# Análise das propostas de certificação de repositórios ao Core Trust Seal: o que podemos aprender com elas?

*Samile Andrea de Souza Vanz[1], Rene Faustino Gabriel Junior[2], Marcel Garcia de Souza[3], Washington Segundo[4], Caterina Groposo Pavão[5]*

## Palabras claves

Certificação, Repositório confiável, Core Trust Seal.

Certification, Trustworthy data repositories, Core Trust Seal.

## Eje temático

Datos abiertos

## Resumen

Observa-se o crescimento do movimento da ciência aberta e maior adesão da comunidade científica à prática de compartilhamento e reuso de dados de pesquisa. Paulatinamente, a infraestrutura para possibilitar estas práticas vem sendo desenvolvida no Brasil, e a partir de seu estabelecimento, surge a necessidade de aprimorar sua qualidade. Neste sentido, inúmeras instituições têm discutido as características e os requisitos de um repositório confiável. Este estudo tem como objetivo analisar as respostas de três repositórios certificados em 2024 pelo Core Trust Seal, com o intuito de explorar possibilidades de atendimento aos 16 requisitos da instituição certificadora. Observou-se que a possibilidade de respostas é bastante ampla e respeita as características do repositório em si, da área de pesquisa e da tipologia do dado.

## Introdução

A maior adesão da comunidade científica à ciência aberta é paulatina à disponibilidade de repositórios confiáveis. Assim como o acesso aberto às revistas científicas despertou desconfiança e inúmeras críticas nos anos iniciais, a abertura dos dados de pesquisa tem gerado as mesmas incertezas. Soma-se a isso o desconhecimento acerca do significado do compartilhamento e reuso de dados, como parte do movimento da ciência aberta (Caregnato et al., 2019). Nesse contexto, a existência de repositórios certificados por instituições reconhecidas pode despertar maior confiança dos pesquisadores e fortalecer as iniciativas de compartilhamento.

1 UFRGS, samile.vanz@ufrgs.br
2 UFRGS, rene.gabriel@ufrgs.br
3 IBICT, marcelsouza@ibict.br
4 IBICT, washingtonsegundo@ibict.br
5 UFRGS, caterina@ufrgs.br



Este documento se distribuye bajo una licencia **Creative Commons Attribution 4.0 International**

Um dos argumentos em prol do compartilhamento de dados recai sobre o custeio da pesquisa científica. No ecossistema da ciência, os financiadores desempenham papel fundamental ao fornecer recursos à pesquisa básica e aplicada. No Brasil, estudos relatam que quase 70% das publicações são financiadas por dez agências públicas, principalmente Capes, CNPq e FAPESP, e oitenta empresas (principalmente industriais e farmacêuticas) também foram identificadas por financiar pesquisas no Brasil, nenhuma delas nacional (McManus & Baeta Neves, 2021). Considerando que a ciência abrange grande parcela de investimento público, é mister garantir a perenidade e a otimização desse investimento. Nesse sentido, o compartilhamento possibilita o aproveitamento e utilização do dado de pesquisa por mais tempo, por diferentes cientistas e grupos de pesquisa, e permite que o mesmo seja analisado a partir da aplicação de diferentes métodos de pesquisa.

Do ponto de vista do pesquisador responsável pela coleta de dados, um repositório confiável garante que seus dados estão em local seguro, permanecendo acessíveis, utilizáveis e inteligíveis ao longo do tempo. Do ponto de vista do pesquisador que deseja utilizar novamente dados coletados em uma pesquisa anterior, o interesse é por dados de alta qualidade que vêm sendo preservados adequadamente (Rezende, 2021). A infraestrutura para arquivamento, preservação e disseminação dos dados de pesquisa vêm sendo desenvolvida há alguns anos, estruturada em repositórios criados em softwares próprios para esta finalidade (Rocha et al., 2021).

No contexto da América Latina, a criação de repositórios vem acontecendo no interior das instituições de pesquisa e universidades (Gabriel Junior et al., 2019; Silveira, Pavão & Vanz, 2023). No caso do Brasil, a exigência do depósito de dados e da apresentação de um Plano de Gestão de Dados começou a ser feita pelas agências de financiamento há alguns anos. A Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) e o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) estão entre as fundações que tornaram o planejamento do compartilhamento de dados mandatório. Esta demanda se tornou um estímulo ao desenvolvimento dos repositórios nacionais e, ultrapassado o momento inicial de desenvolvimento destes repositórios, torna-se premente buscar critérios de qualidade para o estabelecimento desta infraestrutura.

Dada a importância dos repositórios confiáveis, inúmeras instituições vêm desenvolvendo critérios e metodologias de certificação (Santos & Vanz, 2023). Especificamente, a certificação promete que o atendimento a esses critérios e padrões tornará as partes interessadas dos repositórios mais confiantes de que os dados que eles contêm serão protegidos, gerenciados adequadamente e estarão disponíveis para reutilização futura. Até o momento, o Brasil ainda não possui nenhum repositório de dados certificado. Nesse sentido, observar a experiência de repositórios estrangeiros pode ser considerada uma excelente oportunidade para apoiar a certificação dos repositórios nacionais. Este estudo tem como objetivo analisar as respostas de três repositórios certificados em 2024 pelo Core Trust Seal, com o intuito de explorar possibilidades de atendimento aos 16 requisitos da instituição certificadora.

