

# Workshop

5º encontro das  
**bibliotecas**  
de ensino superior  
Covilhã, 30 junho e 1 julho 2022



**Património digital**

**Preservar, salvaguardar e  
promover o acesso aos acervos**



associação portuguesa de  
**bibliotecários, arquivistas,  
profissionais da informação  
e documentação**



Património digital. Preservar, salvaguardar e promover o acesso aos acervos

# Salvaguarda e acesso à informação

Arquivo.pt. Preservação dos resultados de  
investigação de projetos científicos. Do FP4 ao H2020

ricardo.basilio@fccn.pt



Património digital. Preservar, salvaguardar e promover o acesso aos acervos

# Agenda

O Arquivo.pt (breve apresentação)

Preservação dos resultados de investigação de projetos científicos (europeus)

Preservação de resultados da investigação científica nacional

Participação dos bibliotecários e gestores de ciência na preservação



5º encontro das  
**bibliotecas**  
de ensino superior  
Covilhã, 30 junho e 1 julho, 2022



Património digital. Preservar, salvaguardar e promover o acesso aos acervos







# Arquivo.pt

- Serviço de público de preservação da Web operado pela Fundação para a Ciência e a Tecnologia, I.P.
- Missão – preservar a Web portuguesa (Decreto-Lei 55/2013, art. 3º)
- Instalações físicas – Av. do Brasil – Lisboa (Campus do LNEC)
- Acessível na Web em <https://arquivo.pt>





# Arquivo.pt

- 28 milhões de websites
- 852 TeraBytes
- 73 servidores
- 17 TB de memória RAM

Cf. [O Arquivo.pt em números, junho 2022](#)



Sala Grid – Unidade FCCN -FCT (pormenor)



## Património digital. Preservar, salvaguardar e promover o acesso aos acervos

The screenshot shows the website <https://sobre.arquivo.pt/pt/acerca/equipa/>. The page features the Arquivo.pt logo and a search bar. The main content is titled "Equipa" and lists six team members with their roles:

- João Gomes: Diretor Serviços Avançados
- Daniel Gomes: Gestor do serviço Arquivo.pt
- Francisco Esteveira: DevOps
- Pedro Gomes: Investigação e desenvolvimento
- Vasco Rato: Desenvolvimento Web
- Ricardo Basílio: Curador Digital

On the right side, there is a "Notícias" section with several articles:

- Criptomoedas e preservação da Web nos 15 anos do Arquivo.pt em Viseu
- Acabe com as "Páginas não encontradas" no teu website
- Grava páginas no Arquivo.pt com o SavePageNow
- Como preservar as referências na Wikipédia?
- Continua as sessões online Café com o Arquivo.pt

A link "Mais notícias ..." is visible at the bottom of the news section.

Equipa do Arquivo.pt (junho de 2022)



# O Arquivo.pt é uma infraestrutura de apoio às instituições do Ensino Superior

- Recolhe a Web para servir em todas as áreas de investigação
- Dá acesso aos conteúdos preservados
  - Pesquisa por endereço, texto e imagem
  - Pesquisa automáticas por APIs (Application Programming Interface)





## Património digital. Preservar, salvaguardar e promover o acesso aos acervos

The screenshot displays the Arquivo.pt search interface. At the top, the search bar contains the text "Encontro Bibliotecas do Ensino Superior" and a "Pesquisar" button. Below the search bar, there is a date range selector showing "1991 6 Ago" and "2022 29 Jun", along with buttons for "Páginas", "Imagens", and "Narrativa". A "Pesquisa avançada" button is also visible. The search results are displayed in a grid of 30 thumbnails, each representing a different image from the search results. The thumbnails include various logos, photos of people, and documents related to the event. A "Seguinte" button is located at the bottom right of the grid.

Exemplo da interface de pesquisa do Arquivo.pt, neste caso por imagens



## Património digital. Preservar, salvaguardar e promover o acesso aos acervos

A screenshot of a web browser displaying the homepage of the Major Minors project. The browser's address bar shows 'minors.ilch.uminho.pt'. The website has a dark sidebar on the left with navigation options like 'Homepage', 'Sobre o Projeto', 'Minorias', 'Ontologia', and 'Pesquisa Guiada'. The main content area features a large title 'Major Minors' and a paragraph describing the project as an ontology of newspaper clippings. Below this, there are four colorful boxes with statistics: 16286 articles, 4406 comments, 16286 images, and 5178169 ontological triples. A gold medal icon is visible in the top right corner of the page content.

← → ↻ 🏠 minors.ilch.uminho.pt 90% ☆ 📧 📄 🔄 📱 ☰

⚙️ Most Visited 🌟 Getting Started 📄 New Tab

**MajorMinors** | 🏠 Homepage | 📄 Contactos | 📄 Publicações | 📄 1º Prémio Arquivo.pt | 🏠 Universidade do Minho

# Major Minors

Ontologia de recortes de imprensa de jornais portugueses com referência a minorias, entre os quais um corpus de artigos, comentários e imagens. O corpus abarca as duas primeiras décadas do século XX. Este projeto conta com uma base de dados ontológica, ou seja, uma árvore de relações semânticas e referências hierarquizadas, que pretende contribuir para um retrato da representação de minorias no contexto português.

<b>16286</b> Artigos que referem minorias Galeria de Artigos	<b>4406</b> Comentários em artigos Galeria de Comentários	<b>16286</b> Imagens associadas a artigos Galeria de Imagens	<b>5178169</b> Triplos ontológicos Estrutura da ontologia
--	---	--	---

Projeto Major Minors, Universidade do Minho, vencedor do Prémio Arquivo.pt 2021. Ontologia de recortes de imprensa com referência às minorias, usaram o Arquivo.pt.



