

ACESSO ABERTO A DADOS DE PESQUISA NO BRASIL

Planejamento do Protótipo de Repositório
Relatório 2018



Coordenador

Rafael Port da Rocha (Universidade Federal do Rio Grande do Sul)

Coordenadora adjunta

Sônia Elisa Caregnato (Universidade Federal do Rio Grande do Sul)

Pesquisadores

Universidade Federal do Rio Grande do Sul

Caterina Groposo Pavão

Paula Caroline Schifino Jardim Passos

Rene Faustino Gabriel Junior

Samile Andréa de Souza Vanz

Universidade Federal de Rio Grande

Eduardo Nunes Borges

Luís Alberto Barbosa Azambuja

Bolsistas

Iván Andrés Fornos Angues

Victor Andrews Garcia Lima

Contato

e-mail: dadosdepesquisa@rnp.br

site: <https://dadosdepesquisa.rnp.br>

Rua Ramiro Barcelos, 2705 - Campus Saúde

Sala: 106 - Anexo 1

Brasil - Porto Alegre - RS - CEP 90.035-007

Telefone: +55(51)3308.5942



Este relatório é licenciado sobre a licença CC BY - Creative Commons Attribution 4.0 International License.

Como citar:

ROCHA, Rafael Port da; AZAMBUJA, Luís Alberto Barbosa; BORGES, Nunes Borges; CAREGNATO, Sônia Elisa; GABRIEL JUNIOR, Rene Faustino; PAVÃO, Caterina Groposo; PASSOS, Paula Caroline Schifino Jardim; VANZ, Samile Andrea de Souza. Acesso aberto a dados de pesquisa no Brasil: Planejamento de Protótipo do Repositório: relatório 2018.

R672 ROCHA, Rafael Port da

Acesso aberto a dados de pesquisa no Brasil : Planejamento de Protótipo do Repositório : relatório 2018 / Rafael Port da Rocha et al. - Porto Alegre, RS : UFRGS, 2018.

39p.

Relatório do projeto RDP Brasil: Rede de Dados de Pesquisa Brasileira, – Universidade Federal do Rio Grande do Sul; Universidade Federal do Rio Grande.

1. Dados abertos de pesquisa. 2. Repositório de dados.
3. Tecnologias. I. Título



SUMÁRIO

1 Introdução	2
3 Ambiente do Repositório	9
4 Conjuntos de Dados do Repositório	14
6 Produção dos Conjuntos de Dados	18
7 Acesso e Uso dos Conjuntos de Dados	20
8 Armazenamento e Preservação	22
9 Implantação em Dataverse	22
Referências	28
Anexo 1 bn – Dataset das Medições IPÊ	29

1 Introdução

A Rede Nacional de Ensino e Pesquisa (RNP) lançou, em 2017, um edital para promover pesquisas sobre compartilhamento dos dados coletados, gerados e utilizados pelos pesquisadores brasileiros. O projeto contemplado no edital é fruto da parceria entre a Universidade Federal do Rio Grande do Sul (UFRGS) e a Universidade Federal do Rio Grande (FURG), envolvendo o Centro de Documentação e Acervo Digital de Pesquisa (CEDAP), órgão auxiliar da Faculdade de Biblioteconomia e Comunicação (FABICO), e o Centro de Processamento de Dados (CPD), ambos da UFRGS, e o Grupo de Pesquisa em Gerenciamento de Informações do Centro de Ciências Computacionais (C3), da FURG. O projeto RDP Brasil - Rede de Dados de Pesquisa Brasileira tem cinco objetivos planejados pela equipe, que são os seguintes:

1. Identificação de práticas de AADP em instituições brasileiras;
2. Mapeamento de potenciais usuários nacionais de serviços de AADP;
3. Elaboração do portal web para a comunidade nacional em AADP;
4. Levantamento comparativo dos serviços e soluções tecnológicas para compartilhamento de dados;
5. Desenvolvimento de uma proposta de solução tecnológica.

Este documento apresenta o Planejamento para implantação do Protótipo de Repositório da Rede de Dados da Pesquisa - RDP Brasil, que corresponde a proposta de solução tecnológica desenvolvida no objetivo 5.

2 Considerações sobre o Planejamento do Repositório

O Planejamento do Repositório foi desenvolvido a partir da seguinte definição para Repositório:

Ambiente de armazenamento e gerenciamento de materiais digitais. Esse ambiente constitui-se de uma solução informatizada em que os materiais são capturados, armazenados, preservados e acessados. Um repositório digital é, então, um complexo que apoia o gerenciamento dos materiais digitais, pelo tempo que for necessário, e é formado por elementos de hardware, software e metadados, bem como por uma infraestrutura organizacional e procedimentos normativos e técnicos. [i]

O Planejamento do Repositório foi elaborado com base no modelo de referência para Repositório, OAIS [ii] (Open Archival Information System), e na metodologia ligada à OAIS para identificar a interface entre o produtor e o repositório:

- CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS. Producer-Archive Interface Methodology Abstract Standard. CCSDS 651.0-M-1. 2004
<https://public.ccsds.org/Pubs/651x0m1.pdf>
- CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS. Reference model for an open archival information system (OAIS). CSDS 650.0-M-2. 2012
<https://public.ccsds.org/pubs/650x0m2.pdf>

OAIS é um modelo de referência que proporciona uma linguagem comum para facilitar a discussão entre comunidades interessadas no desenvolvimento de repositórios com responsabilidade em preservação digital. OAIS descreve o **ambiente** em que um repositório reside, os seus **componentes funcionais** e a infraestrutura de **informação** que suporta o processo de arquivamento. A figura 2A apresenta o ambiente do repositório (produtor, consumidor e gestão), os componentes funcionais (ingestão, acesso, armazenamento, gestão de dado e planejamento da preservação), e os tipos de pacotes de informação (pacote de submissão de informação – SIP, pacote de armazenamento de informação – AIP, e pacote de disseminação de informação – DIP). Definições sobre OAIS e seus elementos são apresentadas no Quadro 2C.

Figura 2A - Ambiente, Entidades Funcionais e Pacotes de Informação do Modelo OAIS.



