

Compartilhamento de dados de infraestrutura no suporte a pesquisa científica

Ibirisol Fontes Ferreira
Graduate School of Informatics,
Kyoto University,
Kyoto, Japan

WTR do PoP-BA/RNP 2025

29 Setembro – 01 Outubro 2025 // Salvador, Brasil

Sumário

- Motivação
- Ciência Aberta
- Desafios no compartilhamento
- Plano de Gestão de Dados
- Dados de infraestrutura e suas utilidades
- Exemplos
- Caso prático
- Compartilhamento dos dados
- Conclusão

Motivação

- O crescente uso de dados reais em pesquisas, impulsionados pela disponibilidade de **novas tecnologias**, que permitem **coletar**, **processar** e **analisar** grandes volumes de informações, direciona os requisitos de validação e aplicação dos resultados obtidos.
- Continuidade dos avanços exige **integração** ininterrupta entre os diversos **setores da sociedade** (administração pública, academia, indústria, etc.).
 - Democratização da pesquisa científica e seus resultados é colocada como um direcionamento global para cumprimento do direito humano de acesso à ciência.
 - Experiência durante a pandemia impulsiona ainda mais a necessidade de maior transparência.

Fontes: <https://www.unesco.org/pt/fieldoffice/brasilia/expertise/open-science-brazil>

Motivação

- Especialmente para a área de tecnologia da informação (TI), alguns **avanços** de pesquisa acadêmica são mais **tangíveis**.
 - Resultantes de estudos em software, algoritmos e infraestrutura.
- Novos protocolos e serviços proporcionam que os avanços científicos sejam aplicados no dia a dia da operação de infraestruturas tecnológicas como **data centers**, **redes ópticas**, **nuvens computacionais**, entre outros.
 - Performance dos serviços (Observabilidade, tuning, etc.)
 - Eficiência operacional
 - Consumo de recursos (pessoal, energético, etc.)
- Embora seja um setor precursor em práticas de transparência e democratização (ex.: software livre), pode se beneficiar ainda mais com maior adoção das práticas de **ciência mais transparentes** e **integradas** aos atores para **impulsioneamento** de **inovações** em **produtos** e **serviços** coerentes com cenários específicos.
 - Maior transparência, praticabilidade, colaboração e acessibilidade à pesquisa científica.

Ciência Aberta

- Ciência Aberta (Open Science) é um termo abrangente que descreve práticas de pesquisa que se tornaram mais **abertas**, **acessíveis** e **transparentes** na era digital.
- Combinação do aspectos ideológico e prático:
 - Crença de que a ciência deve ser inerentemente aberta.
 - Garantia de responsabilização e transparência na pesquisa.
- Principais objetivos incluem:
 - **Democratizar** o acesso ao conhecimento científico.
 - Aumentar a **colaboração** entre disciplinas (campos do conhecimento) e fronteiras (geográficas e culturais).
 - Acelerar a descoberta e a inovação nos processos de **criação**, **avaliação** e **comunicação** do conhecimento científico.

Fontes: <https://www.unesco.org/en/open-science/about>
<https://rcos.nii.ac.jp/en/document/openscience/>

Ciência Aberta – Brasil

- Discussões em andamento para consolidação da ciência aberta no país.
 - Governança voltada a ter uma **ciência** mais **transparente**, **reprodutível**, **colaborativa** e **acessível** ao conhecimento científico.
 - Alinhamento com iniciativas globais e com o diálogo com a sociedade (diversos setores).
 - Ponto de atenção consolidados em outros lugares:
 - Relação entre o **acesso aberto** a **artigos acadêmicos** financiados por bolsas públicas e a **gestão/utilização de dados de pesquisa**.