# Preservação dos resultados de investigação de projetos científicos (UE)





# Preservação dos resultados de investigação de projetos científicos - problema

- Os investigadores publicam fazem referências a informação complementar em websites, páginas de projeto, notícias e plataformas de publicação
- Os conteúdos Web desaparecem ou mudam em pouco tempo
- Os resultados científicos dos projetos perdem sustentabilidade



# O caso dos programas europeus

(Criados em 1984)

FP4 – 1994-1998

FP5 – 1998-2002

FP6 - 2002-2006

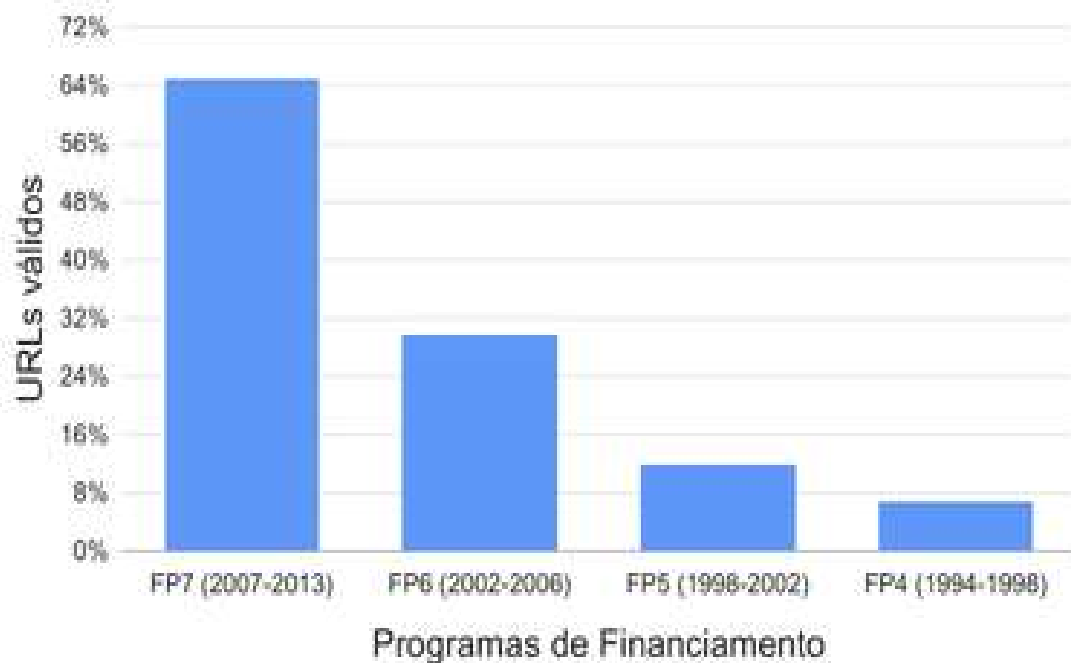
FP7 – 2007-2013

(FP8) H2020 – 2014-2020

(FP9) Horizon Europe – 2021-2027



## O caso dos programas europeus



URLs nos relatórios dos programas de financiamento entre 1998 e 2013. Fonte: Arquivo.pt





## Património digital. Preservar, salvaguardar e promover o acesso aos acervos

The screenshot shows the homepage of data.europa.eu. At the top, there is a navigation bar with 'Dados', 'Estudos', 'data.europa academy', 'Notícias', and 'Contacto'. Below this is a banner for 'Solidariedade da UE com a Ucrânia'. The main heading is 'Portal oficial dos dados europeus'. On the left, statistics show 174 Catalogos, 36 Países, and 1 481 065 Conjuntos de dados. A 'Trending datasets' section lists several datasets, with a yellow arrow pointing to the 'COROS' entry. To the right is a search section for 'Pesquisar conjuntos de dados' with a search bar and various category icons like 'ambiente', 'ciência e tecnologia', 'economia e finanças', etc.

174 **Catálogos** 36 **Países** 1 481 065 **Conjuntos de dados**

**Trending datasets**

- Lista consolidada de pessoas, grupos e entidades sujeitos a sanções financeiras da UE
- Número de identificação do contribuinte (NIF)
- Bolson petroliero da Comissão Europeia
- Base de dados de ingredientes cosméticos (CosIng) — Ingredientes e inventário de fragâncias
- COROS** — Projetos de investigação da UE no âmbito do Horizonte 2020 (2014-2020)

