

CAFÉ
COM O
ARQUIVO.PT





Arquivar um website: Faça você mesmo!

Ricardo Basílio, curador digital

15 de abril, 15 às 16 horas



Objetivos:

- Gravar páginas Web com o **Webrecorder.io** e guardar o conteúdo no formato normalizado WARC
- Conhecer os requisitos para utilizar o **Browsertrix** na recolha de um site inteiro



Agenda:

- Demo **Webrecorder.io**
- Demo **Browsertrix**
- Partilha de caso de uso
- Debate e questões



O Arquivo.pt

Preserva milhões de ficheiros recolhidos da web **desde 1996** e disponibiliza um **serviço público** de pesquisa sobre esta informação.

WR Webrecorder | Homepage

+

←

→

↺

🏠

🔒

https://webrecorder.io

90%

⋮

🔒

★

📖

📄

👤

🔄

☰

WR

About

Report Bug

Help

Support Us

Sign Up

Login

WR

pelicanbomb / Pelican Bomb / Art Review

About Report Bug Help Support Us Sign Up Login

http://pelicanbomb.com/art-review 11/26/2018, 11:46:59 AM

1 Art Review - Pelican Bomb

2 Art Review Archives: 2011 - Pelican Bomb

3 Finding and Losing New Orleans - By the Pelican Bomb

4 Amy Mackie - Pelican Bomb

5 Untitled Document

PELICAN BOMB

Art Review What We Do Who We Are Calendar Get Involved Contact

ESSAY

The Future of Arts Criticism: Part 1

ESSAY

A Dissonance of Parts and People: Publishing from the Provinces

Art Review - Pelican Bomb

THE FUTURE OF ARTS CRITICISM: PART 2

NYC BIKERIE AZIZ, MARJORIE RAWLE

NOTE

NOTE

REVIEW

Events

Exhibitions

CSA Shop Page

Critic in Residence

What We Do

Who We Are

Get Involved

News

Pelican Bomb

Pelican Bomb was an arts organization based in New Orleans. From 2011 to 2018, its programs included the Art Review, an online publication about contemporary art in South Louisiana; exhibitions and public projects across the city; a Community Supported Art initiative offering affordable editioned artworks to local artists; and a residency for art critics from across the country.

This is an archive of Pelican Bomb's website as of November 2018.

Art Review

1 Art Review - Pelican Bomb

2 Art Review Archives: 2011 - Pelican Bomb

4 Amy Mackie - Pelican Bomb

Versão on line – a mais rápida!

Versão desktop – a mais consistente!

Webrecorder.io is a web archiving service to collect and revisit web pages.

Webrecorder creates an interactive copy of any web page that you browse, including content revealed by your interactions such as playing video and audio, scrolling, clicking buttons, and so forth.

Create a Free Account

Existing Users Login

Webrecorder.io offers free accounts with 5GB of storage. Get more and contribute to Webrecorder's development by becoming a supporter.

Don't want to register? Download Desktop App to collect and access archived web pages on your own computer, no account necessary.