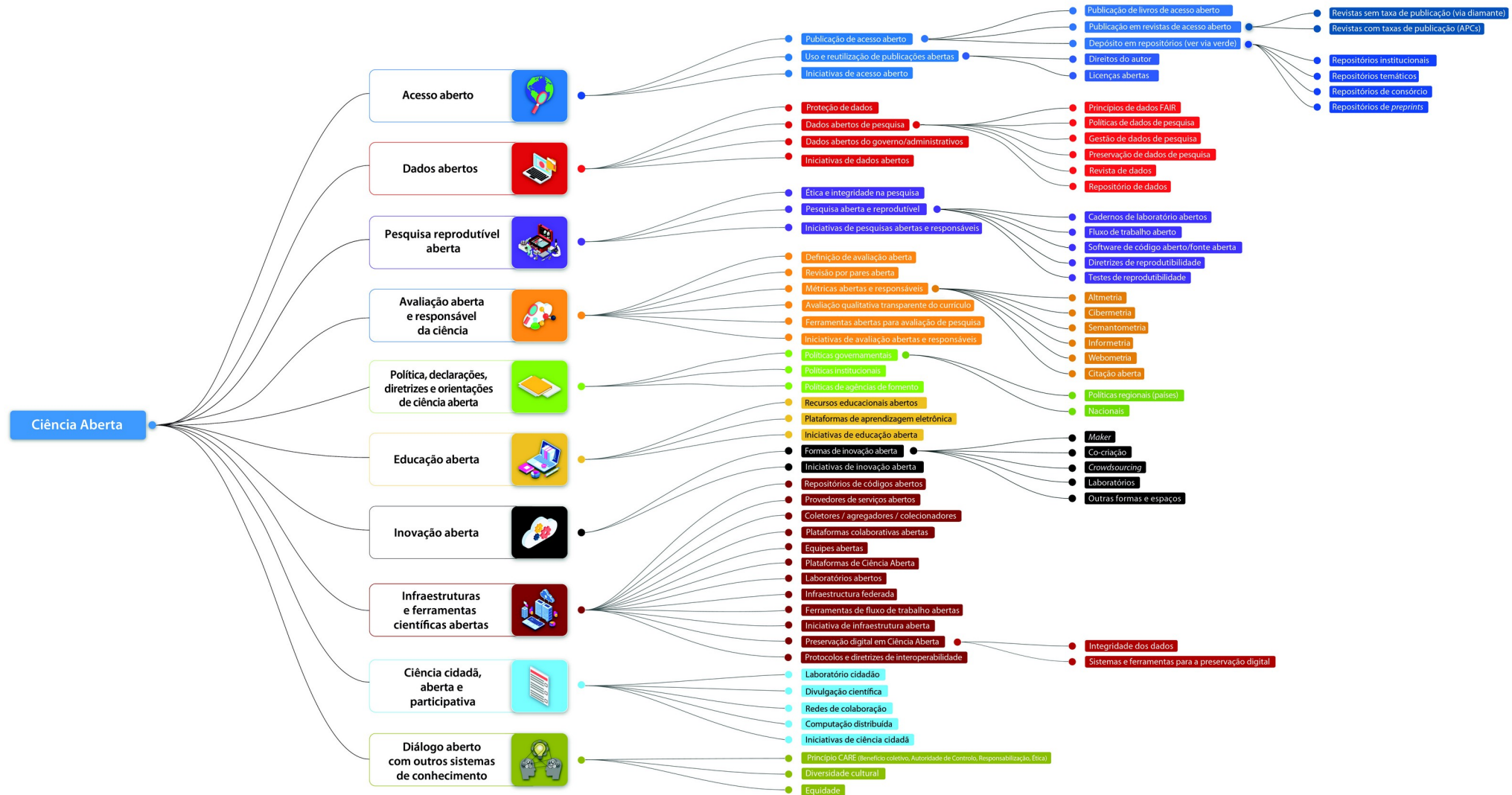
Fontes: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/noticias/2025/06/mcti-discute-diretrizes-para-consolidar-a-ciencia-aberta-no-brasil>

<https://www.gov.br/cgu/pt-br/governo-aberto/artigos/ciencia-aberta-uma-nova-forma-de-fazer-ciencia-mais-colaborativo-transparente-e-sustentavel/ciencia-aberta-um-novo-modo-de-fazer-ciencia-mais-colaborativo-transparente-e-sustentavel>

<https://www.kulib.kyoto-u.ac.jp/openaccess/1404342?lang=en>

<https://www.reprodutibilidade.org/en/post/brazilian-reproducibility-network-develops-recommendations-for-incorporating-open-and-reproducible-r>

Ciência Aberta – Abrangência



Inclusão de termos na taxonomia da Ciência Aberta (Silveira et al, 2023)