## Certificação de repositório digital confiável

Um repositório digital confiável é aquele capaz de manter a autenticidade, preservar e fornecer acesso a materiais digitais ao longo do tempo. Conforme o relatório **Trusted Digital Repositories: attributes and responsibilities** (RLG/OCLC, 2002), os repositórios digitais confiáveis devem assumir a responsabilidade pela manutenção dos documentos digitais em nome de seus depositantes. Para honrar este compro-

misso, é necessário que possuam uma estrutura organizacional que garanta não apenas a viabilidade do próprio repositório a longo prazo, mas também a preservação dos materiais digitais sob sua responsabilidade.

Cada vez mais as instituições e as bibliotecas sentem a necessidade de mostrar aos pesquisadores a qualidade dos repositórios nos quais eles depositam os seus dados de pesquisa. Nos Estados Unidos, agências federais como os National Institutes for Health (NIH) estão trabalhando em políticas e orientações para guiar a seleção de repositórios confiáveis para o depósito de dados de pesquisa quando estas são financiadas com recursos públicos (Key, Llebot & Boock, 2023).

O resultado da certificação, de acordo com Donaldson (2020), é que ela fornece validação externa de que os repositórios que afirmam preservar e fornecer acesso a recursos digitais, tanto agora como no futuro, estão realmente à altura do desafio. Ou seja, o repositório, foi avaliado por outra pessoa, para determinar se ele pode preservar, gerenciar e fornecer acesso a diversos tipos de materiais digitais em vários formatos, fazer a curadoria desses materiais para permitir a pesquisa, a descoberta e a reutilização e controlar para que o material digital seja autêntico, confiável, acessível e utilizável continuamente. Os méritos da certificação do repositório devem ser divulgados para financiadores, produtores de dados, depositantes e consumidores e os repositórios devem utilizar selos e marcas de certificação para comunicar a sua certificação e passar segurança para a comunidade designada.

Um repositório digital confiável é "[...] aquele cuja missão é fornecer acesso confiável e de longo prazo a recursos digitais gerenciados para sua comunidade designada, agora e no futuro" (RLG/OCLC, 2002). Há quase 30 anos, um grupo internacional de representantes de arquivos e bibliotecas nacionais, faculdades e universidades, agências de financiamento, indústria, editoras, e outras organizações governamentais e do setor privado articularam pela primeira vez pela necessidade de repositórios digitais confiáveis (Donaldson, 2020). Desde então, vários padrões para certificação de repositórios foram desenvolvidos em todo o mundo. O cenário global de certificação de repositórios está em constante mudança e adaptação, no Quadro 1 pode-se verificar um resumo das principais certificações para repositórios de dados.

**Quadro 1 –** Principais certificadores de repositórios

| NOME | DESCRIÇÃO | ANO | CORRENTE |
|---|---|---|---|
| Open Archival Information System (OAIS). ISO 14721:2012 | Criado pelo Consultative Committee for Space Data Systems (CCSDS). É um framework, não uma certificação. A maioria das certificações, atuais ou não, são baseadas nele. | 2012; versão obsoleta de 2003 | Sim |
| NESTOR, DIN 31644 | Adaptado ao contexto alemão, publicado pelo Working Group Trusted Repositories – Certification. Consiste em um catálogo de critérios que são destinados, principalmente, para as organizações que têm o compromisso de preservar a memória, como, por exemplo, arquivos, bibliotecas e museus. Orienta a elaboração, o planejamento e a implementação de um repositório digital confiável a longo prazo. | 2007 | Sim |

| NOME | DESCRIÇÃO | ANO | CORRENTE |
|---|---|---|---|
| Trustworthy Repositories Audit and Certification (TRAC) | Criado pelo Center for Research Libraries (CRL), of Research Libraries Group (RLG) e o National Archives and Records Administration (NARA). Baseado em OAIS, consistia em produzir e esquematizar um processo genérico para auditar e certificar repositórios digitais. Substituído pela ISO 16363. | 2007 | Não |
| Data Seal of Approval (DSA) | Desenvolvido pela Dutch Data Archiving and Networked Services (DANS) para repositórios de dados. Fundido com a Certificação WDS para se tornar o Core Trust Seal. | 2008 | Não |
| Audit and Certification of Trustworthy Digital Repositories (ACTDR) | Baseada em OAIS. Desenvolvido pelo Consultative Committee for Space Data Systems (CCSDS). Destina-se a administradores de repositórios e certificadores. Avalia a infraestrutura organizacional, sustentabilidade financeira, gerenciamento dos objetos digitais e gestão de riscos dos repositórios. Tornou-se a norma ISO 16363:2012 | 2011 | Sim |
| ISO 16363 - Audit and Certification of Trustworthy Digital Repositories | Define práticas recomendadas para avaliar a confiabilidade dos repositórios digitais. É aplicável a toda a gama de repositórios digitais, não apenas para repositórios de dados. | 2012 | Sim |
| WDS Certification | O World Data System (WDS) é um órgão afiliado do International Science Council (ISC) para repositórios de dados. Sua missão é melhorar as capacidades, o impacto e a sustentabilidade dos repositórios de dados e serviços de dados dos seus membros. Fundiu-se com o Data Seal of Approval (DAS) para tornar-se o Core Trust Seal. | 2011 | Não |
| Core Trust Seal | Mescla a certificação WDS e DAS. Baseia-se numa autoavaliação que é revista por pares e em evidências publicamente disponíveis. Para repositórios de dados em todas as disciplinas. Abrange aspectos importantes como governança, infraestrutura técnica, modos de acesso e licenciamento, gestão de metadados, preservação digital e sustentabilidade financeira. | 2017 | Sim |

Fonte: Adaptado de (Key, Llebot & Boock, 2023; Santos, 2018).

