

# Compartilhamento de dados de pesquisa na Fiocruz: diagnóstico e percepção do pesquisador

## **Viviane Santos de Oliveira Veiga**

Doutora em Informação e Comunicação em Saúde pela Fundação Oswaldo Cruz (Fiocruz) - Brasil, com período sanduíche em Universidade de Coimbra (UC) – Portugal. Tecnologista em Saúde Pública da Fundação Oswaldo Cruz (Fiocruz) - Brasil. Coordenadora da Rede de Bibliotecas da Fundação Oswaldo Cruz (Fiocruz) – Brasil.

<http://lattes.cnpq.br/4983074089687751>

E-mail: [vivianesantosveiga@gmail.com](mailto:vivianesantosveiga@gmail.com)

## **Rejane Ramos Machado**

Doutora em Ciências pela Fundação Oswaldo Cruz (Fiocruz/Icict) - Rio de Janeiro, RJ - Brasil. Pesquisadora da Fundação Oswaldo Cruz (Fiocruz/Icict/Licts) - Rio de Janeiro, RJ – Brasil.

<http://lattes.cnpq.br/2263131940838871>

E-mail: [rejane.machado@icict.fiocruz.br](mailto:rejane.machado@icict.fiocruz.br)

## **Vanessa de Arruda Jorge**

Doutora em Ciência da Informação pelo Programa de Pós-graduação em Ciência da Informação (PPGCI), convênio Universidade Federal do Rio de Janeiro e Instituto Brasileiro de Informação em Ciência e Tecnologia (UFRJ/Ibict) - Rio de Janeiro, RJ - Brasil. Pesquisadora da Fundação Oswaldo Cruz (Fiocruz/INCQS) - Rio de Janeiro, RJ – Brasil.

<http://lattes.cnpq.br/0218139692140149>

E-mail: [vanessa.arruda@incqs.fiocruz.br](mailto:vanessa.arruda@incqs.fiocruz.br)

## **Erick Penedo**

Especialização em Gestão de Projetos pela Universidade Estácio de Sá (Unesa) - Brasil. Analista de Informação e Comunicação da Fundação Oswaldo Cruz (Fiocruz/Icict/Licts) - Rio de Janeiro, RJ - Brasil.

<http://lattes.cnpq.br/6085077532205244>

E-mail: [erick.penedo@icict.fiocruz.br](mailto:erick.penedo@icict.fiocruz.br)

Data de submissão: 27/09/2019. Data de aprovação no ConfOA:12/09/2019. Data de publicação:

## **RESUMO**

Este artigo apresenta os resultados parciais de pesquisa em andamento para caracterizar os dados de pesquisa produzidos na Fiocruz e analisar a percepção dos pesquisadores quanto ao compartilhamento e abertura desses dados. Foi utilizado como instrumento de coleta de dados um questionário on-line, aplicado aos pesquisadores da Fiocruz. A análise dos dados obtidos foi baseada no Modelo de Fatores que influenciam a prática de compartilhar dados de pesquisa. Os resultados parciais indicam as barreiras que interferem na coparticipação de dados de pesquisa e os estímulos à adesão ao movimento Ciência Aberta. É traçado um diagnóstico sobre os dados de pesquisa gerados pelos participantes, que pode subsidiar estratégias para o desenvolvimento de infraestrutura para o compartilhamento de dados de pesquisa na Fiocruz. Verificou-se que a principal barreira é a desinformação dos tipos de compartilhamento das plataformas existentes e das normas vigentes. A maioria dos pesquisadores da Fiocruz compreende a importância do compartilhamento e abertura dos dados de pesquisa, principalmente para a transparência na conduta da pesquisa e otimização do avanço científico.

**Palavras-chave:** Comportamento de compartilhamento de dados. Ciência Aberta. Dados de pesquisa.

## **Research data sharing at Fiocruz: researcher diagnosis and perception**

### **ABSTRACT**

*This article aims to present the partial results of ongoing research to characterize the research data produced at Fiocruz and to analyze the researchers' perception regarding the sharing and opening of these data, identifying the barriers that interfere with the sharing and opening of research data and the stimuli that can optimize adherence to the Open Science movement, building a diagnosis to support strategies for data management, sharing and openness at Fiocruz. The online questionnaire was used as a data collection instrument. In this paper we will present results and analysis of some questions in the questionnaire. The analysis of the obtained data was based on the Model of Factors that influence the research data sharing behavior. The instrument was applied to Fiocruz researchers. As a result, it was found that the main barrier is the desinformación, the forms of sharing, the existing platforms, and the current rules. Most Fiocruz researchers understand the importance of sharing and opening research data mainly for transparency in research conduct and optimization of scientific advancement. **Keywords:** Data Sharing Behavior; Open Science; Search data.*

