

# Mãos na massa para arquivar a web

Coimbra, 6 de maio de 2025



# Objetivo

- Gravar, armazenar e reproduzir páginas Web no próprio computador num formato normalizado, utilizando ferramentas gratuitas do [Webrecorder.net](http://Webrecorder.net)

# Agenda

Introdução: O Arquivo.pt e a promoção das práticas de preservação da Web

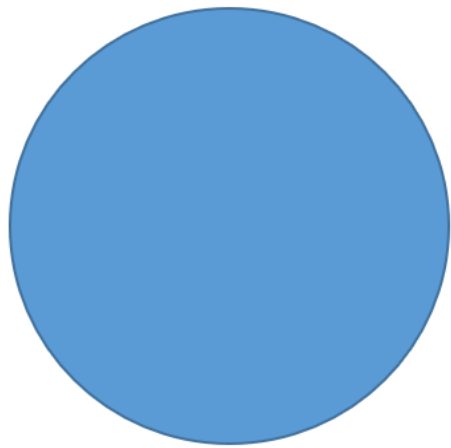
1. Demonstrações básicas de gravação com a extensão ArchiveWeb.page
2. Mãos na massa ou “*gravathon*”
3. Partilha das gravações
4. Extra: Gravação automática com o Browsertrix-crawler

# O Arquivo.pt

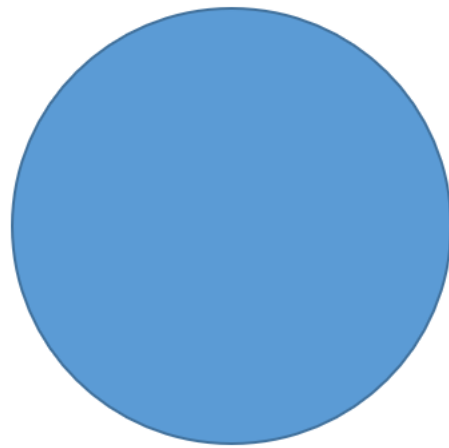
- Serviço público da FCCN Unidade de Computação Científica da Fundação para a Ciência e a Tecnologia, I. P.
- Missão: **“Promover a preservação de conteúdos disponíveis na Internet nacional, garantindo a disponibilização deste à comunidade científica e ao público em geral”** (Decreto-Lei nº 55/2013)
- Acessível na Web em <https://arquivo.pt>
- Instalações físicas – Av. do Brasil – Lisboa (Campus do LNEC)



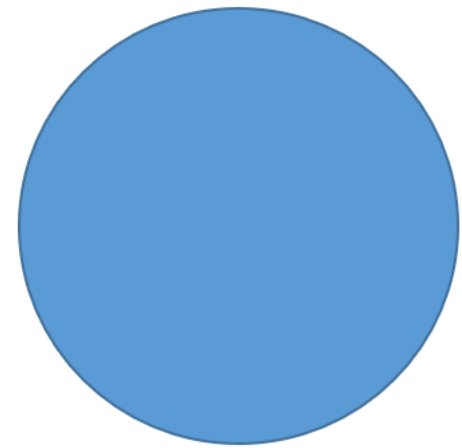
# Arquivar a Web



Gravar



Armazenar



Reproduzir

# Arquivar a Web



# Webrecorder.net

“Web archiving for all”

