

Research the Past Web using Web archives

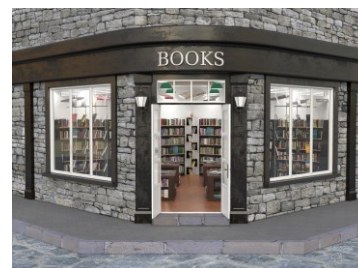
Daniel.Gomes@fccn.pt

80%

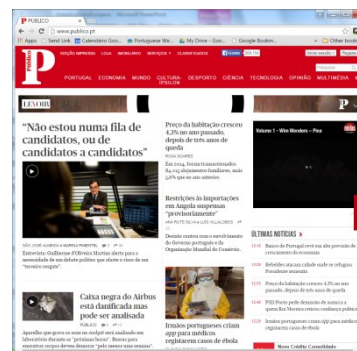
of web pages
change their content or disappear
within 1 year

Archives are also require to preserve information published online

Preserve the Press



Preserve the Web

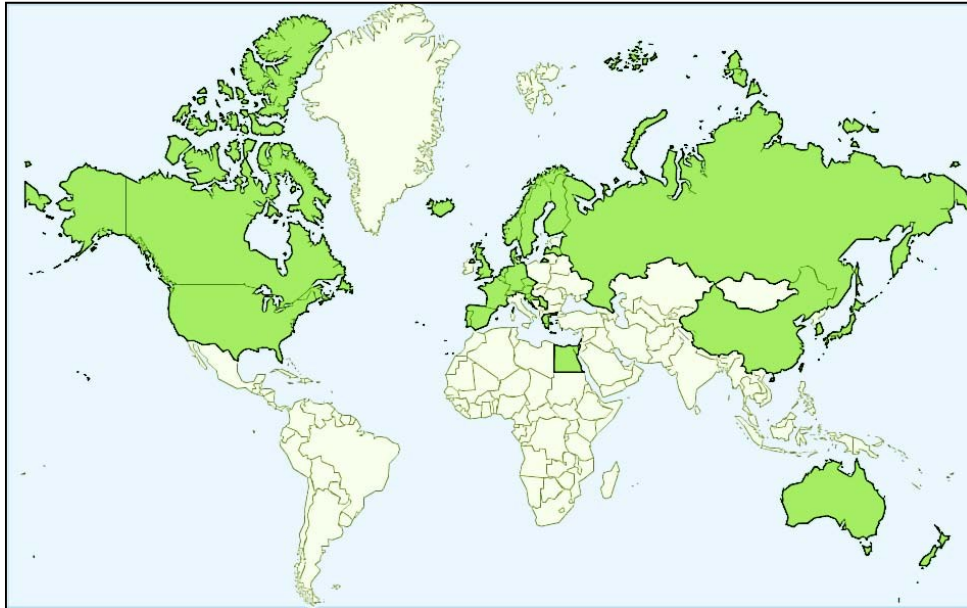


[Archive.org](https://archive.org): the first web archive




Brewster Kahle founded the **Internet Archive** in 1996, a non-profit organization to provide “Universal access to all knowledge”.

93 web archiving initiatives around the World (2020)



*List of Web archiving initiatives,
Wikipedia*



INTERNATIONAL
INTERNET
PRESERVATION
CONSORTIUM

Menu

PROJECTS > COLLABORATIVE COLLECTIONS

Collaborative collections

IIPC members are collaborating to build public web archive collections based on themes or events that align with the following criteria:

- of high interest to IIPC members
- broader than any one member's responsibility or mandate
- of higher value to research due to broader perspective provided by multiple institutions
- transnational in scope.

The project is administered by the IIPC [Content Development Group](#). More information about the project is available on their page.

Collections

- 🔗 [Novel Coronavirus \(Covid-19\) \(2020\)](#)
- 🔗 [Climate change \(2019\)](#)
- 🔗 [Online News Around the World \(2018\)](#)
- 🔗 [Olympics/Paralympics collections \(2010-2018\)](#)
- 🔗 [European Refugee Crisis \(2015-2016\)](#)
- 🔗 [World War One Commemoration \(2015-\)](#)
- 🔗 [Intergovernmental Organizations \(2015-\)](#)

56 members, netpreserve.org
Interesting thematic collections

I work at the **Arquivo.pt web archive**

Research infrastructure provided by Foundation for Science and Technology (Portugal)

Preserves **publicly accessible** information related with:





















- Portugal
- **Research** and **Education** (international)

Free online services to research the Past Web



Arquivo.pt is used word-wide



Country	Users	% Users
1.  Portugal	87,532	 47.69%
2.  United States	34,784	 18.95%
3.  Brazil	6,991	 3.81%
4.  Japan	5,306	 2.89%
5.  United Kingdom	4,996	 2.72%
6.  Russia	3,227	 1.76%
7.  Germany	2,649	 1.44%
8.  Canada	2,475	 1.35%
9.  Spain	2,385	 1.30%
10.  India	2,086	 1.14%

52% of our users were international in 2019

User Interfaces and documentation also in English

Combined with Google Translate enables **cross-lingual access** to preserved content

The Past Web

Examples of international web-archived pages
and collections

Historical web pages: Tim Berners Lee personal page and W3C in the 90s

Tim Berners-Lee

Bio

Tim Berners-Lee is the inventor of the World Wide Web, an internet-based hypermedia initiative for global information sharing. He wrote the first Web clients and server and defined the URL, HTTP and HTML specifications on which the web depends while working at [CERN](#), the European Particle Physics Laboratory, in late 1990.

Tim is now the Director of the [W3 Consortium](#), an open forum of companies with a common mission to realize the full potential of the Web. He is a Principal Research Scientist in Computer Science ([LCS](#)) at the Massachusetts Institute of Technology ([MIT](#)).

Before going to CERN, Tim was a founding director of Image Computer Systems, a consultant in hardware and software system design, real-time communications, and image processing; and a principal engineer with Plessey Telecommunications, in the UK, and a professor of Oxford University.

Tim is married to Nancy Carlson. They have two children, born 1991 and 1993.

www.w3.org/pub/WWW/People/Berners-Lee/, 1996
@Arquivo.pt



www.w3.org, 1996
@Arquivo.pt



The World Wide Web Consortium

The World Wide Web is the universe of network-accessible information. The [World Wide Web Consortium](#) exists to realize the full potential of the Web.

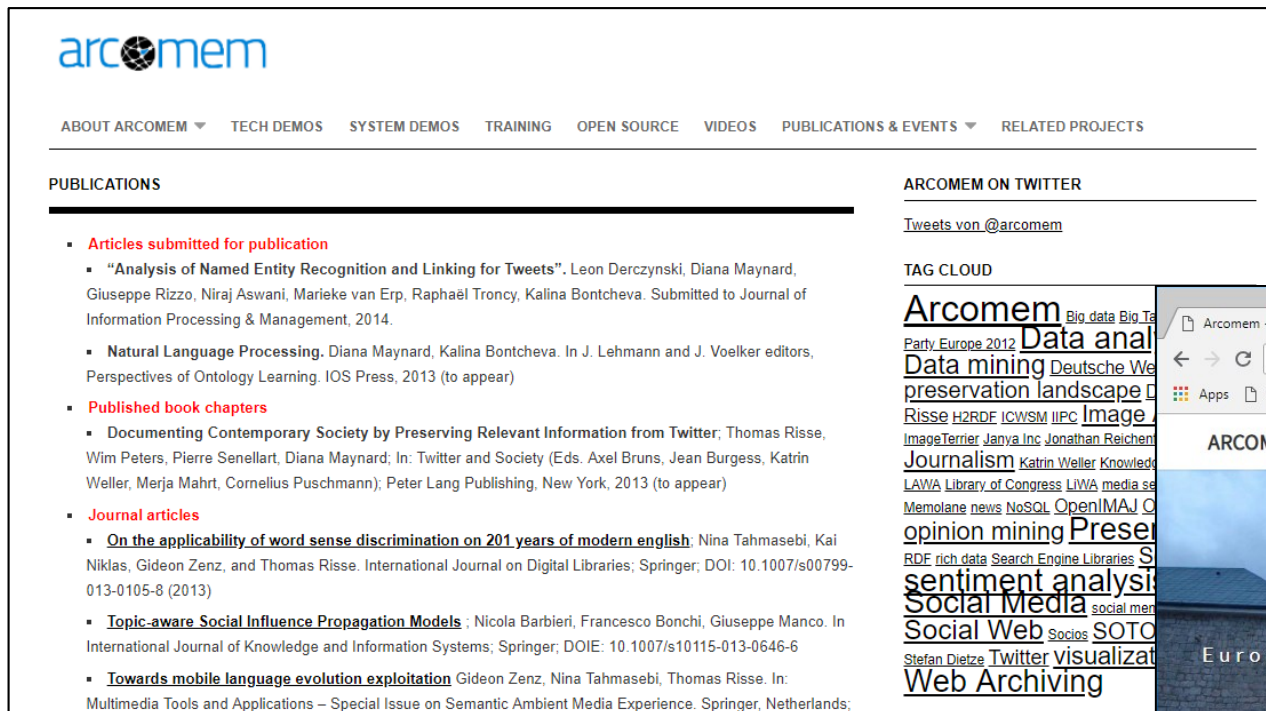
W3C works with the global community to produce [specifications](#) and [reference software](#). W3C is funded by industrial [members](#) but its products are freely available to all. The Consortium is run by [MIT LCS](#) and by [INRIA](#), in collaboration with [CERN](#) where the web originated. Please see the [list of members](#) to learn about individual members and visit their World Wide Web sites.