**Keywords:** Data sharing behavior. Open Science. Search data.

## **Compartir datos de investigación en Fiocruz: diagnóstico y percepción del investigador**

### **RESUMEN**

*Este artículo tiene como objetivo presentar los resultados parciales de la investigación en curso para caracterizar los datos de investigación producidos en Fiocruz y analizar la percepción de los investigadores sobre el intercambio y la apertura de estos datos, identificando las barreras que interfieren con el intercambio y la apertura de datos de investigación y los estímulos. Que puede optimizar la adhesión al movimiento Open Science, construyendo un diagnóstico para apoyar estrategias de gestión de datos, intercambio y apertura en Fiocruz. El cuestionario se utilizó como instrumento de recopilación de datos. En este artículo presentaremos resultados y análisis de algunas preguntas en el cuestionario. El análisis de los datos obtenidos se basó en el Modelo de Factores que influye en el comportamiento de intercambio de datos de investigación. El instrumento se aplicó a investigadores de la Fiocruz. Como resultado, se descubrió que la barrera principal es la falta de conocimiento, las formas de compartir, las plataformas existentes y las reglas actuales. La mayoría de los investigadores de Fiocruz entienden la importancia de compartir y abrir datos de investigación principalmente para la transparencia en la conducta de investigación y la optimización del avance científico.*

**Palabras clave:** Comportamiento de intercambio de datos. Ciencia Abierta. Datos de encuesta.

## INTRODUÇÃO

Os dados de pesquisa adquiriram maior destaque na comunicação científica contemporânea, demandando novos estudos sobre sua gestão e compartilhamento, assim como a necessidade de considerar a percepção do pesquisador, responsável por sua produção e/ou coleta. Práticas inovadoras carecem de novos estudos e processos de escuta, promovendo uma reflexão crítica de sua adoção.

A literatura ressalta uma resistência do pesquisador em compartilhar os seus dados, mesmo com a imposição de políticas mandatárias (VEIGA, 2017). Assim, surgem diversas questões de estudo sobre essas temáticas. Por exemplo, quais barreiras têm afastado o pesquisador do compartilhamento e abertura de dados de pesquisa? Quais seriam os benefícios, na visão do pesquisador, ao compartilhar seus dados?

Trazendo essa perspectiva para o setor da saúde, considera-se que o acesso à informação e aos dados de pesquisa é fundamental para produção e disseminação do conhecimento, favorecendo o progresso da ciência nesse setor, o que pode ajudar a minimizar as iniquidades em saúde. Uma das maneiras de maximizar o acesso é por meio do compartilhamento de informação, como proposto por Hall (2001). O movimento de Ciência Aberta está sendo discutido e implementado em diversas instituições no mundo.

É fundamental sublinhar que o acesso aos dados de pesquisa em saúde pode diminuir o sofrimento de pessoas e suas doenças e contribuir para a garantia do direito à saúde. Portanto, compreender as barreiras à adesão ao movimento de Ciência Aberta para tentar minimizá-las significa contribuir para a otimização da produção de conhecimento em saúde no Brasil e possivelmente salvar vidas.

A Fiocruz é uma organização de ensino, pesquisa e assistência que promove, no território nacional, a saúde e o desenvolvimento social, gerando e difundindo conhecimentos científicos e tecnológicos. Em termos institucionais, está comprometida em oferecer respostas ágeis às questões de saúde da população através de projetos colaborativos e abertos. A instituição tem desenvolvido nos últimos 10 anos uma série de estudos e debates sobre a abertura do conhecimento científico. Em 2007 iniciou o seu repositório, o Arca, fruto de projeto de pesquisa. Em 2011 esse repositório foi institucionalizado, e em 2014 a *Política de Acesso Aberto ao Conhecimento* foi publicada.

No âmbito da gestão e compartilhamento de dados, desde 2015 há estudos em desenvolvimento, sendo o primeiro deles a tese de doutorado do Programa de Pós-Graduação em Informação Científica e Tecnológica em Saúde, intitulada “Percepção dos pesquisadores portugueses e brasileiros da área de neurociências quanto ao compartilhamento de artigos científicos e dados de pesquisa no acesso aberto verde: custos, benefícios e fatores contextuais”<sup>1</sup>, defendida em 2017 e ganhadora do prêmio Oswaldo Cruz de Tese<sup>2</sup>. Ainda em 2017 foi criado o Grupo de Trabalho em Ciência Aberta, coordenado pela Vice-presidência de Ensino, Informação e Comunicação da Fiocruz e o Grupo de trabalho e pesquisa de Dados de pesquisa da Rede de Bibliotecas Fiocruz, mostrando a necessidade de se conhecer o universo dos dados de pesquisa da instituição para apoiar ações e decisões estratégicas, posicionando-se diante desse movimento de abertura dos dados de pesquisa.