**Pesquisar conjuntos de dados**

CATÁLOGOS DE DADOS | TODOS OS CONJUNTOS DE DADOS | EU INSTITUTIONS DATASETS

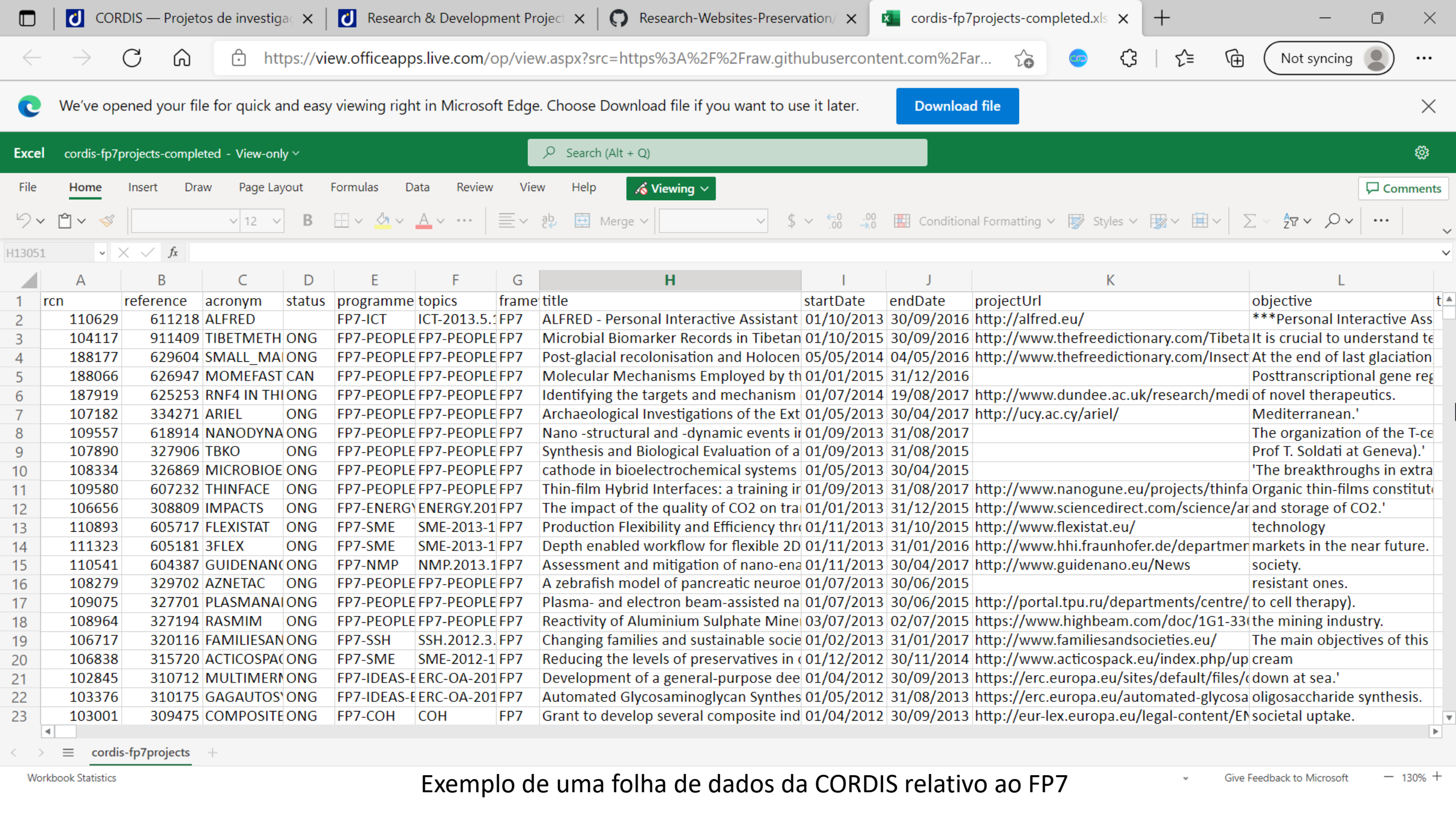
[Portal data.europa.eu](https://data.europa.eu)



## O caso dos programas europeus

92% dos projetos FP4 a FP7 listados na base de dados CORDIS\*  
estavam vazios (2015)

\* Community Research and Development Information Service



We've opened your file for quick and easy viewing right in Microsoft Edge. Choose Download file if you want to use it later. [Download file](#)

Excel cordis-fp7projects-completed - View-only

File Home Insert Draw Page Layout Formulas Data Review View Help **Viewing** Comments