- [W3C Activity areas and directions](#)
- [How to contact W3C](#)
- [Frequently Asked Questions about W3C](#)
- [W3C meetings, newsletter, mailing lists \[W3C Members only\]](#)
- [Help](#)

In this document:

- [News and Updates](#)
- [Web Specifications and Development Areas](#)
- [W3C Software](#)
- [The World Wide Web and the Web Community](#)
- [Getting involved with the W3C](#)

Websites that document R&D projects funded by the European Union



arcomem

ABOUT ARCOMEM ▼ TECH DEMOS SYSTEM DEMOS TRAINING OPEN SOURCE VIDEOS PUBLICATIONS & EVENTS ▼ RELATED PROJECTS

PUBLICATIONS

- Articles submitted for publication
 - "Analysis of Named Entity Recognition and Linking for Tweets". Leon Derczynski, Diana Maynard, Giuseppe Rizzo, Niraj Aswani, Marieke van Erp, Raphaël Troncy, Kalina Bontcheva. Submitted to Journal of Information Processing & Management, 2014.
 - Natural Language Processing. Diana Maynard, Kalina Bontcheva. In J. Lehmann and J. Voelker editors, Perspectives of Ontology Learning. IOS Press, 2013 (to appear)
- Published book chapters
 - Documenting Contemporary Society by Preserving Relevant Information from Twitter; Thomas Risse, Wim Peters, Pierre Senellart, Diana Maynard; In: Twitter and Society (Eds. Axel Bruns, Jean Burgess, Katrin Weller, Merja Mahrt, Cornelius Puschmann); Peter Lang Publishing, New York, 2013 (to appear)
- Journal articles
 - On the applicability of word sense discrimination on 201 years of modern english; Nina Tahmasebi, Kai Niklas, Gideon Zenz, and Thomas Risse. International Journal on Digital Libraries; Springer; DOI: 10.1007/s00799-013-0105-8 (2013)
 - Topic-aware Social Influence Propagation Models; Nicola Barbieri, Francesco Bonchi, Giuseppe Manco. In International Journal of Knowledge and Information Systems; Springer; DOI: 10.1007/s10115-013-0646-6
 - Towards mobile language evolution exploitation Gideon Zenz, Nina Tahmasebi, Thomas Risse. In: Multimedia Tools and Applications – Special Issue on Semantic Ambient Media Experience. Springer, Netherlands;

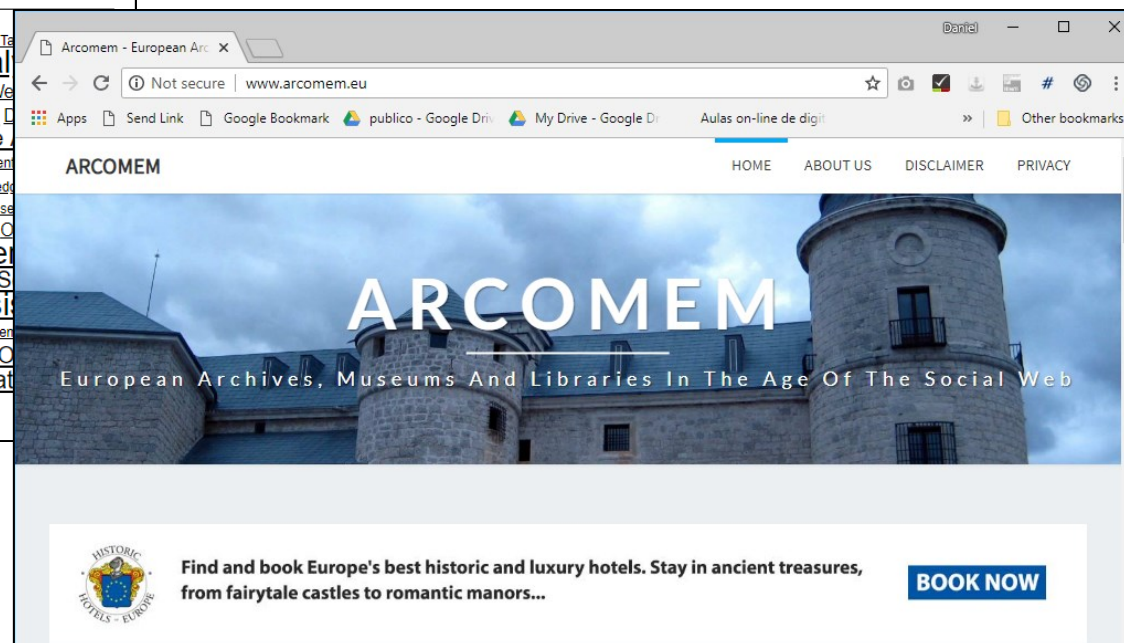
ARCOMEM ON TWITTER

Tweets von @arcomem

TAG CLOUD

Arcomem Big data Big Te Party Europe 2012 Data anal Data mining Deutsche We preservation landscape D Risse H2RDF ICWSM IIPC Image ImageTerrier Janya Inc Jonathan Reichen Journalism Katrin Weller Knowledge LWA Library of Congress LIWA media se Memolane news NoSQL OpenIMAJ O opinion mining Preser RDF rich data Search Engine Libraries S sentiment analysis Social Media social me Social Web Socios SOTO Stefan Dietze Twitter visualizat Web Archiving

arcomem.eu, 2014
@Internet Archive



Arcomem - European Arc X

Not secure | www.arcomem.eu

ARCOMEM

HOME ABOUT US DISCLAIMER PRIVACY

ARCOMEM

European Archives, Museums And Libraries In The Age Of The Social Web

HISTORIC HOTELS - EUROPE

Find and book Europe's best historic and luxury hotels. Stay in ancient treasures, from fairytale castles to romantic manors...

BOOK NOW

arcomem.eu, 2018

Cross-lingual collaborative web collections (e.g. European Elections 2019)

Keywords - Post-European Elections 2019 - Palavras-chave Pós-Eleições Europeias 2019

File Edit View Insert Format Data Tools Add-ons Help All changes saved in Drive

100% £ % .0 .00 123 Arial 10 B I S A [Grid] [List] [Table] [Chart] [Filter] [Sum]

	A	B	C	D
1	Sheet description			
2	Given the set of expressions in Portuguese - the same we tried in the Crawl 1 before elections - we translated it to all other 23 EU languages, using the GOOGLETRANSLATE formula on google sheets. Example: GOOGLETRANSLATE(A7, "pt", "bg"). But we only use 12 expressions. So, we've used 288 expressions in 24 languages. Using Bing Web API, we collected 10 top results, and we got about 2880 seeds. Then we've ran the Crawl 2 . Results were good. So we did the same again using all research the expressions on Bing and we've got more than 9000 seeds. So we've done the Crawl 3 .			
3	Language codes reference: http://publications.europa.eu/code/en/en-5000800.htm			
4				
5				
6	PT - Portuguese	BG - Bulgarian	HR - Croatian	CS - Czech
7	eleições parlamento europeu 2019	изборите за Европейски парламент 2019	izbori za Europski parlament 2019. godine	Volby do Evropského parlamentu 2019
8	eleições europeias 2019	Европейски избори 2019	Europski izbori 2019	Volby do Evropského parlamentu 2019
9	abstenção europeias 2019	Европейската въздържал се през 2019 г.	Europska apstinencija u 2019	Evropská zdržení v roce 2019
10	resultados eleitorais 2019	результатите от изборите 2019	Rezultati izbora 2019	Výsledky voleb 2019
11	vencedores europeias 2019	Европейските победители 2019	Europski pobjednici 2019	Evropští vítězové 2019
12	brexit eleições europeias 2019	brexit европейските избори 2019	brexit europski izbori 2019.	Volby brexit evropští 2019

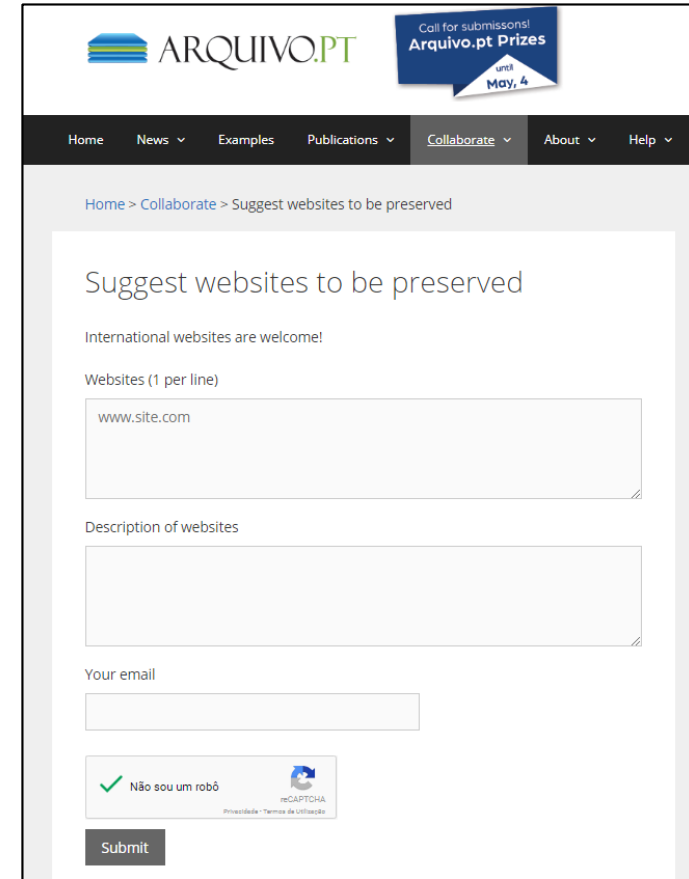
Suggest **any website** before it vanishes and you miss it!

Any domain

Any language

Broad interest for research and
education

Save your research outputs!



The screenshot shows the 'Suggest websites to be preserved' form on the Arquivo.pt website. The header includes the Arquivo.pt logo and a banner for 'Arquivo.pt Prizes' ending on May 4. The navigation bar has links for Home, News, Examples, Publications, Collaborate (selected), About, and Help. The breadcrumb trail is 'Home > Collaborate > Suggest websites to be preserved'. The form itself has a title 'Suggest websites to be preserved' and a subtitle 'International websites are welcome!'. It contains three main input fields: 'Websites (1 per line)' with the placeholder 'www.site.com', 'Description of websites', and 'Your email'. At the bottom, there is a reCAPTCHA widget with the text 'Não sou um robô' and a 'Submit' button.

arquivo.pt/suggest

Use cases for a web archive

Fix citations to broken links that jeopardize Science in all areas

Data sets and tools applied in experiments are online.

When cited resources disappear, experiments cannot be reproduced

- Not Reproducible,
Not Science.

Web archives preserve online scientific information.