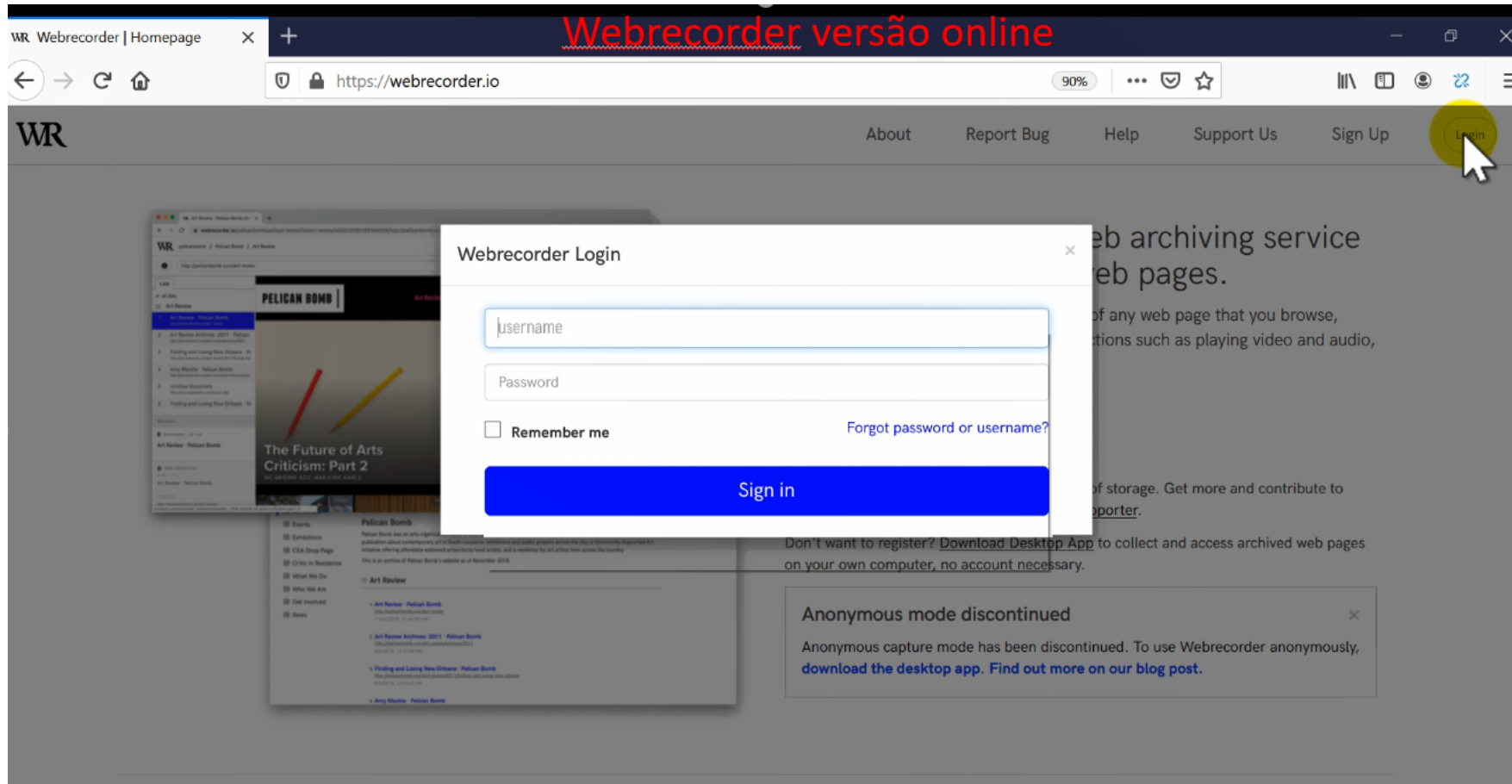
Anonymous mode discontinued

Anonymous capture mode has been discontinued. To use Webrecorder anonymously, download the desktop app. Find out more on our blog post.