Fonte: CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS. Producer-Archive Interface Methodology Abstract Standard. CCSDS 651.0-M-1. 2004

Tendo como base o modelo de referência OAIS, o Planejamento do Repositório foi organizado nos seguintes tópicos:

- **Ambiente do Repositório:** identifica e caracteriza o ambiente do repositório, seus produtores, consumidores, gestores, unidades organizacionais que compõem o repositório, políticas para criar e gerir unidades

- **Conjuntos de Dados:** identifica tipos de conjuntos de dados, suas estruturas de representação dos conjuntos de dados e os formatos dos arquivos;
- **Descrição e Documentação dos Conjuntos de Dados:** documentação, metadados e padrões para descrever e usar os conjuntos de dados;
- **Produção dos Conjuntos de Dados:** submissão dos conjuntos de dados ao Repositório, incluindo produção, validação e verificação de pacotes de submissão, por pessoas autenticadas e autorizadas;
- **Acesso e Uso dos Conjuntos de Dados:** Políticas de acesso e uso dos conjuntos de dados, incluindo restrições de acesso e licenças;
- **Armazenamento e Planejamento da Preservação:** armazenamento a longo prazo dos conjuntos de dados, planejamento e execução da preservação digital

O planejamento do repositório também relaciona requisitos FAIR^[iii] e critérios para obtenção de Certificação de Repositório Digital Confiável de Core Trust Seal^[iv]:

- Princípios FAIR - <https://www.force11.org/group/fairgroup/fairprinciples>^[v]
- CTS - Core Trustworthy Data Repositories Requirements v01.
<https://www.coretrustseal.org/wp-content/uploads/2017/01/Core-Trustworthy-Data-Repositories-Requirements-0100.pdf>

Repositório Digital Confiável é aquele

cuja missão é fornecer acesso de longo prazo a recursos digitais gerenciados; que aceita a responsabilidade pela manutenção a longo prazo dos recursos digitais em nome de seus depositantes e em benefício dos usuários atuais e futuros; que projeta seu(s) sistema(s) de acordo com as convenções e os padrões comumente aceitos para garantir o gerenciamento, acesso e segurança contínuos dos materiais depositados nele; que estabelece metodologias para avaliação de sistemas que atendem às expectativas de confiabilidade da comunidade; que cumpre suas responsabilidades com depositantes e usuários de forma aberta e explícita; e cujas políticas, práticas e desempenho podem ser auditados e medidos ^[vi]

Os princípios FAIR ^{iii [vii]} indicam que dados devem ser localizáveis (*Findable*), acessíveis (*Accessible*), interoperáveis (*Interoperable*) e Reusáveis (*Reusable*). Para dados serem FAIR, estes devem ser atribuídos a identificadores únicos, persistentes e globais (F). Devem ser descritos por metadados indexáveis e ricos (F), representados em linguagens formais (I), aceitos pela comunidade (R), com atributos relevantes, precisos e úteis ao contexto (R), incluindo metadados de proveniência (R) e usando vocabulários controlados que seguem princípios FAIR (I). Estes dados devem ser recuperáveis pelo seu identificador através de um

protocolo de comunicação padronizado, aberto, gratuito (A). Também devem ser acompanhados de licenças claras e acessíveis (R), e referências qualificadas devem ligar (meta)dados para enriquecer o conhecimento sobre os mesmos (I).

Este planejamento objetiva na elaboração de um protótipo de Repositório. Em função disso, alguns pontos levantados, principalmente com relação ao seu ambiente (infraestrutura e comprometimento organizacional de gestão), são propostas a serem atendidas caso o protótipo passe a ser um repositório digital que venha a assumir as responsabilidades de preservação a longo prazo.

O Quadro 2A apresenta os princípios FAIR, cuja codificação indicadora será citada no Planejamento do Repositório, indicando aspectos relacionados com esses princípios. O Quadro 2B apresenta os critérios para repositório digital confiável da certificação Core Trust Seal[viii], cujas codificação também é citada no Planejamento do Repositório. O Quadro 2C apresenta conceitos usados neste Planejamento.

Quadro 2A - Princípios FAIR

<p>F Findable</p> <p>F1 Meta(dados) com identificador persistente</p> <p>F2 Metadados ricos</p> <p>F3 (Meta)dados usados por recursos de busca</p> <p>F4 Metadados com identificador do conjunto de dados</p>
<p>A Accessible</p> <p>A1 Meta(dados) recuperável por identificador</p> <p>A1.1 Por protocolo de recuperação aberto</p> <p>A1.2 Autenticação e autorização</p> <p>A2 Metadados persistem quando dados são removidos</p>
<p>I Interoperable</p> <p>I1 Representação do conhecimento, vocabulários, ontologias</p> <p>I2 Vocabulários FAIR</p> <p>I3 Referências a outros meta(dados), contexto</p>
<p>R Re-usable</p> <p>R1 (Meta)dados descritos ricamente por atributos relevantes e precisos</p> <p>R1.1 Licença</p> <p>R1.2 Proveniência</p> <p>R1.2 Padrões relevantes para a comunidade</p>

Fonte: FAIR ^v

Quadro 2B - Critérios Core Trust Seal

R1	Missão
R2	Licenças
R3	Acesso contínuo, a longo prazo
R4	Confidencialidade e ética – em conformidade com normas
R5	Infraestrutura organizacional
R6	Orientações e feedback de especialistas para manter valor dos dados
R7	Integridade (fixidez, completeza, rastreabilidade, versões) e autenticidade (proveniência, auditoria) dos dados na ingestão, armazenamento e acesso
R8	Dados e metadados aceitos mediante critérios que garante compreensão e relevância aos usuários
R9	Procedimentos documentados para armazenamento
R10	Planejamento da preservação
R11	Dados e metadados com qualidade, e informações para avaliações de qualidade
R12	Atividades de acordo com fluxos de trabalhos definidos
R13	Descoberta e identificação de dados (busca, metadados, harvesting, citação, identificação persistente)
R14	Compressão e uso dos dados no futuro, mesmo com mudanças tecnológicas
R15	Segurança, proteção das instalações e seus dados, produtos, serviços e usuários
R16	Infraestrutura técnica (software e hardware apropriados, maximizando disponibilidade)

Fonte: Core Trust Seal ^{viii}

Quadro 2C - Conceitos

Repositório Digital	Ambiente de armazenamento e gerenciamento de materiais digitais. Esse ambiente constitui-se de uma solução informatizada em que os materiais são capturados, armazenados, preservados e acessados. Um repositório digital é, então, um complexo que apoia o gerenciamento dos materiais digitais, pelo tempo que for necessário, e é formado por elementos de hardware, software e metadados, bem como por uma infraestrutura organizacional e procedimentos normativos e técnicos. ⁱ
----------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Repositório Digital Confiável	Repositório cuja missão é fornecer acesso de longo prazo a recursos digitais gerenciados; que aceita a responsabilidade pela manutenção a longo prazo dos recursos digitais em nome de seus depositantes e em benefício dos usuários atuais e futuros; que projeta seu(s) sistema(s) de acordo com as convenções e os padrões comumente aceitos para garantir o gerenciamento, acesso e segurança contínuos dos materiais depositados nele; que estabelece metodologias para avaliação de sistemas que atendem às expectativas de confiabilidade da comunidade; que cumpre suas responsabilidades com depositantes e usuários de forma aberta e explícita; e cujas políticas, práticas e desempenho podem ser auditados e medidos. ^{vi}
OAIS	Modelo de referência que proporciona uma linguagem comum para facilitar a discussão entre comunidades interessadas na preservação. Descreve o ambiente em que um arquivo (repositório) reside, os componentes funcionais de um arquivo e a infraestrutura de informação que suporta o processo de arquivamento.
Ambiente OAIS	O ambiente de um repositório OAIS é composto pelo produtor, que fornece a informação a ser preservada, pelo consumidor, que interage com o ambiente para localizar e obter a informação preservada de seu interesse, e pela gestão, que determina as políticas, controlando as responsabilidades de gestão
Componentes funcionais de OAIS	O modelo funcional de um repositório OAIS compreende em entidades funcionais que compõem um repositório OAIS: ingestão, acesso, planejamento da preservação, armazenamento, gestão de dados e administração.
Produtor	Papel desempenhado pelas pessoas ou sistemas clientes que fornecem as informações a serem preservadas, podendo incluir outros sistemas OAIS. ⁱⁱ
Consumidor	Papel desempenhado por pessoas ou sistemas clientes que interagem com os serviços OAIS para encontrar informações preservadas de interesse e para acessar essas informações em detalhes, podendo incluir outros sistemas OAIS ⁱⁱ
Comunidade Designada	Um grupo identificado de potenciais consumidores que devem ser capazes de entender um conjunto específico de informações. ⁱⁱ

