

Proposta de modelo de serviço de gestão de dados de pesquisa*

Proposal for a model of research data management service

Luana Farias Sales¹, Luís Fernando Sayão²

¹ Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT, Universidade Federal do Rio de Janeiro - Programa de Pós-graduação em Ciência da Informação – PPGCI, Rio de Janeiro, RJ, Brasil. ORCID: <http://orcid.org/0000-0002-3614-2356>

² Comissão Nacional de Energia Nuclear – CNEN, Universidade Federal do Rio de Janeiro-Programa de Pós-graduação em Ciência da Informação – PPGCI, Rio de Janeiro, RJ, Brasil. ORCID: <http://orcid.org/0000-0002-6970-0553>

Autor para correspondência/Mail to: Luana Farias Sales, luanasales@ibict.br

Recebido/Submitted: 27 de abril de 2022; Aceito/Approved: 30 de agosto de 2022



Copyright © 2022 Sales & Sayão. Todo o conteúdo da Revista (incluindo-se instruções, política editorial e modelos) está sob uma licença Creative Commons Atribuição 4.0 Internacional. Ao serem publicados por esta Revista, os artigos são de livre uso em ambientes educacionais, de pesquisa e não comerciais, com atribuição de autoria obrigatória. Mais informações em <http://revistas.ufrpr.br/atoz/about/submissions#copyrightNotice>.

Resumo

O planejamento, desenvolvimento, implantação e operação de plataformas de gestão de dados de pesquisa envolvem muitas variáveis que precisam ser adequadamente articuladas. O projeto dessas plataformas constitui-se um problema complexo e multifacetado, cuja solução precisa ser manifestada na forma de serviços que atendam às necessidades dos fluxos de pesquisa de domínios disciplinares específicos, parâmetros informacionais, tecnológicos e científicos, perfis profissionais variados, condicionantes políticas, éticas e legais, e sustentabilidade econômica e temporal. Neste ambiente multifacetado próprio da ciência contemporânea, a questão que se pretende responder é: que tipos de serviços de gestão de dados podem ser oferecidos em uma articulação que permeie os fluxos de trabalhos do ambiente científico, considerando as peculiaridades dos domínios disciplinares e servindo como mecanismo de auxílio para o desenvolvimento de pesquisas e incentivo ao compartilhamento e reuso de dados científicos? Como resposta, desenvolveu-se uma pesquisa de cunho teórico-exploratório que tomou como base metodológica a análise da literatura da área, com ênfase especial em artigos, relatórios, manuais e projetos de infraestrutura de dados, elaborados por pesquisadores e instituições de pesquisa, dando proeminência aos pontos de observação dos atores mais proximamente envolvidos na questão. A partir do inter-relacionamento das diversas variáveis levantadas, apresenta-se como resultado a descrição de um conjunto de serviços que podem ser oferecidos, organizados em um modelo de serviços. Espera-se que esse modelo possa servir como norte para instituições de pesquisa que desejem implantar plataformas de gestão de dados mais próximas das suas comunidades e que sejam aderentes aos padrões e tradições de compartilhamento e o reuso de dados científicos de seus pesquisadores.

Palavras-chave: Serviços de gestão de dados de pesquisa; Dados FAIR; Gestão do conhecimento científico.

Abstract

The planning, development, implementation, and operation of research data management platforms involves many variables that need to be properly addressed. The design of these platforms constitutes a complex and multifaceted problem, of which the solution needs to be manifested in the form of services that meet the needs of research flows in specific disciplinary domains, informational, technological and scientific parameters, varied professional profiles, political, legal and ethical constraints, and economic and temporal sustainability. In this multifaceted environment of contemporary science, the question we intend to answer is: What types of data management services can be offered based on the workflows of the scientific environment, considering the peculiarities of the disciplinary domains and serving as a support mechanism for the development of research and encouraging the sharing and reuse of scientific data? As an answer to the question, theoretical-exploratory research was developed that took as a methodological basis the analysis of the literature in the area, with special emphasis on articles, reports, manuals and data infrastructure projects, prepared by researchers and research institutions, giving prominence to the observation points of the actors most closely involved in the issue. From the interrelationship of the diverse variables raised, the result is the description of a set of services that can be offered, organized in a service model. It is hoped that this model can serve as a guide for research institutions that wish to implement data management platforms closer to their communities and that follow the standards and traditions of sharing and reusing scientific data from their researchers.

Keywords: Research data management services; FAIR data; Scientific knowledge management.

INTRODUÇÃO

A gestão de dados de pesquisa se configura, atualmente, como um foco de interesse e um dos maiores desafios para as instituições de pesquisa. Em escala planetária, as ações de gestão e curadoria de dados se destacam com proeminência no cenário da pesquisa do século XXI, acompanhando a ubiquidade das tecnologias digitais para a coleta, análise e arquivamento em quase todos os domínios disciplinares (Mayernik et al., 2012). Assim sendo, as instituições de pesquisa, em gradações distintas, estão reconceituando a gestão de dados e identificando-a como parte integrante dos processos de pesquisa, reconsiderando ou ampliando as suas estratégias de tratamento dos dados, implementando plataformas de gestão e curadoria, adquirindo ferramentas de análises e visualização, e desenvolvendo programas de capacitação para as suas equipes.

*O presente artigo é uma parte atualizada e reorganizada do capítulo Sayão, L. F., Sales, L. F. Gestão de dados como serviço: proposta de um modelo. Publicado na coletânea organizada por Sales, L. F., Violla, C. M. M. (2021). *Informação digital e suas diversas abordagens pela ótica de um cientista da informação*. Rio de Janeiro: PPGCI IBICT-UFR. doi: 10.22477/9786589167136 – que reúne as publicações mais citadas do Professor Luís Fernando Sayão.

Temos ultrapassado algumas décadas na compreensão e equacionamento da gestão de dados, e percebe-se haver um interesse crescente por parte dos vários *stakeholders* – pesquisadores, agências de fomento, instituições de pesquisa e profissionais de informação, casas editoras científicas e tantos outros. No entanto, o nível de sucesso em implementar e oferecer serviços de gestão de dados, no âmbito das várias organizações de pesquisa intensiva em dados, não tem sido consistente e ainda está, de forma geral, nos estágios preliminares. O planejamento, o desenvolvimento e a implantação de plataformas de gestão de dados de pesquisa, devido ao número de variáveis que precisam ser equacionadas, são um problema complexo e multifacetado. A solução para esse problema precisa ser articulada em torno de fluxos de trabalho de domínios disciplinares específicos, parâmetros informacionais, tecnológicos, políticos, éticos e legais de sustentabilidade e de expertise numa odisseia marcada por constantes mudanças, cuja principal característica é a heterogeneidade. Neste sentido, duas questões se colocam como proeminentes, as quais tentamos responder na pesquisa, a saber: que tipos de serviços de gestão de dados de pesquisa podem ser oferecidos no decorrer do desenvolvimento das pesquisas considerando as peculiaridades dos domínios disciplinares? E ainda, que tipos de serviços de gestão de dados podem ser oferecidos durante os fluxos de trabalhos do ambiente científico, a fim de auxiliar o desenvolvimento de pesquisas e incentivar o compartilhamento e reuso de dados de pesquisa?

O ambiente complexo que se apresenta pode ser um terreno adequado para a adoção dos Princípios FAIR¹ (Mons et al., 2017) como horizonte para a implementação de serviços de gestão que tornem os dados de pesquisa encontráveis, acessíveis e interoperáveis, para que possam ser reusados por longo prazo. Desta maneira, criam-se condições para a transição de uma pesquisa autocontida para uma pesquisa mais colaborativa, em rede e relacional, que, ao mesmo tempo, atende a requisitos disciplinares e beneficia comunidades de culturas e restrições específicas.

Tentando equacionar essa diversidade, o presente trabalho tem como objetivo apresentar uma arquitetura genérica para apoiar o projeto de plataformas de serviços de dados, definindo, realinhando, agregando e articulando os vários módulos conceituais – diretrizes, políticas, serviços, ferramentas, infraestruturas etc. – em torno de um modelo de camadas, que como blocos de construção podem ser ajustados de acordo com a profundidade, alcance e filosofia de cada instituição ou disciplina.

