

Linking web archiving with arts and humanities: the collaboration ROSSIO and Arquivo.pt

ricardo.basilio@fccn.pt

How to make Web archives useful and accessible to digital humanities researchers, and by extension to citizens?

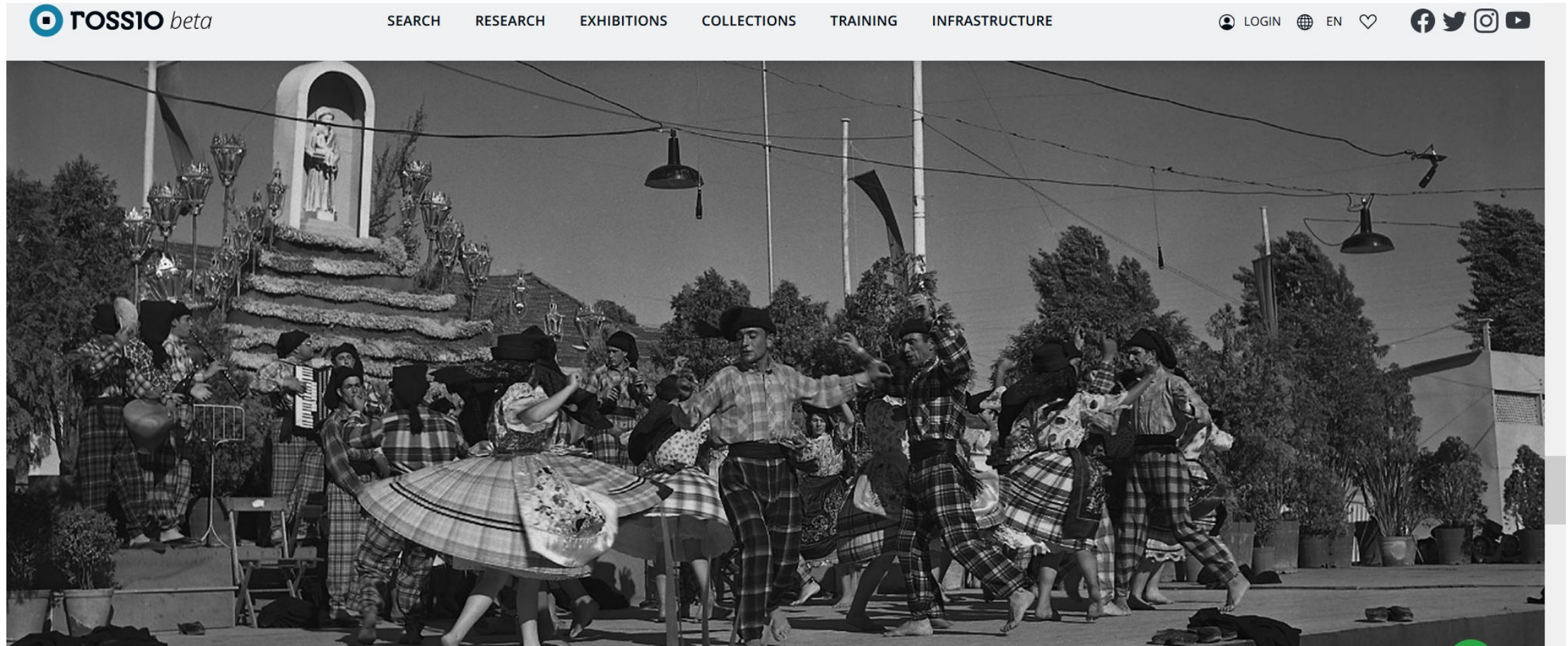
Challenge answered in three ways

- Dissemination
- Training
- Collaborative Curation of Websites

Agenda

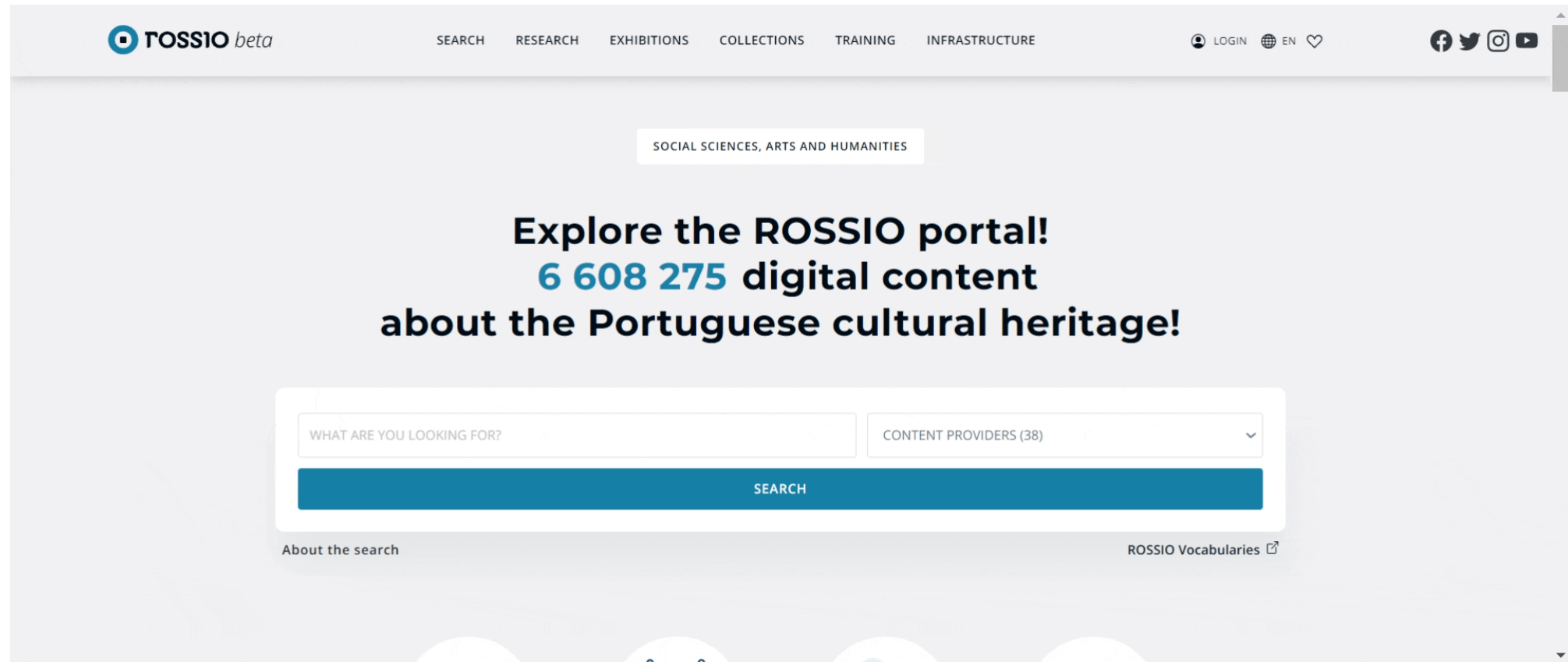
- What is ROSSIO?
- What is Arquivo.pt?
- Digital Humanities - a growing field of study
- Challenge – make a web archive useful