Undo Copy Paste Font Size: 12 Bold Merge Conditional Formatting Styles

H13051

	A	B	C	D	E	F	G	H	I	J	K	L
1	rcn	reference	acronym	status	programme	topics	frame	title	startDate	endDate	projectUrl	objective
2	110629	611218	ALFRED		FP7-ICT	ICT-2013.5.	FP7	ALFRED - Personal Interactive Assistant	01/10/2013	30/09/2016	<a href="http://alfred.eu/">http://alfred.eu/</a>	***Personal Interactive Ass
3	104117	911409	TIBETMETH	ONG	FP7-PEOPLE	FP7-PEOPLE	FP7	Microbial Biomarker Records in Tibetan	01/10/2015	30/09/2016	<a href="http://www.thefreedictionary.com/Tibeta">http://www.thefreedictionary.com/Tibeta</a>	It is crucial to understand te
4	188177	629604	SMALL_MA	ONG	FP7-PEOPLE	FP7-PEOPLE	FP7	Post-glacial recolonisation and Holocen	05/05/2014	04/05/2016	<a href="http://www.thefreedictionary.com/Insect">http://www.thefreedictionary.com/Insect</a>	At the end of last glaciation
5	188066	626947	MOMEFAST	CAN	FP7-PEOPLE	FP7-PEOPLE	FP7	Molecular Mechanisms Employed by th	01/01/2015	31/12/2016		Posttranscriptional gene reg
6	187919	625253	RNF4 IN THI	ONG	FP7-PEOPLE	FP7-PEOPLE	FP7	Identifying the targets and mechanism	01/07/2014	19/08/2017	<a href="http://www.dundee.ac.uk/research/medi">http://www.dundee.ac.uk/research/medi</a>	of novel therapeutics.
7	107182	334271	ARIEL	ONG	FP7-PEOPLE	FP7-PEOPLE	FP7	Archaeological Investigations of the Ext	01/05/2013	30/04/2017	<a href="http://ucy.ac.cy/ariel/">http://ucy.ac.cy/ariel/</a>	Mediterranean.'
8	109557	618914	NANODYNA	ONG	FP7-PEOPLE	FP7-PEOPLE	FP7	Nano -structural and -dynamic events in	01/09/2013	31/08/2017		The organization of the T-ce
9	107890	327906	TBKO	ONG	FP7-PEOPLE	FP7-PEOPLE	FP7	Synthesis and Biological Evaluation of a	01/09/2013	31/08/2015		Prof T. Soldati at Geneva).'
10	108334	326869	MICROBIOE	ONG	FP7-PEOPLE	FP7-PEOPLE	FP7	cathode in bioelectrochemical systems	01/05/2013	30/04/2015		'The breakthroughs in extra
11	109580	607232	THINFACE	ONG	FP7-PEOPLE	FP7-PEOPLE	FP7	Thin-film Hybrid Interfaces: a training in	01/09/2013	31/08/2017	<a href="http://www.nanogune.eu/projects/thinfa">http://www.nanogune.eu/projects/thinfa</a>	Organic thin-films constitut
12	106656	308809	IMPACTS	ONG	FP7-ENERG	ENERGY.201	FP7	The impact of the quality of CO2 on tra	01/01/2013	31/12/2015	<a href="http://www.sciencedirect.com/science/ar">http://www.sciencedirect.com/science/ar</a>	and storage of CO2.'
13	110893	605717	FLEXISTAT	ONG	FP7-SME	SME-2013-1	FP7	Production Flexibility and Efficiency thr	01/11/2013	31/10/2015	<a href="http://www.flexistat.eu/">http://www.flexistat.eu/</a>	technology
14	111323	605181	3FLEX	ONG	FP7-SME	SME-2013-1	FP7	Depth enabled workflow for flexible 2D	01/11/2013	31/01/2016	<a href="http://www.hhi.fraunhofer.de/departmen">http://www.hhi.fraunhofer.de/departmen</a>	markets in the near future.
15	110541	604387	GUIDENAN	ONG	FP7-NMP	NMP.2013.1	FP7	Assessment and mitigation of nano-ena	01/11/2013	30/04/2017	<a href="http://www.guidenano.eu/News">http://www.guidenano.eu/News</a>	society.
16	108279	329702	AZNETAC	ONG	FP7-PEOPLE	FP7-PEOPLE	FP7	A zebrafish model of pancreatic neuroe	01/07/2013	30/06/2015		resistant ones.
17	109075	327701	PLASMANA	ONG	FP7-PEOPLE	FP7-PEOPLE	FP7	Plasma- and electron beam-assisted na	01/07/2013	30/06/2015	<a href="http://portal.tpu.ru/departments/centre/">http://portal.tpu.ru/departments/centre/</a>	to cell therapy).
18	108964	327194	RASMIM	ONG	FP7-PEOPLE	FP7-PEOPLE	FP7	Reactivity of Aluminium Sulphate Mine	03/07/2013	02/07/2015	<a href="https://www.highbeam.com/doc/1G1-33">https://www.highbeam.com/doc/1G1-33</a>	the mining industry.
19	106717	320116	FAMILIESAN	ONG	FP7-SSH	SSH.2012.3.	FP7	Changing families and sustainable socie	01/02/2013	31/01/2017	<a href="http://www.familiesandsocieties.eu/">http://www.familiesandsocieties.eu/</a>	The main objectives of this
20	106838	315720	ACTICOSPAC	ONG	FP7-SME	SME-2012-1	FP7	Reducing the levels of preservatives in	01/12/2012	30/11/2014	<a href="http://www.acticospack.eu/index.php/up">http://www.acticospack.eu/index.php/up</a>	cream
21	102845	310712	MULTIMER	ONG	FP7-IDEAS-E	ERC-OA-201	FP7	Development of a general-purpose dee	01/04/2012	30/09/2013	<a href="https://erc.europa.eu/sites/default/files/c">https://erc.europa.eu/sites/default/files/c</a>	down at sea.'
22	103376	310175	GAGAUTOS	ONG	FP7-IDEAS-E	ERC-OA-201	FP7	Automated Glycosaminoglycan Synthes	01/05/2012	31/08/2013	<a href="https://erc.europa.eu/automated-glycosa">https://erc.europa.eu/automated-glycosa</a>	oligosaccharide synthesis.
23	103001	309475	COMPOSITE	ONG	FP7-COH	COH	FP7	Grant to develop several composite ind	01/04/2012	30/09/2013	<a href="http://eur-lex.europa.eu/legal-content/EN">http://eur-lex.europa.eu/legal-content/EN</a>	societal uptake.

cordis-fp7projects +

Exemplo de uma folha de dados da CORDIS relativo ao FP7





## Coluna “projectURL”

	G	H	I	J	K	L
1	frameworkProgram	title	startDate	endDate	projectUrl	object
2	FP7	ALFRED - Personal Interactive Assistant for Independent Living and Active Ageing	01/10/2013	30/09/2016	http://alfred.eu/	***Pe
3	FP7	Microbial Biomarker Records in Tibetan Peats: Monsoon Variability and its Impact o	01/10/2015	30/09/2016	http://www.thefreedictionary.com/Tibetan	It is cr
4	FP7	Post-glacial recolonisation and Holocene anthropization impact on populations of st	05/05/2014	04/05/2016	http://www.thefreedictionary.com/Insectiv	At the
5	FP7	Molecular Mechanisms Employed by the Newly Assigned RNA-binding Protein FASTK	01/01/2015	31/12/2016		Posttr
6	FP7	Identifying the targets and mechanism of action of the SUMO targeted ubiquitin liga	01/07/2014	19/08/2017	http://www.dundee.ac.uk/research/media/	The
7	FP7	Archaeological Investigations of the Extra-Urban and Urban Landscape in Eastern M	01/05/2013	30/04/2017	http://ucy.ac.cy/a	'The
8	FP7	Nano -structural and -dynamic events in the T-cell activation	01/09/2013	31/08/2017		The or
9	FP7	Synthesis and Biological Evaluation of a Potent Antituberculosis Natural Product and	01/09/2013	31/08/2015		'This
10	FP7	Microbially catalyzed electricity driven bioproduction from CO2 at the	01/05/2013	30/04/2015		'The b

