continua a guerra sem quartel. E, por cá, mais ...

26 minutes ago — A Norte, a cidade está cercada e há combates, campos refugiados continuam a ser bombardeados. É uma **guerra** sem quartel. Aguarda-se pela chegada ...

youtube.com
https://www.youtube.com › watch

Guerra de Israel: Novos bombardeios na Faixa de Gaza; veja



Guerra de Israel: Novos bombardeios na Faixa de **Gaza**; veja | CNN 360°. 168K views · 13 hours ago #CNNBrasil ...more. CNN Brasil. 4.44M.

YouTube · CNN Brasil · 13 hours ago

Menu

ARQUIVO.PT

Opções

Pesquisar

1991
6 Ago

2005
31 Dez

Páginas

Imagens

Narrativa

Pesquisa avançada

Cerca de 53.338 resultados desde 1991 até 2005

jornal.publico.pt/publico/2000/11/21/Destaque/X01.html

POL | Destaque | Israel intensifica guerra em Gaza

21 Novembro 2000

POL | Destaque | Israel intensifica guerra em Gaza SECÃŁS 1ª PÃgina Destaque Editoriais PolÃ ... Guerra em Gaza Por GRAHAM USHER, em JerusalÃm Terça-feira, 21 de Novembro de 2000 Enganaram-se os que ... OUTROS TÃTULOS EM DESTAQUE Israel intensifica guerra em Gaza Shas longe de Barak "O ExÃrcito tem ...

blog-19.blogspot.com/2003/10/amnistia-acusa-o-exrcito-israelita-de.html

:: BLOG 19 :: Amnistia acusa o exército israelita de "crime de guerra" em Gaza

1 Junho 2005

:: BLOG 19 :: Amnistia acusa o exército israelita de "crime de guerra" em Gaza BlogThis! :: BLOG ... acusa o exército israelita de "crime de guerra" em Gaza Na LUSA : A Amnistia Internacional condenou vigorosamente as operações militares lançadas sexta-feira no sul de Gaza pelo exército israelita ...



Pesquisa também imagens do passado!

Menu

ARQUIVO.PT

Opções

Pesquisar

1991
6 Ago

2005
3 Nov

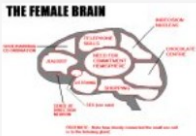
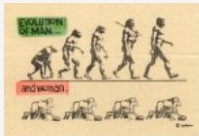
Páginas

Imagens

Narrativa

Pesquisa avançada

Cerca de 14.225 resultados desde 1991 até 2005

 → http://barnabe.w... 2 Outubro 2003	 → http://tsf.sapo.pt... 4 Julho 2002	 → http://www.rna.pt... 2 Maio 2003	 → http://www.geb-p... 25 Junho 2004	 → http://clientes.net... 9 Janeiro 2004	 → http://dn.sapo.pt/... 6 Fevereiro 2005	 → http://www.elpais... 11 Agosto 2005	 → http://tsf.sapo.pt... 13 Junho 2002
 → http://tsf.sapo.pt... 4 Julho 2002	 → http://www.tsf.pt... 8 Janeiro 2005	 → http://tsf.sapo.pt... 1 Outubro 2004	 → http://tsf.sapo.pt/... 4 Outubro 2001	 → http://eventos.cli... 20 Dezembro 2001	 → http://tsf.sapo.pt... 26 Janeiro 2005	 → http://www.geb-p... 23 Julho 2004	 → http://tsf.sapo.pt... 23 Dezembro 2004
 → http://www.terrav... 9 Março 2002	 → http://www.terrav... 9 Março 2002	 → http://www.terrav... 7 Março 2002	 → http://www.terrav... 8 Março 2002	 → http://www.terrav... 10 Março 2002	 → http://www.terrav... 8 Março 2002		

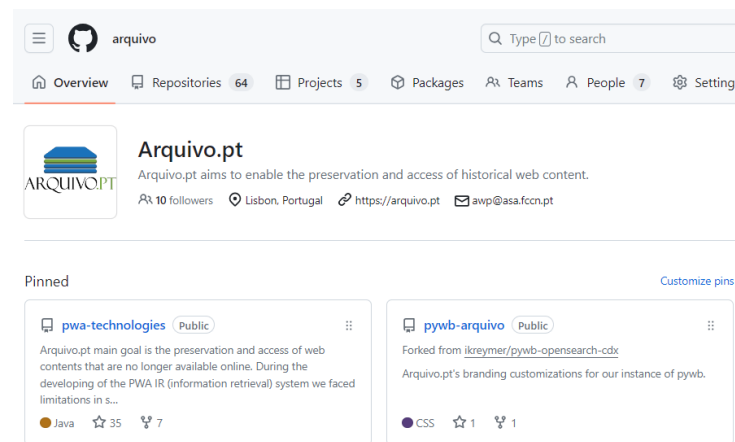
Application Programming Interfaces e código-aberto para **desenvolver novas aplicações**