Ciência Aberta – Exemplos

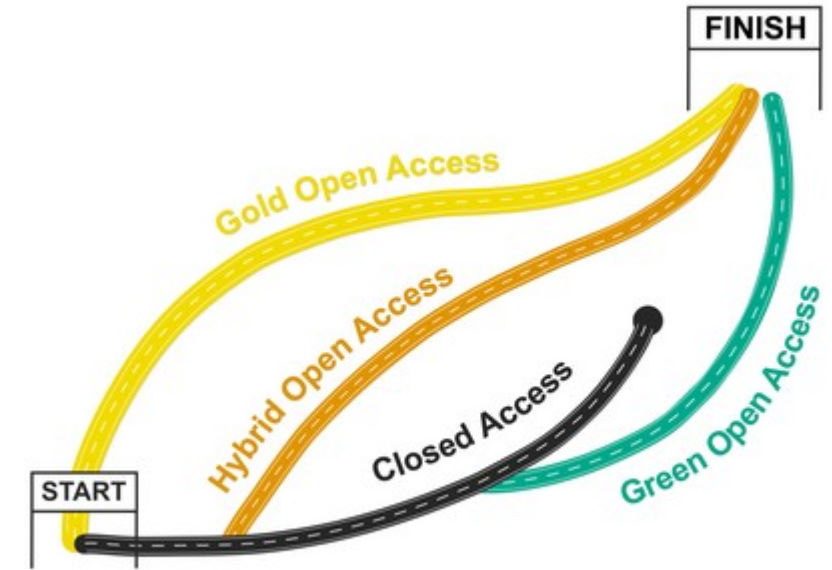
- Exemplos de práticas de Ciência Aberta incluem:
 - Acesso Aberto – AA (Open Access – OA)
 - **Acesso** online **gratuito** e **irrestrito** a publicações acadêmicas
 - Materiais Aberto (Open Materials)
 - **Disponibilização** dos **componentes** da metodologia de pesquisa necessários **para reproduzir** o procedimento e a análise relatados.
 - Dados Abertos (Open Data)
 - **Compartilhamento** de **dados de pesquisa** com padrões claros para permitir a reutilização

Fontes: <https://help.osf.io/article/452-open-practice-badges>
<https://www.abcd.usp.br/noticias/open-science-os-caminhos-da-producao-cientifica-aberta-no-brasil/>

Ciência Aberta – Acesso Aberto

- Acesso Aberto nos veículos de publicação
 - **Ouro**: A **publicação** final é **disponibilizada abertamente** pelo editor após a publicação. Financiada por taxas pagas pelos autores (**APCs**) para cobrir custos editoriais.
 - **Verde**: Autores depositam a versão final do artigo em um **repositório** de **acesso aberto**, mantendo os direitos autorais.
 - **Híbrido**: Suporte ao OA **mediante taxas** ou **assinaturas**, assumidas ou financiadas.
 - **Diamante**: Similar ao Ouro, mas com os **custos** na **disponibilização garantidos** por governos, fundos e agências, etc.

Pathways to Open Access



Source: open-access.network (2021), Pathways of Open Access (CC BY 4.0 International)

Fontes: <https://openeconomics.zbw.eu/en/knowledgebase/the-meaning-of-the-different-types-and-colours-of-open-access/>
<https://www.acessoaberto.usp.br/acesso-aberto-um-guia-para-pesquisadores>
<https://www.sibi.ufscar.br/espacodopesquisador/divulgacaodapesquisa/o-que-e-o-acesso-aberto>

Ciência Aberta – Materiais Aberto

- **Disponibilização** pública de todos os **materiais** e **métodos** utilizados em uma pesquisa científica.
 - **Permite** a **validação** e **replicação** da pesquisa por outros(as) pesquisadores(as) e outros grupos.
 - Precisa de **licenciamento** adequado
 - O **armazenamento** dos materiais deve ocorrer em **plataformas** de **acesso aberto** (ex. Open Science Framework (OSF)). Depósito em repositórios, garantem que os materiais tenham um registro e formato permanente.
 - É crucial **documentar** os materiais de forma **clara** e **completa**, incluindo informações sobre o contexto da pesquisa.
 - Pode envolver componentes de softwares externos, precisando especificação das versões específicas.

Fontes: <https://riseopenjournal.org/article/id/2623/>

Ciência Aberta – Dados Abertos

- Referem-se a dados de pesquisa **disponibilizados gratuitamente** para qualquer pessoa **acessar**, **usar** e **redistribuir**.
 - Dados quantitativos, qualitativos, experimentais, observacionais, etc.
- Também promove a reprodutibilidade, a transparência e a eficiência na pesquisa
- **Permite** que outros possam **validar** descobertas, **reutilizar** conjuntos de dados para novas análises ou **combiná-los** com outros dados.
- Incentiva a colaboração entre disciplinas.
 - Nos casos de **dados sensíveis** aplica-se o princípio de torná-los "o mais abertos possível e o mais fechados o necessário".
 - Pode envolver o uso de repositórios específicos que garantem a **proteção da privacidade**.
- Dados Abertos para a Ciência Aberta atendem aos **princípios FAIR**.

