

INTRODUÇÃO

Esta matriz de critérios de avaliação para repositórios de dados e publicações foi elaborada no contexto do 5º Plano de Ação Nacional, na Parceria para Governo Aberto (*Open Government Partnership - OGP*). Ela que está vinculada ao Compromisso 8 -Construir uma proposta de modelo de avaliação que fomente a Ciência Aberta. Apresentam-se os resultados da atividade proposta pelo Marco 2, do referido compromisso. Este Marco tem como objetivo realizar uma proposição de critérios de qualificação de repositórios de dados e de repositórios de publicações. A execução do Marco 02 ficou sob responsabilidade de oito instituições, a saber: Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict, coordenador do Marco), Comissão Nacional de Energia Nuclear (CNEN), Empresa Brasileira de Pesquisa Agropecuária (Embrapa), Fundação Oswaldo Cruz (Fiocruz), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Scientific Electronic Library Online (SciELO) e a Associação Brasileira de Editores Científicos (ABEC) Universidade Estadual Paulista "Júlio de Mesquita Filho" (Unesp).

Além de funcionar como ferramenta para avaliação de repositórios, a matriz também tem como objetivo auxiliar gestores de repositórios de dados e de publicações a identificar serviços que podem ser oferecidos à comunidade científica por meio das plataformas de gestão de conjuntos de dados de pesquisa e de publicações científicas. A matriz permitirá que o gestor possa fazer um levantamento dos serviços que já disponibilizam e dos serviços que ainda não são oferecidos na plataforma utilizada por sua instituição.

Para alcançar as metas acima, a matriz está organizada por tipo de serviço, classe de requisito, requisito e propósito. As classes caracterizam grupos de atividades mais abrangentes e os requisitos são atribuições relacionadas a uma dada classe. Para cada requisito, foram descritos propósitos que têm como finalidade justificar o porquê daquele requisito.

Para identificar os critérios que compõem a matriz, o grupo teve como ponto de partida a matriz da *Resource Description and Access* (RDA), os critérios do *Trustworthy Repositories Audit and Certification: Criteria and Checklist - TRAC / TRUST / FAIR* para Repositórios de Dados, e o quadro para boas práticas em repositórios da comunidade COAR. A partir da identificação de critérios já existentes na literatura, os integrantes do grupo apresentaram sugestões de novos elementos que pudessem compor a matriz.

PARTE 1 - REQUISITOS DO REPOSITÓRIO

	CLASSE	REQUISITO	PROpósito
GOVERNANÇA, INFRAESTRUTURA ORGANIZACIONAL	CERTIFICAÇÃO	O repositório é certificado.	Garantir verificação, confidencialidade, integridade, autenticidade e estabilidade ao longo do tempo.
		Qual o nível de certificação?	Identificar o nível de confiabilidade do repositório (auto certificação, certificação por uma terceira parte etc.).
		Qual a instituição certificadora/norma utilizada.	Identificar quais os requisitos/normas aplicados (OAIS, DIN, etc.).
	POLÍTICA	O repositório tem um documento que explicita a política que ele adota.	Orientar os usuários que depositam e acessam os dados sobre as regras que regem o depósito, acesso e reuso.
		A política está publicada e acessível diretamente pelos usuários de forma explícita.	Tornar transparente as regras de uso do repositório.
		A política considera os códigos éticos e legais e o alinhamento às normativas em vigor no país.	Garantir que o repositório não esteja infringindo a legislação concernente ao acesso à dados (LGDP, LAI, Lei de copyrigth etc.).
		Copyright.	Informar a comunidade o tipo de direito de uso.

