

INFRAESTRUTURA E SERVIÇOS DE GESTÃO DE DADOS DE PESQUISA: UMA AVALIAÇÃO NO ÂMBITO DA UNIVERSIDADE ESTADUAL PAULISTA (Unesp)

Research data management infrastructure and services: an assessment within the scope of São Paulo State University (Unesp)

Emanuelle Torino

Doutoranda em Ciência da Informação. Programa de Pós-Graduação em Ciência da Informação, Universidade Estadual Paulista (UNESP), Marília, Brasil
Universidade Tecnológica Federal do Paraná (UTFPR), Curitiba, Brasil
Universidade Estadual Paulista - Unesp, Programa de Pós-Graduação em Ciência da Informação, Marília, Brasil
emanuelle@utfpr.edu.br

<https://orcid.org/0000-0002-3791-9884> 

Gustavo Lunardelli Trevisan

Doutorando em Ciência da Informação. Programa de Pós-Graduação em Ciência da Informação, Universidade Estadual Paulista - Unesp, Marília, Brasil
Universidade Estadual Paulista - Unesp, Programa de Pós-Graduação em Ciência da Informação, Marília, Brasil
g.trevisan@unesp.br

<https://orcid.org/0000-0002-4175-7910> 

Jean Fernandes Brito


Doutorando em Ciência da Informação. Programa de Pós-Graduação em Ciência da Informação, Universidade Estadual Paulista (UNESP), Marília, Brasil
Universidade Estadual Paulista - Unesp, Programa de Pós-Graduação em Ciência da Informação, Marília, Brasil
j.brito@unesp.br

<https://orcid.org/0000-0002-9258-8205> 

Silvana Aparecida Borsetti Gregorio Vidotti

Doutora em Educação. Programa de Pós-Graduação em Ciência da Informação, Universidade Estadual Paulista (UNESP), Marília, Brasil
Professor Assistente Doutor MS3.2, Regime de Dedicção Integral à Docência e à Pesquisa. Universidade Estadual Paulista - Unesp, Departamento de Ciência da Informação. Programa de Pós-Graduação em Ciência da Informação, Marília, Brasil
silvana.vidotti@unesp.br

<https://orcid.org/0000-0002-4216-0374> 

A lista completa com informações dos autores está no final do artigo 

RESUMO

Objetivo: Compreender o processo de autoavaliação da infraestrutura e oferta de serviços de gestão de dados de pesquisa, por meio do diagnóstico da Universidade Estadual Paulista utilizando a ferramenta de autoavaliação *Research Infrastructure Self-Evaluation*.

Método: A pesquisa de cunho exploratória, aplicada e de natureza qualitativa, que se constituiu na autoavaliação realizada junto à Universidade Estadual Paulista, com o uso do *Research Infrastructure Self-Evaluation*, gerando um *spider chart*, visando avaliar a infraestrutura e a oferta de serviços de gestão de dados de pesquisa. A coleta de dados consistiu na aplicação do *Research Infrastructure Self-Evaluation*, tendo como resultado o mapeamento das práticas de gestão de dados de pesquisa na Instituição.

Resultado: A partir do mapeamento foi possível a análise dos pontos fortes e necessidades de melhoria da infraestrutura de apoio à gestão de dados de pesquisa na Universidade Estadual Paulista, possibilitando à Instituição verificar aspectos que demandam investimentos, visando aperfeiçoar a infraestrutura e os serviços de apoio ao pesquisador.

Conclusões: A partir da autoavaliação e da análise dos aspectos relacionados à gestão de dados de pesquisa, foi possível identificar os pontos fortes, dentre os quais se destaca a atenção institucional para políticas que, embora generalista, pode ser aplicada à gestão de dados de pesquisa. Por outro lado, para que a infraestrutura e os serviços de gestão de dados de pesquisa sejam adequadamente estruturados é necessário um olhar global da Universidade Estadual Paulista para todos os aspectos que compõem a gestão dos dados de pesquisa, conforme estruturado no modelo *Research Infrastructure Self-Evaluation* e, por meio do resultado da autoavaliação investir no preenchimento de lacunas ora existentes e saná-las em benefício do pesquisador, dos seus resultados de pesquisa, da comunidade científica, da sociedade e da própria instituição.

PALAVRAS-CHAVE: Dados de Pesquisa. Gestão de Dados de Pesquisa. Research Infrastructure Self-Evaluation. RISE.

ABSTRACT

Objective: Understand the process of infrastructure self-assessment and the provision of research data management services, through the diagnosis of the São Paulo State University, using the Research Infrastructure Self-Evaluation Framework.

Methods: Exploratory, applied and qualitative research, which consisted of the self-assessment carried out at the São Paulo State University, using the Research Infrastructure Self-Evaluation, generating a spider chart, in order to assess the infrastructure and the offer of services of research data management. Data collection consisted of applying the Research Infrastructure Self-Evaluation, resulting in the mapping of research data management practices in the Institution.

Results: Based on the mapping, it was possible to analyze the strengths and needs to improve the infrastructure to support research data management at São Paulo State University, enabling the Institution to verify aspects that require investments, in order to improve the infrastructure and support services for researchers.

Conclusions: Starting from the self-assessment and analysis of aspects related to research data management, it was possible to identify the points highlighted, among which institutional attention to policies stands out, which, although general, can be applied to research data management. However, for the infrastructure and research data management services to be properly structured, it is necessary to have a global look by São Paulo State University for all aspects that make up the management of research data, as structured in the *Research Infrastructure Self-Evaluation* model and through the As a result of self-assessment, invest in filling existing gaps and remedying them for the benefit of the researcher, their research results, the scientific community, society and the institution itself.

KEYWORDS: Research Data. Research Data Management. Research Infrastructure Self-Evaluation. RISE.

1 INTRODUÇÃO

Fazer e comunicar ciência são atividades rotineiras de pesquisadores em todo o mundo. Contudo, discussões recentes no bojo da Ciência Aberta trazem à tona processos e práticas realizadas pelos pesquisadores, dentre as quais, neste estudo, enfatiza-se a gestão e a disponibilização de dados de pesquisa.

A necessidade de gestão e disponibilização de dados de pesquisa pode se constituir em uma das respostas às demandas sociais aplicadas à ciência e aos dados de pesquisa, sua transparência, reuso, validação e replicação. Fato que requer que sejam propostos e consolidados novos modelos de infraestruturas e serviços específicos visando o suporte às pesquisas e aos pesquisadores, culminando em formas otimizadas de gerenciamento, disponibilização e acesso aos dados de pesquisa.

Nesse sentido, instituições acadêmicas, de pesquisa e agências de fomento em todo o mundo estão trabalhando na disponibilização de diretrizes, políticas e infraestruturas que possam apoiar o pesquisador na atividade de pesquisa, bem como em ações que otimizem os processos de tratamento, disponibilização, recuperação e uso dos dados de pesquisa gerados. O que torna de interesse comum a todos os campos da ciência, a criação e a padronização de modelos para o tratamento, o uso e o reuso de dados de pesquisa. Para tanto, o envolvimento interdisciplinar não é mais interpretado como facultativo, mas condição básica para a gestão de dados de pesquisa (GDP).

No Brasil, a oferta de serviços e infraestruturas de apoio à gestão de dados se apresentam como um desafio, assim, as instituições iniciaram as atividades com a implantação de repositórios de dados de pesquisa, a definição de políticas, e, aos poucos, atuam no suporte aos pesquisadores na tarefa de planejar as diferentes etapas que compõem o plano de gestão e o ciclo de vida dos dados de pesquisa.

Assim, é importante que as instituições possam analisar se as ações planejadas e realizadas estão consonantes com as necessidades de apoio à atividade de pesquisa.

Desta forma, com o intuito de compreender o processo de autoavaliação da infraestrutura e oferta de serviços de gestão de dados de pesquisa, o presente estudo realizou o diagnóstico da infraestrutura e oferta de serviços de gestão de dados de pesquisa na Universidade Estadual Paulista (Unesp), utilizando a ferramenta de autoavaliação *Research Infrastructure Self-Evaluation* (RISE).

A metodologia utilizada na pesquisa apresenta características de natureza qualitativa, do tipo exploratória e aplicada, utilizando-se como técnicas de coleta e análise de dados a aplicação da autoavaliação da infraestrutura e da oferta de serviços de apoio ao pesquisador por meio da ferramenta *Research Infrastructure Self-Evaluation* (RISE). Para o embasamento teórico foram realizados levantamentos bibliográficos sobre os temas pesquisados, bem como o levantamento dos documentos oficiais da Universidade pesquisada acerca da gestão de dados de pesquisa e acesso aberto. Ainda, foi realizada a autoavaliação RISE por um gestor da Universidade analisada. A análise dos resultados da autoavaliação foi realizada com base nos elementos do RISE e sob o olhar da literatura pertinente.