Fontes: <https://openeconomics.zbw.eu/en/knowledgebase/the-meaning-of-the-different-types-and-colours-of-open-access/>
<https://www.acessoaberto.usp.br/acesso-aberto-um-guia-para-pesquisadores>
<https://www.sibi.ufscar.br/espacodopesquisador/divulgacaodapesquisa/o-que-e-o-acesso-aberto>

Ciência Aberta – Dados Abertos (cont.)

- **Compartilhamento** de dados que envolve a **consonância** com os princípios **FAIR** (Findable, Accessible, Interoperable, and Reusable), que definem princípios para compartilhar dados e metadados.

- **Encontrável**

- Meta(dados) possuem um **identificador único** e **persistente** e **meta(dados) detalhados** que incluem esse identificador.
- Eles são registrados ou indexados em um recurso pesquisável.

- **Acessível**

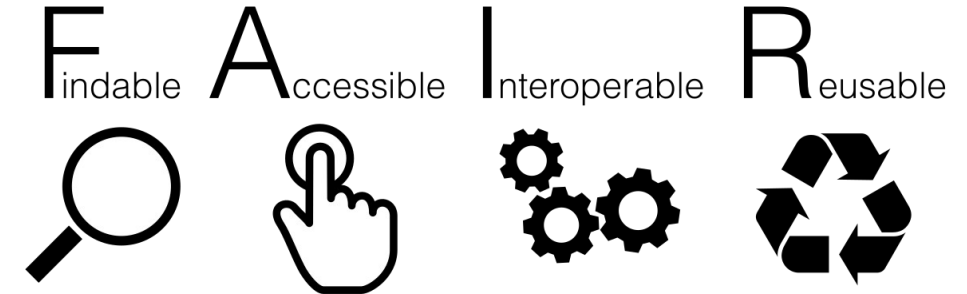
- (Meta)dados são **recuperáveis** pelo seu identificador usando um protocolo de comunicação padronizado e aberto. Com suporte ao controle de acesso.

- **Interoperável**

- Os dados **permitem** a **integração** a **outros dados**. Precisam interoperar com aplicativos ou fluxos de trabalho para análise, armazenamento e processamento.

- **Reutilizável**

- Meta(dados) devem ser **bem descritos** e **licenciados** para que possam ser replicados/combinados em diferentes cenários.



FAIR guiding principles for data resources (User:SangyaPundir, CC BY-SA 4.0, via Wikimedia Commons)

Fontes: <https://openscience.eu/article/infrastructure/guide-fair-principles>
<http://go-fair-brasil.ibict.br/>
<https://www.go-fair.org/fair-principles/>

Ciência Aberta – Processo de pesquisa

- Relação entre o processo de pesquisa e práticas de Ciência Aberta

	Processo de Pesquisa	Práticas de Ciência Aberta
Planejamento	Formulação de questões de pesquisa, definição de metodologia	Pré-registro de estudos, planos de gerenciamento de dados (DMP - data management plans)
Coleta/Aquisição de Dados	Coleta de dados primários ou secundários para o estudo	Dados Abertos, Crowdsourcing
Processamento e Análise de Dados	Limpeza, organização e análise de dados (por exemplo, codificação, análise estatística)	Dados Abertos, Código Aberto, Crowdsourcing
Publicação e Disseminação	Submissão a periódicos, conferências e repositórios	Revisão Aberta, Acesso Aberto
Avaliação e Impacto	Avaliando alcance, citações e influência social mais ampla	

Adaptação de "Research Process & Open Science Practices" (Chifumi Nishioka, 2025)

Plano de Gestão de Dados

- O PGD **descreve** como os dados de pesquisa serão **coletados, organizados, armazenados, compartilhados** e **preservados**.
- Exigido por muitos financiadores e instituições de pesquisa.
- Garantem que os dados estejam bem organizados e acessíveis.
- Aumentam a transparência e a reprodutibilidade da pesquisa.
- Elementos típicos de um PGD
 - **Tipos** de dados a serem **coletados** (por exemplo, textos, entrevistas, imagens)
 - **Formatos** de **arquivo** e documentação
 - Estratégia de **armazenamento** e **backup**
 - **Condições** de **compartilhamento** e **acesso** aos dados
 - **Preservação** a longo prazo