Gestor	Papel desempenhado por aqueles que definem a política geral do OAIS como parte em um domínio de política mais amplo, por exemplo, como parte de uma organização maior. [ix]
Pacote de Informação	Um contêiner lógico composto de Informações de Conteúdo opcionais e Informações de Descrição de Preservação opcionais associadas. Contém informação empacotada usada para delimitar e identificar as Informações de Conteúdo e de Descrição do Pacote usadas para facilitar as pesquisas pelas Informações de Conteúdo.
Pacote de Submissão de Informação	SIP - Submission Information Package Pacote de Informação que contém as informações a serem submetidas a um Repositório OAIS
Pacote de Armazenamento de Informação	AIP - Archival Information Package Pacote de Informação que contém as informações que são armazenadas em um Repositório OAIS
Pacote de Disseminação de Informação	DIP - Dissemination Information Package Pacote de Informação que contém as informações a serem disseminadas a um Repositório OAIS
Metadados	Dados estruturados que provém informação sobre um ou mais aspectos de outros dados. Podem ser classificados de diferentes formas, tais como descritivos, estruturais e administrativos (técnicos, direitos e de preservação digital).
Estudo	Coleção de arquivos de dados resultantes da coleta intencional de dados por meio de solicitação, observação ou coleta de fontes secundárias para um propósito; utilizando instrumentos e metodologias de coleta descritos; e expressos como arquivos de dados com estruturas lógicas relacionadas ao instrumento de coleta de dados
Conjunto de Dados	Arquivos de dados descritos por uma única estrutura

<p>Livro de Códigos</p> <p>Codebook</p>	<p>Livro que descreve o conteúdo, a estrutura e o layout de uma coleta de dados. Um livro de códigos bem documentado deve conter informações destinadas a ser completas e autoexplicativas para cada variável em um arquivo de dados [x]</p>
<p>Proveniência dos Dados</p>	<p>Informações sobre entidades, atividades e pessoas envolvidas na produção de um dado ou coisa, que podem ser usadas para formar avaliações sobre sua qualidade, confiabilidade ou fidedignidade [xi]</p>

3 Ambiente do Repositório

Essa seção caracteriza o ambiente do Repositório. Envolve a identificação e a caracterização:

- da **Missão** do repositório;
- do **Gestor do Repositório**: órgão responsável pela gestão do repositório, que assumirá a responsabilidade pela preservação e acesso às informações armazenadas a longo prazo. Esse órgão deve dotar de responsabilidades pela gestão do repositório, ter equipe com qualificação e número suficientes e definidas, planos de treinamento e de atualização tecnológica, infraestrutura, orçamento e transparência em gestão;
- dos **Produtores e Consumidores de Informação do Repositório**: usuários que produzem e usam as informações a serem preservadas no repositório. Envolve identificar, por exemplo, se os produtores são:
 - **grupos** de pesquisa que necessitam publicar e gerenciar seus vários estudos;
 - **estudos** que necessitam publicar seus conjuntos de dados;
 - **autores** de publicações, que necessitam publicar os dados dessas publicações (chamados de conjuntos de dados de replicação);
 - **instituições** que desejam prover serviços de dados aos seus grupos, pesquisas ou pesquisadores;
 - **instituições** que desejam publicar seus dados para a pesquisa.
- da **Estrutura Organizacional do Repositório**: definição de quais e como serão criadas as unidades que compõem o Repositório. Compreende em identificar:
 - A **política** do Repositório para criar as unidades organizacionais que irão compor o Repositório. Envolve identificar, por exemplo, se o Repositório será composto por unidades como:
 - unidades que representam **grupos de pesquisa**, que armazenam seus vários estudos/conjuntos de dados;
 - unidades que representam conjuntos de dados de **estudos** de pesquisas, multidisciplinares ou disciplinares, de produtores ligados a várias instituições;

- unidades que armazenam conjuntos de dados associados a **publicações** (conjuntos de dados de referência)
- unidades que representam **instituições ou organizações**, com autonomia para definir suas próprias estruturas organizacionais e políticas
- A **estrutura organizacional** básica, inicial do Repositório;

Os quadros 3A, 3B, 3C, 3D e 3E apresentam o planejamento do ambiente do Repositório. O quadro 3A apresenta a **missão** do Repositório e sua infraestrutura que demonstra que o mesmo é capaz de garantir acesso a longo prazo aos dados. O quadro 3B apresenta os **produtores** e os **consumidores** dos dados do Repositório, e caracteriza sua **comunidade indicada**, isto é, o grupo de potenciais consumidores que tem a capacidade de compreender os conjuntos de dados.

Os Quadros 3C e 3D dizem respeito a **Estrutura Organizacional do Repositório**. O Quadro 3C apresenta a **política** do repositório para criar unidades organizacionais. Os Quadro 3D e 3E apresentam a **estrutura** organizacional inicial do Repositório, criada de acordo com as políticas definidas no Quadro 3C.