Considerando que a autoavaliação RISE busca retratar a capacidade institucional na oferta da infraestrutura e serviços de apoio à gestão de dados de pesquisa, é necessário que seja realizada de forma gerencial, buscando avaliar os aspectos institucionais de forma geral. Assim, o avaliador, acessa a ferramenta e tem acesso às áreas e recursos a serem autoavaliados e, com base nas políticas, serviços e infraestrutura institucional indica o nível em que a instituição se encontra em cada um dos recursos avaliados.

Assim, a coleta de dados do presente estudo foi realizada em junho de 2021 e se constituiu na autoavaliação RISE, junto à Universidade Estadual Paulista (Unesp), com o envolvimento do Comitê Gestor de Acesso Aberto da Unesp, responsável pela supervisão e acompanhamento do acesso Aberto na Unesp, tendo como resultado o mapeamento das práticas de gestão de dados de pesquisa na Instituição, exibidas por meio de três *spider*

chart, que retratam a avaliação da infraestrutura e da oferta de serviços de gestão de dados de pesquisa na Unesp.

Vale destacar que o referido Comitê Gestor é o órgão responsável pelo acesso aberto na Unesp, o que inclui a infraestrutura e os serviços de gestão de dados de pesquisa, e a autoavaliação RISE propiciará à Instituição perceber quais os seus pontos fortes e fracos no que tange à gestão de dados de pesquisa.

Após o processo de autoavaliação as respostas do avaliador foram reunidas para a construção da análise e discussão dos resultados de acordo com o objetivo proposto. Esse mapeamento possibilitou a análise dos pontos fortes e das necessidades de melhoria na infraestrutura e nos serviços de apoio à gestão de dados de pesquisa na Unesp, possibilitando à Instituição verificar aspectos que demandam investimentos, visando aperfeiçoar a infraestrutura e os serviços de apoio ao pesquisador.

2 GESTÃO DE DADOS DE PESQUISA

No contexto da pesquisa científica, a comunicação dos resultados se consolidou ao longo dos anos por meio dos periódicos científicos, tidos como os principais canais formais de comunicação científica. Torino, Roa-Martinez e Vidotti (2020) destacam que as discussões recentes acerca da Ciência Aberta deram luz à abertura dos dados de pesquisa, cuja gestão e armazenamento eram comumente realizados pelos próprios pesquisadores, culminando em uma relevância histórica menor a estes objetos de pesquisa.

Aos poucos este cenário está sendo alterado e, pesquisadores, impulsionados por agências de fomento, fontes de publicação e instituições de ensino e pesquisa, passam a adotar práticas de gestão, armazenamento e disponibilização de dados de pesquisa.

Para a *Organisation for Economic Co-Operation and Development* (OECD, 2007, p. 13, tradução nossa) “dados de pesquisa são definidos como registros factuais (números, textos, imagens e sons) utilizados como fontes primárias de pesquisa científica, e que são comumente aceitos na comunidade científica para validar os resultados da pesquisa.”. Ao que esclarecem Sales e Sayão (2019, p. 36) “Dado de pesquisa é todo e qualquer tipo de registro coletado, observado, gerado ou usado pela pesquisa científica, tratado e aceito como necessário para validar os resultados da pesquisa pela comunidade científica.”

O ciclo de vida dos dados de pesquisa, cujas etapas compreendem: planejar, coletar, assegurar a qualidade, descrever, preservar, descobrir, integrar e analisar (STRASSER,

2012), em geral, é superior ao tempo do projeto ao qual estão vinculados, do tempo de financiamento, quando disponível, e ainda do próprio desenvolvimento da pesquisa. Dessa forma, é possível que estes dados sejam utilizados pelos pesquisadores que podem gerar novas análises sobre eles ou ainda adicionar novos elementos, além de fornecer bases para outras pesquisas e pesquisadores. (SAYÃO; SALES, 2015).

Assim, “no contexto da ciência contemporânea, os dados de pesquisa deixam de ser meros subprodutos das atividades de pesquisa e ressurgem como protagonistas na busca por novos conhecimentos” (SALES; SAYÃO, 2019, p. 31). Segundo os autores, os dados de pesquisa, quando gerenciados adequadamente tornam-se uma fonte importante para outras pesquisas.

Para a OECD (2007, p. 9, tradução nossa) “Os bancos de dados científicos estão rapidamente se tornando uma parte crucial da infraestrutura do sistema científico global.”. A gestão e a disponibilização dos dados de pesquisa auxiliam na otimização do investimento público realizado em pesquisa, possibilita criar cadeias de valor de inovação e aprimora a cooperação nacional e internacional.

No Brasil, a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) foi a primeira agência de fomento a reconhecer a relevância da gestão dos dados de pesquisa e

[...] reconhece a importância da gestão adequada dos dados de pesquisa como parte essencial das boas práticas de pesquisa. Para tanto, considera necessário que os dados resultantes de projetos financiados pela Fundação sejam gerenciados e compartilhados de forma a garantir o maior benefício possível para o avanço científico e tecnológico. (FAPESP, 2020).

Neste sentido, destaca que a gestão adequada dos dados beneficia a reprodutibilidade da pesquisa, possibilita o reuso e o compartilhamento e, com isso, o avanço de novas pesquisas. Assim, o plano de gestão de dados tornou-se um documento obrigatório para os pesquisadores que desejam submeter projetos para pleitear fomento junto à FAPESP.

O *research data management* (RDM), cuja tradução para o português é gestão dos dados de pesquisa (DGP), “pode ser definido como um conjunto de práticas para lidar com as informações coletadas e criadas durante a pesquisa” (HIGMAN; BANGERT; JONES, 2019, tradução nossa). Embora o planejamento dos dados de pesquisa possa parecer uma burocratização da atividade científica, estabelecer, no início do projeto, as formas de coleta, organização, documentação, armazenamento, preservação, direitos de uso, entre outros, guiará o processo da pesquisa e as condutas do pesquisador, trazendo ganhos durante a

execução da pesquisa e posteriormente, por meio da facilidade de localização, reuso e citação. Esta ação, realizada de forma sistemática, atendendo às boas práticas e especificidades do domínio dos dados e do pesquisador que conduziu o processo, é que possibilitam aos dados de pesquisa uma formalização capaz de elevar seu reconhecimento como objeto científico, que pode ser localizável, acessível, interoperável e reutilizável.

De acordo com a OECD (2007), o compartilhamento e a disponibilização pública, em acesso aberto, dos dados de pesquisa, além de maximizar o potencial das pesquisas e seus resultados, também possibilita maiores retornos aos investimentos públicos realizados em pesquisas.

Inegavelmente esta ação tomada por uma agência de fomento, no caso a FAPESP, alterou a perspectiva dos pesquisadores do Estado de São Paulo diante da necessidade de gerenciar e disponibilizar seus dados de pesquisa e, de igual maneira, levou as instituições de ensino e pesquisa a se posicionarem no sentido de oferecer ações de apoio ao pesquisador, nas condutas relacionadas ao planejamento e gestão dos dados de pesquisa, bem como na definição de ambientes para o seu armazenamento e disponibilização.

Marín-Arraiza e Vidotti (2019) afirmam que instituições e grupos de pesquisas percebem a necessidade de contar com profissionais capacitados para atuar nestas demandas. Neste sentido, propõem um conjunto de três fases para a implementação de serviços institucionais de gestão de dados de pesquisa, constituídos por: elaboração de política de gestão de dados de pesquisa, estabelecimento de unidade de informação e integração de especialistas em gestão de dados de pesquisa.

É possível assim, vislumbrar que a consolidação da alteração nos processos de planejamento, coleta, gestão, armazenamento, disseminação e recuperação de dados de pesquisa podem culminar na melhoria do seu reconhecimento como objeto resultante das atividades de pesquisa, o que requer conhecimentos específicos do domínio da pesquisa e do processo de gestão de dados de pesquisa. Para tanto, o apoio institucional, no sentido de fornecer suporte ao pesquisador nas suas demandas é imprescindível, ao mesmo tempo em que requer uma alteração na forma de atuação das instituições para que a infraestrutura e os serviços ofertados sejam especializados e condizentes com as demandas do pesquisador, do domínio da pesquisa e da comunidade científica.

Neste sentido, é importante que as instituições que disponibilizam serviços de apoio à gestão de dados de pesquisa, possam avaliar constantemente a infraestrutura disponível, visando a melhoria dos processos.

3 RESEARCH INFRASTRUCTURE SELF-EVALUATION

O *Research Infrastructure Self-Evaluation* (RISE) é uma estrutura de autoavaliação desenvolvida na Universidade de Potsdam da Alemanha e disponibilizada pelo *Digital Curation Center* (DCC) (RANS; WHYTE 2017). O RISE tem como proposta fornecer às instituições e aos gestores de dados de pesquisa um diagnóstico do apoio à gestão de dados de pesquisa que pode auxiliar no desenvolvimento de estratégias institucionais. Por isso, segundo Rans e Whyte (2017) o RISE deve ser conduzido por profissional interno à instituição, com conhecimento profundo da infraestrutura e suporte à pesquisa da instituição além de domínio do processo de gestão de dados de pesquisa (GDP).

