



O IMPACTO DA
CURADORIA
DIGITAL DE
**DADOS DE
PESQUISA** NA
COMUNICAÇÃO
CIENTÍFICA

LUANA FARIAS SALES
INSTITUTO DE ENGENHARIA NUCLEAR – IEN/CNEN
PPGCI DO IBICT/UFRJ
LSALES@IEN.GOV.BR

LUIS FERNANDO SAYÃO
COMISSÃO NACIONAL DE ENERGIA NUCLEAR – CNEN
LSAYAO@CNEN.GOV.BR

AGENDA

A IMPORTÂNCIA DOS DADOS DE PESQUISA
CURADORIA DIGITAL
CICLO VIDA DA CURADORIA DIGITAL
NOVOS DOCUMENTOS
IMPACTO NA COMUNICAÇÃO CIENTÍFICA
CONCLUSÕES

IMPORTÂNCIA DA PRESERVAÇÃO DOS DADOS DE PESQUISA

REGISTROS ÚNICOS IMPOSSÍVEIS DE SEREM REPRODUZIDOS

OU REPRODUZÍVEIS A UM CUSTO MUITO ALTO

No período de 1918 a 1919 a **gripe espanhola** se espalhou pelo mundo inteiro, matando de 20 a 80 milhões de pessoas. De origem viral, não havia tratamento conhecido. Como veio, se extinguiu. Só depois de muito tempo, foi encontrada uma **amostra de tecido humano infectado pelo vírus num hospital militar da Inglaterra**. A partir desses vestígios estão sendo desenvolvidas pesquisas para se descobrir vacinas e meios de tratamento da gripe espanhola. As pesquisas em torno da amostra só se tornaram possíveis graças à **preservação dos arquivos científicos**, datados de 1916, daquele hospital militar (DITADI, 2003).

**GRIFE
ESPANHOLA**



IMPORTÂNCIA DA PRESERVAÇÃO DOS DADOS DE PESQUISA

REGISTROS ÚNICOS IMPOSSÍVEIS DE SEREM REPRODUZIDOS

OU REPRODUZÍVEIS A UM CUSTO MUITO ALTO

E OS DADOS DE PESQUISA EM FORMATO DIGITAL ?



o arquivamento persistente, a preservação digital e o estabelecimento de modelos de informação para a preservação de registros científicos estão se tornando questões-chave para as áreas de pesquisa

DADOS DE PESQUISA SÃO MUITO SUSCETÍVEIS A PERDAS

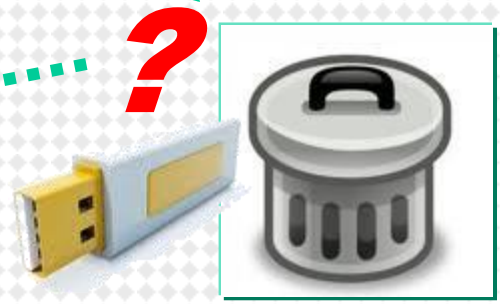
PESQUISA
TESE

publicação

dados
nascidos
digitais



Biblioteca convencional
Biblioteca de teses e
dissertações
Repositórios



REPOSITÓRIO DE DADOS

SÃO ESPAÇOS DE MEMÓRIA QUE FAZEM PARTE TAMBÉM DA HERANÇA DIGITAL

Testemunhos da atividade de pesquisa que se desenrola nas instituições acadêmicas precisam ser preservados

ERBS SOLAR MONITOR

Year	1992	1993	1994	1995	1996	1997	1998	1999	2000
Solar Irradiance (W/m²)	1360	1360	1360	1360	1360	1360	1360	1360	1360

Normalized Intensity (a.u.) vs Wavelength (nm)

Wavelength (nm)	Intensity
250	0.1
300	0.2
350	0.3
400	0.4
450	0.5
500	0.6
550	0.7
600	0.8
650	0.9
700	1.0

Wavelength (µm) vs GaN Well Width (nm)

