

# Arquivo.pt Memorial and other goodies

[Daniel.Gomes@fccn.pt](mailto:Daniel.Gomes@fccn.pt)



ARQUIVO.PT

(Ar – Kee – Voo .PT)

# Who are we?

**Free and public online** service to preserve the Past Web

Preserves **publicly accessible** information related with:



















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

**Public institution** belonging to the Ministry of Science and Higher Education

A **digital preservation infrastructure**

# Arquivo.pt is used word-wide



Country	Users	% Users
1.  Portugal	38,456	 22.08%
2.  United States	35,805	 20.56%
3.  Russia	33,846	 19.43%
4.  Brazil	6,565	 3.77%
5.  Kazakhstan	5,808	 3.33%
6.  Japan	4,902	 2.81%
7.  United Kingdom	4,621	 2.65%
8.  Ukraine	2,739	 1.57%
9.  Canada	2,572	 1.48%
10.  Germany	2,403	 1.38%

78% of our users were international in 2018

User Interfaces and documentation also in English

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

# Arquivo.pt Memorial

So that your old website may live forever...

# Old websites: the IT department **nightmare**

Nobody cares about them

- **No longer updated** with new content
- **No significant resources** to maintain them

But also nobody let's them go

- Provide **important information**
- **Increasing** costs of maintainance
  - Software obsolescence
  - Old hardware with inefficient power consumption
- Dangerous **security vulnerabilities**
  - Websited neglected get hacked
  - Die shamefully and information gets lost