Quadro 3A – Missão e Gestão do Repositório

<p>Missão do Repositório</p> <ul style="list-style-type: none"> ● Critérios e princípios relacionados: CTS.R1
<ul style="list-style-type: none"> ● Investigar a implantação e operação de um repositório para o acesso aberto a dados de pesquisa no Brasil (via prototipação)
<p>Responsabilidades, Comprometimentos, Equipe e Recursos</p> <ul style="list-style-type: none"> ● Critérios e princípios relacionados: CTS.R3, CTS.R4, CTS.R5, CTS.R6, CTS.R15, CTS.R16
<ul style="list-style-type: none"> ● GT-RDP Brasil (fase de Prototipação)

Quadro 3B - Usuários e Comunidade do Repositório

<p>Produtores</p> <ul style="list-style-type: none"> • Critérios e princípios relacionados: CTS.R6. FAIR.R.1.2
<ul style="list-style-type: none"> • Grupos de Pesquisa na área de Redes de Computadores que necessitam compartilhar seus vários estudos • Pesquisadores da área de Redes de Computadores que desejam compartilhar conjuntos de dados de seus estudos • Pesquisadores da área de Redes de Computadores que necessitam compartilhar conjuntos de suas publicações • Grupo de Monitoração Rede IPÊ, da RNP, que deseja compartilhar dados de medições de rede de computador
<p>Consumidores</p> <ul style="list-style-type: none"> • Critérios e princípios relacionados: CTS.R6
<ul style="list-style-type: none"> • Pesquisadores em Redes de Computadores • Operadores e Gestores de Redes de Computadores • Federações de Repositórios de Dados da Pesquisa
<p>Comunidade Designada</p> <ul style="list-style-type: none"> • Critérios e princípios relacionados: CTS.R6
<p>Comunidade composta por pesquisadores, operadores e gestores de redes de computadores familiarizados com técnicas de medição de desempenho de redes, qualidade do serviço de conectividade, com dados de pesquisa em formatos associados a essas técnicas</p>

Quadro 3C - Política Organizacional do Repositório

<p>Política Organizacional do Repositório</p> <ul style="list-style-type: none"> • Critérios e princípios relacionados: CTS.R6, CTS.R12
<p>O repositório é constituído por unidades organizacionais (comunidades produtoras) que possuem autonomia para definir suas políticas, para a criação de políticas de submissão e acesso, e para criação e gestão de subunidades.</p> <p>O repositório será inicialmente composto por:</p> <ul style="list-style-type: none"> • Dados da pesquisa para a comunidade de Redes de Computadores • Dados da pesquisa provenientes de medições da Rede Ipê • Dados da Pesquisa do GT RDP Brasil

Quadro 3D - Unidades Organizacionais do Repositório: Produtores e gestores

Unidades Organizacionais do Repositório: Produtores e Gestores ● Critérios e princípios relacionados: CTS.R6, CTS.R12		
Unidade	Produtores	Gestores
RDP-RC Repositório de Dados da Pesquisa em Redes de Computadores	<ul style="list-style-type: none"> ● Pesquisadores em Redes de Computadores ● Grupos de Pesquisa de Redes de Computadores ● GT-RDP e Rede Ipê são as primeiras comunidades desse repositório 	Gestores do Repositório
GT-RDP	<ul style="list-style-type: none"> ● GT-RDP Brasil 	GT-RDP Brasil
REDE-IPÊ Repositório de Medições da Rede IPÊ	<ul style="list-style-type: none"> ● Equipe de Monitoramento da Rede Ipê 	Equipe da Rede IPE

Quadro 3E - Unidades Organizacionais do Repositório: Estrutura e política organizacional

Unidades Organizacionais do Repositório: Estrutura e política organizacional Critérios e princípios relacionados: CTS.R6, CTS.R12	
Unidade	Estrutura e Política Organizacional
RDP-RC Repositório de Dados da Pesquisa em Redes de Computadores	<p>Contém</p> <ul style="list-style-type: none"> ● Grupos de Pesquisa, que contém conjuntos de dados de seus estudos ● Conjuntos de dados de estudos individuais ou relacionados à publicações <p>Política para cada Grupos de Pesquisa (comunidade)</p> <ul style="list-style-type: none"> ● Papéis: coordenador, colaborador e membro ● Curador: autorizado a publicar conjuntos de dados, e para restringir de acesso a conjuntos de dados ● Colaborador: autorizado a criar, editar e remover conjuntos de dados não publicados ● Membro: autorizado a acessar conjuntos de dados não publicados e arquivos de acesso restrito de conjuntos de dados publicado

	<p>Política para Conjuntos de Dados de Pesquisas Individuais ou de publicações</p> <ul style="list-style-type: none"> ● Papéis: Curador ● Curador: Autorizado a editar e publicar conjuntos de dados
<p>RDP-IPÊ Repositório de Medições da Rede IPÊ</p>	<p>Contém</p> <ul style="list-style-type: none"> ● Conjunto de Dados para pesquisas referentes à monitoração da Rede Ipê. Dados referentes a medições da REDE Ipê. <ul style="list-style-type: none"> ○ O indicador 3 expressa a qualidade do serviço de conectividade, através de pontuação combinada sobre medidas de desempenho da rede. Os pontos são atribuídos a duas características da rede: taxa média de perda de pacotes e retardo médio de entrega de pacotes. (fonte: https://www.rnp.br/file/21361/download?token=XIsIHBCW) <p>Política</p> <ul style="list-style-type: none"> ● Papéis: coordenador, colaborador e membro ● Curador: autorizado a publicar conjuntos de dados, e para restringir de acesso a conjuntos de dados ● Colaborador: autorizado a criar, editar e remover conjuntos de dados não publicados ● Membro: autorizado a acessar conjuntos de dados não publicados e arquivos de acesso restrito de conjuntos de dados publicado
<p>GT-RDP-Brasil</p>	<p>Contém</p> <ul style="list-style-type: none"> ● Conjuntos de dados dos estudos realizados pelo GT-RDP Brasil <p>Política</p> <ul style="list-style-type: none"> ● Papéis: coordenador, colaborador e membro ● Curador: autorizado a publicar conjuntos de dados, e para restringir de acesso a conjuntos de dados ● Colaborador: autorizado a criar, editar e remover conjuntos de dados não publicados ● Membro: autorizado a acessar conjuntos de dados não publicados e arquivos de acesso restrito de conjuntos de dados publicado

4 Conjuntos de Dados do Repositório

Essa seção identifica as informações (conjuntos de dados) a serem preservados pelo repositório e suas estruturas de representação. Envolve a definição dos:

- **Tipos de dados:** tabulares, texto, imagens, sons, vídeos, etc;
- **Formatos dos arquivos:** txt, pdf, csv, excell, mp3, mp4, jpg, etc;
- **Estrutura do pacote :** arquivos com documentação, arquivos de dados, arquivos para outras finalidades, estrutura para organizar e nomear os arquivos;
- **Esquemas de metadados:** esquemas usados para descrever o conjunto de dados (ou compatibilidade dos metadados com esses esquemas)
- **Versionamento:** gerenciamento de versões decorrentes de atualizações e/ou produções periódicas de dados;

