**Ciência Aberta no Brasil: Conquistas e Desafios**

Capítulo XX: **A contribuição da RNP para a Ciência Aberta no Brasil**

Autores: Leandro N. Ciuffo, Carolina H. Felicíssimo

**1. Sobre a RNP**

A Rede Nacional de Ensino e Pesquisa (RNP) é uma organização social sem fins lucrativos vinculada ao Ministério da Ciência, Tecnologia e Inovação (MCTI) e mantida por esse, em conjunto com os ministérios da Educação (MEC), das Comunicações (MCom), Cultura, Saúde (MS) e Defesa (MD).

Em 2019, a RNP completou 30 anos de pioneirismo no desenvolvimento da Internet no Brasil, sendo a instituição responsável pela rede de Internet acadêmica brasileira de alta velocidade, conectando mais de 800 instituições e beneficiando milhares de alunos, professores e pesquisadores. Atualmente o “Sistema RNP[[1]](#footnote-2)” inclui universidades, institutos educacionais e culturais, agências de pesquisa, hospitais de ensino, parques e polos tecnológicos.

A missão da RNP é “promover o uso inovador das TIC e redes avançadas”, e seu propósito visa “impulsionar a ciência e educação para todos”. Para cumprir a sua missão e apoiar a pesquisa e o desenvolvimento das Tecnologias da Informação e Comunicação (TIC) no Brasil, desde 2002 a RNP abre editais[[2]](#footnote-3) de projetos de P&D para impulsionar o desenvolvimento de soluções inovadoras para uso avançado da Internet, nas mais diversas áreas de aplicação, como por exemplo: saúde digital, educação à distância, cibersegurança, videocolaboração etc.

**2. Da transmissão à gestão de dados**

Além da gestão de projetos de P&D para o desenvolvimento de novos serviços inovadores sobre a Internet, a Diretoria de Pesquisa, Desenvolvimento de Inovação da RNP (DPDI) também se notabilizou pelo suporte à e-Ciência[[3]](#footnote-4), em especial por ações e projetos com a finalidade de apoiar grupos de pesquisa a movimentarem, de/para infraestruturas de pesquisa[[4]](#footnote-5), volumes massivos de dados pela Internet com alto desempenho e segurança.

Em dezembro de 2014 o então diretor de pesquisa e desenvolvimento da RNP, Michael Stanton, foi indicado pelo MCTI para ser membro do Conselho da *Research Data Alliance* (RDA[[5]](#footnote-6)), preenchendo uma vaga que havia sido destinada para representantes oriundos da América Latina. A RDA é uma organização internacional e não governamental, que promove ações para impulsionar o compartilhamento de dados como boa prática para o incentivo à pesquisa e à disseminação de conhecimento científico.

Durante seu mandato de três anos como conselheiro, a RNP foi exposta aos ideais e à visão da RDA, que previa um futuro em que pesquisadores seriam capazes de compartilhar abertamente seus dados de pesquisa e terem acesso à dados de pesquisa gerados por outros pesquisadores, para serem reutilizados entre diferentes disciplinas, formatos, tecnologias e países, possibilitando o enfrentamento dos grandes desafios da sociedade.

Essa exposição levou à reflexão de que a RNP poderia apoiar a e-Ciência não apenas na transmissão de dados, mas em várias fases de seu ciclo de vida. Desta forma, um novo tópico de interesse, voltado para investigar soluções para apoiar a gestão do ciclo de vida de dados de pesquisa, foi introduzido na DPDI, resultando no lançamento de um edital em 2017 para a seleção de um Grupo de Trabalho (GT) para atuar na identificação de práticas e ferramentas para acesso aberto a dados de pesquisa[[6]](#footnote-7).



Figura 1. Ciclo de vida do dado de pesquisa, tendo os repositórios de dados de pesquisa como um elemento fundamental para possibilitar o compartilhamento e o reuso de dados

**3. Grupo de Trabalho Rede de Dados de Pesquisa Brasileira (GT-RDP Brasil)**

O edital lançado em 2017 foi fruto de um Acordo de Cooperação Técnica (ACT) celebrado entre a RNP e o IBICT, que previa a promoção de atividades de P&D, iniciando-se com a investigação de práticas nacionais de Acesso Aberto à Dados de Pesquisa (AADP[[7]](#footnote-8)). Como a temática ainda era nova para a RNP, o ACT permitiu à DPDI contar com a experiência do IBICT para contribuir na elaboração do edital, participar do comitê de seleção e no acompanhamento técnico das entregas realizadas pelo Grupo de Trabalho.