What is ROSSIO?



Web portal of the ROSSIO Infrastructure: <https://rossio.pt>

What is ROSSIO?

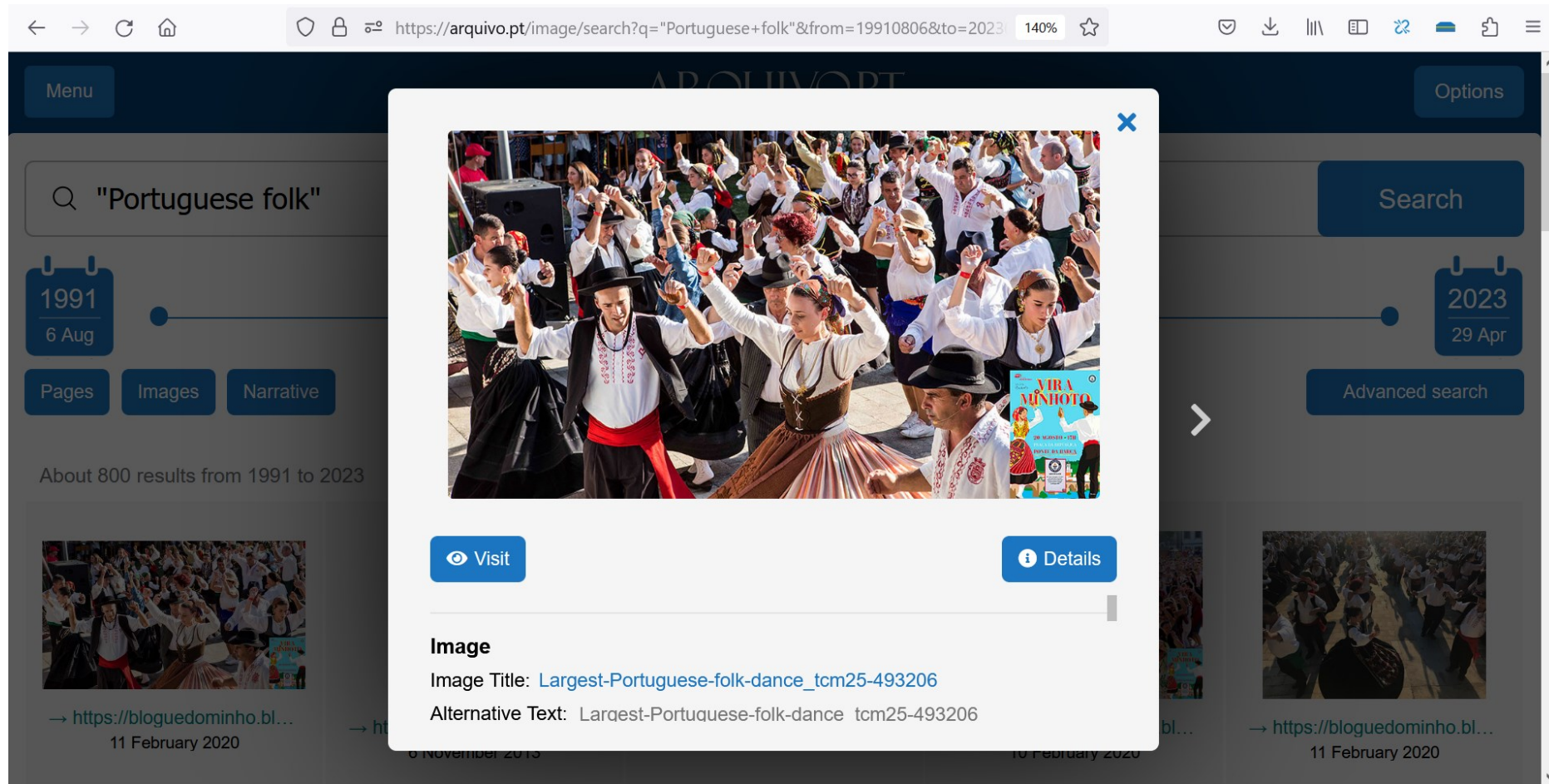


Web portal of the ROSSIO Infrastructure: <https://rossio.pt>

What is ROSSIO?

- Consortium of cultural and educational institutions
- Coordinated by the Faculty of NOVA School of Social Sciences and Humanities of Lisbon
- Aggregates, contextualizes, enriches and disseminates digital content
- 6 billion digital content about the Portuguese cultural heritage

What is Arquivo.pt?



[Example: a search for “Portuguese folk” on Arquivo.pt, the Portuguese Web Archive](#)

What is Arquivo.pt?