APIs

Daniel Gomes edited this page on 27 Jul · 29 revisions

- [Arquivo.pt API \(Full-text & URL search\)](#)
- [CDX-server API \(URL search\)](#)
- [Memento API \(URL search\)](#)
- [Image Search API v1.1 \(beta version\)](#)

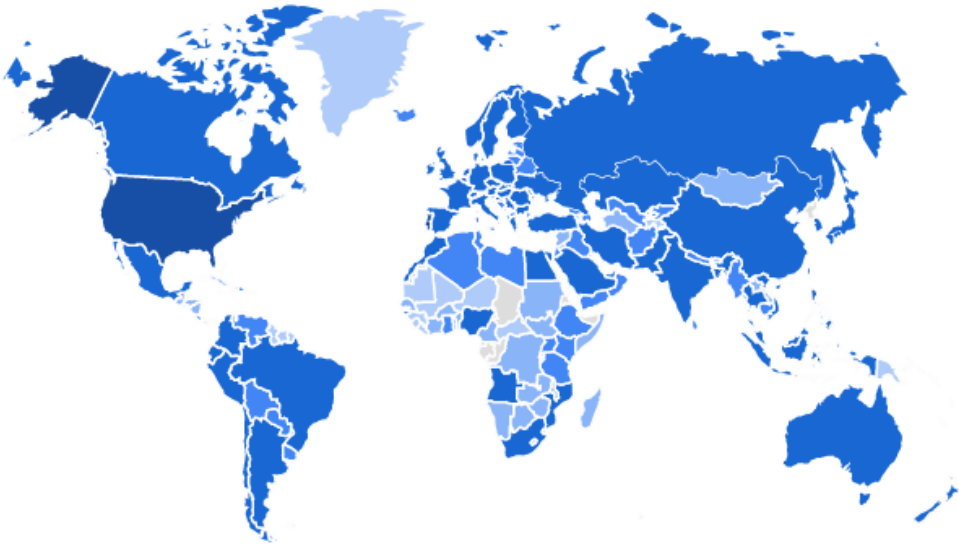
arquivo.pt/api



github.com/arquivo

Arquivo.pt é um serviço **internacional** e **interdisciplinar**

Metade dos utilizadores são internacionais



Fonte de dados em várias áreas
(ex. Inteligência Artificial)

Generating a European Portuguese BERT
Based Model Using Content from Arquivo.pt
Archive

[Nuno Miquelina](#) , [Paulo Quaresma](#) & [Vitor Beires Nogueira](#)

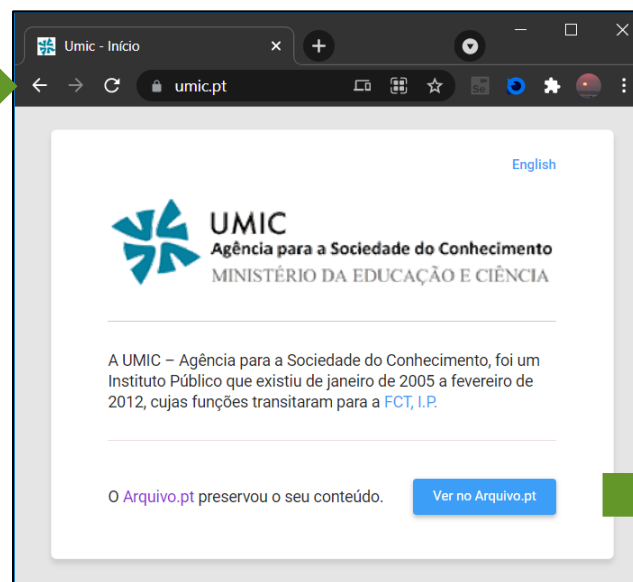
Conference paper | [First Online: 21 November 2022](#)

Serviços **complementares**
para apoio às **organizações**

Não mate o seu **website histórico!** Preserve-o no **Memorial** do Arquivo.pt: arquivo.pt/memorial



- ✓ Informação pesquisável e acessível
- ✓ Cibersegurança
- ✓ Reduz CO2
- ✓ Poupa custos



Arquivo404: apresente **versões arquivadas** em vez de páginas não encontradas



Página não encontrada no website atual



Página preservada no Arquivo.pt

Conheça todos os serviços do Arquivo.pt

Catálogo de serviços do Arquivo.pt

Última atualização em 13 de Fevereiro de 2023 às 14:50

Públicos

- [Pesquisa e acesso](#) sobre a informação preservada da web
- [Interfaces para Programação de Aplicações \(APIs\)](#)
- [Sugira websites](#) para serem preservados
- [SavePageNow](#): archive 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](#)

arquivo.pt/catalogo