

**Novos Caminhos
para a Preservação
e o Acesso à Informação**

7 novembro



Universidade do Minho

Serviço de Documentação e Bibliotecas

PRESERVAÇÃO DIGITAL DOS RESULTADOS DE INVESTIGAÇÃO

realizações, lacunas e desafios em Portugal

Pedro Príncipe | 07-11-2024

TÓPICOS EM DISCUSSÃO

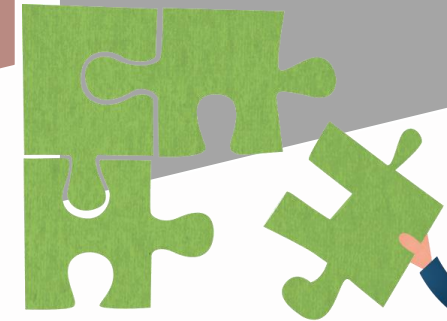
COM FOCO NO ACESSO ABERTO

Revistas
académicas e
científicas

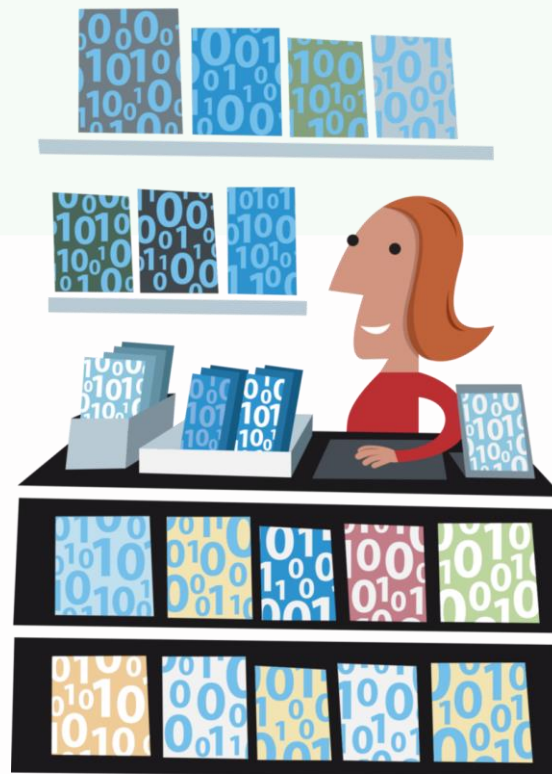
Repositórios
de literatura
científica

Dados de
investigação

Projetos de
investigação



REALIZAÇÕES, LACUNAS E DESAFIOS EM PORTUGAL



Digitalbevaring.dk

REVISTAS ACADÉMICAS E CIENTÍFICAS

Caminho sustentado em iniciativas institucionais, com recursos escassos mas alavancado por projetos guiados pela comunidade e iniciativas recentes nacionais (PUB IN)...

Preservação do conteúdo das revistas

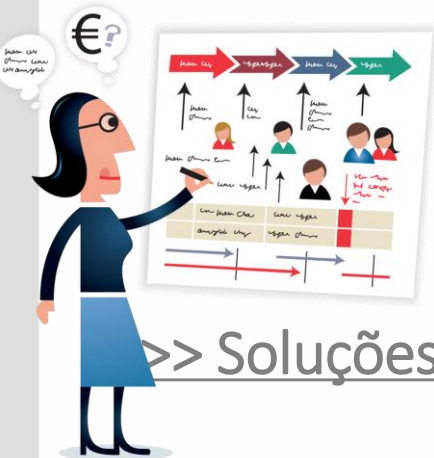
- ✓ As revistas são ainda centrais na comunicação científica e são (quase) todas publicadas apenas em formato digital;
- ✓ Os editores têm que assegurar não só a sua publicação, mas também a sua continuidade, garantindo a sua **preservação digital**, para que os autores e leitores possam aceder hoje, mas também no futuro.