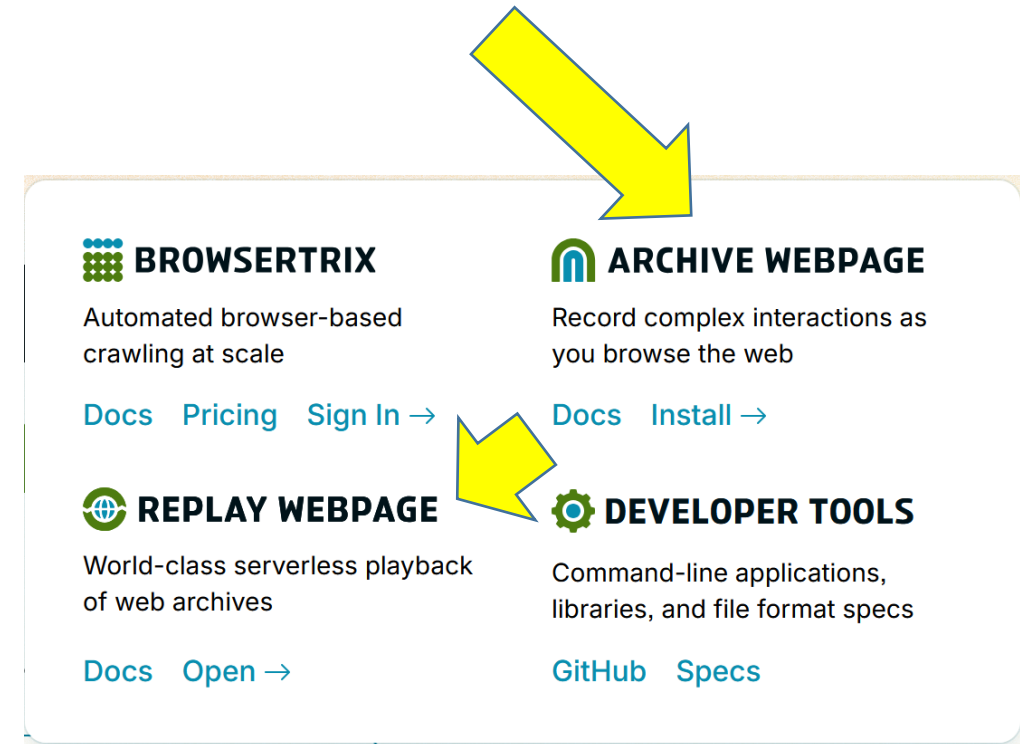
- 2016 - Parte do Projeto Rhizome.org (Webrecorder.io)
- 2019 - Independente do Projeto Rhizome, dá suporte ao serviço Conifer
- 2020 - Webrecorder.net

# Webrecorder.net

- Código aberto
- Centrado no utilizador não especializado
- Formatos normalizados
- **Compatibilidade com os arquivos da Web**
- Apoio do International Internet Preservation Consortium (IIPC)










# Webrecorder.net



[Webrecorder.net](https://webrecorder.net)

# Webrecorder.net

**Trusted by**

Webrecorder software is used by libraries, universities, governments, arts organizations, and more around the world.

Página <https://webrecorder.net/community>

# 1. Demonstrações básicas de gravação utilizando o ArchiveWeb.page





# Instalar a extensão ArchiveWeb.page

# Instalar a extensão ArchiveWeb.page no navegador Chrome

ArchiveWeb.page

https://archiveweb.page

## ARCHIVE WEBPAGE

Interactive browser-based web archiving from [Webrecorder](#).

The ArchiveWeb.page browser extension and standalone application allows you to capture web archives interactively as you browse!

After archiving your webpages, your archives can be viewed using [ReplayWeb.page](#) — no extension required!

For those who need to crawl whole websites with automated tools, check out [Browsertrix](#).

[View the ArchiveWeb.page User Guide](#)

[Install Extension from Chrome Web Store](#)

[Download ArchiveWeb.page Desktop App](#) - Latest version available for download!