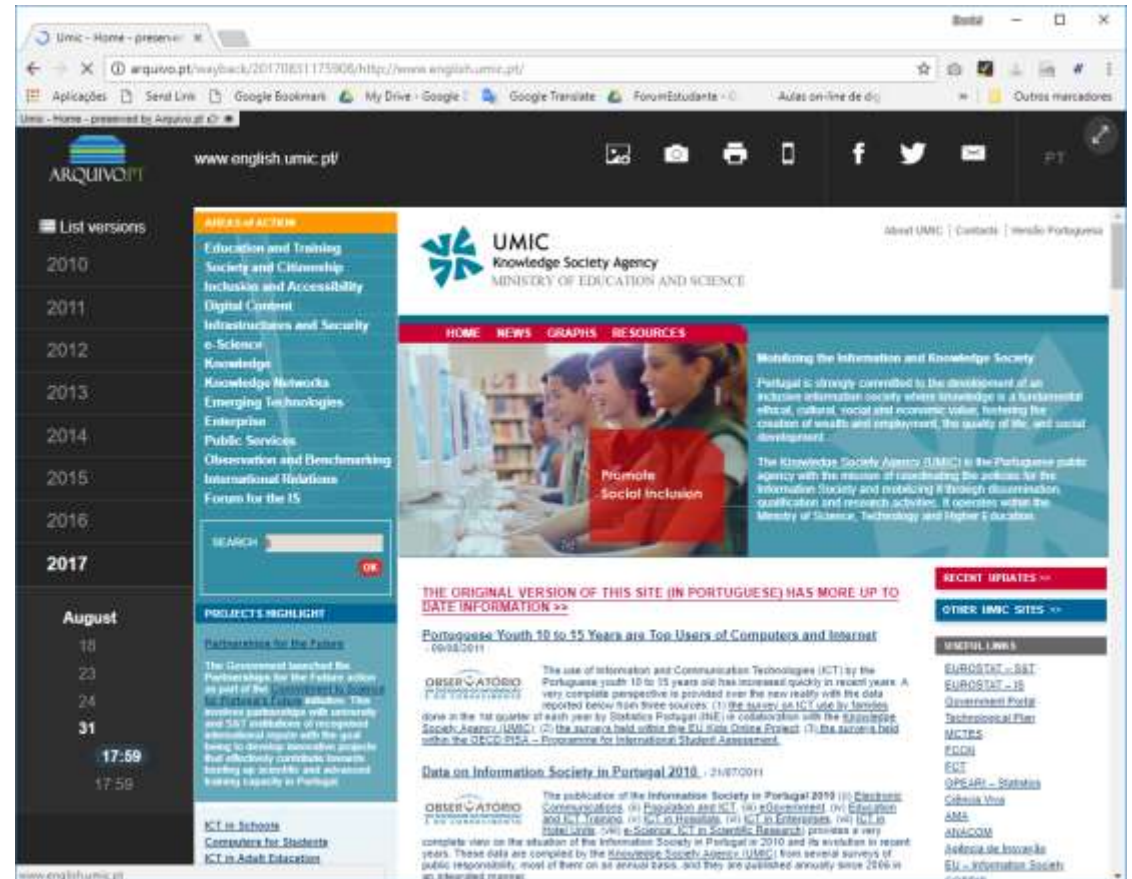


# Arquivo.pt Memorial: Stop worrying about maintaining your old websites



**website content** is maintained  
accessible at Arquivo.pt

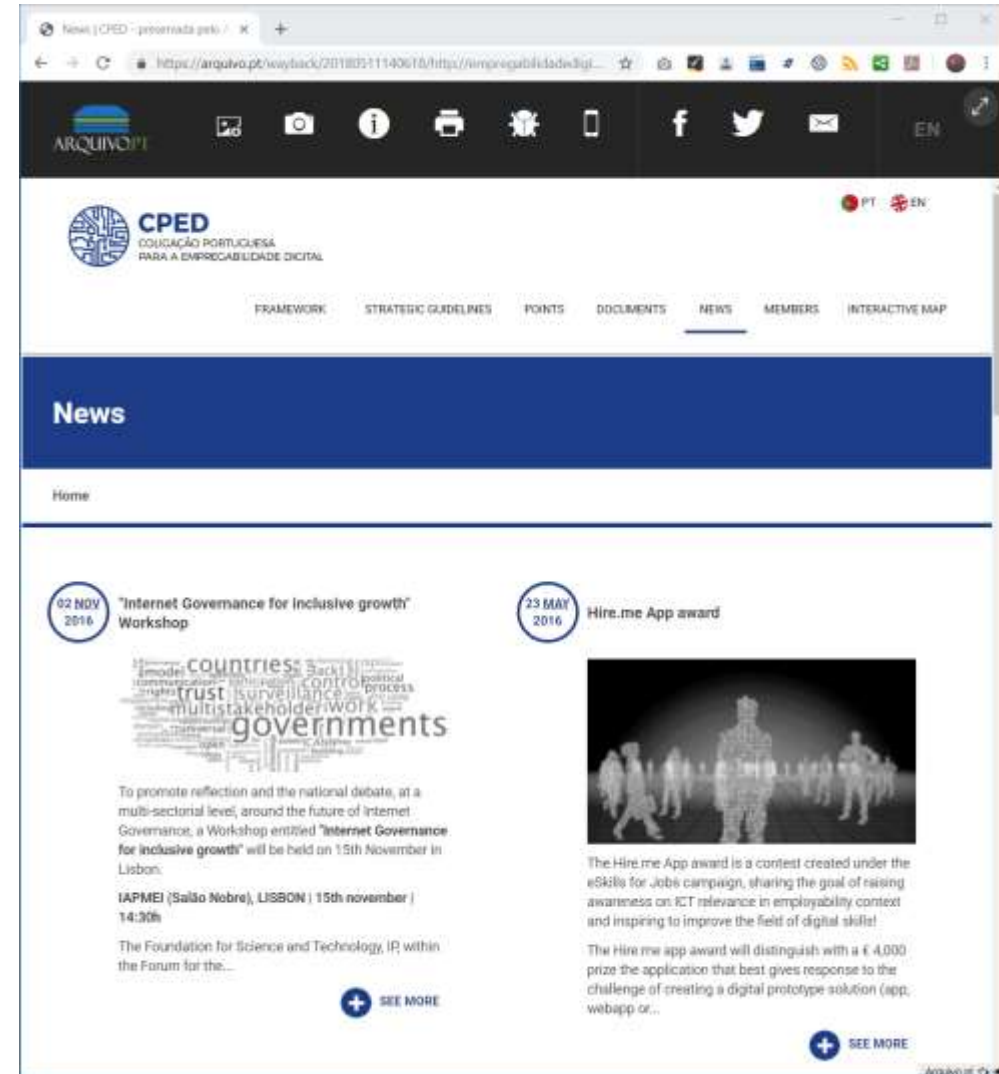
[english.unic.pt](http://english.unic.pt): **website address** is  
maintained accessible on the live-web



# Also maintains links to inner pages



<http://empregabilidadedigital.pt/en/news>





# Workflow to preserve a website in the Memorial

1. **Request** from website owner
2. **Check quality** of existing web-archived version
3. Patch with **webrecorder.io** or new high-quality crawl with **Brozzler**
4. **Feedback** from website owner
5. Website owner **redirects domain** to our name servers

## 6. **Disconnect** annoying old website



Website owners must **renew their domains!**

# Initial results about Arquivo.pt Memorial

Testing during 2018

Service announced in February 2019

14 websites in the Memorial

25 underway in cooperation with the  
Management Center for the  
Government Computer Network

- Manages .GOV.PT
- Government Portal

## Memorial of Arquivo.pt

### Is it a problem to maintain your old website?

There are many frozen websites without new contents, that may be maintained online, because they host useful information, like for example sites of finished projects).

Organizations have to invest important resources for the long term maintainance of old websites.

Costs grows with time in old websites, because of obsolescence in supporting technologies. This also causes dangerous vulnerabilities on security.

### The Memorial of Arquivo.pt preserves your old website

This service offers a high quality preservation of frozen contents and gives the chance to maintain the original domain name.

For example, [unplace.org](http://unplace.org) domain is maintained but redirects to a preserved version on Arquivo.pt. Owners of this website don't have to maintain the infrastructure (e.g. servers, electricity, content management system). They only have to maintain the domain name.

If you want your website to be on the Memorial of Arquivo.pt, contact us.

### Examples of preserved websites



UMIC – Agência para a  
Sociedade do  
Conhecimento



Coligação Portuguesa para  
a Empregabilidade Digital



Ligar Portugal



Iniciativa Nacional GRID

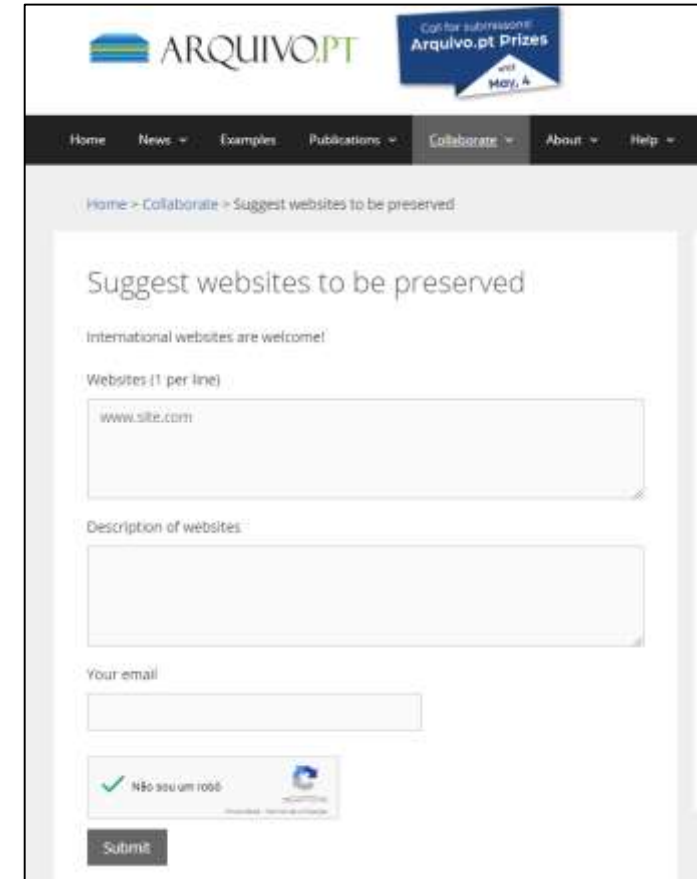
<https://arquivo.pt/memorial>

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

Any domain

Any language

Broad interest for research  
and education



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' with a deadline of May 4. The navigation bar contains links for Home, News, Examples, Publications, Collaborate, About, and Help. The breadcrumb trail reads 'Home > Collaborate > Suggest websites to be preserved'. The form itself has a title 'Suggest websites to be preserved' and a sub-header 'International websites are welcome!'. It includes a label 'Websites (1 per line)' above a text input field containing 'www.site.com'. Below this is a 'Description of websites' text area. Further down is a 'Your email' input field. At the bottom, there is a 'Não sou um robô' (I am not a robot) checkbox with a reCAPTCHA logo, and a 'Submit' button.

