



# Preservação de Conteúdo Web

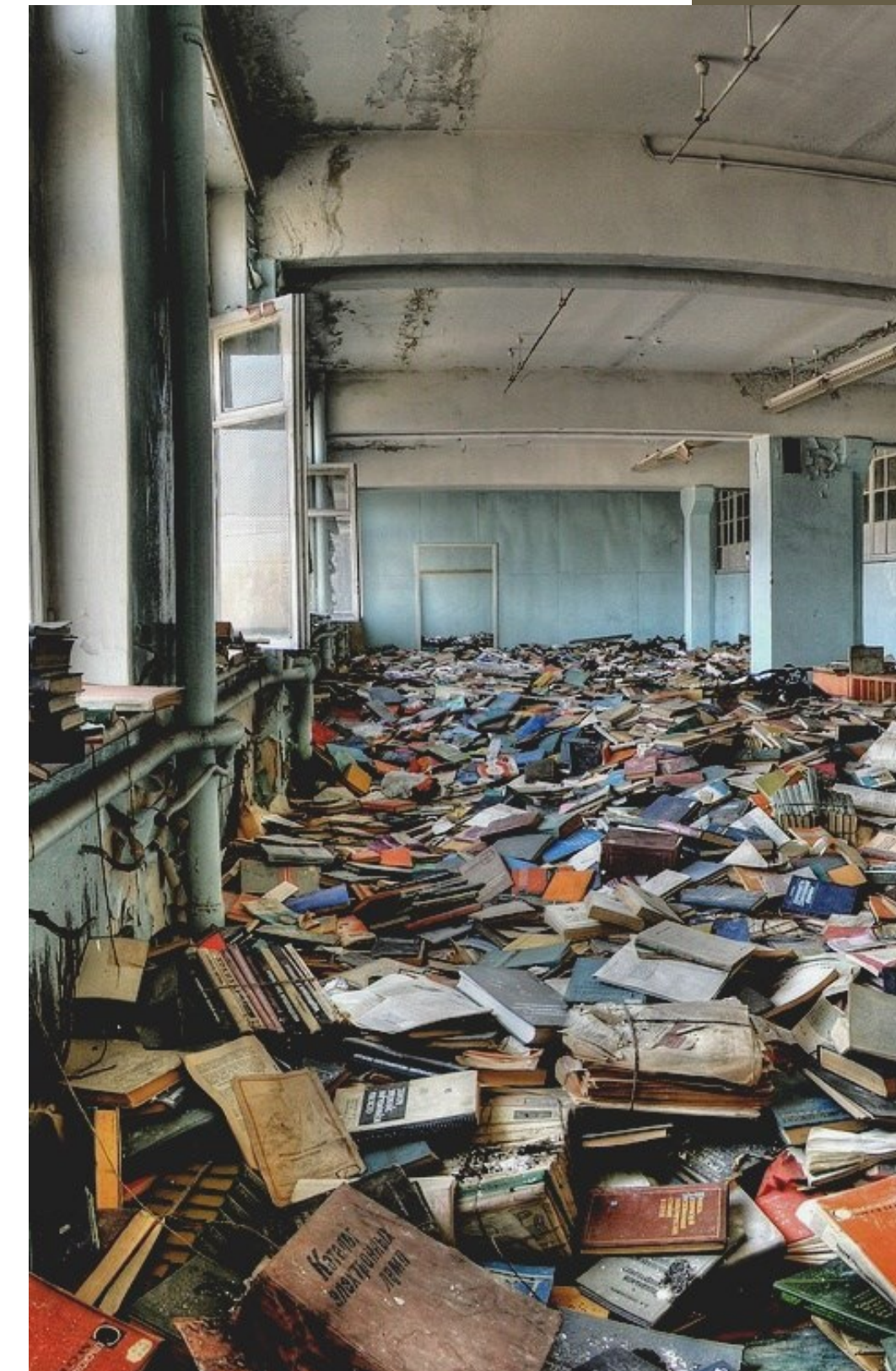
Aquiles Alencar Brayner

Universidade Federal do Cariri (UFCA)