GaN Well Width (nm)	Wavelength (µm)
1	0.1
2	0.2
3	0.3
4	0.4
5	0.5
6	0.6
7	0.7

Market Indices - Year 2000

Month	Tech Sector	Bucks	Oil
Jan	2000	2000	2000
Feb	2100	2050	2000
Mar	2200	2000	2000
Apr	2300	2050	2000
May	2400	2100	2000
Jun	2500	2150	2000
Jul	2600	2200	2000
Aug	2700	2250	2000
Sep	2800	2300	2000
Oct	2900	2350	2000
Nov	3000	2400	2000
Dec	3100	2450	2000

HÁ UMA PREOCUPAÇÃO EM TORNO DO PROBLEMA

Os dados de pesquisa são hoje quase sempre objetos informacionais nascidos digitais com toda a fragilidade intrínseca da informação digital

Muita informação importante já foi perdida

Há clamor geral no mundo acadêmico em torno desse problema

The screenshot shows the D-Lib Magazine website interface. At the top, there is a blue header with the logo and the text "D-Lib Magazine". Below this is an orange banner with the text "The Magazine of Digital Library Research". A dark blue navigation bar contains links for "HOME", "ABOUT D-LIB", "CURRENT ISSUE", "ARCHIVE", and "INDEXES". The main content area is white and features the title "D-Lib Magazine" followed by the issue information: "January/February 2011", "Volume 17, Number 1/2", and a link for "Table of Contents". Below this is a section titled "Access to Research Data" with a sub-heading "Introduction From the Guest Editors" and the name "Jan Brase" from the "German National Library of Science and Technology". A white box with a blue border contains the DataCite logo and the text "Helping you to find, access, and reuse data". At the bottom, a red and orange banner features the DataCite logo and the slogan "because good research needs good data".

ACESSO LIVRE AOS DADOS DE PESQUISA

A Declaração de Berlin sobre o Acesso Aberto ao Conhecimento em Ciências e Humanidades, publicada em 2003, amplia o escopo do que se entende por acesso livre ao definir que as...



Contribuições de acesso livre incluem resultados de pesquisas científicas originais, dados não processados e metadados, fontes originais, representações digitais de materiais pictóricos e gráficos e materiais acadêmicos multimídia

NOVOS PARADIGMA NA CIÊNCIA

eScience

Ciência orientada por dados

O quarto paradigma

- ciência **experimental**
- ciência **teórica**
- ciência baseada **simulação**
- ciência baseada no **uso intensivo** de dados



Como traduzir em significado e conhecimento a torrente de dados que caracteriza certos domínios da ciência contemporânea ?

- EXCESSO DE DADOS**
- DESENVOLVIMENTO DE INSTRUMENTOS**
- SOFTWARE/SIMULAÇÃO**

A NECESSIDADE DE TRANSMITIR CONHECIMENTO

É nesse ambiente que surge o conceito de **curadoria digital** de **dados científicos**, cujo principal desafio recai na necessidade de se preservar não somente o **conjunto de dados**, mas de preservar, sobretudo, a capacidade que ele possui de **transmitir conhecimento** para uso futuro das comunidades interessadas.

Isto significa que os ativos genuínos da pesquisa científica devem permitir que futuros usuários reanalise os dados dentro de novos contextos.



SIGNIFICADO & ESTRUTURA

ePRINTS



**SIGNIFICADO E
ESTRUTURA
EXPLÍCITOS**

DADOS DE PESQUISA

323 32 322 35 65 99 6565
6656 56 54 54 4 55 56 433
353 66 352 232 64 0 05030
32302 02 30 04 35 00 60
50 4 54 50 34 04 043 340
43 4 40 340 40 54 040

**PRESSÃO
TEMPERATURA
SEÇÃO DE CHOQUE
POPULAÇÃO**



SIGNIFICADO & ESTRUTURA

ePRINTS



**SIGNIFICADO E
ESTRUTURA
EXPLÍCITOS**

DADOS DE PESQUISA

323 32 322 35 65 99 6565
6656 56 54 54 4 55 56 433
353 66 352 232 64 0 05030
32302 02 30 04 35 00 60
50 4 54 50 34 04 043 340
43 4 40 340 40 54 040

**É PRECISO CONFERRIR
SIGNIFICADO AOS DADOS**

METADADOS

AFINAL, O QUE É CURADORIAL DIGITAL

Atividade que envolve a manutenção, a preservação e a agregação de valor a dados de pesquisa durante o seu ciclo de vida (DCC)