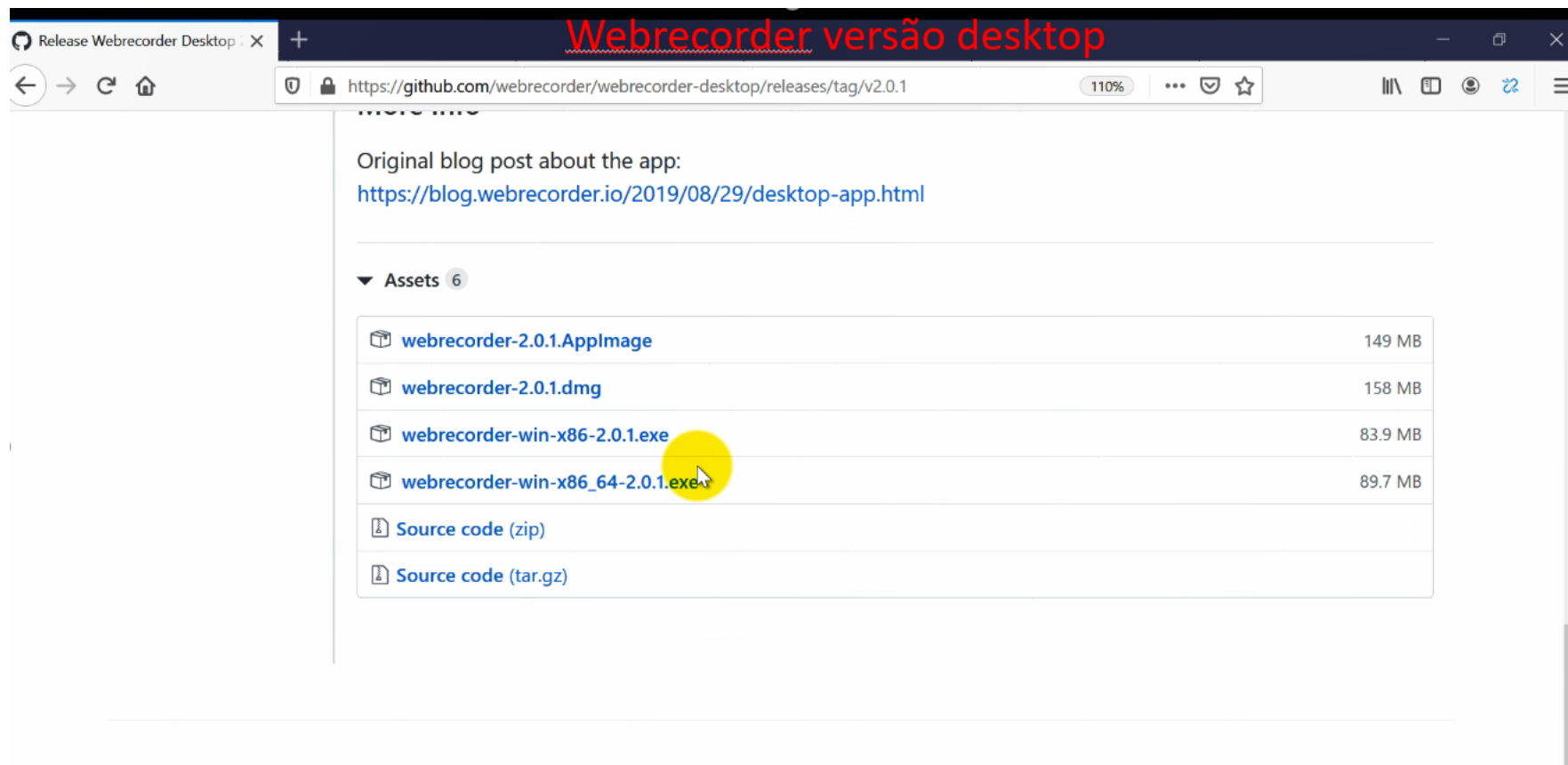


Webrecorder.io

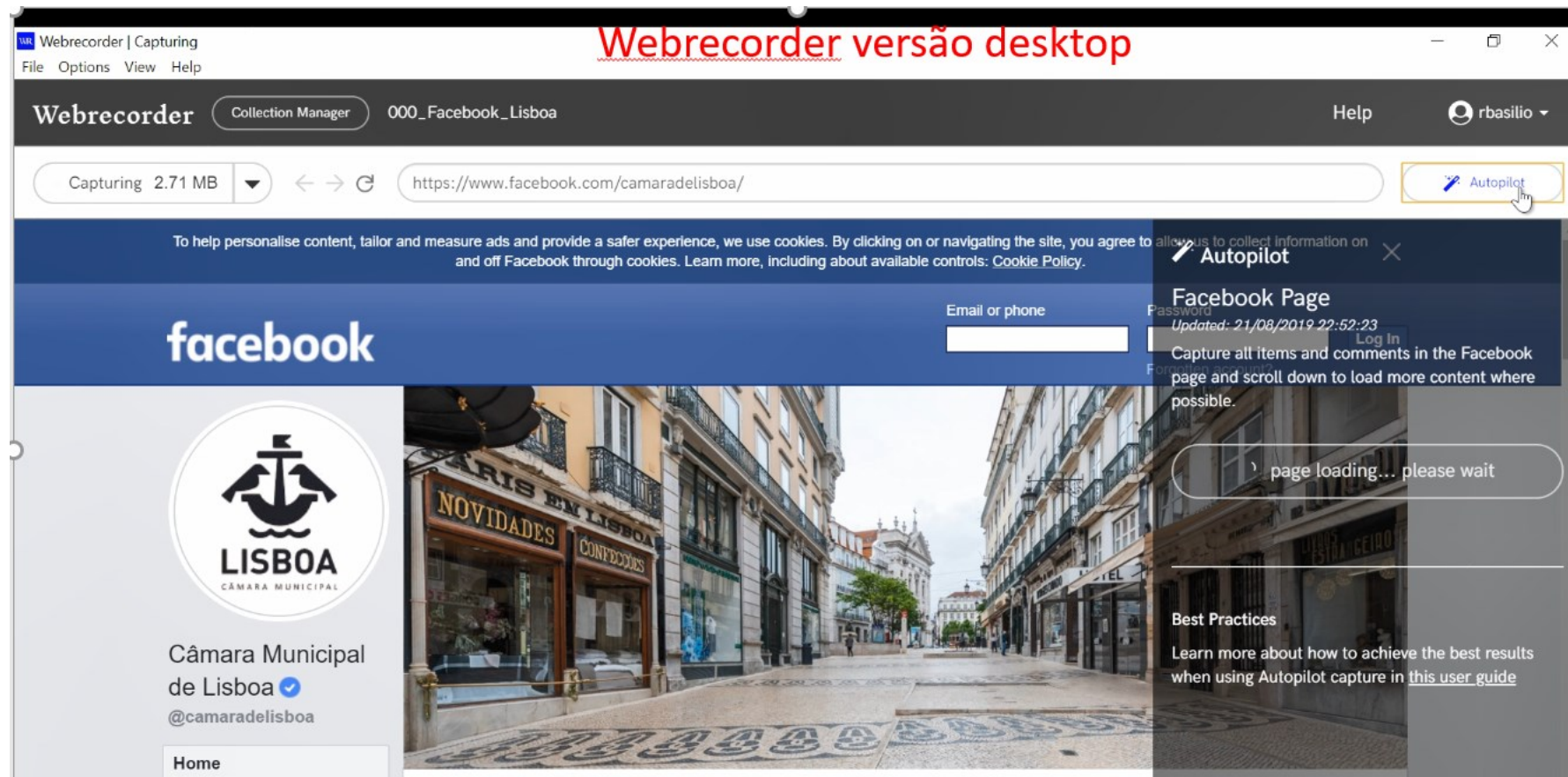
- É fácil de utilizar
- Utiliza um formato normalizado (WARC)
- Grava áudio e vídeo em sites e redes sociais
- Permite fazer download dos ficheiros gravados
- Reproduz as páginas gravadas