O edital solicitava o cumprimento dos objetivos específicos:

* Identificação das atuais práticas de AADP em instituições brasileiras;
* Mapeamento de potenciais usuários nacionais de serviços de AADP e dos principais requisitos e funcionalidades para o atendimento dos mesmos;
* Levantamento comparativo dos serviços e soluções tecnológicas existentes para compartilhamento de dados, com identificação de lacunas e adaptações necessárias para atender aos usuários mapeados;
* Desenvolvimento de um protótipo de repositório de dados de pesquisa, baseado em uma das soluções tecnológicas levantadas, e com o potencial de atender aos usuários mapeados.

O GT selecionado, dentre 10 candidaturas recebidas, foi o “Rede de Dados de Pesquisa Brasileira” (GT-RDP Brasil), constituído por professores e bolsistas dos Centros de Documentação e Acervo Digital de Pesquisa (CEDAP) e de Processamento de Dados (CPD), ambos da Universidade Federal do Rio Grande do Sul (UFRGS), além do Grupo de Pesquisa em Gerenciamento de Informações do Centro de Ciências Computacionais (C3) da Universidade Federal do Rio Grande (FURG). Atualmente, o GT-RDP Brasil mantém o site: <https://www.ufrgs.br/rdpbrasil/>, onde conta parte de sua história.

Uma das atividades iniciais do GT-RDP Brasil foi a execução de uma pesquisa para o mapeamento de práticas para abertura de dados de pesquisa no Brasil. A pesquisa foi conduzida através da disponibilização de um questionário *online* a ser respondido com 8 questões demográficas e 19 sobre a práticas e percepções dos pesquisadores sobre compartilhamento de dados de pesquisa. Com a ajuda do IBICT e do CNPq, o questionário foi enviado para 71.712 e-mails, organizados de acordo com os seguintes perfis:

* 48.335 líderes + 27.869 vice-líderes de grupos de pesquisa registrados no Diretorio de Grupos de pesquisa do CNPq;
* 4.677 e-mails institucionais de coordenadores de programas de pós-graduação;
* 249 e-mails de coordenadores de Institutos Nacionais de Ciência e Tecnologia (INCTs);

O questionário ficou aberto pelo período de um mês para a coleta de respostas, e atingiu o impressionante número de 4.703 respondentes. O relatório completo, com mais detalhes sobre a metodologia e com as análises e conclusões sobre os dados levantados, está disponível em: <http://hdl.handle.net/10183/185195> . Os dados coletados pela pesquisa estão disponíveis em: <https://hdl.handle.net/20.500.11959/5> .

Para o estudo das soluções tecnológicas existentes, o grupo estabeleceu uma extensa lista de critérios e requisitos para a aplicação de uma análise comparativa entre ferramentas, com ênfase maior na comparação entre as soluções específicas para repositórios digitais: CKan, DSpace e Dataverse. O relatório com a análise comparativa está disponível em: <http://hdl.handle.net/10183/185126>

Adicionalmente, ao longo de dois anos de projeto (2018 – 2020), o GT-RDP Brasil executou ações de P&D, tanto nas vertentes conceitual, quanto tecnológica, incluindo a execução de provas de conceitos (PoCs), resultando em diversas publicações importantes e orientadoras para a comunidade brasileira que, na época, começava a experimentar repositórios institucionais de dados de pesquisa.