“Aberto não é para sempre”*

**“lack of preservation infrastructure
for many smaller academic journals”**

M. Laakso, M. Matthias, N. Jahn. *Open is not forever: A study of vanished open access journals*. Journal of the Association for Information Science and Technology, Feb. 2021. <https://doi.org/10.1002/asi.24460>

>> Soluções partilhadas como boas práticas na comunidade PUB IN (revistas de AA em PT)



ALGUNS DADOS DE CONTEXTO

(dados de Peter Burnhill - advisor to ISSN-IC)

Digital Scholarly Journals Are Poorly Preserved: A Study of 7 Million Articles (Martin Eve, 2024)

CrossRef database used to extract details of 7.4m DOIs checked against metadata from 8 larger archiving organisations of the 15 reporting into Keepers Registry.

Findings

Publishers ["CrossRef members"] classified into four bands:

	Definition	N	%
Gold	75% of titles in 3+ archives	204	1.0
Silver	50% 'preserved' in 2+ archives	1,797	8.5
Bronze	25% of titles in at least one archive	12,257	57.7
Unclassified	no preservation found	6,982	32.9

Martin Eve: 58.38% (4.34/7.4m) "present in at least one archive"
(but he excludes 14% DOIs as not journal articles or too recent/insufficient date metadata.)

So, in study of 6m articles, 67.86% (4.34m/6.40m) "present in at least one archive"

This is to the credit of the 'Big Bad Publishers' who pay Portico & CLOCKSS!!

Journal of Librarianship and Scholarly Communication, 12(1) <https://doi.org/10.31274/jlsc.16288>

Success Criteria for KEEPERS REGISTRY

Number of titles recorded in the Registry: 84,715

Celebrating Progress in Archival Coverage, 2012 - 2023

	Ingested & archived (1+)	Ingest Ratio %	Ingested & archived (3+)	KeepSafe Ratio %	Keepers reporting	Keepers then not reporting
June 2012	17,240	.	.	.	7	
June 2015	28,103	16.9	11,328	6.8	9	e-depot (KB)
August 2018	39,117	.	15,305	.	13	
June 2023	66,704	18.6	.	.	14	ADS, BL, Swiss
December 2023	69,878	17.7	21,077	5.3	15	ADS, BL, Swiss
Indexed as 'Scholarly'						
December 2023	43,335	37.6	21,076	18.3	15	ADS, BL, Swiss

Progress with 'scholarly', but still, 60+ % could be at risk of loss!

Preservação das revistas de acesso aberto em Portugal

- ✓ A **inexistência** de iniciativas e orientações gerais no contexto da preservação digital a nível nacional **obriga à exploração de soluções e alternativas viáveis** para a comunidade acompanhar as tendências e boas práticas.
- ✓ No âmbito do **iniciativa nacional PUB IN** (projeto SAMA PIAPC coordenado pela UMinho com colaboração da FCT-FCCN), foi desenvolvido um piloto em 2021/22 para demonstrar a aplicabilidade de um sistema open source para operações básicas de preservação digital.
- ✓ O piloto (2021/22) com revistas selecionadas do serviço SARC (RCAAP, agora PUB IN) – software **LOCKSS** para o contexto de revistas científicas com o **software Open Journal Systems (OJS)**
 - LOCKSS - <https://www.lockss.org/>
 - projeto assenta em pressupostos elementares de preservação e persistência de acesso aos dados, permitindo a criação de redes para assegurar o conceito **LOCKSS (Lots of Copies Keep Stuff Safe)**.



JASPER - JournALS are Preserved forever

O Projeto JASPER (JournALS are Preserved forever), desenvolvido por parceiros que inclui o **DOAJ**, **CLOCKSS**, **PKP**, **Internet Archive**, **Keepers Registry**, definiu 3 vias para os editores poderem preservar digitalmente:

- i. **Via OJS:** a Rede de Preservação do PKP com o plugin do Open Journal System, baseado no **software LOCKSS (Lots Of Copies Keep Stuff Safe)**;
- ii. **Via Upload de ficheiros:** para o **CLOCKSS**, o Arquivo da Internet Archive ou outro 'Keeper';
- iii. **Via Web-crawling:** maior esforço possível para assegurar agregação da web pelo **Arquivo da Internet**.