	LICENÇAS	O repositório explicita os tipos de licenças que ele adota para depósito, acesso e reuso.	Comunicar aos usuários sobre as suas possibilidades de acesso e reuso dos dados; e os direitos mantidos sobre os dados pelo depositante.
		Que licenças o repositório adota.	Comunicar sobre os padrões de licenças adotados pelo repositório (Creative Commons ...).
SUSTENTABILIDADE		Política – o repositório está alinhado com as políticas locais e nacionais pertinentes.	Garantir que as decisões estejam alinhadas com as políticas locais e nacionais.
		Econômica – o repositório conta com fontes de financiamento de longo prazo.	Garantir a manutenção dos serviços do repositório por um longo prazo
		Organizacional – o repositório está inserido na estrutura organizacional de uma instituição comprometida com a pesquisa científica.	Assegurar que o repositório está alinhado com boas práticas nacionais/internacionais de compartilhamento de dados
		Tecnológica - O repositório acompanha os desenvolvimentos tecnológicos e suas possíveis aplicações, tendo como referência a estabilidade de seus ativos.	Antever as inovações tecnológicas e estabelecer planejamento de normas e padrões pertinentes.
		De longo prazo – O repositório tem um plano para gerenciamento de longo prazo para os dados , incluindo a manutenção da integridade, autenticidade e disponibilidade de conjuntos de dados; tem planos de contingência para garantir que os dados estejam disponíveis e mantidos durante e após eventos imprevistos; tem também um plano de sucessão para o caso de descontinuidade do repositório .	Garantir a usabilidade, viabilidade e interpretação dos dados agora e no futuro; para garantir a continuidade dos datasets em outra plataforma .
	GOVERNANÇA	O repositório está inserido num sistema de governança de amplo espectro.	Apoiar as diversas facetas de sustentabilidade : econômica, política, social, tecnológica e de longo prazo.
	ASPECTOS ÉTICOS E LEGAIS	O repositório dispõe de uma documentação que esclarece sob quais licenças os dados depositados estão atribuídos	Garantir aos autores o correto uso e citação dos dados e publicações depositadas no repositório
	GESTÃO DE RISCO	O repositório tem recursos documentados para garantir que as salvaguardas administrativas, técnicas e físicas sejam empregadas para cumprir com a confidencialidade aplicável,	Assegurar as salvaguardas aplicáveis possam ser cumpridas.

		gestão de risco e monitoramento contínuo para dados sensíveis.	
EQUIPE QUALIFICADA		Dispõe de uma equipe qualificada de profissionais de informação, computação e científica.	Atender as múltiplas funções, serviços e processos de uma plataforma de gestão de dados de pesquisa.
PROMOÇÃO		O repositório dispõe de um programa de promoção dos seus serviços/benefícios.	Comunicar aos usuários reais e potenciais do valor da gestão de dados.

PADRÕES BOAS PRÁTICAS	CLASSE	REQUISITO	PROPÓSITO
	PADRÕES COMUNITÁRIOS/ DISCIPLINARES	O repositório alinha seus procedimentos, fluxos de trabalho e boas práticas aos padrões comunitários e/ou disciplinares .	Aumentar o nível de interoperabilidade entre repositórios disciplinares
	PRINCÍPIOS FAIR	O repositório atende totalmente/ parcialmente aos princípios FAIR .	Aumentar o nível de encontrabilidade, acessibilidade, interoperabilidade e reuso .
	OAIS (norma ISO 16363)	O repositório está alinhado à norma OAIS no âmbito do arquivamento de longo prazo.	Garantir estabilidade (integridade, fixidez e autenticidade) aos dados de valor contínuo.
	COAR	O repositório é aderente ao manual de boas práticas do COAR .	Alinhar o repositório com as características/requisitos preconizados internacionalmente .
	INTEROPERABILIDADE	O repositório disponibiliza API aberta .	Facilitar o acesso aos metadados .
		O repositório disponibiliza metadados via protocolo OAI-PMH .	Permitir análises e serviços por terceiros.
		O repositório faz harvesting de metadados em outras fontes usando o protocolo OAI-PMH.	Permitir coleta automática (harvesting) por provedores de serviço (agregadores).
		O repositório lista os atributos comuns.	Agregar informações de outras fontes de dados.
	REGISTROS EM AGREGADORES	O repositório está registrado no Re3data .	Facilitar a interoperabilidade com outros serviços .
		O repositório está indexado em outros dispositivos (Open Doar, Google ...).	Proporcionar visibilidade internacional e encorajar o reuso .
	GESTÃO DE METADADOS	O repositório utiliza padrões de qualidade para metadados.	Aumentar a possibilidade do repositório ser encontrado.
		O repositório empacota metadados no padrão METS .	Exportar de metadados e arquivamentos de pacotes AIP para preservação.
		O repositório assinala metadados legíveis e/ou acionáveis por máquina .	Permitir a exploração por agentes computacionais .