Algumas das publicações mais relevantes do GT-RDP Brasil, além dois relatórios citados acima, são destacadas a seguir:

* Repositórios brasileiros que disponibilizam dados de pesquisa (2018) <http://hdl.handle.net/10183/185138>
* Levantamento sobre a estrutura e funcionamento de grupos nacionais criados no RDA. (2019) <http://hdl.handle.net/20.500.11959/1257>
* Políticas para repositórios de dados de pesquisa. (2019) <http://hdl.handle.net/20.500.11959/1263>
* Planejamento para implantação de comunidade produtora de dados para o repositório Rede de Dados de Pesquisa. (2019) <http://hdl.handle.net/20.500.11959/1204>

**4. Federação piloto de repositórios de dados de pesquisa**

Em outubro de 2018, o Brasil, representado pelo Ministério da Transparência e Controladoria-Geral da União (CGU), publicou seu “4º Plano de Ação Nacional em Governo Aberto[[8]](#footnote-9)”, composto por 11 compromissos, os quais foram cocriados com o envolvimento de representantes de 88 instituições da sociedade civil e da administração pública, nas esferas federal, estadual e municipal. Tal Plano faz parte da Parceria para Governo Aberto (do inglês, *Open Government Partnership* - OGP), uma iniciativa internacional composta por 75 nações e que tem como objetivo difundir e incentivar globalmente práticas governamentais relacionadas à transparência, acesso à informação pública e participação social.

Um dos compromissos do plano, proposto e liderado pela Embrapa, tinha como objetivo “Estabelecer mecanismos de governança de dados científicos para o avanço da Ciência Aberta no Brasil[[9]](#footnote-10)”. Por sua vez, o compromisso foi organizado em um conjunto de nove metas, que foram construídos em oficinas de cocriação que contaram com a participação de órgãos públicos e da sociedade civil, entre eles a RNP, que foi convidada pela Embrapa, no papel de coordenadora do compromisso.

Dentro desse compromisso, coube à RNP, em parceria com o IBICT e com a equipe do GT-RDP Brasil, executar uma meta que visava a “implantação de infraestrutura federada piloto de repositórios de dados de pesquisa”. Basicamente, o trabalho consistiu em apoiar a implantação de um conjunto de repositórios de dados de pesquisa em diferentes instituições, ainda como protótipos, e executar uma prova de conceito de indexá-los na busca do portal Oasisbr[[10]](#footnote-11) operado pelo IBICT. A meta foi concluída com êxito em julho de 2020 e, desde então, o Oasisbr passou a incorporar a funcionalidade de busca por conjuntos de dados em repositórios digitais de dados indexados pela ferramenta.

As lições aprendidas e os subprodutos desenvolvidos durante a execução das atividades para o atingimento da meta foram todos documentados e estão à disposição dos pesquisadores:

* Manual de Instalação do Dataverse (versão 1.0 de 2019, versão 1.4 de 2021) <http://hdl.handle.net/20.500.11959/1202>
* Dataverse documentação: Identificadores persistentes. (2020). <http://hdl.handle.net/20.500.11959/1258>
* Dataverse documentação: Licenças de uso. (2020) <http://hdl.handle.net/20.500.11959/1260>
* Dataverse documentação: Ativação do DOI (2020) <http://hdl.handle.net/20.500.11959/1259>
* Pacote de tradução do Dataverse para Português BR.
<https://github.com/RNP-dadosabertos/dataverse-language-packs> <https://github.com/GlobalDataverseCommunityConsortium/dataverse-language-packs/tree/develop/pt_BR>
* Guia de instalação Shibboleth para autenticação institucional no Dataverse
[https://www.rd-alliance.org/group/rda-brazil/outcomes/autenticação-federada-no-dataverse](https://www.rd-alliance.org/group/rda-brazil/outcomes/autentica%C3%A7%C3%A3o-federada-no-dataverse)

**5. Recomendações para uma política nacional para a ciência aberta**

Em 2018 o Ministério da Ciência, Tecnologia e Inovação, por meio da Portaria nº 6.072 de 23/11/2018, instituiu um Grupo de Trabalho para discutir uma proposta de Política Nacional para Ciência Aberta[[11]](#footnote-12). Tal grupo foi originalmente composto por representantes de cinco diferentes secretarias e diretorias da estrutura do MCTI e quatro instituições vinculadas: IBICT, CNPq, FINEP e RNP.

O resultado final gerado pelo grupo de trabalho foi a elaboração de uma proposta de minuta de decreto, a ser apresentada à presidência da república, com a finalidade de instituir uma Política Nacional de Ciência Aberta que dispunha sobre o acesso aberto às publicações científicas e aos dados de pesquisas científicas custeadas com recursos públicos federais. Entretanto, vale ressaltar que um dos principais princípios da Política proposta foi a observância do preceito de que o ciclo de pesquisa científica e seus resultados deveriam ser “tão abertos quanto possível e tão fechados quanto necessário”, adotando-se o acesso aberto como regra geral e o sigilo como exceção.