Preservação de revistas - qual a melhor opção?

- Como a maioria das revistas portuguesas tem **equipas editoriais e orçamento reduzidos**, mas muitas utilizam já o *Open Journal Systems* como plataforma de gestão editorial e estão registadas no DOAJ,
- **LOCKSS (Lots Of Copies Keep Stuff Safe)** é uma excelente forma de garantir a continuidade do conteúdo destas revistas.

Configurações Website

Aparência

Configurar

Plugins

Plugins Instalados

Galeria Plugin

Plugins Genéricos (1)

▼ PKP PN Plugin

The PKP PN plugin will deposit your published content into the PKP Preservation Network.



Configurações

Situação

Eliminar

Actualizar

Preservação de revistas - qual a melhor opção?

Já ativei o meu plugin e agora?

- Quando as revistas são preservadas digitalmente, aparecem na lista do e terão no portal do ISSN, o logotipo do **Keepers**.

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INTERNATIONAL CENTRE

ISSN PORTAL The Global Index for Continuing Resources

Identifiers

ISSN :2795-5915

Linking ISSN (ISSN-L): 2795-5915

Links

URL

Key-title Child studies (Braga)

Resource information	Archival Status
 Title proper: Child studies. Country: Portugal Medium: Online	

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INTERNATIONAL CENTRE

ISSN PORTAL The Global Index for Continuing Resources

Identifiers

ISSN :2183-3575

Linking ISSN (ISSN-L): 1645-2089

Links

URL

Google

Bing

Yahoo

MIR@BEL

ISSN Portugal

Key-title Comunicação e sociedade (Braga. Em linha)

Status	Publisher	Keeper	From	To	Updated	Extent of archive
Preserved	Centro de Estudos de Comunicação e Sociedade (CECS) da Universidade do Minho	Public Knowledge Project PLN	1999	2024	04/11/2024	

Preserved : 0; 1; 2; 3; 4; 5; 6; 7; 8; 9 (10); 11; 12; 13; 14; 15; 16; 17; 18; 19; 20; 21; 22; 23; 24; 25; 26; 27; 28; 29; 30; 31; 32; 33; 34; 35; 36; 37; 38; 39; 40; 41; 42; 43; 44; 45; 46

Related titles

Resource network

Resource history

Has other medium version: Comunicação e sociedade, 1645-2089

Record information

Type of record: Confirmed
Last modification date: 26/05/2023
ISSN Center responsible of the record: [ISSN National Centre for Portugal](#)
For all potential issues concerning this bibliographic record (missing or wrong data etc.), please contact the ISSN National Centre mentioned above by clicking on the link.
Record creation date: 28/11/2014
Original ISSN Centre: ISSN National Centre for Portugal

JASPER: JOURNALS PRESERVED

About the service partners

CLOCKSS

CLOCKSS is a not-for-profit collaboration of leading academic publishers and research libraries. Our mission is to ensure the long-term survival of digital scholarly content and instill confidence in authors, scholars, policymakers, libraries and publishers worldwide that their content will be safely and securely preserved for future generations. We are entrusted with the long-term preservation of more than 51 million journal articles and 400,000 scholarly books. CLOCKSS operates 12 archive nodes at leading academic institutions worldwide. This secure, robust, and decentralized infrastructure can withstand threats from technological, economic, environmental, and political failures. A destructive event in one location doesn't jeopardize the survival of preserved digital content because the 11 other locations serve as mirror sites to back up and repair the disrupted location's archive. Content is triggered from our dark archive if it disappears from the web and is made available to everyone under an Open Access license.