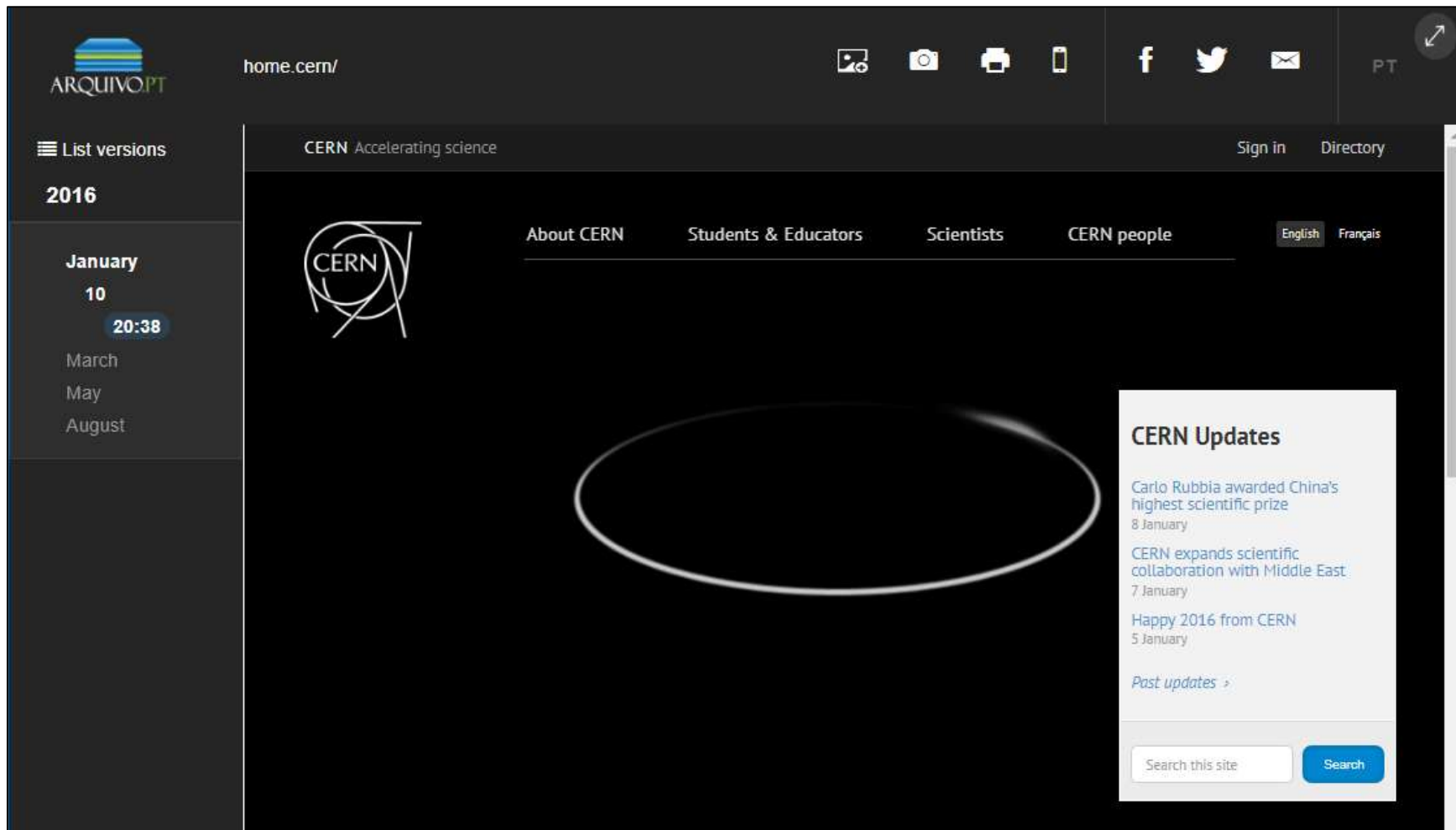
[arquivo.pt/suggest](https://arquivo.pt/suggest)

Other Arquivo.pt goodies

# URL Search: Wayback Machine

[illegible]

# [www.cern.ch](http://www.cern.ch), 2016



# [www.cern.ch](http://www.cern.ch), 1996

## CERN European Laboratory for Particle Physics

[Lab](#) - [News](#) - [Activities](#) - [Physics](#) - [Other Subjects](#) - [Index](#) - [Search](#) - [Shrink](#) - Expand

Welcome to the European Laboratory for [Particle Physics](#), located near [Geneva](#) in [Switzerland](#) and [France](#). CERN is the birthplace of the [World-Wide Web](#).

The [WWW support team](#) provides a set of [Services](#) to the physics experiments and the lab.

## About the Laboratory

### General::

[Help](#) and [General information](#), [Divisions, groups and activities](#) (structure), [Scientific Committees](#), [Scientific Information Service](#) (Library and Archives), [Visits Service](#), [Housing Service](#), [Microcosm exhibition](#), [CERN Courier](#), [Safety guide](#) for experiments, [Education Services](#), [Recruitment and Vacancies](#), [Clubs and Associations](#).

### Databases:

[Directories](#) (phone & email, services & people), [Buildings](#), [Alice](#) (library and preprints catalogue), [Preprint Server](#), [other directories](#).