---

<b>View Source</b>	<b>Language</b>
<a href="#">Extension and App</a>	<a href="#">English</a>
<a href="#">This Site</a>	<a href="#">Español</a>

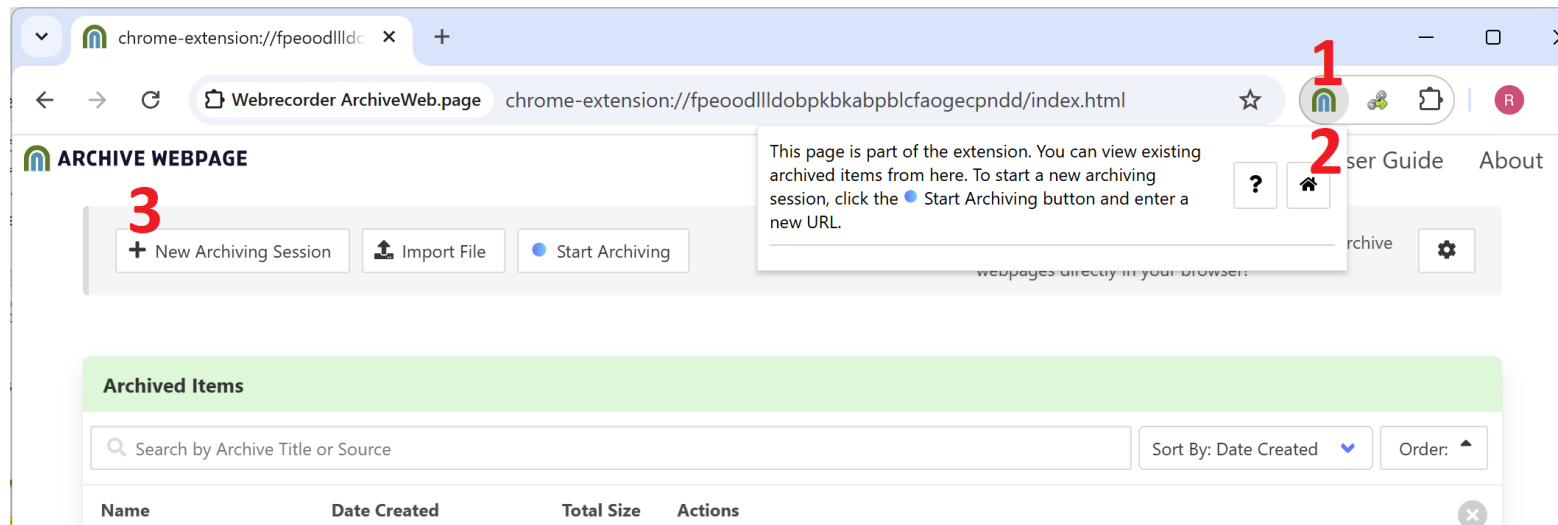


# Gravar uma página única

# Gravar uma página única

The screenshot shows the website of the Câmara Municipal de Coimbra. The browser address bar displays <https://www.cm-coimbra.pt>. The page features a sidebar menu on the left with the following items: Notícias, Serviços Online, Balcão Virtual, Agenda, Viver, Turismo, Estudar, Investir, Obras, Transparência, and Comunicação. The main content area is dominated by a large video player showing a scene with stained glass windows and the text "à volta da mesa de um café." Below the video, there is a section titled "Atividade municipal" with navigation arrows and a "ver +" link. This section contains three promotional cards: "Encontro Nacional de Provedores dos Municípios" on 6 MAIO 25, "viagem ao mundo da escrita: em tempo de camões" from 2 DE MAIO A 29 DE AGOSTO, and "Save the date" for 2 e 3 julho at Convento São Francisco. The footer contains a news ticker with the following items: "tagem à Rua Couraça da Estrela", "Coimbra recebe grande Regata Internacional Queima das Fitas no sábado", "Coimbra formou Laço Azul humano para lembrar a importância da proteção de crianças e jovens", and "Comissão de Gestão Integrada de Fogos Rurais de Coimbra reuniu ontem nos Paços do Co". The footer also includes logos for Coimbra+, CENTRO 2020, and the Câmara Municipal de Coimbra, along with social media icons and the text "© 2025 . Câmara Municipal de Coimbra . Política de privacidade".

Para começar o ArchiveWeb.page usar siga estes passos







# Gravar várias páginas

# Gravar várias páginas

The screenshot shows the website of the Câmara Municipal de Coimbra. The browser address bar displays <https://www.cm-coimbra.pt>. The main content area features a large aerial photograph of the city's ruins, with the text "Ouvi-la falar de histórias" overlaid. Below the main image is a section titled "Atividade municipal" with a navigation bar containing arrows and "ver +".