	CITAÇÃO	O repositório oferece/formula citação padronizada no padrão DataCite .	Garantir que os créditos sejam atribuídos corretamente de maneira que possibilite o acesso direto ao conjunto de dados utilizado.
	ANOTAÇÃO	O repositório permite que usuários emitem comentários em forma padronizada	Aumentar o nível de colaboração e de enriquecimento dos registros .
	CURADORIA/GARANTIA DE QUALIDADE	O repositório disponibiliza ou facilita a curadoria especializada e garantia de qualidade.	Melhorar a precisão e integridade dos conjuntos de dados e metadados.
	PLANO DE PRESERVAÇÃO	O repositório elabora/adota planos de preservação compatíveis com as normas e boas práticas correntes e que contemple os critérios da comunidade disciplinar e que incluem os acervos digitais e físicos (amostras, herbários, manequins etc.)	Assegurar a persistência temporal dos dados de valor contínuo com as qualidades estabelecidas pela comunidade.
	TREINAMENTO DE USUÁRIO	O repositório elabora/adota guias, manuais ou cartilhas sobre princípios básicos de gestão.	Conscientizar o pesquisador sobre a importância da gestão de dados .
		O relatório oferece treinamento em gestão de dados para o pesquisador (PGD, organização, proteção etc.)	Inserir a gestão de dados no ciclo de pesquisa .
	FORMATO COMUM	O repositório permite que os datasets e metadados sejam acessados, baixados ou exportados do repositório em formatos amplamente usados , de preferência não proprietários, consistentes com os padrões usados nas disciplinas que o repositório atende.	Atender os requisitos disciplinares.
	NORMAS/PADRÕES DISCIPLINARES	O repositório utiliza/compatibiliza em seus diversos procedimentos os padrões das comunidades disciplinares.	Aderir o repositório a cultura disciplinar de coleta, processamento, compartilhamento e reuso do domínio disciplinar.
	RESTRIÇÕES DE ACESSO E USO	O repositório tem regras documentadas e implementa restrições de acesso tipificada por “ tempo de embargo ”.	Garantir que o acesso aos conjuntos de dados respeite as autorizações de acesso e o prazo de embargo
		O repositório emprega procedimentos documentados para restringir acesso e uso a comunidades específicas – por exemplo, no contexto de pesquisa sobre uma determinada doença.	Garantir quem poderá ter acesso a determinados conjuntos de dados
	USABILIDADE	O repositório adota padrões de usabilidade (interfaces, níveis de conhecimento exigido, documentação, ajuda online etc).	Facilitar a consulta e depósito.
	ORIENTAÇÃO CLARA DE ACESSO E USO	O repositório oferece informação para o usuário sobre seus serviços.	Garantir que os usuários encontrem facilmente na interface dos repositórios quais serviços são ofertados e como eles funcionam.

		O repositório assegura que os datasets são acompanhados por documentação descrevendo os termos de acesso e uso pertinentes.	Informar ao usuário as possibilidades e extensão da exploração dos dados.
		O repositório dispõe de procedimentos para endereçar violações de termos de uso.	Prevenir violações e mal uso dos datasets e do repositório.