### Agenda and News:

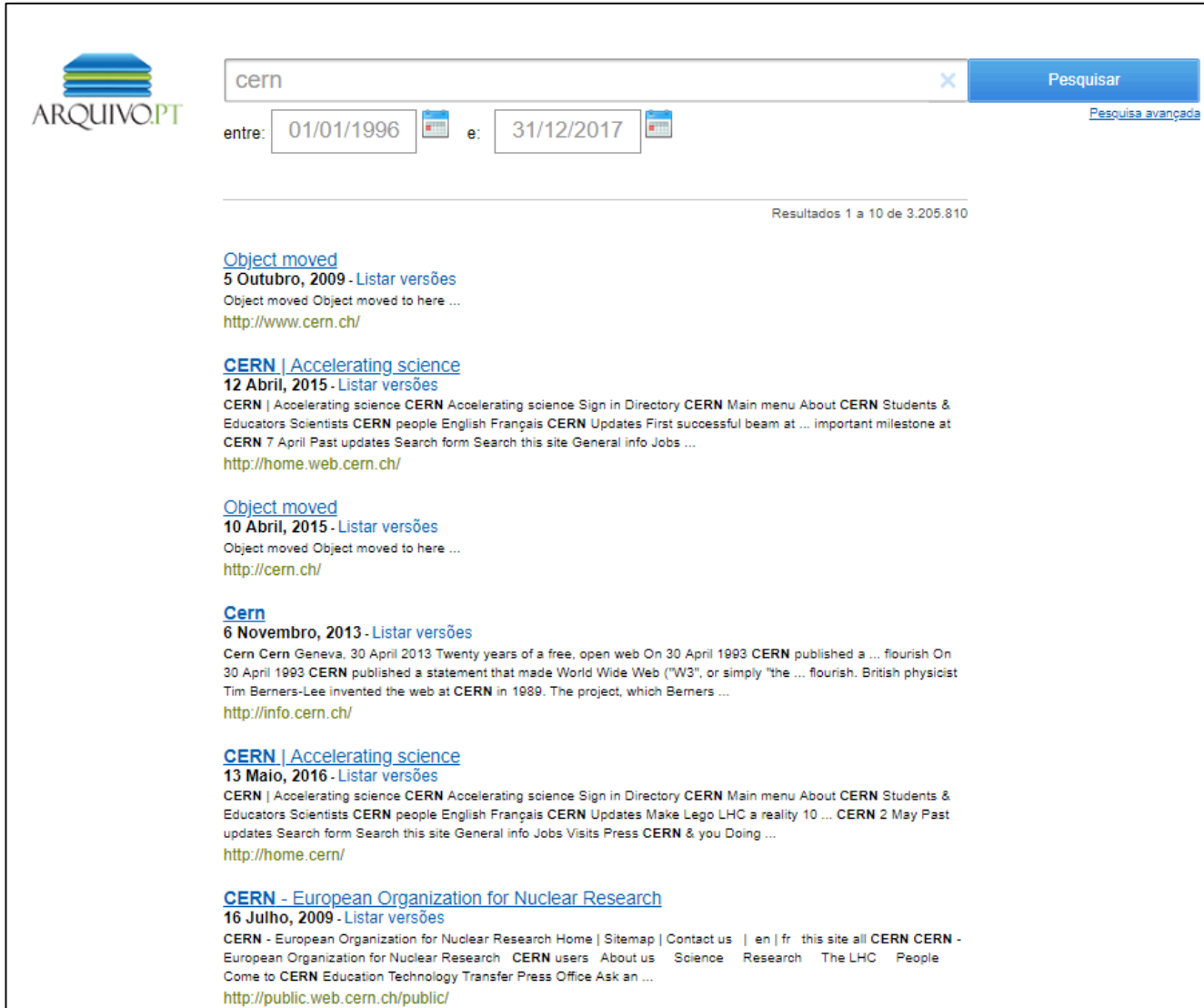
[Press Office](#), News from the [Users' Office](#), [ACCU](#), current [Seminars](#), CERN [Schools & Conferences](#), internal [Newsletters](#), internal [CERN news](#).

## Hot News

[New President of CERN Council](#), Professor Luciano Maiani is new President of CERN Council  
[Seminars on WWW Authoring](#), Bebo White of SLAC will give seminars on Web authoring  
[Symposium on Hadron therapy](#), Second International Symposium on Hadron Therapy, CERN  
[Older News](#), Hot News (older items)



# Full-text search over all pages with Arquivo.pt



The screenshot shows the Arquivo.pt search interface. The search bar contains the text 'cern'. Below the search bar, there are date filters: 'entre: 01/01/1996' and 'e: 31/12/2017'. The search button is labeled 'Pesquisar'. Below the search bar, there is a link to 'Pesquisa avançada'. The search results are displayed below the search bar, showing 'Resultados 1 a 10 de 3.205.810'. The first result is titled 'Object moved' and dated '5 Outubro, 2009'. The second result is titled 'CERN | Accelerating science' and dated '12 Abril, 2015'. The third result is titled 'Object moved' and dated '10 Abril, 2015'. The fourth result is titled 'Cern' and dated '6 Novembro, 2013'. The fifth result is titled 'CERN | Accelerating science' and dated '13 Maio, 2016'. The sixth result is titled 'CERN - European Organization for Nuclear Research' and dated '16 Julho, 2009'.

ARQUIVO.PT

cern

Pesquisar

entre: 01/01/1996 e: 31/12/2017

Pesquisa avançada

Resultados 1 a 10 de 3.205.810

[Object moved](#)  
5 Outubro, 2009 - [Listar versões](#)  
Object moved Object moved to here ...  
<http://www.cern.ch/>

[CERN | Accelerating science](#)  
12 Abril, 2015 - [Listar versões](#)  
CERN | Accelerating science CERN Accelerating science Sign in Directory CERN Main menu About CERN Students & Educators Scientists CERN people English Français CERN Updates First successful beam at ... important milestone at CERN 7 April Past updates Search form Search this site General info Jobs ...  
<http://home.web.cern.ch/>

[Object moved](#)  
10 Abril, 2015 - [Listar versões](#)  
Object moved Object moved to here ...  
<http://cern.ch/>