Neste estudo aborda-se o Core Trust Seal, uma certificação baseada em autoavaliação e projetada especificamente para repositórios de dados de todas as áreas do conhecimento. O Core Trust Seal (CTS) é um certificado internacional que atesta a confiabilidade e a qualidade dos repositórios de dados de pesquisa. O CTS fornece critérios de avaliação para verificar se um repositório emprega boas práticas de gestão e preservação de dados, que abrangem aspectos como governança, infraestrutura técnica, modos de aces-

so e licenciamento, gestão de metadados, preservação digital e sustentabilidade financeira (Santos & Vanz, 2023). A terminologia utilizada pelo CRS é baseada no Modelo OAIS e os Princípios FAIR estão implícitos nos requisitos.

O certificado Core Trust Seal é concedido após uma avaliação criteriosa por meio do preenchimento de um formulário online específico onde o repositório deve provar que cumpre 16 requisitos. A equipe gestora do repositório deve indicar um nível de conformidade para cada um dos 16 requisitos: 0 – Não se aplica; 1 – O repositório ainda não considerou este requisito; 2 – O repositório possui um conceito teórico sobre este requisito; 3 – O repositório está em fase de implementação deste requisito; 4 – Este requisito foi totalmente implementado no repositório.

Ao longo do formulário de submissão, todo em língua inglesa, o responsável deve informar os níveis de conformidade, que são indicadores do processo auto avaliativo do repositório. A avaliação de conformidade declarada nas respostas é feita pelos avaliadores a partir das evidências comprobatórias apresentadas, como links de documentos formais, websites, entre outros.

Exige-se justificativa detalhada para resposta "não se aplica" conferida a algum requisito. A certificação pode ser concedida se alguns dos requisitos estiverem em fase de implementação (3) e o restante implementado (4). No entanto, níveis de conformidade na fase (1) ou (2) não são suficientes para a obtenção da certificação. Todos os 16 requisitos são mandatórios pois refletem as características desejáveis de um repositório confiável, por isso possuem pesos iguais no processo avaliativo. Algumas sobreposições e/ou informações duplicadas são inevitáveis.

A certificação é válida por 3 anos e desde fevereiro de 2024 a taxa para submissão é de EUR 3.000, para cobrir custos com avaliação por pares e a administração da organização sem fins lucrativos. A taxa administrativa dá direito a cinco revisões.

## Procedimentos metodológicos

O Core Trust Seal disponibiliza um arquivo PDF contendo as respostas e evidências fornecidas pelos repositórios acreditados, bem como comentários dos avaliadores para cada item avaliado. Foram analisadas as respostas concedidas por três repositórios certificados em fevereiro de 2024, momento da coleta de dados:

1. GAMS (https://gams.uni-graz.at), repositório específico para Ciências Humanas coordenado pelo Centre for Information Modelling at the University of Graz (Alemanha); https://doi.org/10.34894/BYIJIP
2. Level-1 Atmosphere Archive & Distribution System (LAADS) Level-1 and Atmosphere Archive & Distribution System Distributed Active Archive Center - LAADS DAAC (nasa.gov), Distributed Active Archive Center (DAAC), vinculado à NASA; https://doi.org/10.34894/ZMSMZI
3. IPSL Computing and Data Center Institut Pierre-Simon Laplace – Sciences du climat (ipsl.fr), dedicado ao gerenciamento, coleta, distribuição e serviços de dados em ciências climáticas para observação e modelagem de dados, vinculado ao Institute Pierre Simon Laplace (IPSL); https://doi.org/10.34894/C4OTAF

O LAADS DAAC submeteu proposta conforme o CoreTrustSeal Requirements 2023-2025. O GAMS e o IPSL utilizaram a versão CoreTrustSeal Requirements 2020-2022. As versões dos requisitos apresentam pequenas diferenças que foram compatibilizadas, e apresentadas em um quadro de acordo com a versão mais atual, qual seja, CoreTrustSeal Requirements 2023-2025.

Além da análise das respostas, todos os documentos indicados como evidência comprobatória para os 16 requisitos foram acessados a partir do link fornecido no documento PDF. Uma síntese de seu conteúdo foi registrada em um quadro comparativo.

## Resultados

A leitura dos formulários preenchidos pelos três repositórios revelou inúmeras possibilidades de preenchimento das informações para atender aos 16 requisitos do Core Trust Seal. As respostas foram sintetizadas e organizadas em um quadro, conforme os requisitos. Observou-se que as particularidades dos repositórios são respeitadas. Os três casos analisados apresentam características bastante distintas, tanto em relação à área de pesquisa quanto em relação ao tipo e formato de dados.