Este artigo objetiva apresentar resultados parciais de uma pesquisa coordenada por pesquisadores do Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT) e Instituto Nacional de Controle de Qualidade em Saúde (INCQS) para caracterização dos dados de pesquisa produzidos na Fiocruz e percepção dos pesquisadores quanto ao compartilhamento desses dados.

## METODOLOGIA

O percurso metodológico perpassa por reflexões a partir de conceitos e métodos das questões de pesquisa anteriormente mencionadas. Para tal, busca-se, a partir do estofo teórico, criar possibilidades para ratificar as questões de pesquisa. Nesse processo, faz-se necessário mapear as características dos dados de pesquisa produzidos pela Fiocruz e identificar custos e benefícios que interferem no compartilhamento de dados de pesquisa.

Foram analisados instrumentos utilizados nacional e internacionalmente para a elaboração de diagnóstico dos dados de pesquisa de uma instituição e análise da percepção de seus pesquisadores quanto à gestão, compartilhamento e abertura de dados de pesquisa.

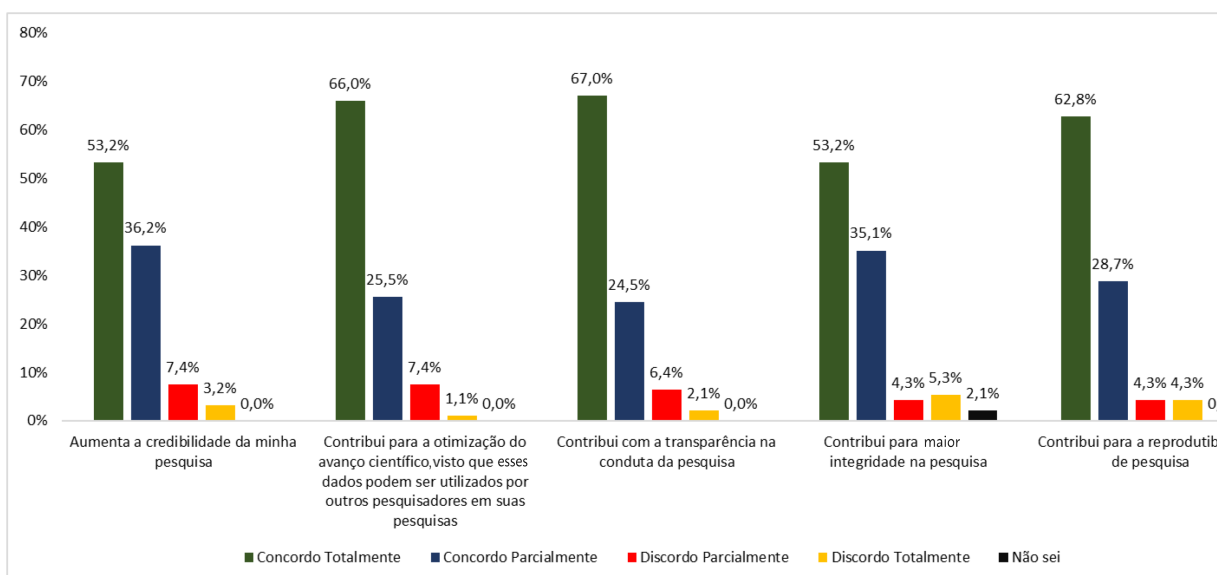
Elaborou-se como instrumento de coleta de dados um questionário on-line, utilizando o Google Forms. O questionário foi aplicado de 22 de dezembro de 2018 a 7 de junho de 2019 (ainda em processo de coleta de dados) para os pesquisadores pertencentes aos grupos de pesquisa da Fiocruz, obtendo-se 94 respostas até o momento.

As análises dos dados foram baseadas no Modelo de Fatores que influenciam no comportamento de compartilhamento de dados de pesquisa, o MFDados (VEIGA, 2017). Para apresentação neste artigo, elegeram-se seis questões do instrumento para análise e discussão.

## ANÁLISE E DISCUSSÃO DOS RESULTADOS

A primeira análise refere-se à percepção dos pesquisadores da Fiocruz sobre as atitudes e valores em relação à abertura de dados na comunicação científica. Verificou-se que 67% concordam plenamente que a abertura contribui com a transparência na conduta da pesquisa; 66% concordam plenamente que a abertura de dados contribui para a otimização do avanço científico, visto que esses dados podem ser utilizados por outros pesquisadores; 62% concordam plenamente que a abertura de dados contribui para a reprodutibilidade da pesquisa; 53% concordam plenamente que a abertura de dados contribui para o aumento da credibilidade da própria pesquisa e para maior integridade na pesquisa. Esses achados mostram que a maioria dos pesquisadores da Fiocruz percebe os benefícios da abertura de dados de pesquisa no processo de comunicação científica.

