

Arquivo.pt

Taking a Web Archive to the next level

João Gomes

Director, Internet Advanced Services / FCT

15th November 2018

Arquivo.pt: 10 years of na alive archive



Photo courtesy of Eric BakerCC BY-NC-SA 2.0

2017 and 2018: bet on dissemination



Easy and fast access



Insira palavra ou URL

Pesquise páginas do

Milhões de conteúdos da web preservados



Responsive interfaces



Trainings and Lectures



Event Arquivo.pt and presence in scientific events



Digital Marketing



Colaboration and sharing



INTERNATIONAL
INTERNET
PRESERVATION
CONSORTIUM



RHIZOME

BdG Biblioteca de Galicia



Press Releases and Partnerships



● A TSF chega à Internet 1996

Veja aqui os primeiros passos do site da TSF. Sob a direção de Carlos Andrade, a atualização do site era, na altura, da responsabilidade do jornalista António Jorge Branco.

● 1998 Expo 98 e o trânsito em Lisboa

A Expo 98 provocou grandes alterações na capital. A informação de trânsito, conteúdo essencial na antena, não podia faltar no site da TSF.

● Timor 1999

A independência de Timor foi um marco na história contemporânea mundial. O antes e o depois do referendo foi acompanhado dia e noite pela TSF.

● 2001 Sinais de Fernando Alves

Desde 1995, a voz de Fernando Alves enchia a antena da TSF com a crónica Sinais. Em janeiro de 2001, o programa faz a sua estreia no online da TSF. A crónica "Artur, meu Presidente" é um dos registos mais antigos

Scientific publications and good practices videos

Publicações Científicas

- Daniel Bicho, Daniel Gomes, Preserving Websites Of Research & Development Projects, International Conference on Digital Preservation, Outubro de 2016 (ppt, bibtex).
- Miguel Costa, Daniel Gomes, Mário J. Silva, The Evolution of Web Archiving, International Journal of Digital Library Systems, Abril de 2016 (bibtex).
- Miguel Costa, Information Search in Web Archives, Tese de Doutoramento, Universidade de Lisboa, Dezembro de 2014 (vídeo, ppt, bibtex).
- Miguel Costa, Mário J. Silva, Francisco M. Couto, Learning Temporal-Dependent Ranking Models, 37th international ACM SIGIR conference on Research and Development in Information Retrieval, Setembro 2014 (ppt, bibtex).
- Daniel Gomes and Miguel Costa, The Importance of Web Archives for Humanities, International Journal of Humanities and Arts Computing, Abril 2014 (bibtex).
- Daniel Gomes, David Cruz, João Miranda, Miguel Costa, Simão Fontes, Acquiring and providing access to historical web collections (demo), 10th International Conference on Preservation of Digital Objects, Setembro 2013 ("minute madness" slides ppt, bibtex).
- David Cruz, Daniel Gomes, Adapting search user interfaces to web archives, 10th International Conference on Preservation of Digital Objects, Setembro 2013 ("minute madness" slides ppt, poster, bibtex).
- Miguel Costa, João Miranda, David Cruz, Daniel Gomes, Query suggestion for web archive search, 10th International Conference on Preservation of Digital Objects, Setembro 2013 ("minute madness" slides, poster, bibtex).
- Miguel Costa, Daniel Gomes, Francisco M Couto, Mário J. Silva, A Survey of Web Archive Search Architectures, Temporal Web Analytics Workshop 2013, Rio de Janeiro, Brazil, Maio de 2013 (apresentação, bibtex)
- Daniel Gomes, Miguel Costa, David Cruz, João Miranda, Simão Fontes, Creating a Billion-Scale Searchable Web Archive, Rio de Janeiro, Brasil, Maio de 2013 (apresentação, bibtex)
- Daniel Gomes, David Cruz, João Miranda, Miguel Costa, Simão Fontes, Search the Past with the Portuguese Web Archive, 22nd International World Wide Web Conference, Rio de Janeiro, Brasil, Maio de 2013 (poster, apresentação, bibtex)
- Miguel Costa, Mário J. Silva, Evaluating Web Archive Search Systems, Web Information Systems Engineering 2012, Chipre, Novembro de 2012 (apresentação, bibtex).
- Daniel Gomes, João Miranda, Miguel Costa, A survey on web archiving initiatives, International Conference on Theory and Practice of Digital Libraries 2011, Berlin, Alemanha, Setembro de 2011 (apresentação, poster, bibtex, vídeo)
- Miguel Costa, Mário J. Silva, Characterizing Search Behavior in Web Archives, Temporal Web Analytics Workshop 2011, Hyderabad, Índia, Março de 2011 (apresentação, bibtex)



Closer and more relevant to researchers



Grants for Research over Arquivo.pt

- 17 videos
- 15 presentations
- 7 written communications
- 7 online channels

Transformações da presença dos jornais portugueses na web (1996-2016):