alencarbrayner@gmail.com

# Desafio das instituições de memória na era digital

- Selecionar, capturar, descrever e preservar objetos digitais, oferecendo novos sistemas de busca e plataformas de acesso e re-uso destes objetos;
- Promover seus acervos entre diferentes tipos de usuários, compreendendo e integrando a necessidade informacional de diferentes grupos às coleções institucionais (democratização da informação);
- Proporcionar a integração de distintos formatos eletrônicos e criação de ferramentas que facilitem o acesso e análise de conteúdos digitais;
- Participação ativa nos projetos de pesquisa digital, gerenciando as atividades e disseminação de resultados obtidos;
- Instituições de memória como “laboratórios” de novas idéias e modelos em pesquisa e educação no ambiente digital;
- Atuar como centros de capacitação de estudantes, docentes e pesquisadores no manejo de objetos e dados eletrônicos.



# Novos comportamentos informacionais: Serendipidade

serendipidade | s. f.

## se·ren·di·pi·da·de

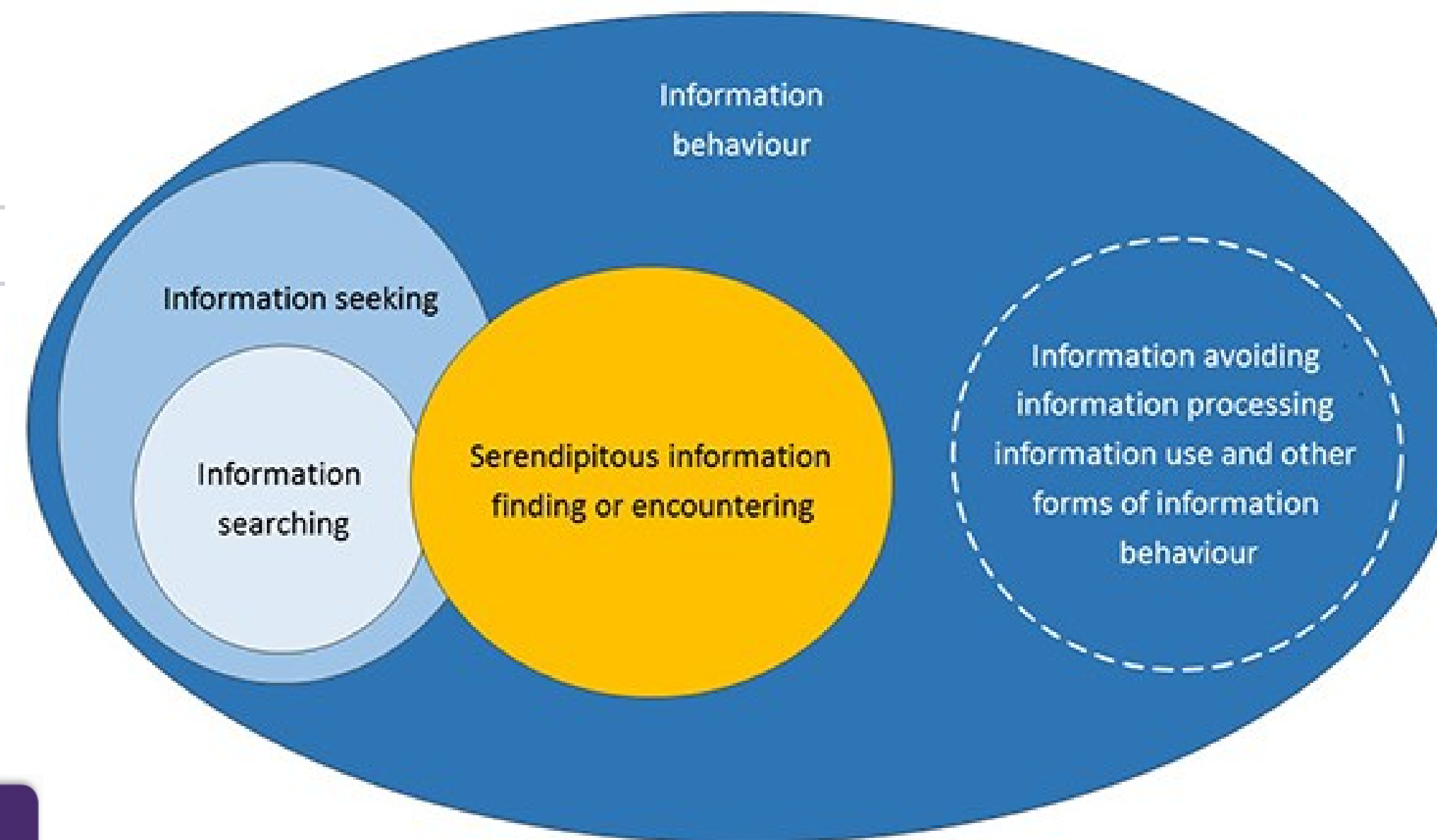
(inglês *serendipity*)