DOAJ

DOAJ is the largest, most diverse, free directory of open access journals in the world. DOAJ's mission is to increase the visibility, accessibility, reputation, usage and impact of quality, peer-reviewed, open access scholarly research journals globally, regardless of discipline, geography or language. By assisting journals to become archived in long-term digital preservation schemes, DOAJ fulfils its mission by ensuring accessibility, visibility and usage of these crucial scholarly records.

Internet Archive

Internet Archive is a non-profit digital library with the mission of "Universal Access to All Knowledge." Since 1996, it has worked with over 1,000 universities, libraries, archives, governments, non-profits, and other memory and social good organizations for preservation and access to the cultural and scholarly record in digital form. It owns and operates its own non-profit data centers, stewards over 50 petabytes of archived data, and provides open-source tools and services for digitization, web archiving, digital preservation, web, and data services, including the Wayback Machine and IA Scholar.

Keepers Registry

Keepers Registry is the global monitor on the archiving arrangements for continuing resources managed by the ISSN International Centre and fuelled by reports from contributing archiving agencies.

PKP

A research and development initiative within Simon Fraser University's Core Facilities Program, the Public Knowledge Project (PKP) develops (free) open-source software. It carries out research to improve the quality and reach of scholarly publishing. As part of its mandate to make knowledge public, PKP understands that preserving and maintaining long-term access to published content is paramount. To support this, PKP can be a valuable partner.



KEEPERS REGISTRY

Information about archiving agencies which act as Keepers

The term 'Archiving Agency' refers to an organisation or an initiative that is running a programme for the long-term archiving of e-journals and other continuing resources and can report about the archived materials using the ISSN. Each agency participating in the Keepers Registry regularly provides the ISSN International Centre with metadata on their holdings. This metadata is then included in the ISSN Register which is the information source for the ISSN Portal

The Keepers Registry is collecting data from the following archiving agencies:

Archaeology Data Service
British Library
Cariniana Network
CLOCKSS Archive
Global LOCKSS Network
HathiTrust
Internet Archive
Library of Congress
Merritt Preservation Repository at the California Digital Library
National Digital Preservation Program, China
National Library of France
National Library of Spain
National Library of the Netherlands
PKP PN
Portico
Scholars Portal
Swiss National Library
TIB Leibniz Information Centre for Science and Technology
ZBW – Leibniz Information Centre for Economics

Each participating agency provides below a description of their organisation's approach. This highlights the different operating models and approached to preservation and access conditions. In order to understand each archiving agency in detail, it is recommended that you visit each initiative's website directly.

keepers.issn.org



CLOCKSS 5 ESSENTIAL FACTS

1. It's a Dark Archive

A dark archive is a secure and restricted repository designed for the long-term preservation of digital content. Unlike traditional archives that are actively accessed and used, a dark archive remains largely inaccessible to the public, only being used for restoration or recovery purposes in the event of a disaster, data corruption, or other significant disruptions.

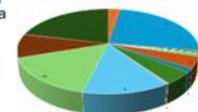


2. It's International

CLOCKSS is dedicated to preserving digital scholarly content from around the world, ensuring its long-term accessibility and integrity. By collaborating with a global network of libraries and publishers, CLOCKSS safeguards a diverse array of academic and scientific resources, protecting them against loss or obsolescence.

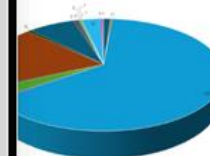
3. It's Multidisciplinary

CLOCKSS preserves a vast array of resources that span multiple disciplines, reflecting its commitment to supporting a multidisciplinary academic landscape. By safeguarding content from diverse fields such as the sciences, humanities, social sciences, and more, CLOCKSS ensures that researchers from various disciplines have access to essential scholarly materials.



4. It's multilingual

CLOCKSS is dedicated to preserving scholarly materials in multiple languages, ensuring the accessibility and longevity of global academic content. By collaborating with international publishers and institutions, CLOCKSS safeguards a diverse array of research and publications, reflecting the rich linguistic diversity of the academic community.