I	J	K	L
startDate	endDate	projectUrl	object
01/10/2013	30/09/2016	http://alfred.eu/	***Pe
01/10/2015	30/09/2016	http://www.thefreedictionary.com/Tibetan	It is cr
05/05/2014	04/05/2016	http://www.thefreedictionary.com/Insectiv	At the
01/01/2015	31/12/2016		Posttr



## Iniciativa do Arquivo.pt para identificar e preservar resultados dos projetos de I&D

- Uso de dados abertos disponibilizados pelo EU data portal
- Recurso ao serviço de pesquisa automática Bing (Microsoft) (versão gratuita, 2016)



## Iniciativa do Arquivo.pt para identificar e preservar resultados dos projetos de I&D

- Uso de expressões de pesquisa avançada em motor de busca
- Acrónimo + Título do projeto
- Acrónimo + Título do projeto –Cordis +termos comuns

Exemplo:

ALFRED + “ALFRED – Personal interactive assistant” –Cordis +projeto





# Iniciativa do Arquivo.pt para identificar e preservar resultados dos projetos de I&D - Resultados

- 53 993 sites
- 52 milhões de ficheiros
- 7 Terabytes
- 86,6% de melhoria



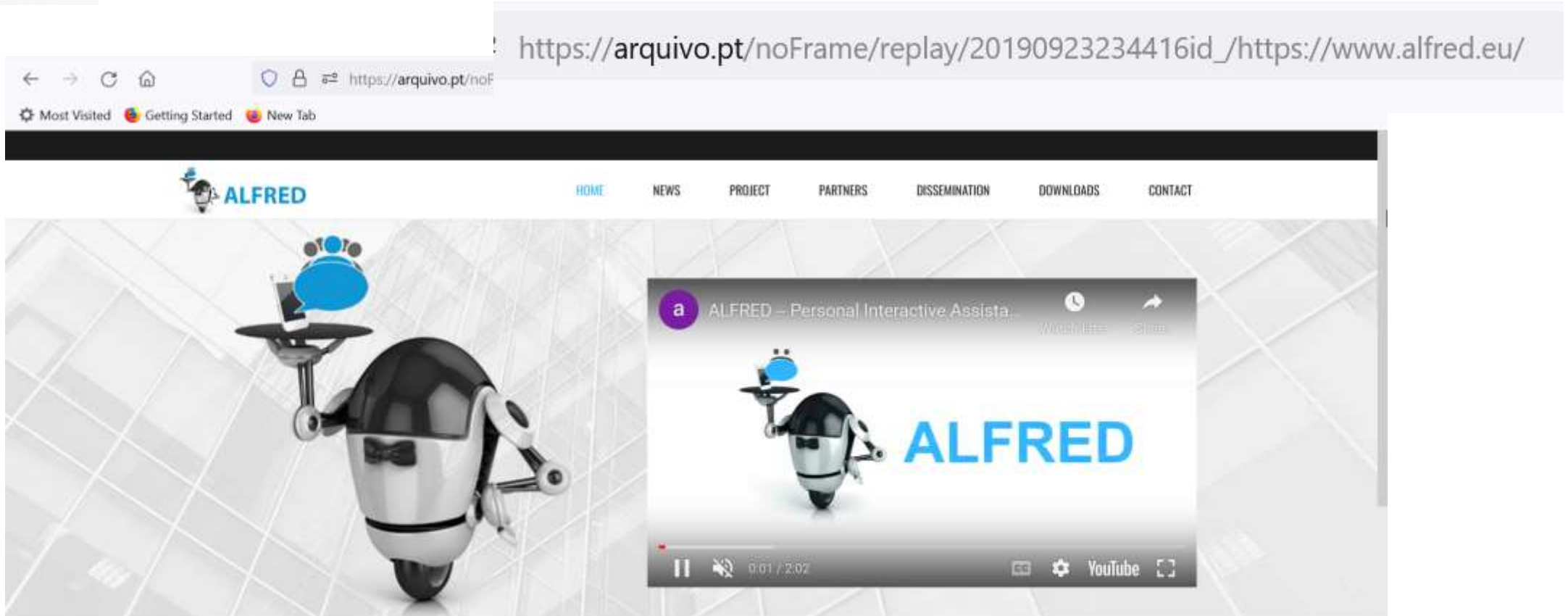
## Património digital. Preservar, salvaguardar e promover o acesso aos acervos

The screenshot shows a web browser window displaying the WEZARD website. The browser's address bar shows the URL <https://arquivo.pt/wayback/20160515030359/http://www.wezard.eu/>. The website header includes the 'ARQUIVO.PT' logo and the date '15 May at 03h03, 2016'. The main content area features the WEZARD logo and the title 'WEather haZARDs for aeronautics'. A navigation menu on the left lists various sections like 'Home', 'The Project', and 'Contact'. The main content area includes an 'Overview' section with text describing the project's goals and a list of consortium members. Logos for the European Commission and the Portuguese Ministry of Education and Science are also visible.