Correio da Manhã, Diário de Notícias, Expresso e Público
Relatório final de um estudo de caso do projecto "Investiga XXI"

Transformations of the Portuguese newspaper web presence (1996-2016): *Correio da Manhã, Diário de Notícias, Expresso and Público*

Final report of a case study of the project "Investiga XXI"

The image shows a YouTube channel page for 'Investiga XXI' and a mobile application interface. The YouTube page displays the channel name, a search bar, and a video player. The video title is 'Investiga XXI: investigação utilizando o Arquivo.pt'. Below the video player, there are several video thumbnails with titles like 'Investiga XXI no estúdio' and 'A FCSH de Arqu...'. The mobile app interface shows a menu with options like 'Entrada > Colabore > Investigação e Desenvolvimento > Investiga XXI > O Straight-Edge na área metropolitana de Lisboa'. The app also displays a video player with a thumbnail of a man speaking.

2018: Launched Annual Award Arquivo.pt



Photo courtesy of FPF Press Gallery

Candidaturas abertas!
Prémios Arquivo.pt

até
4 de Maio

VIAJE NO TEMPO E GANHE 10.000€

1.º Prémio: 10.000€
2.º Prémio: 3.000€
3.º Prémio: 2.000€

Crie um trabalho inovador, individual ou em grupo,
que use a informação histórica preservada
pelo **Arquivo.pt** para concorrer.



FCT Fundação
para a Ciência
e a Tecnologia
Comissão Científica Nacional
FCN

Sabe mais em: arquivo.pt/premios



ARQUIVO.PT

Call for submissions!
Arquivo.pt Prizes

until
May, 4



Pesquisar

[Pesquisa avançada](#)

Pesquise páginas do passado

Milhões de conteúdos da web preservados desde 1996

Arquivo.pt Award in the news



Antena2 / há 155 dias

Antena 2 Ciência

Prémio Arquivo.pt, primeira edição: a informação histórica p...

Shifter

GLOBAL WEB

Prémios no total de 15 mil euros para quem melhor aproveitar o Arquivo.pt

Candidaturas abertas até 4 de Maio. Redação 12 MARÇO 2018



OUVIR EMISSÃO

TSF

NOTICIÁRIOS POLÍTICA SOCIEDADE ECONOMIA DESPORTO INTERNACIONAL CULTURA VÍDEOS OPINIÃO HUMOR

PROGRAMAS

Ambiente Ciência e Tecnologia Educação Saúde Justiça Segurança

CONSULTE CONDIÇÕES

ARQUIVO.PT

Se existisse um museu da internet, Daniel Gomes seria arqueólogo

17 DE ABRIL DE 2018 - 16:33

O Arquivo.pt guarda informação que está disponível na Internet para que possa ser consultada no futuro.

FOTOGALERIA DO DIA



New Applications born out of Arquivo.pt

Conta-me Histórias (Tell me Stories)

an online tool that allows users to automatically generate timeline summarizations of news articles maintained by Arquivo.pt.

<https://goo.gl/dNhp1W>



Explore examples

Economy and Society

Crise Oriente Médio

Política Internacional

Crise Coreia do Norte

Política Internacional

Crise Crimeia

Política Internacional

II Guerra Mundial

História

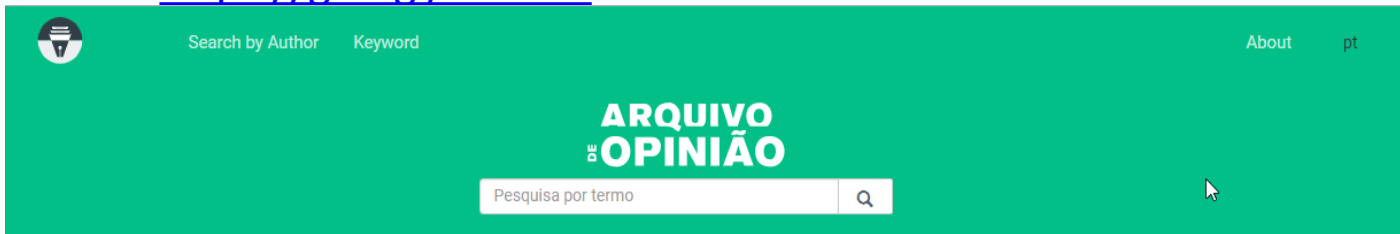


New Applications born out of Arquivo.pt

Arquivo de Opinião

A web application that provides a digital archive of opinion articles published in the main newspapers in Portugal. It offers a search engine that allows searching by word or exact phrase, as well as filter by author, date or newspaper.

