

Slides com notas
e informação adicional



APDSI
APRESENTA

CICLO DE WEBINARS

 **arquivo.pt** fct

**PRESERVAR A INFORMAÇÃO
PUBLICADA NA WEB**

4 SESSÕES
20,25,27 de MARÇO e 1 de ABRIL
15H - 16.30H
@ZOOM

APOIO INSTITUCIONAL:
FCCN serviços digitais fct | **fct** Fundação para a Ciência e a Tecnologia

Webinar 4 e conclusão do ciclo de webinars

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)



- Webinar 1 – 20 de março – Arquivo.pt: uma nova ferramenta para pesquisar o passado. Daniel Gomes, Gestor do Arquivo.pt
- Webinar 2 – 25 de março – Bem publicar, para bem preservar. Pedro Gomes, Encarregado pelas recolhas do Arquivo.pt
- Webinar 3 – 27 de março – Acesso e processamento automático de informação preservada da Web através de APIs. Vasco Rato, Web developer do Arquivo.pt
- Webinar 4 – 1 de abril – Arquivar a Web: faça-você-mesmo!. Ricardo Basílio, Curador digital do Arquivo.pt

- [Formação do Arquivo.pt com a APDSI \(página do Arquivo.pt\)](#)
- [APDSI promove o Ciclo de Webinars Arquivo.pt \(Notícia no site da APDSI\)](#)
- [Prémio Transformação digital 2024 da APDSI: lista de finalistas](#)
- [Arquivo.pt venceu o Prémio Transformação Digital 2024](#)



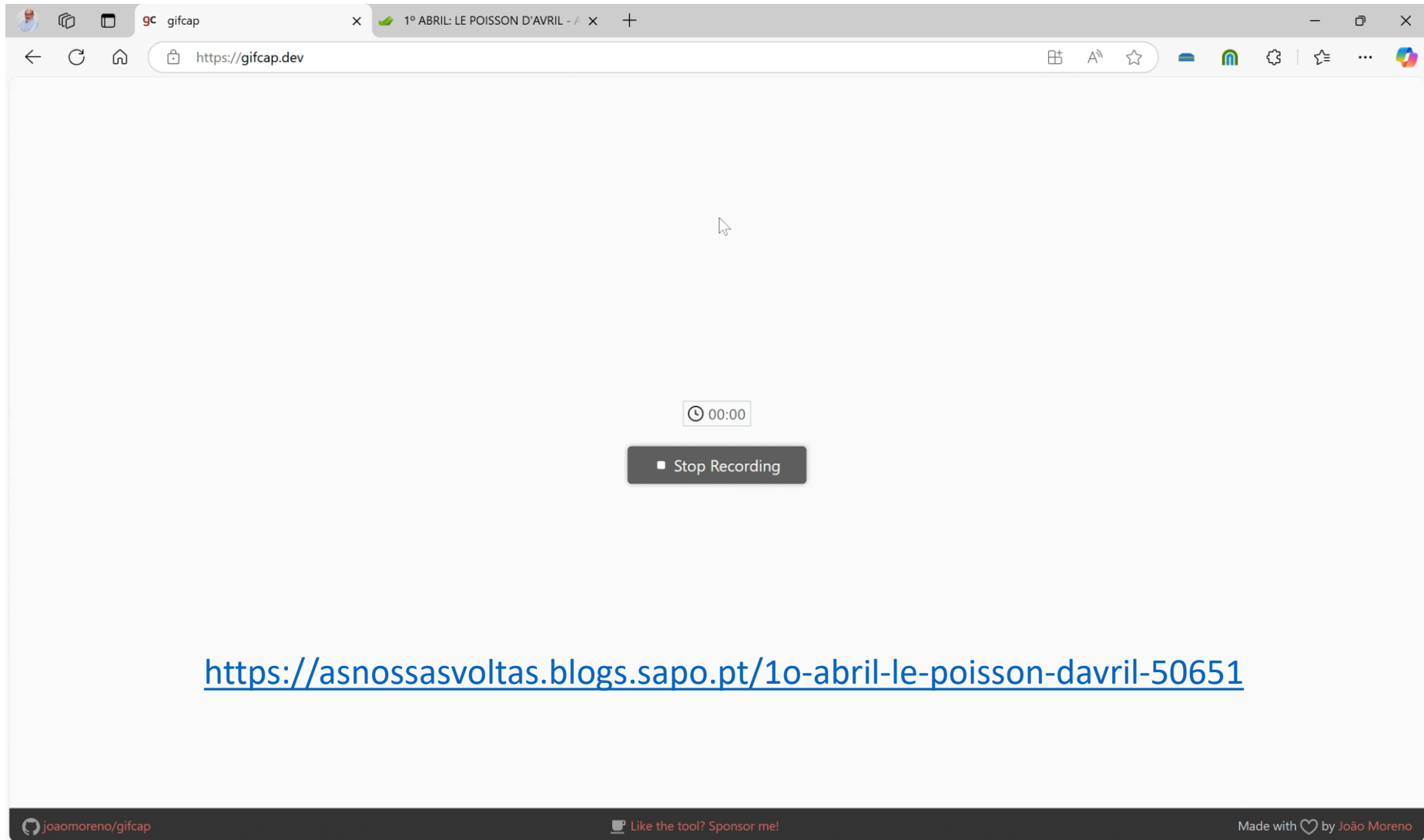
Arquivar-a-Web: faça-você-mesmo!