5. It's dedicated to Diversity, Equity and Inclusion

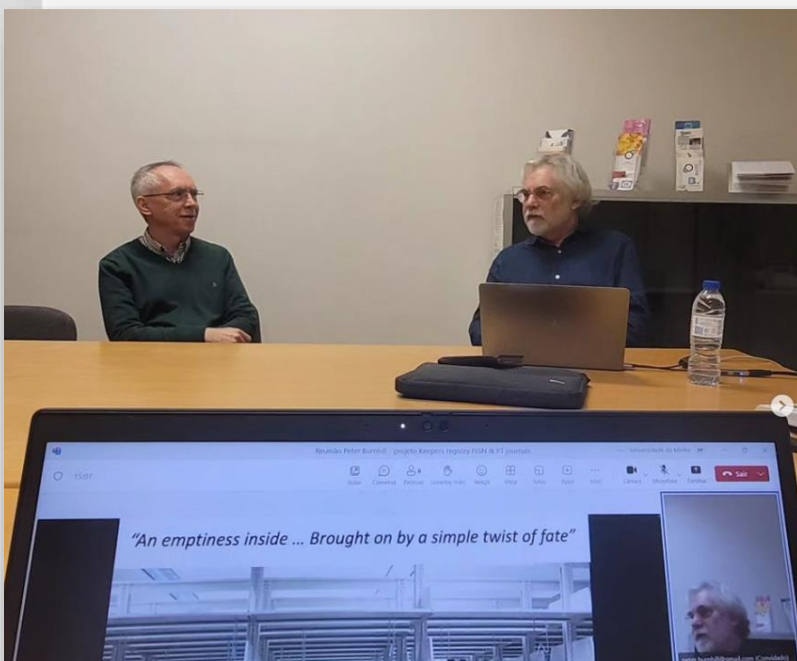
CLOCKSS is dedicated to Diversity, Equity, and



**ABORDAGENS EM REDE (MULTI INSTITUCIONAL / MULTI CUSTODIAIS) PARA A
PRESERVAÇÃO DIGITAL DE RESULTADOS DE INVESTIGAÇÃO**

E a nível nacional, o que pode ser feito?

- Disseminar pelas centenas de revistas existentes em Portugal quanto à necessidade das revistas serem preservadas digitalmente.
- Por outro lado, vários países encontram-se já na lista do Keepers Registry, com agências nacionais de preservação digital.



pedroprincipe
Universidade do Minho

pedroprincipe Today we welcomed @peter.burnhill for a brief visit but very useful for scientific journals in Portugal #KeepersRegistry #archivingjournals
Editada · 37 sem · Ver tradução

PORTUGAL TEM UMA AGÊNCIA NACIONAL DE ISSN, MAS NÃO É AINDA UM DOS PAÍSES COM PRESENÇA NO KEEPERS REGISTRY.

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INTERNATIONAL CENTRE

K
Keepers Registry

Good News, Bad News & Something To Do ...

Peter Burnhill
Advisor to ISSN-IC

Em Portugal...



- ✓ Vamos abrir a discussão com a Agência Nacional de atribuição de ISSN, a **Biblioteca Nacional**, ou no quadro do Arquivo.pt, para que a nível nacional seja possível ter uma política de preservação digital comum,
 - não só para as revistas científicas, mas para todas as publicações periódicas publicadas em Portugal
- ✓ e que esta fizesse o envio sistemático para o **The Keepers' Registry**.