Todas as atividades envolvidas na **gestão de dados**, desde o **planejamento da sua criação** – quando os sistemas são projetados -, passando pelas boas práticas na digitação, na seleção dos formatos e na documentação, e na garantia dele estar disponível e adequado para ser **descoberto e reusado no futuro**. A curadoria digital também inclui a gestão de grandes conjuntos de dados para uso diário, assegurando, por exemplo, que eles possam ser pesquisados e continuem viáveis, ou seja, capazes de serem lidos e interpretados continuamente. (Daisy Abbott , 2008)

CURADORIA DIGITAL

**MANTER
PRESERVAR
ADICIONAR VALOR
REDUZIR AS AMEAÇAS
RISCO DE OBSOLESCÊNCIA
OTIMIZAR O COMPARTILHAMENTO
REDUZIR A DUPLICAÇÃO DE ESFORÇOS
TORNAR OS DADOS DISPONÍVEIS PARA FUTURAS PESQUISA**

CURADORIA DIGITAL

Repositórios digitais confiáveis

AFINAL, O QUE É CURADORIAL DIGITAL

InterPro: IPR006408 Calcium-translocating P-type ATPase, PMCA-type - Microsoft Internet Explorer

Arquivo Editar Exibir Favoritos Ferramentas Ajuda Links

Simple | Full HTML Version
Click here for help!

InterPro IPR006408 Calcium-translocating P-type ATPase, PMCA-type

Matches	Overview: sorted by AC , sorted by name , of known structure , proteins with splice variants Detailed: sorted by AC , sorted by name , of known structure , proteins with splice variants Table: For all matching proteins , of known structure Architectures
Accession	IPR006408 ATPase-IIB_Ca Matches: 190 proteins
Type	Family
Signatures	Database ID Name Proteins TIGRFAMs TIGR01517 ATPase-IIB_Ca 190
Parent Tree	IPR001757 ATPase, E1-E2 type
Contains	IPR004014 Cation transporting ATPase, N-terminal IPR005834 Haloacid dehalogenase-like hydrolase IPR006068 Cation transporting ATPase, C-terminal IPR008250 E1-E2 ATPase-associated region
Process	GO:0006816 calcium ion transport
Function	GO:0005388 calcium-transporting ATPase activity GO:0005509 calcium ion binding GO:0005524 ATP binding GO:0015085 calcium ion transporter activity
Component	GO:0016020 membrane
Abstract	This family describes the P-type ATPase responsible for translocating calcium ions across the plasma membrane of eukaryotes [1], out of the cell. In some organisms, this type of pump may also be found in vacuolar membranes [2]. In humans and mice, at least, there are multiple isoforms of the PMCA pump with overlapping but not redundant functions. Accordingly, there are no human diseases linked to PMCA defects, although alterations of PMCA function do elicit physiological effects [3]. The calcium P-type ATPases have been characterized as Type IIB based on a phylogenetic analysis which distinguishes this group from the Type IIA SERCA calcium pump [4].
Database links	Enzyme: 3.6.3.8