A pesquisa exploratória, de acordo com Gil (2002), se caracteriza por ser um tipo de pesquisa realizada quando há pouco conhecimento da temática a ser abordada, buscando-se conhecer com maior profundidade o assunto de modo a torná-lo mais claro para futuras formulações de hipótese. Por ser voltada para temáticas novas, a pesquisa exploratória não se compromete com métodos rígidos, sendo seu planejamento “bastante flexível, de modo que possibilite a consideração dos mais variados aspectos relativos ao fato estudado” (Gil, 2002, p. 44). A presente pesquisa se caracteriza como exploratória e de cunho teórico. Para sua realização, explorou-se a literatura internacional sobre a temática “*Research data management services*”, com ênfase especial em artigos, relatórios, manuais e projetos de infraestrutura de dados, elaborados por pesquisadores e instituições de pesquisa, dando proeminência aos pontos de observação dos atores mais proximamente envolvidos na questão.

COMPREENDENDO OS SERVIÇOS DE GESTÃO DE DADOS DE PESQUISA

Num ambiente científico em que há uma intensa geração e consumo de dados pela odisseia cotidiana dos cientistas na busca por novos conhecimentos, há uma necessidade crítica de dispositivos para controle e organização desses ativos que, idealmente, se concretizam por um conjunto de serviços alinhados a uma estratégia coerente de gestão de dados. Nesse contexto em constante transição, a gestão de dados de pesquisa, em escala mundial, se torna cada vez mais um serviço essencial que deve ser intermediado pelas unidades de informação das organizações de pesquisa (Tang & Hu, 2019).

A amplitude das ações proporcionadas pelos sistemas técnicos e gerenciais das plataformas de gestão de dados propicia às diversas partes interessadas, que orbitam em torno dos fluxos de geração de conhecimento científico, uma série de benefícios que dependem das suas matrizes de interesse, por exemplo: para um coordenador ou gestor acadêmico, a plataforma se configura em um dispositivo de avaliação de produtividade de um programa, área ou grupo de pesquisa; para um editor científico e seu corpo de revisores, a plataforma serve como um instrumento de avaliação e validação de registros submetidos; para os formuladores de políticas científicas e financiadores de pesquisa, as plataformas tornam-se uma cartografia que apoia a ordenação e os níveis de incentivos necessários a uma área de estudo; e poderíamos dizer que para o cidadão comum e para a grande imprensa serviria como um elemento a mais de transparência dos investimentos públicos em ciência. Todavia, o pesquisador tem a percepção sobre os dados de um instrumento de pesquisa e de geração de novos conhecimentos, que tem um potencial transformacional, conforme enfatizam Sara Jones, Prior, e White (2013).

Assim sendo, a gestão de dados não é um fim em si mesmo e se concretiza aos olhos do pesquisador na forma de um amplo espectro de serviços e ferramentas que apoiam todo o ciclo de vida dos dados, no âmbito de um projeto de pesquisa, cujos benefícios são diretamente perceptíveis por ele, como por exemplo: maior visibilidade para a

¹FAIR, um acrônimo para *Findable, Accessible, Interoperable e Reusable*

pesquisa; mais citações e prestígio; reconhecimento da autoria dos dados; maior nível de colaboração em escala global; organização dos dados para o próprio uso do pesquisador e de seus colegas próximos; reconhecimento em termos de promoção e financiamento; proteção lógica e física dos dados; e preservação de longo prazo. Isto indica fortemente que no âmbito da gestão de dados a conexão orgânica com as comunidades científicas se concretiza via serviços ressonantes com a dinâmica e a cultura disciplinar, porém, sem perder a perspectiva de inserção global e interdisciplinar.

O estudo profundo sobre a avaliação da maturidade – um enfoque comum para a determinação do nível de sofisticação de serviços e produtos – dos serviços de dados conduzido por Inna Kouper, Fear, Ishida, Kollen, e Williams (2017) indica que os serviços de gestão de dados mais avançados são provavelmente aqueles cujas ações privilegiam as necessidades das comunidades de instituições individuais, mas também estão cientes das comunidades de pesquisa mais amplas nas quais os indivíduos se inserem. Os programas de serviços de gestão de dados de pesquisa mais consolidados não são necessariamente aqueles que oferecem as maiores carteiras de serviços ou que empregam as equipes maiores, “mas aqueles cujas atividades são mais profundamente conectadas à missão da biblioteca e da instituição como um todo [...], em outras palavras, onde os serviços são cuidadosamente escolhidos, organizados, monitorados e otimizados”, concluem os autores (Kouper et al., 2017, p. 164).

Nessa diversidade de enfoques, os princípios FAIR oferecem uma base conceitual e um horizonte mais substantivo para o delineamento e consecução de serviços de gestão de dados, posto que eles estão focados em assegurar que os objetos de pesquisa sejam encontráveis, acessíveis, interoperáveis e reusáveis, sintetizando dessa forma os possíveis objetivos de um serviço de gestão de dados, ou, num nível de abstração mais elevado, o alcance de uma Internet de Dados e Serviços FAIR. Os princípios FAIR “descrevem características e aspirações aplicadas aos sistemas e serviços voltados para apoiar a criação de resultados valiosos de pesquisa que podem, então, ser rigorosamente avaliados e amplamente reutilizados [...]” (Mons et al., 2017, p. 50). Os princípios deliberadamente não especificam requisitos técnicos e não são padrões, mas constituem um conjunto de diretrizes orientadoras que estabelecem um continuum crescente de possibilidades de reuso dos objetos de pesquisa, por meio de uma variedade de implementações diferentes. Nesta direção, os princípios FAIR descrevem as qualidades ou comportamentos dos recursos de dados para conseguir – possivelmente de forma incremental – um ótimo nível de descoberta e de reuso acadêmico, abrindo possibilidade para muitos e diferentes enfoques na concepção de serviços e no reuso de dados.

Levando em conta os argumentos acima e tomando os princípios FAIR como núcleo agregador conceitual, propomos a seguinte definição para **serviços de gestão de dados de pesquisa**: é o conjunto de serviços informacionais, computacionais, científicos e administrativos oferecidos no âmbito da gestão de dados de pesquisa e ancorados nas necessidades específicas das comunidades acadêmicas e científicas, que têm como propósito tornar os dados localizáveis, acessíveis, interoperáveis e reusáveis, de forma que eles se traduzam em benefícios para a ciência e para todos os seus *stakeholders*.

A partir desta definição, a seguir propomos um modelo de serviço de gestão de dados que procura representar os diferentes aspectos da gestão de dados de pesquisa, sem perder de vista a natureza interrelacionada da dinâmica das atividades que se desenrolam num ambiente científico intensivo de dados, cujo objetivo é torná-los FAIR.

DELINEANDO OS CONTORNOS DOS SERVIÇOS DE GESTÃO DE DADOS

Como ratificado por Sara Jones et al. (2013, p. 5), “para dar apoio efetivo à gestão e ao compartilhamento de dados, uma instituição necessita de uma estratégia coerente e de um conjunto de serviços”. Mas o que poderia significar este conjunto de serviços de gestão de dados? Naturalmente ele tem um espectro contínuo que varia em termos disciplinares, cultural e epistemológico, institucional e político, e ainda depende das bases tecnológicas disponíveis para a gestão de dados. De fato, as instituições de pesquisa podem oferecer serviços de dados numa grande multiplicidade, que varia não somente nos tipos de serviço, mas também na profundidade e alcance que esses serviços são disponibilizados, nos níveis de especificidade e comprometimento e para quem e com que objetivos esses serviços são oferecidos (Choudhury et al., 2018).