Ricardo Basílio, Curador Digital do Arquivo.pt

Objetivo

- Gravar, armazenar e reproduzir páginas Web no próprio computador num formato normalizado, utilizando ferramentas gratuitas do Webrecorder.net

Objetivo



<https://asnossasvoltas.blogs.sapo.pt/1o-abril-le-poisson-davril-50651>

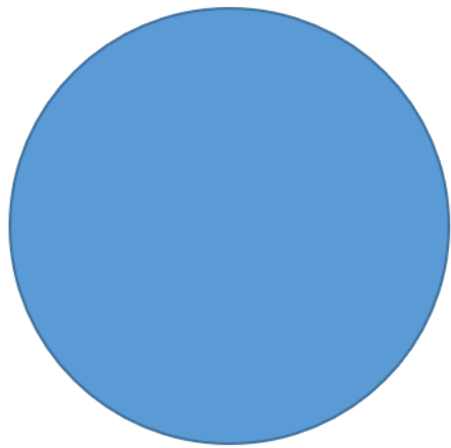
Agenda

Introdução

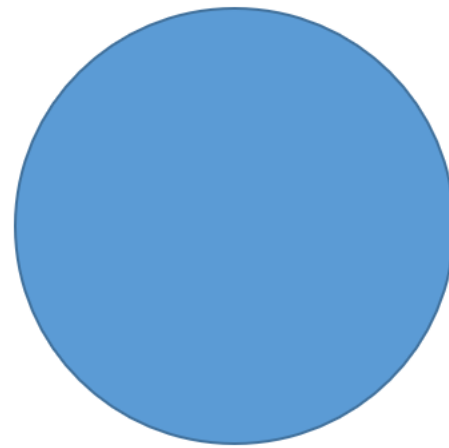
1. O Webrecorder.net. Porque razão utilizamos as suas ferramentas
2. Gravar utilizando a extensão ArchiveWeb.page
3. Armazenar num formato normalizado: WARC
4. Reproduzir conteúdos em qualquer tempo ou lugar (interoperável)
5. Gravar diretamente no Arquivo.pt

Conclusão

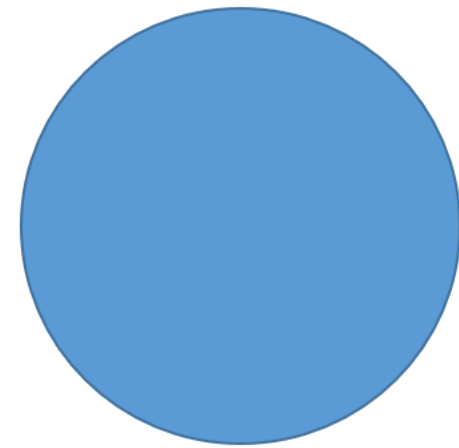
Preservação da Web



Gravar



Armazenar



Reproduzir

Preservação da Web



1 . Webrecorder.net

Porque razão utilizamos as suas ferramentas

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)

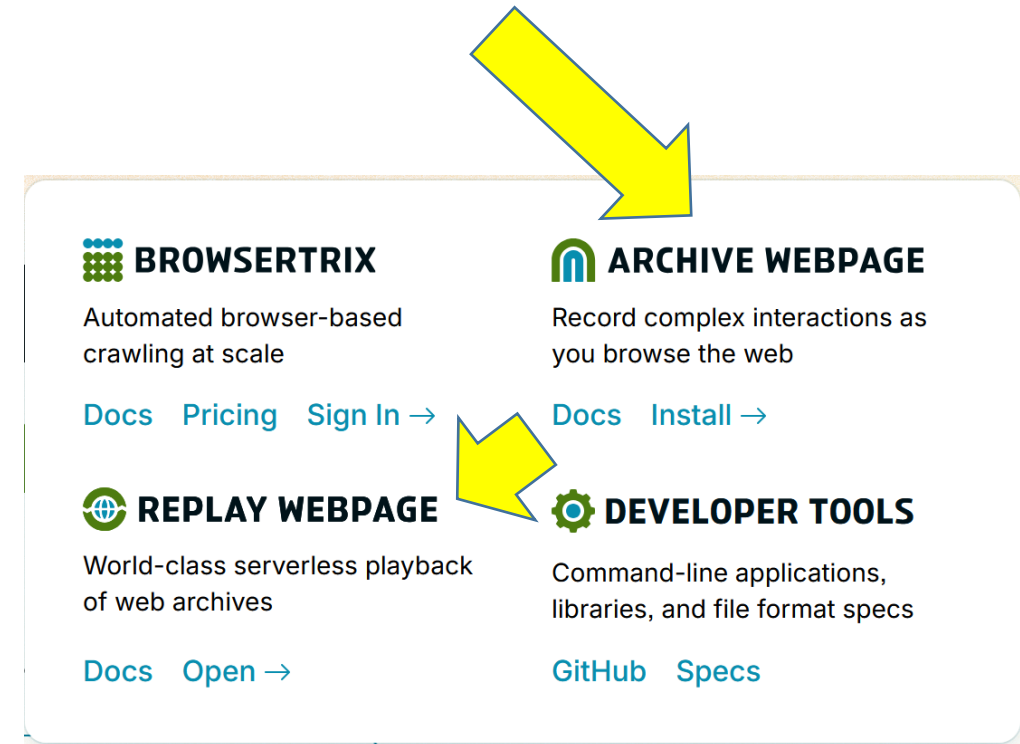
Nas ações de formação do Arquivo.pt valorizamos o que pode contribuir para a colaboração com a comunidade, independentemente das ferramentas utilizadas.





Presentemente, o Webrecorder.net tem as ferramentas que consideramos mais adequadas.

ArchiveWeb.page

Neste tutorial vamos usar ArchiveWeb.page na versão de extensão de browser. No final, para demonstrar como reproduzir os ficheiros gravados em ambiente totalmente local, instalamos o ArchiveWeb.page Desktop App (a razão desta opção, é estar a funcionar melhor do que o ReplayWeb.page)

O projeto Webrecorder.net tem um conjunto de ferramentas que todos podem usar para gravar e reproduzir páginas Web.

 BROWSERTRIX Automated browser-based crawling at scale Docs Pricing Sign In →	 ARCHIVE WEBPAGE Record complex interactions as you browse the web Docs Install →
 REPLAY WEBPAGE World-class serverless playback of web archives Docs Open →	 DEVELOPER TOOLS Command-line applications, libraries, and file format specs GitHub Specs

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

Webrecorder.net

O Webrecorder.net tem o apoio de entidades internacionais importantes no domínio do Web archiving.

O Arquivo.pt, por exemplo, usa o software pywb do Webrecorder para reproduzir as páginas Web, usa-o para o novo serviço de gravação de páginas na hora SavePageNow.

O ArchiveWeb.page é utilizado em recolhas de alta qualidade de websites. O que as outras ferramentas não conseguem gravar, o ArchiveWeb.page provavelmente conseguirá.

O Webrecorder.net é utilizado para dar formação sobre preservação da Web.