<https://goo.gl/PHucTF>



Grande ministro

Não havia memória de nada muito marcante na passagem de Alberto Martins pelo Ministério da Justiça. Tirando a história mal contada de ter autorizado o pagamento de um subsídio especial de dezenas de milhares de euros à mulher, procuradora – e arriscava-se a não ter direito a uma simples citação.

...

Written by Manuel Catarino, Dec. 8, 2011, midnight



A teia de Sócrates

Não será de estranhar que todos aqueles que tenham beneficiado com a estrutura do poder de Sócrates respondam à sua chamada.

Written by Paulo Morais, Jan. 24, 2015, midnight





Results

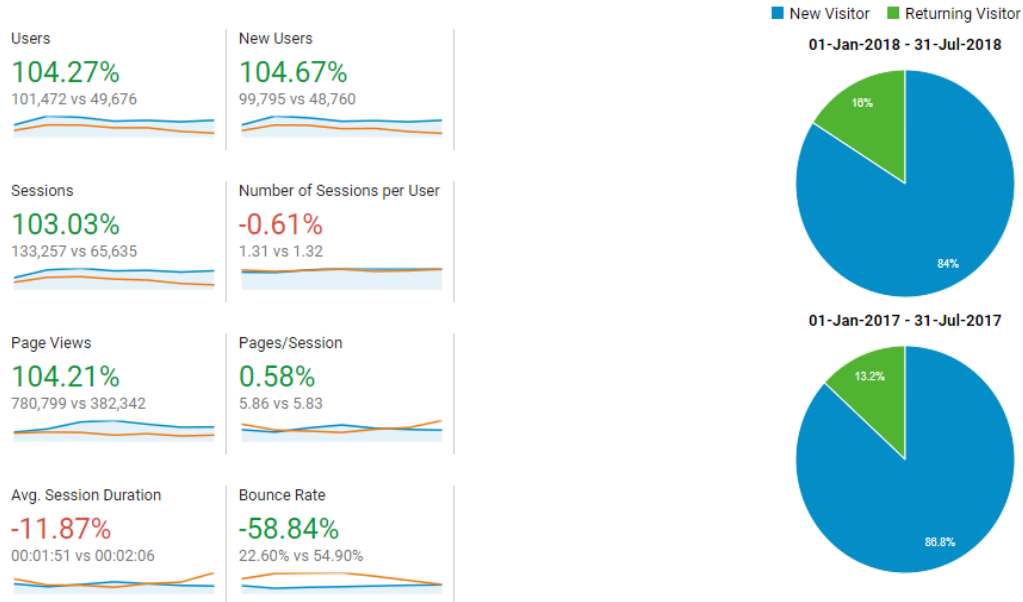


Fig.2 –Users and new users. Lower bounce rate. Source: Google Analytics

Bounce Rate decreases 59% comparing 2017 and 2018.
Users access the website deliberately and stay longer to explore the service.

Results

Overview

Users VS Select a metric

Hourly Day Week **Month**

01-Jan-2018 - 31-Jul-2018: ● Users

01-Jan-2017 - 31-Jul-2017: ● Users

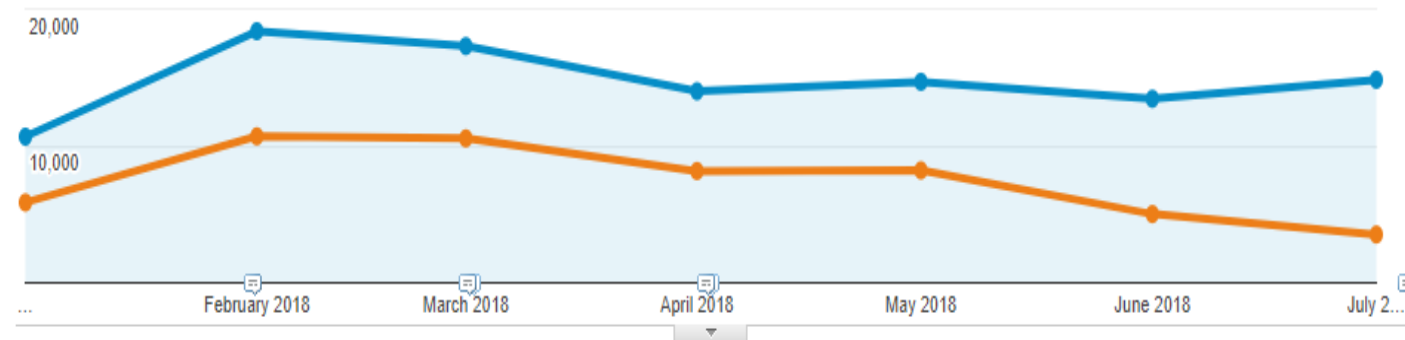


Fig. Comparison of Arquivo.pt users for the first 6 months of 2017 and 2018. Source: Google Analytics

104% Increase in website users overall comparing 2017 and 2018



ARQUIVO.PT