- Public Web preservation service
- Provides a public search interface (by URL, text, image, and API)
- Aimed at academic research, but also at all citizens

What is Arquivo.pt?

[Menu](#)

ARQUIVO.PT

Search

1991
6 Aug

2023
10 May

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

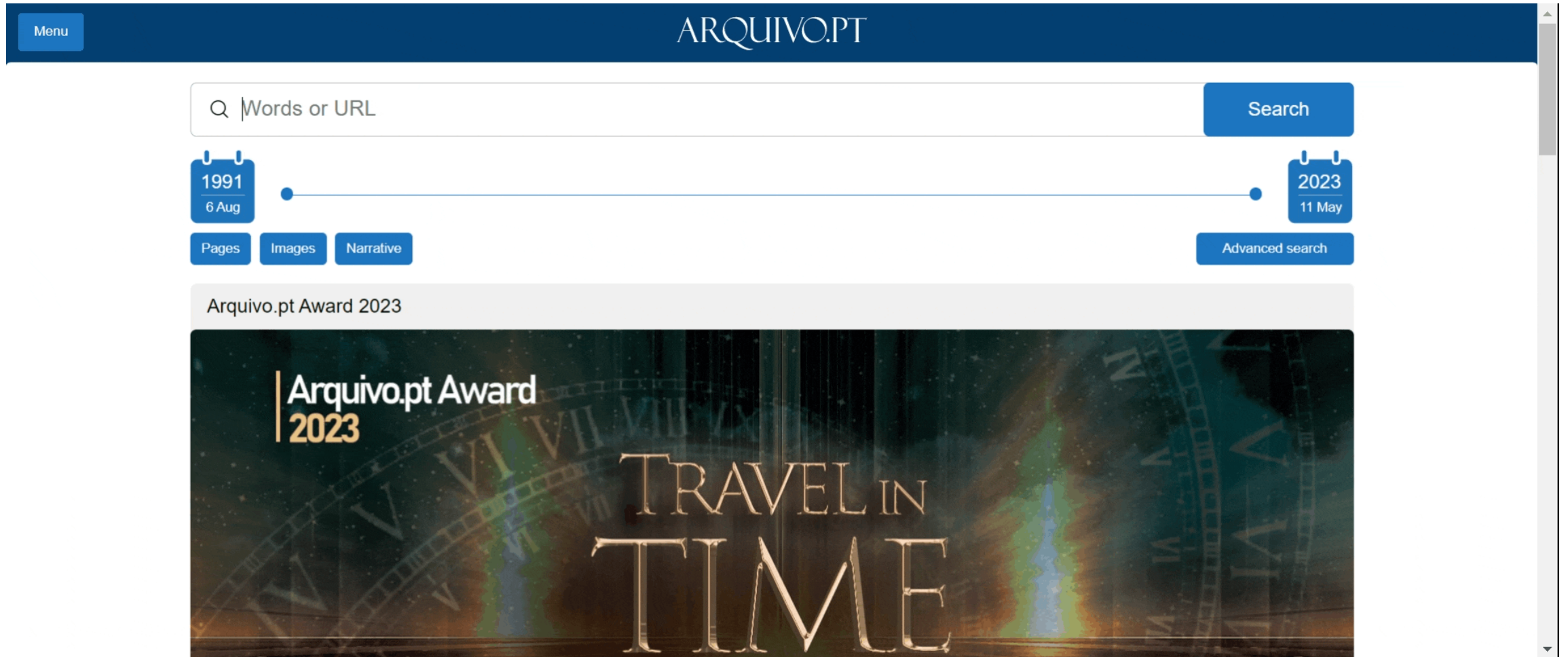
Advanced search

Arquivo.pt Award 2023

Arquivo.pt Award
2023

TRAVEL IN
TIME

What is Arquivo.pt?



Arquivo.pt interface

Digital Humanities - a growing field of study

The digital humanities, also known as **humanities computing**, is a field of study, research, teaching, and invention concerned with the **intersection** of computing and the disciplines of the humanities. It is **methodological** by nature and **interdisciplinary** in scope.

Matthew G. Kirschenbaum, ADE Bulletin, 150 (2010), p 2.

See also the Jason Heppler website "What is digital humanities?": <http://whatisdigitalhumanities.com/>

Digital Humanities - a growing field of study

«Dealing with a greater volume of data will be a common occurrence that historians will have to face from now on. This opens up **a world of possibility for historians**, but it also forces us to ask difficult questions»

Daniela Major, Early Stage Researcher in Digital Humanities at the School of Advanced Study in London ["It Takes a Village to Raise an Archive. How the Use of Web Sources Fosters Collaboration"](#).



Digital Humanities - a growing field of study

About the new Master in NOVA FCSH,
the headquarters of ROSSIO:

«The Master in Digital Curation and Humanities offers a solid and comprehensive training in **digital methodologies, tools and techniques** suitable for new approaches in research topics in the Arts, Humanities and Social Sciences.»

<https://www.fcsch.unl.pt/cursos/curadoria-e-humanidades-digitais/>

Digital Humanities

- Introduction to Programming for the Humanities
- Spatial Data Analysis in the Humanities
- Computational Methods and Tools for Exploring Texts
- Digital Technologies for Heritage Analysis
- Network Analysis

Digital Curation

- Ethics and Sustainability of Information in the Digital World
- Information Curation
- Terminology and Knowledge Organization
- Digital Culture and Convergence

Collaboration between ROSSIO and Arquivo.pt

ROSSIO funded Infrastructure

- Web curator
- Community manager
- Web developer
- Research team

Arquivo.pt data provider

- Know-how
- New services
- Open data sets

Dissemination

- Communication plan on social media (Facebook, Twitter)
- Mailing list
- Exhibitions in loco
- Exhibitions online

Dissemination

- Communication plan on social media
- 800 post on Facebook in 4 years
- 200 posts/year



Dissemination

- Mailing list - The preferred way to communicate
- Online form arquivo.pt/suggest
- Paper sheet during face-to-face training