O RISE utiliza uma estrutura de autoavaliação, incorporando em seus processos as melhores práticas e/ou aperfeiçoando seus próprios métodos (*benchmarking*), numa perspectiva própria para a instituição. Desse modo, é possível direcionar ações futuras à existência de uma infraestrutura mínima para que os envolvidos no processo disponham de forma organizada e estruturada suas contribuições em todos os aspectos da gestão, sempre observados numa perspectiva holística.

Para Rans e Whyte (2017), algumas questões são relevantes no que tange aos propósitos como suporte ao planejamento dos serviços de apoio institucional em seu arcabouço conceitual. Os autores destacam a necessidade da integração entre membros e processos para o *design* de serviço, visto que essa prática pode reduzir riscos e auxiliar no alcance dos resultados, ao contrário de quando os elementos integrantes de gestão de dados de pesquisa são tratados de forma isolada.

O RISE também oferece aos gestores de dados de pesquisa indicativos que fornecem referências necessárias para que o direcionamento dos esforços e as discussões do desenvolvimento de serviços não sejam direcionados ao uso de tecnologias e produtos específicos de forma precipitada, uma vez que, para os não familiarizados com as novas tecnologias, isso poderia se constituir em um fator dissonante à proposta conceitual.

A estrutura de autoavaliação traz subsídios que possibilitam a análise dos elementos pertinentes à gestão de dados de pesquisa, no intuito de contribuir com a otimização e a profissionalização dos processos, bem como com políticas, serviços e infraestrutura da instituição (SPARC EUROPE, 2020). Ainda, a autoavaliação por meio do RISE, tem como

instrumento para aplicação um *framework* para que os gestores designados por suas instituições avaliem os diversos aspectos implicados à gestão de dados de pesquisa.

O processo de autoavaliação por meio da ferramenta RISE consiste na resposta a uma espécie de questionário (Anexo A), constituído por 10 áreas de serviço de suporte de dados de pesquisa, a saber: *research data management policy and strategy/política e estratégia de gestão de dados de pesquisa*, *business plans and sustainability/planos de negócios e sustentabilidade*, *advisory services/serviços de consultoria*, *training/capacitação*, *data management planning/planejamento de gerenciamento de dados*, *active data management/gerenciamento ativo de dados*, *appraisal and risk assessment/avaliação e avaliação de risco*, *preservation/preservação*, *access and publishing/acesso e publicação*, e *discovery/descoberta*.

Essas áreas são subdivididas em 22 recursos (conforme Anexo A) que permitem avaliar a infraestrutura e serviços de apoio à gestão de dados de pesquisa na instituição, por meio da escolha de um dos três níveis de atendimento ao recurso avaliado. Cada recurso permite a indicação de uma das três alternativas possíveis, consideradas três níveis de avaliação (nível 1, 2 ou 3) que indicam a 'capacidade' da Instituição avaliada de agregar um valor ao serviço. Desse modo, por padrão, determina-se que: o nível 1 retrata a conformidade da instituição ao recurso avaliado, o nível 2 indica que a instituição fornece serviços personalizados ao seu contexto e o nível 3 considera que a instituição é referência no fornecimento do recurso avaliado. Desta forma, quando selecionado o nível 3, os demais níveis estão contemplados pela instituição. Observa-se, no entanto que o RISE não conta com uma alternativa que retrate a ausência de atividade de suporte institucional ao recurso avaliado, sendo assim, a 'capacidade' é descrita como estando no nível 0 (RANS; WHYTE, 2017).

A aplicação do RISE é realizada por meio de aplicativos ou de ferramentas que auxiliam a autoavaliação. Neste estudo, a ferramenta utilizada foi a disponibilizada pela SPARC Europe.

4 AVALIAÇÃO E DISCUSSÃO DOS RESULTADOS

O presente estudo realizou a autoavaliação *Research Infrastructure Self-Evaluation* (RISE) por meio de uma ferramenta disponibilizada pelo SPARC *Europe*¹ para avaliar a oferta de diferentes aspectos relacionados à gestão de dados de pesquisa, com a participação de um representante do Comitê Gestor de Acesso Aberto da Unesp, o que gera credibilidade ao processo, que deve envolver profissional interno à Instituição (RANS; WHYTE, 2017).

Como resultados foram gerados três gráficos² que sintetizam a análise da infraestrutura e a oferta de serviços de gestão de dados de pesquisa na Unesp e permitem compreender a capacidade Institucional em fornecê-los.

O Gráfico 1 objetiva apresentar a oferta e a gestão de dados de pesquisa, sinalizando a variedade de políticas, suporte e serviços que a instituição oferece. No caso da Unesp, o gráfico evidencia a presença de política e estratégia de gestão de dados de pesquisa e o gerenciamento ativo de dados, fator decorrente do processo de implantação da Política de Acesso Aberto à Produção Intelectual da Unesp, aprovada pelo Conselho Universitário, por meio da Resolução Unesp-8, de 4-2-2021³.

Gráfico 1 – Oferta e Gestão de Dados de Pesquisa na Unesp



Fonte: Dados da pesquisa (2021).

¹ Disponível em: <https://sparceurope.org/evaluate-your-rdm-offering/>.

² A representação gráfica da análise RISE apresenta quantidade de raios diferente entre os três gráficos apresentados. O que se deve ao peso que a ferramenta atribui para cada resposta assinalada na autoavaliação. O peso de cada alternativa em relação ao elemento avaliado tem relação com o nível RISE avaliado. Assim, quando da autoavaliação, quanto mais distante do eixo do gráfico a instituição esteja, maior a quantidade de raios apresentada.

³ Disponível em: https://bibliotecaunespsjbv.files.wordpress.com/2021/02/pg_0091.pdf.

A discussão e recente aprovação da Política de Acesso Aberto à Produção Intelectual da Unesp justifica a inexpressividade dos demais serviços, tendo em vista o momento de planejamento de um conjunto de ações que serão decorrentes.

O Gráfico 2 é utilizado para representar o grau em que a instituição adapta a gestão de dados de pesquisa às necessidades de seus usuários. No que tange à Unesp, reforçando o que foi constatado no Gráfico 1, o Gráfico 2 evidencia a presença de política e estratégia de gestão de dados de pesquisa, o gerenciamento de dados ativos e o acesso e publicação. Sendo o último aspecto relacionado à disponibilidade de ambiente informacional para a disponibilização e ampla disseminação dos dados de pesquisa proveniente das investigações realizadas na Instituição, o Repositório Institucional da Unesp. Embora haja convergência com a análise anterior, é possível verificar que a representação gráfica da personalização dos serviços possui ramificação menor, mantendo todos os aspectos mais próximos do elemento central.

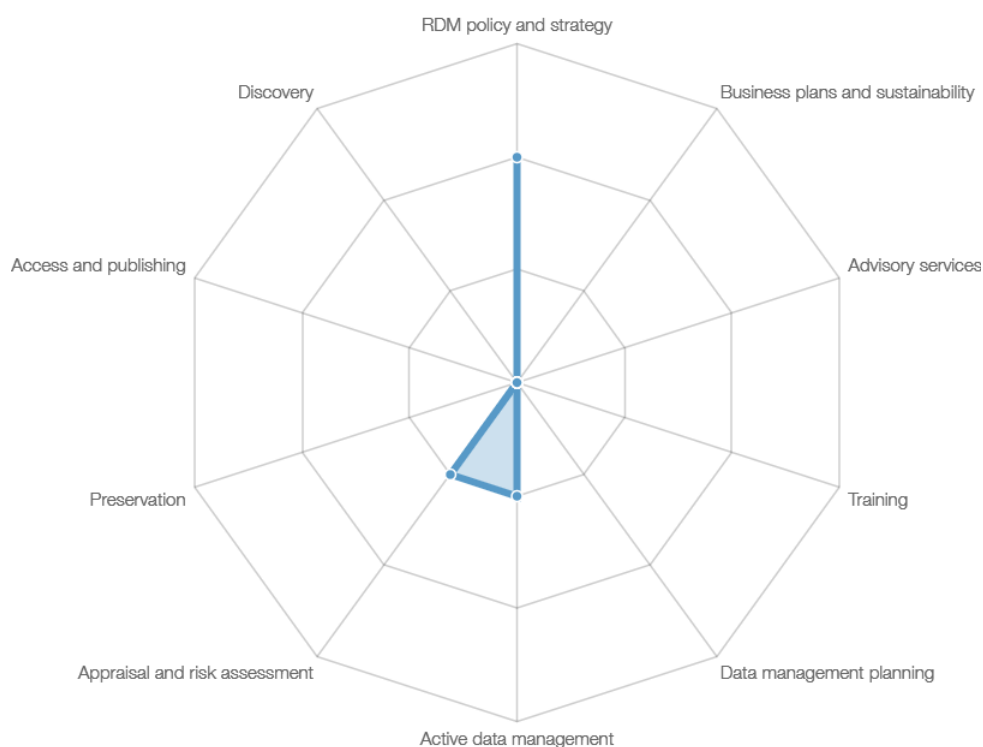
Gráfico 2 – Serviços Personalizados de Gestão de Dados de Pesquisa na Unesp



Fonte: Dados da pesquisa (2021).