**Quadro 2** – Resposta dos repositórios analisados e organizados conforme Requisitos CTS

| | Requisito e Escopo | GAMS | LAADS DAAC | IPSL |
|---|---|---|---|---|
| | R0 Contexto do repositório<br>Informar: tipo de repositório; breve descrição do repositório e comunidade; nível de curadoria; | Informaram o identificador Re3data, tipo de repositório, vínculo institucional do repositório e comunidade depositante, infraestrutura e equipe responsável. | Informaram o identificador Re3data, tipo de repositório, vínculo institucional do repositório, dados estatísticos como volume de arquivos depositados, usuários e visitas; instituições parceiras. | Informaram e apresentaram as diversas comunidades científicas que utilizam o repositório, e apresentaram a rede de parceiros. |
| Infraestrutura organizacional | R1 Missão/escopo<br>Descrever a missão da organização em preservar e fornecer acesso aos dados. | Informaram a missão, o âmbito da coleção e a comunidade designada. Reconheceram a responsabilidade pela criação do repositório o que garante as perspectivas de continuidade e preservação da infra-estrutura. | Informaram a missão do repositório e encaminharam por meio de links para Requisitos para arquivamento, distribuição e serviços ao usuário no sistema de informações e dados, assim como para informações adicionais. | Informaram a missão e os vários campos relativos a aspectos do ciclo de vida dos dados, produtos de dados de diferentes interesses para a comunidade das ciências da terra e apresentaram links para diversos sites e plataformas. |
| | R2 Licenças<br>Informar as licenças aplicáveis | Informaram que o Instituto promove o acesso aberto e a disponibilidade gratuita de dados de pesquisa; tipo de licença e a regulamentação sobre dados pessoais ou dados confidenciais, assim como os procedimentos em casos de violação de propriedade intelectual ou de direitos pessoais. | Informaram que o repositório está em conformidade com a Política de Dados e Informações de Ciências da Terra da NASA, com acesso gratuito e aberto às suas coleções de dados. Indicaram site da Política de dados, Direitos de dados e instruções de citação. | A maioria das informações dizem respeito a licenças, que os dados não estão sujeitos a qualquer regulamentação de direitos autorais, patentes, marcas registradas ou segredos comerciais. Informaram que não dispõem de uma Política de Dados geral, que é elaborado um contrato com cada fornecedor de dados. |
| | R3 Continuidade do serviço<br>Ter um plano para garantir o acesso contínuo e a preservação de seus dados e metadados | Apresentaram a política institucional da Universidade de Graz que se responsabiliza pela manutenção do repositório, com Plano de preservação para 10 anos. Em caso de descontinuidade, está prevista devolução aos proprietários dos dados ou entrega a repositórios temáticos ou institucionais. | Apresentaram declaração da NASA, dizendo acreditar que a administração de longo prazo dos dados de sensoriamento remoto e de campanha de campo coletados pela NASA é essencial. | A ampla rede de parceiros garante a sucessão e disponibilidade dos dados. O texto informa as responsabilidades de cada parceiro e a correspondente previsão de anos de preservação. |
| | R4 Questões éticas e legais<br>Informar as disposições éticas e de privacidade que afetam a criação, curadoria e uso dos dados | Indicaram que o depositante é responsável por seguir normas éticas e legais nacionais e internacionais e específicas de cada área; dados pessoais e sensíveis são de responsabilidade do depositante e podem ser excluídos do repositório a qualquer momento; requerem a assinatura do termo de depósito. | Os dados arquivados se referem a Terra e seu ambiente e, portanto, não são suscetíveis a risco, são disponíveis gratuitamente e o depositante assina o termo de depósito. | Os dados arquivados se referem a observações e simulações sobre clima, por isso não há preocupação com questões éticas. As questões que exigem atenção se referem a datasets protegidos por embargo e dados questionados pela comunidade científica. O repositório mantén uma equipe preparada para estes casos. |
| | R5 Governança e recursos<br>Ter financiamento adequado e número suficiente de funcionários gerenciados por meio de um sistema claro de governança para realizar a missão com eficácia | Informaram o número de funcionários em tempo integral que cuidam da infra-estrutura, realizam tarefas técnicas, e curadoria de conteúdo; não há riscos de financiamento, mas a perda de membros-chave do pessoal é um risco, que é minimizado com documentação e transferência de conhecimento. | Informaram o gasto médio anual que permite cumprir a sua missão, incluindo apoio ao pessoal, recursos de TI, formação e viagens; número de funcionários; treinamentos anuais exigidos pela NASA para a equipe lidar com dados. | Detalharam a quantidade de profissionais, sua formação, suas atividades e regime de trabalho; orçamento anual e o valor utilizado para equipamentos; em termos de governança esclareceram a composição e o trabalho de comitês e comissões. |
| | R6 Orientação de especialista<br>Adotar mecanismos para garantir conhecimento, orientação e feedback contínuos de especialista | Informaram como é realizada a comunicação interna, por meio de reuniões de pesquisa e grupos de trabalho. Os comentários e solicitações dos usuários são geralmente direcionados ao endereço de e-mail do departamento e distribuídos à equipe técnica. | O repositório mantém um Escritório de Atendimento ao Usuário para interagir com a comunidade. Informaram que participam anualmente de pesquisa que fornece um índice de satisfação do cliente. Esclareceram que participam de uma comunidade em rede que reúne profissionais de ciência, dados e tecnologia da informação de mais de 120 organizações, incluindo agências federais dos EUA, universidades e entidades comerciais. | Busca feedback da comunidade, em apresentações, reuniões e sessões práticas organizadas durante os treinamentos. Produzem tutoriais que incluem melhores práticas e uso passo a passo dos serviços e ferramentas para seus próprios projetos de pesquisa. Acrescentam que mantém a comunidade informada por meio de seu site. Possui um link "Contato" que permite aos usuários fazer perguntas ou fornecer comentários ao suporte ao usuário. |