Trusted by



ARQUIVO.PT



Perma.cc ∞



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

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

2. Gravar utilizando a extensão [ArchiveWeb.page](#)



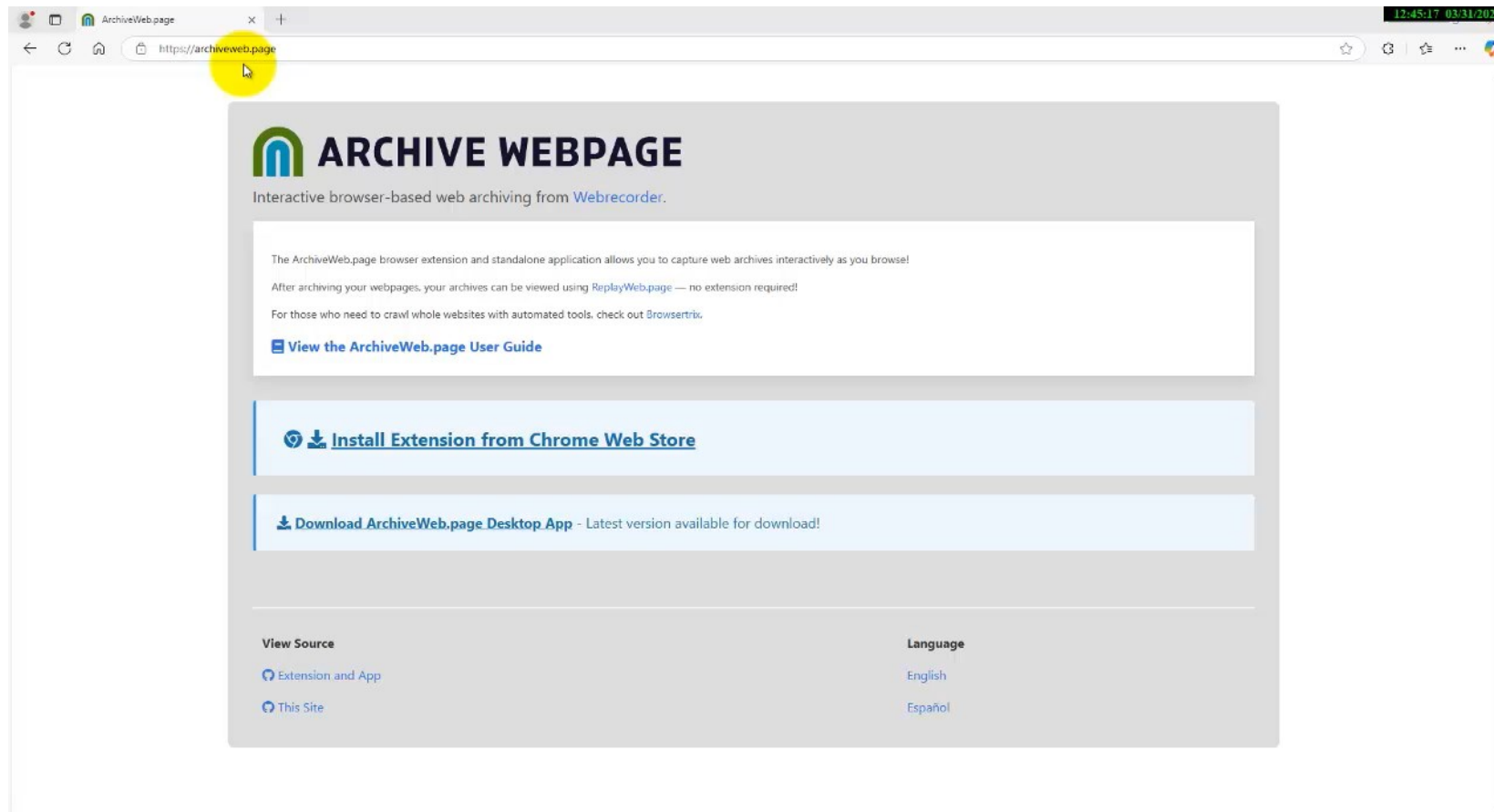
Tutorial – casos de uso do ArchiveWeb.page



Instalar a extensão [ArchiveWeb.page](#)

<https://youtu.be/LbF5cCN0eql?si=xZxNLm9gkgyagFRI&t=1065>

Instalar a extensão ArchiveWeb.page no navegador Chrome



<https://youtu.be/LbF5cCN0eql?si=xZxNLm9gkgyagFRI&t=1065>

Tutorial – casos de uso do ArchiveWeb.page



Gravar uma página única

https://youtu.be/LbF5cCN0eql?si=dFwdGxXguoqQ_-kb&t=1225

Gravar uma página única

APDSI - Associação para a Promoção da Saúde e do Bem-Estar

APDSI Notícias Áreas Grupos Iniciativas Produtos

Associe-se!

PRÉMIO TRANSFORMAÇÃO DIGITAL 2024 APDSI

POWERED BY ORACLE

Notícias da APDSI

- 15.º aniversário da SPMS | Digitalizar a Saúde, Transformar o Futuro**
26 Março 2025
- APDSI promove o Ciclo de Webinars Arquivo.pt**
7 Março 2025
- APDSI esteve presente no lançamento do 3º volume do livro "Raparigas na Ciência"**
7 Março 2025
- APDSI assina carta conjunta de defesa da adoção imediata do Esquema Europeu de Certificação de**

Próximos eventos

ABR 15:00 - 16:30
1 APDSI promove o Ciclo de Webinars Arquivo.pt

Ver calendário

Patrocinadores

Platina
Google

Ouro
aws INCM

Seja também nosso patrocinador

Sócios Coletivos

INTEGRITY

https://youtu.be/LbF5cCN0eqI?si=dFwdGxXguoqQ_-kb&t=1225



Gravar várias páginas

<https://youtu.be/LbF5cCN0eql?si=tlp61W23MKCzXA-g&t=1391>

Gravar várias páginas

APDSI - Associação para a Promoção de Serviços e Produtos de Software

APDSI Notícias Áreas Grupos Iniciativas Produtos

PRÉMIO TRANSFORMAÇÃO DIGITAL 2024

APDSI

POWERED BY ORACLE

Notícias da APDSI

- 15.º aniversário da SPMS | Digitalizar a Saúde, Transformar o Futuro**
26 Março 2025 0
- APDSI promove o Ciclo de Webinars Arquivo.pt**
7 Março 2025 0
- APDSI esteve presente no lançamento do 3º volume do livro "Raparigas na Ciência"**
7 Março 2025 0
- APDSI assina carta conjunta de defesa da adoção imediata do Esquema Europeu de Certificação de**