Archiving organization	Date of latest update	Total
Merritt Preservation Repository	2024-11-04	125
Public Knowledge Project PLN	2024-11-04	6,554
HathiTrust	2024-11-01	13,408
Technische Informationsbibliothek	2024-10-29	768
CLOCKSS Archive	2024-10-28	25,536
LOCKSS Archive	2024-10-28	14,996
National Digital Preservation Program, China	2024-10-21	8,798
Gallica	2024-10-11	19,544
Cariniana Network	2024-09-30	2,226
ZBW - Leibniz Information Centre for Economics	2024-09-05	877
Library of Congress	2024-08-30	10,283
Portico	2024-04-28	35,503
Internet Archive	2024-04-26	15,015
Scholars Portal	2024-04-02	21,209
Archaeology Data Service	2024-03-15	31
National Library of the Netherlands	2023-10-20	10,760
BNE - Biblioteca Nacional de España	2023-08-08	2,799
Swiss National Library	2020-01-08	370
British Library	2019-09-18	2,385

RCAAP COMO RESPOSTA ROBUSTA E SUSTENTÁVEL

Principais Serviços RCAAP

Alojamento de
repositórios
(SARI)



28 repositórios de
publicações alojados

Alojamento
de revistas
(SARC)



38 revistas
científicas alojadas

Repositório
Comum



100 Instituições
aderentes

Portal de
Pesquisa



377 recursos
agregados

Formação, Disseminação e Helpdesk

- A rede de repositórios institucionais assegura preservação e acesso.
- São boa prática e um caso de estudo europeu.
- Feito trabalho relevante há 10 anos: *“Auditoria ISO 16363 a repositórios institucionais – 28 RCAAP”*

Auditoria aos Repositórios Científicos Institucionais de Portugal

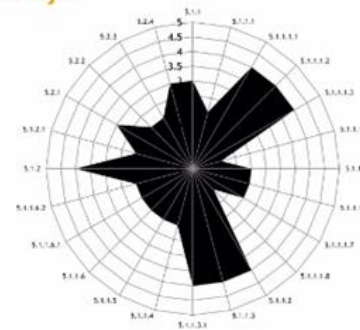
Plano geral (duração 1 ano)

Fase	Designação	Descrição	Data
1	Diagnóstico	Auditoria preliminar a todos os repositórios	Nov. 2013 - Fev. 2014
2	Plano de Ações	Identificação das ações para melhorar a conformidade do repositório com a norma	Março 2014
3	Implementação		
4	Auditoria Final		

Principais dificuldades

- Existência de um plano de sucessão
- Ausência de uma política de preservação digital
- Desenvolver a política de preservação
- Prever auditorias internas
- Garantir os direitos de propriedade intelectual

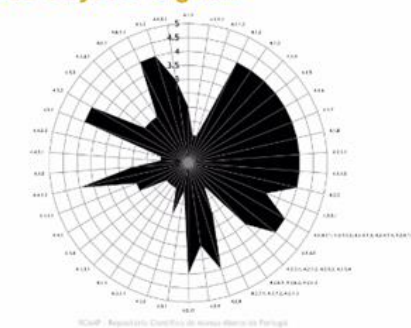
INFRAESTRUTURA E GESTÃO DA SEGURANÇA



Principais Dificuldades

- Definir e documentar processos para, por exemplo, atualizar as aplicações/software ou substituição de hardware (discos, memória, etc.)
- Falta de evidências de financiamento para atualização de hardware, software.

Gestão de Objetos Digitais



Principais Dificuldades

- Registo de mais informação sobre os conteúdos existentes e a forma como são alterados, transformados ou eliminados.
- Definição de processos documentados
- Falta de ferramentas / processos de Preservação Digital (gestão de AIPs)