Quadro 4A- Tipos de Dados

Informação a ser Preservada pelo Repositório e suas Estruturas de Representação <ul style="list-style-type: none"> ● Critérios e princípios relacionados: CTS.R7, CTS.R8, CTS.R9, CTS.R11, FAIR.R1 	
Conjunto de Dados	Estruturas
Conjunto de Dados de: Pesquisa em Redes de Computadores GT-RDP Brasil	Tipos de dados: <ul style="list-style-type: none"> ● tabelas, textos, imagens, sons, vídeos, ... Formatos de Arquivos: <ul style="list-style-type: none"> ● csv, excell, pdf, Estrutura do Pacote: <ul style="list-style-type: none"> ● Arquivos com documentação: <ul style="list-style-type: none"> ○ Arquivos que documentam os dados com relação a suas estruturas, de acordo com as especificações definidas pelo coordenador do grupo de pesquisa ou produtor do conjunto de dados ○ Livro de Códigos ● Arquivos com dados: <ul style="list-style-type: none"> ○ Arquivos que contém dados, de acordo com as especificações definidas pelo coordenador do grupo de pesquisa ou produtor do conjunto de dados ● Outros documentos, de acordo com as especificações definidas pelo coordenador do grupo de pesquisa ou produtor do conjunto de dados. Versionamento: <ul style="list-style-type: none"> ● Sim

<p>Conjunto de dados de Medições da Rede Ipê</p>	<p>Tipos de dados:</p> <ul style="list-style-type: none"> ● tabelas e documentos textuais <p>Formatos de Arquivos:</p> <ul style="list-style-type: none"> ● csv, tab, pdf, tab <p>Estrutura do Pacote (detalhes em anexo):</p> <ul style="list-style-type: none"> ● Conjuntos de Dados: Medições AAAA (ex.: Meidições 2019) <ul style="list-style-type: none"> ● Um conjunto para cada ano ● Codebook - Documento que contém a descrição de cada arquivo de dados ● Estratégia de Coleta – Documentação que informa como os dados foram coletados (no Codebook) ● Arquivos de Dados <ul style="list-style-type: none"> ○ Arquivos mensais, cada um contendo dados de medição da rede IPE em formato tabular ○ 01-2019.tab ○ 02-2019..tab ○ <p>Versionamento:</p> <ul style="list-style-type: none"> ● Cada medição mensal leva a uma nova versão do conjunto de dados
---------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

5 Descrição e Documentação dos Conjuntos de Dados do Repositório

Essa seção identifica as informações necessárias para descoberta e uso dos dados, que podem tanto estar representadas através de metadados, quando em documentos. Inclui:

- Informações de **proveniência** dos dados: atividades e pessoas envolvidas na produção do dado, que podem ser usadas para formar avaliações sobre sua qualidade, confiabilidade ou fidedignidade. Registro da origem, fonte, mudanças ocorridas na informação, versões.
- Informações relevantes para **citação** dos dados e para interoperar com serviços de citação: prover meios para permitir a citação dos dados, via metadados de citação e identificadores persistentes, atribuindo créditos aos produtores, disponibilizando informações de proveniência e que permitem a verificação de fixidez, considerando que os dados em questão podem ser versões, séries, subconjuntos, etc
- Informações que **documentam** as estruturas dos dados (como variáveis), livros de códigos/codebooks, como significado de variáveis;
- Informações sobre **direitos de uso** dos dados (licenças);
- Informações **descritivos** necessárias e usadas pela comunidade para descrever seus dados; padrões de esquemas de metadados usados ou compatíveis, vocabulários controlados

Quadro 5A – Descrição e Documentação dos Conjuntos de Dados

Informações de proveniência <ul style="list-style-type: none">• Critérios e princípios relacionados : FAIR.R1.2, CTS.R7, CTS.R11, CTS.R14
<p>https://www.dataone.org/best-practices/provide-citation-and-document-provenance-your-dataset https://www.go-fair.org/fair-principles/r1-2-metadata-associated-detailed-provenance/</p> <p>Para RDP-IPÊ - Rede IPÊ:</p> <p>As medidas coletadas para o indicador da qualidade são realizadas através de máquinas de serviço, instaladas nos 27 PoPs da RNP. Cada máquina de serviço envia pacotes ICMP de teste para todas as demais, gerando uma grande matriz 27x27 de medições. Os pacotes de teste são enviados em intervalos aleatórios de distribuição exponencial. Os valores de RMédio e PPerda são calculados como a média aritmética das medianas obtidas em todas as máquinas de serviço. Os testes são feitos no período entre 08h00min às 18h00min (GMT -3).</p> <p>O script que dispara os pings nas máquinas de serviço gera um arquivo para cada PoP testado diariamente. Os arquivos ficam gravados na pasta em diretório interno em formato txt. O formato de nome de arquivo que este script gera é o seguinte: PoP-XX_AAAAMMDD. O conteúdo das medições tem o formato: Timestamp_GMT : PERC_PERDA : RTT_MS</p> <p>Um script que roda diariamente através do daemon cron na máquina central no Rio de Janeiro faz a coleta do resultado dos pings de todas as máquinas de serviço. Este script executa um outro script que faz a média dos resultados gerando um resultado parcial. O resultado dos dados após o cálculo da média fica armazenado em diretórios locais, sendo um arquivo para cada dia no formato:</p> <pre>POP_DEST:PERDA_MDN:LAT_MIN:LAT_MED:LAT_MAX:STD_DVN:LAT_10_PERC:LAT_MDN:LAT_90_PERC</pre> <p>(fonte: https://www.rnp.br/file/21361/download?token=XlslHBCW)</p> <p>Para Da Pesquisa em Redes de Computadores e GT-RDP Brasil</p> <p>Dados gerados pelas pesquisas dos grupos</p>
Informações para citação das informações <ul style="list-style-type: none">• Critérios e princípios relacionados: FAIR.R1, CTS.R11, CTS.R13, CTS.R14

Metadados obrigatórios, segundo DataCite[xii]: Identificador, Criador, Título, Publicador, Ano de Publicação, Tipo de recurso

- Metadados recomendados, segundo DataCite: Assunto, Colaborador, Data, Identificador Relacionado, Descrição, (GeoLocalização)
- Metadados opcionais, segundo DataCite: Idioma, (identificador alternativo), tamanho, formato, versão, direitos, (referência de financiamento)
- Citração no padrão Dataverse (UNF)

Documentação do significado das estruturas dos dados

- Critérios e princípios relacionados: FAIR.R1, CTS.R7, CTS.R7, CTS.R11, CTS.R13, CTS.R14

Para **RDP-RC - Repositório de Dados da Pesquisa em Redes de Computadores:**

- Livro de Códigos (Codebook), opcional e de acordo com as características de cada conjunto de dados, definidos pelo coordenador do grupo ou produtor do conjunto de dados;
- Outra documentação, se necessária.