*substantivo feminino*

1. A faculdade ou o acto de descobrir coisas agradáveis por acaso.
2. Coisa descoberta por acaso.

Sinónimo Geral: SERENDIPISMO

Palavras relacionadas: [serendipismo](#), [serendipitoso](#).



JISC

## The Digital Information Seeker

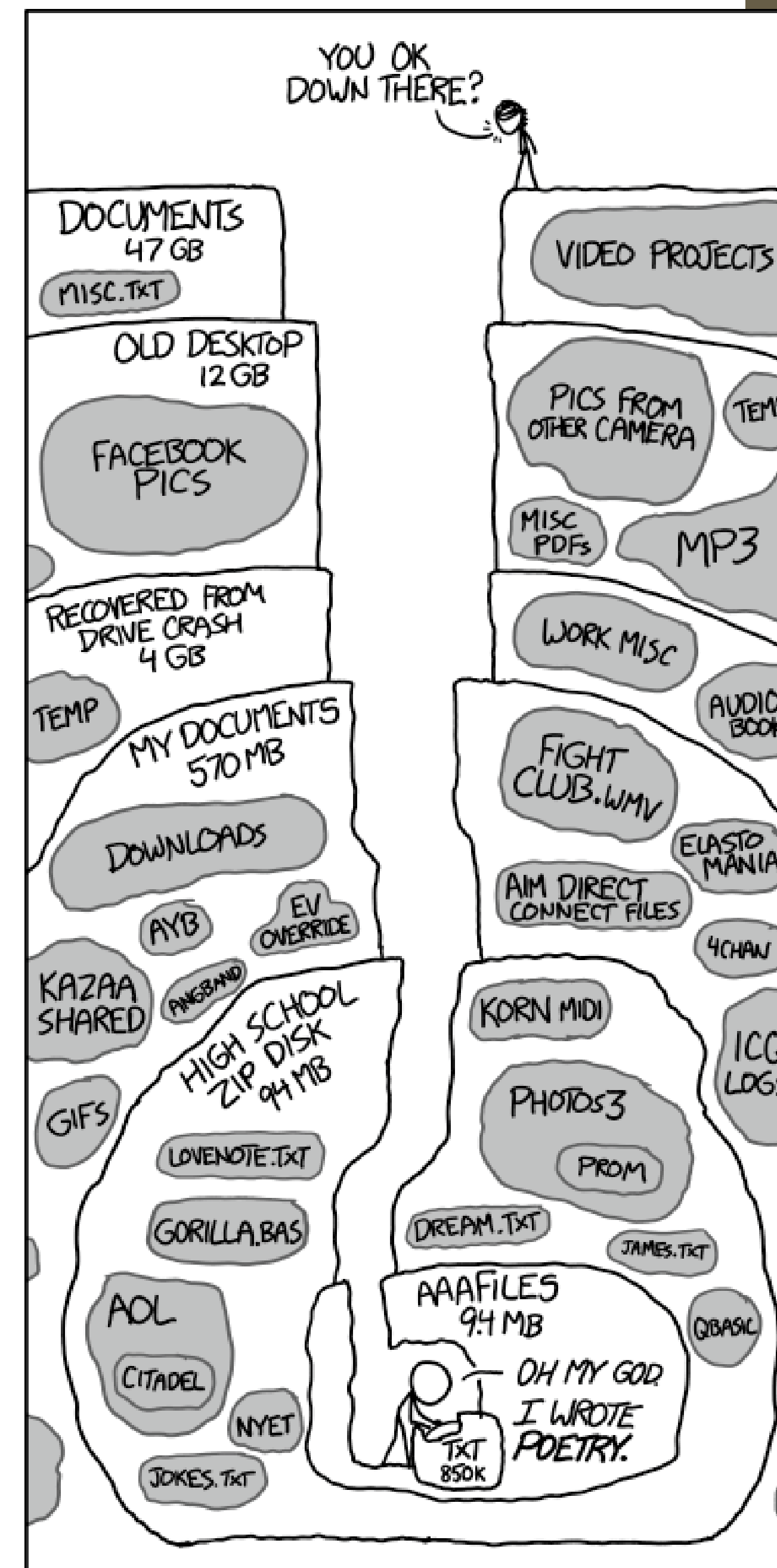
Report of findings from selected OCLC, RIN and JISC user behaviour projects