Próximos eventos

ABR 15:00 - 16:30
1 APDSI promove o Ciclo de Webinars Arquivo.pt

Ver calendário

Patrocinadores

Platina
Google

Ouro
aws INCM

Seja também nosso patrocinador

Sócios Coletivos

INTEGRITY

<https://youtu.be/LbF5cCN0eql?si=tlp61W23MKCzXA-g&t=1391>



Gravar muitas páginas

<https://youtu.be/LbF5cCN0eql?si=I7SYYZ21Milwb06t&t=1568>

Gravar muitas páginas

The screenshot shows the APDSI website's news section. The navigation bar includes 'APDSI', 'Notícias' (highlighted), 'Áreas', 'Grupos', 'Iniciativas', 'Produtos', and 'Associe-se!'. The main content area features a 'Filter by Categories' dropdown and a 'Show all' link. Two news items are displayed:

- 31 Março 2025**
Futures '25 Conference on Artificial Intelligence in Education
De APDSI, Áreas de Intervenção, Inteligência Artificial, Parceiros, Transformação Digital
APDSI divulga a conferência "Futures '25 Conference on Artificial Intelligence in Education". Quer seja apaixonado pelo papel da tecnologia na educação, curioso por experiências de aprendizagem [...]
- 26 Março 2025**
15.º aniversário da SPMS | Digitalizar a Saúde, Transformar o Futuro
De APDSI, Áreas de Intervenção, Parceiros, Saúde, Transformação Digital
A APDSI divulga a Conferência Comemorativa do 15.º Aniversário da

On the right side, there is a 'Próximos eventos' section with a calendar view showing an event on '1 ABR' from 15:00 - 16:30: 'APDSI promove o Ciclo de Webinars Arquivo.pt'. A 'Ver calendário' link is provided below.

<https://youtu.be/LbF5cCN0eql?si=I7SYYZ21Milwb06t&t=1568>



Gravar redes sociais?

<https://youtu.be/LbF5cCN0eql?si=MN2Zgr7SvNQ2PQex&t=1796>

Gravar redes sociais – Facebook – sem login – “de fora para dentro”

facebook

Email or phone Password Log in Forgotten account?

CONFERÊNCIA - DEBATE
COMO REPENSAR A EDUCAÇÃO
COM A INTELIGÊNCIA ARTIFICIAL?
PAVILHÃO DO CONHECIMENTO
CIÊNCIA VIVA
22 NOV
14H

APDSI

APDSI Portugal
2.2K likes • 2.3K followers

Posts About Photos Videos

Intro
Organização Sem Fins Lucrativos

Page - Non-profit organisation

Rua Alexandre Cabral, Nº 2C - Loja A,
Lisbon, Portugal

21 751 0762

APDSI Portugal
5d · 🌐

APDSI divulga a Conferência Comemorativa do 15.º Aniversário da SPMS, que terá lugar hoje, dia 26 de março de 2025, entre as 09:00 e as 17:00 horas, no Pátio da Galé, em Lisboa.

Para assinalar esta data, o evento reunirá especialistas, líderes, gestores, profissionais e parceiros para debater alguns dos principais temas do setor da saúde. A Professora Doutora Maria Helena Monteiro, atual Presidente da Direção da APDSI, será uma das intervenientes.

Há 15 anos, a SPMS coloca ... See more

Log in or sign up for Facebook to connect with friends, family and people you know.





Log in or Create new account

Gravar redes sociais – Facebook – com login - para uso restrito

Facebook extension://fpeodllldobpkbap x | +

https://www.facebook.com/apdsi.portugal/

Pesquisar no Facebook



APDSI Portugal
2,2 mil gostos • 2,3 mil seguidores

[Contacta-nos](#) [Enviar mensagem](#) [Gosto](#)

[Publicações](#) [Sobre](#) [Menções](#) [Críticas](#) [Seguidores](#) [Fotos](#) [Mais](#)

Apresentação

Organização Sem Fins Lucrativos

- [Página](#) · Organização sem fins lucrativos
- Rua Alexandre Cabral, Nº 2C - Loja A, Lisbon, Portugal
- 21 751 0762

Publicações

[Filtros](#)

APDSI Portugal
5 d · [Publicar](#)

APDSI divulga a Conferência Comemorativa do 15.º Aniversário da SPMS, que terá lugar hoje, dia 26 de março de 2025, entre as 09:00 e as 17:00 horas, no Pátio da Galé, em Lisboa.

Para assinalar esta data, o evento reunirá especialistas, líderes, gestores, profissionais e parceiros para debater alguns dos principais temas do setor da saúde. A Professora Doutora Maria Helena

Gravar redes sociais – Instagram

Home | Politécnico de Lisboa x | IPL-Politécnico de Lisboa (@ipl) x | chrome-extension://fpeodllldc x +

Webrecorder ArchiveWeb.page chrome-extension://fpeodllldobpkbkabpblcfaogecpndd/index.html
















ARCHIVE WEB.PAGE User Guide About

+ New Archiving Session Import File Start Archiving

The ArchiveWeb.page Extension allows you to archive webpages directly in your browser!

Archived Items

Search by Archive Title or Source Sort By: Date Created Order: ^

Name	Date Created	Total Size	Actions
04-demo-redes-sociais-instagram	7/1/2024, 3:48:37 PM	0 B	  
03-demo-muitas-paginas	7/1/2024, 12:16:56 PM	27.5 MB	  
02-demo-varias-paginas	7/1/2024, 11:46:03 AM	14.3 MB	  
01-demo-pagina-unica	7/1/2024, 11:25:31 AM	15.9 MB	  
My Archiving Session	7/1/2024, 10:49:08 AM	0 B	  

Windows taskbar: Procurar, File Explorer, Firefox, Chrome, Edge, etc.



Gravar redes sociais?