Para exemplificar, Fearon Jr., Gunia, Pralle, Lake, e Sallans (2013) apontam que serviços de gestão de dados englobam o fornecimento de informações, consultoria, treinamento e ainda o envolvimento ativo no planejamento da gestão de dados, orientação durante a pesquisa (por exemplo, aconselhamento sobre o armazenamento de dados e segurança de arquivos), documentação e metadados, compartilhamento de dados de pesquisa e curadoria (seleção, preservação, arquivamento, citação) de projetos concluídos e dados publicados. Já na perspectiva de Choudhury et al. (2018) (2018), os serviços de gestão de dados incluem a oferta de infraestrutura necessária para realizar a curadoria de dados por meio de licenças para preservação, análises e ferramentas de acesso; a disponibilidade de espaço em sistemas de armazenamento financiados pela organização para dados curados; treinamento e consultoria que permitam o pesquisador explorar os serviços de dados oferecidos pelas várias unidades da instituição. Complementando, Tang e Hu (2019) apontam que no diagrama de componentes de gestão de dados de pesquisa, as atividades abrangentes incluem “política e estratégias de gestão de dados” e “plano de negócios e sustentabilidade”. Subjacente ao estabelecimento de um serviço de gestão de dados de

pesquisa, vários níveis de orientação, treinamento e suporte são necessários. Para esses autores, o ponto focal do processo de gestão de dados deve dar proeminência aos componentes de serviço de gestão de planejamento, gerenciamento de dados ativos, seleção e compartilhamento, bem como repositórios e catálogos de dados. Neste sentido, os repositórios ou os catálogos de dados são apenas mais um serviço dentre inúmeros outros que uma plataforma de serviço de gestão de dados pode oferecer.

É importante observar que são muitas as diferenças entre a gestão de recursos mais tradicional e o nível de exigências técnicas e de infraestruturas e expertises necessárias à gestão de dados de pesquisa. Um livro, por exemplo, tem uma catalogação universal e padronizada, as diferenças de tratamento entre disciplinas são poucas e seus processos estão focados na pós-publicação; o mesmo não se pode dizer de dados de pesquisa e de outros objetos digitais de pesquisa, como base de dados e códigos, cuja gestão tem que se preocupar com o longo e idiossincrático ciclo de vida que se inicia ainda na fase de planejamento – muito antes da publicação e arquivamento, indo até a pós-publicação, mas num processo ainda mais complexo do que era executado na gestão das publicações bibliográficas. Some-se a isso toda a peculiaridade própria que exige a articulação da gestão com o ciclo de vida do projeto de pesquisa. Neste contexto, o que se observa é que “O leque de competências e conhecimentos necessários para entregar serviços de gestão de dados é ditado em grande parte pelas fases individuais do ciclo de vida do projeto”, confirmam Jones et al. (2013, p. 3). Assim, a escala de serviços que as instituições de pesquisa oferecem pode variar não apenas nos tipos de serviços disponibilizados, mas também no nível de profundidade em que eles atuam, e no universo de usuários para quem os serviços são oferecidos (Choudhury et al., 2018). Pesquisadores, professores e estudante de pós-graduação são os clientes-alvo mais prováveis dos sistemas de gestão de dados, porém outros *stakeholders* devem ser considerados, como os gestores de C & T, financiadores e comunidades de práticas específicas – como engenheiros e agrônomos –, que reusam os dados, especialmente os dados com alto grau de processamento nos seus projetos e empreendimentos, como na construção das fundações de uma usina nuclear ou na seleção de cultivares. Os serviços podem estar distribuídos por várias unidades da instituição ou concentrados e coordenados por uma unidade, possivelmente a biblioteca de pesquisa.

A visão fragmentada e heterogênea sobre os serviços de gestão de dados – que por fim reflete as múltiplas faces da atividade de pesquisa – cria um obstáculo no delineamento dos seus contornos e na enumeração do diagrama dos seus componentes. Possivelmente, uma racionalidade partindo dos pesquisadores na qualidade de usuários desses serviços pode ajudar na compreensão e na construção de um possível conceito de serviços de gestão de dados. É o que este ensaio tenta fazer na próxima seção: propor uma definição para serviço de gestão de dados de pesquisa onde se considere as necessidades específicas de seus usuários.

MODELO DE SERVIÇO DE GESTÃO DE DADOS DE PESQUISA FAIR

Como uma abstração conveniente da realidade que se quer compreender, um modelo é uma criação cultural, um “mentefato” destinado a representar uma realidade ou alguns dos seus aspectos, a fim de torná-los descritíveis qualitativa e quantitativamente e, algumas vezes, observáveis (Sayão, 2001). A partir desse ponto, decidiu-se dividir o modelo em quatro camadas representacionais: 1) a governança, onde são discutidos os princípios norteadores do projeto de serviços de gestão de dados; 2) as infraestruturas técnicas, onde se incluem também as categorias de expertises necessárias; 3) os serviços informacionais, computacionais, científicos e administrativos; 4) os resultados da efetivação desses serviços manifestados pela FAIRificação dos dados.

Governança de Dados de Pesquisa: planejamento, política, institucionalização e sustentabilidade

A configuração organizacional e institucional nas quais a gestão de dados é realizada pode variar em relação a vários aspectos, como a intensidade de apoio à gestão e o nível de investimentos aplicados. Algumas instituições, como centros referenciais de dados científicos e agências estatísticas governamentais, podem estar inteiramente dedicadas à gestão de dados, tendo-a como finalidade principal. Em outras configurações, a gestão de dados é parte de uma atividade mais ampla que se conecta a outras atividades de pesquisa, como no caso das universidades (National Research Council, 2015), cuja atividade de gestão de dados é decorrente de suas funções de ensino, pesquisa e extensão. Porém, mesmo no contexto acadêmico, são muitas as formas de planejar e executar as tarefas de gestão de dados, que variam de acordo com referências objetivas, como graus de investimento, sistemas técnicos disponíveis, volume e tipo de dados e de como a gestão de dados está integrada aos seus fluxos de trabalhos e processos; e com percepções mais subjetivas, como cultura disciplinar e prestígio acadêmico. No presente modelo, esses parâmetros são equacionados por um nível mais administrativo compreendido pela categoria “governança de dados”. Num plano mais conceitual, a governança de dados delinea os princípios, políticas e estratégias que são comumente adotados num ambiente que necessita de um programa de gestão de dados coerente; delinea também as ações, funções e papéis que são necessários para implementar essas políticas e estratégias. No âmbito de uma instituição de pesquisa, os princípios operacionalizados pela gestão governam todo o ciclo de vida dos dados – da conceitualização ao arquivamento e possível descarte. Assim sendo, o processo de governança de dados trata os dados não somente em seu aspecto espacial, mas também ao longo da sua dimensão temporal (Solomonides, 2019), e este requisito implica uma ampliação do grau de complexidade e envergadura dos comprometimentos da governança.

Este arcabouço estruturante é necessário posto que dados de pesquisa digitais só podem ser gerenciados e preservados adequadamente ao longo do tempo por meio de um compromisso institucional sustentável (Mayernik et al., 2012, p. 1). Em certa medida, a consolidação dos serviços de gestão de dados reflete o nível de aceitação organizacional incorporada a eles e o grau de planejamento das várias ações necessárias: orçamento sustentável em vigor, política de dados apropriada, conexão orgânica com as comunidades-alvo, conformidade com os códigos éticos e legais, alinhamento com os objetivos estratégicos institucionais e uma estratégia de desenvolvimento que considere os percursos possíveis para cada instituição. É necessário considerar também a inevitabilidade do fato de que as infraestruturas tecnológicas para acessar, interpretar e preservar a informação digital estão continuamente evoluindo; antecipar esses problemas e desenvolver estratégias para mitigá-los é uma atividade relevante para os compromissos de governança (National Research Council, 2015). Sobre esses pilares podem ser desenvolvidos serviços avançados de dados que possam apoiar apropriadamente todo o ciclo de vida desses ativos informacionais, de acordo com os interesses dos vários *stakeholders* envolvidos. A figura 1 a seguir apresenta os pilares da governança de dados que vão sustentar todo o modelo proposto. Em seguida, faz-se uma descrição de cada um dos itens que representa os pilares.

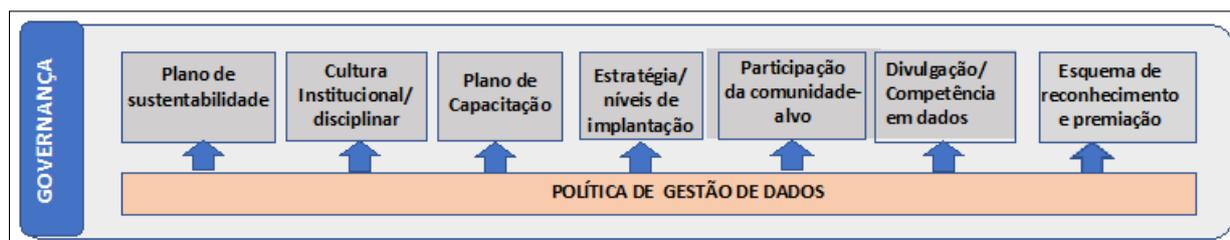


Figura 1. Pilares da Governança de Dados.

