

Aspectos teóricos da Preservação Digital no cenário internacional

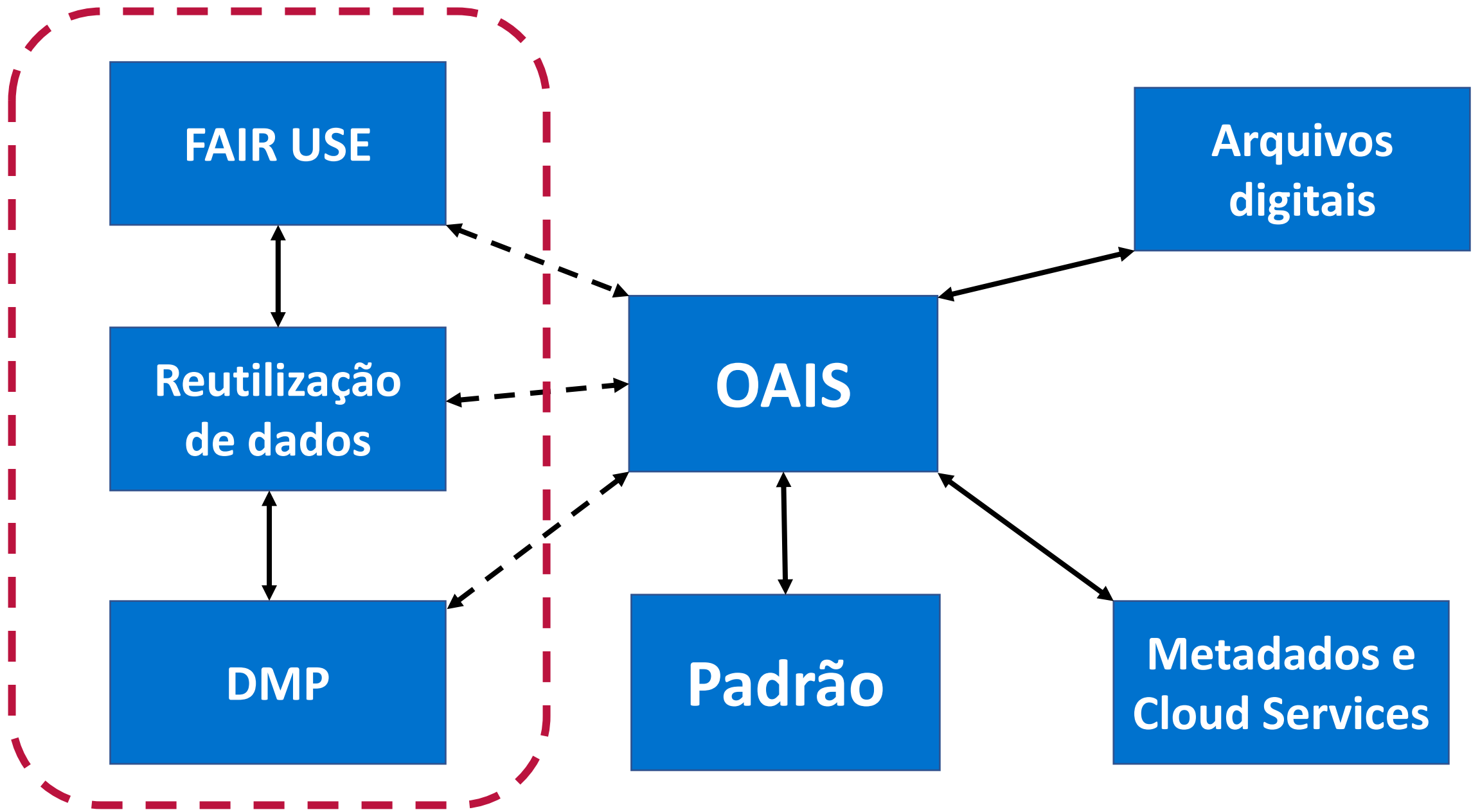
Juan-José Boté juanjo.botev@ub.edu

Universitat de Barcelona



V SEMINÁRIO INTERNACIONAL DE PRESERVAÇÃO
DIGITAL

12-13 MAIO - Universidade Estadual de Campinas - Unicamp



DMP - Research Data Management Plan

Heranças

```
graph TD; A[Heranças] --> B[Cultural]; A --> C[Científica]
```

Cultural

Humanidades
Digitais

Trabalhos artísticos

Obras audiovisuais

O som funciona

Etc,

Científica

Replicação de pesquisa

Preservação do conjunto de dados
Oceanográfico (Wulff, 2020)
Dados Espaciais

Outras

Tendências na produção científica mundial

No cenário internacional, as áreas de pesquisa mais relevantes são:

- Sistemas ponto a ponto, por exemplo, LOCKSS
- Estratégias de preservação digital
- Redes de dados federadas
- Análise de arquivos e repositórios da web

(Patra, Sahoo e Mohanty, 2021)

Padrão

Eles são essenciais na Preservação Digital

- i) Definição de requisitos e resultados mensuráveis
- ii) Facilitar a interoperabilidade

Modelos conceituais

- OAIS (Open Archival Information System) – ISO 14721:2012
 - “Digital Preservation is the act of maintaining information, Independently Understandable by a Designated Community, and with evidence supporting its Authenticity, over the Long Term”
- Archives Performance Model

Eles são modelos de referência

Padrão

Metadados

- PREMIS
- METS
- Australian Government Recordkeeping Metadata Standard (AGRkMS)
- LMER (Long-term preservation Metadata for Electronic Resources) de Alemanha
 - Aborde perguntas específicas: p. por exemplo, migração de formatos.