Nesse processo de construção, coube à RNP fornecer insumos para que fosse inserido no texto do decreto, um conjunto de recomendações acerca da infraestrutura de TIC mínima necessária que deveria ser observada pelas Instituições Científicas, Tecnológicas e de Inovação (ICTs) para permitir o armazenamento, compartilhamento e acesso às publicações científicas e aos dados de pesquisas científicas.

Até onde é de conhecimento dos autores deste capítulo, a proposta de decreto chegou a ser apresentada à diretores de unidades de pesquisa vinculadas ao MCTI em abril de 2021, mas não houveram, desde então, avanços significativos ou consulta pública visando o encaminhamento da proposta para a aprovação da presidência. De qualquer forma, vale a pena compartilhar aqui as recomendações feitas pela RNP como contribuição ao Grupo de Trabalho:

**Recomendações para as ICTs:**

* Instituições de pesquisa devem facilitar o acesso à repositórios de dados ou outras plataformas que permitam a seus pesquisadores o compartilhamento e o armazenamento de dados de pesquisa.
	+ Pesquisadores deverão armazenar dados de pesquisa preferencialmente em repositórios de dados institucionais e certificados, observando e respeitando as diretrizes de sua instituição.
* No caso de uso de repositórios de dados institucionais, existentes ou novos, estes deverão perseguir a implementação de processos de certificação.
	+ Repositórios de dados devem garantir que os conjuntos de dados (datasets) sejam identificáveis através do uso de identificadores persistentes, bem como garantir que os dados sejam descritos adequadamente para facilitar a busca e o uso dos dados.
	+ Repositórios de dados devem ser capazes de representar tanto dados coletados durante a pesquisa como dados a serem preservados após o seu término.
	+ Repositórios de dados devem ser capazes de permitir o gerenciamento de versões de dados decorrentes de transformações que ocorrem no ciclo de vida do dado.
	+ Repositórios de dados deverão ser implantados em infraestrutura de TIC confiável e estável, preferencialmente operada pelo núcleo de TI da instituição e com políticas de backup definidas.
	+ A infraestrutura de hardware e software deve ser adequada para permitir o armazenamento, o compartilhamento e a descoberta do dado, fazendo uso de padrões e modelos de referência conhecidos e relevantes para o propósito do repositório.
	+ É recomendável que os repositórios de dados possuam um plano de gestão de riscos e procedimentos para a restauração do repositório em caso de ataques maliciosos, falha humana, falha na infraestrutura de TI ou desastres.
	+ Repositórios de dados devem ser eficientes no uso da infraestrutura de TIC. Nem todo dado coletado é passível de ser armazenado. Os custos e os benefícios de armazenagem do dado, incluindo o período de tempo em que o dado precisará ser preservado, deverão ser considerados no processo do planejamento da gestão do dado.
		- Para a construção e operação de repositórios institucionais de universidades, bem como para o estabelecimento de políticas de uso e processos de suporte, é recomendado o estabelecimento de parcerias entre o núcleo de TI, biblioteca e departamentos de ciência a informação e ciência da computação, ou equivalentes.
		- Colaborações entre instituições para o compartilhamento de experiências e redução de custos operacionais são encorajadas.
		- O uso de infraestrutura computacional terceirizada, como por exemplo infraestrutura de armazenamento fornecida por provedores de nuvens públicas, pode ajudar na redução de custos operacionais e, portanto, devem ser consideradas.
* A adoção de modelos de federação de repositórios de dados pode levar à economia de escala e ao compartilhamento de custos entre instituições.

**6. Incubação de repositórios digitais institucionais de dados de pesquisa**

Em 2021, a parceria existente entre RNP e IBICT contou com a inclusão do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) por meio de um novo Acordo de Cooperação Técnico Científico, assinado pelas três instituições para a promoção da Ciência Aberta no Brasil.