Fonte: Os autores (2022).

Política de Gestão de Dados da Instituição – Estabelece os fundamentos, diretrizes e compromissos da instituição concernentes à gestão, uso, propriedade, conformidade aos códigos éticos e legais, aderência às políticas das agências de fomento, políticas nacionais de ciência, tecnologia e inovação, às orientações e práticas internacionais e, por fim, mas de importância crítica, à cultura, às práticas e às idiossincrasias das comunidades e domínios disciplinares. Uma política de gestão de dados de pesquisa abrangente deve também identificar as responsabilidades de cada um dos atores – biblioteca, laboratórios, tecnologia da informação, administração etc. – posto que a gestão de dados envolve diferentes setores da instituição (Mushi, Pienaar, & Deventer, 2020) e o projeto é considerado como parte das atividades de pesquisa da instituição. É necessário enfatizar que o processo de desenvolvimento de uma política institucional de gestão de dados requer uma consulta extensiva a todos os agentes envolvidos e a aprovação das comunidades e organizações científicas relevantes (Wilson, Martinez-Uribe, Fraser, & Jeffreys, 2011). As orientações da política devem permear todo o ciclo da gestão. “Políticas podem ser um importante fator motivador para dados FAIR e outros objetos de pesquisa (*software, workflow, modelos, protocolos* etc.). Portanto, é essencial que esforços “*bottom-up*” baseados na comunidade sejam combinados com políticas com enfoque “*top-down*”, completam o European Commission (2020).

Cultura Institucional/disciplinar – A implantação de uma plataforma de serviços de gestão de dados de pesquisa deve ser precedida de uma análise de requisitos que considere o contexto e a cultura institucional, comunitária e disciplinar, bem como suas características únicas. Espera-se que este processo ajude a definir uma carteira de serviços de gestão de dados mais efetiva para apoiar as práticas de pesquisa da instituição e de suas comunidades (Coates, 2014, Reed, 2015, Mushi et al., 2020). É importante também reconhecer que algumas disciplinas necessitam de diferentes tipos de soluções técnicas para obter os mesmos benefícios dos dados FAIR (European Commission, 2020).

Plano de sustentabilidade – Um dos grandes desafios de um programa de implementação e manutenção de uma infraestrutura de gestão de dados é assegurar que cada fase do projeto seja sustentável como um serviço contínuo ao longo do tempo (Wilson et al., 2011). Uma vez que a gestão de dados de pesquisa é reconhecida como algo necessário para as atividades de pesquisa, os seus custos devem ser estimados e suas fontes de financiamento – especialmente as perenes – identificadas. A escolha sobre onde investir os recursos, tipicamente limitados, e as projeções sobre as necessidades futuras se tornam essenciais (Choudhury et al., 2018). Desta forma, um projeto de implementação de serviços de gestão de dados de pesquisa precisa estar associado a um plano de sustentabilidade que delinear um comprometimento possível com o agora e com o futuro. A criação e o comprometimento com uma estratégia de longo prazo para os serviços podem revelar com mais clareza os recursos necessários à continuidade dos serviços e das infraestruturas necessárias para tal. Isso, portanto, pode incluir, um plano de sucessão (Mushi et al., 2020).

Divulgação/Competência em dados – Para a implementação de um ambiente de pesquisa FAIR é necessário que as comunidades envolvidas desenvolvam uma compreensão compartilhada do que está circunscrito pelo conceito FAIR e pelos seus princípios. De uma forma geral, os pesquisadores e outros *stakeholders* têm baixo nível

de percepção sobre a importância das práticas de gestão de dados e das exigências de gestão e compartilhamento das agências de fomento e dos compromissos de depósito dos dados firmados com os editores científicos, além das questões éticas e legais envolvidas na publicação dos dados. Por exemplo, em relação ao conceito FAIR, [European Commission \(2020\)](#) observam que o pesquisador não sabe o que é dado FAIR e muitas vezes acha que é o mesmo que dado aberto. Isto indica que é necessário planejamento e ações de divulgação e de conscientização que tragam à tona essas questões. Um programa de divulgação nesta direção deve contemplar a elaboração de material didático (cartilhas e guias), cursos, eventos, oficinas, entre outros.

Conhecimento/participação da comunidade-alvo – Como criadores e usuários de dados de pesquisa, o engajamento dos pesquisadores é crucial no desenvolvimento de serviços de gestão de dados. O provisionamento de qualquer serviço precisa ser baseado numa compreensão próxima dos padrões e fluxos das pesquisas que se desenvolvem na instituição, das suas motivações, características e prioridades. Assim sendo, a definição precisa dos requisitos dos serviços necessita ser estabelecida com o comprometimento e a contribuição da comunidade de pesquisadores, sem essas considerações, as características dos serviços podem não estar em harmonia com os objetivos dos pesquisadores. A comunidade deve ser acompanhada nas mudanças de interesse sobre os dados, e a participação dela no desenvolvimento e escolha de padrões compartilháveis para as práticas e para as infraestruturas FAIR deve ser reconhecida e institucionalizada. A proximidade, interação e alinhamento das comunidades com as organizações nacionais e internacionais que lidam diretamente com a gestão de dados FAIR, como GO FAIR, RDA, CODATA, DCC e outras, devem ser incentivados.

Plano de capacitação – Para oferecer serviços completos em gestão de dados, as bibliotecas precisam ter pessoal tecnologicamente qualificado ou aumentar muito as oportunidades de treinamento tecnológico para o pessoal existente. ([Tenopir, Birch, & Allard, 2012](#)) A sustentabilidade humana é uma condição crítica para assegurar a continuidade e a consistência da oferta de serviços ao longo do tempo. Entretanto, poucos programas formais em estudos informacionais incluem em seus currículos gestão de dados, desta forma, os gestores de dados de pesquisa são normalmente treinados em serviço nas disciplinas específicas onde trabalham ([Borgman, 2007](#), p. 155).

Estratégia/níveis de implantação – O desenvolvimento e a implantação de infraestrutura de gestão de dados, além de muitos recursos, requerem tempo para alcançar sua plena maturidade e espelhar as demandas das comunidades científicas. Isto implica a necessidade do estabelecimento de níveis de implantação de infraestruturas e serviços. As bibliotecas de pesquisa, por exemplo, têm, em muitos casos e de forma proativa, procurado suprir as necessidades de gestão de dados para as suas comunidades de usuário. Frequentemente isso acontece sem aporte financeiro adicional destinado ao desenvolvimento e disponibilização de serviços de dados. Assim sendo, as bibliotecas têm que começar numa escala mais simples, construindo uma base sobre a qual possa desenvolver serviços mais sofisticados ([Erway, Horton, Nurnberger, Otsuji, & Rushing, 2016](#)), começando com serviços básicos que exijam apenas recursos da própria biblioteca, até alcançarem serviços mais complexos que exijam alto nível de compromisso institucional e mais recursos financeiros, tecnológicos e humanos ([Kouper et al., 2017](#)).

Reconhecimento e premiação – A gestão de dados de pesquisa consome tempo, recursos e exige grande dedicação do pesquisador. Entretanto, esse esforço raramente é reconhecido pelos sistemas de recompensa acadêmicos, exceto quando linkado com publicações em periódicos científicos. Portanto, para incentivar essa nova tarefa dos pesquisadores e destacar sua importância, é essencial que ela seja apropriadamente reconhecida e que seja considerada nos critérios de avaliação, promoção e de financiamento.

Infraestruturas de Dados de Pesquisa

Infraestrutura é uma noção de grande amplitude e multidimensional. Ela pode ter uma conotação técnica, legal, organizacional e, em muitos casos, é imprescindível considerar também os aspectos sociais, culturais e políticos. De fato, é assim no domínio da ciência: o projeto de infraestrutura de pesquisa é simultaneamente uma questão tecnológica, uma questão de identificação das necessidades da pesquisa em áreas disciplinares específicas e uma questão política. Essa ótica mais geral se aplica às infraestruturas institucionais de gestão de dados de pesquisa que precisam oferecer tecnologias e ferramental, processos, políticas, recursos e treinamento para os vários e diversificados estágios da gestão de dados.