Malaysian Journal of Library & Information Science, Vol. 16, no. 3, December 2011: 17-29

Death of web citations: a serious alarm for authors

The FASEB Journal

The Journal of the Federation of American Societies for Experimental Biology

HOME | CURRENT ISSUE | NEW ARTICLES | ARCHIVE | ALERTS | RSS | SUBMIT | HELP

Unavailability of online supplementary scientific information from articles published in major journals

« Previous | Next Article »
Table of Contents

This Article

doi:
10.1096/fj.05-4784lsf
December 2005
The FASEB Journal
vol. 19 no. 14 1943-1944

» Abstract
Full Text
Full Text (PDF)

+ Classifications
+ Services

JOURNAL OF THE ASSOCIATION FOR INFORMATION SCIENCE AND TECHNOLOGY

Research Article

Link decay in leading information science journals

Dion Hoe-Lian Goh and Peng Kin Ng

Article first published online: 3 NOV 2006

DOI: 10.1002/asi.20513

Copyright © 2007 Wiley Periodicals, Inc., A Wiley Company

Persistence and Decay of Web Citations Used in Theses and Dissertations Available at the Sokoine National Agricultural Library, Tanzania

ACADEMIC JOURNAL ARTICLE

By Sife, Alfred S.; Bernard, Ronald

International Journal of Education and Development using Information and Communication Technology, Vol. 9, No. 2, July 1, 2013

Management: documentation/audit of finished project



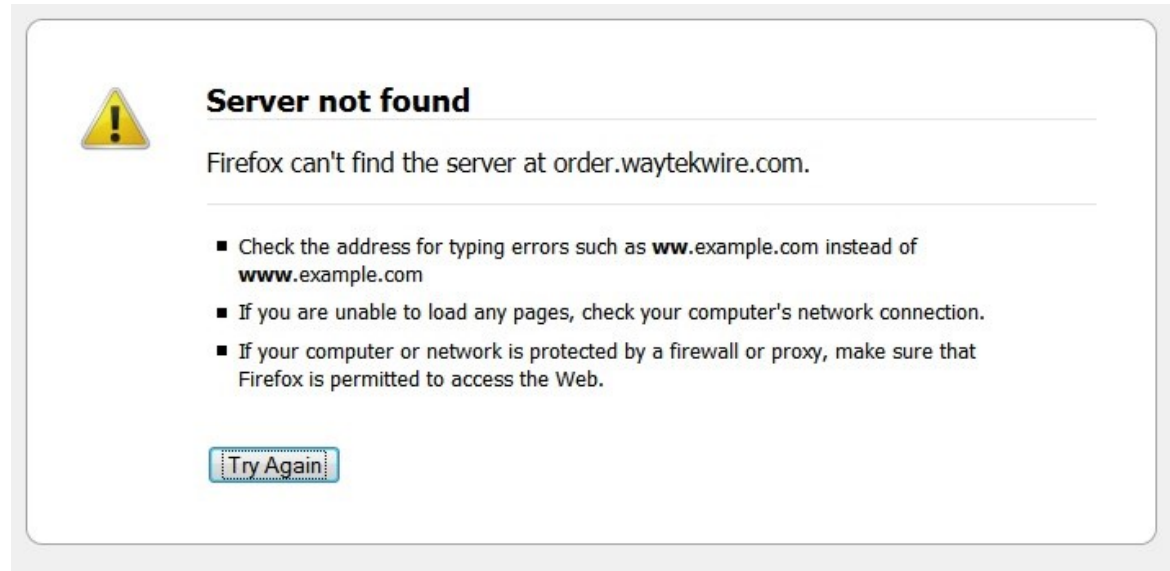
“Everything was on the website”

Linguateca Structure Team Overview Access to resources Evaluation contests Catalogue of resources Catalogue of tools Catalogue of actors Catalogue of publications <ul style="list-style-type: none">Publication searchPublication additionSUPeRB Interesting information FAQ Gramateca Português Norsk  Comments and suggestions <small>Please use the navigation options on the left handside frame (clicking on the links on the right handside does not update the left handside).</small>	Publications on the computational processing of portuguese Linguateca Our goal is to gather a collection of bibliographic references on the computational Processing of portuguese in this pages. We appreciate any contribution that you as an author or reader can give us, by providing us with additional information or to direct us to pages with relevant content. To look for publications we suggest the use of our search interface . To contribute with new references you can use this form . To use SUPeRB as your publications manager see here . We currently list 2375 references under the following categories: <ul style="list-style-type: none">• Publications in journals (184 ordered alphabetically, by date; In BibTeX: orderer alfaphabetically, by date)• Books (1 publications, orderer alphabetically, by date; In BibTeX: orderer alphabetically, by date)• Book chapters (189 publications, orderer alphabetically, by date; In BibTeX: orderer alphabetically, by date)• Book (Author) (57 publications, orderer alphabetically, by date; In BibTeX: orderer alphabetically, by date)• Book (Compilation) (53 publications, orderer alphabetically, by date; In BibTeX: orderer alphabetically, by date)• teshis (152 publications, orderer alphabetically, by date; In BibTeX: orderer alphabetically, by date)• Articles in international conferences (1134 publications, orderer alphabetically, by date; In BibTeX: orderer alphabetically, by date)• Reports and monographies (192 publications, orderer alphabetically, by date; In BibTeX: orderer alphabetically, by date)• Other academical work (3 publications, orderer alphabetically, by date; In BibTeX: orderer alphabetically, by date)• Publications only available on the web (95 publications, orderer alphabetically, by date; In BibTeX: orderer alphabetically, by date)• Presentations (297 publications, orderer alphabetically, by date; In BibTeX: orderer alphabetically, by date) Publications inserted last month . Last update 30 November 2015. Comments e suggestions
---	--

Maintain Portfolio/CV

“My best work was a website that no longer exists...”

60% of the websites disappear just after 2 years.



Search the Past Web

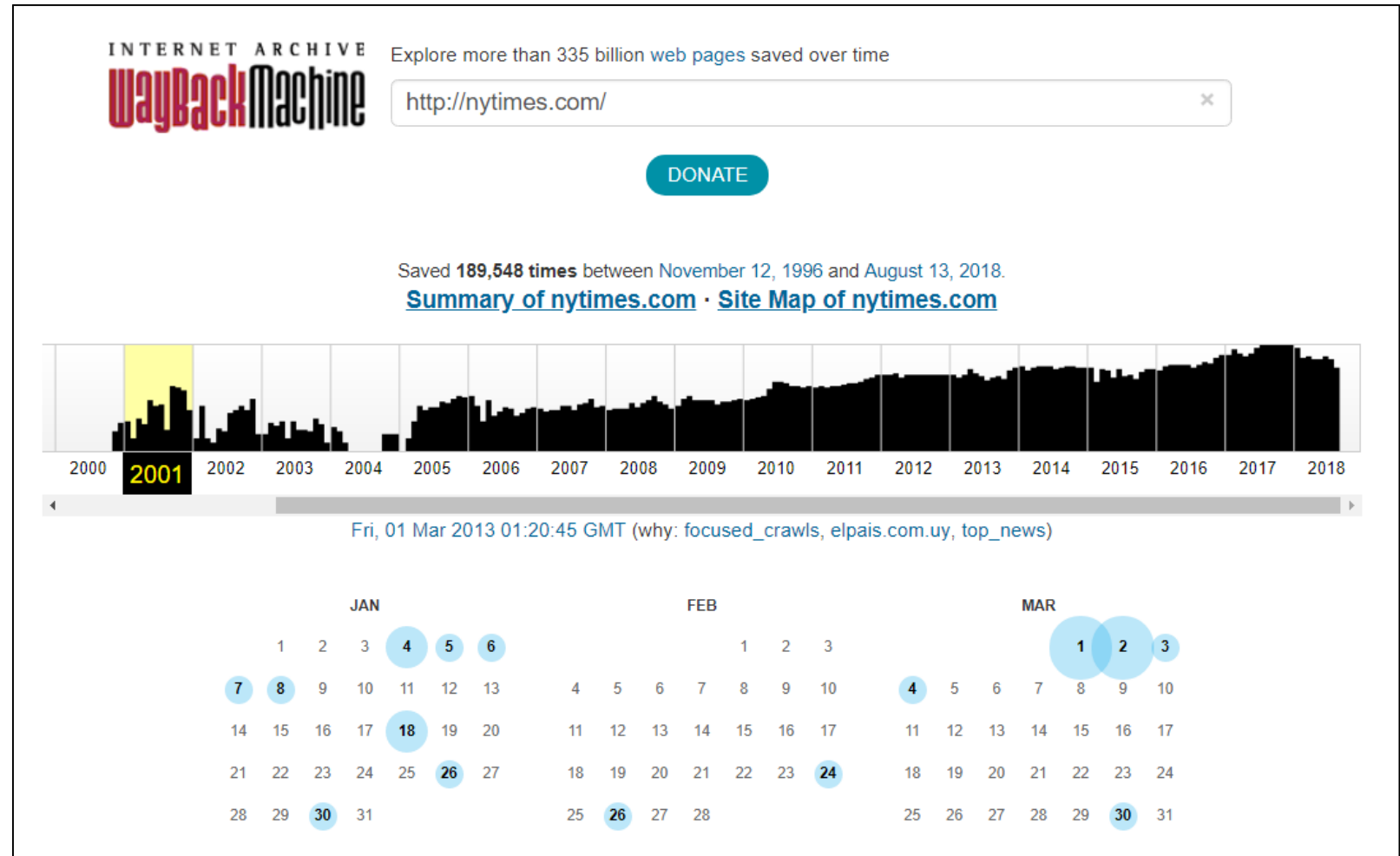
Anyone can do it.

URL Search:

Internet Archive Wayback Machine

[Archive.org](https://archive.org)

World wide content
URL search



New York Times news on February 14th, 2009

The New York Times

Saturday, February 14, 2009 Last Update: 1:11 AM ET

Try Our
EXTRA
Home Page

 Get Home Delivery | Personalize Your Weather

JOB
REAL ESTATE
AUTOS
ALL CLASSIFIEDS

WORLD
U.S.
POLITICS
N.Y./REGION
BUSINESS
TECHNOLOGY
SPORTS
SCIENCE
HEALTH
OPINION
ARTS
Books
Movies
Music

Stimulus Plan Sets New Limits on Executive Pay

By EDMUND L. ANDREWS and ERIC DASH

A provision buried in the stimulus bill passed by Congress would impose restrictions on executive bonuses at companies reaping bailout money.

Stimulus Plan Receives Final Approval in Congress



Stephen Farrell/The New York Times

Iraqi Militants Become Patrons of the Arts

By STEVEN LEE MYERS

OPINION »

TALK SHOW
When Updike and Cheever Came to Visit
Dick Cavett recalls the rare thrill of having two great writers as guests on the same show.



- Collins: Stock a Cabinet
- Comments
- Editorial: California Prisons
- Op-Ed: No Price on Love
- Op-Ed: Baby Bailout
- Op-Ed: Stagflation
- Opinionator: Renditions

MOVIES »

Lesser-Known Mobsters
“Gomorrah” takes an unsentimental look at Italian organized crime.

BUSINESS »

A Stress Test for the Latest Bailout Plan
The stress test finds out how crippled big banks would be

New York Times on 9/11, 2001

The New York Times on the Web

UPDATED TUESDAY, SEPTEMBER 11, 2001 4:43 PM ET | [Personalize Your Weather](#)

Search Past 30 Days [Go to Advanced Search](#) [Sign Up](#) [Log In](#)

[World Trade Center Toppled in Attack](#)

[Thousands Feared Dead in New York and Washington](#)
By JAMES BARRON
In what appeared to be coordinated attacks, hijacked jetliners crashed into both towers of the World Trade Center and the Pentagon Tuesday morning. [Go to Article](#)
☐ [Slide Show: Destruction in New York](#)

[In Washington, an Attack on a Symbol of American Power](#)
By DAVID STOUT
Shortly after the attacks in New York, another aircraft crashed into the Pentagon, the massive building that is the very symbol of the American defense establishment.
☐ [Throughout U.S., Shock After Attacks](#)
☐ [Bush Vows to Hunt Down Terrorists](#)



TECHNOLOGY

[Dell on Top](#)

Everyone is feeling the slump, but Michael Dell says his company can tolerate it better than most.