Site do projeto europeu WEZARD ([www.wezard.eu](http://www.wezard.eu)), financiado pelo 7º Programa-Quadro da União Europeia entre 2011 e 2013, disponível no Arquivo.pt.



## Património digital. Preservar, salvaguardar e promover o acesso aos acervos



### Interactive Assistant for Independent Living and Active Ageing

ALFRED - Personal Interactive Assistant for Independent Living and Active Ageing - is a project funded by the Seventh Framework Programme of the European Commission under Grant Agreement No. 611218. It will allow older people to live longer at their own homes with the possibility to act independently and to actively participate in society by providing the

O site projeto [alfred.eu](https://www.alfred.eu) foi preservado pelo Arquivo.pt



# Disponibilização dos dados recolhidos pelo Arquivo.pt no European Open Data Portal

- Para serem melhorados
- Para serem reutilizados





## Património digital. Preservar, salvaguardar e promover o acesso aos acervos

A screenshot of a web browser displaying the 'data.europa.eu' website. The browser's address bar shows the URL 'https://data.europa.eu/data/datasets?query=FP4&locale=en'. The website's navigation bar includes 'Data', 'Studies', 'data.europa academy', 'News', and 'Contact'. The main content area is titled 'Datasets' and features a search bar with 'FP4' entered. Below the search bar, there are tabs for 'Datasets', 'Catalogues', and 'Editorial Content', along with sorting options: 'Last Modified', 'Relevance', and 'More'. A map on the left side shows the European Union. The search results list two datasets:

- CORDIS - EU research projects under FP4 (1994-1998)**: This dataset contains all projects funded by the European Union under the fourth framework programme for research and technological development (FP4) from 1994 to 1998. The file 'FP4 Projects' contains the public grant information for each project, including the following information: Record Control Number (RCN), project ID (grant agreement number), project... Updated: 12.02.2018 Created: 27.07.2015. Source: Publications Office. Available formats: CSV, Excel, ZIP.
- Research & Development Project Websites funded by the European Commission: FP4, FP5, FP6, FP7**: This data set contains the lists of Research Project Websites of the Framework Programmes 4th, 5th and 7th or Framework Programms FP4, FP5, FP6, FP7, funded by the European Commission. In 2016, it was found that online information on R & D projects was not fully documented. For example, the information on the website addresses of projects funded... Updated: 31.08.2021 14:02 Created: 31.08.2021 13:50. Source: dados.gov.pt. Available format: Excel.

EU Open Data Portal disponibiliza conjuntos de dados complementados pelo Arquivo.pt



## Património digital. Preservar, salvaguardar e promover o acesso aos acervos

The screenshot shows the website <https://dados.gov.pt> with the following content:

- Header:** Includes the **dados.gov** logo, a "CONTRIBUIR!" button, and the user profile "RICARDO BASILIO". Navigation links for "DOCUMENTAÇÃO", "DADOS ABERTOS", "NOTÍCIAS", and "CONTACTOS" are present.
- Main Title:** "Websites dos projetos de Investigação & Desenvolvimento financiados pela Comissão Europeia: FP4, FP5, FP6, FP7".
- Description:** A paragraph explaining that the data comes from a certified public service and details the methodology used by Arquivo.pt to collect and preserve research website data from the 4th, 5th, 6th, and 7th Framework Programs.
- Resources:** A section with a "Pré-visualizar" button and a list of three XLSX files:
  - "Lista de projetos do 7º Programa Quadro (FP7): websites recolhidos pelo Arquivo.pt" (34.01 MB)
  - "Lista de projetos do 6º Programa Quadro (FP6): websites recolhidos pelo Arquivo.pt" (10.4 MB)
  - "Lista de projetos do 5º Programa Quadro (FP5): websites recolhidos pelo Arquivo.pt" (10.4 MB)
- Provider:** "Arquivo.pt - pesquisa páginas do passado" with a "dados.gov" certification seal.
- Actions:** Buttons for "VER PERFIL", "CONTACTOS", and "DEIXAR DE SEGUIR".
- Informações:** License "Creative Commons Attribution 4.0 - CC BY 4.0", date range "2004-2013", location "Portugal", and update date "31 de agosto de 2021".

Dados.gov (Portugal) disponibiliza conjuntos de dados sobre os projetos europeus



## Património digital. Preservar, salvaguardar e promover o acesso aos acervos

The screenshot shows a GitHub repository page for 'arquivo / Research-Websites-Preservation'. The repository is public and has 2 forks and 3 stars. The main content is a list of folders representing datasets, each with a commit message and a date. The folders are: Bing experiments, Cleaning golden dataset, FCT, FP4, FP5, FP6, and FP7. The commit messages describe the actions taken for each dataset, such as 'Sync repository, cordis fp7 data completed with the retriev...' or 'add FCT projects dataset'. The repository also has a 'Code' button and a 'Go to file' button. On the right side, there is an 'About' section with a description: 'Study about preserving R&D website projects. Methods to automatically identify these URLs.' and a 'Releases' section.