SERVIÇOS INFORMACIONAIS	CLASSE	REQUISITO	PROpósito
	IDENTIFICAÇÃO PERSISTENTE	O repositório aplica identificador persistente às coleções de dados (DOI, Handles, URN etc.).	Apoiar a descoberta de dados, contextualização (linkage), citação, arquivamento de longo prazo e métricas de uso.
		O repositório aplica identificador aos pesquisadores (ORCID, etc).	Aumentar o nível de identificação dos dados ; e fazer citações mais precisas.
	METADADOS	O repositório garante que os datasets são acompanhados de metadados ricos e de qualidade que descrevem os termos de reuso, licenças, processos científicos disciplinares, contexto, proveniência, dependências técnicas etc.	Assegurar que os dados sejam encontrados, acessados, reusados, citados e contextualizados.
		O repositório opera serviços de gestão de metadados .	Garantir/assegurar/avaliar a qualidade dos metadados e dos vocabulários controlados .
		O repositório desenvolve/aplica esquemas disciplinares de metadados, ontologias e vocabulários .	Garantir níveis de representação compatíveis com o domínio disciplinar.
	PROVENIÊNCIA	O repositório tem mecanismos terminológicos e de representação para registrar a origem, cadeia de custódia, controle de versão e outras modificações submetidas aos datasets e seus metadados.	Assegurar a proveniência dos datasets.
	LINKING/CONTEXTUALIZAÇÃO	O repositório tem procedimentos para fazer o linkage dos dados com outros recursos relacionados (artigos, projetos, outros datasets, modelos, instrumentos etc).	Estabelecer o contexto dos datasets .
	CATALOGAÇÃO	Têm/desenvolvem manuais de referência para catalogação descritiva dos dados.	Melhorar a descrição e encontrabilidade dos dados e controlar as entradas na catalogação/indexação.
	EXPORTAÇÃO/IMPORTAÇÃO ONTOLOGIA	O repositório oferece ferramenta que permite a exportação dos metadados	Facilitar a exportação dos metadados e resultados da pesquisa
	CITAÇÃO REFERÊNCIA PADRONIZADA	O repositório dispõe de serviço de geração de citação em formatos padronizados , por exemplo, Datacite.	Aumentar o nível de citação e reuso dos dados.
	ANOTAÇÃO	O repositório oferece mecanismos para que sejam inseridas contribuições (preferencialmente padronizadas) de outros pesquisadores aos dados.	Apoiar a dinâmica de cooperação entre pesquisadores e demais stakeholders.

	TREINAMENTO DE USUÁRIO/ MATERIAL DIDÁTICO	O repositório elabora material didático (guias, tutoriais, manuais) sobre gestão de dados. O repositório oferece cursos sobre PGD – Plano de Gestão de Dados .	Conscientizar o pesquisador sobre a importância da gestão de dados; para inserir a gestão de dados no ciclo da pesquisa. Alinhar os projetos de pesquisa as diretrizes dos financiadores, editores e da organização.
	RECUPERAR/ACESSAR DADOS	O repositório oferece recursos avançados para formulação de buscas (consulta à tesouro, busca facetada etc.). Dispõe de interfaces de busca que atendam à requisitos comunitários/ disciplinares.	Apoiar a formulação de buscas com maior capacidade de recuperação . Facilitar formulação de buscas com maior capacidade de expressão disciplinar e granularidade
		Permite buscas integradas com outros sistemas.	Buscar transversalmente por domínios diferentes .
	REVISÃO	O repositório dispõe de serviços de revisão/registros/recursos e seus respectivos metadados .	Certificar da acurácia e garantir recuperação, visibilidade e acesso aos recursos.
	BALCÃO DE REFERÊNCIA DE DADOS	O repositório disponibiliza material que auxilia os usuários a encontrar os conjuntos de dados	Facilitar a encontrabilidade de conjuntos de dados
	PRESERVAÇÃO (GESTÃO DE LONGO PRAZO)	O repositório possui política de preservação	Garantir o ciclo de vida dos conjuntos de dados a longo prazo
	APOIO A PUBLICAÇÃO DE DADOS	O repositório permite a publicação de dados e o versionamento.	Permitir o reúso e a reprodutibilidade.
	APOIO AO DESENVOLVIMENTO DE DATASETS	O repositório disponibiliza material que auxilia o pesquisador a organizar os datasets	Garantir que os datasets disponibilizados estejam organizados e reutilizáveis