Em linhas gerais, o objetivo do Acordo foi a união de esforços para estruturar uma federação nacional de repositórios de dados de pesquisa e implementar ações de transferência de conhecimentos relacionadas à criação de repositórios de dados de pesquisa. Desta forma, uma das primeiras ações foi o lançamento de um edital[[12]](#footnote-13) para a incubação de repositórios institucionais de dados de pesquisa, com os seguintes resultados esperados:

1. Selecionar até 4 ICTs interessadas em implantar repositórios de dados de pesquisa em âmbito institucional.
2. Apoiar as ICTs selecionadas na implantação de seus repositórios institucionais de dados de pesquisa por meio de ações de capacitação e transferência de conhecimento, com o compromisso das ICTs selecionadas de manter e operar os repositórios de dados gerados após um período de incubação de 9 meses.
3. Ampliar o conjunto de instituições participantes do Consórcio Nacional para Ciência Aberta (CoNCienciA[[13]](#footnote-14)), com a inclusão das ICTs selecionadas, que poderiam ter identificadores persistentes DOI (*Digital Object Identifier*) associados a seus conjuntos de dados, com o custo inicial da emissão de DOIs patrocinado pelo CNPq.
4. Estruturar de uma Rede brasileira de repositórios institucionais de dados de pesquisa, com a participação dos repositórios de dados das ICTs selecionadas por meio do edital.

O cronograma previsto pelo edital estipulava um prazo de nove meses para atividades de incubação (i.e., transferência de conhecimento e testes com repositório de dados em ambiente de experimentação) e três meses finais dedicados à migração do repositório para a infraestrutura própria de TI da ICT selecionada e ao estabelecimento de políticas institucionais para a operação e governança do repositório. Adicionalmente, ao final do nono mês, o progresso dos repositórios incubados foi avaliado pela equipe do CNPq para confirmar o convite às ICTs para ingressarem no Consórcio CoNCienciA e, portanto, receberem identificadores persistentes para datasets do tipo DOI, emitidos pela DataCite[[14]](#footnote-15), e fornecidos pelo CNPq sem custo às instituições, porém limitado em 10.000 (dez mil) registros por ano.

De forma mais específica, a fase de incubação consistiu em atividades de mentoria e capacitação nas vertentes conceitual e tecnológica, realizadas em reuniões semanais com as equipes das ICTs e com muitas atividades práticas, que foram conduzidas pela equipe do GT-RDP, que teve seu financiamento renovado pela RNP para a execução dessas atividades em apoio ao edital.

Devido à natureza intensiva da mentoria, ao limite da equipe acadêmica e ao limite orçamentário do CNPq para o fornecimento de DOIs aos repositórios selecionados, apenas quatro instituições puderam ser incubadas, em um total de 14 candidaturas. As ICTs selecionadas no edital foram: Centro Brasileiro de pesquisas Físicas (CBPF), Universidade Federal do Ceará (UFC), Universidade Federal de Goiás (UFG) e Universidade Federal Rural do Amazonas (UFRA).

Os resultados do edital foram muitos satisfatórios e o modelo foi considerado um sucesso pelos partícipes. As ICTs incubadas avaliaram muito positivamente a experiência e a mentoria recebida. Posteriormente, a RNP buscou viabilizar recursos para o lançamento de um novo edital, no intuito de selecionar mais dez ICTs para incubação. Entretanto, foi constatado que o modelo, baseado em mentorias individuais e customizadas para cada instituição, tornou-se inviável para atender a um número maior de ICTs e escalar as atividades necessárias para criação de uma rede nacional de repositórios institucionais digitais de dados de pesquisa. Assim, foi necessário repensar o modelo e buscar meios de impulsionar a criação da rede de colaboração desses repositórios de forma escalável e sustentável.

**7. Um modelo de rede de colaboração para Acesso Aberto a Dados de Pesquisa**

No segundo semestre de 2022, a autointitulada “REDEDADOSABERTOS”, uma rede de colaboração para Acesso Aberto a Dados de Pesquisa constituída pelas ICTs que tiveram seus repositórios incubados (CBPF, UFC, UFG e UFRA), contou com o aceite do convite para a representatividade também das instituições fundadoras do Consórcio CoNCienciA (CNPq, IBICT, Embrapa e Fiocruz), além da própria RNP e da UFRGS, representada pelo GT-RDP.