Produced for JISC by Lynn Silipigni Connaway, PhD and Timothy J Dickey, PhD, OCLC Research



# Arquivo de conteúdos Web: Obstáculos

- Preocupação excessiva com conteúdos digitalizados em detrimento aos conteúdos nato-digitais
- Alto índice de dependência tecnológica
- Rápida mudança de formatos e evolução das linguagens de marcação
- Caráter interativo da Web 2.0
- Defasagem das técnicas de coleta, armazenamento e preservação de conteúdos Web em detrimento à evolução do ambiente de publicação digital
- Estrutura rizomática da Web (ex: hiperlincagem)
- Copyright e outros aspectos legais



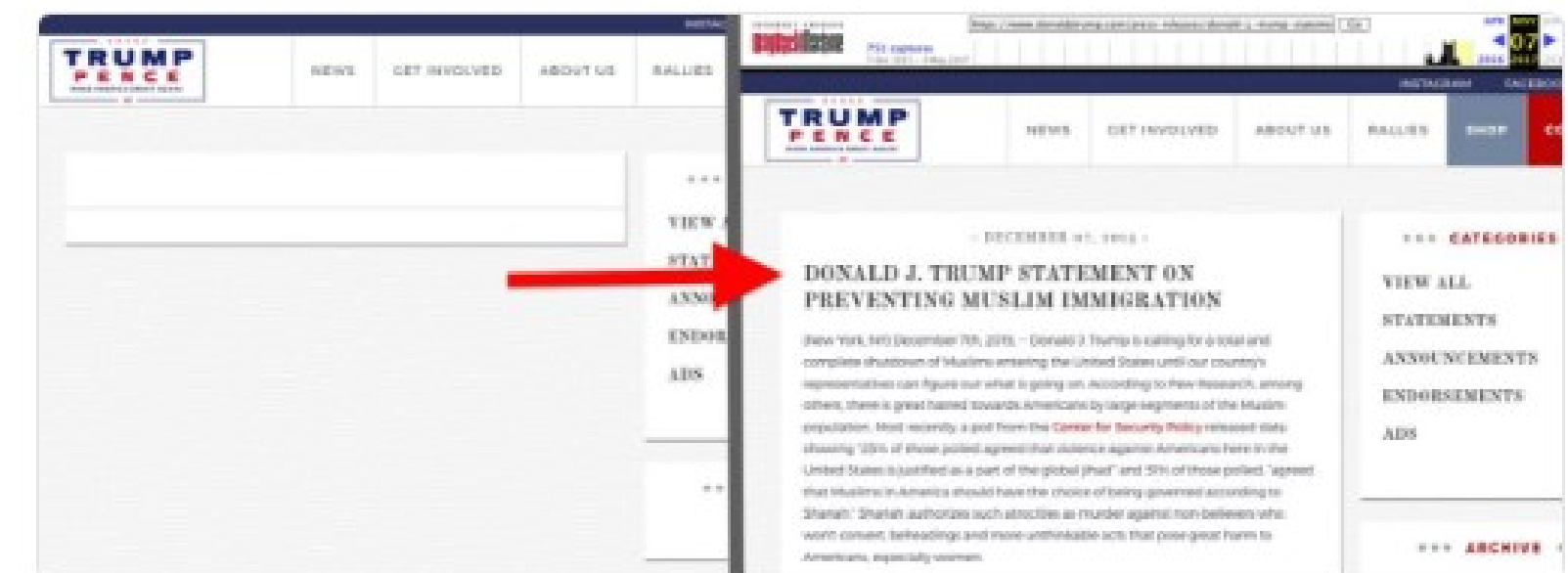
# Arquivos de publicações Web

- Contexto: 80% das páginas web são alteradas ou desaparecem anualmente (“Idade das Trevas” da informação digital);
- Vida útil média: 100 dias
- Importância patrimonial do arquivamento e preservação das publicações nos domínios nacionais;
- Aplicação e desenvolvimento de novas técnicas e ferramentas para recuperação e análise do conteúdo arquivado; Ex: N-gram, Trends, etc.
- Consórcio Internacional de Preservação da Internet (IIPC)  
<http://www.netpreserve.org/>
- Internet Archive: [Wayback machine](#)



Internet Archive @internetarchive · 58s

The deleted press release "Donald J. Trump Statement on Preventing Muslim Immigration" is available in the Wayback: [web.archive.org/web/2017050701...](http://web.archive.org/web/2017050701...)



# Perdendo a nossa memória...

- Segundo dados coletados do site de registros de páginas Web no Brasil, o país conta com mais de 4.5M de endereços criados desde 1996;
- Cada endereço Web brasileiro contém em média 85 páginas
- Lei 10.994/2004 (Art. 2º) Depósito Legal: “exigência estabelecida em lei para depositar, em instituições específicas, um ou mais exemplares, de todas as publicações, produzidas por qualquer meio ou processo, para distribuição gratuita ou venda”