O Gráfico 3 é utilizado para indicar o quanto a instituição é inovadora na gestão de dados de pesquisa. Assim, no Gráfico 3, o destaque da Unesp na definição da Política de Acesso Aberto à Produção Intelectual reforça a vanguarda da Instituição em aspectos relacionados à gestão da produção intelectual, o que já ocorria no Repositório Institucional da Unesp com a produção acadêmico-científica e passa a ser realizado com os dados de pesquisa.

Gráfico 3 – Inovação na Gestão de Dados de Pesquisa na Unesp



Fonte: Dados da pesquisa (2021).

De forma similar, os Gráficos 2 e 3 possuem todos os aspectos avaliados muito próximos ao eixo do gráfico, o que é claramente diferente do Gráfico 1, cuja representação se expande e distancia do centro. A análise dessa divergência aponta para o fato de que, nesta etapa, há uma expressividade no investimento da Unesp na oferta de infraestrutura e serviços de gestão de dados de pesquisa, o que pode ser considerado o primeiro passo para alicerçar ações de personalização para então tornar-se inovadora nos diferentes recursos avaliados pelo RISE.

A partir dos resultados da autoavaliação RISE aplicada na Unesp, representados nos Gráficos 1, 2 e 3, os diferentes aspectos relacionados à infraestrutura e à oferta de serviços de gestão de dados de pesquisa serão discutidos, por área de avaliação RISE, nas próximas seções.

4.1 Política e Estratégia de Gestão de Dados de Pesquisa

A área designada '*research data management policy and strategy*/política e estratégia de gestão de dados de pesquisa' está dividida em três pontos: 'desenvolvimento

de políticas’, ‘conscientização e engajamento das partes interessadas’ e ‘roteiro de implementação de gestão de dados de pesquisa’.

As universidades e centros de pesquisa devem repensar de maneira holística a estrutura de sua gestão de dados. “Para que uma instituição possa aproveitar esta janela de oportunidade e elaborar uma política de gestão de dados, pode começar pela análise e redigir uma política do zero ou adaptar modelos para esse fim” (MARIN-ARRAÍZA; VIDOTTI, 2019, p. 267). Além disso, bibliotecários, cientistas de dados e profissionais da computação podem auxiliar no desenvolvimento dessas políticas.

A Política de Acesso Aberto à Produção Intelectual da Unesp não deixa explícita claramente a questão da gestão dos dados de pesquisa, mas vale salientar que, a página web Acesso Aberto Unesp⁴ disponibiliza informações e ferramentas que demonstram a preocupação da Universidade em pensar e investigar os dados de pesquisa, o que se faz envolvendo outros canais e processos informacionais.

A autoavaliação RISE da Unesp evidencia a implantação da Política de Acesso Aberto à Produção Intelectual da Unesp é um processo recente e que necessita de alteração da cultura institucional, por meio de passos estratégicos para a efetiva estruturação e gestão dos dados de pesquisa.

Nessa discussão, salientam-se os aportes da gestão de dados e a inclusão de profissionais que atuam com esses dados e o papel que eles exercem no desenvolvimento de produtos e serviços para estratégias, preservação, uso e reuso dos dados de pesquisa geridos, repensando o desenvolvimento da política institucional de estratégia de gestão de dados de pesquisa, que inclui, dentre outros, orientações sobre plano de gestão de dados, capacitação de pessoal, planos de negócios, sustentabilidade.

A autoavaliação RISE da Unesp nesta área, indica o nível 2 para o ‘desenvolvimento de políticas’, indica o nível 3 para a ‘conscientização e engajamento das partes interessadas’ e o nível 3 para o ‘roteiro de implementação de gestão de dados de pesquisa’, podendo, a instituição ser considerada referência no fornecimento dos recursos avaliados.

4.2 Planos de Negócio e Sustentabilidade

⁴ Disponível em: <https://www2.unesp.br/portal#!/acessoaberto/>.

A área '*business plans and sustainability*/planos de negócios e sustentabilidade' se refere a 'investimento de equipe', 'investimento em tecnologia' e 'modelagem de custos'.

Refletir sobre os dados de pesquisa à luz da gestão de negócios é [...] “contribuir para a eficiência na gestão, oportunizando tomar decisões acertadas, confiáveis e que mostre os caminhos a serem trilhados” (SANTOS; SILVA, 2012, p. 50).

No âmbito da Unesp podemos citar o investimento da universidade em pessoal capacitado para gerir e trazer os melhores caminhos, quais sejam com o uso das tecnologias, consultorias e capacitações. Destaca-se, neste sentido, que a Unesp investe na formação contínua do quadro de bibliotecários para que possam atuar no auxílio aos pesquisadores no processo de gestão e disponibilização de dados de pesquisa.

A sustentabilidade nesse bojo é a visão macro, a possibilidade de desenvolvimento econômico de uma forma sustentável (SANTOS; SILVA, 2012). Se o modelo de negócios não for sustentável, a instituição deve criar um modelo novo, abrangendo as novas possibilidade de atuação e garantir os recursos exigidos por uma gestão estratégica (BHIDE, 2002). Além disso, a junção e parcerias com as Pró-Reitorias de Graduação, Pós-Graduação, Pesquisa, Planejamento e Gestão Estratégica da universidade podem ser uma alternativa para o encaminhamento dessa demanda.

Nesse sentido, além do entendimento de que o acesso aos dados de pesquisa devem ser democratizados e disponíveis à sociedade, os dados da pesquisa realizada apontam que 'o serviço Gestão de Dados de Pesquisa é entregue dividindo as responsabilidades entre os funcionários existentes, não havendo o redesenho das funções da equipe nas atribuições destinadas à Gestão de Dados de Pesquisa', uma vez que a estrutura da instituição é bastante enxuta no que se refere à disponibilidade de recursos.

Quanto aos investimentos em tecnologia, registra-se que a Unesp faz a gestão de dados de pesquisa, considerando os recursos previstos em orçamento do projeto conforme Plano de Desenvolvimento Institucional (PDI).

A autoavaliação RISE da Unesp nesta área, indica nível 1 para 'investimento de equipe', nível 2 para 'investimento em tecnologia' e nível 1 para 'modelagem de custos', que consiste na conformidade da instituição no fornecimento dos recursos avaliados.

4.3 Serviços de Consultoria

A área destinada aos '*advisory services*/serviços de consultoria' consiste no suporte para o pesquisador em todos os aspectos relacionados à gestão de dados de pesquisa, como a necessidade de conhecimento das normas, diretrizes e políticas da instituição, da área de domínio do pesquisador e das agências de fomento à pesquisa. Esta área apresenta como recurso 'serviços de consultoria'.

É pertinente considerar que os profissionais e pesquisadores de todos os campos da ciência têm se engajado no intuito de buscar novos conhecimentos no que se refere ao gerenciamento de dados para suporte à gestão de dados de pesquisa (AKERS; DOTY, 2013). As compreensões das diferenças disciplinares e das necessidades em campos de domínio específicos trazem à tona algumas particularidades, revelando importantes distinções disciplinares em ações de gerenciamento de dados de pesquisa, e, por conseguinte, provocando ações pontuais em serviços de suporte.

O processo construtivo à gestão de dados de pesquisa requer tempo, conhecimento daqueles que se pretende atender e maturidade dos participantes para que o modelo de gestão contemple todos os quesitos necessários.

Nesse sentido, conforme avaliação realizada, na Unesp o suporte aos envolvidos no processo de gestão de dados de pesquisa ocorre por meio de orientação on-line, abordando as principais áreas de gestão de dados de pesquisa, no intuito de que os envolvidos possam atuar no processo de forma genérica, autoinstrucional ou individualizada, o que é detalhado no momento da capacitação.

Assim, a autoavaliação RISE da Unesp nesta área, indica nível 1 para 'serviços de consultoria', que consiste na conformidade da instituição no fornecimento do recurso avaliado.

4.4 Capacitação

A área RISE designada '*training/capacitação*' avalia a oferta de 'capacitação on-line' e 'capacitação presencial'.