Dados de infraestrutura e suas utilidades

- Na operação de infraestruturas a **telemetria** e **observabilidade** dos componentes ativos e passivos fazem parte da rotina.
 - **Monitoramento** da utilização e performance são os mais comuns para **visualização da saúde** dos recursos de redes e sistemas.
 - Visualização histórica é importante para planejamento e tomada de decisão organizacional.
- **Exemplos** de **dados** coletados da infraestrutura:
 - Fluxos de rede (volume de pacotes, bytes, tipo de protocolo, origens/destinos e portas)
 - Status dos protocolos de roteamento
 - Status e desempenho de serviços e aplicações
 - Status e desempenho de máquinas físicas e virtuais
 - Acesso aos serviços e sistemas (registro de logs também ao nível de aplicação)
- **Exemplos** de **uso** desses dados:
 - Engenharia de tráfego para otimizar a infraestrutura com múltiplos objetivos (QoS, controle de congestionamento, etc.).
 - Análise de segurança para controle do uso correto dos recursos computacionais (SIEM, IDS, IPS, etc.).
 - Projeto de novas arquiteturas de serviços e protocolos.

Exemplos – Engenharia de trafego

- Temos diversos **modelos** de **roteamento**, que, baseados em informações da **demanda** de **tráfego**, podem ser aplicados para otimizar o roteamento na rede com métricas objetivas.
- Modelo de demanda de tráfego
 - **Pipe**: Quando a matriz de tráfego é completamente descrita entre os pontos da rede.
 - **Hose**: Quando a matriz de tráfego é sumarizada pelo total de tráfego por nó (demanda agregada).
 - **HSDT**: Quando a matriz de tráfego é estimada pela demanda agregada com base no tráfego dos links e dados anteriores.
 - **HLT**: Quando a matriz de tráfego é estimada pela demanda agregada e pelos limites máximos dos links.
- O **uso** recorrente de **modelagem sintética** dos padrões de tráfego e topologias.
 - Grafos de referência ou aleatórios (ex: Waxman) e tráfego seguindo modelos estatísticos (ex: gravity model)
- **Dados** que podem ser **compartilhados** abrangem:
 - Informações de **topologia** e protocolos usados.
 - Topologias disponibilizadas publicamente para tipos de redes específicos (SNDlib, CAIDA, Internet Topology Zoo, etc.).
 - **Matriz de tráfego** (ex.: SNDlib)
 - Entre os componentes da rede.
 - Estatísticas dos enlaces conectando os dispositivos de rede.

Exemplos – Serviços de rede

- Performance de serviços de rede pode envolver **diversos componentes** que impactam diretamente na **performance** do **usuário** (WWW, DNS, CDN, etc.).
 - Modelagem arquitetural do software para **cargas de trabalho específicas**.
- Comum o uso de alguns poucos **conjunto de dados públicos** e/ou **tráfego sintético** modelando padrão de requisição dos usuários.
- **Dados** para **compartilhamento** na visão do mantenedor da infraestrutura abrangem:
 - Tráfego de usuários até o servidor local (Captura de tráfego).
 - Registro transacional ao nível do protocolo ou da aplicação (Logging).

Caso prático – DNS Recursivo

- É possível usar alguns conjuntos de **dados já publicizados**, porém alguns **não** são **padronizados** e não **permitem extração** de algumas **informações**.
 - **Parceria** com organizações que provêm infraestruturas de **serviços reais** como a com o PoP-BA/RNP é um caminho **essencial** para maior diversidade de conjunto de dados.
- Como providenciar os dados úteis?
 - **Captura de tráfego** não escala e requer um grande esforço no tratamento da captura para utilização (correlacionamento e montagem de transações de protocolo).
 - Uso de registro de **log da aplicação** (depende da ferramenta usada).
 - Interfaces padronizadas para **exportação** de dados de **transações** DNS (ex. DNSTap).
- É necessária a **configuração** do serviço e **manutenção** da coleta nos **processos** internos.
 - Cooperação com as equipes técnicas.
- A utilização exige **extração** e **formatação** para uso e reprodução posterior dos **dados**.
 - Precauções de **segurança** (Dados possíveis de compartilhamento)
 - **Processamento** dos **dados** de transação do serviço precisa ser definido, ex:
 - Como lidar com múltiplas solicitações em uma requisição.
 - Considerar a comunicação cliente-serviço, serviço-autoritativo, ambas.