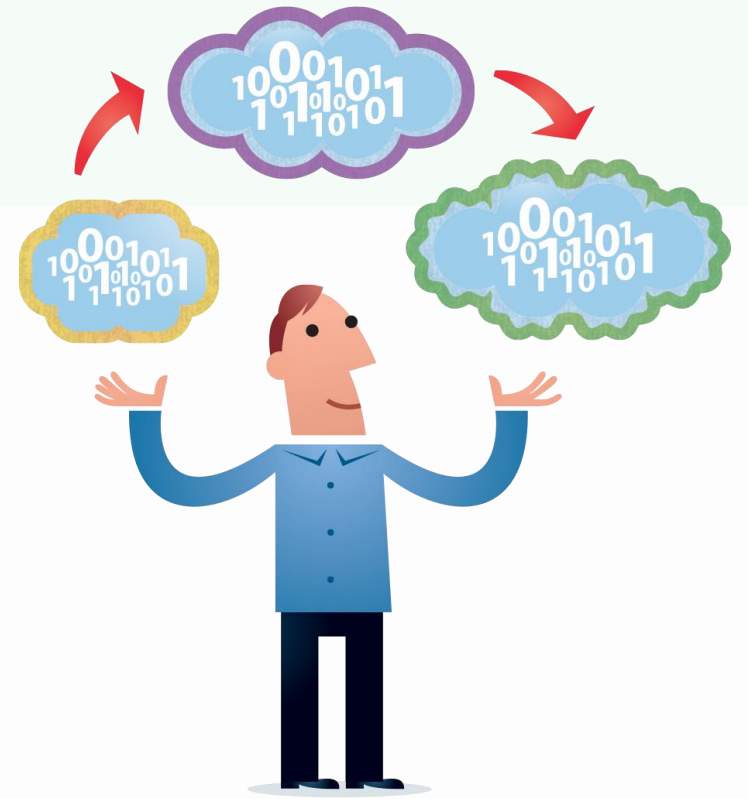
FALTAM RECURSOS E UM RUMO PARA A CERTIFICAÇÃO



Digitalbevaring.dk

- Estudo europeu (OpenAIRE, COAR, LIBER, SPARC) recente prova falta de investimento e estruturas robustas, em Portugal também estamos necessitados apesar do RCAAP.
- Processo sério e necessário de certificação é o rumo certo, com políticas e estratégias institucionais bem definidas.





Digitalbevaring.dk

DADOS DE INVESTIGAÇÃO

Caminho a alavancar iniciativas relevantes, suportado pela preocupação de financiadores de ciência, muito desafiante para as instituições, carente de infraestruturas confiáveis...

DADOS DE INVESTIGAÇÃO

diagnóstico por fazer, problemas comuns identificados

DIFICULDADE EM ENCONTRAR DADOS

- Fez o trabalho árduo de produzir a sua investigação e os resultados estão registados em dados. Seria um desperdício se não pudesse ser encontrado por outros, certo?

PERDA DE DADOS OU ACESSO NÃO AUTORIZADO

- Os dados de investigação armazenados incorretamente podem ser suscetíveis de perda de dados ou de acesso não autorizado.

BARREIRA À INTEROPERABILIDADE

- A combinação e a integração dos seus dados de investigação com outros dados de investigação gerará mais conhecimentos.

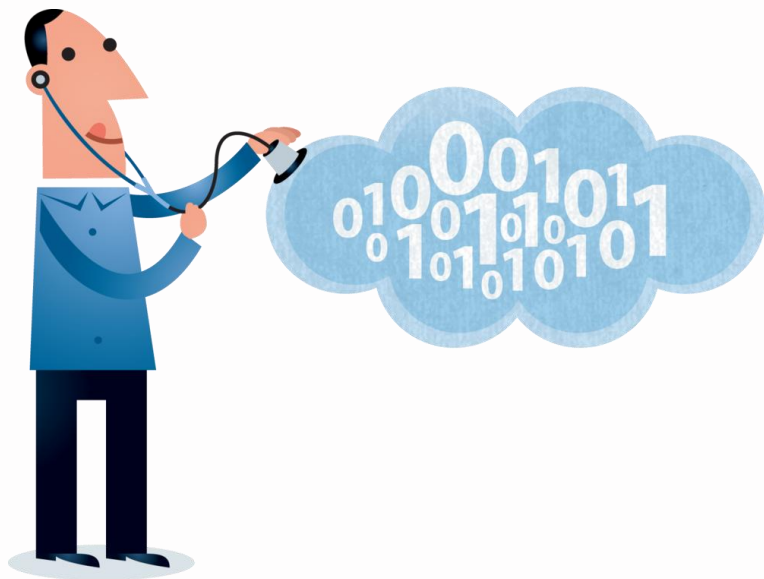
FALTA DE PERSISTÊNCIA DOS DADOS

- É frequente a necessidade de mover ficheiros de um espaço de armazenamento para outro (dados de investigação ou ficheiros pessoais).