Subscribe the Arquivo.pt mailing list

We will send only the most important information about Arquivo.pt (e.g. offers of grants, training, projects, collaborations, events, new collections or services).

sawfccn@gmail.com [Switch accounts](#)

Not shared

* Indicates required question

Name

Your answer

E-mail *

Your answer

Dissemination

On-site Exhibitions: printed Web pages in a public space in Lisbon



Dissemination

On-site Exhibition: printed Web
pages in the Library of the
University of Minho - Portugal



Dissemination



MemoriaFCSH – University websites



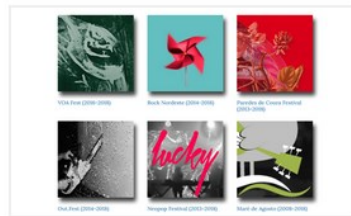
Memory of Portuguese Press (1996-2017)



Memory of Portuguese radio websites (1996-2018)



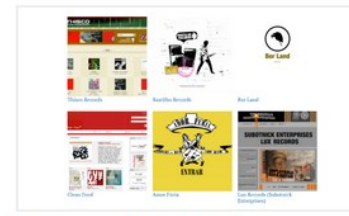
Memory of Portuguese TV Websites



Memory of Music Festivals
Websites in Portugal



Memory of Portuguese
Municipalities on the Internet



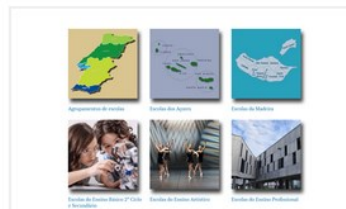
Portuguese Indie Labels



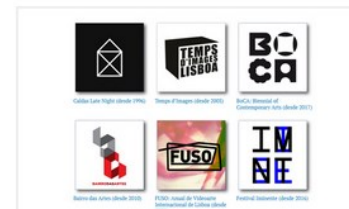
Memory of Scientific Research
(Portugal)



Memorial of Arquivo.pt



Memory of Portuguese schools on
the Internet



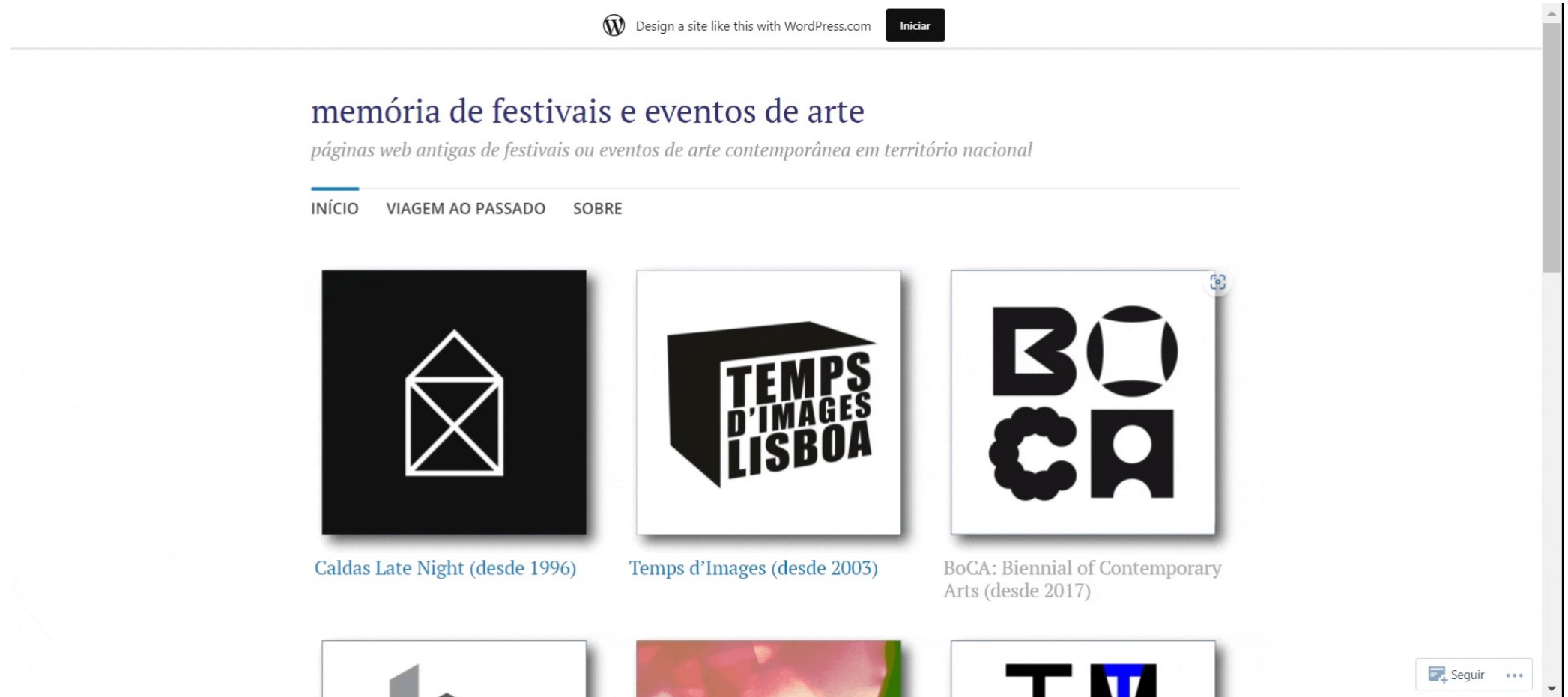
Memory of art festivals and events



Memory of Museums on the Web

Exhibitions on line: arquivo.pt/exbibitions

Dissemination



Memory of art and event festivals: arquivo.pt/exbitions

Dissemination



Time travel about art festivals and events in Portugal: arquivo.pt/festivalsdearte