Padrão

Formatos de arquivo

- ISO/IEC 26300: Open Document Format for Office Applications
- ISO/IEC 15948: Portable Network Graphics
- TIFF (Tagged Image File Format) Revision 6.0

Propiedades significativas

- PRONOM (UK) - <https://www.nationalarchives.gov.uk/PRONOM/Default.aspx>
- Formato Audiovisual: MPEG-21

Padrão

Análise de organizações (riscos, salvaguardas e outros)

TRAC – ISO 16363 (Trusted Repository Audit Checklist)

NDSA – Define 4 níveis de análise

DIN 31644 – Critérios de arquivo digital confiável

DIN 31645 – Transferência de informação

DIN 31646 – Confiabilidade do identificador persistente

Padrão

Análise de organizações (riscos, salvaguardas e outros)

ISO 20652:2006

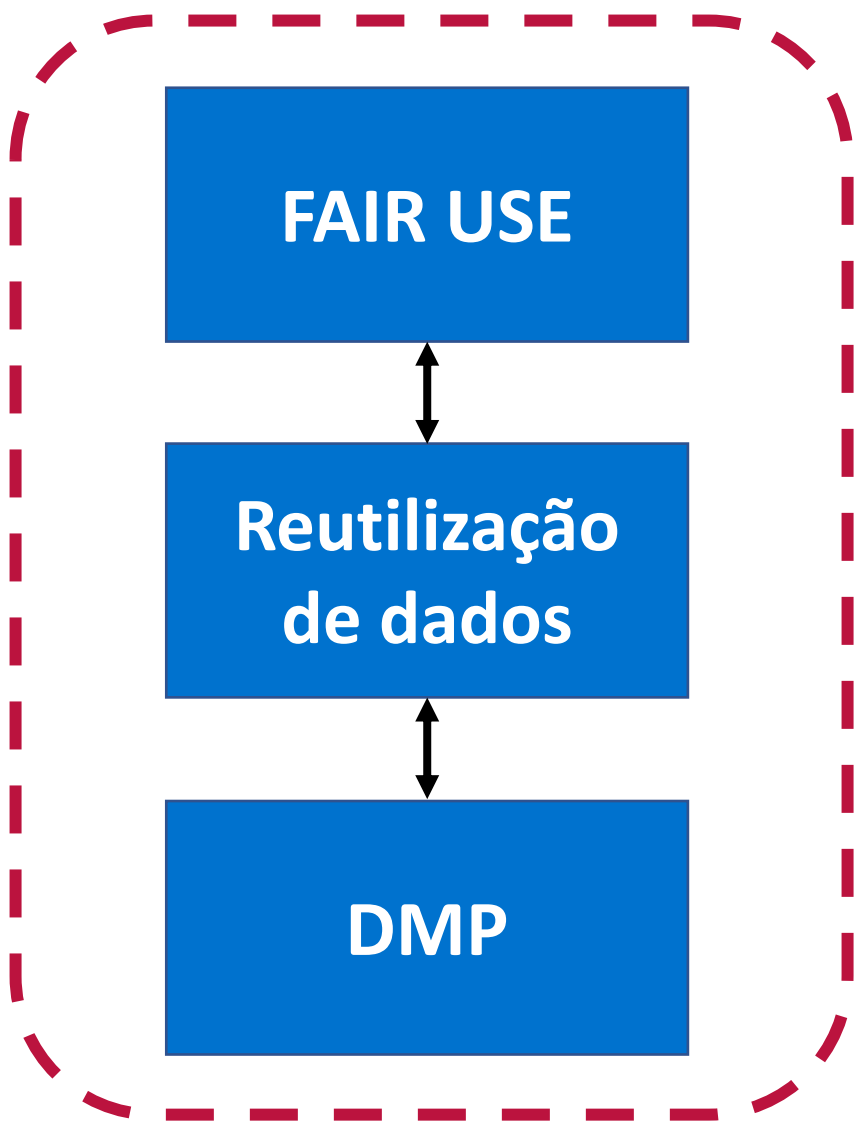
Productor-Interfície del Archivo

ISO/TR 18492:2005

Preservação de informações eletrônicas de longo prazo

AS/NZS 27001:2006, 27002:2006, 27005:2012

Requisitos de segurança



FAIR data (Findable, Accessible, Interoperable, and Reusable)

(Wilkinson et al., 2016)

Eles também são um requisito para a retenção de dados de pesquisa a longo prazo.