Taxonomic coverage

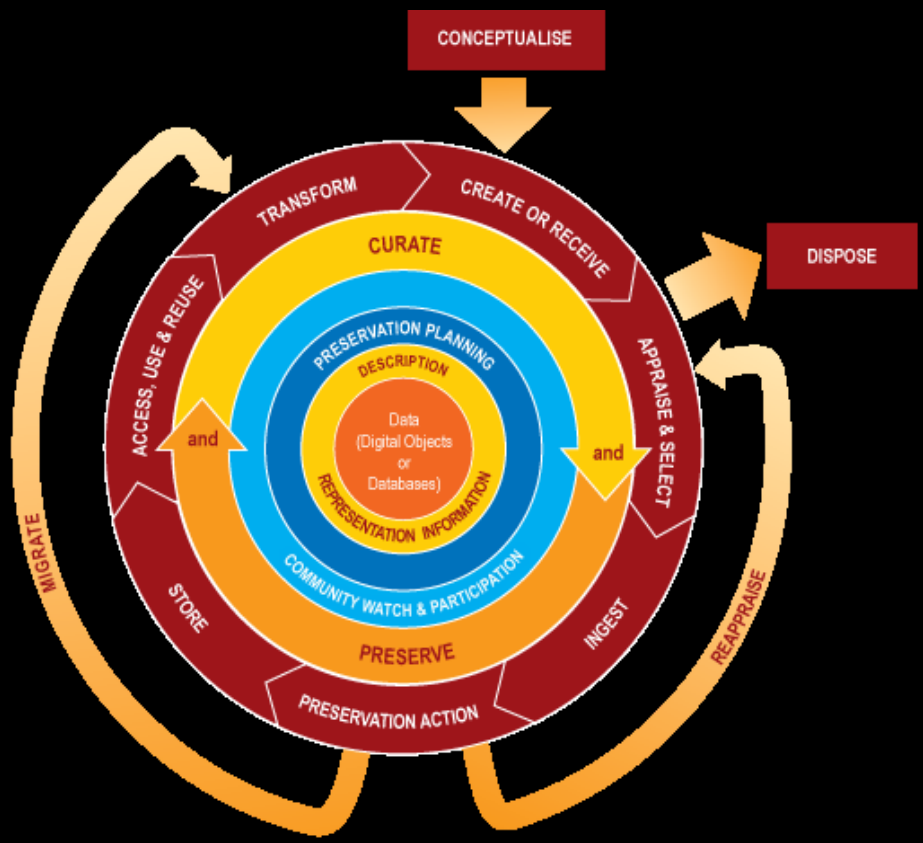
Intranet local

09:22

DCC – DIGITAL CURATION CENTRE - UK

 because good research needs good data

- Home
- Digital Curation
- About Us
- News
- Events
- Resources
- Training
- Projects
- Community
- Contact Us



MODELO DO CICLO DE VIDA DA CURADORIA DIGITAL

Estágios necessários para o sucesso da curadoria e da preservação

CICLO DE VIDA DA CURADORIA DIGITAL

Ações PERMANENTES

Descrição e a representação da informação - é efetivada pela atribuição de metadados administrativos, técnicos, estruturais e de representação de acordo com os padrões apropriados; visa assegurar a descrição adequada e o controle de longo prazo; compreende também a coleta e a atribuição de informações de representação necessárias para o entendimento do dado e para a sua apresentação (ou renderização).

Planejamento da Preservação - é necessária a definição de um plano de preservação cujo espectro englobe todo o ciclo de vida da curadoria do material digital, incluindo gestão, administração, políticas, e tecnologias.

Participação e monitoramento - enfatiza a necessidade de atenção para as atividades que se desenrolam no âmbito das comunidades envolvidas com o problema de curadoria

Curadoria e preservação - estar continuamente alerta e empreender as ações administrativas e gerenciais planejadas para a curadoria e preservação

CICLO DE VIDA DA CURADORIA DIGITAL

Ações SEQUENCIAIS

CONCEITUALIZAÇÃO
Conceber e planejar a criação de dados

CRIAÇÃO
Criar ou receber dados, incluindo metadados

AVALIAÇÃO E SELEÇÃO

CAPTURA
Transferência para um repositório, data center...