| | Requisito e Escopo | GAMS | LAADS DAAC | IPSL |
|---|---|---|---|---|
| Gerenciamento de objetos digitais | R7 Proveniência e autenticidade<br>garantir a integridade e autenticidade dos dados | Informaram que o Fedora suporta versionamento de todos os aspectos do recurso digital, porém às vezes falta transparência desses processos para o usuário final. Por esse motivo, até não conseguirem a implementação completa, o nível de conformidade do requisito foi definido como 3. | Informaram os procedimentos de Integridade do arquivo de dados, garantia de qualidade, validação do produto, gerenciamento de versões, documentação do produto e descoberta e acesso a dados. | Informaram que, devido à origem dos dados, não é necessário verificar sua autenticidade. Utilizam histórico de versões e acesso privilegiado. Acrescentam informações sobre rastreabilidade e identificadores persistentes. |
| | R8 Depósito e avaliação<br>aceitar dados e metadados baseados em critérios previamente definidos garantindo relevância e inteligibilidade aos usuários | Informaram que na sua maioria os dados serão depositados como parte de um projeto de pesquisa, complementados por um plano de gestão de dados. Cada projeto é apoiado por um gestor de metadados que auxilia no fluxo de trabalho, modelagem de dados, processo de depósito e publicação. O controle de qualidade é apoiado pelo uso de vocabulários controlados e arquivos de autoridade. Informaram que utilizam o padrão de Dublin Core para recuperação de informações e identificação de recursos e operam com um conjunto limitado de formatos e tipos de dados para manter o repositório sustentável por longo período de tempo. | Informaram que os produtores são obrigados a enviar dados que estejam em conformidade com os padrões da comunidade no que se refere a formatos de dados, metadados e interfaces e estes são verificados para garantir que atendem padrões suficientes para auxiliar na sua interpretação, utilização e garantir que todos os requisitos de metadados sejam atendidos durante o ciclo de vida de dados e a preservação a longo prazo. Disponibilizam o link com uma lista de formatos de dados e metadados aceitáveis. | Esclareceram que os dados são validados pelas equipes científicas responsáveis pelo conjunto de dados antes da divulgação, verificando compressão dos dados através dos seus metadados e a conformidade com as sintaxes de dados que definem formatos de arquivos, nomes de arquivos, estruturas de diretórios, etc. Informaram sobre as diretrizes para auxiliar a formatar os dados e completar os metadados e que o depósito de dados em formato não preferencial ou não suportado é permitido apenas para dados de cauda longa, pois para permitir a descoberta dos dados é dada especial atenção ao formato dos dados. |
| | R9 Plano de preservação<br>Assumir a responsabilidade pela preservação a longo prazo e gerência essa função de forma planejada e documentada | Apresentaram uma síntese dos aspectos principais e diversos links para documentos comprobatórios do plano de preservação. | Apresentaram uma síntese dos aspectos principais do plano de preservação e o link para o texto integral. | O ESPRI arquiva datasets de diferentes naturezas e provenientes de diversas fontes de referência, que requerem diferentes níveis de arquivo a longo prazo. Apresentaram link para todos esses documentos. |
| | R10 Qualidade dos dados<br>possuir experiência apropriada para lidar com dados técnicos e qualidade de metadados e garantir que informações suficientes estejam disponíveis para os usuários finais fazerem avaliações relacionadas à qualidade | Detalharam procedimentos para garantir a qualidade dos dados e dos metadados, baseado principalmente no rigoroso trabalho da equipe com formação na área de Humanas e Ciência da Informação, o que possibilita um diálogo entre equipe do repositório e pesquisadores. | Informaram que o repositório utiliza mecanismos automatizados para controle da qualidade dos dados e metadados. Contato por telefone e e-mail é utilizado para dirimir dúvidas dos usuários. | Informaram que as ferramentas desenvolvidas permitem o controle dos metadados e garantem autodescrição dos metadados (cada variável no arquivo tem uma descrição associada, incluindo unidades físicas, localização espacial, proveniência, citação, etc.). O repositório delega ao depositante o controle da qualidade, e não se responsabiliza por controle adicional. |
| | R11 Workflows (Fluxos de trabalho)<br>A gestão de objetos digitais ocorre de acordo com fluxos de trabalho definidos, desde o depósito até o acesso | Apresentaram o link para o workflow do repositório, e sintetizaram os procedimentos principais. | Mencionaram que os processos de ingesta, curadoria, exportação e arquivamento possuem workflows. | Detalharam os processos de ingesta, curadoria e disseminação dos dados. |
| | R12 Identificação e descoberta de dados<br>Permitir a descoberta dos dados e consulta de forma persistente por meio de citações adequadas | Detalharam todas as possibilidades de pesquisa e acesso aos dados, indicando os respectivos sites. Detalhes sobre identificador persistente também foram fornecidos. | Detalharam todas as possibilidades de pesquisa e acesso aos dados, indicando os respectivos sites. Detalhes sobre identificador persistente também foram fornecidos. | Detalharam procedimentos de harvesting e catálogos interoperáveis; o repositório arquiva diversos tipos de dados que demandam ferramentas distintas para descoberta e acesso. |

| | Requisito e Escopo | GAMS | LAADS DAAC | IPSL |
|---|---|---|---|---|
| Tecnologia da Informação e Segurança | R13 Reuso dos dados<br>Permitir a reutilização dos dados ao longo do tempo, garantindo que os metadados apropriados estejam disponíveis para apoiar a compreensão e o uso dos dados | Informaram que o primeiro ponto de acesso geralmente será a interface gráfica do usuário criada especificamente para o Repositório. Esclareceram sobre o padrão de metadados, os Elementos Semânticos e o Modelo de Dados da Europeana utilizados. | Esclareceram que fornecem metadados em nível de coleção e de arquivo para facilitar a pesquisa e descoberta de dados. Detalharam os formatos de dados e o fornecimento de links para documentos mais detalhados. | Esclareceram que um objetivo da descoberta de dados é permitir que o repositório seja coletado por outros catálogos de dados interoperáveis. Detalharam as pesquisas do catálogo para dados geoespaciais e forneceram uma lista de links para auxiliar no entendimento deste requisito. |
| Tecnologia da Informação e Segurança | R14 Armazenamento e integridade<br>Permitir a reutilização dos dados ao longo do tempo, garantindo que os metadados apropriados estejam disponíveis para apoiar a compreensão e o uso dos dados | Esclareceram que dados e metadados serão migrados, caso haja necessidade no planejamento da preservação. | Informaram que existem vários documentos de Controle de Interface, Documentos de Requisitos e Acordos de Operações que regem a dinâmica do fluxo de dados e especificam requisitos para segurança da informação. | Esclareceram que aplicam os princípios FAIR na gestão dos seus dados e infraestruturas e, em particular, no princípio da reutilização de dados. Especificaram a utilização do formato formato netCDF. |
| | R15 Infraestrutura técnica<br>O repositório é gerenciado em sistemas operacionais bem suportados e outros núcleos software e hardware de infraestrutura apropriados aos serviços que presta aos seus Comunidade Designada | Informaram os servidores existentes e localização dos mesmos, a lista de softwares e tecnologias envolvidas. | Informaram a infraestrutura técnica e localização dos servidores. | Forneceram todo o detalhamento dos equipamentos e da infraestrutura disponível, informando tamanhos dos servidores e das redes disponíveis. |
| | R16 Segurança<br>O repositório protege a instalação e seus dados, metadados, produtos, serviços e usuários | Informaram volume de servidores e locais, volume de backups, quem faz a gestão do repositório, planos de substituição de pessoas ao longo dos anos. | Detalharam a segurança física e de software. | Informaram em detalhes todos os procedimentos técnicos para manutenção do repositório em segurança. |