Nesse sentido, a necessidade de capacitação aos pesquisadores e envolvidos na gestão de dados de pesquisa é parte dos desdobramentos do processo, pois proporciona uma maior autonomia quanto à forma de operacionalização e instrumentalização da atividade do pesquisador, pois cada instituição, organização, país e projeto são individuais, com diferentes abordagens relacionadas ao gerenciamento de dados de pesquisa.

Nessa perspectiva os serviços de capacitação devem considerar que:

As atividades de capacitação são um complemento importante para as ferramentas e orientações de gerenciamento de dados. Existem recursos de treinamento em uma variedade de formatos para ajudar os pesquisadores a preparar o gerenciamento de dados de pesquisa com sucesso e, mais importante, a implementar esses planos. (JONES et al., 2020, tradução nossa).

Sob essa perspectiva, o investimento em capacitação dos envolvidos na instituição avaliada é possível tanto na forma presencial, quanto na modalidade on-line. Conforme os dados da avaliação, observam-se a escolha da Unesp pela modalidade on-line devido à sua dispersão geográfica e por se constituir por diversas áreas do conhecimento, corroborados pela plasticidade da produção dos conteúdos destinados à orientação dos pesquisadores e facilidade de disponibilização nos ambientes informacionais digitais. As capacitações são realizadas sob demanda, não descartando, quando necessário, a modalidade presencial.

A autoavaliação RISE da Unesp nesta área, indica nível 2 para ‘capacitação on-line’ e nível 1 para ‘capacitação presencial’, que consiste na conformidade da instituição no fornecimento dos recursos avaliados, personalizando-os ao contexto institucional.

4.5 Planejamento de Gerenciamento de Dados

No que tange à área ‘*data management planning*/planejamento de gerenciamento de dados’, a Unesp, por sua vinculação ao Governo do Estado de São Paulo, tem como uma das agências de fomento a Fapesp que, conforme explicitado anteriormente, foi pioneira na obrigatoriedade da apresentação do plano de gestão de dados, o que deu à Fapesp o encargo burocrático da exigência.

Por outro lado, visando instrumentalizar o pesquisador, a Unesp rapidamente disponibilizou aos pesquisadores informações que visam apoiá-los nessa ação. Essas informações, mais recentemente foram incorporadas à página ‘Acesso Aberto Unesp’ em menu próprio designado ‘Plano de Gestão de Dados’⁵. A partir desse ponto central de apoio on-line ao pesquisador, estão disponíveis definições, informações, materiais próprios e de diferentes autores, links para acesso às agências de fomento, além de ferramentas para a elaboração de plano de gestão de dados, a exemplo da DMPTool, DMPonline e PGDonline.

⁵ Disponível em: <https://www2.unesp.br/portal#!/acessoaberto/plano-de-gestao-de-dados/>.

Destacam-se ainda facilidades como links de acesso a modelos de planos de gestão de dados, modelos já implementados pela Unesp nas ferramentas disponíveis e o acesso a elas utilizando as credenciais institucionais.

Quanto à gestão de dados de pesquisa, vale destacar que a própria Fapesp também disponibiliza em sua página informações de interesse do pesquisador⁶ no momento do planejamento da sua pesquisa.

De acordo com Whyte e Tedds (2011) a gestão adequada dos dados de pesquisa possibilita sinergia entre os diferentes grupos de pesquisa da instituição quer seja na criação de conhecimentos, no envolvimento de diferentes atores, bem como na potencialização do (re)uso de dados de pesquisa nas atividades de ensino, pesquisa, extensão, definição de políticas, entre outras.

Neste aspecto, considerando os dados da avaliação, a Unesp fornece orientação aos pesquisadores sobre a elaboração de planos de gestão de dados exigidos pelo financiador. Apesar disso, nos Gráficos 1, 2 e 3 é visível que essa ação ainda não é expressiva, e, embora tenha caráter inovador no Brasil, merece destaque e investimento por parte da Instituição, visando sobretudo generalizar a prática entre os pesquisadores para que ocorra de forma sistemática e internalizada, independente da cobrança de agentes externos.

A Política de Acesso Aberto da Unesp é omissa em aspectos relacionados à gestão de dados de pesquisa, embora seja possível inferir que sejam cobertos a partir dos seus princípios gerais que estabelece no item VI “A Política tem caráter mandatório para toda produção intelectual, que deverá ser depositada no Repositório Institucional Unesp, respeitando-se o período de embargo, quando couber, resguardando-se aos autores os direitos morais e usos comerciais de suas obras.”.

Recomenda-se que a Unesp inclua em sua Política de Acesso Aberto, de forma explícita, a gestão dos dados de pesquisa, objetivando institucionalizar a necessidade de planejamento e disponibilização desses dados. Tal sugestão reforça a atuação do Repositório Institucional da Unesp⁷ que disponibiliza informações acerca do autoarquivamento de planos de gestão de dados e dados de pesquisa.

⁶ Disponível em: <https://fapesp.br/gestaodedados>.

⁷ Disponível em: <https://repositorio.unesp.br/>.

A autoavaliação RISE da Unesp nesta área, indica nível 1 para 'planejamento de gerenciamento de dados', que consiste na conformidade da instituição no fornecimento do recurso.

4.6 Gerenciamento Ativo de Dados

A área '*active data management*/gerenciamento ativo de dados' está relacionada às características técnicas e à infraestrutura de tecnologia da informação, sobretudo, quanto à oferta de espaço de armazenamento de dados de pesquisa para pesquisadores da instituição e seus colaboradores externos, bem como o gerenciamento de segurança. Nesse sentido, avalia 'escalabilidade e sincronização', 'suporte de colaboração' e 'gerenciamento de segurança'.

Segundo dados da autoavaliação RISE, a Unesp oferece serviços importantes, como espaço de armazenamento de dados de pesquisa, com possibilidade de ampliação de espaço e desempenho mediante solicitação. Além disso, permite que os dados de pesquisa armazenados em ambientes digitais sejam compartilhados e estejam acessíveis a pesquisadores colaboradores da pesquisa externos à Instituição. Contudo, o gerenciamento se faz por meio do acesso autenticado ao espaço de armazenamento, visando proteção contra acesso não autorizado aos dados de pesquisa.

Embora este aspecto não se relacione com o armazenamento de longo prazo, vale destacar que a Política de Acesso Aberto à Produção Intelectual da Unesp institucionaliza e reforça o papel do Repositório Institucional da Unesp, cuja gestão é realizada pela Coordenadoria Geral de Bibliotecas, de forma compartilhada com outras instâncias Institucionais. Determina ainda que este ambiente deverá ter capacidade de integração com sistemas nacionais e internacionais, o que pode contribuir com necessidades do pesquisador, quer seja no momento da disponibilização dos dados de pesquisa ou ainda para fins de ampliação de visibilidade.

Considerando que a Política de Acesso Aberto da Unesp é omissa com relação aos dados de pesquisa, não foi possível identificar questões específicas ao gerenciamento ativo de dados. Contudo, a inclusão da gestão de objetos cujos autores tenham vínculo com a Universidade, ainda que em parceria com autores externos, reforça o espírito colaborativo próprio do desenvolvimento das pesquisas.

No que tange aos espaços de armazenamento dos dados de pesquisa durante a sua execução, recomenda-se a inclusão, na área de 'Recursos' disponível na página Acesso

Aberto Unesp, de informações sobre os espaços institucionais de armazenamento de dados de pesquisa disponíveis ao pesquisador.

Considerando a análise RISE efetuada na Unesp, reafirma-se novamente que a Política de Acesso Aberto da Unesp explicita aspectos relacionados à gestão, armazenamento e disponibilização de dados de pesquisa.

A autoavaliação RISE da Unesp nesta área, indica nível 2 para ‘escalabilidade e sincronização’, nível 3 para ‘suporte de colaboração’ e nível 1 para ‘gerenciamento de segurança’, que representam, respectivamente, a conformidade da instituição no fornecimento do recurso, a customização da oferta do recurso ao contexto institucional e que a instituição é referência no recurso avaliado.

4.7 Avaliação de Risco

A ‘*appraisal and risk assessment*/avaliação e avaliação de risco’ se caracteriza por aspectos concernentes às definições de suporte à coleta dos dados de pesquisa para disponibilização e entrada no ciclo de vida de dados em um ambiente institucional. Possui como recursos ‘política de coleta de dados’, ‘avaliação de risco de segurança, legal e ético’ e ‘coleta de metadados para informar a tomada de decisão’.

Na Unesp, a definição dos critérios para a inclusão de coleções para conjuntos de dados no Repositório Institucional da Unesp é prerrogativa do Grupo Gestor do Repositório Institucional da Unesp, assegura atender às necessidades dos pesquisadores em consonância à Política de Acesso Aberto da Unesp.

No que tange aos aspectos éticos e legais, na Unesp a responsabilidade é mantida a cargo do pesquisador que deve assegurar que os dados de pesquisa foram gerados, criados ou coletados seguindo critérios éticos e legais cabíveis.