[Cern](#)  
6 Novembro, 2013 - [Listar versões](#)  
Cern Cern Geneva, 30 April 2013 Twenty years of a free, open web On 30 April 1993 CERN published a ... flourish On 30 April 1993 CERN published a statement that made World Wide Web ("W3", or simply "the ... flourish. British physicist Tim Berners-Lee invented the web at CERN in 1989. The project, which Berners ...  
<http://info.cern.ch/>

[CERN | Accelerating science](#)  
13 Maio, 2016 - [Listar versões](#)  
CERN | Accelerating science CERN Accelerating science Sign in Directory CERN Main menu About CERN Students & Educators Scientists CERN people English Français CERN Updates Make Lego LHC a reality 10 ... CERN 2 May Past updates Search form Search this site General info Jobs Visits Press CERN & you Doing ...  
<http://home.cern/>

[CERN - European Organization for Nuclear Research](#)  
16 Julho, 2009 - [Listar versões](#)  
CERN - European Organization for Nuclear Research Home | Sitemap | Contact us | en | fr this site all CERN CERN - European Organization for Nuclear Research CERN users About us Science Research The LHC People Come to CERN Education Technology Transfer Press Office Ask an ...  
<http://public.web.cern.ch/public/>

Found new pages about CERN on different URLs.

[home.web.cern.ch](http://home.web.cern.ch)

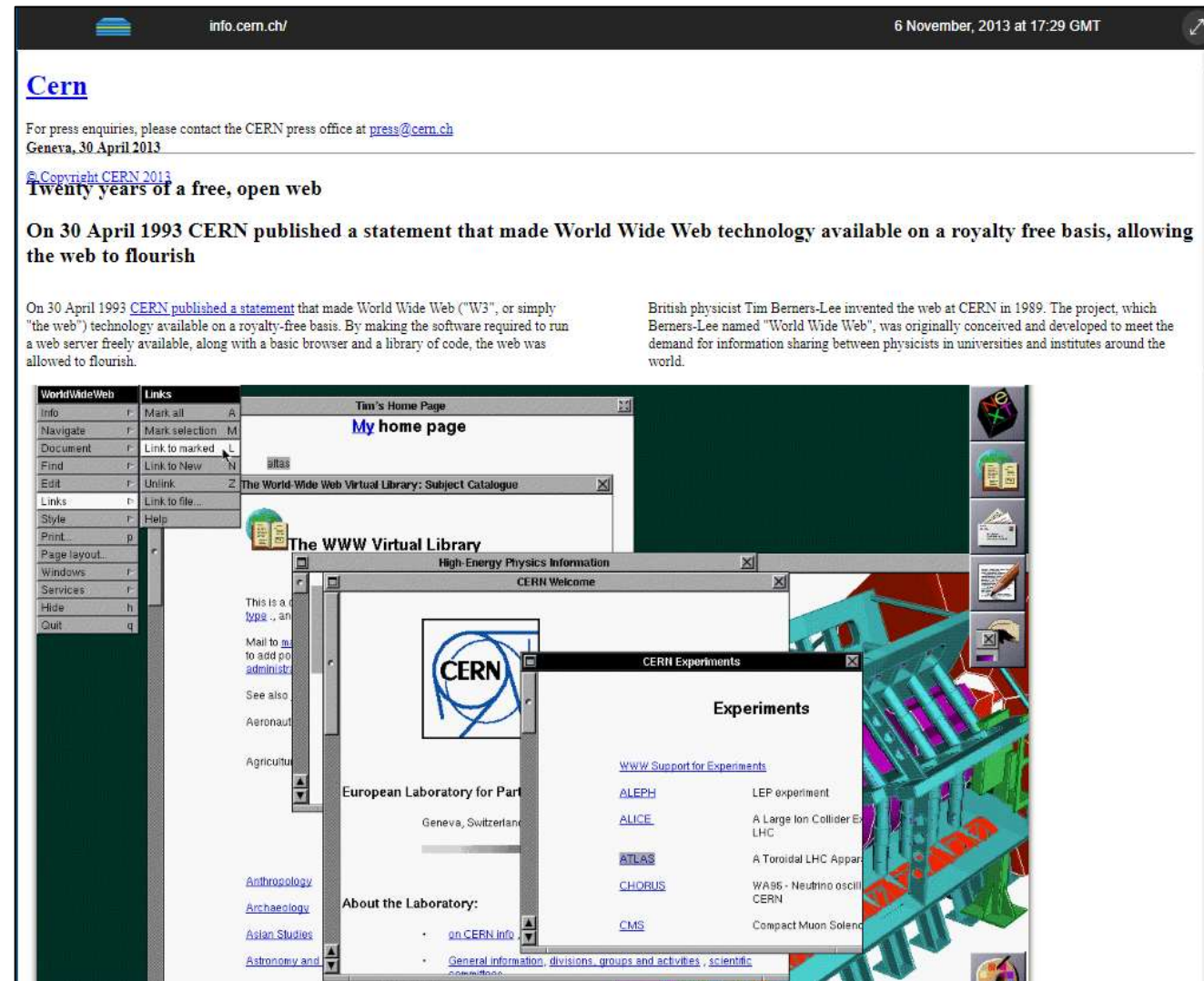
[cern.ch](http://cern.ch)

[info.cern.ch](http://info.cern.ch)

[home.cern](http://home.cern)