Na prática, foi estabelecida uma lista de discussão e uma agenda quinzenal de encontros regulares para a troca de experiências relacionadas à implantação, configuração e operação de repositórios de dados de pesquisa baseados no software Dataverse. Tomou-se os requisitos do Core Trust Seal[[15]](#footnote-16) como guia do que precisava ser alcançado para que os repositórios recém incubados ganhassem maturidade e sustentabilidade como repositórios digitais confiáveis. Discussões sobre desafios e conquistas com pesquisadores, representantes institucionais, elaboração e publicação de políticas institucionais de dados, além de várias outras temáticas, fizeram parte desses encontros quinzenais.

O sucesso do edital para a incubação de repositórios de dados de pesquisa, incluindo as ações de mentoria e os encontros para a troca de experiências, fez com que novas ICTs procurassem a RNP para receberem orientação e suporte na implantação de seus repositórios de dados de pesquisa. Entretanto, a falta de capacidade de atendimento para atender a diversas ICT e a falta de regras claras para o acolhimento de novas ICTs na REDEDADOSABERTOS, evidenciou a necessidade de discutir novos modelos para impulsionar a criação e sustentação de repositórios digitais de dados de pesquisa no Brasil.

Isso motivou o planejamento das ações da RNP para a REDEDADOSABERTOS em 2023, que teve como diretriz as seguintes perguntas norteadoras:

* Como escalar as atividades de mentoria e capacitação para um modelo não individualizado por instituição?
* Como criar uma rede de colaboração escalável, a qual não dependa da RNP, para a troca de experiências sobre repositórios digitais de dados de pesquisa?
* Como criar um modelo de governança que englobe o Consórcio CoNCienciA, capitaneado pelo CNPq, e que não compita com outras iniciativas semelhantes existentes, como a Rede Brasileira de Repositórios Digitais (RBRD), coordenada pelo IBICT?

Dada essas questões, o ano de 2023 foi marcado por estudos de modelos de governança praticados por redes colaborativas de repositórios digitais existentes em outros países (como, por exemplo, o consórcio europeu CESDDA[[16]](#footnote-17)) e no Brasil. As representantes regionais Rede Brasileira de Repositórios Digitais (RBRD) foram convidadas para os encontros quinzenais da REDEDADOSABERTOS para compartilharem lições aprendidas, desafios vencidos, expectativas e requisitos para a proposição de um modelo de governança para uma rede brasileira de repositórios institucionais de dados de pesquisa.

No total, nove encontros foram realizados ao longo de 2023, sendo que os dois últimos foram dedicados à construção colaborativa de uma proposta de modelo governança. Para esses dois encontros, foi adotada a metodologia de *Design Thinking* para a dinâmica. Representantes de todas as ICTs foram divididos em três grupos. Cada grupo tinha um quadro para serem adicionados *post-it* com ideias e sugestões para as seguintes questões:

* **Sustentabilidade**: Quais ações devem ser priorizadas para as ICTs sustentarem institucionalmente seus repositórios de dados?
* **Maturidade**: Quais ações devem ser priorizadas para que os repositórios de dados institucionais ganhem maturidade em suas respectivas ICTs?
* **Escala**: Quais ações devem ser priorizadas para conseguirmos escalar a rede nacional de repositório de dados abertos com novas ICTs integrantes?

Depois dos dois encontros conjunto, as contribuições foram consolidadas e organizadas em cinco eixos: (i) Transformação Digital das Bibliotecas; (ii) Gestão e Governança; (iii) Capacitação e Treinamentos; (iv) Financiamento e Parcerias e (v) Divulgação e Eventos.

**8. Conclusões**

O movimento da Ciência Aberta e, em especial, o movimento de Acesso Aberto à Dados de Pesquisa é uma consequência natural do desenvolvimento da Internet e da ciência orientada a dados. É possível observar que, gradativamente, a prática de compartilhamento de dados de pesquisa vem sendo estimulada por agências de fomento, além de ser incluída em políticas públicas de alguns países como também recomendada por organizações internacionais multilaterais como a OCDE e UNESCO.