On the left side, there is a vertical menu with the following items:

- Notícias
- Serviços Online
- Balcão Virtual
- Agenda
- Viver
- Turismo
- Estudar
- Investir
- Obras
- Transparência
- Comunicação

At the bottom of the page, there is a footer with several news items and logos:

- Comissão de Gestão Integrada de Fogos Rurais de Coimbra reuniu ontem nos Paços do Concelho
- Casa da Cidadania da Língua inaugura no sábado mostra evocativa do Dia da Língua Portuguesa
- Nova edição do Coimbra BD já confirmada para os dias 24, 25 e 26 de abril de 2026
- Logos for COIMBRA, CENTRO 2020, and other partners.



# Gravar muitas páginas

# Gravar muitas páginas

The image shows a browser window displaying the website of the Câmara Municipal de Coimbra. The browser's address bar shows the URL <https://www.cm-coimbra.pt>. The website features a navigation menu on the left with categories such as Notícias, Serviços Online, Balcão Virtual, Agenda, Viver, Turismo, Estudar, Investir, Obras, Transparência, and Comunicação. The main content area is dominated by a large banner image of a historic town reflected in a river, with the text "ouço-o nos seus risos!" overlaid. Below the banner, there is a section titled "Atividade municipal" which contains a grid of four activity cards. The first card shows a meeting, the second is a poster for "Encontro Nacional de Provedores dos Municípios" on May 6, 2025, the third is a poster for "viagem ao mundo da escrita: em tempo de camões" from May 2 to August 20, 2025, and the fourth shows a group of people standing together. At the bottom of the page, there are several news snippets: "roteção de crianças e jovens", "Comissão de Gestão Integrada de Fogos Rurais de Coimbra reuniu ontem nos Paços do Concelho", "Casa da Cidadania da Língua inaugura no sábado mostra evocativa do Dia da Língua Portuguesa", and "Nova edição do Coimbra BD já confirmada para os dias 24, 25 e 26 de ab". The footer includes logos for Coimbra+, CENTRO 2020, and the Câmara Municipal de Coimbra, along with social media icons and a copyright notice for 2025.

Câmara Municipal de Coimbra

extension://fpeodllldobpkbap

4:50:54 05/02/2025

<https://www.cm-coimbra.pt>

Atualizar

Notícias

Serviços Online

Balcão Virtual

Agenda

Viver

Turismo

Estudar

Investir

Obras

Transparência

Comunicação

ouço-o nos seus risos!

Atividade municipal

ver +

roteção de crianças e jovens

Comissão de Gestão Integrada de Fogos Rurais de Coimbra reuniu ontem nos Paços do Concelho

Casa da Cidadania da Língua inaugura no sábado mostra evocativa do Dia da Língua Portuguesa

Nova edição do Coimbra BD já confirmada para os dias 24, 25 e 26 de ab

coimbra+ CENTRO 2020

CÂMARA MUNICIPAL COIMBRA

© 2025 . Câmara Municipal de Coimbra . Política de privacidade

Armazenar em formato  
normalizado: WARC

# Formato WARC

- ISO 28500:2017
- Independente de plataforma
- Reproduzido por software específico (geralmente designado *wayback*)

WACZ ficheiro comprimido e processado para facilitar a reproduzidos com o ReplayWeb.page

Gravar diretamente no Arquivo.pt  
com o *ArchivePageNow*

# ArchivePageNow para gravar no Arquivo.pt



The screenshot shows a web browser window with the URL `arquivo.pt/services/archivepagenow?l=pt`. The page features a dark blue header with the 'ARQUIVO.PT' logo and a 'Menu' button. The main content area has a yellow circular icon with a play symbol and the text 'ArchivePageNow Arquive páginas no Arquivo.pt'. Below this is a search form with the placeholder text 'Endereço da página' and a blue 'Arquivar' button. An example URL 'ex. www.fccn.pt' is provided. The text explains that the service allows archiving a page at the moment of request, and that archived content will be integrated into the Arquivo.pt collection. A blue link 'Saber mais' is also present.