AÇÕES DE PRESERVAÇÃO
Assegura autenticidade, integridade e confiabilidade

ARMAZENAMENTO
Armazenamento seguro seguindo padrões

ACESSO, USO E REUSO

TRANSFORMAR
Criar novos dados a partir do original

CICLO DE VIDA DA CURADORIA DIGITAL

Ações OCASIONAIS

REAVLIAÇÃO

Avaliação falha, voltar para possível curadoria

DESCARTE

Liberar o sistema de dados não selecionados

MIGRAÇÃO

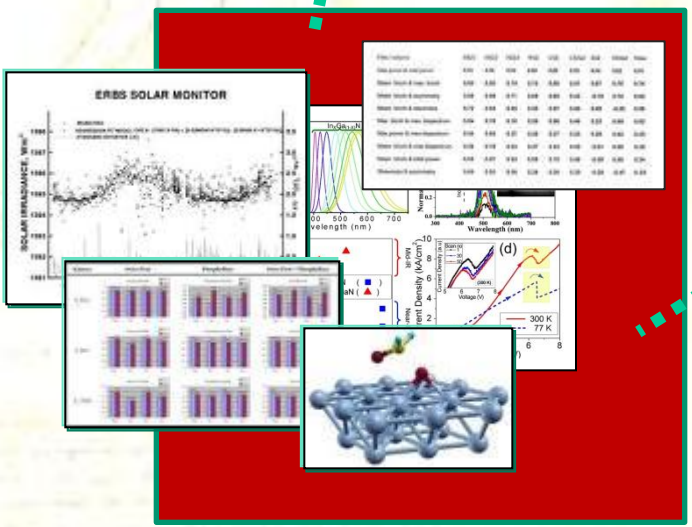
Contornar a obsolescência tecnológica

NOVOS MODELOS DE DOCUMENTOS

REPOSITÓRIO DE ePRINTS



REPOSITÓRIO DE DADOS ou DATA SETS



- DOCUMENTOS AMPLIADOS
- DOCUMENTOS COMPOSTOS
- DOCUMENTOS SOBREPOSTOS

OAI-ORE (OBJECT REUSE AND EXCHANGE)

ALTERAÇÃO DO CICLO DA COMUNICAÇÃO CIENTÍFICA

Acesso Limitado = Impacto Limitado

(Traduzido e adaptado de Harnad)



Figura1: Modelo da Comunicação Científica Tradicional (RODRIGUES, 2008)



ALTERAÇÃO DO CICLO DA COMUNICAÇÃO CIENTÍFICA – 1º encurtamento

Impacto e acesso à investigação maximizado pelo “auto-arquivo”