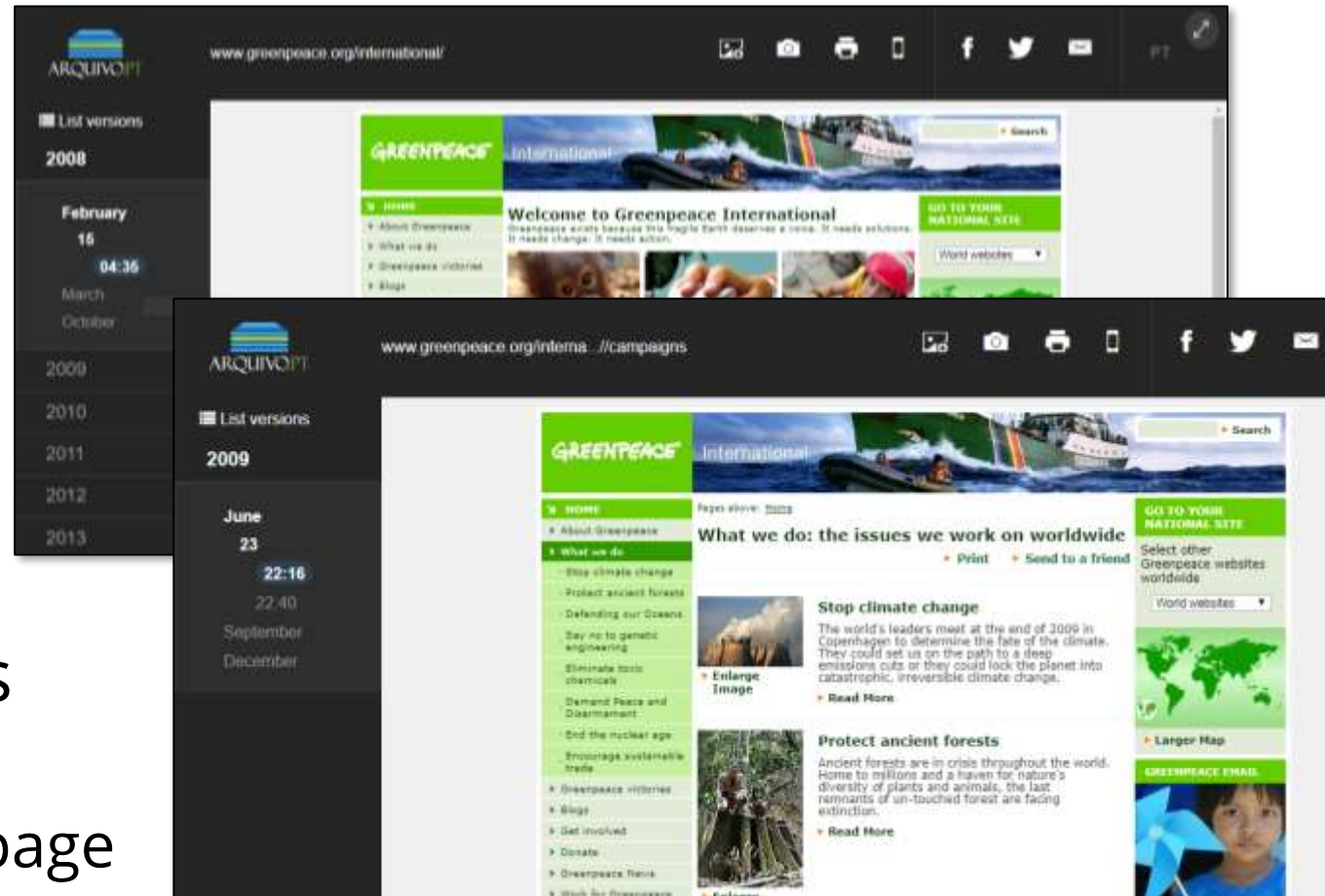
[public.web.cern.ch/public/](http://public.web.cern.ch/public/)

# Found cool pages from the Past: [public.web.cern.ch/public/](http://public.web.cern.ch/public/) “Twenty years of a free, open web, 2013”