Vídeos no Youtube

<https://youtu.be/LbF5cCN0eql?si=l0oEUdFe0VmWqfK-&t=2236>

Gravar redes sociais? – Vídeos no Youtube

The screenshot shows a web browser window displaying the APDSI website. The page title is "CONCLUSÕES | Conferência "A Regulação da Inteligência Artificial em Portugal"". The date is "8 Outubro 2024". The main image shows a conference room with a speaker at a podium and an audience. Below the image, there is a text block stating: "A APDSI, no âmbito do seu Grupo de Missão "Regulação Digital", realizou a Conferência "A Regulação da Inteligência Artificial em Portugal", no passado dia 19 de setembro, cujas conclusões estamos agora a difundir." Below this text, there are four multimedia links: "Conclusões" (with a PDF icon), "Multimédia: Sessão de abertura", "Multimédia: Painel 1", and "Multimédia: Encerramento". On the right side, there is a sidebar with the heading "A APDSI" and a description: "Entidade sem fins lucrativos, de utilidade pública e ONGD que mobiliza sócios individuais e coletivos em torno de causas orientadas à transformação digital." Below this, there is a section titled "Últimas Notícias" with three items: "Futures '25 Conference on Artificial Intelligence in Education" (31 Março 2025), "15.º aniversário da SPMS | Digitalizar a Saúde, Transformar o Futuro" (26 Março 2025), and "APDSI promove o Ciclo de Webinars Arquivo.pt" (7 Março 2025). At the bottom of the sidebar, there is a "REDE APDSI" logo.

<https://youtu.be/LbF5cCN0eql?si=l0oEUdFe0VmWqfK-&t=2236>

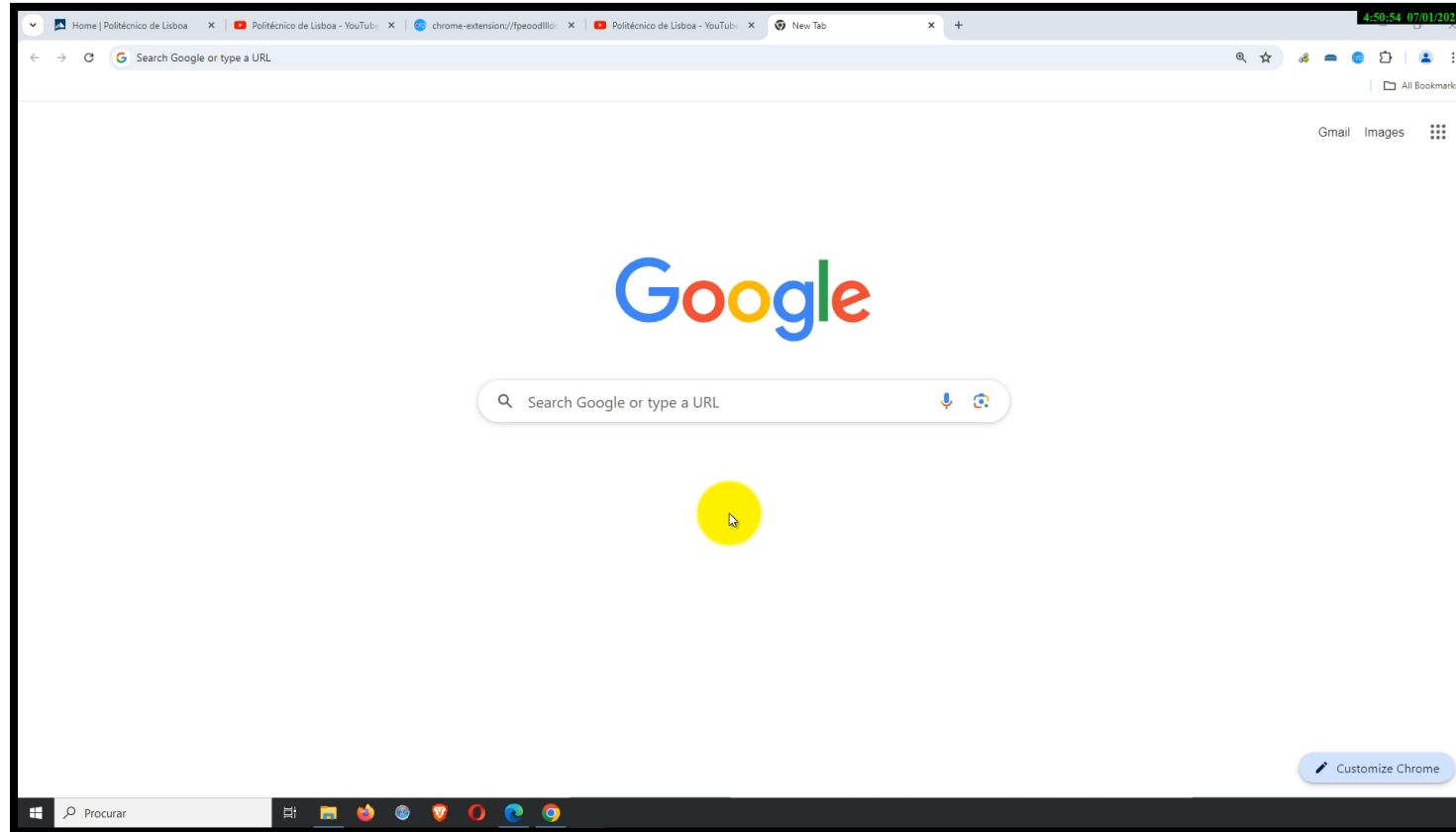


Gravar vídeos com sucesso

Exemplo do site “O corvo.pt”