[Demo 1: Gravar páginas com o Webrecorder versão online – ver no Youtube](#)

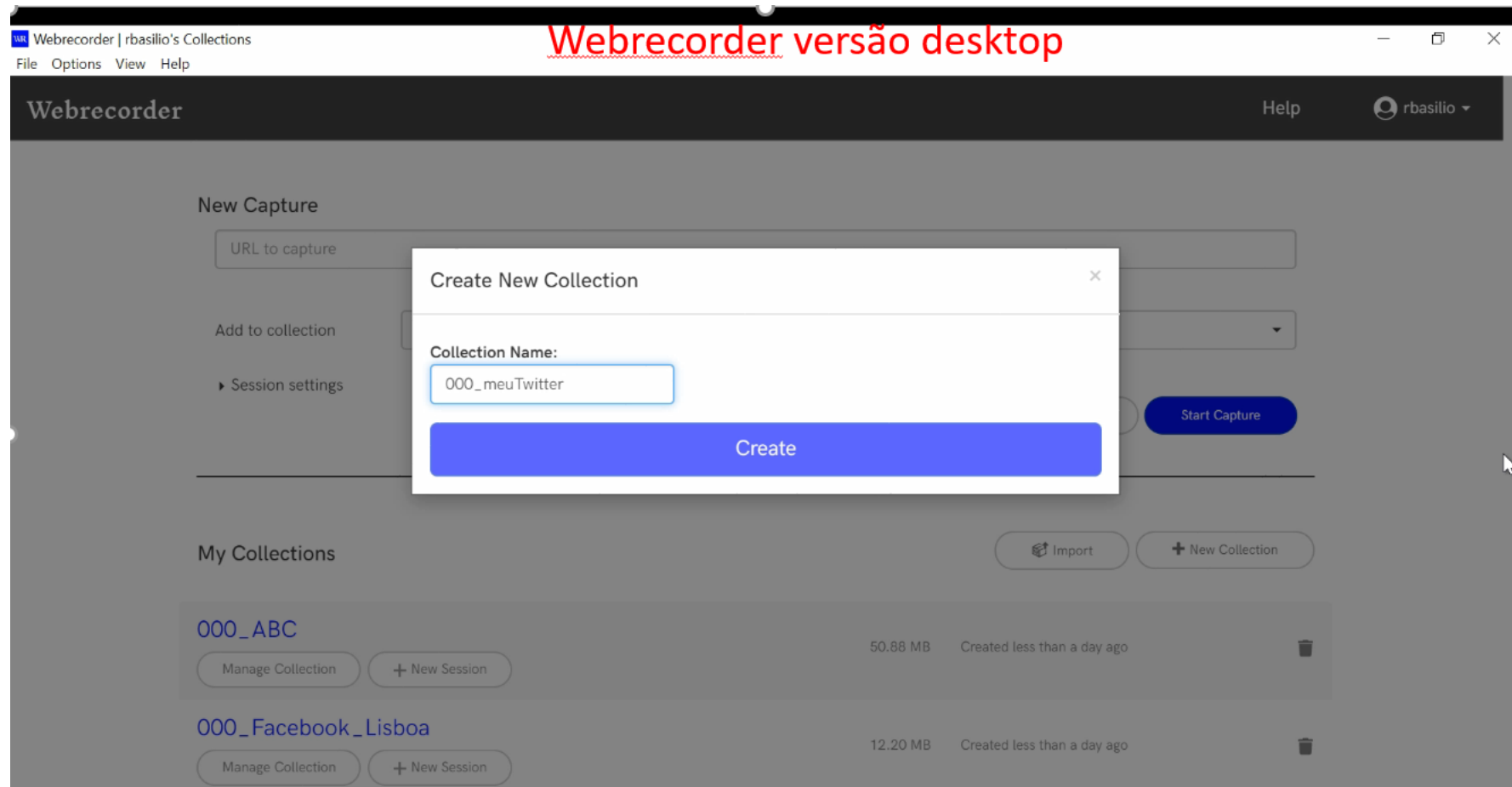


[Demo 2: Gravar páginas com o Webrecorder na versão desktop – ver no Youtube](#)



[Demo 3: Gravar páginas sociais com o Webrecorder na versão desktop – ver no Youtube](#)

Webrecorder versão desktop



[Demo 4: Gravar páginas sociais com o Webrecorder na versão desktop – ver no Youtube](#)



Browsertrix

- Recolhe sites inteiros e listas de páginas
- Relativamente fácil de implementar localmente
- Útil para recolhas em pequena escala e projetos exploratórios
- Utiliza a mesma tecnologia do Webrecorder