Da mesma forma que as instituições devem providenciar infraestruturas básicas para a pesquisa – tais como laboratórios, instrumentação, computação de alto desempenho, redes, reagentes e muito mais – elas devem também tomar medidas para uma gestão adequada dos dados. Isto pressupõe um amplo espectro de atividades gerenciais, tecnológicas e informacionais que inclui profissionais de informação treinados para apoiar pesquisadores no planejamento e gestão de seus dados, no acesso a dispositivos de armazenamento seguro e backups durante o desenvolvimento do projeto e disponibilidade de plataformas de acesso e de preservação de longo prazo, necessárias após o fim da pesquisa ([Strasser, 2015](#)). É imprescindível também um corpo de normas, padrões e boas práticas que permitam, principalmente, uma interlocução em níveis variados dos sistemas e serviços, tanto local quanto global, que pode ser traduzida por interoperabilidade. Nesta categoria, à guisa de exemplo, estão os **padrões de modelo de dados** – geralmente estabelecidos por um domínio disciplinar ou repositório – que determinam a estrutura dos vários componentes de uma coleção de dados, que, por fim, têm efeito sobre as

interfaces de interação com os usuários humanos e computacionais e sobre os níveis de interoperabilidade do *dataset* (Choudhury et al., 2018).

Quando comparamos a publicação acadêmica tradicional com a publicação de dados verificamos que as infraestruturas subjacentes à publicação acadêmica criam uma ponte epistemológica entre disciplinas, tendo como ponto agregador as bibliotecas de pesquisa que selecionam, coletam, organizam e tornam acessíveis publicações de todo o tipo e de todas as áreas. Por sua natureza, as instituições sociais trabalham para estabilizar práticas particulares e formas de conhecimentos. Em certo sentido, as instituições são infraestruturas sociais em si mesmas. Nessa direção, as infraestruturas técnicas estão entrelaçadas com as infraestruturas sociais das instituições, muitas vezes mediadas por padrões, protocolos, documentos e artefatos que ligam os aspectos sociais e técnicos das infraestruturas (Leonardi, 2010). Entretanto, não existe ainda infraestrutura dessa magnitude para os dados. Algumas poucas áreas têm mecanismos consolidados para publicar dados; outras estão nos estágios de desenvolvimento de padrões e práticas para agregar seus dados e torná-los amplamente acessíveis. Um problema-chave nas instituições de pesquisa, como observam Mayernik et al. (2012, p. 158), “é a falta de uma infraestrutura confiável que possa ser implantada num nível institucional”, essa “falta de infraestrutura para dados amplifica a descontinuidade na publicação acadêmica”, acrescenta Borgman (2007, p. 155).

Os arcabouços infraestruturais voltados para a gestão de dados são diversos e fragmentados em termos de fluxos, complexidade, aplicação e topologia, e organizados de forma diferente pelas várias disciplinas e em diferentes países (Graaf & Waaijers, 2011). Contudo, crescentemente as infraestruturas moldam os padrões e as práticas da gestão de dados. Diante desse fato, o conhecimento sobre a origem, domínio disciplinar, grau de processamento, sistemas de coleta, *workflows* etc. parecem ser de importância crítica na concepção de infraestruturas voltadas para a gestão de dados (Sayão & Sales, 2020). A figura 2 a seguir revela a parte do modelo que representa a infraestrutura necessária para a gestão de dados de pesquisa, sendo explicado, na sequência, cada um dos itens.

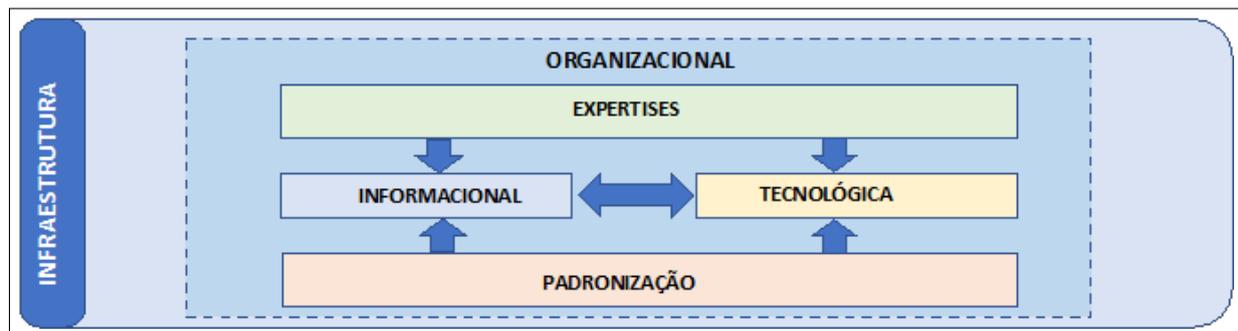


Figura 2. Infraestrutura necessária para a implementação de Serviços de Gestão de Dados.
Fonte: Os autores (2022).

Na presente proposta de modelo, consideramos cinco instâncias de infraestruturas necessárias à implantação de sistemas de gestão de dados: **de padronização** – em que normas e padrões se apresentam como formas consensuais de codificar o conhecimento que circula transversalmente por comunidades, assegurando uniformidade e similitude aos seus produtos e processos através do tempo e do espaço; **tecnológicos** – que compreende um vasto conjunto de atividades, equipamentos, processos e expertises que possam viabilizar os requisitos tecnológicos operacionais necessários às ciberinfraestruturas de gestão de dados; **informacionais**, que compreende todo o arcabouço conceitual e teórico materializado nas práticas da Ciência da Informação, Biblioteconomia e Arquivologia, que são plenamente aplicadas à gestão de dados, como seleção, catalogação, indexação, classificação e descarte, e os instrumentos que viabilizam essas práticas; **profissionais**, que compreende equipes de apoio compostas por diferentes profissionais, como por exemplo, administrador de dados e cientistas de dados, bibliotecários de dados, arquivistas de dados, informacionista de dados, curador de dados, entre outros; **organizacionais**, que pressupõe uma estrutura organizacional voltada para a pesquisa, como uma universidade, instituto de pesquisa, ou mesmo uma empresa cujos empreendimentos dependem da gestão de dados.

Essas vertentes infraestruturais – que possibilitam imbricamento de saberes e práticas que estão subjacentes a equipamentos, instalações, metodologias e principalmente a pessoas – proporcionam uma vasta carteira de serviços, ferramentas e processos que continuamente levam os objetos de pesquisa para um alinhamento com os Princípios FAIR.

O que se percebe até aqui é que, para se chegar na categoria que representa os serviços de gestão de dados de pesquisa, faz-se necessário o estabelecimento de uma governança de dados eficiente e a construção de uma infraestrutura adequada à formulação de serviços. Assim, o item a seguir descreve a categoria que é implicitamente o coração do modelo.

Serviços de Gestão de Dados de Pesquisa

Na construção do presente modelo, consideramos uma matriz de serviços baseada em dois eixos principais: um eixo temporal, que considera o desenrolar dos serviços de dados ao longo do tempo interligando o ciclo de vida dos dados ao ciclo de vida da pesquisa; o segundo eixo considera o ponto de ancoragem dos serviços, significando que eles podem estar fundamentados em processos informacionais, computacionais, científicos ou administrativos. Fica claro que esses contornos não são sempre bem definidos e as sobreposições estão presentes em ambos os eixos, o que demonstra a necessidade de interconexão de várias expertises para a consecução das atividades de gestão de dados de pesquisa. Do ponto de vista temporal, podemos considerar que a atuação da gestão na forma de serviços se efetiva em três momentos: antes da pesquisa começar, durante a pesquisa e depois que a pesquisa finaliza (Jones et al., 2013).

Para atender o amplo espectro das necessidades de gestão de dados, materializadas em serviços específicos, é necessária a colaboração de diversas áreas e a integração de expertises, infraestruturas, práticas e metodologias. Mesmo identificando que há sobreposições importantes – e desejáveis –, consideramos quatro tipos de serviços: serviços científicos, serviços computacionais, serviços informacionais e serviços administrativos (Figura 3).