Fonte: dados da pesquisa.

Os Repositórios Digitais Confiáveis devem dispor de governança sustentável e estruturas organizacionais, infraestrutura confiável e políticas abrangentes de apoio às práticas acordadas pela comunidade; devem demonstrar capacidades essenciais e duradouras para permitir o acesso e reutilização de dados ao longo do tempo para as comunidades que atendem (Lin et al., 2020). Os 16 requisitos do Core Trust Seal estão organizados no sentido de avaliar as capacidades do repositório em relação às características de confiabilidade exigidas. A confiabilidade é demonstrada por meio de evidências que dependem de transparência, desta forma, os repositórios devem fornecer evidências transparentes, honestas e verificáveis de sua prática. No caso dos três repositorios analisados, as evidências consistiram no link para documentos importantes, como o Preservation Plan GAMSRepository (https://gams.uni-graz.at/o:gams.preservationplan); NASA Data and Information Policy (https://www.earthdata.nasa.gov/engage/open-data-services-and-software/data-and-information-policy); e fluxos de trabalho mapeados no ESPRI (https://cloud.ipsl.fr/index.php/s/ZiZWeeetRYw8CSG).

Observou-se que no caso do IPSL Computing and Data Center - ESPRI (Ensemble de Services Pour la Recherche à l'IPSL), considerando ser um repositório dedicado a dados observados e de modelagem nas ciências climáticas, a proposta envolve inúmeros parceiros institucionais, todos eles evidenciados através de documentos como convênios e contratos de parceria. É possível perceber que a equipe reuniu documentos que vão além daqueles relacionados de forma direta ao repositório, abrangendo a documentação de origem dos convênios e contratos institucionais de colaboração científica e parceria. Sem dúvida, ao evidenciar o compromisso institucional com os dados e o próprio repositório, a equipe avaliadora do Core Trust Seal pode confiar na garantía de integridade, autenticidade, precisão, confiabilidade e acessibilidade dos dados por longo prazo.

A questão da segurança é tema recorrente e perpassa vários dos 16 requisitos do Core Trust Seal (CORETRUSTSEAL, 2023). Para atender tais requisitos, o repositório deve analisar ameaças potenciais, avaliar riscos e criar um sistema de segurança consistente. Deve descrever cenários de danos com base em ações maliciosas, erro humano ou falha técnica que representam uma ameaça ao repositório e seus dados, produtos, serviços e usuários. A probabilidade e o impacto de tais cenários devem ser mensurados, para possibilitar a avaliação acerca de quais níveis de risco são aceitáveis e determinar quais medidas devem ser tomadas para combater as ameaças ao repositório e sua comunidade designada. Para este requisito deve ser descrito: o sistema de segurança de TI, funcionários com funções relacionadas à segurança e quaisquer ferramentas de análise de risco utilizados; os níveis de segurança que são exigidos e como eles são suportados; quaisquer procedimentos de autenticação e autorização empregados para gerenciar com segurança o acesso aos sistemas em uso.

## Conclusão

O processo de certificação serve como catalisador para reunir grupos de pessoas em torno da biblioteca com interesses comuns. O grupo responsável pela preservação digital deve responder aos requisitos da estratégia de preservação, e grupos diversos de usuários dos repositórios devem ser consultados para tomar decisões relativas a serviços do repositório. Além disso, a equipe que trabalha diretamente na customização do repositório e a equipe que coleta, prepara a documentação deve contribuir para o compartilhamento de conhecimento permitindo solucionar inconsistências e problemas antes e depois da submissão da certificação.

Na certificação de um repositório destaca-se a importância da documentação disponibilizada de forma pública e aplicação das boas práticas, assim como a sua documentação. Além disso, manter uma equipe dedicada para os serviços do repositório torna-se essencial para satisfazer a maioria dos requisitos. O efeito combinado desses requisitos se traduz no aumento significativo na confiança e na qualidade dos repositórios.

No estudo desenvolvido por Donaldson (2020) foram levantados alguns questionamentos sobre a certificação de repositórios. A pesquisa mostrou que os repositórios podem preservar efetivamente a informação digital sem necessariamente seguir as melhores práticas e que aqueles que tentam seguir práticas recomendadas específicas muitas vezes têm dificuldade em fazê-lo. O autor coloca não saber se a auditoria e a certificação dos repositórios digitais confiáveis realmente são importantes, por exemplo, se os repositórios digitais são na verdade, melhores na preservação de informações digitais após a certificação do que eram antes e se os repositórios digitais confiáveis preservam melhor a informação digital do que os seus homólogos, embora os padrões dos repositórios digitais confiáveis promulguem esta suposição. Uma forma de avaliar se a auditoria e a certificação dos repositórios digitais confiáveis são importantes consistiria em examinar o seu impacto nas partes interessadas dos repositórios digitais confiáveis.