Dissemination

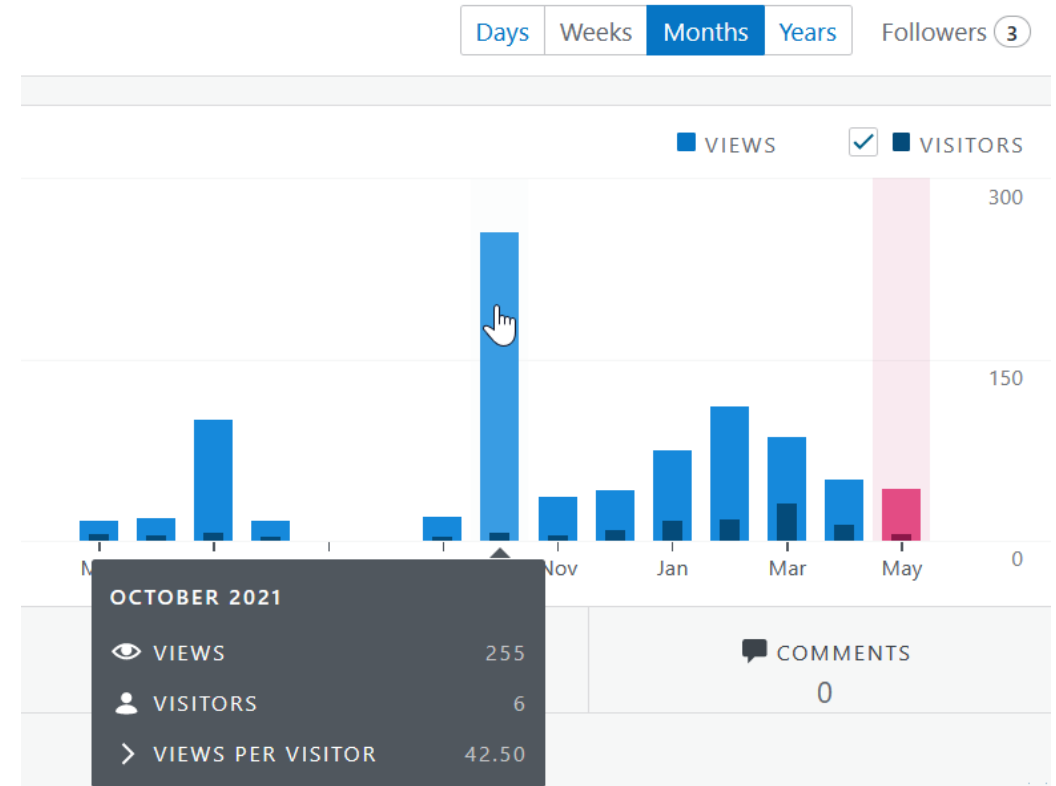


Time travel about art festivals and events in Portugal: arquivo.pt/festivalsdearte

Dissemination

From 2018 to April 2022:

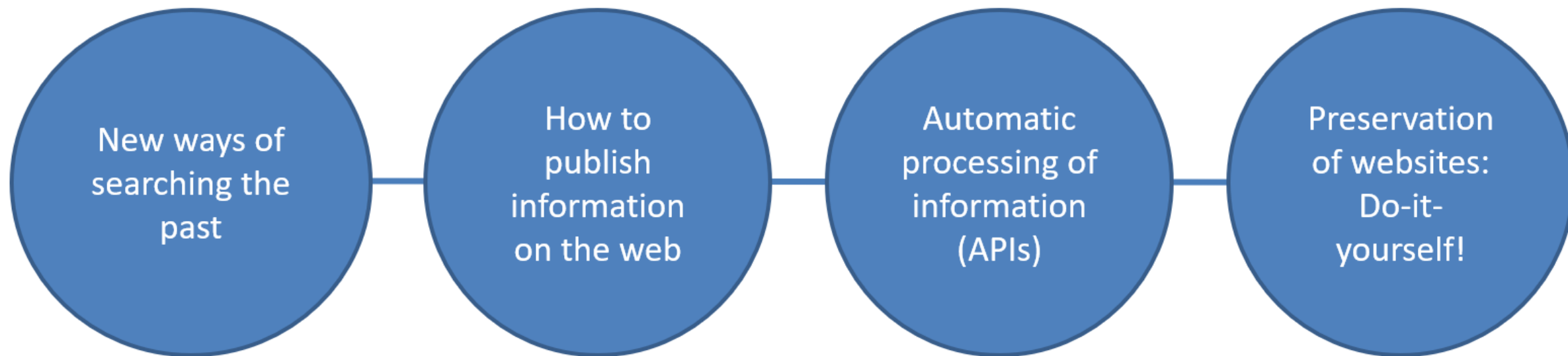
- 11 exhibitions on Wordpress.com
- 5 chronologies (time travels)
- 4 000 views
- **> 2 000 visitors**