Q&A

[Ottawa, Canada](#)

susz99 asks: Does anyone have any suggestions for things to do in Ottawa for a group of about 20?
[answer the question](#)

☐ [More News From Reuters](#)
☐ [More News From AP](#)

New York Times on the early days of the Web, 1996

Welcome! First-time users [register here](#) or [take a QuickTour](#) of the site.

NEW
AUTO
CLASSIFIEDS

The New York Times

ON THE WEB

knicks NEW!
celtics
one on one
newyork.boston.com

"All the News That's Fit to Print" Tuesday, November 12, 1996

SECTIONS

- Front Page • CyberTimes
- Politics • Business
- Editorials/Letters • Op-Ed
- Arts & Leisure • Travel
- Real Estate • Job Market
- Diversions • Web Specials

NEWS BY CATEGORY

CLASSIFIEDS | **FORUMS**

SERVICES | **SEARCH**

TABLE OF CONTENTS



Crowds Looking for Food in Zaire

LATE NEWS UPDATE

Two Aircraft Collide In Air Near Delhi, Killing About 350

IN CYBERTIMES

Europe Betting on Self-Regulation to Control the Internet


[Help Center](#) | [Low Graphics](#)

PLEASE OPEN YOUR WINDOW TO THE WIDTH OF THIS LINE OF TEXT.

[Copyright 1996 The New York Times](#)

URL search across **several web archives**

Memento Time Travel Portal
TimeTravel.mementoweb.org

Time TravelAbout API Privacy Terms

FindReconstruct◀11 Sep 2011 16:13:07 GMT▶

Mementos closest to the requested date 11 Sep 2011 15:48:01 GMT

[Internet Archive Memento, 1 hour after](#) Embed
<https://web.archive.org/web/20110911161307/http://www.nytimes.com/>
11 Sep 2011 16:13:07 GMT [<1 hour from requested date]

Previous Memento data not provided	Next Memento data not provided
First Memento data not provided	Last Memento data not provided



[Portuguese Archive Memento, 1 hour before](#) Embed
<http://arquivo.pt/wayback/20110911151946/http://www.nytimes.com/>
11 Sep 2011 15:19:46 GMT [<1 hour from requested date]

Previous Memento data not provided	Next Memento data not provided
First Memento data not provided	Last Memento data not provided

[Bibliotheca Alexandrina Web Archive Memento, 3 hours after](#) Embed
<http://web.archive.bibalex.org:80/web/20110911191236/http://nytimes.com>
11 Sep 2011 19:12:36 GMT [+3 hours from requested date]


Previous Memento 11 Sep 2011 07:12:26 GMT [-8 hours]	Next Memento 12 Sep 2011 07:12:37 GMT [+15 hours]
--	---

experience web time travel



install Memento for Chrome


enable web time travel



MEDIAWIKI

install Memento for MediaWiki

say **no** to "404 Not Found"



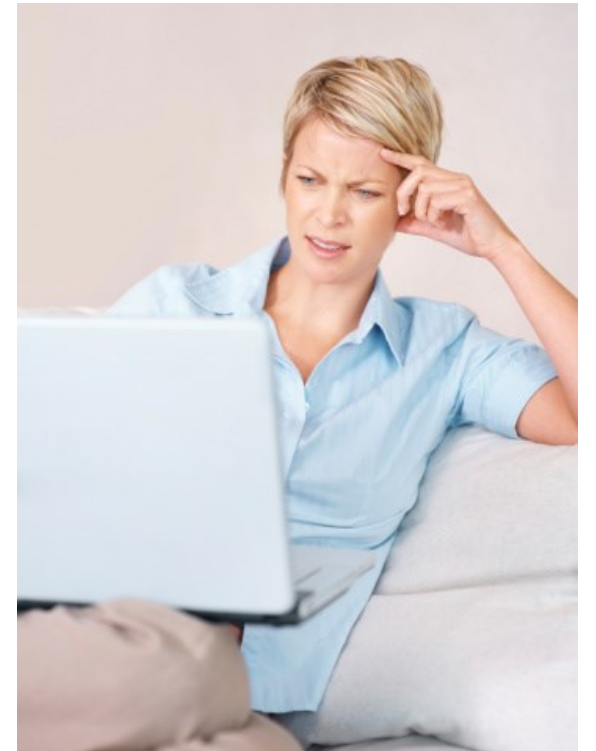
Robust Links

use links that refuse to die

Limitation of URL search

Users **do not know** the **exact URL** of the **web page of the Past** that contained the information that they search for.

Search for: Michael Jackson concert before 1998.
Which **was** the address of the web page?



Keyword search related to home pages of websites

[Archive.org](https://archive.org)

Keywords from the links to the homepages of websites.

Search results **do not include:**

- Home pages that do not receive links
- Internal webpages (not home pages)
- Cannot define time span for search

The screenshot shows the WayBack Machine search interface. At the top, the logo 'INTERNET ARCHIVE WayBack Machine' is displayed next to the tagline 'Explore more than 335 billion web pages saved over time'. A search bar contains the text 'michael jackson' and is highlighted with a red rectangle. Below the search bar is a 'DONATE' button. The search results are listed in a table-like format with four rows, each showing a website thumbnail, the URL, the search term, the number of captures, and the time span.

Thumbnail	URL	Search Term	Captures	Time Span
	www.aboutmichaeljackson.com	<i>michael jackson</i> news	29,655 captures, 79 images, 0 videos	49,561 web captures from 2003 to 2013
	www.michaeljackson.com	<i>michael jackson</i>	501,388 captures, 7,717 images, 78 videos	752,022 web captures from 2001 to 2017
	michaeljacksonconspiracy.blogspot.com	the <i>michael jackson</i> conspiracy	48 captures, 0 images, 0 videos	125 web captures from 2007 to 2017
	www.malibufanclub.de	malibu <i>michael jackson</i> family fan club	1,847 captures, 1,828 images, 1 video	8,273 web captures from 2002 to 2017

Full-text search over all pages with Arquivo.pt

ARQUIVO.PT

michael jackson concert

Search

Advanced search

between: 01/01/1996 and: 31/12/1998

December 1998

Su Mo Tu We Th Fr Sa

1 2 3 4 5

6 7 8 9 10 11 12

13 14 15 16 17 18 19

20 21 22 23 24 25 26

27 28 29 30 31

OK Cancel

from 24

60

Oct 24

Untitled

3 February, 1998 - List versions

Untitled ASSUNTO NÚMERO DIST DATA/HORA DE CHEGADA STATUS LINHAS

RTUovru Thu Oct 24 16:09 1996 WIRE 3 256 Malaysia's Mahathir defends

(Reuter) - Prime Minister Mahathir Mohamad said on Thursday he had all

<http://rorty.ist.utl.pt/news/malasia2.html>

Not site home page
“Untitled”

There was a controversial
Michael Jackson concert in
Malaysia in 1996.

ARQUIVO.PT

ASSUNTO NÚMERO DIST DATA/HORA DE CHEGADA STATUS LINHAS

BC-Malaysia-Jackson REU1960 RTUovru Thu Oct 24 16:09 1996 WIRE 3 256

Malaysia's Mahathir defends Jackson concert

KUALA LUMPUR, Oct 24 (Reuter) - Prime Minister Mahathir Mohamad said on Thursday he had allowed a Michael Jackson concert in Kuala Lumpur to give Malaysians the chance to watch a "star in action", the national Bernama news agency reported.

Advanced search:
did Michael Jackson ever
came to Portugal?



Advanced Search

Refine the details of your search using the options bellow.

Search pages by:

Search

Words

With these words:

Portugal

ex.: group draw

With this phrase:

michael jackson

ex.: euro 2004

Without any of these words:

albums

ex.: rugby

Date

Between

01/01/1996



and

01/12/2017



Sort by:

▼

- All formats
- Adobe PDF (.pdf)
- Adobe PostScript (.ps)
- HTML (.htm, .html)
- Microsoft Excel (.xls)
- Microsoft PowerPoint (.ppt)
- Microsoft Word (.doc)
- Rich Text Format (.rtf)

▼

Format

Show the pages in the format:

All formats

Results in Portuguese can be automatically translated to answer our question.

The image shows a web browser window with a search engine interface. The search bar contains the text "Portugal 'michael jackson' -albums". Below the search bar, there are filters for "between: 01/01/1996" and "and: 01/12/2017". The search results are displayed in Portuguese. The first result is "Revista TIME Michael Jackson - Abitureiras (Santarém) - Livr..." dated 10 April, 2015. The second result is "Michael Jackson – Wikipédia, a enciclopédia livre" dated 11 April, 2015. A context menu is open over the second result, showing options like "Back", "Forward", "Reload", "Save as...", "Print...", "Cast...", "Translate to English", "Show Anchors", "View page source", "View frame source", "Reload frame", and "Inspect". The "Translate to English" option is highlighted. The browser's address bar shows the URL "arquivo.pt/wayback/20150411005521/http://pt.wikipedia.org/wiki/Michael_Jackson". The browser's search bar contains the text "portugal". The browser's toolbar shows various icons for navigation and social media. The browser's status bar shows the text "Daniel".

Portugal "michael jackson" -albums

between: 01/01/1996 and: 01/12/2017

Revista TIME Michael Jackson - Abitureiras (Santarém) - Livr...
10 April, 2015 - List versions
Rochas Ornamentais Portuguesas - 3 Volumes - Fotos: Fotos de Revista TIME Michael Jackson - Abitureiras > Revista TIME Michael Jackson Classificados grátis - Portugal Copyright © 2015 TIME Michael Jackson - Abitureiras (Santarém) - Livros - Revistas Revista TIME Michael Jackson - Abitureiras (Santarém) - Livros - Revistas
<http://abitureiras.olx.pt/revista-time-michael-jackson-iiid-460216091>

Michael Jackson – Wikipédia, a enciclopédia livre
11 April, 2015 - List versions
Michael Jackson – Wikipédia, a enciclopédia livre Michael Jackson Origem: Wikipédia, a enciclopédia livre. Nota: Para outros significados, veja Michael Jackson ... , livros e artigos sobre Michael Jackson durante performance ao vivo de Wanna Be ...
http://pt.wikipedia.org/wiki/Michael_Jackson

originals de Michael foram remixadas por Justin Timberlake e Kevin Antunes (produtor de Rihanna) ao longo de um trabalho de anos no estúdio, sendo liberados pela Epic Records em conjunto com o espólio de Michael Jackson. Apesar de mais de 60 canções serem usadas para a The Immortal World Tour, a versão simples contendo 15 faixas e versão Deluxe com do

Em Portugal

Michael Jackson actuou apenas uma vez em Portugal.

O espectáculo foi a 26 de Setembro de 1992 no Estádio João

duas vezes, dizendo "I love you" ("amo vocês") a começar e

No Brasil

Michael esteve três vezes no Brasil²⁶.

A primeira vez em que Jackson veio ao Brasil foi em setem

grupo Jackson 5, durante uma turnê pela América Latina. N

Paulo, Brasília, Porto Alegre, Salvador, Rio de Janeiro e Belo Horizonte", além de se apresentar para a Rede Tupi.