FALTA DE PRESERVAÇÃO A LONGO PRAZO

- Não se trata apenas de garantir que os investigadores no futuro possam utilizar os seus resultados.

ESTAMOS A FAZER CAMINHO... FALTAM ESTRATÉGIAS INSTITUCIONAIS DE GDI



Digitalbevaring.dk

- Sinais +: crescimento do nº de repositórios de dados de instituições e infraestruturas, iniciativa Polén, PNCA – DAI.
- Definição de estratégias institucionais, roadmaps, políticas – urgente.



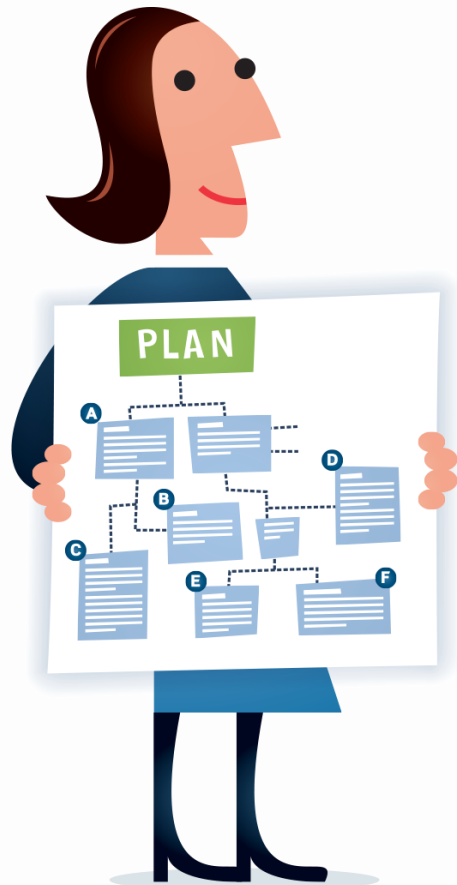


Digitalbevaring.dk

PROJETOS DE INVESTIGAÇÃO

Relevar a necessidade das instituições e financiadores providenciarem instrumentos e requisitos que assegurem boas práticas de gestão e preservação dos resultados...

PLANOS DE GESTÃO DE DADOS E ARQUIVO.PT PARA WEB DOS PROJETOS



Digitalbevaring.dk

- Plano de Gestão de Dados como requisito dos financiadores e instituições será um excelente recurso!
- Arquivo.PT vem salvar as páginas web dos projetos.

OS DESAFIOS EM PORTUGAL: **o que falta fazer...**

Um caminho com foco, ambição e necessário investimento.



Capacitação

Desenvolvimento de **competências** e programas para apostar em mais profissionais habilitados.

Data Stewards, Data Curators, Gestores de Repositórios e de Revistas, Editores



OS DESAFIOS EM PORTUGAL: o que falta fazer...

Políticas

**Estabelecimento de
políticas de Ciência Aberta
(para além do acesso
aberto a publicações).**

Políticas do financiador (FCT),
de organismos nacionais e nas
Instituições.



**OS DESAFIOS EM
PORTUGAL: o que
falta fazer...**

Sustentabilidade

Ambicionar **infraestruturas abertas** mais sustentáveis, partilhadas, baseadas em modelo de rede/consórcio, “com menos tutela”.

Projetos dedicados, mas estamos no bom caminho.



OS DESAFIOS EM
PORTUGAL: o que
falta fazer...



GESTÃO DE INFORMAÇÃO CIENTÍFICA,
REPOSITÓRIOS E CIÊNCIA ABERTA
openscience@usdb.uminho.pt



Universidade do Minho

Serviço de Documentação e Bibliotecas

OBRIGADO

pedro.principe@usdb.uminho.pt