Caso prático – DNS Recursivo

```
---
type: MESSAGE
identity: "81c9a3aa77a7"
version: "unbound 1.17.1"
message:
  type: CLIENT_RESPONSE
  response_time: !!timestamp 2024-01-11 10:53:22.860767
  socket_family: INET
  socket_protocol: UDP
  query_address: 172.17.0.1
  response_address: 0.0.0.0
  query_port: 41033
  response_port: 53
  response_message: |
    ;; opcode: QUERY, status: NOERROR, id: 44853
    ;; flags: qr rd ra; QUERY: 1, ANSWER: 6, AUTHORITY: 0, ADDITIONAL: 1

    ;; OPT PSEUDOSECTION:
    ; EDNS: version 0; flags: ; udp: 1232

    ;; QUESTION SECTION:
    ;www.cnn.co.jp.      IN      A

    ;; ANSWER SECTION:
    www.cnn.co.jp.      7199    IN      CNAME    e57395d3fa5ee6dd1caf4489bc955d4e.cdnext.stream.ne.jp.
    e57395d3fa5ee6dd1caf4489bc955d4e.cdnext.stream.ne.jp.  300     IN      CNAME    cdnext-svg001-ipb001.stream.ne.jp.
    cdnext-svg001-ipb001.stream.ne.jp.  30      IN      A        202.79.241.200
    cdnext-svg001-ipb001.stream.ne.jp.  30      IN      A        202.79.240.201
    cdnext-svg001-ipb001.stream.ne.jp.  30      IN      A        101.102.235.200
    cdnext-svg001-ipb001.stream.ne.jp.  30      IN      A        202.247.51.200
  ---
```

Exemplo de saída de texto de uma exportação transacional no DNS recursivo

Caso prático – DNS Recursivo

- Campos do protocolo e **anonimização**
 - Timestamp (segundos)
 - Client IP address do usuário (client)
 - Consegue **identificar** a **origem** que é o usuário?
 - Server IP address do serviço (server)
 - Consegue **identificar** o **destino** do tráfego DNS do usuário?
 - Query type (A, NS, TXT, etc.)
 - Query string (domain)
 - Pode mostrar **dados do usuário**, como domínios com **informações exclusivas** do usuário, é necessário?
 - RR type of first answer (A, SOA, CNAME, etc.)
 - Answer string
 - Pode mostrar o **destino** do **tráfego** do usuário. É necessário?
 - TTL
 - Campo mais ligado a gerência do registro

Caso prático – DNS Recursivo

- Existem muitas **opções** de **anonimização** consideradas **aceitáveis**.
 - **Campo higienizado** usando função hash (como SHA-1)
 - (campo_original + segredo_do_campo_dedicado)
 - Valor do **subcampo higienizado** criando um hash em alguma parte do campo.
 - domain_hash1.domain_hash2.domain_hash3.net
 - domain_hash1.domain_hash2.domain_hash3.jp
 - **Função** de **tradução** aleatória.
 - Tradução aleatória de um IP original em um IP privado (RFC 1918).
 - Tradução para um identificador único aleatório.