https://youtu.be/LbF5cCN0eql?si=9z0L1ZmOIK0WMe_E&t=2460

Gravar vídeos com sucesso



https://youtu.be/LbF5cCN0eqI?si=9z0L1ZmOIK0WMe_E&t=2460



Quando era possível gravar o Twitter

https://youtu.be/LbF5cCN0eqI?si=-0_l1iPj8pHyETMZ&t=2586

Quando era possível gravar o Twitter

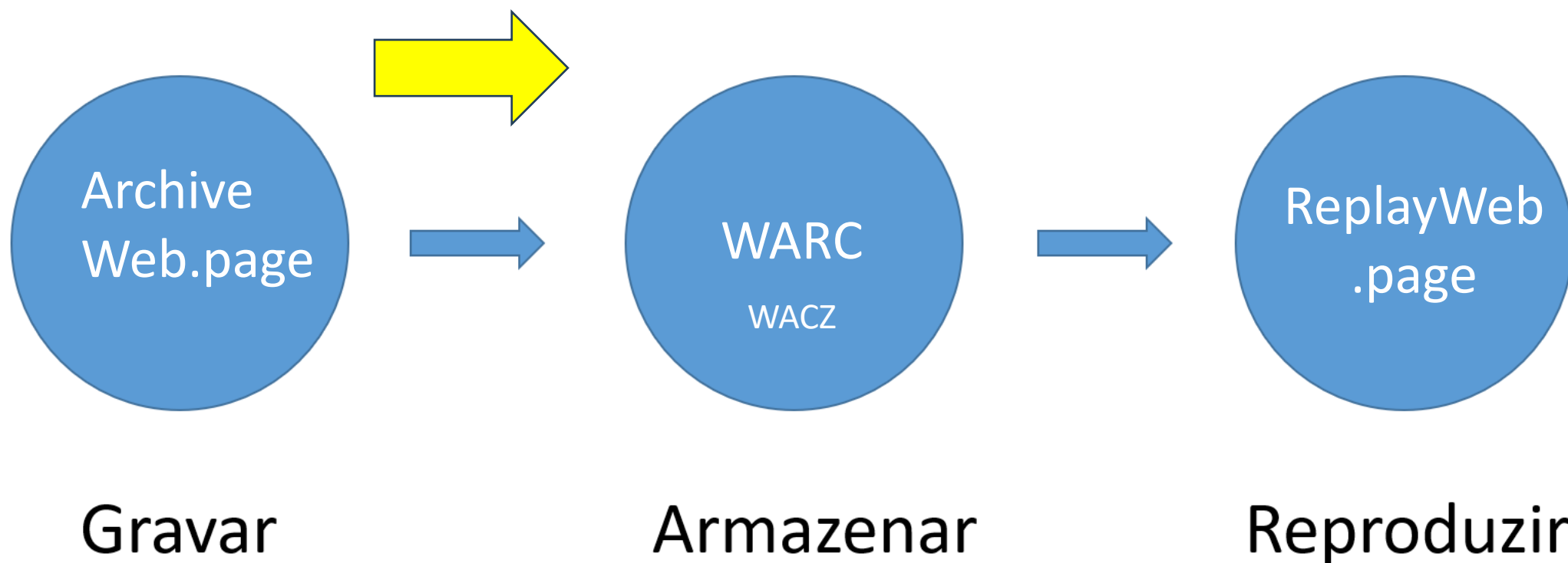
The screenshot displays the ArchiveWebPage extension interface. At the top, there are buttons for '+ Create New...', 'Import Archive...', and 'Start Recording...'. Below this is a section titled 'Current Web Archives' with a search bar and sorting options. The main content is a table with the following data:

Title	Date Created	Total Size	Actions	Sharing (via IPFS)
demo1-gravar muitas páginas	08/06/2022, 18:29:10	10.2 MB	[Icons]	[Start Sharing]
demo1-gravar várias páginas	08/06/2022, 17:57:25	5.58 MB	[Icons]	[Start Sharing]
demo1-gravar uma página	08/06/2022, 17:22:05	3.97 MB	[Icons]	[Start Sharing]
My Web Archive	08/06/2022, 16:47:04	0 B	[Icons]	[Start Sharing]

The system tray at the bottom indicates the temperature is 24°C, the location is Ger. Ensolado, and the date and time are 08/06/2022, 19:06.

https://youtu.be/LbF5cCN0eqI?si=-0_l1iPj8pHyETMZ&t=2586

Preservação da Web



<https://youtu.be/LbF5cCN0eql?si=ljJ2TSF87T6xQfHg&t=2680>

3. Armazenar em formato normalizado: WARC

<https://youtu.be/LbF5cCN0eql?si=ljJ2TSF87T6xQfHq&t=2680>

Formato WARC

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

Porque utilizamos o Webrecorder.net?

Razão principal: Porque grava os conteúdos no formato WARC.

O ArchiveWeb.page, por exemplo, grava os conteúdos e exporta-os no formato WARC. Esta é a sua capacidade mais importante, pelas razões seguintes:

- trata-se de um formato normalizado - ISO 28500:2017

- independente de plataforma

- reproduzível (geralmente designado *wayback*)

O resto vem por acréscimo. O formato WACZ, por exemplo, um formato preparado especificamente para o ArchiveWeb.page e o ReplayWeb.page lerem. Se o descomprimos poderemos ver que contém um ficheiro WARC e índices dos endereços gravados. A base de tudo, mais importante, é o ficheiro WARC que é atualmente a norma para os arquivos da Web.

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

Armazenar num formato normalizado – WARC ou em WACZ (comprimido)

extension://fpeodllldobpkbkabp x Descarregamentos x | +

extension://fpeodllldobpkbkabplcfaogecpndd/index.html





























ARCHIVE WEBPAGE User Guide About


+ New Archiving Session Import File Start Archiving

The ArchiveWeb.page Extension allows you to archive webpages directly in your browser!











Archived Items

Search by Archive Title or Source Sort By: Date Created Order: ^

Name	Date Created	Total Size	Actions	
06-gravar-youtube	31/03/2025, 23:27:32	200 MB	  	
05 -gravar redes sociais - Facebook com login	31/03/2025, 22:55:37	58.9 MB	  	
04 -gravar redes sociais - Facebook	31/03/2025, 22:30:07	16.2 MB	  	
03 -gravar muitas páginas	31/03/2025, 22:00:53	19.8 MB	  	
02 -gravar várias páginas	31/03/2025, 21:29:01	11.2 MB	  	
01 -gravar página única	31/03/2025, 13:20:06	5.2 MB	  	
My Archiving Session	31/03/2025, 12:46:13	0 B	  	



Armazenar num formato normalizado – Como é um ficheiro WARC por dentro?