Porém, os instrumentos de certificação estão melhorando ao longo do tempo e estes, por sua vez, estão sendo utilizados para melhorar políticas, procedimentos, competências, ferramentas e serviços dos repositórios de dados para permitir a reutilização de dados de forma contínua. O processo de certificação permite refletir sobre as melhorias necessárias no repositório e nos serviços de dados de pesquisa, incluindo eficiência de fluxo de trabalho, melhorias nos recursos de software e na preservação dos conjuntos de dados. A partir da primeira certificação, a literatura consultada mostra que, as bibliotecas percebem a necessidade de elaborar novos manuais como por exemplo, sobre os princípios FAIR e as características desejáveis para os dados da pesquisa financiada pela instituição, visando aumentar a conformidade com os requisitos na renovação da sua certificação.

Este estudo ajudou a entender não apenas o processo de submissão da proposta de certificação de repositório digital confiável ao Core Trust Seal, mas principalmente, verificar o tipo de resposta que os avaliadores esperam dos gestores do repositório, o nível, a quantidade e a disponibilidade da documentação relativa ao repositório e suas boas práticas à qual os certificadores devem ter acesso para confirmar as respostas dos 16 requisitos. Verificou-se que em alguns casos as respostas eram bastante sucintas e foram consideradas satisfatórias mesmo não remetendo para links de documentação adicional. Também, foi possível avaliar o nível de conformidade e os comentários.

## Bibliografía

Caregnato, S. E., Vanz, S. A. De S., Pavão, C. M. G., Passos, P. C. S. J., Borges, E. N., Gabriel Junior, R. F., Azambuja, L. A. B. & Rocha, R. P. (2019). Práticas e percepções dos pesquisadores brasileiros sobre serviços de acesso aberto a dados de pesquisa. LIINC EM REVISTA, 15, 121-141. https://doi.org/10.18617/liinc.v15i2.4771

CORETRUSTSEAL. Coretrustseal Trustworthy Data Repositories Requirements 2023-2025. 2023. https://zenodo.org/records/7051012

Donaldson, D. R. (2020). Certification information on trustworthy digital repository websites: A content analysis. PLoS ONE, 15(12): e0242525. https://doi.org/10.1371/journal. pone.0242525

Gabriel Junior, R. F., Rocha, R. P., Caregnato, S. E., Pavão, C. M. G., Passos, P. C. S. J., Borges, E. N., Vanz, S. A. de S. & Azambuja, L. A. B. (2019). Acesso aberto a dados de pesquisa no Brasil: mapeamento de repositórios, práticas e percepções dos pesquisadores e tecnologias. Ciência da Informação (Online), 48, 87-101. https://doi.org/10.18225/ci.inf.v48i3.4958

Key, C., Llebot, C. & Boock, M. (2023). Building a Trustworthy Data Repository: CoreTrustSeal Certification as a Lens for Service Improvements. Journal of eScience Librarianship, 12(3): e761. https://doi.org/10.7191/jeslib.761

Lin, D., Crabtree, J., Dillo, I. et al. The TRUST Principles for digital repositories. Sci Data 7, 144 (2020). https://doi.org/10.1038/s41597-020-0486-7

McManus, C. & Baeta Neves, A.A. (2021). Funding research in Brazil. Scientometrics 126, 801–823. https://doi.org/10.1007/s11192-020-03762-5

Rezende, L. V. R. (2021). Preservação e certificação de repositórios Dataverse. Revista Brasileira de Preservação Digital, 2(00): e021001. https://doi.org/10.20396/rebpred.v2i00.15810

RLG/OCLC Working Group on Digital Archive Attributes. (2002). Trusted digital repositories: Attributes and responsibilities. Mountain View, CA: Research Libraries Group. http://www.oclc.org/research/ activities/past/rlg/trustedrep/repositories.pdf

Rocha, R. P., Gabriel Junior, R. F., Vanz, S. A. de S., Borges, E. N., Azambuja, L. A. B., Caregnato, S. E., Pavão, C. M. G., Passos, P. C. S. J. & Felicissimo, C. H. (2021). Análise dos sistemas DSpace e Dataverse para repositórios de dados de pesquisa com acesso aberto. Revista Brasileira de Biblioteconomia e Documentação (Online), 17, 1-25. https://rbbd.febab.org.br/rbbd/article/view/1572

SANTOS, H. M. dos. (2019). Auditoria de repositórios arquivísticos digitais confiáveis. Informação em Pauta, 4(2), 156-172. https://doi.org/10.32810/2525-3468.ip.v4i2.2019.41787.156-172

Santos, D. B. & Vanz, S. A. de S. (2023). Repositórios de dados de pesquisa: confrontação dos princípios, critérios e requisitos internacionais de avaliação da confiabilidade. Revista Brasileira de Preservação Digital, 4: e023003. https://doi.org/10.20396/rebpred.v4i00.17355

SILVEIRA, J. I. & Vanz, S. A. de S. (2023). Diretrizes para políticas de depósito, acesso e uso de dados de pesquisa: proposta a partir da análise de repositórios de dados universitários internacionais. ATOZ: Novas Práticas em Informação e Conhecimento, 12. https://doi.org/10.5380/atoz.v12i0.87331

Silveira, Pavão e Vanz, 2023, ANAIS DO BIREDIAL ainda não publicado.