Para **RDP-IPÊ - Rede IPÊ:**

- Livro de Códigos elaborado a partir de <https://www.rnp.br/file/21361/download?token=XIsIHBCW>

Informações de Direitos de Uso – Licença

- Critérios e princípios relacionados: FAIR.R1, CTS.R7, CTS.R8, CTS.R11, CTS.R13, CTS.R14

Licenças de Uso

-

Informações Descritivas

- Critérios e princípios relacionados: FAIR.F2, FAIR.F3, FAIR.F4, FAIR.A2, FAIR.I1, FAIR.I3, FAIR.R1, FAIR.R1.1, FAIR.R1.3, CTS.R7, CTS.R8, CTS.R11, CTS.R13, CTS.R14

Esquemas mínimos necessários:

- Dublin Core Básico (metadados mínimos para OAI-PMH)
- DataCite (Metadados obrigatórios para registro de DOI em DataCite)

<p>Vocabulários Controlados e Controles de Autoridades</p> <ul style="list-style-type: none"> ● Critérios e princípios relacionados: FAIR.F3 , FAIR.I2, CTS.R11,CTS.R13, CTS.R14
<p>Vocabulários Controlados</p> <ul style="list-style-type: none"> ● Formato de Arquivo: mime
<p>Recursos para Fixidez e Integridade das informações</p> <ul style="list-style-type: none"> ● CTS.R7
<ul style="list-style-type: none"> ● Arquivo: MD5 ● Dataset: UNF

6 Produção dos Conjuntos de Dados

Envolve identificar e definir as políticas para submissão dos dados ao repositório, observando:

- **Propriedade intelectual e acordos de submissão:** Gerenciamento de acordos de submissão realizados entre o repositório e o produtor; os direitos dados ao repositório, as limitações impostas pelos detentores dos direitos, as licenças envolvidas, as obrigações do Repositório com relação aos dados;
- **Transferência legal da custódia dos dados ao repositório:** Procedimentos e fluxos definidos para recepção, aceitação/rejeição, verificação e validação dos dados transferidos ao Repositório. Confidencialidade e autenticação da informação na transferência Repositório-Consumidor: registro dos acordos de submissão e das ações realizadas na transferência, técnicas para transferência segura (criptografia, https), técnicas para garantir a autenticidade da informação (como assinatura), procedimentos para autenticação e verificação de autoridade para usuários envolvidos na transferência. Validação de identificadores globais e persistentes para o conjunto de dados e para recursos relacionados, como ORCID dos produtores;
- **Submissão por máquina, em ambientes distribuídos e/ou em lote:** Recepção de pacotes submetidos por outros ambientes e/ou depósito de materiais aprovados em outros ambientes, via protocolos e estruturas de pacotes de submissão aceitos e padronizados, como Sword , OAI-PMH (alimentação da BD via colheita, BD é agregador), OAI-ORE

Propriedade Intelectual e Acordos de Submissão

- FAIR.R1.1, CTS.R2, CTS.R12

Para RDP-RC - Repositório de Dados da Pesquisa em Redes de Computadores:

- Termo de Submissão, em que o depositante declara que é detentor dos direitos sobre o conjunto de dados e autoriza a publicação dos dados de acordo com licença de acesso livre Creative Commons.

Para RDP-IPÊ - Rede IPÊ

Transferência legal da custódia dos dados ao repositório: Aceitação (fluxo), Verificação e Validação dos Dados Submetidos

- CTS.R7, CTS.R12

Para RDP-RC - Repositório de Dados da Pesquisa em Redes de Computadores

- Dados são aceitos, aprovados e validados de acordo com a política apresentada na tabela 3D
- Formatos são validados automaticamente pelo software

Para RDP-IPÊ - Rede IPÊ

- Dados são aceitos, aprovados e validados pela equipe de Produção da rede Ipê, de acordo com a política apresentada na tabela 3C
- Formatos são validados automaticamente pelo software

Transferência legal da custódia dos dados ao repositório: Confidencialidade e autenticação da informação na transferência Consumidor-Repositório

- CTS.R7, CTS.R12

- Os usuários são cadastrados e autenticados pelo sistema.
- Somente usuários com papéis autorizados podem realizar as funções definidas pela política apresentada na tabela 3D
- Os dados são transferidos do Repositório ao consumidor via protocolo seguro HTTPS

Transferência legal da custódia dos dados ao repositório: Submissão por Máquinas e/ou Batch

- Dados alimentados pelo protocolo Sword

7 Acesso e Uso dos Conjuntos de Dados

Essa seção identifica as informações necessárias para descoberta e uso dos dados. Inclui:

- **Identificação dos Dados e Recursos Relacionados:** Identificadores únicos, globais e persistentes para conjuntos de dados e arquivos, considerando versionamento e usando serviços, protocolos e padrões aceitos pela comunidade (como os DOI e Handle System e os serviços DataCite , etc). Identificadores únicos, globais e persistentes para recursos relacionados aos conjuntos de dados e usando serviços, protocolos e padrões aceitos pela comunidade (como os ORCID , para pesquisadores)
- **Informações ao consumidor para busca e uso:** metadados descritivos, metadados de proveniência, citação, direitos/licenças de uso
- **Acesso aos Dados e Metadados:** Acesso aos dados e aos metadados via identificador e protocolo aberto. Restrições de acesso (nível de arquivo, conjunto de dados, grupos, unidades, etc., a determinados grupos), embargos, acesso mediante registro do usuário, registro de solicitações de uso e relacionamento com usuários (Livro de Visitas / Guestbook)
- **Transferência dos dados ao consumidor:** Confidencialidade e autenticação da informação na transferência . Gerenciamento e autenticação de usuários envolvidos com acesso. Entrega de dados e dos metadados ao consumidor (pessoa ou sistema) nas representações (formatos e estruturas) adequadas, acesso via API. Protocolos de colheita de metadados (OAI-PMH, Z39-50), acesso via API

Quadro 7 - Acesso e Uso dos Conjuntos de Dados

Identificador Único, Global e Persistente

- FAIR.F1, FAIR.A1, CTS.R13

Identificadores para:

- Conjuntos de Dados
- Arquivos
- Versões

Padrão

- Handle System

<p>Identificador Persistentes em Recursos Referenciados</p> <ul style="list-style-type: none"> ● CTS.R13
<p>Identificadores para:</p> <ul style="list-style-type: none"> ● Autoridades (ORCID)
<p>Informações aos consumidores para busca e uso</p> <ul style="list-style-type: none"> ● FAIR.R1.1, CTS.R2
<ul style="list-style-type: none"> ● Informar ao consumidor: <ul style="list-style-type: none"> ■ Metadados Descritivos e Proveniência ■ Citação ■ Direitos/Termos de Uso
<p>Transferência dos dados e metadados ao consumidor: Confidencialidade e autenticação da informação na transferência Repositório-Consumidor. Entrega de dados e dos metadados ao consumidor</p> <ul style="list-style-type: none"> ● FAIR.A1, FAIR.A1.2, FAIR.I1
<ul style="list-style-type: none"> ● Os usuários são cadastrados e autenticados pelo sistema. ● Somente usuários com papéis autorizados podem realizar as funções definidas pela política apresentada na tabela 3D ● Os dados são transferidos do Repositório ao consumidor via protocolo seguro HTTPS
<p>Transferência dos dados e metadados ao consumidor: transferência por máquinas</p>
<ul style="list-style-type: none"> ● Transferência de Metadados via OAI-PMH