Did Michael Jackson ever came to Portugal?

Yes, in 1992.


The screenshot shows a web browser window with the title "Michael Jackson - Wikipé". The address bar displays "arquivo.pt/wayback/20150411005521/http://pt.wikipedia.org/wiki/Michael_Jack...". A "Translated" pop-up window is visible over the article content. The article text includes: "by Justin Timberlake and Kevin Antunes (producer of Rihanna) over a years work in the studio, being released by Epic Records in conjunction with Michael's estate Jackson. Although more than 60 songs were used for The Immortal World Tour, only 22 songs were chosen. The CD will have two versions the simple version containing 15 tracks and Deluxe version with two CD's containing 22 tracks."

In Portugal

Michael Jackson performed only once in [Portugal](#) .



The show was the [26 of September](#) of [1992](#) at the [Estádio José Alvalade](#) in the capital [Lisbon](#) . The singer only spoke to the audience twice, saying "I love you" ("*I love you*") the start and "peace" ("*Peace*") in farewell. ³³

Search **images** from the past




× Search


[Advanced image search](#)

between:  and: 


Pages Images Options




www.netsurf.pt/si...
7 June, 2010




www.londontown.co...
7 September, 2014




www.nzhistory.net...
13 August, 2015




www.nzhistory.net...
13 August, 2015



www.nzhistory.net...
13 August, 2015

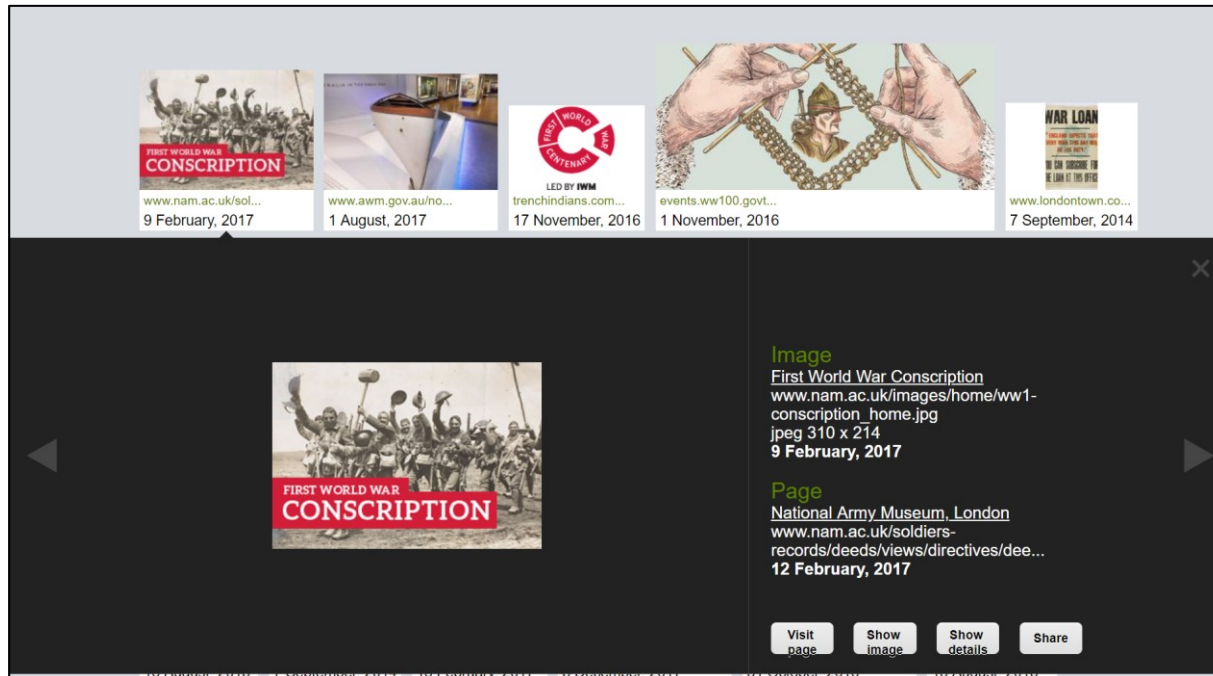


www.nzhistory.net...
13 August, 2015



www.nzhistory.net...
13 August, 2015

Image source and details



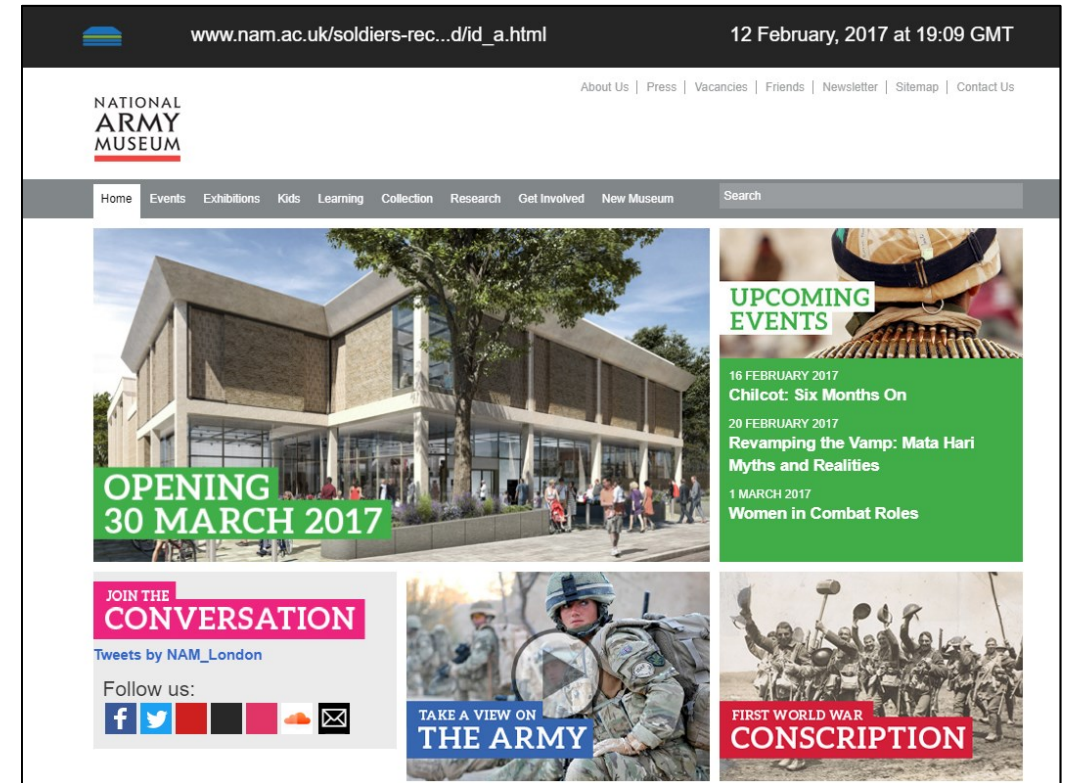
Image

First World War Conscription
www.nam.ac.uk/images/home/ww1-conscription_home.jpg
jpeg 310 x 214
9 February, 2017

Page

National Army Museum, London
www.nam.ac.uk/soldiers-records/deeds/views/directives/dee...
12 February, 2017

Visit page Show image Show details Share



www.nam.ac.uk/soldiers-rec.../id_a.html 12 February, 2017 at 19:09 GMT

NATIONAL ARMY MUSEUM

About Us | Press | Vacancies | Friends | Newsletter | Sitemap | Contact Us

Home Events Exhibitions Kids Learning Collection Research Get Involved New Museum Search

OPENING 30 MARCH 2017

UPCOMING EVENTS

16 FEBRUARY 2017
Chilcot: Six Months On

20 FEBRUARY 2017
Revamping the Vamp: Mata Hari Myths and Realities

1 MARCH 2017
Women in Combat Roles

JOIN THE CONVERSATION

Tweets by NAM_London

Follow us:

Facebook Twitter YouTube Instagram Email

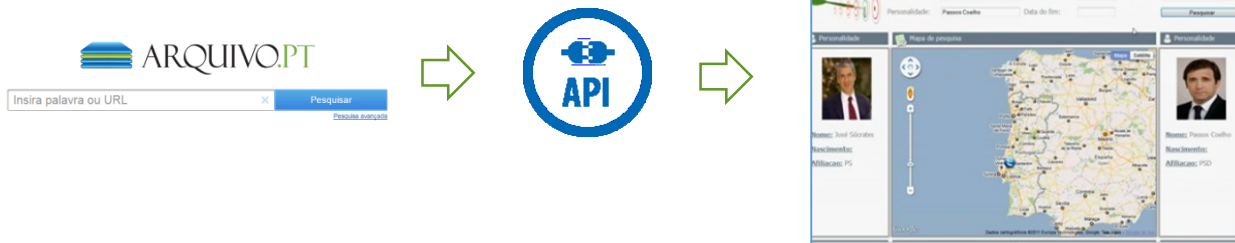
TAKE A VIEW ON THE ARMY

FIRST WORLD WAR CONSCRIPTION

Application Programming Interfaces to develop new applications

Arquivo.pt APIs

Your new app



arquivo / pwa-technologies

<> Code ⓘ Issues 48 🔗 Pull requests 0 📁 Projects 0

APIs

Fernando-Melo edited this page 38 minutes ago · 9 revisions

- URL search: CDX-server API
- URL search: Memento Timetravel API
- Full-text search: OpenSearch based Arquivo.pt API
- Image Search API (under development)

arquivo.pt/api


Custom search engine through “collection” API parameter over 3 crawls of .EU domain (2015-16-17)



[Search pages from the past](#)

[Meet the service](#)

arquivo.pt/resawdev/



[Search](#)

[Advanced search](#)

between:

and:

[Pages](#) [Images](#)

[European elections 2014](#)
22 November, 2014 - [List versions](#)
European elections 2014 Main menu (press 'Enter') Access to page content (press 'Enter')
Direct access to list of other websites (press 'Enter') **European** Parliament / **European Elections**
(22-25 May ... content Secondary navigation **European elections** 2014 Latest products
Court of Auditors: pushing ...
<http://www.elections2014.eu/en>