Os aspectos éticos e legais que envolvem pesquisas estão regidos por instrumentos jurídicos que podem ser divergentes nas diferentes jurisdições, por se fundamentarem nas bases legais do país. Neste sentido, no Brasil, alguns dispositivos devem ser considerados, em função do seu relacionamento com o planejamento e desenvolvimento de pesquisas.

No que tange aos procedimentos éticos, destacam-se a Resolução nº 466, de 12 de dezembro de 2012 (BRASIL, 2012) que trata de “diretrizes e normas regulamentadoras de pesquisas envolvendo seres humanos”; e a Resolução nº 510, de 7 de abril de 2016 (BRASIL, 2016), que “dispõe sobre as normas aplicáveis a pesquisas em Ciências Humanas e Sociais cujos procedimentos metodológicos envolvam a utilização de dados

diretamente obtidos com os participantes ou de informações identificáveis ou que possam acarretar riscos maiores do que os existentes na vida cotidiana, na forma definida nesta Resolução”, ambas relacionadas à pesquisas com seres humanos.

Há ainda a necessidade de atentar-se à Lei Geral de Proteção de Dados Pessoais (LGPD) que “dispõe sobre o tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural.” (BRASIL, 2018), sobretudo quando do armazenamento de dados pessoais.

No Brasil esses dispositivos legais podem ser tidos como condutores de diferentes aspectos relacionados à coleta, armazenamento e disponibilização de dados de pesquisa durante as diferentes etapas do ciclo de vida da pesquisa, cuja negligência no momento do planejamento pode colocar em risco o trabalho desenvolvido, o que reforça a necessidade de planejar adequadamente a execução da pesquisa e a gestão de dados de pesquisa.

Vale destacar ainda a Lei de Direitos Autorais (BRASIL, 1998) que determina aspectos legais acerca da proteção e as formas de aproveitamento das obras intelectuais. Neste dispositivo legal, diferentes objetos são mencionados, considerando as tecnologias da época, dentre os quais as bases de dados, quando sua constituição puder ser tida como uma criação intelectual nova. Dado o tempo em que a lei está em vigor, de forma inalterada, não há menção direta a formas de tratamento a dados de pesquisa, nem sempre criados por humanos, mas sim coletados e ou gerenciados por eles, o que pode se constituir em uma lacuna para discussão.

Em meio a essas discussões, as características e limitações impostas pelo *copyright* cederam espaço para o desenvolvimento de formas criativas de colaboração, a partir do estabelecimento de licenças de uso concedidas pelos autores ou titulares de direitos autorais da obra, visando explicitar ao leitor os usos previamente autorizados da sua produção. Neste sentido, a forma mais comumente utilizada para o licenciamento é o *Creative Commons*⁸, que estabelece instrumentos jurídicos gratuitos legíveis por humanos e aplicações computacionais. A versão 4.0 das licenças é a mais atual e considerada de acreditação internacional, em função do seu alinhamento com as leis de direitos autorais de diferentes jurisdições. Possibilita ainda, o licenciamento de bases de dados, quando protegidas por direitos autorais.

⁸ Disponível em: <https://creativecommons.org/licenses/by-nc-sa/4.0/deed.pt>.

Tendo em vista a discussão realizada, é importante que a instituição disponibilize serviços de suporte especializado ao pesquisador em todos os meandros éticos e legais que podem estar relacionados à sua pesquisa, visando o auxílio e a oferta de informações especializadas que podem ser cruciais para o desenvolvimento e posterior divulgação dos diferentes objetos gerados.

Neste sentido, a Unesp, no momento da coleta dos dados para armazenamento no Repositório Institucional, entende que as exigências éticas e legais que envolvem a pesquisa são etapas anteriores à divulgação e, portanto, responsabilidade do pesquisador.

A autoavaliação RISE da Unesp nesta área, indica nível 3 para 'política de coleta de dados', nível 1 para 'avaliação de risco de segurança, legal e ético' e nível 1 para 'coleta de metadados para informar a tomada de decisão', que representam que a instituição é referência no recurso avaliado e a conformidade da instituição no fornecimento do recurso.

4.8 Preservação

A área de '*preservation*/preservação' é categorizado nos recursos 'planejamento e ação de preservação' e 'suporte de continuidade', apresentando questões fundamentais para estruturação e garantia efetiva da preservação na instituição.

Os suportes, dispositivos e softwares utilizados para armazenamento tornam-se obsoletos cada vez mais rápido. Nesse sentido, esses objetos digitais necessitam de uma representação efetiva para auxiliar na preservação e autenticidade das informações e registros descritivos.

Pensar em preservação digital exige refletir sobre aplicação dos metadados nesse contexto, principalmente os de preservação digital. Os metadados estão relacionados com a preservação digital, pois são eles que resolvem um dos problemas de longo prazo, pois conseguem identificar um conjunto de dados e informações, fazendo com que fiquem armazenados de forma adequada para futura busca e acesso (SAYÃO, 2010).

Para Hedstrom (1998, p. 190, tradução nossa), a preservação digital é “[...] o planejamento, alocação de recursos e aplicação de métodos e tecnologias para assegurar que a informação digital de valor contínuo permaneça acessível e utilizável.”. Ao que complementa Márdero Arellano (2004, p. 17) “A preservação digital compreende os mecanismos que permitem o armazenamento em repositórios de dados digitais que garantiriam a perenidade dos seus conteúdos.”

Para minimizar o risco de perder os objetos digitais, a instituição deve-se utilizar *software*, *hardware*, e métodos de preservação com qualidade e que armazenem as informações e os dados de maneira adequada para permanecer em longo prazo, utilizando-se padrões de metadados.

Neste sentido, a Unesp realiza algumas ações de preservação digital, conduzida pela Comissão Permanente de Preservação Digital⁹ e embasada em uma política¹⁰.

A autoavaliação RISE da Unesp nesta área, indica nível 1 para 'planejamento e ação de preservação' e nível 2 para 'suporte de continuidade', que consistem na conformidade da instituição no fornecimento do recurso e na customização da oferta do recurso ao contexto institucional.

4.9 Acesso e Publicação

No âmbito do serviço 'access and publishing/acesso e publicação', os recursos são: 'monitoramento de conjuntos de dados produzidos localmente', 'mandato de publicação de dados' e 'nível de Curadoria de dados'.

No que diz respeito ao ponto de monitoração de conjunto de dados produzidos na Unesp, os dados da avaliação realizada indicam que as informações acerca das formas de acesso e publicação de dados de pesquisa são coletadas dos projetos de pesquisa para permitir a conformidade com os requisitos das agências de fomento. No que concerne ao mandato de publicação de dados, o Repositório Institucional da Unesp oferece suporte aos padrões de melhores práticas da comunidade para acesso a dados, citações e interoperabilidade de metadados. E por fim, sobre o nível de curadoria de dados, se compromete a manter ou aumentar o valor por meio de ações de rotina ou nas coletas de dados.

A Unesp mantém, desde o ano de 2013, o Repositório Institucional da Unesp, que reúne, armazena e dissemina a produção intelectual da Universidade, ao qual estão sendo incorporados os dados de pesquisa, por meio do autoarquivamento realizado pelo pesquisador, visando contribuir com a gestão e a visibilidade dos dados de pesquisa. Destaca-se a indexação do Repositório Institucional da Unesp em diferentes mecanismos

⁹ Disponível em: <https://www2.unesp.br/portal#!/cppd>.

¹⁰ Disponível em: <https://www2.unesp.br/portal#!/noticia/21084/politica-de-preservacao-digital-de-documentos/>.

de busca, acadêmicos e gerais, bem como a presença expressiva em *rankings* específicos, o que reforça sua reputação como ambiente de disponibilização dos dados de pesquisa.

Nesse sentido, considerando o tempo de implantação da Política de Acesso Aberto da Unesp e a característica recente do início da disponibilização dos planos de gestão de dados e dos dados de pesquisa no Repositório Institucional da Unesp, a gestão e a curadoria de dados de pesquisa no âmbito da Unesp está em processo de estruturação, visando instrumentalizar os pesquisadores e demais envolvidos no processo de gestão e disponibilização de dados, para que a disponibilização e o acesso aos dados de pesquisa ocorram de forma efetiva. Sayão e Sales (2016, p. 92) afirmam que para a [...] “execução dos processos de gestão é necessário um arcabouço tecnológico e gerencial que compreenda todo o ciclo de vida dos dados”.