8 Armazenamento e Preservação

Recursos para gerenciar o armazenamento dos dados, Segurança da Informação <ul style="list-style-type: none">• Cópias de segurança, dispositivos de armazenamento• CTS.R9, CTS.R16
<ul style="list-style-type: none">• Cópias de Segurança:• Exportação de Pacotes

9 Implantação em Dataverse

Dataverse	
Identificador único e persistente	Handle - Prefixo:
Citação Dataset	Autores, Ano, Título do Conjunto de Dados, DOI ou Handle, Repositório, Versão
Citação Arquivo	Autores, Ano, Título do Conjunto de Dados, DOI ou Handle, Repositório, Versão, UNF ou MD5
Autenticação - Identidade	Usuários cadastrados no Dataverse
Site	https://dadosabertos.rnp.br

Unidades

- Dataverse Raiz
 - Dataverse Rede Ipê
 - Dataset Medições AAAA (um dataset para cada ano)
 - Dataverse GT RDP BRASIL
 - Dataset Acesso aberto a dados de pesquisa no Brasil : repositórios brasileiros de dados de pesquisa
 - Dataset Acesso aberto a dados de pesquisa no Brasil : práticas e percepções dos pesquisadores
 - Dataverse Pesquisas em Redes de Computadores
 - Dataverse Grupo de Pesquisa X
 - Dataset Dados de Pesquisa X
 - Dataset Dados de Pesquisa X

Dataverse Raiz

Nome	Repositório de dados abertos - versão beta
Identificador	dvrnp https://dadosabertos.rnp.br/dataverse/dvrnp
Categoria	Sem categoria
Grupos e Papéis	Grupo: Dataverse Admin Papel: admin
Metadados Descritivos	Metadados de Citação de Dataverse (compatível com DDI Codebook, Datacite, Dublin Core)
Dataverses e Datasets	<ul style="list-style-type: none">● Dataverse Pesquisas em Redes de Computadores● Dataverse Rede Ipê● Dataverse GT-RDP

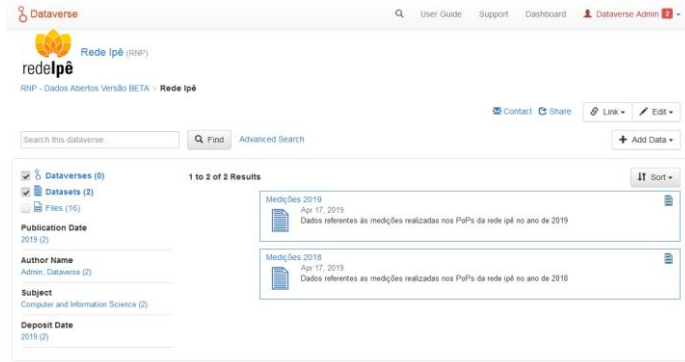
Apresentação

The screenshot shows the Dataverse website interface. At the top, there's a navigation bar with 'Dataverse', 'User Guide', 'Support', 'Dashboard', and 'Dataverse Admin'. Below that, the RNP logo and 'RNP - Dados Abertos Versão BETA (RNP)' are visible. A 'Metrics' section shows '16 Downloads'. There are three featured datasets: 'GT RDP BRASIL', 'redelipe Rede Ipê', and 'Pesquisas em Redes de Computadores'. A search bar is present with 'Find' and 'Advanced Search' buttons. On the left, there are filters for 'Dataverses (3)', 'Datasets (4)', and 'Files (17)'. The main content area shows '1 to 7 of 7 Results' with three dataset entries: 'Pesquisas em Redes de Computadores (rnp.br)', 'GT RDP BRASIL (GT RDP Brasil)', and 'Rede Ipê (RNP)'. Each entry includes a date of 'Mar 13, 2019' and a small icon.

Dataverse Rede IPÊ

Nome	Rede Ipê
Identificador	Dvipe https://dadosabertos.rnp.br/dataverse/dvipe
Categoria	Organização ou Instituição
Grupos e Papéis	<ul style="list-style-type: none"> ● CuradorRedeIPE, papel Curator (adicionar, editar, aprovar) ● ColaboradorRedeIPE, papel Contributor (adicionar, editar) ● MembroRedeIPÊ, papel Membro (ter acesso a datasets não aprovados e a arquivos de acesso restrito)
Metadados Descritivos	<ul style="list-style-type: none"> ● Metadados de Citação de Dataverse
Dataverses e Datasets	<ul style="list-style-type: none"> ● Datasets <ul style="list-style-type: none"> ○ Medições 2019 ○ Medições 2018

Apresentação



Datasets Medições Rede IPÊ (Medições AAAA)

Metadados Mínimos	<p>Título: Medições 2018</p> <p>Autor: RNP</p> <p>Contato: GT-RDP Brasil</p> <p>Descrição: Dados referentes às medições realizadas nos PoPs da rede ipê no ano de 2019</p> <p>Assunto: Computer and Information Science</p> <p>Palavra Chave: Network performance evaluation</p>
Termos	Licença CCO
Permissões ao Dataset	<p>Grupo/Papel</p> <ul style="list-style-type: none"> ● CuradorRedeIPE, papel Curator (adicionar, editar, aprovar) ● ColaboradorRedeIPE, papel Contributor (adicionar, editar) ● MembroRedeIPÊ, papel Membro (ter acesso a datasets não aprovados e a arquivos de acesso restrito)
Permissões aos Arquivos	Acesso a arquivo _____ restrito a grupo _____ -
Arquivos	AA-MMMM.tab

Dataverse GT RDP BRASIL	
Nome	GT RDP BRASIL
Identificador	gtrdpbrasil https://dadosabertos.rnp.br/dataverse/gtrdpbrasil
Categoria	Grupo de pesquisa
Grupos e Papéis	<ul style="list-style-type: none"> ● CuradorRDP, papel Curador (criar e aprovar datasets no grupo) ● ColaboradorRDP, papel Contributor (criar datasets no grupo) ● MembroRDP, papel Membro (ter acesso a datasets não aprovados e a arquivos de acesso restrito)
Metadados Descritivos	Metadados de Citação de Dataverse
Dataverses e Datasets	

Dataverse Pesquisas em Redes de Computadores	
Nome	Pesquisa em Redes de Computadores
Identificador	dvredes
Categoria	Grupo de Pesquisa
Grupos e Papéis	<ul style="list-style-type: none"> ● CuradorRPC (Cria, publica, dataverses e datasets de grupos de pesquisa a pedido do pesquisador) ● CuradorDeGrupoX, papel Curador (criar e aprovar datasets no grupo) ● ColaboradorDeGrupoX, papel Contributor (criar datasets no grupo) ● MembroDeGrupoX, papel Membro (ter acesso a datasets não aprovados e a arquivos de acesso restrito)