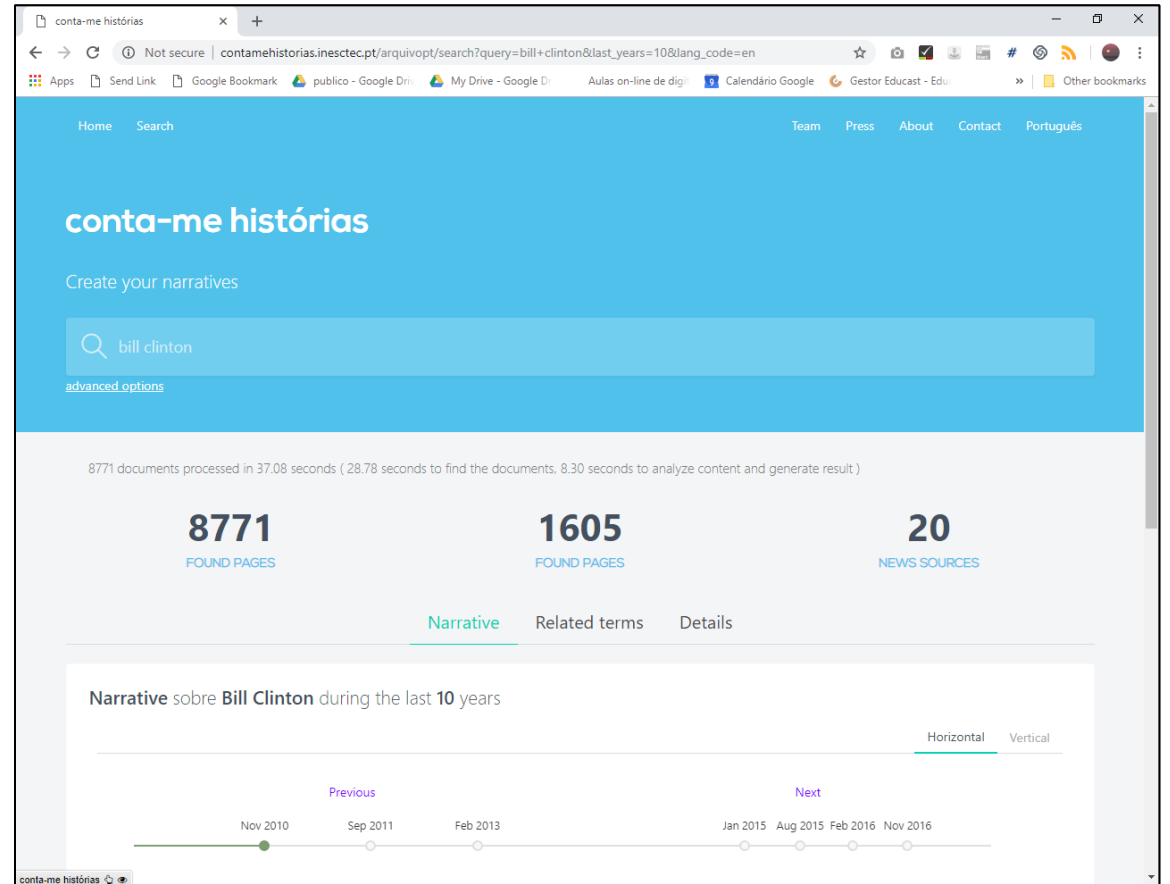
[301 Moved Permanently](#)
8 January, 2016 - [List versions](#)
301 Moved Permanently Moved Permanently The document has moved here ...
<http://www.elections2014.eu/en>

[European Elections 2014: anger or disenchantment? - The Euro...](#)
1 December, 2014 - [List versions](#)
European Elections 2014: anger or disenchantment? - The Euros News feed 19/11
EUROSBLOG FSE: €2 ... Europa dell'era digitale 8/07 The Euros **European Elections** 2014:
anger or disenchantment? 8/07 Gli Euros ... Special files Members EUROSBLOGS Les Euros
Services **European Elections** 2014: anger or ...
<http://www.theeuros.eu/European-Elections-2014-Anger-or,6150.html>

Innovative service based on Arquivo.pt API: “Tell Me Stories”

Automatically generates narratives about any subject based on online news from the Past.

Winners of the Arquivo.pt Award 2018.



ContaMeHistorias.pt (with English UI)
TellMeStories.pt



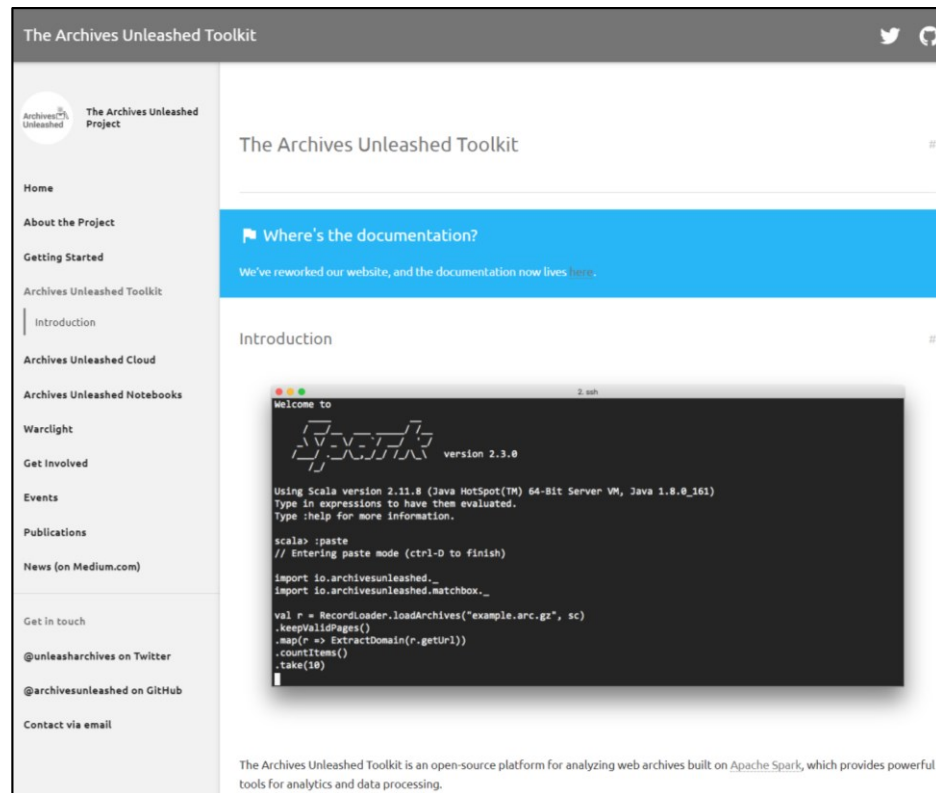
- Any subject
- Arquivo.pt as main source of information
- Submissions in Portuguese
- arquivo.pt/awards