Nome	Data de modificação	Tipo	Tamanho
 01-demo-pagina-unica	01/07/2024 18:18	Ficheiro WACZ	12 692 KB
 02-demo-varias-paginas	01/07/2024 18:18	Ficheiro WACZ	10 338 KB
 03-demo-muitas-paginas	01/07/2024 18:20	Ficheiro WACZ	68 044 KB
 04-demo-redes-sociais-instagram	01/07/2024 18:20	Ficheiro WACZ	38 404 KB
 05-demo-redes-sociais-youtube-videos	01/07/2024 18:20	Ficheiro WACZ	15 671 KB
 01-demo-pagina-unica	01/07/2024 18:18	Ficheiro WARC	12 639 KB
 02-demo-varias-paginas	01/07/2024 18:19	Ficheiro WARC	10 194 KB
 03-demo-muitas-paginas	01/07/2024 18:20	Ficheiro WARC	67 238 KB
 04-demo-redes-sociais-instagram	01/07/2024 18:20	Ficheiro WARC	36 847 KB
 05-demo-redes-sociais-youtube-videos	01/07/2024 18:20	Ficheiro WARC	15 502 KB

4. Reproduzir os conteúdos em qualquer tempo e lugar (interoperável)



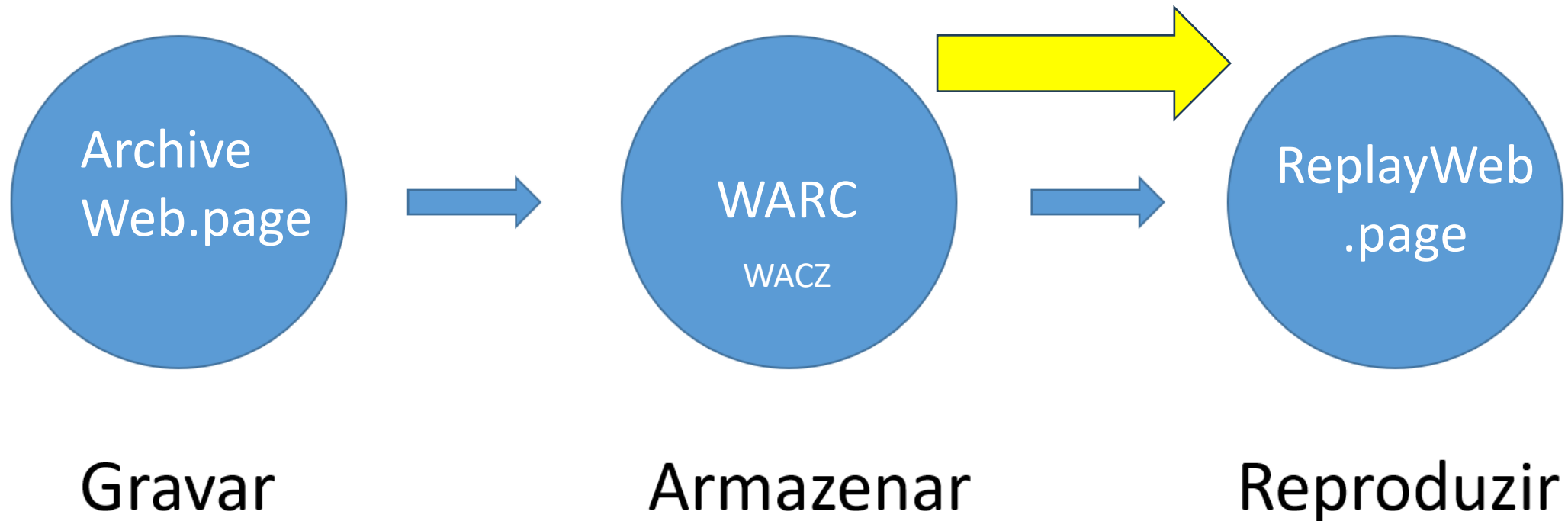
https://youtu.be/LbF5cCN0eql?si=5c_mVhbdEeFZ60DC&t=3089

Preservação da Web

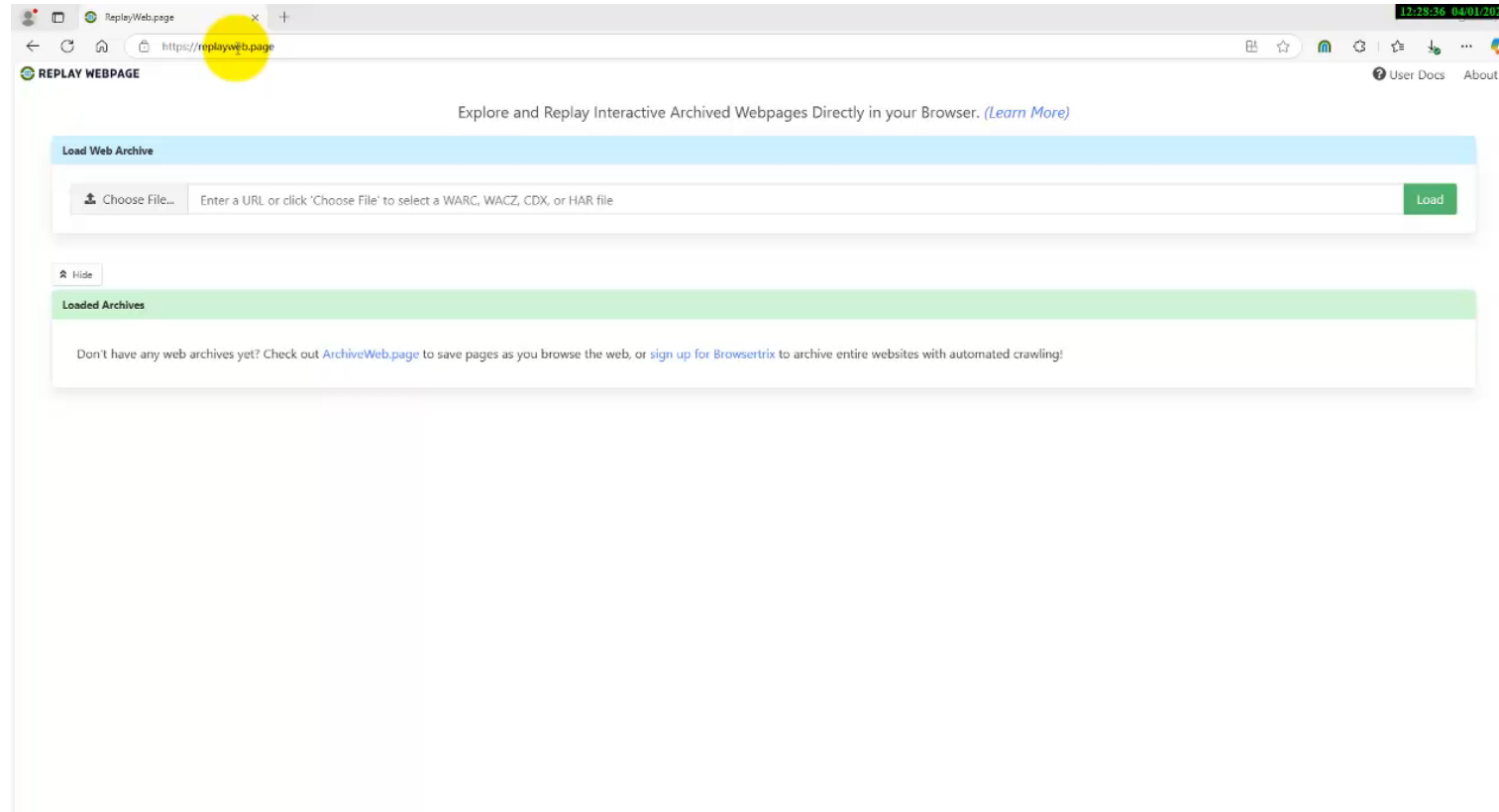
Gravar, Armazenar e Reproduzir são, de forma simplificada, as três fases da preservação (poderíamos falar de outras, tal como a seleção, controle de qualidade, etc)