Papel atribuído ao criador do dataset ou dataset	<ul style="list-style-type: none"> • Criador do Dataset, possui o papel de Curador, estando habilitado a criar grupos e autorizações no escopo dataset criado • Criador do Dataverse possui o papel de Curador, estando habilitado a criar grupos, autorizações, (dataverses), datasets, no escopo do dataverse criado
Metadados Descritivos	Metadados de Citação de Dataverse
Dataverses e Datasets	<ul style="list-style-type: none"> • Dataverse (Grupo de Pesquisa X) • Dataset (Dados Pesquisa X)

Dataverse Grupo de Pesquisa X	
Nome	Grupo de Pesquisa
Identificador	dvg...
Categoria	Grupo de Pesquisa
Grupos e Papéis	
Metadados Descritivos	Citação de Dataverse
Dataverses e Datasets	<ul style="list-style-type: none"> • Dataset (Dados Pesquisa X)

Dataset Dados de Pesquisa X	
Metadados Mínimos	Título: Autor: Contato: Descrição: Assunto:
Termos	Licença CCO
Permissões ao Dataset	Grupo/Papel

Permissões aos Arquivos	Acesso a arquivo _____ restrito a grupo _____ -
Arquivos	

Referências

- [i] CONARQ. DIRETRIZES PARA A IMPLEMENTAÇÃO DE REPOSITÓRIOS DIGITAIS CONFIÁVEIS DE DOCUMENTOS ARQUIVÍSTICOS, 2014
- [ii] OAIS - CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS. Reference model for an open archival information system (OAIS). CSDS 650.0-M-2. 2012
<https://public.ccsds.org/pubs/650x0m2.pdf>
- [iii] Wilkinson, M. D., et al.: The FAIR Guiding Principles for scientific data management and stewardship. Scientific data, vol. 3 (2016). <https://www.nature.com/articles/sdata201618>
- [iv] CTS - Core Trustworthy Data Repositories Requirements v01.
https://www.coretrustseal.org/wp-content/uploads/2017/01/Core_Trustworthy_Data_Repositories_Requirements_01_00.pdf
- [v] Princípios Fair - <https://www.force11.org/group/fairgroup/fairprinciples>
- [vi] Jantz, R., Giarlo, M.: Digital Preservation - Architecture and Technology for Trusted Digital Repositories. Dlib Magazine, vol. 11, n. 6 (2006)
<http://www.dlib.org/dlib/june05/jantz/06jantz.html>
- [vii] Wilkinson, M. D., et al.: The FAIR Guiding Principles for scientific data management and stewardship. Scientific data, vol. 3 (2016). <https://www.nature.com/articles/sdata201618>
- [viii] CTS - Core Trustworthy Data Repositories Requirements v01.
https://www.coretrustseal.org/wp-content/uploads/2017/01/Core_Trustworthy_Data_Repositories_Requirements_01_00.pdf
- [ix] CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS. Producer-Archive Interface Methodology Abstract Standard. CCSDS 651.0-M-1. 2004
<https://public.ccsds.org/Pubs/651x0m1.pdf>
- [x] UCPSR - What is a Codebook?
<https://www.icpsr.umich.edu/icpsrweb/content/shared/ICPSR/faqs/what-is-a-codebook.html>
- [xi] Groth, P., Moreau, L.: PROV-Overview - An Overview of the PROV Family of Documents. W3C Working Group, Note 30 (2013). <https://www.w3.org/TR/prov-overview>
- [xii] Esquema DataCite - https://schema.datacite.org/meta/kernel-4.1/doc/DataCite-MetadataKernel_v4.1.pdf

Anexo 1 bn – Dataset das Medição IPÊ

Representação Estrutural Atual

Dataset no formato atual (<https://www.rnp.br/sites/default/files/dadosbrutos012018.tgz>)

- dadosbrutos012018.tgz
- dadosbrutos012018.tar
 - o home
 - § dbdisp
 - pm
 - parciais
 - § 20180101
 - PoP-AC_20180101
 - PoP-AL_20180101
 - ...
 - PoP-TO_20180101
 - 20180102
 - ...
 - § 20180131

Arquivo: PoP-AC_20180101

PoP-AL=0:89.407:89.8915234375:129.040:2.66867035546057:89.527:89.5895:89.678
PoP-AM=0:94.013:102.79664893617:138.834:4.93088073681811:94.311:103.8355:103.935
PoP-AP=0:87.810:87.9407210300429:91.214:0.252539079126813:87.884:87.915:87.956
PoP-BA=0:80.695:84.2853636363637:115.503:6.72714469399621:80.728:80.7745:96.542
PoP-CE=0:72.725:73.0801216730038:109.833:2.39743267750283:72.790:72.835:72.904
PoP-DF=0:42.586:43.4859568345324:119.533:5.68225188984483:42.676:42.729:43.092
PoP-ES=0:72.242:88.8573261648745:149.805:15.1168797634211:72.305:93.376:111.354
PoP-GO=0:39.777:40.1768308270677:64.873:2.2319513903779:39.877:39.946:40.011
PoP-MA=0:86.600:87.0264414715719:136.555:3.13930439850233:86.684:86.748:86.807
PoP-MG=0:60.366:63.2865946969697:90.850:6.75745585509707:60.408:60.462:68.947
PoP-MS=0:34.190:34.8211964285714:80.080:3.40129966841992:34.273:34.337:34.411
PoP-MT=0:24.782:25.5254732824428:68.196:3.88022463569067:24.910:24.993:25.061
PoP-PA=0:69.669:70.6084308300396:125.758:5.68616988494608:69.719:69.765:69.815
PoP-PB=0:89.679:90.8077940074906:144.008:6.3531330976663:89.731:89.819:89.880
PoP-PE=0:85.360:85.9531565836299:119.590:3.56419905103471:85.401:85.441:85.482
PoP-PI=0:85.521:86.2212592592592:124.066:3.83085298035212:85.554:85.597:85.641
PoP-PR=0:53.165:53.6750176678445:125.430:4.64010174629778:53.231:53.273:53.320
PoP-RJ=0:62.128:62.4046582733813:79.253:1.28547615605372:62.168:62.214:62.347
PoP-RN=0:95.414:95.7556598639456:118.466:1.88436257677924:95.487:95.549:95.595
PoP-RO=0:6.705:6.90352142857143:7.565:0.0645076715014209:6.837:6.904:6.957
PoP-RR=0:103.962:112.184630901288:125.908:3.45361720435693:104.101:113.687:113.772
PoP-RS=0:68.135:69.1185087108014:144.573:6.25622049820223:68.174:68.215:68.252
PoP-SC=0:72.635:74.3214492753623:107.390:2.20740882920163:72.728:74.001:75.255
PoP-SE=0:85.848:88.6222713754646:139.339:6.26988355220217:85.920:85.968:97.868
PoP-SP=0:59.992:60.7480977443609:100.129:3.54665306454227:60.052:60.098:60.202
PoP-TO=0:54.666:55.230986440678:94.493:3.59291676744296:54.743:54.785:54.833