No que tange à publicação dos dados de pesquisa Torino, Roa-Martinez e Vidotti (2020) discutem uma questão conceitual relevante, que consiste nas divergências entre o processo de publicação e disponibilização de dados de pesquisa, no que se refere sobretudo à formalização, que possibilita benefícios aos múltiplos envolvidos. Vale ressaltar, nesse sentido, que a disponibilização pública de um conjunto de dados de pesquisa, ainda que em repositório digital não assegura que os dados sejam reutilizados, sobretudo em razão da disponibilização não assegurar que os dados sejam contextualizados, o que pode prejudicar o potencial reuso. Ao passo que a publicação de dados de pesquisa, envolve a documentação dos dados, que auxilia no processo de formalização necessário à interpretação, compreensão e reuso, o que possibilitará aos dados de pesquisa aumento de reconhecimento como objeto resultante das atividades de pesquisa.

Assim, de acordo com Torino, Roa-Martinez e Vidotti (2020), o processo de publicação dos dados de pesquisa, aumenta seu potencial de reuso, visto que a documentação que o acompanha esclarece elementos contextuais relevantes para os consumidores de dados.

A autoavaliação RISE da Unesp nesta área, indica nível 1 para ‘monitoramento de conjuntos de dados produzidos localmente’, nível 2 para ‘mandato de publicação de dados’ e nível 2 para ‘nível de Curadoria de dados’, que consistem na conformidade da instituição no fornecimento do recurso e na customização da oferta do recurso ao contexto institucional.

4.10 Descoberta



Por fim, o último serviço analisado diz respeito à ‘*discovery*/descoberta’, tendo como recurso ‘escopo e catalogação de metadados’. Os dados da avaliação realizada apontam que, a Unesp faz ‘o uso de Metadados de catálogos de serviços para conjuntos de dados financiados publicamente pela organização, de acordo com as expectativas do financiador de que eles podem ser descobertos, citados e vinculados a conteúdo relacionado.’

Destaca-se que o processo de descoberta dos dados de pesquisa está diretamente relacionado à representação no repositório, considerando as necessidades de exaustividade e granularidade, bem como o uso de metadados do domínio dos dados de pesquisa representados, visando favorecer a descoberta.

Os metadados são comumente utilizados em bibliotecas e arquivos, sendo que, nesses ambientes possuem funções além da identificação e descrição dos documentos. Nesse sentido, para Gilliland (2008, tradução nossa), metadados:

- certificam a autenticidade e o grau de integridade do conteúdo;
- estabelecem e documentam o contexto do conteúdo;
- identificam e exploram as relações estruturais que existem dentro e entre os objetos de informação;
- fornecem uma gama de pontos de acesso intelectual para uma gama cada vez mais diversificada de usuários; e
- apresentam algumas das informações que um profissional da informação pode ter fornecido em um ambiente tradicional de referência pessoal ou de pesquisa.

No contexto deste estudo, os metadados descrevem e representam os recursos informacionais de acordo com as expectativas do financiador, para que possam ser descobertos, citados e vinculados a conteúdo relacionado.

Além dos metadados, os identificadores persistentes auxiliam no processo de desambiguação, descoberta, citação e vinculação de conjuntos de dados de pesquisa, bem como o seus relacionamentos com pessoas e instituições. Nesse sentido, o identificador persistente inicialmente mais difundido foi o *Digital Object Identifier* (DOI), massivamente utilizado para artigos de periódicos e que, atualmente está sendo atribuído para outros objetos, a exemplo dos dados de pesquisa.

Posteriormente pesquisadores, instituições e agências de fomento adotam como prática o uso de identificador persistente para pessoas, visando dentre outras coisas, desambiguar homônimos e realizar relacionamentos unívocos entre pesquisadores e suas publicações, tem-se como identificador mais utilizado o *Open Researcher and Contributor ID* (ORCID iD). E, como boa prática, é crucial que as instituições atentem-se para a

necessidade de desambiguar seu nome e utilizar identificadores persistentes a exemplo do *Global Research Identifier Database* (GRID).

Nesse contexto a Unesp tem utilizado os identificadores persistentes mencionados, contudo, o Repositório Institucional da Unesp utiliza como identificador persistente para objetos o Handle, inclusive para conjuntos de dados de pesquisa. Destaca-se o caráter inovador da Unesp na afiliação à ORCID como uma das primeiras instituições brasileiras, o que permite, dentre outras coisas, alimentar automaticamente o perfil do pesquisador no ORCID iD, a partir da produção intelectual disponibilizada no Repositório Institucional da Unesp.

A autoavaliação RISE da Unesp nesta área, indica nível 1 para 'escopo e catalogação de metadados', que consiste na conformidade da instituição no fornecimento do recurso.

5 CONSIDERAÇÕES FINAIS

Tendo em vista o intuito de compreender a aplicação da autoavaliação *Research Infrastructure Self-Evaluation* (RISE), que consiste em uma ferramenta que auxilia no diagnóstico do desenvolvimento de infraestrutura e serviços de gestão de dados de pesquisa e fornece estrutura para discussão e contribuição em todos os aspectos relacionados ao serviço, sua aplicação no contexto da Unesp, com a participação de um gestor vinculado à Universidade e que atua diretamente no Comitê de Acesso Aberto, possibilitou emergir no cenário da gestão de dados de pesquisa da Universidade.

A partir da autoavaliação e da análise dos aspectos relacionados à gestão de dados de pesquisa, foi possível identificar os pontos fortes, dentre os quais se destaca a atenção institucional para políticas que, embora generalista, pode ser aplicada à gestão de dados de pesquisa.

Desta forma, dos 22 recursos presentes na autoavaliação RISE, conforme análise realizada nas seções 4.1 a 4.10, a Unesp se caracteriza como referência no fornecimento de 5 recursos (nível 3), enquanto fornece serviços personalizados ao seu contexto em 5 recursos (nível 2) e possui conformidade na avaliação de 12 recursos avaliados (nível 1). Destaca-se que, como mencionado anteriormente, na autoavaliação RISE, o nível 1 indica que a instituição está em conformidade no fornecimento do recurso, o nível 2 indica que a instituição está em conformidade no fornecimento do recurso e o customiza ao contexto institucional e o nível 3 indica que a instituição está em conformidade no fornecimento do

recurso, o customiza ao contexto institucional e, com isso é considerada referência no fornecimento do recurso.

Diante do exposto, é relevante destacar que a Unesp está em conformidade com a oferta dos 22 recursos avaliados pelo RISE, sendo que, em 5 deles customiza o recurso ao contexto institucional e, em outros 5 além da customização, é referência na oferta do recurso.

Assim, para que a infraestrutura e os serviços de gestão de dados de pesquisa sejam adequadamente estruturados é necessário um olhar global da Unesp para todos os aspectos que compõem a gestão dos dados de pesquisa, conforme estruturado no modelo RISE e, a partir do resultado da autoavaliação investir no preenchimento de lacunas ora existentes e saná-las em benefício do pesquisador, dos seus resultados de pesquisa, da comunidade científica, da sociedade e da própria instituição.

Por fim, destacam-se nessa discussão os aportes da gestão de dados de pesquisa e inclusão de profissionais e todos os serviços analisados por meio do RISE, na ecologia de pesquisa da Unesp e o papel que exercem no desenvolvimento de produtos e serviços para estratégias, preservação e (re)uso dos dados de pesquisas.

Assim, recomenda-se que as instituições que ofertam infraestrutura e serviços de apoio à gestão de dados de pesquisa utilizem estruturas de autoavaliação, a exemplo do RISE, como forma de diagnóstico institucional, uma vez que permite analisar os pontos fortes e aqueles que necessitam de investimento, e com isso otimizar o processo.

Neste sentido, podem ser realizados trabalhos futuros aplicando a autoavaliação RISE a outras instituições; com a aplicação de outras ferramentas de autoavaliação; e ainda com estudos comparativos de avaliação em diferentes ferramentas.

Com isso, espera-se que este estudo contribua com os gestores da Instituição avaliada, na apresentação e discussão dos resultados apresentados na autoavaliação, bem como a partir das proposições indicadas. E ainda, que haja contribuição a demais instituições que atuem no contexto da gestão de dados de pesquisa, para que possam realizar autoavaliações da sua infraestrutura e dos serviços de gestão de dados de pesquisa ofertados.

REFERÊNCIAS

AKERS, K. G.; DOTY, J. Research data management practices and perspectives: Differences among the arts and humanities, social sciences, medical sciences, and basic

sciences. **International Journal of Digital Curation**, v. 8, n. 2, p. 5-26, 2013. Disponível em: <https://doi.org/10.2218/ijdc.v8i2.263>. Acesso em: 20 nov. 2020.

BHIDE, A. Como os empreendedores constroem estratégias que dão certo. *In*: Harvard Business Review. **Empreendedorismo e estratégia**. Rio de Janeiro: Campus, 2002.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). **Diário Oficial da União**, Brasília, DF, 15 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/ato2015-2018/2018/lei/L13709.htm. Acesso em: 17 nov. 2020.