Caso prático – DNS Recursivo

- Exemplo de dados anonimizados (Allman, 2021)

```
604706.316754 c:32178d3bb55d6d33a89b6d3b87c12e5977a44340 r:42f234f0dbeb02bef102b752d24624d3efc62fcf A  
n:fcd1f6cc70809e67165baafce33615733b57f4dd.c1f4a388bfa5af9936e33fc2e619584a0e9521b3.0c02b49c3500e363ffb0  
06d76f0ed1ab9a81514b.org A i:c65e644481a09ec6a2c10c524bafb553202b09ba 10 A  
i:bc8c52505ccce2243224a3c16dc7034b6ea1939b 10 A i:7e6c5e10c6e0754c2d50e3e98301158da76136fa 10 A  
i:370b90c65e4a695c682480c35e240dc64e98ecb4 10
```

```
604706.316754 c:32178d3bb55d6d33a89b6d3b87c12e5977a44340 r:42f234f0dbeb02bef102b752d24624d3efc62fcf A  
n:fcd1f6cc70809e67165baafce33615733b57f4dd.c1f4a388bfa5af9936e33fc2e619584a0e9521b3.0c02b49c3500e363ffb0  
06d76f0ed1ab9a81514b.org A i:c65e644481a09ec6a2c10c524bafb553202b09ba 10 A  
i:bc8c52505ccce2243224a3c16dc7034b6ea1939b 10 A i:7e6c5e10c6e0754c2d50e3e98301158da76136fa 10 A  
i:370b90c65e4a695c682480c35e240dc64e98ecb4 10
```

```
604706.677000 c:5ba638a1731f0d9e86f375777ccb9f498d41a037 r:42f234f0dbeb02bef102b752d24624d3efc62fcf A  
n:06f1385979232111dd4aeda1e4bf8c29eabe5970.eb6c27514cb41a77d0e16a9e0b779890ba5153cd.com CNAME  
n:717f0ec9bab8f8d87d5b73b7d753f8562662401f.e80e7afb6af2560a7eaf189570e93469bd5f419e.23f689eeba26ef06362b  
1dfcf69ad65f239ccc84.3ff51baf92d014241338573e1cfb525e7cc971ad.com 136 A  
i:b20ae140923f9ab50bbc0c72fe8f3ee296ab1d79 26 A i:6fda8829cd44159cbf6f10f17af9119331a4b04e 26 A  
i:1de5ded2f1ff9bf978785a43722883f5467b5cb2 26 A i:941b6260a36cd9f004e0bc6ed5cc069b7e3debf0 26 A  
i:a8726f70910c5f518b72c3f6a187d48268fd7f08 26 A i:85b2cce103c54529b3c2a32eab7c6293b438acfd 26 A  
i:92d0d6f8e22505950f8b97882f13d35ea837b005 26 A i:e9064e92962c768fb632db6d359ecd81ff4d392e 26
```

Caso práctico – DNS Recursivo

- Exemplo de dados anonimizados (Allman, 2021)

```
604706.316754 c:32178d3bb55d6d33a89b6d3b87c12e5977a44340 r:42f234f0dbeb02bef102b752d24624d3efc62fcf A
n:fcd1f6cc70809e67165baafce33615733b57f4dd.c1f4a388bfa5af9936e33fc2e619584a0e9521b3.0c02b49c3500e363ffb0
06d76f0ed1ab9a81514b.org A i:c65e644481a09ec6a2c10c524bafb553202b09ba 10 A
i:bc8c52505ccce2243224a3c16dc7034b6ea1939b 10 A i:7e6c5e10c6e0754c2d50e3e98301158da76136fa 10 A
i:370b90c65e4a695c682480c35e240dc64e98ecb4 10
```

```
604706.316754 c:32178d3bb55d6d33a89b6d3b87c12e5977a44340 r:42f234f0dbeb02bef102b752d24624d3efc62fcf A
n:fcd1f6cc70809e67165baafce33615733b57f4dd.c1f4a388bfa5af9936e33fc2e619584a0e9521b3.0c02b49c3500e363ffb0
06d76f0ed1ab9a81514b.org A i:c65e644481a09ec6a2c10c524bafb553202b09ba 10 A
i:bc8c52505ccce2243224a3c16dc7034b6ea1939b 10 A i:7e6c5e10c6e0754c2d50e3e98301158da76136fa 10 A
i:370b90c65e4a695c682480c35e240dc64e98ecb4 10
```

```
604706.677000 c:5ba638a1731f0d9e86f375777ccb9f498d41a037 r:42f234f0dbeb02bef102b752d24624d3efc62fcf A
n:06f1385979232111dd4aeda1e4bf8c29eabe5970.eb6c27514cb41a77d0e16a9e0b779890ba5153cd.com CNAME
n:717f0ec9bab8f8d87d5b73b7d753f8562662401f.e80e7afb6af2560a7eaf189570e93469bd5f419e.23f689eeba26ef06362b
1dfcf69ad65f239ccc84.3ff51baf92d014241338573e1c9fb525e7cc971ad.com 136 A
i:b20ae140923f9ab50bbc0c72fe8f3ee296ab1d79 26 A i:6fda8829cd44159cbf6f10f17af9119331a4b04e 26 A
i:1de5ded2f1ff9bf978785a43722883f5467b5cb2 26 A i:941b6260a36cd9f004e0bc6ed5cc069b7e3debf0 26 A
i:a8726f70910c5f518b72c3f6a187d48268fd7f08 26 A i:85b2cce103c54529b3c2a32eab7c6293b438acfd 26 A
i:92d0d6f8e22505950f8b97882f13d35ea837b005 26 A i:e9064e92962c768fb632db6d359ecd81ff4d392e 26
```

Timestamp in seconds from
a reference point

Cipher domain

Time to Live (TTL)

