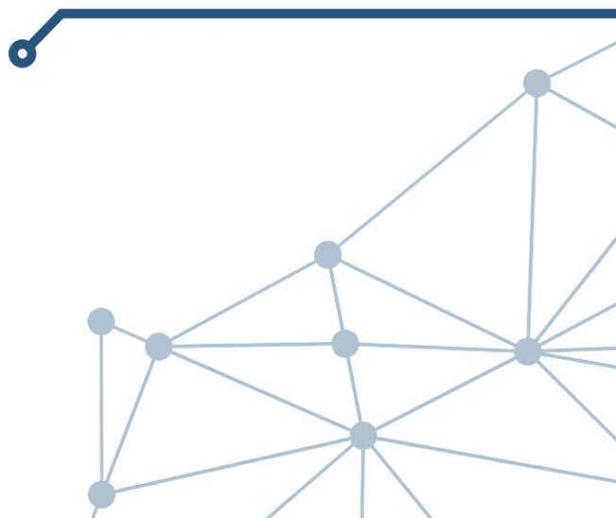


CAPÍTULO 9

INFRAESTRUTURA FEDERADA

MAICON ANÇA DOS SANTOS
MILTON SHINTAKU



9.1 INTRODUÇÃO

A Ciência Aberta, desde seus primeiros estudos, tem gradativamente incorporado novas proposições, visto que sempre se compreendeu esse movimento como um guarda-chuva que abrange diversos temas de investigação. Em todos os casos, a Ciência Aberta atua em todas as etapas da pesquisa, tornando-a mais colaborativa, transparente e democrática por meio do uso de tecnologias livres. Uma das evidências da variedade de temas abordados pela Ciência Aberta é sua taxonomia, revisada por Silveira et al. (2023), que revela a ampla gama de temas englobados por esse movimento.

Entre os diversos temas abrangidos pelo guarda-chuva da Ciência Aberta, este capítulo destaca a infraestrutura federada, inserida em um tema mais amplo: a infraestrutura e as ferramentas de suporte à Ciência Aberta. Isso exige um entendimento aprofundado, já que, ao situar-se no âmbito da Ciência Aberta, essa infraestrutura assume novos contornos, sendo destinada ao contexto de pesquisas, o que pode implicar restrições ou ampliar compreensões.

A palavra *infraestrutura*, por si só, assume diferentes aspectos quando utilizada em linguagens especializadas, tornando-se um termo amplamente empregado na informática e na gestão pública. No Brasil, houve duas ocasiões em que existiu um Ministério da Infraestrutura, voltado principalmente para temas de transporte em todas as suas dimensões. Já na informática, é comum referir-se como infraestrutura ao conjunto de equipamentos e artefatos que possibilitam o funcionamento de redes.

No caso de infraestrutura para pesquisa, o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), em sua página de perguntas frequentes¹, descreve-a da seguinte forma:

¹ Infraestruturas de Pesquisa. Disponível em: https://lattes.cnpq.br/web/diip/ajuda.jsessionid=Nty7O5aJr15dgmYYZkx0pn7G.undefinid?p_p_id=54_INSTANCE_K5JVRIZKWcuF&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-2&p_p_col_count=1&_54_INSTANCE_K5JVRIZKWcuF_struts_action=%2Fwiki_display%2Fview. Acesso em: 21 out. 2024.



Conjunto de instalações físicas e condições materiais de apoio (equipamentos e recursos) utilizados pelos pesquisadores para a realização de atividades de P&D. Esse conceito envolve os seguintes elementos:

- Instalações físicas (usualmente imóveis) que abrigam os equipamentos e instrumentos usados nas atividades de P&D,
- Principais equipamentos e instrumentos utilizados em atividades de P&D;
- Recursos baseados em conhecimento (como bibliotecas, coleções, arquivos e base de dados) utilizados em pesquisas científicas;
- Recursos de tecnologia da informação e comunicação (como grids, redes de alto desempenho e softwares específicos).” (Conselho Nacional de Desenvolvimento Científico e Tecnológico, 2013, p. 1).