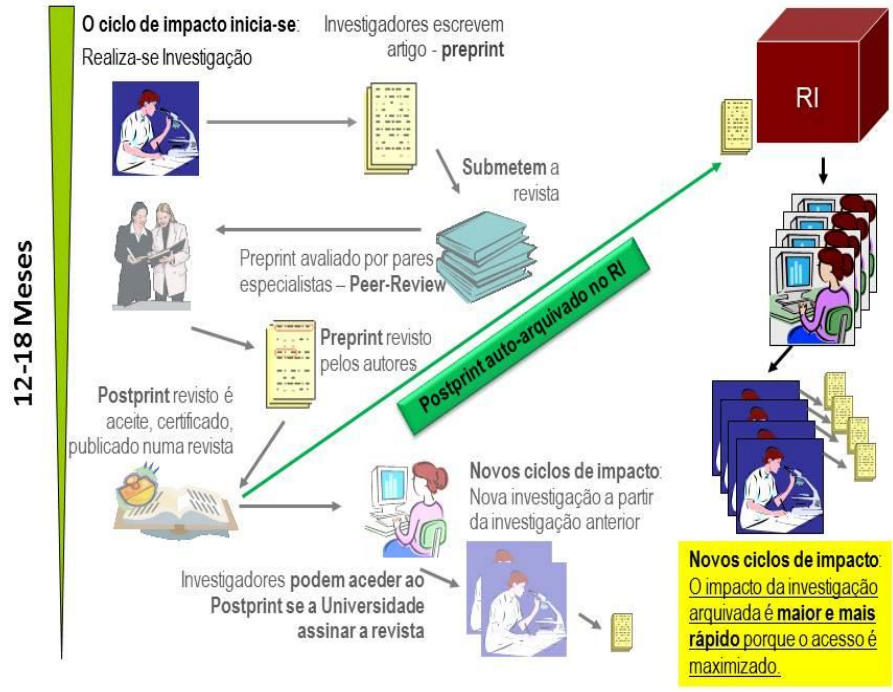
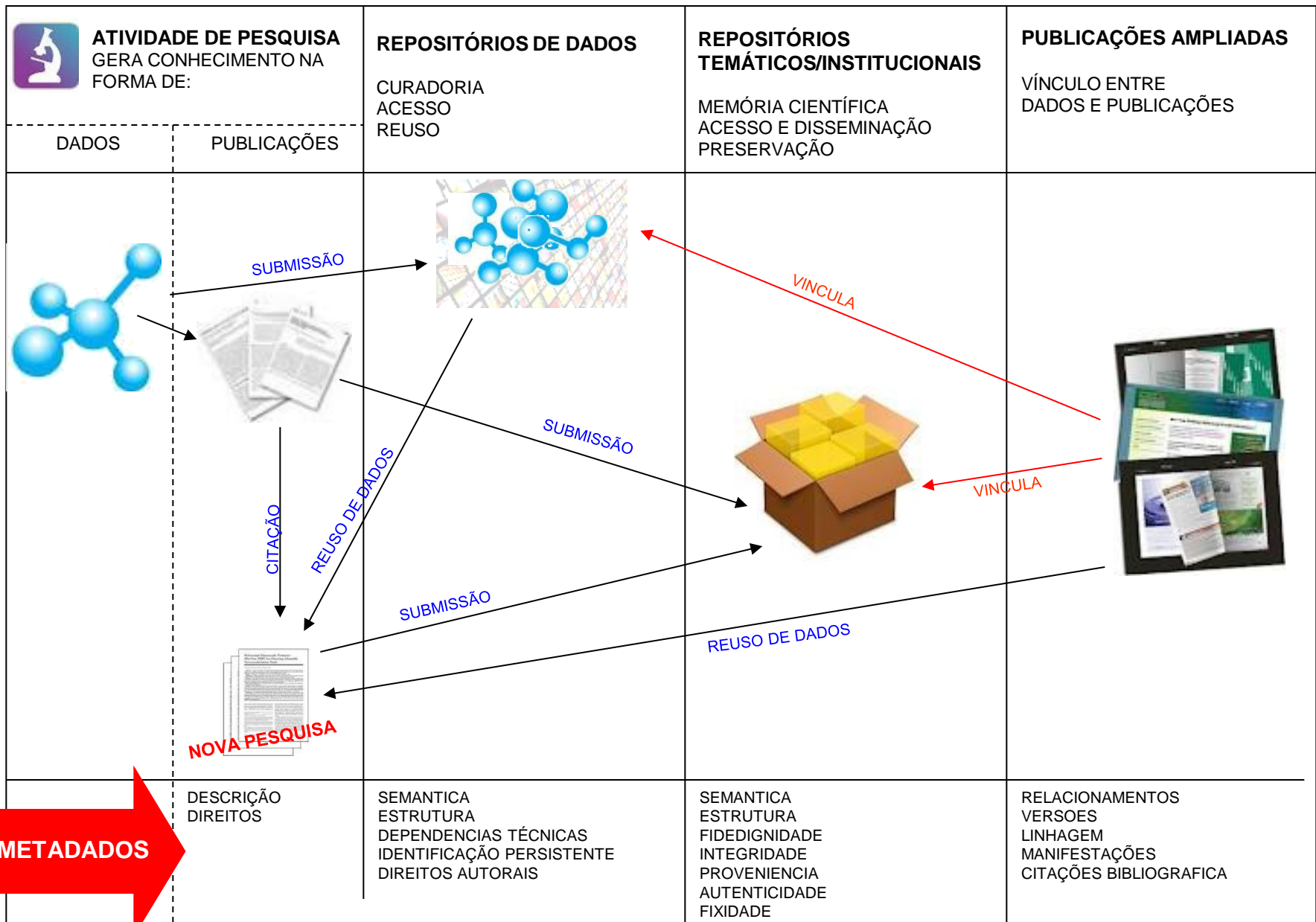


Figura 2: Modelo da Comunicação Científica tradicional alterado pelo Acesso Aberto (RODRIGUES, 2008)



COMUNICAÇÃO CIENTÍFICA



ALTERAÇÃO DO CICLO DA COMUNICAÇÃO CIENTÍFICA

A alteração do ciclo da comunicação científica não está apenas no âmbito do encurtamento do mesmo, mas também na esfera da qualidade da informação acessada, tendo em vista que, a partir da proposta da curadoria digital, pode-se ter acesso à dados científicos que antes não eram disseminados, impossibilitando a replicação dos procedimentos que levaram ao resultado de determinada pesquisa.



Muitas questões a serem discutidas

Quem serão os responsáveis pela curadoria desses dados?

Que tipos de pesquisas deverão conceder os dados?

Como adquirir esses dados?

Quem poderá ter acesso?

Qual o nível de qualidade e confiabilidade?

Ou qual o nível de transparência?





À GUISA DE
CONCLUSÃO

**NOVA LINHA DE ESTUDO
PARA A C.I.**

**PAPEL PARA AS
BIBLIOTECAS DE PESQUISA**

PUBLICAÇÕES AMPLIADAS

MODELOS DE INFORMAÇÃO