Folder Name	Commit Message	Commit Date
Bing experiments	Sync repository, cordis fp7 data completed with the retriev...	6 years ago
Cleaning golden dataset	Sync repository, cordis fp7 data completed with the retriev...	6 years ago
FCT	add FCT projects dataset	6 years ago
FP4	Completed Datasets	6 years ago
FP5	Adding cordisfp5 xlsx format	5 years ago
FP6	add fp6 completed dataset	6 years ago
FP7	Sync repository, cordis fp7 data completed with the retriev...	6 years ago

Dados e ferramentas estão disponíveis publicamente no [canal GitHub do Arquivo.pt](https://github.com/arquivo/Research-Websites-Preservation).



## O caso programas europeus (resumo)

- Links dos projetos **não estavam a ser preservados** sistematicamente
- O Arquivo.pt identificou e recolheu conteúdos relativos aos projetos
- Publications Office da UE integrou a informação adicional do Arquivo.pt
- Os conteúdos estão **acessíveis no Arquivo.pt**





## Programa H2020 - continuamos a preservar

- 31 129 projetos
- 46% apresentavam informação na coluna “urlProject”



## Programa H2020 – identificação e gravação

- 270 650 URLs únicos identificados (“identifiedProjectURL”)
- 197 milhões de ficheiros da web
- 17 TB



## Património digital. Preservar, salvaguardar e promover o acesso aos acervos

The screenshot shows a GitHub repository page for 'arquivo / Research-Websites-Preservation'. The repository is public and has 2 forks and 3 stars. The current branch is 'master'. The file list shows several files related to updating datasets, with the most recent update on 10 Nov 2021. The files are:

File Name	Commit Message	Time Ago
..		
Cordis-h2020organizationsComplementedByArq...	Update Datasets	8 months ago
Cordis-h2020projectDeliverablesComplemented...	Update Datasets	8 months ago
Cordis-h2020projectsComplementedByArquivoP...	Update Datasets	8 months ago
Cordis-h2020reportsComplementedByArquivoPT...	Update Datasets	8 months ago
H2020.py	Add H2020 with BingSearch results	10 months ago
H2020_Complete_With_BingSearchAPI.xlsx	Add H2020 with BingSearch results	10 months ago

[Canal GitHub do Arquivo.pt](https://github.com/arquivo/Research-Websites-Preservation), uma das plataforma onde dados das recolhas e as ferramentas foram disponibilizados publicamente

EXTMOS | EXTended Model of Organic Semiconductors

ARQUIVO.PT

extmos.eu 27 Abril às 18h26, 2017

Menu

Opções

Tabela

2017

Abril

27 Abril 18h26, 2017

2020

EXTMOS

EXTended Model of Organic Semiconductors

AIMS AND OBJECTIVES

OUTREACH AND PUBLICATIONS

METHODOLOGY

PARTNERS

CONTACT

CONFERENCES

AIMS AND OBJECTIVES

OUTREACH AND PUBLICATIONS

METHODOLOGY

PARTNERS


CONTACT

CONFERENCES

Search...

## EXTended Model of Organic Semiconductors

EXTMOS is a €5 million collaborative research project funded by the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 646176. Twelve partners from industry and academia across Europe are using world-leading computational modelling techniques to theoretically predict which of the myriad of possible materials will work best as organic semiconductors, so that only the most promising ones will need to be made for experimental testing.



EXTMOS: EXTended Model of Organi...

[O website do projeto EXTMOS do H2020 já não se encontra online mas preservado pelo Arquivo.pt](https://arquivo.pt/wayback/20170427182603/http://extmos.eu/)





# Preservação de resultados da investigação científica nacional



## Projetos financiados pela FCT

- **2016** – identificação automática de conteúdos pelo Arquivo.pt
- **2020** – adição de um campo de formulário nos relatórios dos projetos para mencionar os websites, as páginas Web e outras referências



## Dados de projetos em plataformas nacionais

- **2021** – colaboração com o PTCRIS para a extração de URLs fornecidos pelos investigadores
- **2022** – colaboração com o RCAAP para extração das referências a conteúdos Web existentes nos trabalhos depositados (em curso)



## Património digital. Preservar, salvaguardar e promover o acesso aos acervos

### Memória Web da Investigação Científica

*Sites de Unidades I&D e de Projetos financiados pela FCT*

[Início](#) [Lista de Unidades I&D](#) [Lista de projetos](#) [Sobre](#)



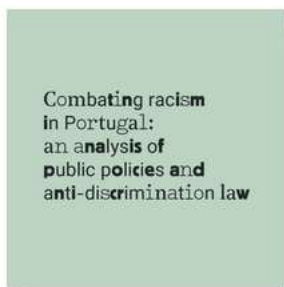
Dourozone



Quimioterápicos – Navigating through marine-derived fungi



2Bio4Cartilage



Combat



E-Dicionário de Termos Literários



Projeto MycoStone

[Memória Web da Investigação Científica](#)



## Património digital. Preservar, salvaguardar e promover o acesso aos acervos



DOCUMENTAÇÃO ▾ **DADOS ABERTOS ▾**

### Universidades e de Institutos Politécnicos: websites e histórico no Arquivo.pt

Este conjunto de dados vem de um serviço público certificado

Lista de websites das Universidades e dos Institutos Politécnicos dos setores público e privado. Identificou-se os websites destas instituições e verificou-se se já existe um histórico de versões no [Arquivo.pt](#) fornecendo para isso uma ligação. Além de uma ferramenta para aferir a preservação de conteúdos de interesse público, a lista é útil para qualquer pessoa que queira fazer um estudo recorrendo aos websites atuais ou do passado.