Uma consequência são os planos de gerenciamento de dados de pesquisa. Na Europa 2021-2027, há mudanças nos requisitos

https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

<https://webgate.ec.europa.eu/funding-tenders-opportunities/display/OM/Online+Manual>

Requisitos básicos

- Manter uma política de preservação nas Instituições
- Garantindo integridade e autenticidade
- Mantenha o contexto (talvez nem sempre seja possível)
- Manter a proveniência - (de onde vem, a cadeia de custódia)
- Manter a recuperação

Desafios de preservação

- Adoção de formatos abertos
- Estratégias institucionais
- Controle de custo
- Que está em conformidade legal
 - Leis da Espanha - Esquema Nacional de Seguridad RD 3/2010
 - Leis do Brasil - Lei.12527 de 2011
- Existem países como a Índia (e outros) onde ainda não existe legislação (Katre, 2012)

Desafios de preservação

- Treinamento dos empregados
- Metadados de preservação digital
- Redes de colaboração

Ej: Rede Cariniana en LOCKSS

(Márdero, 2018; Formenton e de Souza, 2020)

O que deve ser garantido no PD?

- Accesibilidad
- Integridad
- Autenticidad
- Seguridad
- Bitstreams
- ...

E a experiência do usuário ? (Abrams, 2021)

Desafios da preservação digital

- Obsolescência
- Propriedade intelectual e obrigações legais (afeta vários processos: armazenamento, cópia e outros)
- Validação e autenticidade dos conteúdos
- O uso de formatos apropriados
- Serviços na nuvem (Cloud Services)
- Propriedades significativas
- Recursos financeiros

Metadatos y Cloud Services

PDS Cloud

Permite a conexão entre OAIS e diferentes nuvens

Ex: permite que você armazene AIP em diferentes nuvens

Em um sistema de Cloud Services, os metadados garantem a identidade e preservação dos documentos digitais

(RABINOVICI-COHEN, S. et al., 2013; de Castro y Vesú, 2021)

Algumas iniciativas (inter)nacionais

JISC (UK) – **Preservation**. Serve para preservar qualquer objeto digital. É baseado em Artefactual e Preservica, (SaaS) <https://www.jisc.ac.uk/preservation> (2021)

Uma consequência da rede da preservação

Archives in the UK/Republic of Ireland & AI (AURA) – Rede de projetos de pesquisa multidisciplinares na área cultural(<https://www.aura-network.net/>)

WARCnet – Web Archives Studies Network (<https://cc.au.dk/en/warcnet/>)

National and State Libraries Australia (NSLA)

State Library of South Australia - **Jester**, software de otimização de processo de ingestão (Said, 2021)

Conclusões

Ainda há desafios pendentes

- O papel da Inteligência Artificial
- Processamento de linguagem natural em coleções nascidas digitais (Goodman, 2019)
- 3D
- Tecnologia Blockchain

Referencias

CASTRO, F. F. de; ALVES, R. C. V. Cloud services e o padrão PREMIS : rumos para a preservação digital. RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação, Campinas, SP, v. 19, n. 00, p. e021001, 2021. DOI: 10.20396/rdbci.v19i00.8661384. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8661384> .

FORMENTON, D.; GRACIOSO, L. de S. . Preservação digital: desafios, requisitos, estratégias e produção científica. RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação, Campinas, SP, v. 18, n. 00, p. e020012, 2020. DOI: 10.20396/rdbci.v18i0.8659259. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8659259> .

Goodman, M. An Exploration of BitCurator NLP: Incorporating New Tools for Born-Digital Collections. Disponível em: <https://saaers.wordpress.com/2019/07/02/an-exploration-of-bitcurator-nlp-incorporating-new-tools-for-born-digital-collections/>

Katre, D. (2012). Need of Legislation and Digital Preservation Policy Framework in Indian Context. DESIDOC Journal of Library & Information Technology, 32(4). <https://doi.org/10.14429/djlit.32.4.2526>

Márdero, Miguel Angel, (2018). Cariniana network methodology. Disponível em: <https://www.slideshare.net/redecariniana/cariniana-network-methodology>

Patra, Sasmita; Sahoo, Jyotshna Dr.; and Mohanty, Basudev, "Research on Digital Preservation: An empirical analysis" (2021). Library Philosophy and Practice (e-journal). 4693. <https://digitalcommons.unl.edu/libphilprac/4693>

S. Rabinovici-Cohen, J. Marberg, K. Nagin and D. Pease, "PDS Cloud: Long Term Digital Preservation in the Cloud," 2013 IEEE International Conference on Cloud Engineering (IC2E), 2013, pp. 38-45, doi: 10.1109/IC2E.2013.23 .

Said, A. (2021). NSLA's Digital Preservation Network: Looking ahead in 2021. Disponível em: <https://www.dpconline.org/blog/nsla-s-digital-preservation-network>

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). <https://doi.org.sire.ub.edu/10.1038/sdata.2016.18>

Wulff, E. (2020). Open Standards used in Oceanography Research Spatial Data Repositories in Spain. DESIDOC Journal of Library & Information Technology, 40(05), 306-312. <https://doi.org/10.14429/djlit.40.05.15924>