Example: “memory of festivals and art events”
views and visitors

Exhibitions on line: arquivo.pt/exbitions

Training



Arquivo.pt training program in four modules: arquivo.pt/training

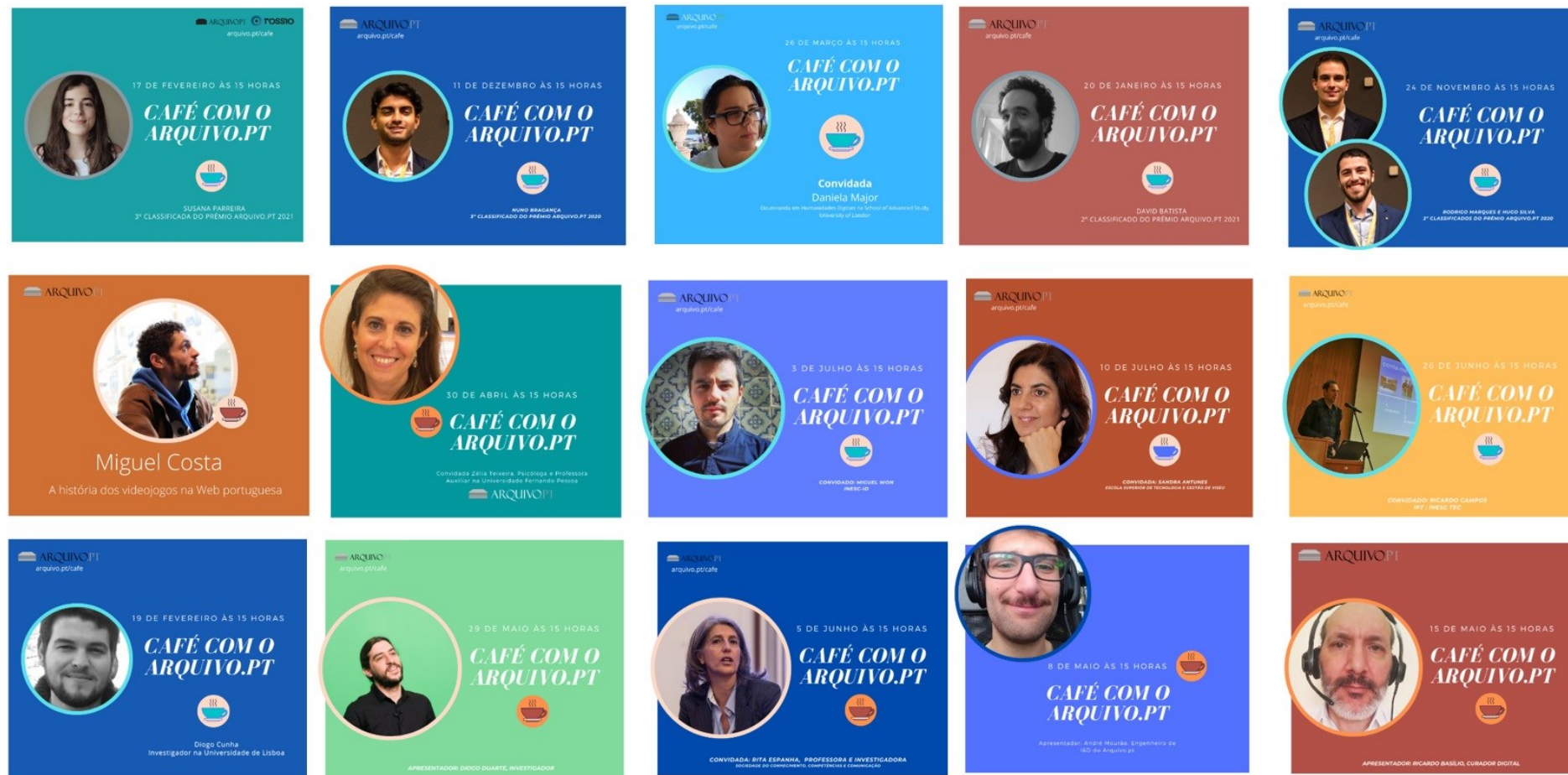
Training

Webinar cycle with the Portuguese Museum Network about museum websites (held in may and june 2022)



Portuguese Museums Network

Training



Café with Arquivo.pt. Online sessions held with ROSSIO partners and their communities

Training initiative - Café with Arquivo.pt

- 23 sessions
- 680 participants
- 30 participants/session
- 90% satisfaction

<https://arquivo.pt/onlinecafe>

Collaborative curation of websites



Collaborative curation of websites

	A	B	C	D	E	
1	URL (Address)	Website Channel Social media	Authors vs gallery owners	Link to the preserved version at Arquivo.pt	Name	Type of recording date and tool used
2	http://www.adrianamolder.com/	website	Autores	https://arquivo.pt/wayback/www.adrianamolder.com	Adriana Molder	Domain crawl with Brozz
3	http://www.adrianorangel.com/	website	Autores	https://arquivo.pt/wayback/www.adrianorangel.com	Adriano Rangel	Domain crawl with Brozz
4	http://www.albuquerquemendes.com/	website	Autores	https://arquivo.pt/wayback/www.albuquerquemendes.com	Albuquerque Mendes	Domain crawl with Brozz
5	http://www.ampersand-ampersand.com/	website	Galerias	https://arquivo.pt/wayback/www.ampersand-ampersand.com	Ampersand	Domain crawl with Brozz
6	http://www.anabiscaia.com/	website	Autores	https://arquivo.pt/wayback/www.anabiscaia.com	Ana Biscaia	Domain crawl with Brozz
7	http://anacardoso.net/	website	Autores	https://arquivo.pt/wayback/anacardoso.net/	Ana Cardoso	Domain crawl with Brozz
8	https://many-islands.com/	website	Autores	https://arquivo.pt/wayback/many-islands.com	Ana Freitas	Domain crawl with Brozz
9	https://many-islands.com/	website	Autores	https://arquivo.pt/wayback/www.anajaneiro.org	Ana Janeiro	Domain crawl with Brozz
10	https://anajoaoromana.wordpress.com/	website	Autores	https://arquivo.pt/wayback/anajoaoromana.wordpress.com	Ana João Romana	Domain crawl with Brozz
11	http://anamanso.net/	website	Autores	https://arquivo.pt/wayback/anamanso.net	Ana Manso	Domain crawl with Brozz
12	https://www.anamarchand.com/	website	Autores	https://arquivo.pt/wayback/www.anamarchand.com	Ana Marchand	Domain crawl with Brozz
13	http://anamata.pt/	website	Autores	https://arquivo.pt/wayback/anamata.pt	Ana Mata	Domain crawl with Brozz
14	http://www.anamendes.com/	website	Autores	https://arquivo.pt/wayback/www.anamendes.com	Ana Mendes	Domain crawl with Brozz
15	http://anaperezquioga.com/	website	Autores	https://arquivo.pt/wayback/anaperezquioga.com	Ana Pérez-Quiroga	Domain crawl with Brozz
16	https://anasimoes.myportfolio.com/	website	Autores	https://arquivo.pt/wayback/anasimoes.myportfolio.com	Ana Simões	Domain crawl with Brozz
17	https://www.anavidigal.com/	website	Autores	https://arquivo.pt/wayback/www.anavidigal.com	Ana Vidigal	Domain crawl with Brozz

Example of a collaborative curation of Portuguese artists' and galleries' websites – ROSSIO Infrastructure

Collaborative curation of websites

Using ArchiveWeb.page
(Webrecorder.net) for small-scale
recordings

ARQUIVAR A WEB
FAÇA-VOCÊ-MESMO!