Ao receber o complemento "federada", o termo ganha novos contornos, pois uma infraestrutura federada incorpora as restrições sugeridas pelo adjetivo. Nesse contexto, *federada* é uma forma particular de derivar o adjetivo *federal* para um adjetivo-substantivo, por meio do sufixo "-ada", que indica filiação. Assim, *federada* refere-se a alguém ou algo que faz parte de uma federação. Portanto, para compreender o termo *federada*, é necessário primeiro entender o conceito de federação.

Em grande parte, falar em federação remete às ciências políticas, nas quais muitos países são considerados estados federais. Nessas configurações, as federações são compostas por unidades territoriais autônomas, mas vinculadas por meio de legislação. Em outras áreas do conhecimento, o entendimento se mantém: federação refere-se à união de unidades autônomas, unidas por um estatuto, geralmente motivadas por atuação ou interesses comuns.



No contexto da Ciência Aberta, a infraestrutura federada pode ser entendida como um conjunto de recursos que apoia atividades de pesquisa pertencentes a membros de uma composição de unidades autônomas, unidas oficialmente. Dessa forma, instituições de ensino e pesquisa podem formar federações para compartilhar parte de sua infraestrutura, ampliando as possibilidades de atuação por meio do uso de recursos compartilhados.

■ 9.2 INFRAESTRUTURA FEDERADA

A Ciência Aberta tem promovido, entre outros aspectos, uma colaboração mais estreita entre pesquisadores e instituições de pesquisa. Dessa forma, a abertura das atividades de pesquisa não se limita apenas a conferir transparência aos processos científicos, mas também visa a uma execução mais democrática da pesquisa, com a participação de pessoas externas à academia e de pesquisadores vinculados a diversas instituições. Do mesmo modo, a Ciência Aberta incentiva o compartilhamento de recursos pelas próprias instituições, tornando-os mais acessíveis a indivíduos externos.

Nesse sentido, apresentam-se as infraestruturas federadas, nas quais instituições podem se unir, formando federações para o compartilhamento de recursos. Dessa forma, amplia-se a disponibilização de recursos sem a necessidade de novas aquisições, promovendo um melhor uso dos recursos existentes. Essa otimização gera benefícios mútuos para todos os membros da federação, pois infraestruturas de pesquisa podem não estar em uso constante, permitindo que o tempo ocioso seja aproveitado por outras instituições.

A importância de uma infraestrutura federada é destacada nos apontamentos de Guimarães e Bessa (2023) sobre o Programa de Emergência para a Soberania Digital, uma proposta elaborada por acadêmicos



de diversas universidades brasileiras e formalizada em uma carta entregue ao candidato à presidência em 2022. No documento, os acadêmicos propõem:

[...] a criação de uma infraestrutura federada para repositório de dados acadêmicos, o financiamento de data centers, a formação de equipe multidisciplinar no MCTI para buscar soluções que considerem a tecno-diversidade brasileira [...] (Guimarães; Bessa, 2023, p. 16).

Outro ponto que corrobora essa percepção são os relatos de Fortaleza e Bertin de Drucker (2022) sobre os desafios de implementação das Ciência Aberta no âmbito do governo aberto (*Open Government Partnership*). As autoras ressaltam que a criação de uma infraestrutura nacional de apoio à Ciência Aberta se insere em um cenário desejado, firmado para ser uma das metas do compromisso 3, liderado pela Rede Nacional de Pesquisa (RNP), com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).

Bastiaansen et al. (2020) descrevem, em termos tecnológicos, os requisitos para a criação de infraestruturas federadas baseadas em conjuntos interligados de computadores em cenários heterogêneos, utilizando nuvens de recursos voltados para processamento, armazenamento e comunicação de dados em contextos militares, apresentando soluções práticas. Muitos desses requisitos, evidentemente, podem ser aplicados ao contexto de pesquisa, considerando que a própria internet nasceu em um ambiente militar e atualmente é amplamente utilizada.

Para os requisitos tecnológicos de monitoramento de infraestruturas federadas, Al-Hazmi e Magedanz (2015) defendem um monitoramento distribuído baseado em protocolos que sejam escaláveis, flexíveis e genéricos o suficiente para transportar diversos conjuntos de dados de



medição. Isso permite que cada ente federado colete e interprete conjuntos de dados tanto da federação quanto de cada unidade federativa.