1st place: 10 000 €

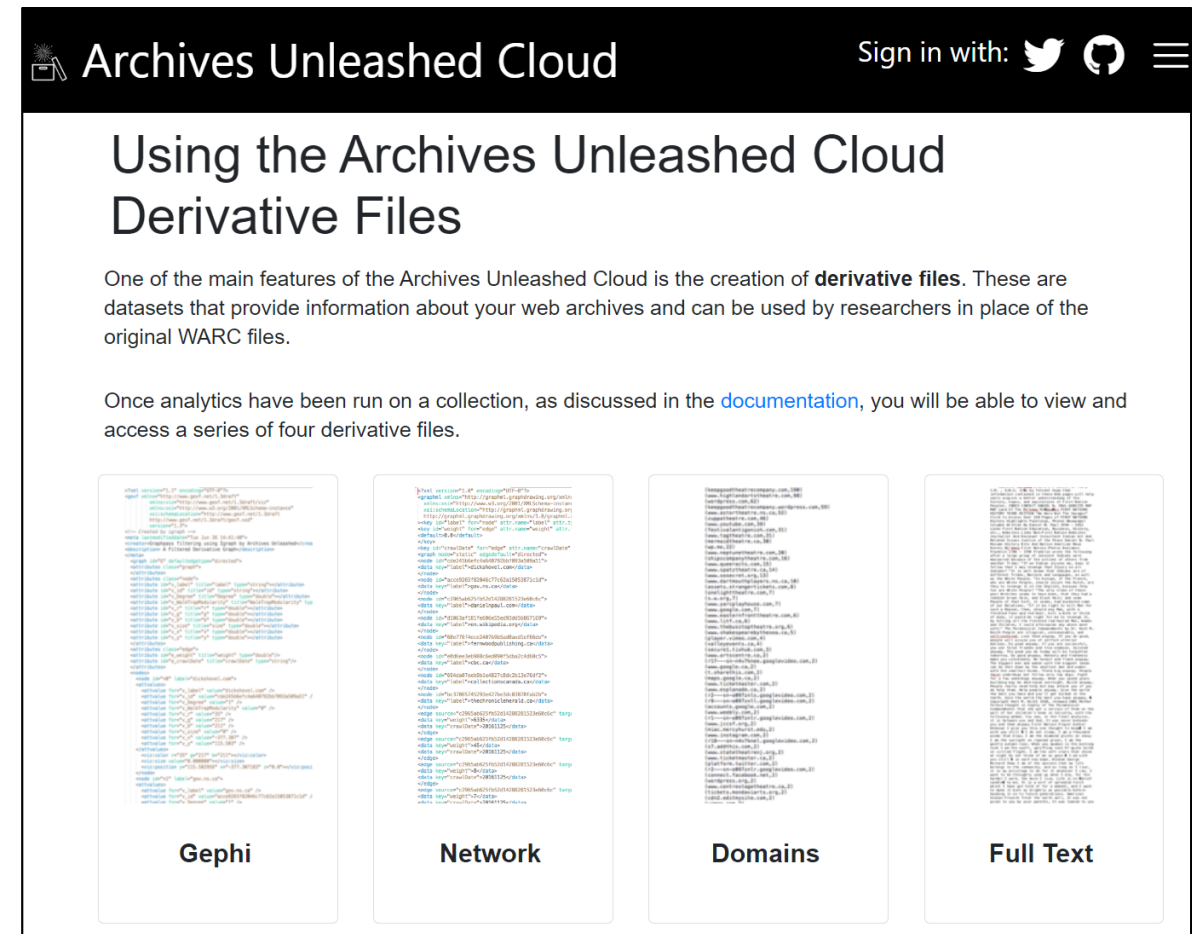
2nd place: 3 000 €

3rd place: 2 000 €

ArchivesUnleashed.org: automatic processing and analytics



Cloud or Toolkit

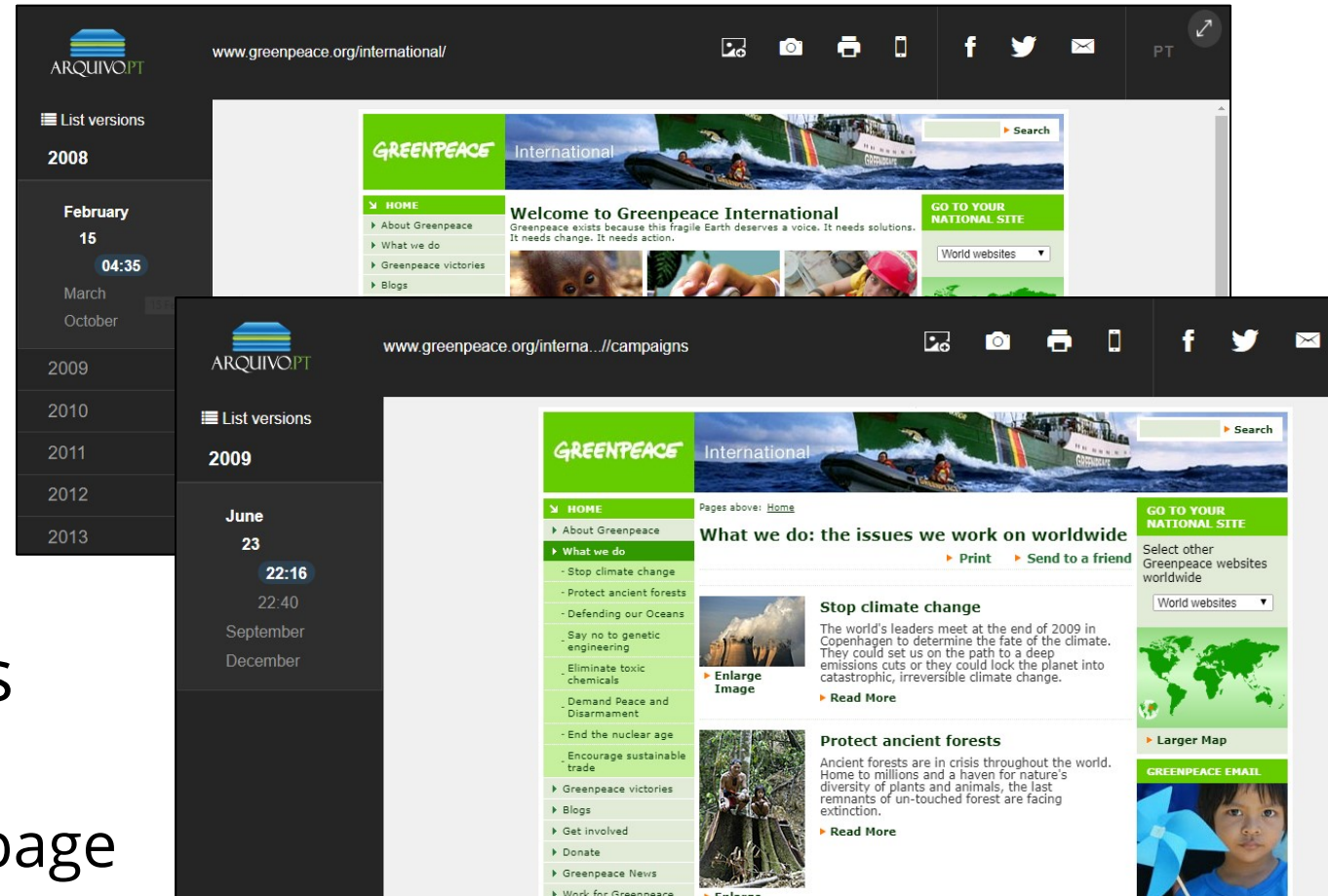


Generates derived results

Access the Past Web

Access web pages from the Past

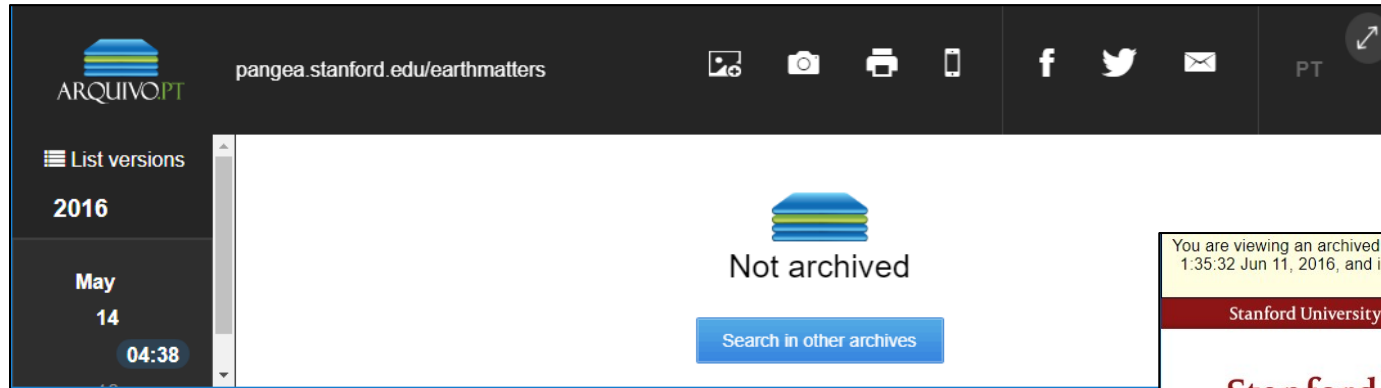
Navigate along **past versions**



Browse links across pages

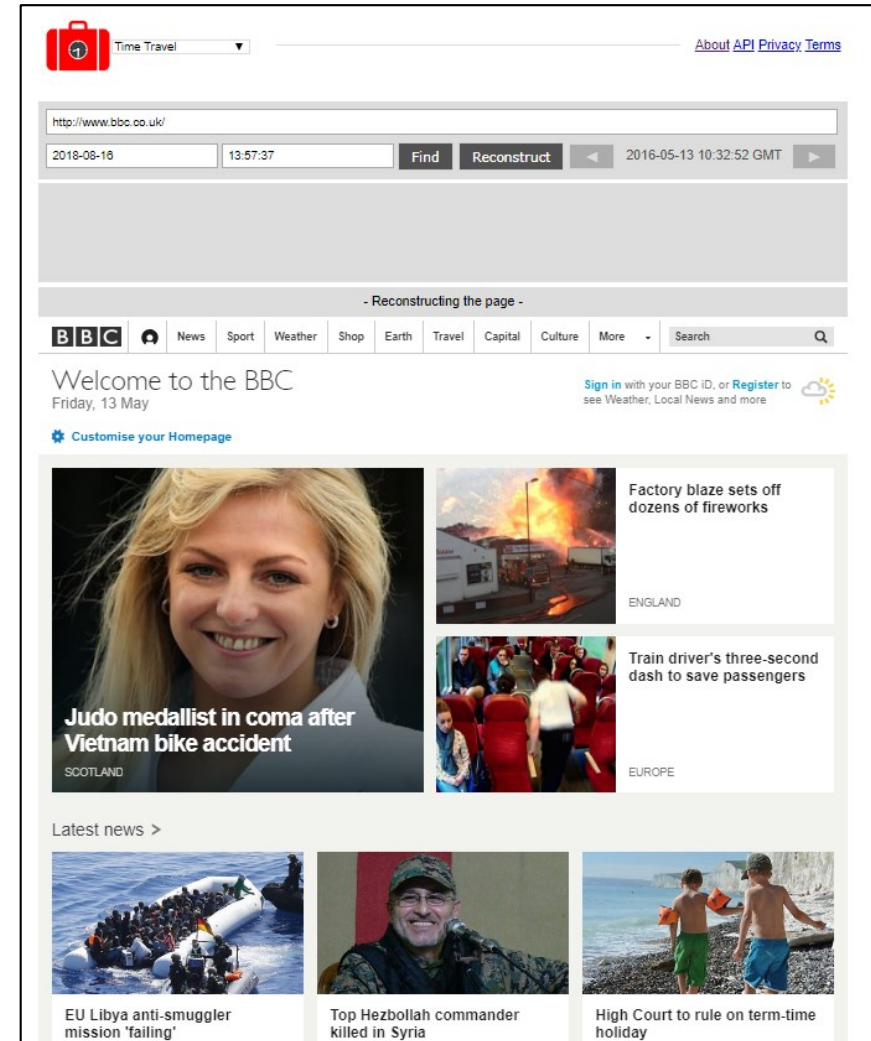
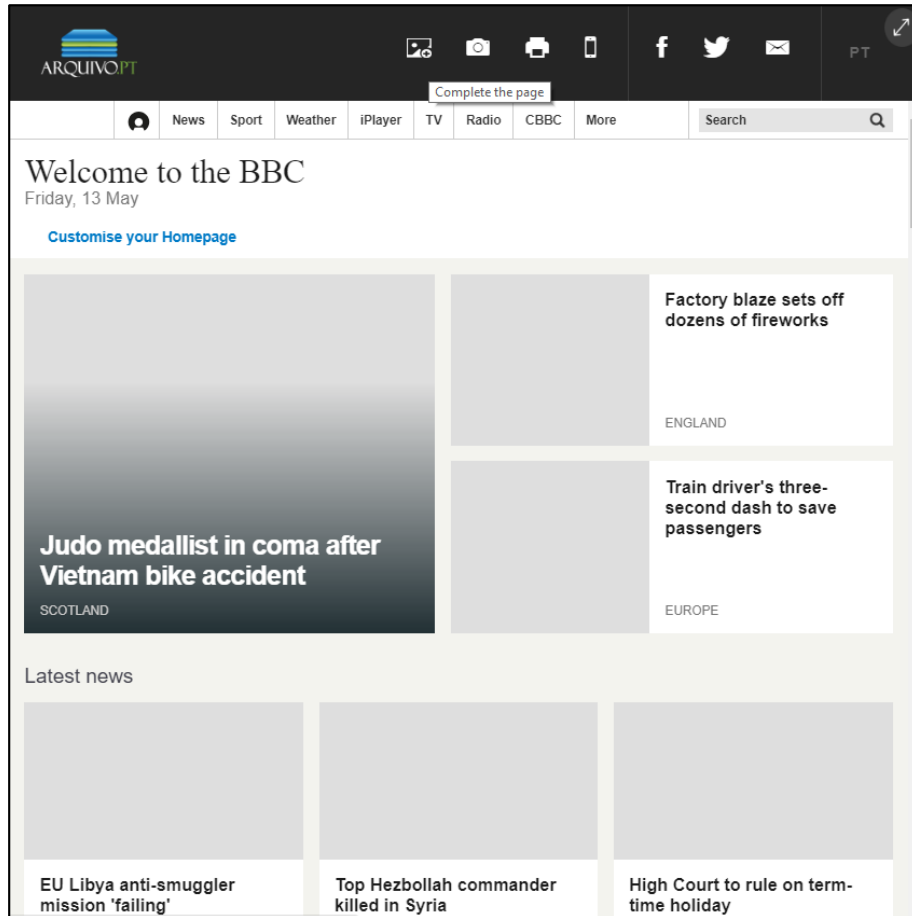
- Destiny version with the closest date to the origin page

Navigate **between** web archives



Powered by:
[TimeTravel.memento.org](https://timetravel.memento.org)

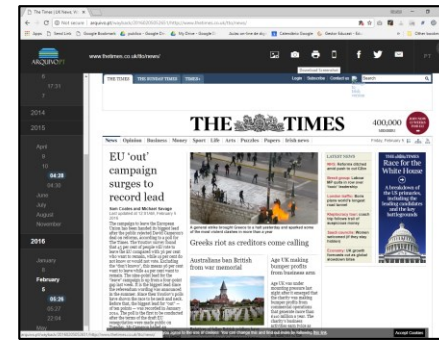
Complete page with resources from external web archives