## Domínios .br registrados até o momento

**4.652.281**

DOMÍNIOS .br

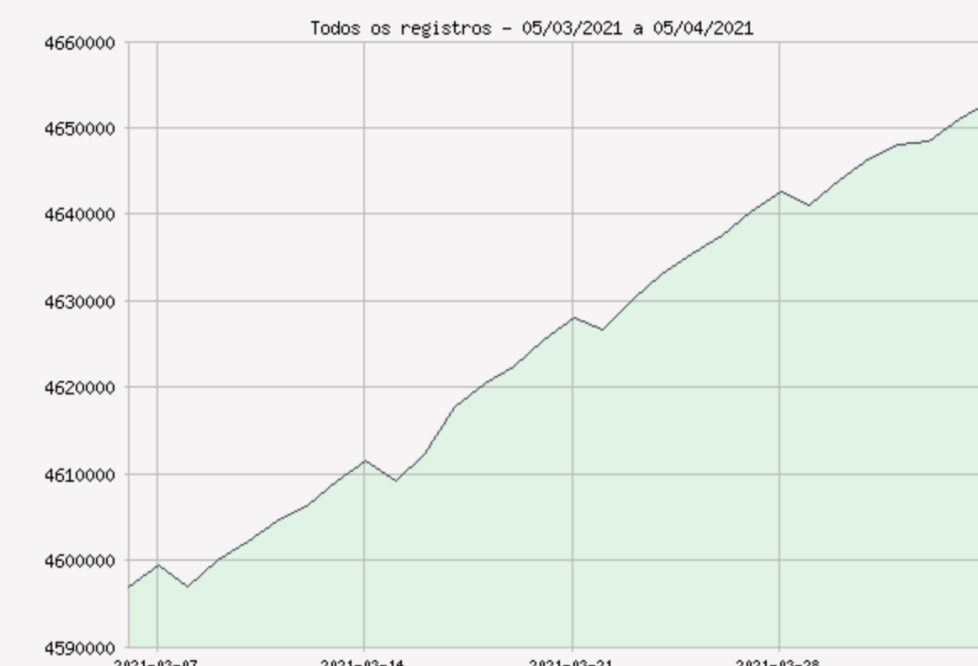
06/04/2021

IDNA **12.114**

» Ver evolução

DNSSEC **1.359.814**

» Ver evolução

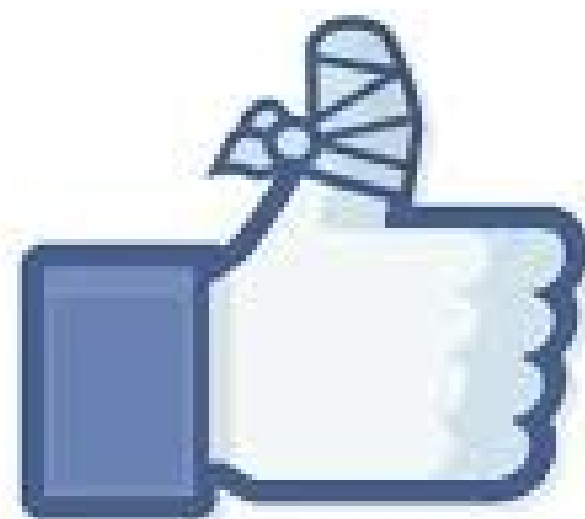


<https://registro.br/estatisticas.html>

# Panorama da história de conteúdos Web no Brasil...

Desculpe, essa página não está disponível.

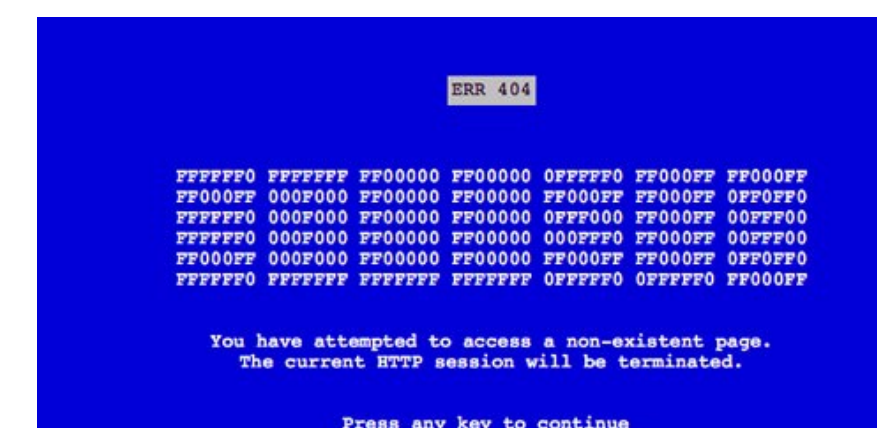
O link que você seguiu pode estar quebrado ou a página pode ter sido removida.



[Retornar à página anterior](#) / [Notar erros e páginas iniciais do Facebook](#) / [Ver a Central de ajuda](#)

## Error 404

Woops. Looks like this page doesn't exist.



Ooops, looks like a ghost!

The page you are looking for can't be found. Go home by [clicking here!](#)

Image: r1haa.com



404. That's an error.

The requested URL /asdasdasdasdasdasd was not found on this server. That's all we know.