Vanz, S. A. de S., Pavão, C. M. G., Caregnato, S. E., Passos, P. C. S. J., Moura, A. M. M. de, Borges, E. N., Gabriel Junior, R. F. & Rocha, R. P. (2021). Diretrizes para o estabelecimento de um checklist para curadoria de dados de pesquisa. Informação em Pauta, 6, 1-18. https://doi.org/10.36517/2525-3468.ip.v6i00.2021.68088.1-18

**Dra. Samile Andréa de Souza Vanz**

Professora associada do Departamento de Ciências da Informação e do Programa de Pós-graduação em Comunicação da Universidade Federal do Rio Grande do Sul (PPGCOM UFRGS). Graduada em Biblioteconomia pela Universidade Federal do Rio Grande do Sul (1999), mestre e doutora em Comunicação e Informação pelo PPGCOM UFRGS (2004 e 2009), com estágio sanduíche na Dalian University of Technology (China, 2007-2008). Pós-doutorado pela Universidad Carlos III de Madrid (Madrid, 2016). Editora da revista Em Questão (2014 – ). Desenvolve pesquisas na área de Comunicação Cientifica, com ênfase na produção de indicadores científicos, bibliometria, colaboração cientifica, análise de citação, análise de co-citação e rankings universitários.Tem experiência acadêmica e profissional na área de Planejamento, gestão e arquitetura de Bibliotecas e Unidades de Informação.

**Dr. Rene Faustino Gabriel Junior**

Graduado em Biblioteconomia e Documentação pela Pontifícia Universidade Católica do Paraná (2008), com mestrado em Ciência, Gestão e Tecnologia da Informação pela Universidade Federal do Paraná (2011) e doutorado em Ciência da Informação pela Universidade Estadual Paulista Júlio de Mesquita Filho (2014). Atualmente é professor adjunto da Universidade Federal do Rio Grande do Sul e do Programa de Pós-Graduação em Ciência da Informação (PPGCIN) da mesma universidade e chefe do Departamento de Ciências da Informação (DCI). Tem experiência na área de Ciência da Informação, com ênfase em Biblioteconomia, atuando principalmente nos seguintes temas: Ciência da Informação, Estudos Métricos da Informação, Bibliometria, Brapci, Comunicação Científica, Dados de Pesquisa e Produção Científica. Implantou e coordena a Base de Dados de Periódicos em Ciência da Informação (BRAPCI). Membro do Grupo de Pesquisa de Comunicação Científica e do Núcleo de Estudos em Ciência, Inovação e Tecnologia da UFRGS.

**Me. Marcel Garcia de Souza**

Doutorando em Ciência da Informação pela Universidade de Brasília. Mestre em Educação em Ciências pela Universidade Federal do Rio Grande do Sul (2016). Graduado em Psicologia pela Universidade Católica de Brasília (2005). Servidor público federal; Analista em Ciência e Tecnologia no Instituto Brasileiro de Informação em Ciência e Tecnologia atuando como Coordenador de Tratamento, Análise e Disseminação da Informação Científica, além de coordenar pesquisas aplicadas voltadas à Ciência da Informação, Informação para Sustentabilidade, Avaliação do Ciclo de Vida, Informação Tecnológica.

**Dr. Washington Lu**ís Ribeiro de Carvalho Segundo

Doutor e Mestre em Informática pela Universidade de Brasília, com Estágio de Doutorado Sanduíche no Kings College London. Possui graduação em Matemática (Bacharelado e Licenciatura) também pela Universidade de Brasília. É Coordenador-geral de Informação Científica e Técnica no Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict / MCTI). É membro e coordena projetos, comitês nas áreas de Ciência Aberta e Ciência de Dados. É líder do Grupo de Pesquisa e Laboratório do Ecossistema da Pesquisa Científica Brasileira (LaEPeCBr) (http://dgp.cnpq.br/dgp/espelhogrupo/9750187028652303, https://pnipe.mctic.gov.br/laboratory/3911). Áreas de interesse em pesquisa: Métodos Formais, Repositórios Digitais Abertos, Repositórios de Dados Científicos, Interoperabilidade entre Sistemas de Informação Abertos, Ciência Aberta e Ciência de Dados.

**Dra. Caterina Marta Groposo Pavão**

Bacharel em Biblioteconomia pela Universidade Federal do Rio Grande do Sul, mestrado e doutorado em Comunicação e Informação pelo Programa de Pós-Graduação em Comunicação e Informação da mesma Universidade e doutorado sanduíche na Universidad Complutense de Madrid. É docente do Departamento de Ciência da Informação da Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul e do Programa de Pós-Graduação em Ciência da Informação da Faculdade de Biblioteconomia e Comunicação da mesma Universidade. Exerceu suas atividades profissionais de 1994-2018 no Centro de Processamento de Dados da Universidade Federal do Rio Grande do Sul, onde dedicou-se à gerência e administração do sistema de automação do SBUFRGS do Lume, repositório institucional da Universidade. Integra a equipe de pesquisadores do Grupo de Pesquisa Comunicação Científica da Faculdade de Biblioteconomia e Comunicação da UFRGS, do Grupo de Estudos e Práticas de Preservação Digital da Rede Cariniana e do grupo de pesquisadores da NUAWEB – Núcleo de Pesquisa em Arquivamento da Web e Preservação Digital. Participa de projetos do Centro de Documentação e Acervo Digital da Pesquisa (CEDAP/UFRGS) e do Centro de Processamento de Dados da UFRGS para implementação do Projeto Repositório de dados científicos da Universidade.