# Access web pages from the Past

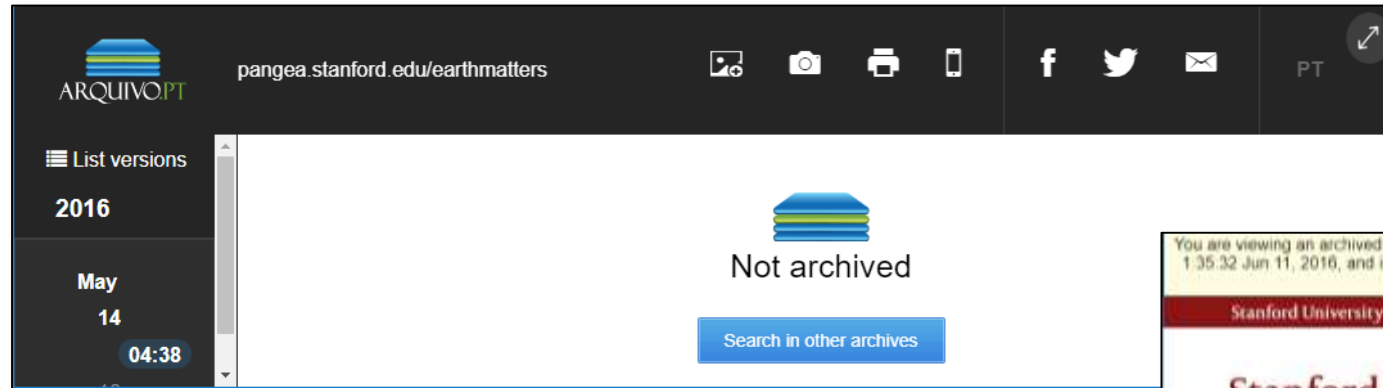
Navigate along **past versions**



**Browse links** across pages

- Destiny version with the closest date to the origin page

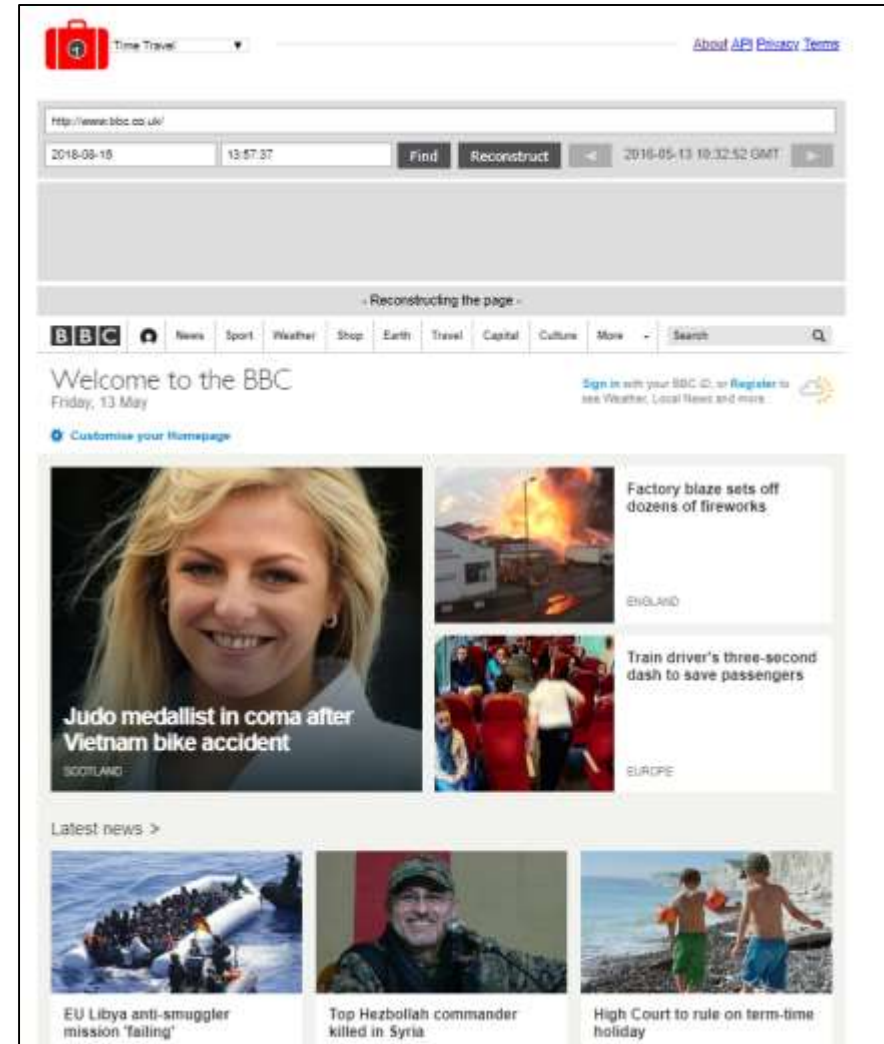
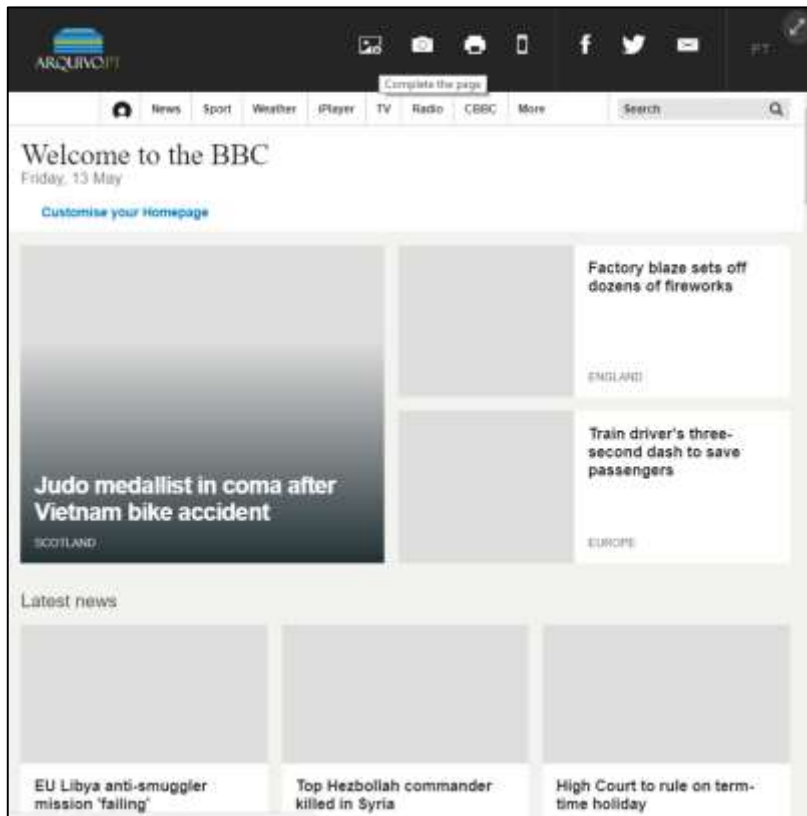
# Navigate across web archives