WEBINÁRIO

25 de Novembro às 10h00

RICARDO BASÍLIO
FCCN PT

LISBOA
CÂMARA MUNICIPAL

FCT
Fundação para a Ciência e a Tecnologia
Computação Científica Nacional
FCCN

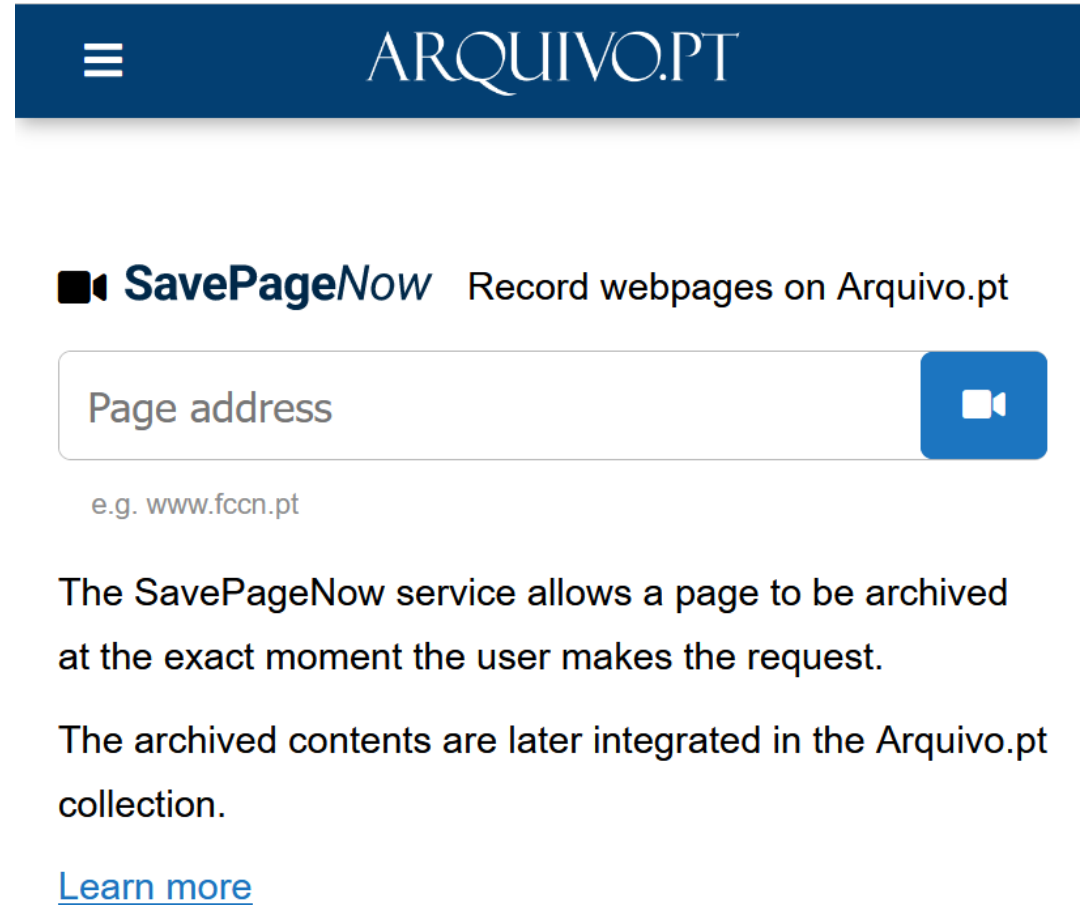
ARQUIVOPT

TOSSIO

foto Arquivo.pt fundo cottonbro no Pexels

Collaborative curation of websites


Services for community
participation – SavePageNow



The screenshot shows the top navigation bar of the Arquivo.pt website, which is dark blue with a white hamburger menu icon on the left and the text "ARQUIVO.PT" in white serif font on the right. Below the navigation bar, there is a section for the "SavePageNow" service. It features a dark blue icon of a camera and the text "SavePageNow" in bold, followed by the description "Record webpages on Arquivo.pt". Below this is a light gray input field with the placeholder text "Page address" and a blue button with a white camera icon. Under the input field, there is a small example text "e.g. www.fccn.pt". Below the input field, there are two paragraphs of text: "The SavePageNow service allows a page to be archived at the exact moment the user makes the request." and "The archived contents are later integrated in the Arquivo.pt collection." At the bottom of this section, there is a blue link that says "Learn more".