Perfis de recolha

- Uma página
- Um site inteiro
- Uma página + links dessa página



Perfis de recolha

- Uma página
- Um site inteiro
- Uma página + links dessa página

No Browsertrix

- Single page
- Same domain links
- All links on page
- Custom depth

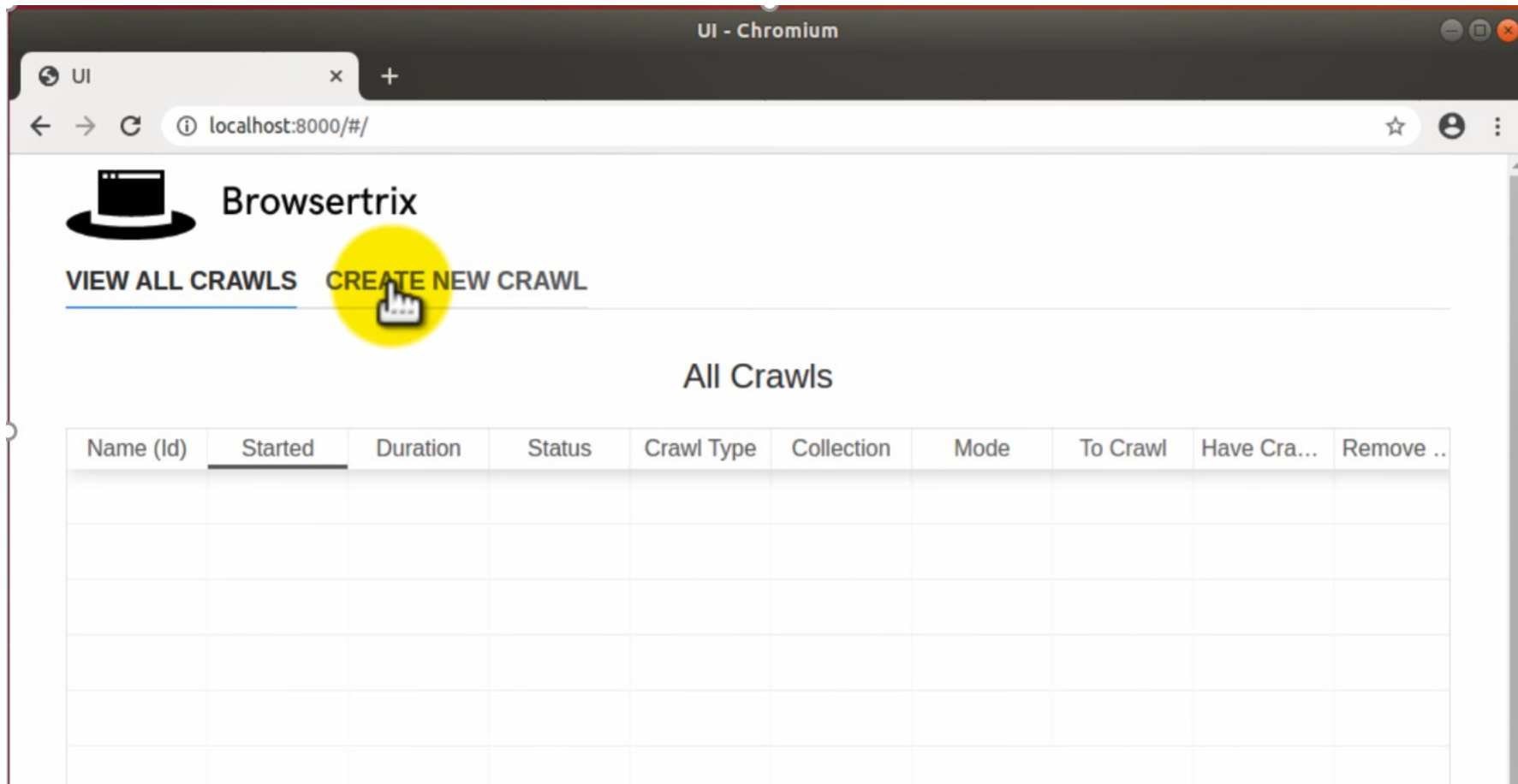


Perfis de recolha

- Uma página