Grande parte dos relatos sobre infraestrutura federada está relacionada a questões tecnológicas. Nesse sentido, Ganzinger et al. (2023) afirmam que, para a implementação de uma infraestrutura federada, o maior desafio reside em sua complexidade, na qual cada membro deve ter sua própria instância, mas com a possibilidade de troca de informações. Assim, cada membro precisa se preocupar com questões de hardware e software para garantir que os serviços possam ser executados de forma distribuída, assegurando a qualidade da oferta.

Em resumo, pode-se afirmar que uma infraestrutura federada possibilita o compartilhamento de recursos, principalmente computacionais, mantidos por diversas instituições, com o intuito de oferecer serviços a usuários e gestores, aumentando a disponibilidade de recursos e permitindo o monitoramento do desempenho e uso. Nesse sentido, cada membro da federação disponibiliza seus recursos com hardware e software adaptados a esse tipo de integração, que permite o compartilhamento enquanto mantém a independência de cada instituição.

9.3 INICIATIVAS DE INFRAESTRUTURA FEDERADA

A partir do entendimento sobre infraestrutura federada, é possível identificar suas iniciativas, especialmente no Brasil. Isso permite verificar como as instituições estão atuando de forma colaborativa ao disponibilizar recursos de infraestrutura de maneira federada. É evidente que a criação de infraestruturas federadas pode ocorrer fora do contexto da



pesquisa, voltando-se para outros objetivos. No entanto, o foco desta seção está voltado para a Ciência Aberta e o governo, restrito a esse âmbito.

Esse ponto está alinhado ao que é apresentado no Quarto Plano de Ação Brasileiro para o Governo Aberto (2018-2021), liderado pela Controladoria-Geral da União (CGU), que inicia a discussão sobre o tema como uma das metas a serem cumpridas para mensurar o compromisso. Assim, no Marco Sete, temos a “Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa.” Esse marco teve como responsáveis a Rede Nacional de Pesquisa (RNP), com o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) e da Empresa Brasileira de Pesquisa Agropecuária (Embrapa).

Seguindo essa diretriz, a Embrapa, em seu manual de Normas, incluiu a Infraestrutura Federada na Política de Governança de Dados, Informação e Conhecimento², especificamente na seção que trata de seus princípios norteadores. Dessa forma, a empresa incentiva o compartilhamento de infraestrutura tecnológica entre suas unidades, visando criar uma estrutura federada para a oferta de serviços de gestão de dados, além de fomentar o uso e reúso dessas informações.

Veiga et al. (2021) relatam a criação da *Virus Outbreak Data Network* (VODAN), uma infraestrutura federada composta por diversos membros, incluindo mais de 20 países ao redor do mundo, coordenada pela Fundação Oswaldo Cruz (Fiocruz). Entre os membros estão a Universidade Federal do Rio de Janeiro (UFRJ), a Universidade Federal do Estado do Rio de Janeiro (UNIRIO), o Instituto Nacional do Câncer (INCA) e o Instituto de Cardiologia, além do Hospital Albert Einstein de São Paulo e do Hospital Universitário Gaffrée Guinle da UNIRIO, no Rio de Janeiro. Segundo os autores, o VODAN tem como objetivo:

² Manual de Normas da Embrapa. Disponível em: <https://www.embrapa.br/documents/10180/1600893/Pol%C3%ADtica+de+Governan%C3%A7a+de+Dados%2C+Informa%C3%A7%C3%A3o+e+Conhecimento+da+Embrapa/96873780-8e-04-b67d-be7c-a0b6a4318556>. Acesso em: 21 out. 2024.



[...] estabelecer uma infraestrutura de dados federada alinhada aos princípios FAIR e que apoie a coleta de dados de prontuários de pacientes infectados por vírus de alto contágio (Veiga et al., 2021, p. 51).

Para tanto, oferece uma infraestrutura para compartilhamento de dados, conforme os preceitos do FAIR (*Findable, Accessible, Interoperable e Reusable*).