Existe um reconhecimento generalizado de que repositórios de dados de pesquisa são parte essencial da infraestrutura habilitadora para a prática da ciência aberta e compartilhamento do conhecimento universal. É importante ressaltar que a disponibilização de dados de pesquisa em repositórios de acesso aberto respeita os devidos direitos autorais, períodos de embargo, questões de confidencialidade e privacidade de acordo com especificidades das áreas científicas e de regras de financiamento. Apesar dessa ressalva, ainda são observadas resistências, pois, trata-se também de um tema sensível, ainda em ampla discussão, e que lida com questões de soberania nacional na realização de pesquisa científica e na produção de conhecimento.

Apesar disso, mesmo com políticas tímidas que suportem a Ciência Aberta no Brasil, algumas ICTs já percebem seus benefícios e buscam se adequar às novas práticas mundiais de compartilhamento e gestão de dados de pesquisa. Tal adequação requer investimentos em infraestrutura tecnológica, implantação de novos processos e políticas, além de ações de capacitação continuadas, voltadas, principalmente, para gestores, profissionais de ciência da informação, tecnologia da informação e pesquisadores.

Na ausência de um organismo central que oferte um Serviço Nacional de Dados, a exemplo do que ocorre em países com a Austrália e a Holanda, faz-se necessário alguma orquestração com instituições nacionais chaves, tais como a RNP, IBICT e CNPq. Em nível nacional, é necessário fornecer orientações e capacitações, visando a transformação digital das bibliotecas institucionais, de modo que também possam atuar com a custódia de dados digitais, incluindo aqueles que são coletados e analisados durante todo o processo de produção e publicação de pesquisas científicas.

A tabela abaixo sumariza uma proposta de *roadmap* para a evolução de uma iniciativa para suportar o compartilhamento de dados de pesquisa a nível nacional. Entendemos que algumas iniciativas existentes no Brasil, as quais trabalhamos juntos, se posicionam entre os níveis de maturidade 1 e 2 do modelo.

Tabela 1. Proposta de modelo evolucionário para uma iniciativa nacional de repositórios institucionais de dados de pesquisa. A leitura deve ser feita de baixo para cima.

|  |  |  |
| --- | --- | --- |
| Nível evolutivo de maturidade | Propósito | Serviços |
| 4 | **Serviço Nacional de Dados de Pesquisa** | Prover serviço de armazenamento e gestão de dados de pesquisa, com acreditação para armazenar dados de projetos de pesquisa nacionais e estruturantes.Atendimento à ICTs que ainda não possuem seus próprios repositórios de dados. | Serviço de armazenamento de dados (armazenamento de longo prazo).Repositório como serviço. |
| 3 | **Federação de Repositórios de Dados de Pesquisa confiáveis** | Coordenar uma federação de repositórios confiáveis, com sustentação de longo prazo. Repositórios com acreditação para armazenar dados de projetos de pesquisa. | Serviço de certificação de repositórios.Consultoria para certificação. |
| 2 | **Rede de Repositórios de Dados de Pesquisa** | Coordenar uma Rede de Repositórios com objetos globalmente identificáveis por identificadores persistentes. Engloba instituições com repositórios implantados e em produção. | Serviço de Identificador de Objeto Persistente dos dados. (DOI).Serviço de busca federada envolvendo os repositórios membros |
| 1 | **Rede de Colaboração de Repositórios de Dados de Pesquisa** | Compartilhamento de conhecimento e melhores práticas para a administração de repositórios de dados de pesquisa. Engloba instituições com repositório em implantação. | *Templates* e modelos para auxiliar a construção de repositórios de dados.Organização de workshops e encontros para a troca de experiências.Mentorias e capacitações sobre instalação, configuração e administração de repositórios. |
| 0 | **Iniciativas isoladas de repositórios de dados de pesquisa em algumas ICTs pioneiras.** | Implantar repositório de dados de pesquisa de uma ICT. | Serviços de TI da própria ICT.Esforço de aprendizado individualizado de bibliotecários, técnicos de TI e pesquisadores. |