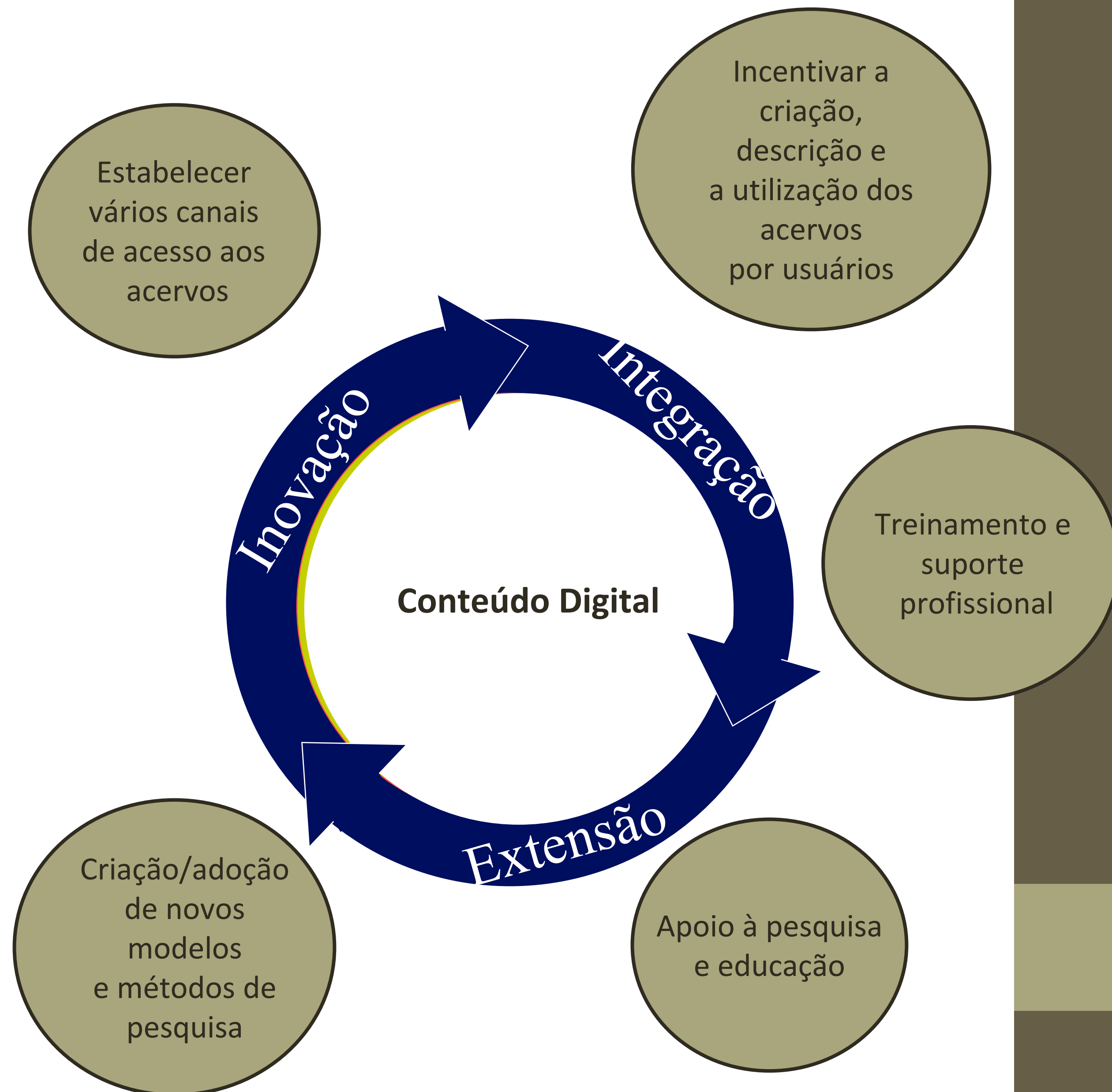
# Gerenciamento de Acervos Digitais

## Objetivos:

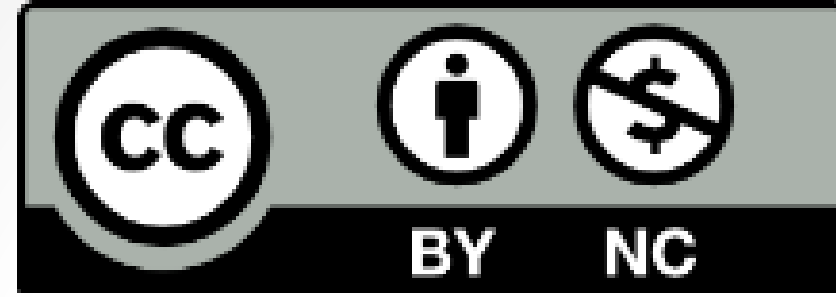
- Infraestrutura para depósito, preservação, acesso e reutilização de conteúdos digitais
- Integração de objetos e dados digitais provenientes de várias fontes
- Promoção e utilização de novas ferramentas aplicadas ao conteúdo digital: anotação, visualização, comparação, análise, etc.
- Aplicação de novos métodos e modelos em pesquisa digital
- Promover uma maior colaboração entre técnicos, pesquisadores, público em geral e instituições de memória

## Alcançados através de:

- Projetos colaborativos
- Plataformas interoperacionais
- Conexão entre base de dados







# Obrigado!

[alencarbrayner@gmail.com](mailto:alencarbrayner@gmail.com)