Outra iniciativa de infraestrutura federada é mencionada por Lima et al. (apud Britto, 2024), denominada *EU Brazil Cloud Connect*, que visa desenvolver pesquisas em computação em nuvem para a criação de uma infraestrutura federada intercontinental, por meio da colaboração entre a União Europeia e o Brasil. Cunha et al. (2016, p. 1) relatam que esse projeto tem:

O objetivo principal é fornecer uma infraestrutura para difusão dos dados obtidos por sensores orbitais e ferramentas para análise multitemporal de tendências de mudanças na superfície terrestre.

Esse ponto revela como o compartilhamento de recursos computacionais pode ocorrer independentemente das fronteiras geográficas. É evidente que a Europa, como uma grande comunidade de membros com interesses políticos e econômicos comuns, possui maior possibilidade de criar infraestruturas federadas, com a participação de instituições dos países membros. Otto (2022), por exemplo, relata sobre o “Projeto Gaya” para dados espaciais europeus, que oferece serviços federados de identificação e confiança, catálogo, intercâmbio de dados e funcionalidades para verificar e ajustar as conformidades das infraestruturas em nuvem dos membros.

Jagtap et al. (2021) relatam sobre o “Projeto Helium,” criado para fornecer acesso sem fio de baixa potência à Internet das Coisas. Essa federação é composta por mais de 40 mil pontos de acesso ativos, com



1.000 novos pontos podendo ser adicionados diariamente, todos atuando em nuvem. Assim, o Helium se apresenta como um provedor de acesso de banda larga em baixa potência, compartilhada e federada (*crowdsourcing*), construindo uma rede expansível, mas robusta o suficiente para garantir a segurança dos usuários.

Wallom et al. (2015) descrevem o cenário do projeto *EGI Federated Cloud*, que oferece infraestrutura federada em nuvem para a comunidade europeia, criado em 2014. Inicialmente voltado para membros acadêmicos, o projeto gradualmente atraiu parceiros comerciais, ampliando sua cobertura. Com a utilização de padrões abertos, busca obter maior impacto social e econômico.

Dessa forma, nota-se que várias iniciativas implementam os preceitos das infraestruturas federadas em todo o mundo, não sendo uma total novidade, uma vez que há relatos sobre esse tema desde 2002, como nos estudos sobre a *Verity Federated Infrastructure* de Choo et al. (2002) e sobre o *Enabling Grids for E-Science (EGGE)*, apresentado por Jones (2002). Em ambos os casos, tratam-se de projetos que envolvem conceitos relacionados à infraestrutura federada para ofertar recursos de forma compartilhada.

Nesse contexto, observa-se um certo amadurecimento sobre o tema, adotado pela Ciência Aberta, que incorpora grande parte dos preceitos de abertura, compartilhamento, democratização e inclusão, entre outros aspectos evidenciados por esse movimento. No Brasil, destaca-se o papel da Controladoria Geral da União (CGU) e os esforços para integrar a Ciência Aberta ao governo, por meio de sua atuação com a *Open Government Partnership (OGP)*, seus compromissos e parceiros, entre os quais se destaca a Fiocruz.



9.4 CONSIDERAÇÕES FINAIS

Na taxonomia proposta por Silveira et al. (2023) sobre a Ciência Aberta, emergem novos temas que se agregam ao movimento, juntamente com ajustes necessários para o novo cenário. Em alguns casos, temas já existentes são adaptados para a ciência, dado que discussões, estudos, tecnologias e outras iniciativas já existiam em contextos distintos. Um exemplo disso é a infraestrutura federada, que apresenta iniciativas na área da informática e, em alguns casos, fornece recursos para várias áreas, inclusive para a pesquisa.

Evidentemente, ao se considerar o compartilhamento de recursos oriundos de múltiplas organizações e instituições, remete-se à formação de consórcios e redes, onde os membros se fortalecem por meio do compartilhamento. Assim, os estudos sobre infraestruturas federadas levam em conta tanto os membros quanto os objetivos que motivam a associação para a oferta de recursos compartilhados. É importante ressaltar que esses recursos não se limitam apenas aos computacionais, podendo englobar diversos tipos, como equipamentos e laboratórios.