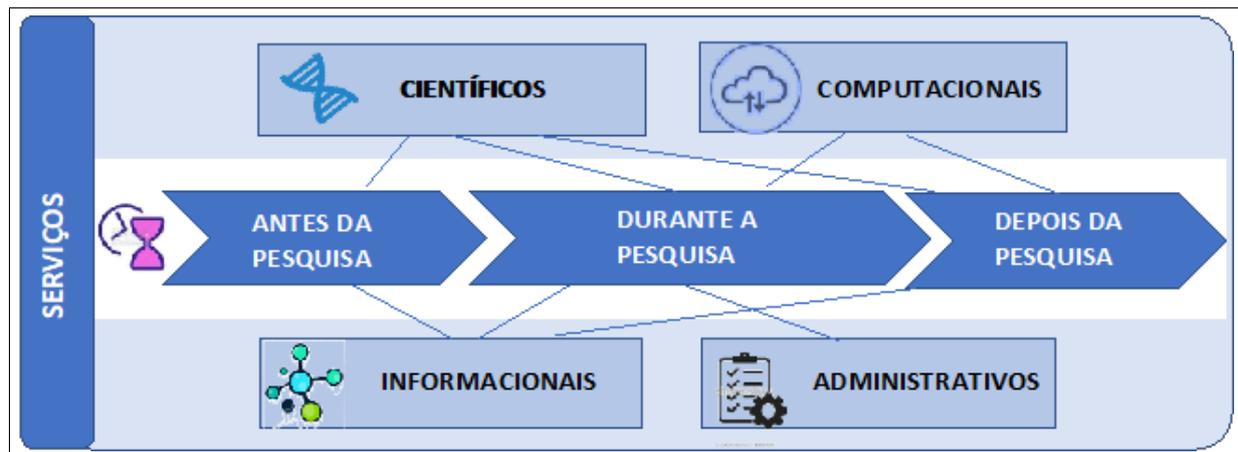


Figura 3. Serviços de Gestão de Dados de Pesquisa.

Fonte: Os autores (2022).

Os **serviços científicos** compreendem os serviços que se desenrolam em ambientes predominantemente científicos, como laboratórios e centros de pesquisa, e que são executados por cientistas, acadêmicos ou especialistas em gestão de dados com profundos conhecimentos disciplinares, e podem ser: descrição disciplinar; avaliação (*appraisal*) das coleções de dados; limpeza dos dados; organização dos dados; transformação; documentação do processamento; anotação; documentação dos códigos; documentação dos *workflows*; análise de dados; apresentação e visualização de dados; e empacotamento dos dados.

Os **serviços computacionais** são aqueles que compreendem a oferta de ferramentas de software e equipamentos de computação para apoiar o processamento, análise e visualização dos dados de pesquisa; apoiar os processos de interoperabilidade e acionamento por máquina de dados e metadados; prover orientação de como os dados podem melhor ser estruturados e armazenados e trabalhar, se necessário, junto aos pesquisadores na estruturação de bases de dados e marcação de texto (Wilson et al., 2011); os serviços podem incluir ainda treinamento específico para a equipe de pesquisadores nos recursos oferecidos e, em situações mais avançadas, oferecer processamento de alto desempenho, armazenamento em nuvem de grandes volumes e computação em grade. São exemplos de serviços computacionais: sistema de armazenamento; proteção de dados; normalização de formatos; serviços de *backup*; e apoio à eliminação segura dos dados.

No que tange aos *serviços informacionais*, entende-se como aqueles oferecidos pelas bibliotecas e executados com o apoio dos profissionais bibliotecários e arquivistas. São exemplos de serviços informacionais que podem ser oferecidos pelas bibliotecas científicas, universitárias e por outros centros de informação para a pesquisa: portal de dados científicos; balcão de referência de dado; apoio na elaboração do plano de gestão de dados; apoio na descoberta e acesso à coleção de dados; desenvolvimento de coleções de dados; desenvolvimento de metadados; criação de referências padronizadas; identificação de dados e pesquisadores; catalogação/indexação das coleções de dados; arquivamento de longo prazo/preservação; publicação de dados; contextualização/*linking*; treinamento para pesquisadores.

Finalmente, os **serviços de administração** são aqueles serviços que dão apoio, sustentabilidade e visibilidade aos demais. Compreendem serviços de orientação sobre custos, orçamento, aquisição de coleções de dados, conformidades ética e legal dos dados – especialmente dos dados sensíveis – às normativas e regulamentos institucionais, nacionais e internacionais; estatísticas de uso e reuso dos dados; esta categoria envolve também as questões de propriedade intelectual, licenças e tempo de embargo.

A ideia de tornar os dados aderentes aos Princípios FAIR – expressa pelo termo “FAIRificação” – não se realiza por si só. Para tal é necessário um processo multidimensional de gestão de dados, consubstanciado por um elenco de serviços, que efetivamente possa ir agregando valor ao longo do tempo aos objetos de pesquisa. Esta é a última camada da arquitetura proposta, que será apresentada a seguir.

FAIRificação dos dados

O que se constata é que o nível de conformidade dos vários objetos digitais de pesquisa aos Princípios FAIR está vinculado ao alcance e profundidade da gestão à qual eles estão submetidos. Conforme visto, isto pressupõe a necessidade de um arcabouço de várias camadas – científica, tecnológica, informacional e administrativa – que endereçam os múltiplos problemas que se interpõem entre o reuso, integridade, proveniência, reprodutibilidade, prestação de contas, bem como entre as novas necessidades e oportunidades de análise e reanálise, em larga escala, necessárias aos fluxos da *eScience* (Wilkinson et al., 2016).

Por conseguinte, o grau de aderência aos Princípios FAIR – a FAIRificação (Mons et al., 2017) – põe em evidência um conjunto de serviços, procedimentos e ferramentas que, mesmo não tendo esse parâmetro como objetivo final, coloca-o como uma parte de uma escala renovada de avaliação da maturidade de sistemas de gestão de dados de pesquisa. A figura 4 a seguir mostra o que vem a ser o topo do modelo, indicando o alvo de todo serviço de gestão de dados científicos, que dever ser o alinhamento com os 15 princípios que se distribuem entre as quatro categorias do FAIR.

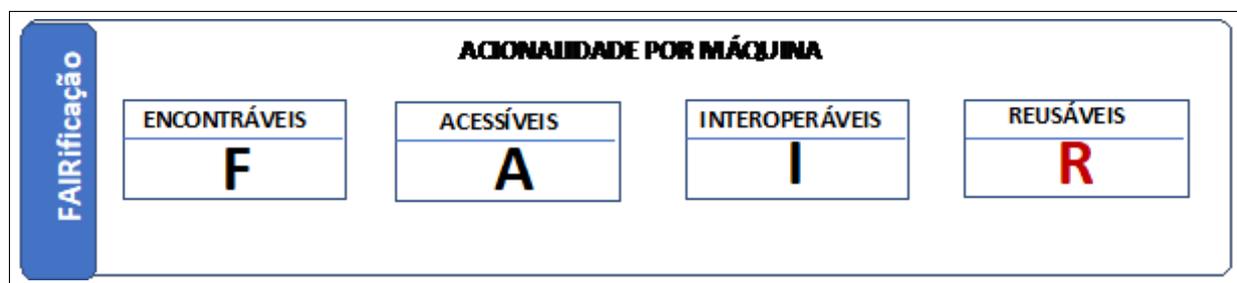


Figura 4. FAIRificação de serviços como alvo da Gestão de Dados.
Fonte: Os autores (2022).

No entanto, o alinhamento e a implementação dos Princípios FAIR em uma instituição de pesquisa exigem investimentos financeiros, mudanças culturais, treinamento e a construção de uma infraestrutura técnica apropriada (Graaf & Waaijers, 2011). Estes fatores podem ser aglutinados em torno do conceito de “plataforma de gestão de dado de pesquisa”, cuja implementação tem o potencial de operacionalizar as diversas camadas de gestão e estabelecer uma crescente infraestrutura de serviços informacionais, científicos, computacionais e administrativos, que viabilize uma escala de aderência aos princípios FAIR dos objetos de pesquisa, sejam eles dados propriamente ditos ou algoritmos, códigos, *workflow* ou outros dispositivos físicos ou conceituais que levam aos dados, bem como os metadados e as infraestruturas para a gestão de dados.