Resource record type

Caso prático – DNS Recursivo

- O **processamento** dos **dados** pode gerar arquivos apenas com os **campos desejados** (chamados de traces).
- Precisa de atenção ao **tipo** de dados, **unidades** e **intervalo** de valores.
- Otimização para o armazenamento.

```
# time, id, client, size, ttl
604706.316754, 0, 0, 170, 10
604706.316754, 0, 0, 170, 10
604706.316754, 0, 0, 170, 10
604706.316754, 0, 0, 170, 10
604706.399381, 1, 1, 89, 300
604706.399381, 1, 1, 89, 300
604706.677, 2, 2, 261, 136
604706.677, 2, 2, 129, 26
604706.677, 2, 2, 129, 26
604706.677, 2, 2, 129, 26
604706.677, 2, 2, 129, 26
604706.677, 2, 2, 129, 26
```

Compartilhamento dos dados

- Na prática, como **tornar** os dados de infraestrutura usados em pesquisas **abertos**?
- Avaliações prévias:
 - Quais dados são **necessários** para pesquisa?
 - Tipo da pesquisa envolve dados sensíveis.
 - **Como** compartilhar esses dados (meio, periodicidade, etc.).
 - **Processamento** dos dados e **previsão** de **publicização**.
 - Estudo de caso na infraestrutura (**aplicação da pesquisa**).
- Construção de um **plano de gestão de dados**.
- Definições posteriores:
 - **Compartilhamento** das avaliações práticas e **publicização** de **resultados**.

Como e quando compartilhar (cont.)

- **Plataformas** para **disponibilização** e **consulta** de dados abertos incluem:
 - Open Science Framework (OSF) (<https://osf.io/>)
 - Zenodo (<https://zenodo.org/>)
 - Figshare (<https://figshare.com/>)
 - Mendeley Data (<https://data.mendeley.com/>)
 - Registry (<https://www.re3data.org/>)
 - Self-hosted (ex.: <https://dataverse.org/>)
- Possibilidade de utilização de **serviços institucionais** existentes.

Conclusão

- Ciência Aberta traz uma **visão** diferente sobre o **ciclo da pesquisa científica** e da interação entre os atores envolvidos.
 - Descreve **diversas práticas** de pesquisa que promovem abertura, acessibilidade e transparência.
 - Aumentar a capacidade de **publicização** da pesquisa **globalmente** e torná-la acessível além do meio acadêmico.
- **Compartilhamento** de **dados** da **infraestrutura** permite que em um mundo cada vez mais orientado a dados, **casos práticos** sejam usados na pesquisa para uma maior orientação ao impacto dos serviços.
 - Diversos tipos de dados podem ser compartilhados.
 - Formato do compartilhamento é necessário.
- A **acessibilidade** dos dados é **necessária** para que outros estudos possam **reproduzir** e **verificar** resultados.

Referências

- BEZJAK, Sonja; et al.. **Open Science Training Handbook**. Zenodo, 2018. doi: 10.5281/zenodo.1212496.
- CHIFUMI, Nishioka. **Open Science in the Japanese Humanities**. 2025.
- CHIFUMI, Nishioka. **Citation Advantages of Green Open Access Articles: A Case Study at Kyoto University**. 2021.
- SILVEIRA, Lúcia da et al. **Taxonomia da Ciência Aberta: revisada e ampliada**. Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação, Florianópolis/SC, Brasil, v. 28, p. 1–22, 2023. DOI: 10.5007/1518-2924.2023.e91712. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/91712>.
- VICENTE-SAEZ, Ruben; MARTINEZ-FUENTES, Clara. **Open Science now: A systematic literature review for an integrated definition**. Journal of business research, v. 88, p. 428-436, 2018.
- OKI, Eiji; **Linear programming and algorithms for communication networks: a practical guide to network design, control, and management**. CRC Press, 2012.
- FERREIRA, Ibirisol Fontes; OKI, Eiji. **Latency-Aware Cache Mechanism for Resolver Service of Domain Name Systems**. In: NOMS 2024-2024 IEEE Network Operations and Management Symposium. IEEE, 2024. p. 1-4.
- ALLMAN, Mark. **Case Connection Zone DNS Transactions**, 2021.
- AHMED, Mahfooz et al. **Factors influencing open science participation through research data sharing and reuse among researchers: a systematic literature review**. Knowledge and Information Systems, v. 67, n. 3, p. 2801-2853, 2025.
- BOCHYNSKA, Agata . **Open and reproducibleresearch: An overview**. 2025.