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



# Complete page with resources from external web archives

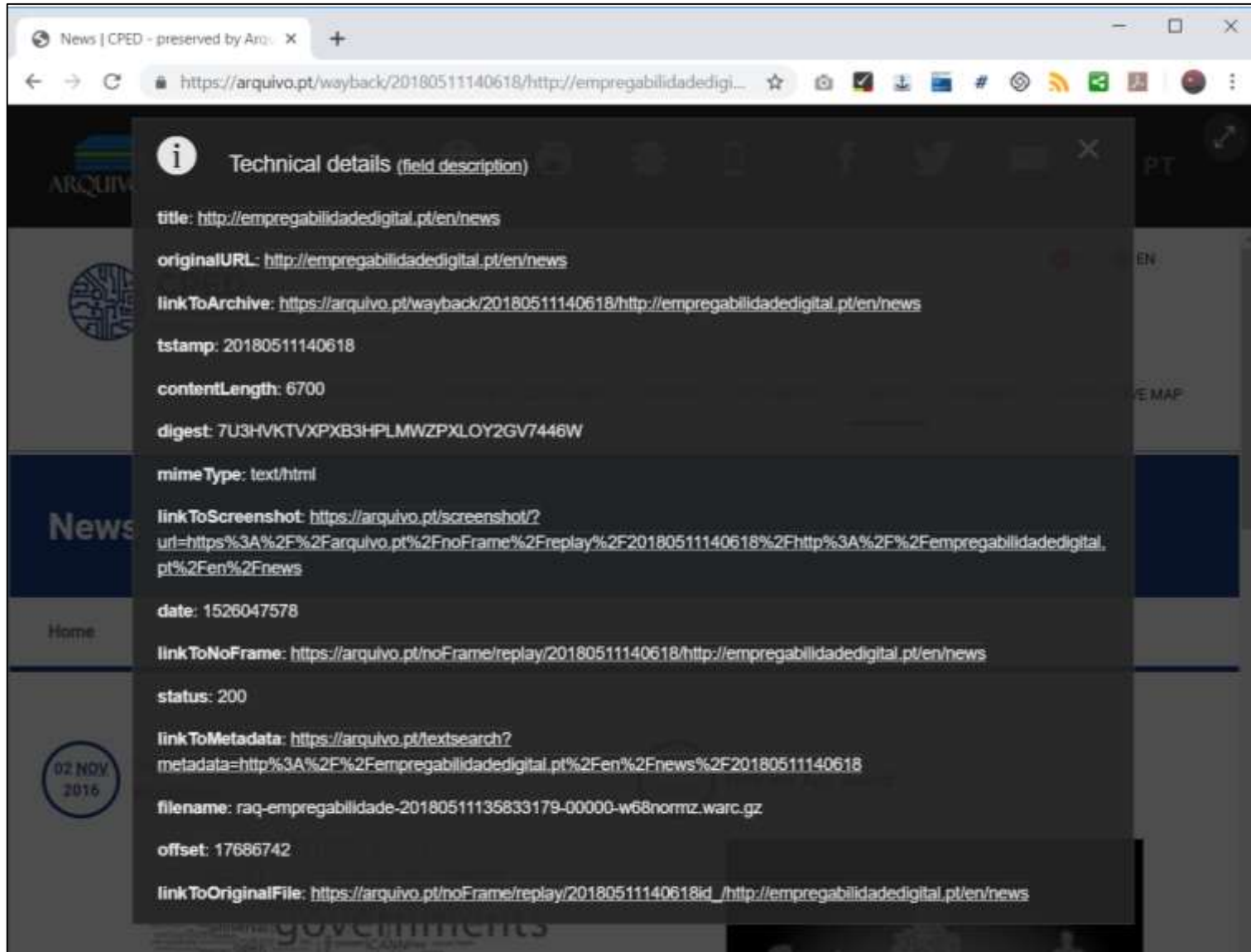


Powered by: [TimeTravel.memento.org](http://TimeTravel.memento.org)

# Download/print **full page screenshot** instead of “Print Screen”

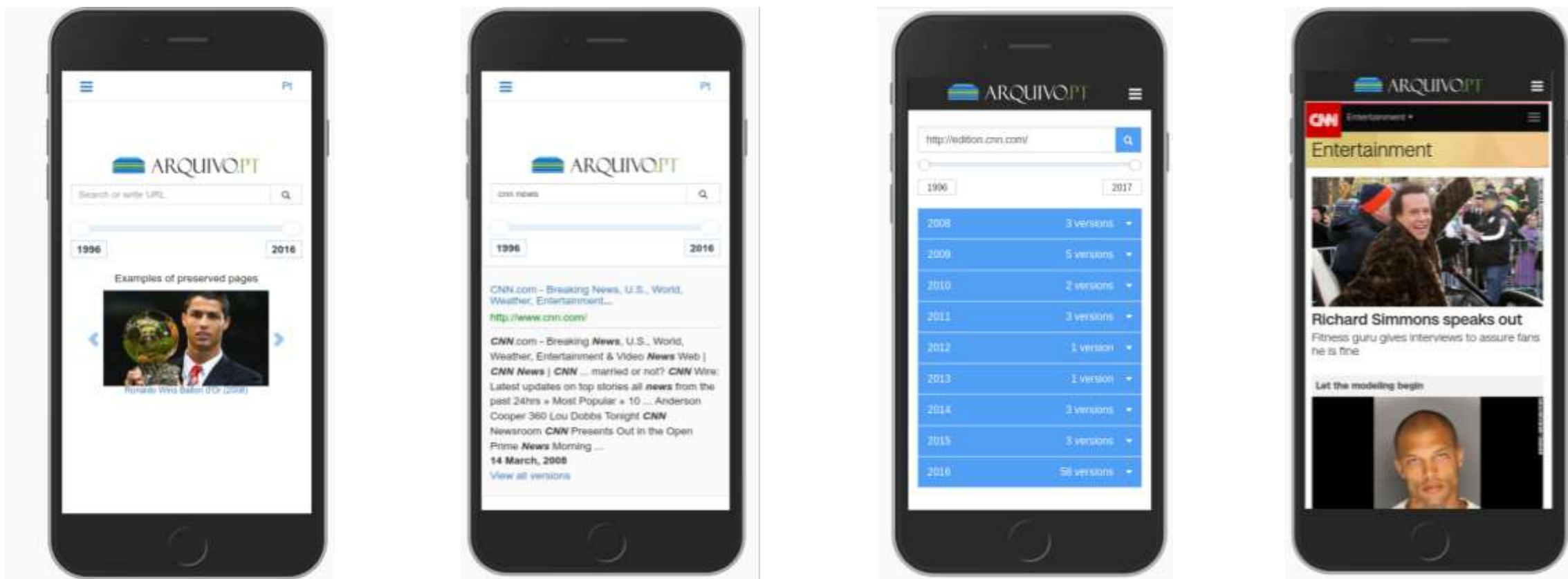


# Access to **technical details**



Useful for debug  
Advanced users  
– E.g. researchers

# Mobile User Interfaces



29% of our Users used mobile devices in 2018



# Application Programming Interface to develop new applications

Your new app



## New Arquivo.pt API

The screenshot shows the GitHub repository page for 'arquivo / pwa-technologies'. The page has tabs for 'Code', 'Issues' (48), 'Pull requests' (0), and 'Projects' (0). The main heading is 'APIs', with a note that 'Fernando-Melo edited this page 38 minutes ago · 9 revisions'. Below this, there is a list of APIs:

- 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](https://arquivo.pt/api)

# Final remarks...

Suggest websites to be preserved

- [arquivo.pt/suggest](https://arquivo.pt/suggest)
- Keep them alive at the Arquivo.pt Memorial

Follow us

- [arquivo.pt/news](https://arquivo.pt/news)
- Subscribe our international mailing list:  
[arquivo.pt/subscribe](https://arquivo.pt/subscribe)