A arquitetura representada na figura 5, a seguir, é proposta com base na combinação dos modelos semânticos de agregação, de associação e classificação – em que apresenta as categorias gerais do modelo semântico de agregação em uma perspectiva *bottom-up*, a saber: governança, infraestrutura, serviços e FAIRificação. Essas categorias se associam a partir de uma relação de viabilidade, isto é, um componente viabiliza o outro. Dentro da categoria governança, seus componentes se associam a partir de uma relação de direcionamento, mostrando como um elemento principal pode direcionar os demais componentes, tendo como parâmetro estruturante uma política de gestão. Já a categoria infraestrutura revela a agregação de componentes necessários para a construção da infraestrutura da plataforma, que precisa estar permeada por uma infraestrutura organizacional, a saber: expertises, tecnologias, recursos informacionais e normas e padrões. Nesta categoria, mais uma vez em uma perspectiva *bottom-up*, observa-se as normas e padrões fundamentando as instâncias tecnológicas e informacionais, que se relacionam com as expertises em uma perspectiva *top down*, mostrando quem as opera. Na categoria *Serviços*, o modelo usado na construção da arquitetura é de agregação, mas os seus componentes se relacionam em uma linha temporal, isto é, os serviços científicos, computacionais, informacionais e administrativos se encontram na linha temporal que contorna todo o processo de pesquisa. Por último, a categoria *FAIRificação*, também em um modelo semântico de agregação, mostra as categorias em que os princípios FAIR se organizam como componentes necessários para promover a FAIRificação e, em um nível mais elevado, um ecossistema de dados e serviços FAIR.

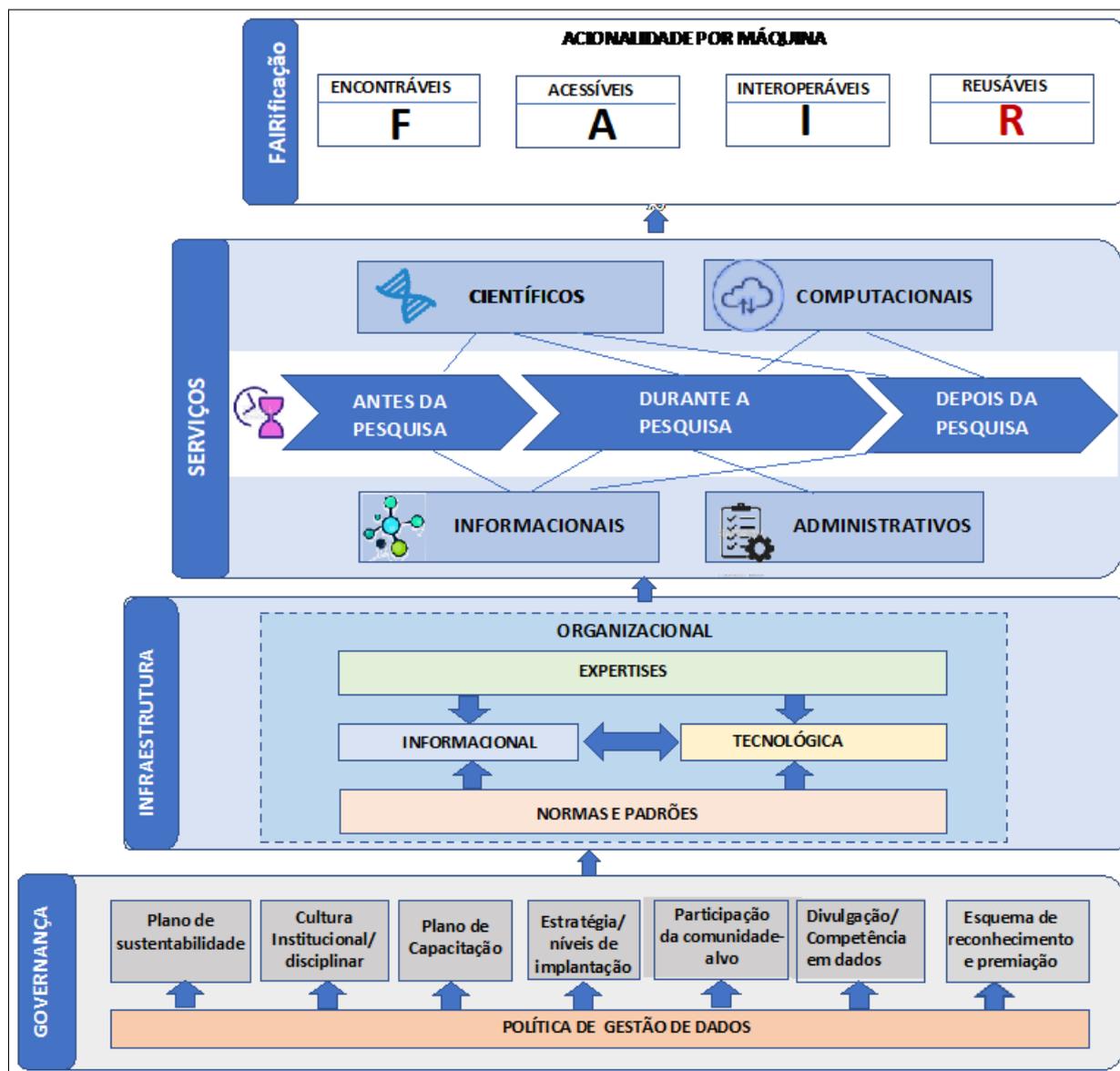


Figura 5. Modelo de Serviço de Gestão de Dados FAIR.

Fonte: Os autores (2022).

À GUIA DE CONCLUSÃO

A figura 5 apresentou uma visão geral dos componentes do modelo de serviço de gestão de dados proposto que foram agrupados em camadas, mostrando suas interrelações, a partir da discussão do texto.

Com esta perspectiva, o modelo proposto procurou desconstruir os blocos que compõem uma arquitetura genérica de uma plataforma de serviços de gestão de dados, articulando os vários módulos conceituais – diretrizes, políticas, serviços, ferramentas, infraestruturas etc. – na forma de peças que podem ser ajustadas de acordo com a profundidade, alcance e filosofia de cada instituição ou disciplina. Para ordenar esses esforços, por vezes entrópicos e difusos, a aplicação do FAIR, dimensionada pelos seus 15 princípios orientadores, aponta um horizonte para a construção eficiente de arquiteturas de plataformas de gestão de dados.

Mesmo tendo em conta o enfoque generalista do modelo, é preciso considerar que na implantação de práticas e infraestruturas de gestão de dados, o contexto específico das comunidades científicas e as possibilidades da adoção devem ser observadas. A importância de cada serviço pode depender das prioridades e da geração e uso de determinados objetos de pesquisa.

Neste viés, o presente modelo pode servir como um norte a ser usado para a constituição de uma possível escala de mensuração do nível de maturidade dos projetos de serviços de gestão, visto que a arquitetura proposta tem como horizonte tornar os dados aderentes aos Princípios FAIR, tornando mais concreta a ideia de uma Internet de Dados e Serviços FAIR.

Um modelo nunca pode ser um fim em si mesmo. Modelos devem ser aplicados, avaliados e melhorados. Novos conhecimentos obtidos após o desenvolvimento da presente pesquisa sobre outros trabalhos relacionados ao tema

serviços de gestão de dados podem ser considerados em uma possível continuação de estudos voltados para a melhoria desta proposta, como é o caso do Modelo DIAMANT (*Design and Information Architecture for Data MANagement Technologies*), que, conforme Marín-Arraiza, Vidotti, et al. (2019, p. 268) explicam, “vincula os serviços de diferentes seções dentro de uma instituição com as fases do projeto de pesquisa e a consequente necessidade de serviços aos pesquisadores em cada fase”. Este artigo é resultado de pesquisas em andamento desenvolvidas no âmbito do grupo de pesquisa BRIET (Biblioteconomia, Representação, Interoperabilidade, E-science e Tecnologia), sendo fruto de projeto de pesquisa “Gestão de dados de pesquisa FAIR: uma proposta de modelo para aceleração das pesquisas científicas no estado do Rio de Janeiro” financiado pela FAPERJ e de projeto de desenvolvimento “Infraestrutura de Apoio à Gestão e Preservação do Conhecimento Nuclear Brasileiro”, financiado pela CNEN e pelo CNPq.