SERVIÇOS COMPUTACIONAIS	CLASSE	REQUISITO	PROpósito
SISTEMA DE STORAGE	SEGURANÇA DOS DADOS	O repositório conta com infraestrutura de armazenamento de dados escaláveis de acordo com os níveis de uso.	Permitir a oferta de serviços adequados.
		O repositório faz uso de infraestrutura externas de armazenamento .	Permitir a ampliação da capacidade de armazenamento do repositório.
		O repositório implementa dispositivos tecnológicos para proteger os dados durante os processos de pesquisa, como sistemas de backups, segurança física e de rede .	Proteger os dados gerados por pesquisas em andamento .
		O repositório implementa procedimentos documentados aderentes aos critérios de cibersegurança para garantir a integridade dos dados apropriados ao seu nível de sensibilidade (checksum, hash, criptografia, etc).	Prevenir acessos não autorizados e modificações que comprometam a integridade e autenticidade dos dados.

	PROTEÇÃO DE DADOS SENSÍVEIS	O repositório implementa dispositivos para proteção dos dados sensíveis (proteção em camadas, credenciamento de usuários, acesso off-line, etc.).	Garantir a confidencialidade, privacidade, sigilo e proteger os dados sensíveis de acessos inadequados.
	TREINAMENTO DE USUÁRIOS	O repositório oferece treinamento aos pesquisadores nas ferramentas computacionais de análise (visualização, estatísticas, etc.).	Apresentar os dados referentes aos objetos digitais.
	FERRAMENTAS DE ANÁLISE/VIZUALIZAÇÃO	O repositório oferece ferramentas para análise e visualização .	Permitir que os usuários tenham acesso aos conjuntos de dados sem a utilização de ferramentas externas
	FERRAMENTAS ESTATÍSTICAS	O repositório utiliza alguma ferramenta de visualização de estatísticas de uso	Auxiliar na tomada de decisões
NORMALIZAÇÃO DE FORMATOS		O repositório dispõe de uma lista de formatos de arquivos aceitos , preferencialmente abertos, não proprietários, padronizados e independentes de software.	Facilitar a gestão corrente e o arquivamento de longo prazo .
		O repositório oferece ferramentas e workflow para transformação (reformatação) dos formatos de dados em formatos padronizados pelo repositório.	Facilitar a construção de pacotes OAIS/SIP .
	DESTRUIÇÃO SEGURA	O repositório tem procedimentos tecnológicos documentados para destruição segura de coleção de dados sensíveis .	Garantir a privacidade, sigilo e confidencialidade.

SERVIÇOS CIENTÍFICOS	CLASSE	REQUISITO	PROPÓSITO
AVALIAÇÃO/REVISÃO		Identificação se a avaliação por pares pode ser realizada de forma integrada à plataforma do repositório.	Facilitar o depósito dos artigos no repositório após receber avaliação por pares positiva.
		O repositório possibilita o acesso reservado e antecipado aos dados para apoiar o processo de revisão por pares dos artigos.	Certificar sobre a existência, completude e qualidade dos dados que apoiam os resultados.
	LIMPEZA DE DADOS	O repositório tem funcionalidades que contribuem com o controle de qualidade de dados, como checagem de congruência entre dados e metadados, alerta de arquivos duplicados e visualização DDI.	Contribuir com o controle de qualidade dos dados.
	SERVIÇO DE ANÁLISE/VISUALIZAÇÃO	Verificar se o repositório possui ferramentas para análise e visualização dos dados via interface direta	Possibilitar que os usuários não precisem utilizar ferramentas externas específicas para visualizar os dados