Recursos **Pré-visualizar**

XLSX	<b>Lista de Universidades e Institutos Politécnicos: websites e historico no Arquivo.pt</b> Lista de websites das universidades e dos Institutos Politécnicos dos setores público e privado. Identif...
100.4 KB	15 <b>Disponível</b> <b>Descarregar</b>



DOCUMENTAÇÃO ▾ **DADOS ABERTOS ▾**

### Unidades de Investigação e Desenvolvimento FCT 2019: websites e histórico de versões no Arquivo.pt

Este conjunto de dados vem de um serviço público certificado

Lista dos websites das Unidades de Investigação e Desenvolvimento com financiamento da FCT (dados de 2019). Identificou-se os websites das unidades e dos centros de investigação e verificou-se se já existe um histórico de versões no [Arquivo.pt](#) fornecendo para isso uma ligação. Além de uma ferramenta para aferir a preservação de conteúdos de interesse público, a lista é útil para qualquer pessoa que queira fazer um estudo recorrendo aos websites atuais ou do passado.

Recursos **Pré-visualizar**

XLSX	<b>Lista dos websites das Unidades de Investigação e Desenvolvimento com financiamento da FCT</b> Lista dos websites das Unidades de Investigação e Desenvolvimento com financiamento da FCT (dado...
126.0 KB	2 <b>Disponível</b> <b>Descarregar</b>





**Património digital. Preservar, salvaguardar e promover o acesso aos acervos**

## Resumo

- Arquivo.pt é um serviço da aberto à comunidade
- Preservação dos resultados de investigação de projetos científicos
  - O caso dos programas europeus
  - Iniciativa do Arquivo.pt de identificar e preservar os conteúdos
  - Disponibilização dos dados obtidos no EU Open Data Portal
  - Acesso aos conteúdos no Arquivo.pt
- Programa H2020 – a preservação é para continuar
- Preservação de resultados da investigação científica nacional
  - Projetos financiados pela Ciência e a Tecnologia
  - Dados de projetos em plataformas nacionais



## Participação dos bibliotecários e gestores de ciência na preservação dos resultados projetos científicos

- **Formação** de utilizadores sobre tema da preservação
- **Menção do Arquivo.pt na página da biblioteca** entre os recursos para a investigação
- **Identificação de conteúdos** para serem preservados (curadoria colaborativa)



## Património digital. Preservar, salvaguardar e promover o acesso aos acervos

The screenshot shows the OPAC (Online Public Access Catalog) website for NOVA-FCT. The page features a search bar at the top with the text "Search Library catalog" and a "Go" button. Below the search bar, there are links for "Advanced search", "Authority search", and "Tag cloud". The main content area displays a list of resources, including NOVA University, ARQUIVO.PT, and various online services. The ARQUIVO.PT logo is prominently displayed in the list. On the right side of the page, there is a "Log in to your account" section with fields for "Login" and "Password", and a "Log in" button. The page also includes a navigation menu with links for "Contacto", "Empréstimos InterBibliotecas", "Recursos On-Line", "Formação", and "Como chegar".

**Menção ao Arquivo.pt** na página da biblioteca da NOVA-FCT entre os recursos e serviços para a investigação. [Pode encontrar o logo do Arquivo.pt neste link](#)



## Futuro...

Professores, alunos e investigadores usam conteúdos Web preservados como material de estudo e investigação.



Património digital. Preservar, salvaguardar e promover o acesso aos acervos

# Conclusão

- O Arquivo.pt salvou do esquecimento páginas que documentam os resultados de projetos científicos
- Os conteúdos preservados estão disponíveis para a investigação científica
- Os bibliotecários e gestores de ciência são fundamentais na formação e sensibilização para a preservação e o estudo do património digital preservado





Menu

ARQUIVO.PT

 **SavePageNow** Grave páginas no Arquivo.pt

Endereço da página

Gravar

ex. [www.fcen.pt](http://www.fcen.pt)

O serviço SavePageNow 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.

#### Sobre

[O que é o Arquivo.pt?](#)  
[Publicações](#)  
[Exposições](#)  
[Área de imprensa](#)  
[API & Código aberto](#)  
[Termos e Condições](#)

#### Notícias

[Subscriver lista de email](#)  
[Últimas novidades](#)  
[Twitter](#)  
[Facebook](#)  
[Vídeos](#)  
[Material de Divulgação](#)

#### Colabore

[Prémios Arquivo.pt](#)  
[Sugira site](#)  
[Memorial](#)  
[Recomendações](#)  
[Grave páginas](#)  
[Investigação](#)

#### Ajuda

[Pesquisa de páginas](#)  
[Pesquisa de imagens](#)  
[Recolha](#)  
[Acesso](#)  
[Formação](#)  
[Contacto](#)



# Referências

[Arquivo.pt preservou sites sobre projetos de Investigação & Desenvolvimento financiados pela Comissão Europeia](#) [Site informativo do Arquivo.pt], 2017.

[Arquivo.pt preservou informação online acerca de projectos europeus financiados pelo H2020](#) [Site informativo do Arquivo.pt], 2021.

[Arquivo.pt certificado como fornecedor de dados abertos](#) [Site informativo do Arquivo.pt]

Gomes, D., Demidova, E., Winters, J., Risse, T. (2021). *The Past Web: exploring web archives*. Springer [versão pré-print]. [Ver mais informação no site informativo do Arquivo.pt](#).

[Save Page Now](#). Serviço de gravação de páginas diretamente no Arquivo.pt para uso da comunidade.



Património digital. Preservar, salvaguardar e promover o acesso aos acervos

Obrigado!

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

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



associação portuguesa de  
**bibliotecários, arquivistas,  
profissionais da informação  
e documentação**