Por fim, observa-se que a Ciência Aberta visa promover a colaboração em todos os aspectos relacionados à pesquisa, envolvendo não apenas as atividades, mas também tudo o que diz respeito aos estudos. No caso específico das infraestruturas federadas, destaca-se a importância do compartilhamento de recursos físicos, permitindo que os membros da federação, e até mesmo toda a população, possam colher os resultados dessa união, na medida em que possibilita um melhor aproveitamento dos recursos disponíveis.



REFERÊNCIAS

AL-HAZMI, Yahya; MAGEDANZ, Thomas. Towards semantic monitoring data collection and representation in federated infrastructures. In: INTERNATIONAL CONFERENCE ON FUTURE INTERNET OF THINGS AND CLOUD, 3., 2015, Roma. **Anais** [...]. Roma: IEEE, 2015. p. 17-24. DOI: doi: 10.1109/FiCloud.2015.40. Disponível em: <https://ieeexplore.ieee.org/document/7300792>. Acesso em: 21 out. 2024.

BASTIAANSEN, Harrie. GEEST, Johan van der Geest; BROEK, Casper Van Den; KUDLA, Thomas. Federated control of distributed multi-partner cloud resources for adaptive c2 in disadvantaged networks. **IEEE Communications Magazine**, [S. l.], v. 58, n. 8, p. 21-27, 2020.

CHOO, Kiam; MUKHERJEE, Rajat; SAMAIR, Rami; ZHANG, Wei. The Verity federated infrastructure. In: INTERNATIONAL CONFERENCE ON INFORMATION AND KNOWLEDGE MANAGEMENT, 11., 2002, New York. **Proceedings** [...]. New York: Association for Computing Machinery, 2002. p. 621-621. DOI: <https://doi.org/10.1145/584792.584897>.

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO. **Infraestruturas de Pesquisa**. Brasília, DF: CNPq, 2013. Disponível em: https://lattes.cnpq.br/web/diip/ajuda?p_p_id=54_INSTANCE_K5JVRIZKWcuF&p_p_lifecycle=0&p_p_state=pop_up&p_p_mode=view&p_p_col_id=column-2&p_p_col_count=1&_54_INSTANCE_K5JVRIZKWcuF_struts_action=%2Fwiki_display%2Fview&_54_INSTANCE_K5JVRIZKWcuF_noDeName=Main&_54_INSTANCE_K5JVRIZKWcuF_title=Infraestruturas+de+pesquisa&_54_INSTANCE_K5JVRIZKWcuF_viewMode=print. Acesso em: 21 out. 2024.

CUNHA, John Elton de Brito Leite; RUFINO, Iana Alexandra Alves; GALVÃO, Carlos de Oliveira; PEREIRA, Thiago Emmanuel; BRASILEIRO, Francisco Vilar; PEREIRA, Esdras Vidal. Difusão de dados orbitais nos estudos de mudanças na cobertura do solo no semiárido brasileiro. In: SIMPÓSIO DE RECURSOS HÍDRICOS DO NORDESTE, 13., 2016, Aracajú. **Anais** [...]. Aracajú: ABRH, 2016. p. 1-10. Disponível em:



<https://files.abrhidro.org.br/Eventos/Trabalhos/27/PAP021727.pdf>.
Acesso em: 21 out. 2024.

FORTALEZA, Juliana Meireles; BERTIN, Patrícia Rocha Bello; DRUCKER, Debora Pignatari. O compromisso pela Ciência Aberta na Open Government Partnership: avanços na governança de dados científicos no Brasil. In: PRÍNCIPE, Eloísa; RODE, S. de M. (org.). **Comunicação científica aberta**. Rio de Janeiro: Instituto Brasileiro de Informação em Ciência, 2022. p. 189-206. 2022. DOI: 10.21452/ABEC.2022. Disponível em: <https://www.alice.cnptia.embrapa.br/alice/bitstream/doc/1151882/1/PL-Compromisso-ciencia-2022.pdf>. Acesso em: 21 out. 2024.