Para concluir, ainda há muito a ser feito, principalmente, nas pontas da cadeia do ciclo de vida dos dados de pesquisa, i.e., com os pesquisadores, os quais são os responsáveis por coletar, gerar e (re)usar dados de pesquisa. Convencê-los a compartilhar seus dados de pesquisa pode ser um desafio. É importante destacar os benefícios do compartilhamento de dados, como a possibilidade de aumentar a visibilidade e o impacto da pesquisa, bem como a oportunidade de colaborar com outros pesquisadores, ou, simplesmente, organizar a gestão dos dados que são produzidos e consumidos pelo grupo de pesquisa. Oferecer crédito e reconhecimento aos pesquisadores que compartilham seus dados também pode ser uma forma eficaz de incentivar o compartilhamento. A ideia de que é necessário revisar o sistema global de métricas e indicadores de produção científica de forma a também reconhecer os pesquisadores pela produção de dados (e códigos-fonte) já é consenso nos fóruns de discussão sobre Ciência Aberta. O ano de 2019 nos mostrou como a abertura da ciência é promissora e acelera o conhecimento humano como um todo. Vencemos uma pandemia em tempo recorde. Tivemos vacina com um esforço global em prol da Ciência Aberta para salvar vidas. Então, por que retroceder?

1. Sistema RNP: https://www.rnp.br/sistema-rnp [↑](#footnote-ref-2)
2. Editais: https://www.rnp.br/inovacao/editais [↑](#footnote-ref-3)
3. e-Ciência: Ciência realizada de forma colaborativa por meio do uso intensivo das Tecnologias da Informação e Comunicação, especialmente computação de alto desempenho, transmissão, análise e armazenamento de grande volume de dados. ​ [↑](#footnote-ref-4)
4. Infraestrutura de Pesquisa: são instalações físicas (equipamentos, instrumentos) ou digitais (arquivos, repositórios, bases de dados) que fornecem à comunidade científica insumos e serviços para realizar atividades de pesquisa e desenvolvimento experimental e fomentar a inovação. [↑](#footnote-ref-5)
5. RDA: https://rd-alliance.org/ [↑](#footnote-ref-6)
6. RNP e IBICT lançam chamada para promover o acesso aberto a dados de pesquisa. 15 Mar 2017. Disponível em: <https://www.rnp.br/noticias/rnp-e-ibict-lancam-chamada-para-promover-o-acesso-aberto-dados-de-pesquisa> [↑](#footnote-ref-7)
7. O Acesso Aberto aos Dados de Pesquisa remete à disponibilização de dados de pesquisa em plataformas digitais de livre acesso, respeitando os direitos autorais, períodos de embargo, questões de confidencialidade, privacidade, especificidades das áreas científicas, entre outros aspectos. [↑](#footnote-ref-8)
8. https://www.opengovpartnership.org/wp-content/uploads/2018/10/Brazil\_Action-Plan\_2018-2020\_POR.pdf [↑](#footnote-ref-9)
9. Mais informações sobre o Compromisso nº 3 do 4º Plano de Ação Nacional em Governo Aberto: http://url.rnp.br/?OGP-Brasil [↑](#footnote-ref-10)
10. Portal Brasileiro de Publicações e Dados Científicos em Acesso Aberto: https://oasisbr.ibict.br [↑](#footnote-ref-11)
11. O grupo de trabalho foi automaticamente suspenso em abril de 2019 pelo Decreto 9.759, que extinguiu centenas de comitês e comissões instituídos pela administração pública federal, sendo depois recriado pela Portaria Nº 3.679 de 7 de agosto de 2019. [↑](#footnote-ref-12)
12. https://www.rnp.br/noticias/rnp-cnpq-e-ibict-divulgam-resultado-do-edital-para-incubacao-de-repositorios-de-dados-de [↑](#footnote-ref-13)
13. Em 2022 o CNPq liderou a criação do Consórcio Nacional para Ciência Aberta (CoNCienciA), que congrega um conjunto de instituições com repositórios institucionais de dados de pesquisa, além dos repositórios de dados LattesData do CNPq e Aleia do IBICT. O Consórcio tem o objetivo de facilitar a intermediação com a DataCite, organização internacional responsável por gerar DOI para Dados de Pesquisa, visando a aquisição de DOIs com maior economicidade. [↑](#footnote-ref-14)
14. https://datacite.org/create-dois/ [↑](#footnote-ref-15)
15. Core Trust Seal: entidade internacional que define requisitos e realiza certificações para “repositórios de dados confiáveis”. https://www.coretrustseal.org/ [↑](#footnote-ref-16)
16. Consortium of European Social Science Data Archives. https://www.cessda.eu [↑](#footnote-ref-17)