REFERÊNCIAS

- Borgman, C. L. (2007). Data: the input and output of scholarship. In *Scholarship in the digital age: Information, infrastructure, and the internet*. London: The MIT Press. doi: 10.7551/mitpress/7434.003.0009
- Choudhury, S., Cowles, E., Croft, H. R., Estlund, K., Fary, M., Faustino, G., ... Menard, K. (2018). Research data curation: A framework for an institution-wide services approach. *EDUCAUSE Working Group on Data Curation*. Recuperado de <https://hsrc.himmelfarb.gwu.edu/libfacpubs/35>
- Coates, H. L. (2014). Building data services from the ground up: Strategies and resources. *Journal of eScience Librarianship*, 3(1), 52–59. doi: 10.7191/jeslib.2014.1063
- Erway, R., Horton, L., Nurnberger, A., Otsuji, R., & Rushing, A. (2016). *Building blocks: laying the foundation for a research data management program*. Dublin: OCLC. Recuperado de <https://files.eric.ed.gov/fulltext/ED589141.pdf>
- European Commission. (2020). *Six recommendations to implementation of fair practices*. Bruxelas: European Commission. doi: 10.2777/986252
- Fearon Jr., D., Gunia, B., Pralle, B. E., Lake, S., & Sallans, A. L. (2013). *Spec kit 334: Research data management services*. Washington, DC: Association of Research Libraries. doi: 10.29242/spec.334
- Gil, A. C. (2002). *Como elaborar projetos de pesquisa* (v. 4). Atlas São Paulo.
- Graaf, M., & Waaijers, L. (2011). *A surfboard for riding the wave: towards a four country action programme on research data*. Copenhagen: Knowledge Exchange. Recuperado de <http://hdl.voced.edu.au/10707/183860>
- Jones, S., Prior, G., & White, A. (2013). *How to develop research data management services – a guide for heis*. Edinburgh: Digital Curation Centre. Recuperado de <https://www.dcc.ac.uk/guidance/how-guides/how-develop-rdm-services>
- Kouper, I., Fear, K., Ishida, M., Kollen, C., & Williams, S. C. (2017). Research data services maturity in academic libraries. In *Curating research data: practical strategies for your digital repository* (v. 1, p. 153–170). Chicago: Association of College and Research Libraries. Recuperado de <https://experts.illinois.edu/en/publications/research-data-services-maturity-in-academic-libraries>
- Leonardi, P. M. (2010). Digital materiality? how artifacts without matter, matter. *First monday*, 15(6-7). doi: 10.5210/fm.v15i6.3036
- Marín-Arraiza, P., Vidotti, S., et al. (2019). Implementação de serviços institucionais de administração de dados. *Liinc em Revista*, 15(2). doi: 10.18617/liinc.v15i2.4819
- Mayernik, M. S., Choudhury, G. S., DiLauro, T., Metsger, E., Pralle, B., Rippin, M., & Duerr, R. (2012). The data conservancy instance: Infrastructure and organizational services for research data curation. *D-Lib Magazine*, 18(9/10). doi: 10.1045/september2012-mayernik
- Mons, B., Neylon, C., Velterop, J., Dumontier, M., da Silva Santos, L. O. B., & Wilkinson, M. D. (2017). Cloudy, increasingly fair; revisiting the fair data guiding principles for the european open science cloud. *Information services & use*, 37(1), 49–56. doi: 10.3233/ISU-170824
- Mushi, G. E., Pienaar, H., & Deventer, M. v. (2020). Identifying and implementing relevant research data management services for the library at the university of do-
doma, tanzania. *Data Science Journal*, 19(1), 1–9. doi: 10.5334/dsj-2020-001
- National Research Council. (2015). *Preparing the workforce for digital curation*. Washington, D.C: The National Academies Press. doi: 10.17226/18590
- Reed, R. B. (2015). Diving into data: planning a research data management event. *Journal of eScience Librarianship*, 4(1). doi: 10.7191/jeslib.2015.1071
- Sayão, L. F. (2001). Modelos teóricos em ciência da informação-abstração e método científico. *Ciência da informação*, 30(1), 82–91. doi: 10.1590/S0100-19652001000100010
- Sayão, L. F., & Sales, L. F. (2020). Afinal, o que é dado de pesquisa? *Biblos*, 34(2). Recuperado de <https://www.seer.furg.br/biblos/article/view/11875>
- Solomonides, A. (2019). Research data governance, roles, and infrastructure. In *Clinical research informatics* (p. 291–310). Springer. doi: 10.1007/978-3-319-98779-8_14
- Strasser, C. (2015). *Research data management*. Baltimore: NISO. Recuperado de <https://wiki.lib.sun.ac.za/images/2/24/PrimerRDM-2015-0727.pdf>
- Tang, R., & Hu, Z. (2019). Providing research data management (rdm) services in libraries: preparedness, roles, challenges, and training for rdm practice. *Data and Information Management*, 3(2), 84–101. doi: 10.2478/dim-2019-0009
- Tenopir, C., Birch, B., & Allard, S. (2012). *Academic libraries and research data services: Current practices and plans for the future*. Chicago, IL: Association of College and Research Libraries. Recuperado de https://trace.tennessee.edu/utk_dataone/20/
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... Bourne, P. E. (2016). The fair guiding principles for scientific data management and stewardship. *Scientific data*, 3(1), 1–9. doi: 10.1038/sdata.2016.18
- Wilson, J. A., Martinez-Urbe, L., Fraser, M. A., & Jeffreys, P. (2011). An institutional approach to developing research data management infrastructure. *The International Journal of Digital Curation*, 6(2). doi: 10.2218/ijdc.v6i2.203

Como citar este artigo (APA):

Sales, L. F. & Sayão, L. F. (2022). Proposta de modelo de serviço de gestão de dados de pesquisa². *AtoZ: novas práticas em informação e conhecimento*, 11, 1 – 13. Recuperado de: <http://dx.doi.org/10.5380/atoz.v11.85765>

²O presente artigo é uma parte atualizada e reorganizada do capítulo Sayão, L. F., Sales, L. F. Gestão de dados como serviço: proposta de um modelo. Publicado na coletânea organizada por Sales, L. F., Violla, C. M. M. (2021). *Informação digital e suas diversas abordagens pela ótica de um cientista da informação*. Rio de Janeiro: PPGCI IBICT-UFR. doi: 10.22477/9786589167136 – que reúne as publicações mais citadas do Professor Luís Fernando Sayão.

NOTAS DA OBRA E CONFORMIDADE COM A CIÊNCIA ABERTA

CONTRIBUIÇÃO DE AUTORIA

Papéis e contribuições	Luana Farias Sales	Luís Fernando Sayão
Concepção do manuscrito		X
Escrita do manuscrito	X	X
Metodologia	X	
Curadoria dos dados	X	
Discussão dos resultados	X	X
Análise dos dados	X	

FINANCIAMENTO

O(s) autor(es) declara(m) que esta pesquisa recebeu financiamento conforme dados indicados a seguir e o documento comprobatório foi anexado como documento suplementar: **Chamada: Universal 2018 – Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) – nº 430553/2018-8**

O(s) autor(es) concorda(m) em interagir diretamente com pareceristas responsáveis pela avaliação do manuscrito, dessa forma tornando a revisão por pares aberta.

Disponibilidade de Dados Científicos da Pesquisa

Os conteúdos subjacentes ao texto da pesquisa já estão disponíveis.

EQUIPE EDITORIAL

Editora/Editor Chefe

Paula Carina de Araújo (<https://orcid.org/0000-0003-4608-752X>)

Editora/Editor Associada/Associado

Helza Ricarte Lanz (<https://orcid.org/0000-0002-6739-2868>)

Editora/Editor de Texto Responsável

Fabiane Führ (<https://orcid.org/0000-0002-3723-050X>)

Seção de Apoio às Publicações Científicas Periódicas - Sistema de Bibliotecas (SiBi) da Universidade Federal do Paraná - UFPR

Editora/Editor de Layout

Felipe Lopes Roberto (<https://orcid.org/0000-0001-5640-1573>)