

# Access policies, challenges and approaches

Daniel Gomes

# Access policy

Only collect **publicly available** information

All web-archived content is **open-access**

**Legal mandate** to collect and provide access to web content interesting to the Portuguese community

Article 3 paragraph 2n) of [Decree-law 55/2013](#)

**Embargo** period of at least 1 year

Content owners may **opt-out**

But must **present proof** that they own the rights of the content

# General Data Protection Regulation made us stronger



## Article 17 EU GDPR "Right to erasure ('right to be forgotten')

1. The data subject shall have the right to obtain from the controller the erasure of **personal data** concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay where one of the following grounds applies:

**we don't keep personal data about our users and only collect published information**

3. Paragraphs 1 and 2 **shall not apply** to the extent that processing is necessary:

...

(b) for **compliance with a legal obligation** which requires processing by Union or Member State law to which the controller is subject or for the performance of a **task carried out in the public interest** or in the exercise of official authority vested in the controller;

**we are complying with the Law of Portugal and we are acting in the public interest**

...

(d) for **archiving purposes in the public interest, scientific or historical research purposes** or statistical purposes in accordance (e) for the establishment, exercise or defence of legal claims

**this is the reason for our existence**

# Access methods

## Human-User interfaces

URL search

Full-text search

Advanced search

Image search

## Application Programming Interfaces

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

# We get some complaints

174 000 users in 2018

Average of **2 removal requests per year**

Some were refused after legal analysis

Frequent complaints about having such a  
**long embargo period**

Including from the website owners

# Digital preservation quiz

Should the authors rights be respected?

Authors have the right to define the accessibility level of their intellectual property

They published their content **openly online**.

Should the **original accessibility level** be preserved?

It was originally open and online.

We do **not take ownership** of a content that belongs to the authors.

We **don't have the right** to close-access to what was born **open**.

# The web-archiving paradox

World-wide **open-access** movement on **closed-born** scholarly publications

Closing-access movement on **open-born** web publications

Libraries make **paid** publications **accessible** to all citizens

Some libraries make **open web** publications **practically inaccessible** to most tax payers that support them.

# Reflection

## Small initial challenge

Being *allowed* to provide access to web-archived content.

## Big challenge

Being *able* to provide access to web archived-content.

Provide access to *only* 6 billion web files preserved since 1996 required:

12 years of development

4 full-time experts

82 servers + maintenance costs