BRASIL. Lei nº 9.610, de 19 de fevereiro de 1998. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. **Diário Oficial da União**, Brasília, DF, 20 fev. 1998. Seção 1. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/L9610.htm. Acesso em: 17 nov. 2020.

BRASIL. Ministério da Saúde. **Resolução nº 466, de 12 de dezembro de 2012**. Aprovar as seguintes diretrizes e normas regulamentadoras de pesquisas envolvendo seres humanos. Disponível em: https://bvsms.saude.gov.br/bvs/saudelegis/cns/2013/res0466_12_12_2012.html. Acesso em: 27 nov. 2020.

BRASIL. Ministério da Saúde. **Resolução nº 510, de 07 de abril de 2016**. Esta Resolução dispõe sobre as normas aplicáveis a pesquisas em Ciências Humanas e Sociais cujos procedimentos metodológicos envolvam a utilização de dados diretamente obtidos com os participantes ou de informações identificáveis ou que possam acarretar riscos maiores do que os existentes na vida cotidiana, na forma definida nesta Resolução. Disponível em: https://www.in.gov.br/materia/-/asset_publisher/KujrW0TZC2Mb/content/id/22917581. Acesso em: 27 nov. 2020.

HEDSTROM, M. Digital preservation: a time bomb for digital libraries. **Computers and the Humanities**, Michigan, U.S.A, n. 31, p. 189–202, 1998. Disponível em: <http://deepblue.lib.umich.edu/bitstream/handle/2027.42/42573/?sequence=1>. Acesso em: 28 nov. 2020.

FAPESP. **Gestão de Dados**. Disponível em: <https://fapesp.br/gestaodedados>. Acesso em: 23 nov. 2020.

GILLILAND, A. J. Setting the stage. *In*: BACA, M. (ed.). **Introduction to metadata**. Los Angeles: Getty, 2008. p. 1-19. Disponível em: <https://www.getty.edu/publications/intrometadata/setting-the-stage/>. Acesso em: 19 nov. 2020.

HIGMAN, R.; BANGERT, D.; JONES, S. Three camps, one destination: the intersections of research data management, FAIR and Open. **Insights**, v. 32, n. 1, p. 18. 2019. DOI: <http://doi.org/10.1629/uksg.468>. Disponível em: <https://insights.uksg.org/articles/10.1629/uksg.468/#>. Acesso em: 19 nov. 2020.

JONES, S. et al. Data management planning: How requirements and solutions are beginning to converge. **Data Intelligence**, v. 2, n. 1-2, p. 208-219, 2020. Disponível em: Disponível em: <http://www.data-intelligence-journal.org/p/49/> Acesso em 20 nov. 2020.

MÁRDERO ARELLANO, M. Á. Preservação de documentos digitais. **Ciência da Informação**, Brasília, DF, v. 33, n. 2, p. 15-27, maio/ago. 2004. Disponível em: <http://www.brapci.ufpr.br/download.php?dd0=9026>. Acesso em: 26 nov. 2020.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **OECD principles and guidelines for access to research data from public funding**. Paris: OECD Publishing, 2007. Disponível em: <https://doi.org/10.1787/9789264034020-en-fr>. Acesso em: 14 nov. 2020.

RANS, J. WHYTE, A. **Using RISE the Research Infrastructure Self-Evaluation Framework**. 2017. Edinburgh: Digital Curation Centre. Disponível em: www.dcc.ac.uk/guidance/how-guides. Acesso: 30 nov. 2020.

SALES, L.; SAYÃO, L. F. Uma proposta de taxonomia para dados de pesquisa. **Conhecimento em Ação**, Rio de Janeiro, v. 4, n. 1, jan./jun. 2019. Disponível em: <https://revistas.ufrj.br/index.php/rca/article/view/26337>. Acesso em: 23 out. 2020.

SANTOS, E. A.; SILVA, C. E. Os modelos de plano de negócios e sua relevância para sustentabilidade das micro e pequenas empresas. **Revista Brasileira de Administração Científica**, Aquidabã, v.3, n.1, p.37-62, 2012. Disponível em: https://www.researchgate.net/profile/Carlos_Silva43/publication/233747518_Os_modelos_de_plano_de_negocios_e_sua_relevancia_para_sustentabilidade_das_micro_e_pequenas_empresas/links/5918bbd6aca27200fe52f207/Os-modelos-de-plano-de-negocios-e-sua-relevancia-para-sustentabilidade-das-micro-e-pequenas-empresas.pdf. Acesso em 30 nov. 2020.

SAYÃO, L. F. Uma outra face dos metadados: informações para a gestão da preservação digital. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 15, n. 30, p. 1-31, 2010. Disponível em: <http://www.brapci.ufpr.br/download.php?dd0=12621>. Acesso em: 18 nov. 2020.

SAYÃO, L. F.; SALES, L. **Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores**. Rio de Janeiro: CNEN/IEN, 2015. Disponível em: <http://www.aben.com.br/Arquivos/420/420.pdf>. Acesso em: 29 out. 2020.

SPARC Europe. **Evaluate your RDM Offering**. Disponível em: <https://sparceurope.org/evaluate-your-rdm-offering/> Acesso: 30 nov. 2020.

STRASSER, C; COOK, R.; MICHENER, W.; BUDDEN, A. **Primer on data management: what you always wanted to know**. California: CDL, 2012. Disponível em: <https://escholarship.org/uc/item/7tf5q7n3#page-1>. Acesso em: 29 out. 2020.

TORINO, E.; ROA-MARTINEZ, S. M.; VIDOTTI, S. A. B. G. Dados de pesquisa: disponibilização ou publicação? In: SHINTAKU, Milton; SALES, Luana Farias; COSTA, Michelli (org.). **Tópicos sobre dados abertos para editores científicos**. Botucatu, SP: ABEC, 2020. p. 183-201. Disponível em: <http://200.19.73.81:8080/jspui/handle/1/4725>. Acesso em: 23 out. 2020.

WHYTE, A.; TEDDS, J. **Making the case for research data management**. DCC Briefing Papers. Edinburgh: Digital Curation Centre. Disponível em: <https://www.dcc.ac.uk/guidance/briefing-papers/making-case-rdm>. Acesso em: 15 nov. 2020.

NOTAS

AGRADECIMENTOS

Agradecemos à Paloma Marin-Arraiza pelas discussões acerca do RISE.

CONTRIBUIÇÃO DE AUTORIA

Os papéis descrevem a contribuição específica de cada colaborador para a produção acadêmica inserir os dados dos autores conforme exemplo, excluindo o que não for aplicável. Iniciais dos primeiros nomes acrescidas com o último Sobrenome, conforme exemplo.

Concepção e elaboração do manuscrito: E. Torino, J. F. Brito, G. L. Trevisan

Coleta de dados: E. Torino, J. F. Brito, G. L. Trevisan, S. A. G. B. Vidotti

Análise de dados: E. Torino, J. F. Brito, G. L. Trevisan, S. A. G. B. Vidotti

Discussão dos resultados: E. Torino, J. F. Brito, G. L. Trevisan, S. A. G. B. Vidotti

Revisão e aprovação: E. Torino, J. F. Brito, G. L. Trevisan, S. A. G. B. Vidotti

Caso necessário veja outros papéis em: <https://casrai.org/credit/>

CONJUNTO DE DADOS DE PESQUISA

Escolha uma das opções e apague as demais.

- 1) Todo o conjunto de dados que dá suporte aos resultados deste estudo foi publicado no artigo e na seção “Materiais suplementares”.

FINANCIAMENTO

Bolsa de Produtividade em Pesquisa PQ do CNPq – Processo nº 311936/2016-4.

LICENÇA DE USO

Os autores cedem à **Encontros Bibli** os direitos exclusivos de primeira publicação, com o trabalho simultaneamente licenciado sob a [Licença Creative Commons Attribution](#) (CC BY) 4.0 International. Esta licença permite que **terceiros** remixem, adaptem e criem a partir do trabalho publicado, atribuindo o devido crédito de autoria e publicação inicial neste periódico. Os **autores** têm autorização para assumir contratos adicionais separadamente, para distribuição não exclusiva da versão do trabalho publicada neste periódico (ex.: publicar em repositório institucional, em site pessoal, publicar uma tradução, ou como capítulo de livro), com reconhecimento de autoria e publicação inicial neste periódico.

PUBLISHER

Universidade Federal de Santa Catarina. Programa de Pós-graduação em Ciência da Informação. Publicação no [Portal de Periódicos UFSC](#). As ideias expressadas neste artigo são de responsabilidade de seus autores, não representando, necessariamente, a opinião dos editores ou da universidade.

EDITORES

Edgar Bisset Alvarez, Ana Clara Cândido, Patrícia Neubert e Genilson Geraldo.

HISTÓRICO

Recebido em: 10-12-2021 – Aprovado em: 15-06-2022 - Publicado em: 08-07-2022.