- Um site inteiro
- Uma página + links dessa página

Outras designações

- Single page
- Domain crawl
- 1 hop
- 1 hop off

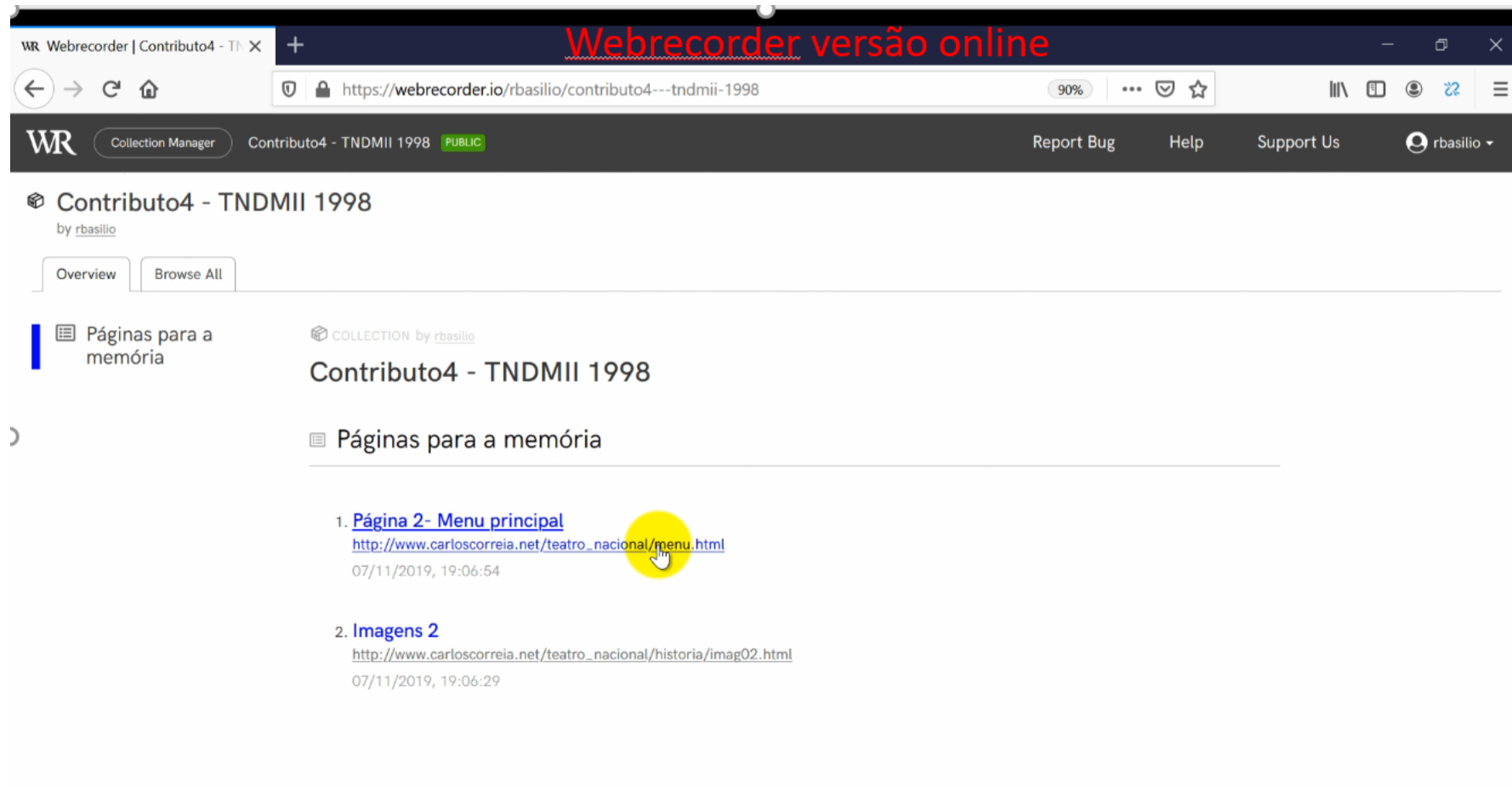


[Demo 5: Gravar sites inteiros com o Browsertrix, um crawler instalado no próprio computador – ver no Youtube](#)

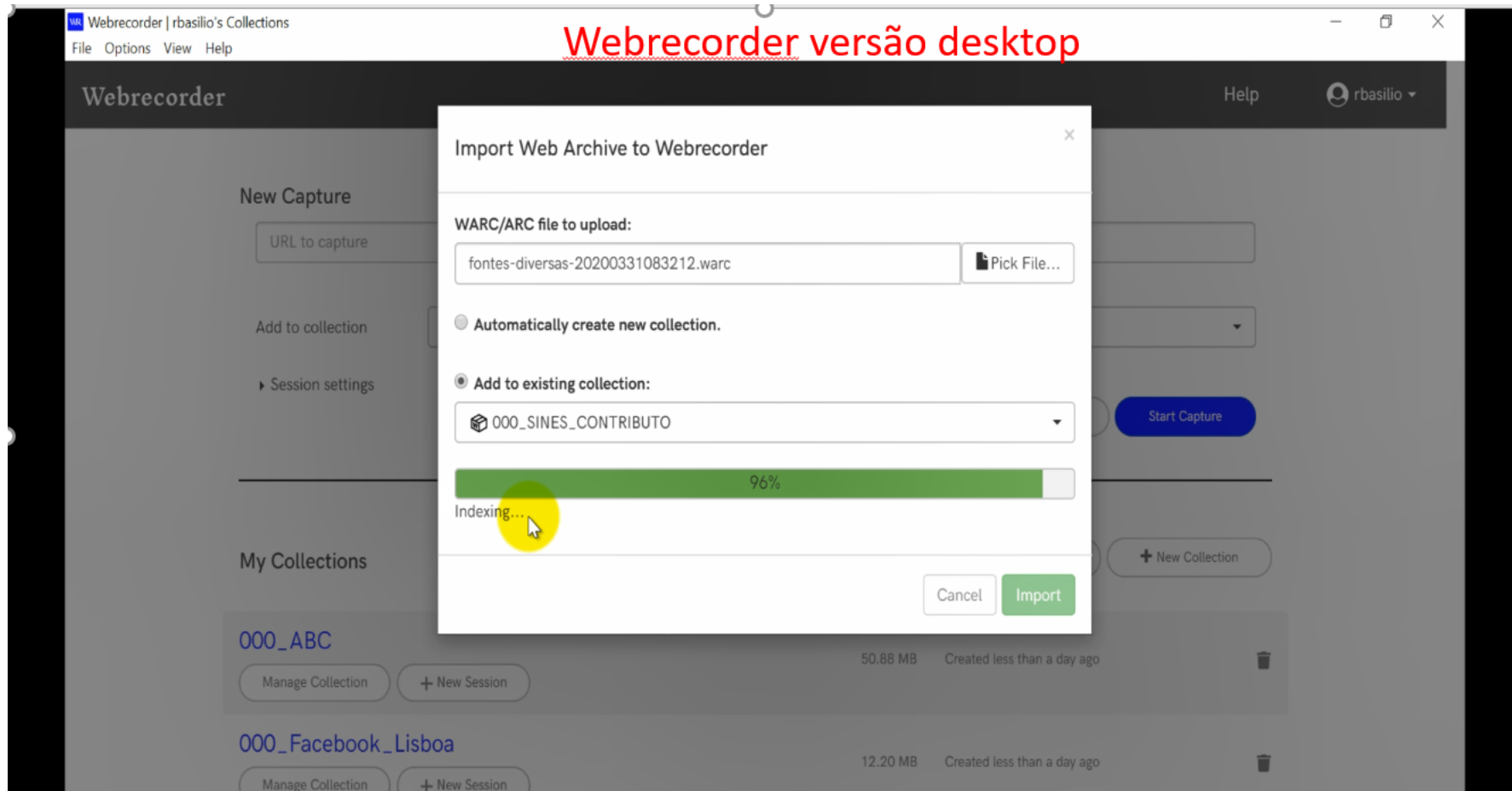


2 casos de uso:

- Teatro Nacional D. Maria II (nesta demo usa-se o Webrecorder online)
- Município de Sines (nesta demo usa-se o Webrecorder na versão desktop)



[Demo 6: Caso de uso, um contributo recolhido com o Webrecorder e enviado ao Arquivo.pt – ver no Youtube](#)



[Demo 7: Caso de uso, um contributo recolhido com o Webrecorder e enviado ao Arquivo.pt – ver no Youtube](#)



Discussão

- Centramo-nos na recolha e no formato WARC
- A reprodução, disponibilização exigem mais
- O armazenamento dos WARCs pode ser um obstáculo

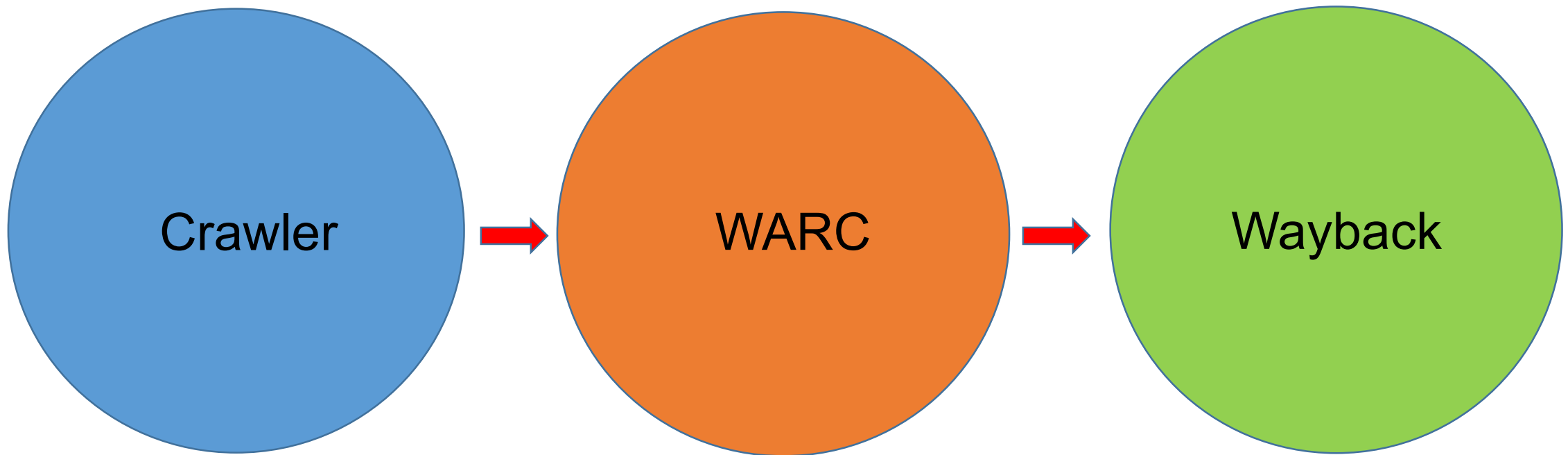


Três fases do processo de preservação:



Home

Três fases do processo de preservação:





Conclusão

- Somos capazes de gravar páginas com o **Webrecorder** e obter guardar o conteúdo em ficheiros WARC
- Acompanhamos o processo de gravação de um site inteiro com o **Browsertrix** e temos um modo concreto de o implementar nos nossos computadores pessoais.



Referências:

- Vídeo completo:
 - Facebook: <https://www.facebook.com/Arquivo.pt/videos/705354763615091/>
 - Youtube: <https://youtu.be/WSO0zvIZ4BA>
- **Guia de instalação do Browsertrix** no próprio computador para pessoas não especialistas em informática (em progresso):
 - <https://docs.google.com/document/d/1QqWfoSVPBQgd34OyhFR6790MmSKbIYmNDOkb6q7eisY/>
 - Mais info - Browsertrix no Git Hub: <https://github.com/webrecorder/browsertrix>
- Cafés com o Arquivo.pt – todas sessões:
 - <https://arquivo.pt/cafe>
 - Consulte a lista Formação e tutoriais no Youtube:
<https://www.youtube.com/watch?v=fNpUllm0dvc&list=PLKfzD5UuSdESJhKBIK3HDM6fk8wBM2RxE>



Email: contacto@arquivo.pt

Prepare-se para Prémio Arquivo.pt do próximo ano:
arquivo.pt/premios