GANZINGER, Matthias et al. Federated electronic data capture (fEDC): architecture and prototype. **Journal of Biomedical Informatics**, [S. l.], v. 138, p. 1-7, 2023. DOI: <https://doi.org/10.1016/j.jbi.2023.104280>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1532046423000011>. Acesso em: 21 out. 2024.

GUIMARÃES, Jacqueline Moreno Gomes Guimarães; BESSA, Luiz Fernando Maceda. A soberania digital construída pela sociedade civil brasileira diante da agenda das Cidades Inteligentes. **Revista Campo de Públicas: conexões e experiências**, [S. l.], v. 2, n. 2, p. 86-114, 2023. Disponível em: <https://revista.fjp.mg.gov.br/index.php/campo-de-publicas/article/view/38/33>. Acesso em: 21 out. 2024.

JAGTAP, Dhananjay; YEN, Alex; WU, Huanlei; SCHULMAN, Aaron; PANNUTO, Pat. Federated infrastructure: usage, patterns, and insights from "the people's network". In: ACM INTERNET MEASUREMENT CONFERENCE, 21., 2021, Nova York. **Proceedings** [...]. Nova York: Association for Computing Machinery, 2021. p. 22-36. DOI: <https://doi.org/10.1145/3487552.3487846>.

JONES, Bob. **EGEE: Enabling grids for E-Science**. Gêneve: CERN, 2002.

LIMA, S. J. et al. Medição da economia baseada em dados: impactos, desafios e oportunidades para o Nordeste brasileiro. In: BRITTO, Jorge. **Nota técnica 4**. Rio de Janeiro: RedeSist; Cicef, 2024.



OTTO, Boris. A federated infrastructure for European data spaces. **Communications of the ACM**, [S. l.], v. 65, n. 4, p. 44-45, 2022. DOI: <https://doi.org/10.1145/3512341>.

SILVEIRA, Lúcia da; RIBEIRO, Nivaldo Calixto; MELERO, Remédios; MORA-CAMPOS, Andrea; PIRAQUIVE-PIRAQUIVE, Daniel Fernando; URIBE-TIRADO, Alejandro; SENA, Priscila Machado Borges; POLANCO-CORTÉS, Jorge; SANTILLÁN-ALDANA, Julio; SILVA, Fabiano Couto Corrêa da; ARAÚJO, Ronaldo Ferreira; ENCISO-BETANCOURT, Andrés Mauricio; FACHIN, Juliana. Taxonomia da Ciência Aberta: revisada e ampliada. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, v. 28, p. 1–22, 2023. DOI: 10.5007/1518-2924.2023.e91712. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/91712>. Acesso em: 13 set. 2024.

VEIGA, Viviane; CAMPOS, Maria Luiza; SILVA, Carlos Roberto Lyra da; HENNING, Patrícia Corrêa; MOREIRA, João. VODAN BR: a gestão de dados no enfrentamento da pandemia coronavírus. **Páginas a&b: arquivos e bibliotecas**, Porto, v. 3, n. especial, p. 51-58, 2021. Disponível em: <https://www.arca.fiocruz.br/handle/icict/46443>. Acesso em: 21 out. 2024.

WALLOM, David; TURILLI, Matteo; DRESCHER, Michel; SCARDARCI, Diego. Federating infrastructure as a service cloud computing systems to create a uniform e-infrastructure for research. In: IEEE International Conference on e-Science, 11., 2015, Sydney. **Proceedings [...]**. Sydney: IEEE, 2015. p. 155-164.

COMO CITAR ESTE CAPÍTULO:

SANTOS, Maicon Ança dos; SHINTAKU, Milton. Infraestrutura federada. In: DRUCKER, Debora Pignatari; CIUFFO, Leandro; SAYÃO, Luis Fernando; SHINTAKU, Milton; VIDOTTI, Silvana Aparecida Borsetti Gregorio (org.) **Infraestruturas de suporte à Ciência Aberta**. Brasília, DF: Editora Ibict, 2025. p. 198-210. DOI: 10.22477/9786589167754. cap9.