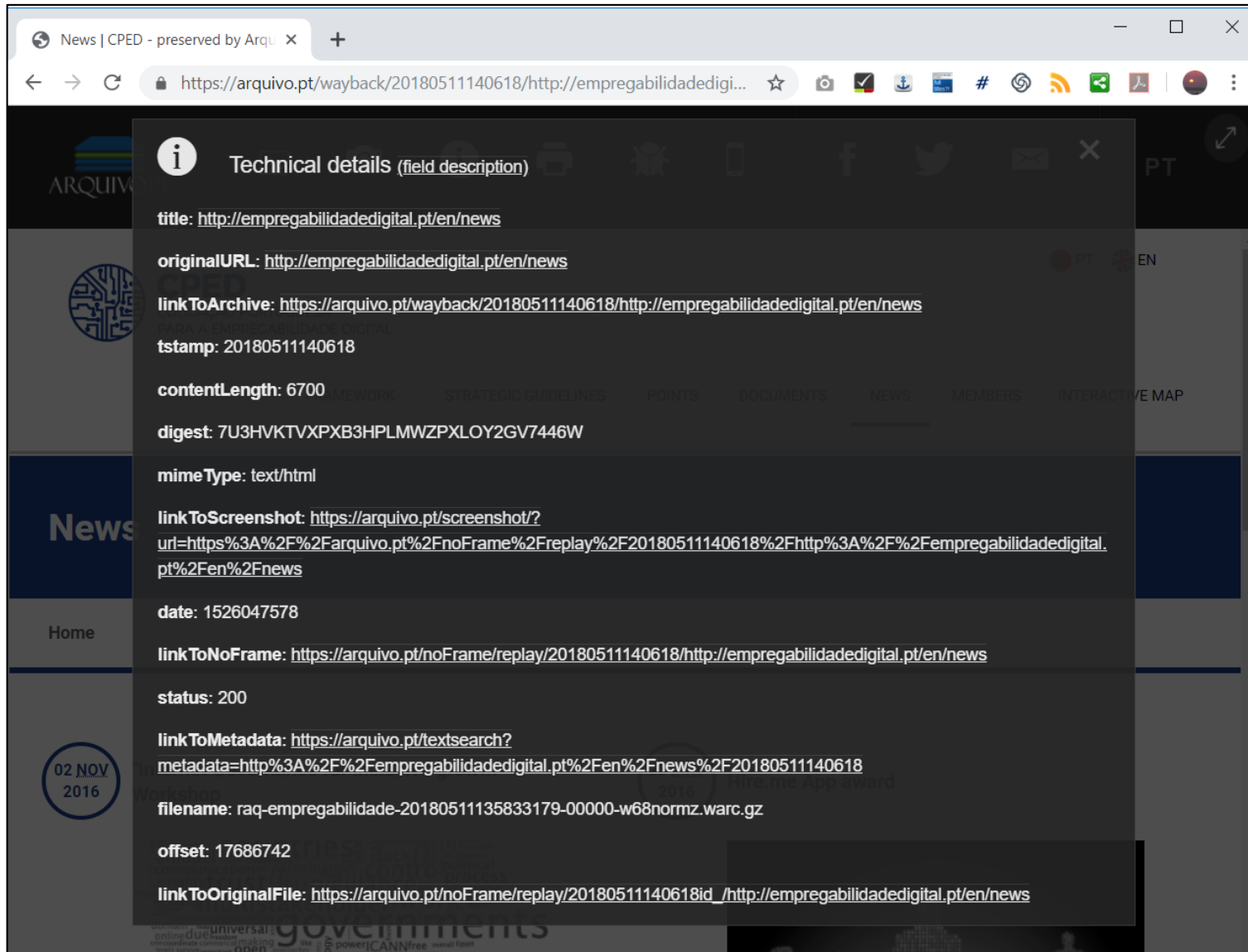
Save **full page** instead of partial "Print Screen"

Embed in documentation
(e.g. thesis, papers)

Keep local copies



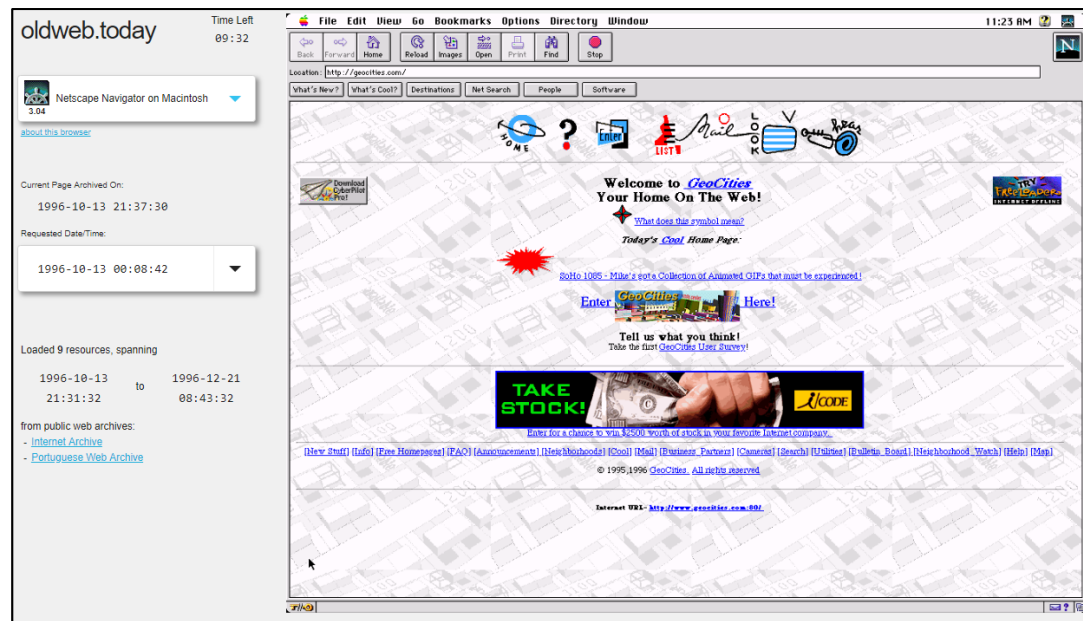
Access to technical details



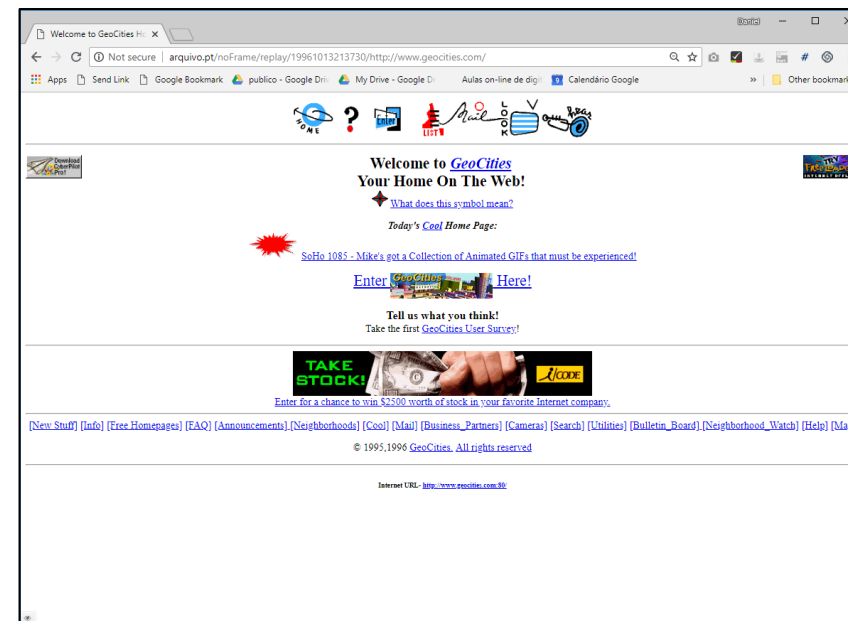
Useful for debug
Advanced users
– E.g. researchers

Oldweb.today: geocities.com, 1996

1997 browser:
Netscape Navigator v.3.04, Macintosh



Modern browser:
Google Chrome, Windows 10



Training and research

Training courses on web preservation and research

New ways of searching the past (module A)

- **Any Internet user**

Publishing preservable information on the web (B)

- Web authors

Automatic processing of information preserved from the Web (C)

- Developers, **Researchers**

Do-it-yourself web archiving (D)

- **Researchers**, Librarians, Archivists



arquivo.pt/training

Investiga XXI (Research XXI)

Communication Studies

Transformations of the Websites of Portuguese Newspapers



Short link to this page: arquivo.pt/newspapers

Information Science

FCSH on the Web: virtual exhibition



Short link to this page: arquivo.pt/fcshontheweb

Social Sciences

Straight-Edge in the Lisbon metropolitan area



Short link to this page: arquivo.pt/straightedgen

All videos, presentations, reports at: arquivo.pt/research

Web archives synergies with CLEOPATRA

[ESR 1. Fact extraction and cross-lingual alignment:](#)

cross-lingual collections (e.g. IIPC Olympics/Paralympics collections, Arquivo.pt EU Elections 2019).

[ESR 3. Crowd quality and training in hybrid multilingual information processing and analytics:](#)

generate test collections and evaluate quality of web archive tools.

[ESR 5. Fact validation across multilingual text corpora:](#)

web archives as additional source of historical information for fact checking.

[ESR 7. Relations of textual and visual information](#) & [ESR 8. Contextualisation of images in multilingual sources:](#)

web archive image search.

[ESR 9. National and transnational media coverage of European parliamentary elections, 2004-2014:](#)

Arquivo.pt EU Elections 2019.

[ESR 10. Nationalism, internationalism and sporting identity: the London and Rio Olympics/Paralympics:](#)

IIPC Olympics/Paralympics collections, Arquivo.pt daily crawls of online news.

[ESR 11. Information propagation with barriers:](#)

study how information propagates across time on the Web, IIPC European Refugee Crisis.

[ESR 12. Cross-lingual news reporting bias:](#) IIPC Online News Around the World collection.

[ESR 13. Multilingual Wikipedia as 'first draft of history':](#) final layout of Wikipedia pages along time (not database dump).

[ESR 14. NLP for under-resourced languages](#) & [ESR 15 Cross-lingual sentiment detection:](#)

under-resourced languages in web archives.

Search and Access the Past Web

Search the Present -> Web Search Engines

Search the Past -> Web Archives!

Useful research tools

- [Archive.org](https://archive.org): URL search
- [TimeTravel.mementoweb.org](https://timetravel.mementoweb.org): search across web archives
- [Arquivo.pt](https://arquivo.pt): full text search
- [ArchivesUnleashed.org](https://archivesunleashed.org): data analytics
- [Oldweb.today](https://oldweb.today): access using old browsers

Keep in touch!

Subscribe our channels:

- arquivo.pt/subscribe: mailing list in English
- arquivo.pt/inscrever: mailing list in Portuguese
- arquivo.pt/news: social networks and videos