O ArchiveWeb.page cumpre as três fases: 1) grava 2) gera um WARC e um WACZ para armazenar em qualquer lugar 3) reproduz o conteúdo.

É isso que os arquivos da web também fazem. Por isso é muito adequado utilizar o ArchiveWeb.page



Reproduzir localmente os conteúdos gravados



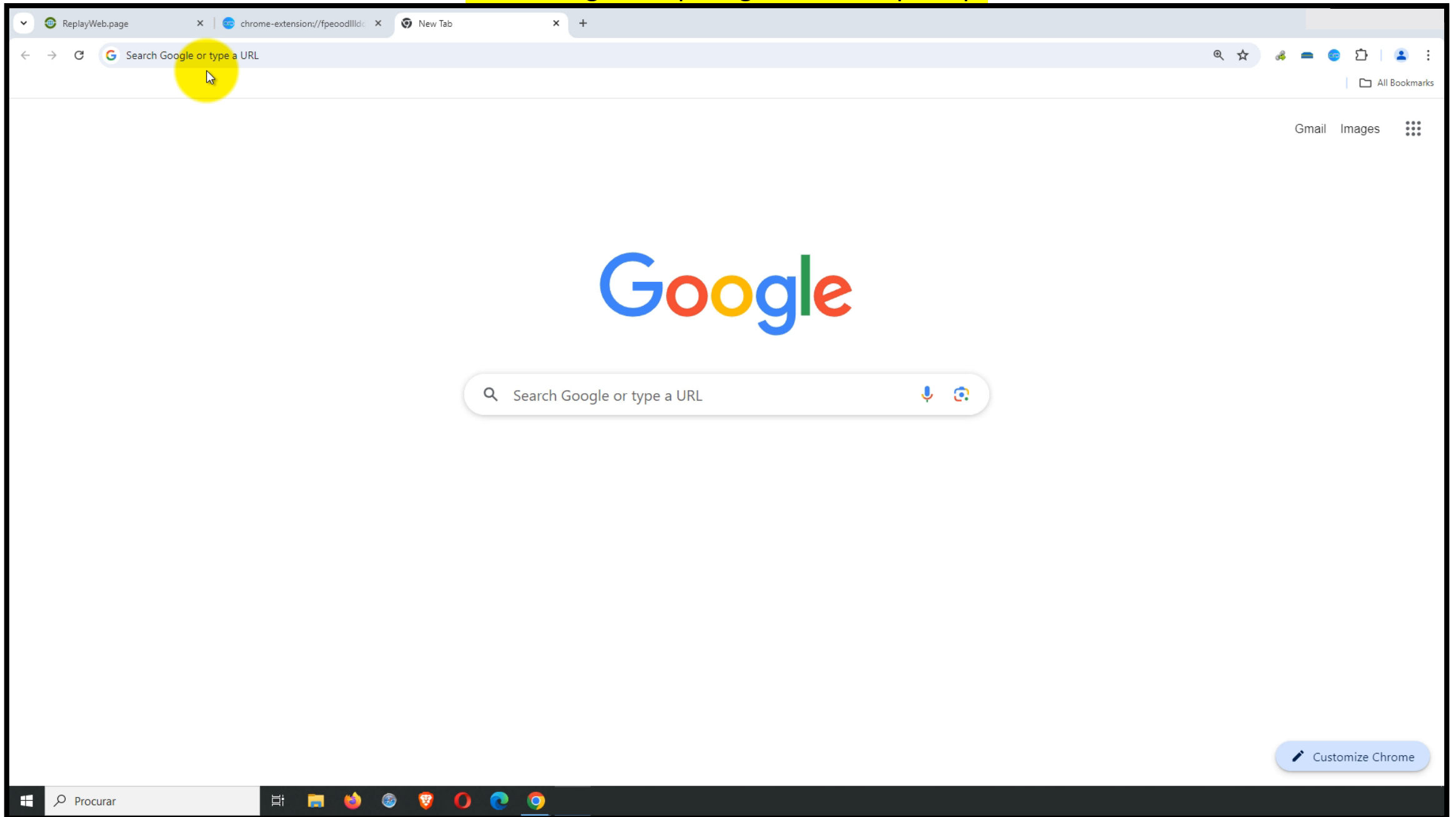
https://youtu.be/LbF5cCN0eql?si=5c_mVhbdEeFZ60DC&t=3089



O Arquivo Municipal de Sines, desde 2020, grava conteúdos Web de interesse para a memória e a história do Município. Usa o Webrecorder ArchiveWeb.page e o serviço Conifer que funciona de forma semelhante. Envia uma cópia dos ficheiros WARC e WACZ obtidos para o Arquivo.pt.

5. Gravar diretamente no Arquivo.pt
com o ArchivePageNow

ArchivePageNow para gravar no Arquivo.pt



arquivo.pt/archivepagenow

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

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

arquivo.pt/inscrever

Concorrer ao Prémio anual Arquivo.pt!



arquivo.pt/premios