	TREINAMENTO DE USUÁRIOS	Disponibilidade de guias, manuais, tutoriais e profissionais qualificados para treinamento sobre a gestão e uso do repositório.	Apoiar os usuários a melhorar a gestão, a busca e o uso da informação; possibilitar amplo uso das fontes de informação disponíveis.
	EMPACOTAMENTO DOS DADOS	O repositório organiza os arquivos dos conjuntos de dados em pacotes de formato conhecido por sua comunidade alvo	Facilitar o download, desempacotamento e reutilização dos conjuntos de dados
	ANOTAÇÃO/COMENTÁRIOS	O repositório permite a inserção de anotações / comentários dos utilizadores dos conjuntos de dados / leitores das publicações	Ter feedback da comunidade na reutilização dos conjuntos de dados / leitura das publicações
	ESPAÇO DE COLABORAÇÃO	O repositório apresenta campo para troca de informações/feedback entre os pesquisadores.	Proporcionar à ampliação de estudos e pesquisas com diversificação de perspectivas teóricas e metodológicas.
	IDENTIFICAÇÃO DE REUSO	O repositório possui aplicações que permitem rastreio/identificação de citações de reuso dos conjuntos de dados. O repositório emprega vocabulário controlado e especificações mínimas.	Apoiar processos de avaliação e estimular o depósito, reconhecimento e recompensa dos pesquisadores. Possibilitar que os dados de pesquisa sejam reutilizáveis entre disciplinas.

SERVIÇOS ADMINISTRATIVOS	CLASSE	REQUISITO	PROpósito
	MÉTRICAS DE USO	O repositório controla e audita os acessos, download e reuso dos dataset.	Apoiar a gestão do repositório e avaliar o atendimento ao público-alvo .
		O repositório mantém disponível claramente os registros de acesso, download e reuso dos datasets.	Aumentar o grau de transparência do repositório para os depositantes de dados .
	POLÍTICA DE CUSTO	O repositório deve ter uma política institucional que assegure o suporte/orçamento financeiro para garantir sua integridade e sustentabilidade.	Viabilizar a gestão do repositório incluindo o próprio sistema, o grupo de trabalho (técnicos e especialistas) e infra-estrutura existente e necessária.
	RESTRIÇÕES	VOLUME DE DADOS	Assegurar o controle da qualidade dos metadados e das informações inseridas de forma a atender aos padrões e requisitos estabelecidos para a entrada dos registros.
		FORMATOS NÍVEL DE QUALIDADE	

			Organizar e manter a integridade dos objetos digitais produzidos pelos autores das instituições.
	COPYRIGHT		Estabelecer licenças dos autores para autorização que concedam ao repositório, permissão para depositar, reunir, divulgar e preservar os metadados e objetos digitais. Produzir documentos importantes sobre a propriedade intelectual que possam formalizar as autorizações dos autores, visando evitar demandas legais onerosas para as instituições.
TEMPO DE EMBARGO	O repositório adota políticas de embargo		Permitir que trabalhos que estão em processo de obtenção de patente e ou similar possam ser depositados
DIVULGAÇÃO/DISSEMINAÇÃO	O repositório deve garantir que as informações produzidas nas instituições de pesquisa possam ser disponibilizadas para a comunidade acadêmica e científica, como também para toda a sociedade.		Permitir a visibilidade dos trabalhos disponibilizados e obter um maior aumento na média de citações e do impacto dos resultados das pesquisas. Institucionalizar a produção intelectual produzida, reunida em um único local. Divulgar informações importantes e relevantes para outros pesquisadores e para a sociedade, como por exemplo, as campanhas de saúde que podem estar vinculadas ao repositório. Promover o reuso dos dados e da pesquisa científica e participar do Movimento Acesso Aberto.
CONTATO COM O USUÁRIO	O repositório adota alguma ferramenta que possibilite contato direto entre o usuário e os responsáveis pelo repositório		Garantir que o usuário tenha acesso direto e facilitado a equipe responsável pelo repositório para que possa sanar dúvidas
REGISTRO E AUTENTICAÇÃO DE USUÁRIOS	O repositório adota dispositivo de autenticação de depositantes de dados e de outros usuários dos serviços.		Garantir o acesso de acordo com as políticas de acesso e uso da plataforma.