Faculdade de Letras x ArchivePageNow: Arquive pági... x + 4:44:33 05/05/2025

arquivo.pt/services/archivepagenow?l=pt

Menu

ARQUIVO.PT

ArchivePageNow Arquive páginas no Arquivo.pt

Endereço da página

Arquivar

ex. www.fccn.pt

O serviço ArchivePageNow permite arquivar uma página no exato momento em que o utilizador faz o pedido.

Os conteúdos arquivados serão posteriormente integrados no acervo do Arquivo.pt.

[Saber mais](#)



[arquivo.pt/archivepagenow](http://arquivo.pt/archivepagenow)

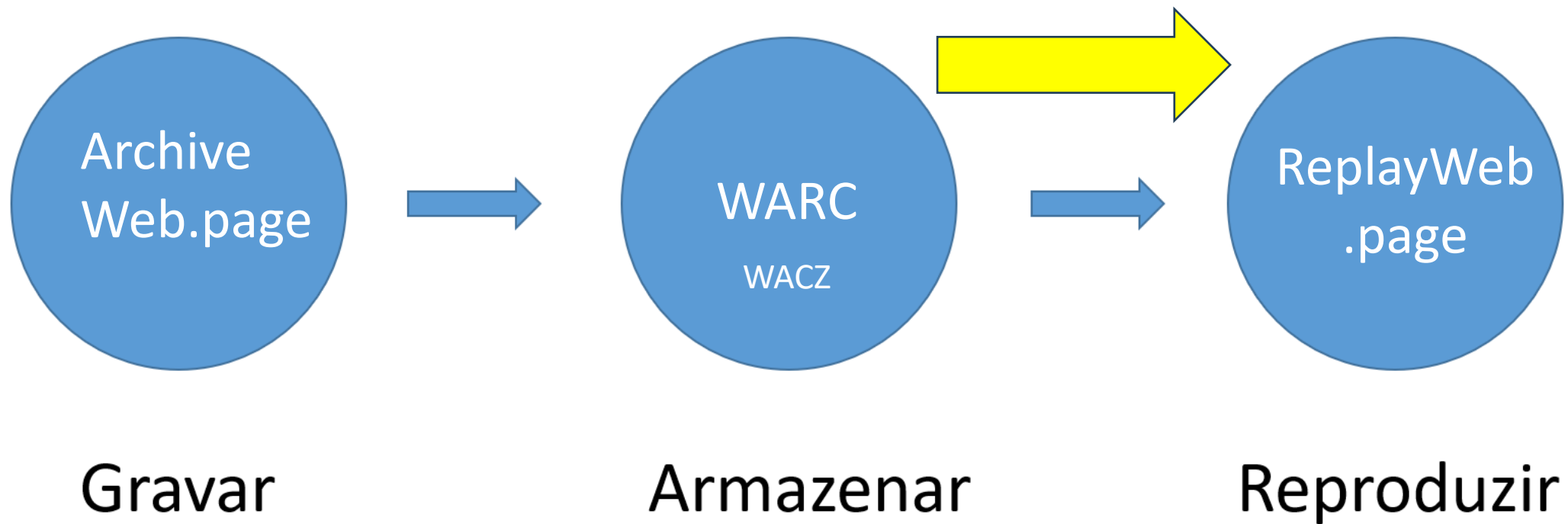
## 2. Mãos na massa – “Gravathon”