ARQUIVO.PT

 **SavePageNow** Record webpages on Arquivo.pt

Page address 

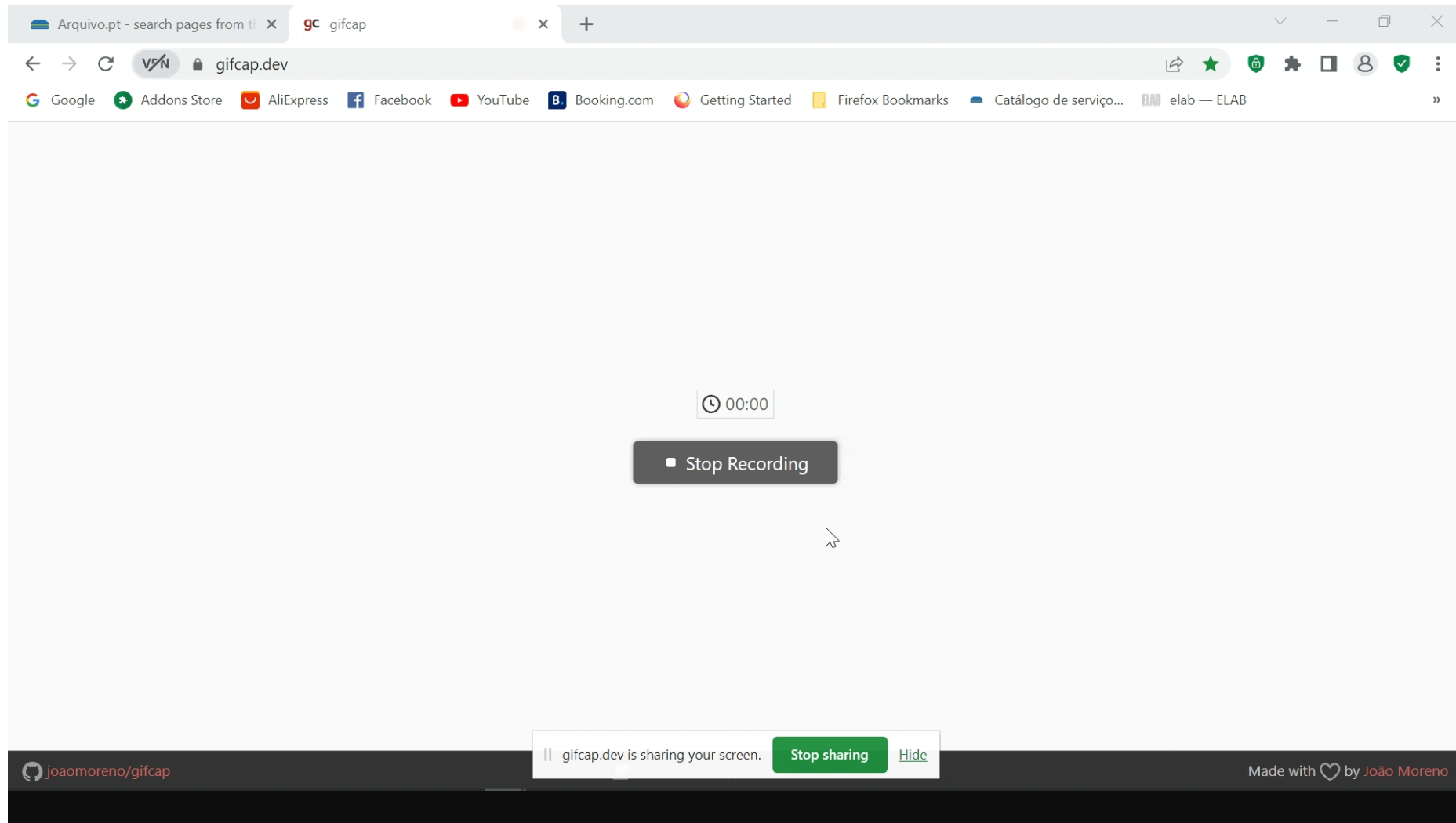
e.g. www.fccn.pt

The SavePageNow service allows a page to be archived at the exact moment the user makes the request.

The archived contents are later integrated in the Arquivo.pt collection.

[Learn more](#)

Collaborative curation of websites



Collaborative curation of websites

Arquivo.pt provides **open data sets** through the Portuguese open data portal Dados.gov

arquivo.pt/dadosabertos



The screenshot shows the Dados.gov website, which is the Portuguese open data portal. The header includes the logo, navigation links (DOCUMENTATION, OPEN DATA, NEWS, CONTACTS), and user options (CONTRIBUTE!, LOGIN / REGISTER, PT). The main content area displays a list of data sets, each with a title, description, and filters. The right sidebar shows a 'Formats' dropdown menu with options like xlsx, api, and csv, and a 'known reuses' section.

Data Set Title	Description	Filters
Publishers and booksellers in Portugal: websites and version history at Arquivo.pt	This data set presents the names of publishers and booksellers in Portugal and the history of their websites, which is accessible to any	1996-2022, One-off, Portugal, countries, 0
Websites of the Regional Government of the Azores and version history at Arquivo.pt	The purpose of listing the websites of the Regional Government of the Azores is to improve the collection of published content by Ar	1996-2021, Yearly, Portugal, countries, 0
Research and Development Units FCT 2019: websites and version history at Arquivo.pt	List of websites of Research and Development Units funded by FCT (2019 data). link.In addition to a tool...	1996-2021, Yearly, Portugal, countries, 0
Portuguese periodicals (newspapers and magazines): websites and version history at Arquivo.pt	List of websites of Portuguese periodical publications or pages related to their dissemination. a link.Besides...	1996-2021, Unknown, Portugal, countries, 0
Portuguese centenary newspapers: websites and links to versions preserved at Arquivo.pt (199...	This document contains a list of the 40 centenary newspapers, data provided by the Portuguese Press Association, in January 2023. 1	1996-2023, One-off, Portugal, countries, 0

Formats

- xlsx (40)
- api (4)
- special-collection-in-international-collaboration-xlsx (1)
- csv (1)
- gov.pt.xlsx (1)

known reuses

- never reused (46)
- little reused (two)
- much reused (1)

Printed screen, translated on Google Chrome

Collaborative curation of websites

- Arquivo.pt makes data available
- Researchers and developers build use cases



arquivo.pt/api#bulk

Collaborative curation of websites



Arquivo.pt Award generated 100+ use cases, between 2018 and 2023

What we learned

- Collaborations with diverse communities enrich Arquivo.pt
- The ROSSIO infrastructure has shown that it is with investment in human resources and means that a web archive can create value
- Digital humanities begin to pay attention to preserved Web content, but there is still a long way to go

Thank you!

Mail: contacto@arquivo.pt

Subscribe: arquivo.pt/subscribe