# 3. Partilhar as gravações

Reproduzir os conteúdos em qualquer tempo e lugar



# Arquivar a Web



# Reproduzir localmente os conteúdos gravados

The image shows a web browser window displaying the ReplayWebpage application. The browser's address bar shows the URL `replayweb.page`. The page content includes a header with the logo and navigation links, a main heading "Explore and Replay Interactive Archived Webpages Directly in your Browser. (Learn More)", and a "Load Web Archive" section. This section features a "Choose File..." button, a text input field with the placeholder "Enter a URL or click 'Choose File' to select a WARC, WACZ, CDX, or HAR file", and a green "Load" button. A file explorer window is overlaid on the browser, showing the directory structure of the local drive. The file explorer is titled "Abrir" and shows the path "utilizador > Ambiente de trabalho > MeuArquivoDaWeb". It contains a list of folders and files, with the file "01--gravar-página-única" selected. The file type is set to "WARC, WACZ, and HAR Files".

ReplayWebpage

replayweb.page

Gostaria de torná-lo o padrão no seu browser diário? [Como posso fazê-lo?](#)

Sim, defini-lo como browser predefinido

REPLAY WEBPAGE

User Docs About

Explore and Replay Interactive Archived Webpages Directly in your Browser. [\(Learn More\)](#)

Load Web Archive

Choose File... Enter a URL or click 'Choose File' to select a WARC, WACZ, CDX, or HAR file Load

Abrir

utilizador > Ambiente de trabalho > MeuArquivoDaWeb

Procurar em MeuArquivoDa...

Organizar Nova pasta

- Ambiente de tr
- 2
- 20250105-coi
- 20250117-pri
- 20250321-vid
- 20250401-wel
- IPL-tutorial
- MeuArquivo
- MeuArquivo
- Nova pasta
- Nova pasta
- reading1
- SharedFolder
- teste-box
- TrainingWG-l
- tutorial-powe

01--gravar-página-única

02--gravar-várias-páginas

03--gravar-muitas-páginas

Nome de ficheiro: 01--gravar-página-única WARC, WACZ, and HAR Files

Abrir Cancelar

to archive entire websites with automated crawling!

# Arquivar a Web



Faça como o Arquivo Municipal de Sines. Seja ativo na preservação dos conteúdos importantes.

# Envie as suas gravações

Short link para voucher do filesender.fccn.pt

[tinyurl.com/enviar2025](https://tinyurl.com/enviar2025)

# 4. Gravação automática com o Browsertrix-crawler



# Gravar sites inteiros com o Browsertrix-crawler

10:55:37 05/05/2025

localhost

localhost:9037



```
ricardo@PretoAsus: ~$  
  
docker run -p 9037:9037 -v $PWD/crawls:/crawls/ -it webrecorder/browsertrix-crawler crawl --url https://www.cm-coimbra.pt/ --screencastPort 9037  
  
docker run  
-p 9037:9037  
-v $PWD/crawls:/crawls/  
-it webrecorder/browsertrix-crawler crawl  
--url https://www.cm-coimbra.pt/  
--screencastPort 9037
```

## Gravar sites inteiros com o Browsertrix-crawler – Opções

11:24:21 05/05/2025

localhost

localhost:9037

```
ricardo@PretoAsus:~$ docker run -p 9037:9037 -v $PWD/crawls:/crawls/ -it webrecorder/browsertrix-crawler crawl --url https://www.cm-coimbra.pt/ --screenshotPort 9037
```



[tinyurl.com/browsertrix-coimbra](https://tinyurl.com/browsertrix-coimbra)

# Conclusão

- Grave conteúdos em formato normalizado com o ArchiveWeb.page
- Crie um arquivo da Web útil para si e para a sua organização
- Use o Arquivo.pt para perpetuar conteúdos da Web importantes

Mail

[contacto@arquivo.pt](mailto:contacto@arquivo.pt)

Inscreva-se na mailing list (enviamos 1 vez por mês)

[arquivo.pt/inscrever](http://arquivo.pt/inscrever)

# Concorrer ao Prémio anual Arquivo.pt!



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

# Obrigado!

[jornadas.fccn.pt](http://jornadas.fccn.pt)

[fccn.pt](http://fccn.pt)