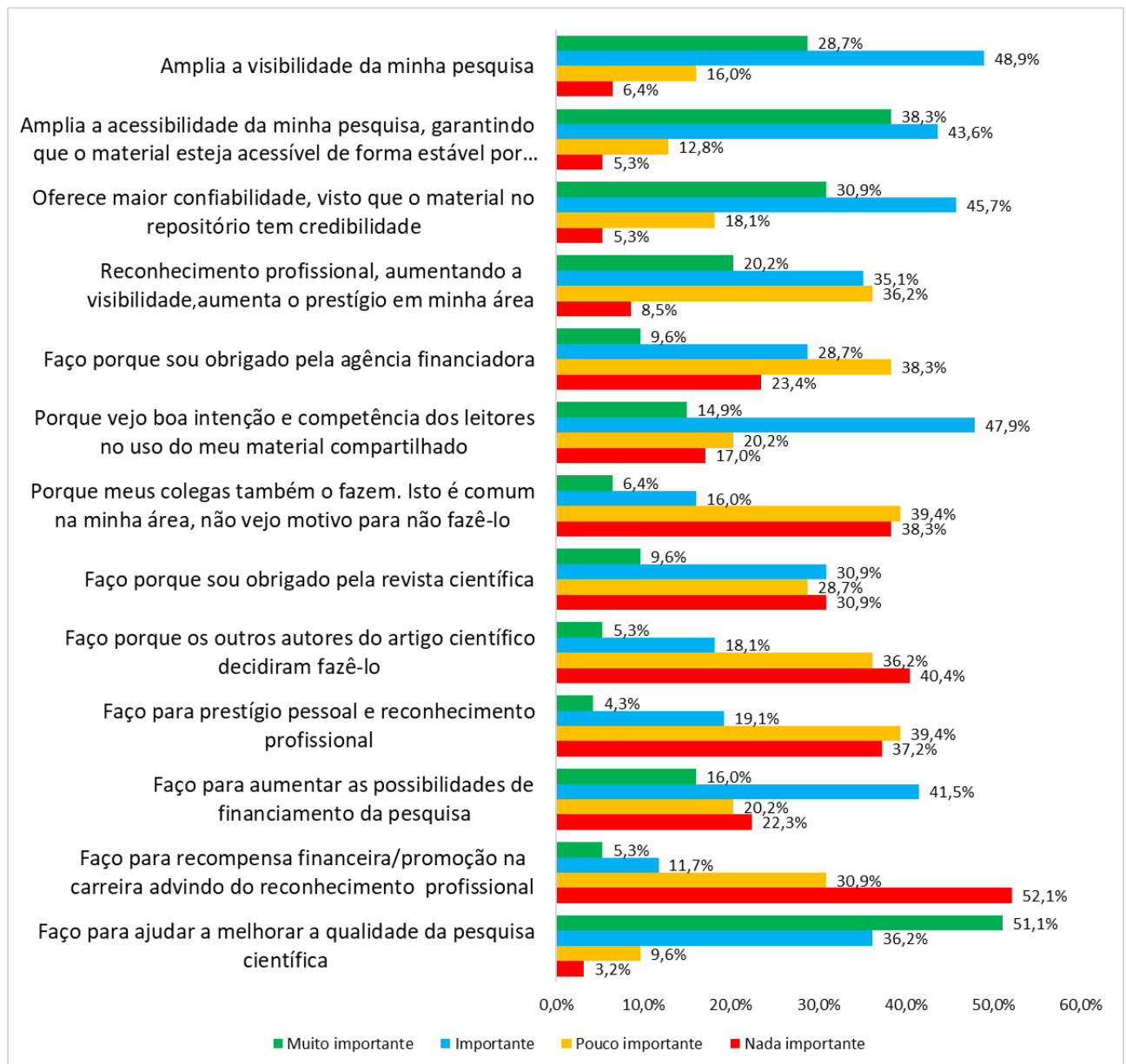
Grafico 1 – Atitudes e valores em relação à abertura de dados



A segunda análise relaciona-se com as motivações que influenciam ou influenciariam os pesquisadores na abertura dos seus dados de pesquisa. Verificou-se que 51% dizem que a principal motivação para a abertura de seus dados é ajudar a melhorar a qualidade da pesquisa científica; 38% afirmam que se sentem muito motivados para compartilhar e abrir seus dados para ampliar a acessibilidade da pesquisa, garantindo que o material esteja acessível de maneira estável por longo tempo;

30% se sentem incentivados a compartilhar seus dados por oferecer maior confiabilidade, visto que o material no repositório tem credibilidade; 28% concordam plenamente que a ampliação da visibilidade da pesquisa é uma motivação para o compartilhamento e abertura dos dados; e 20% asseguram que se sentem motivados a abrir seus dados pelo reconhecimento profissional, observando que o aumento da visibilidade aumenta o prestígio na área.

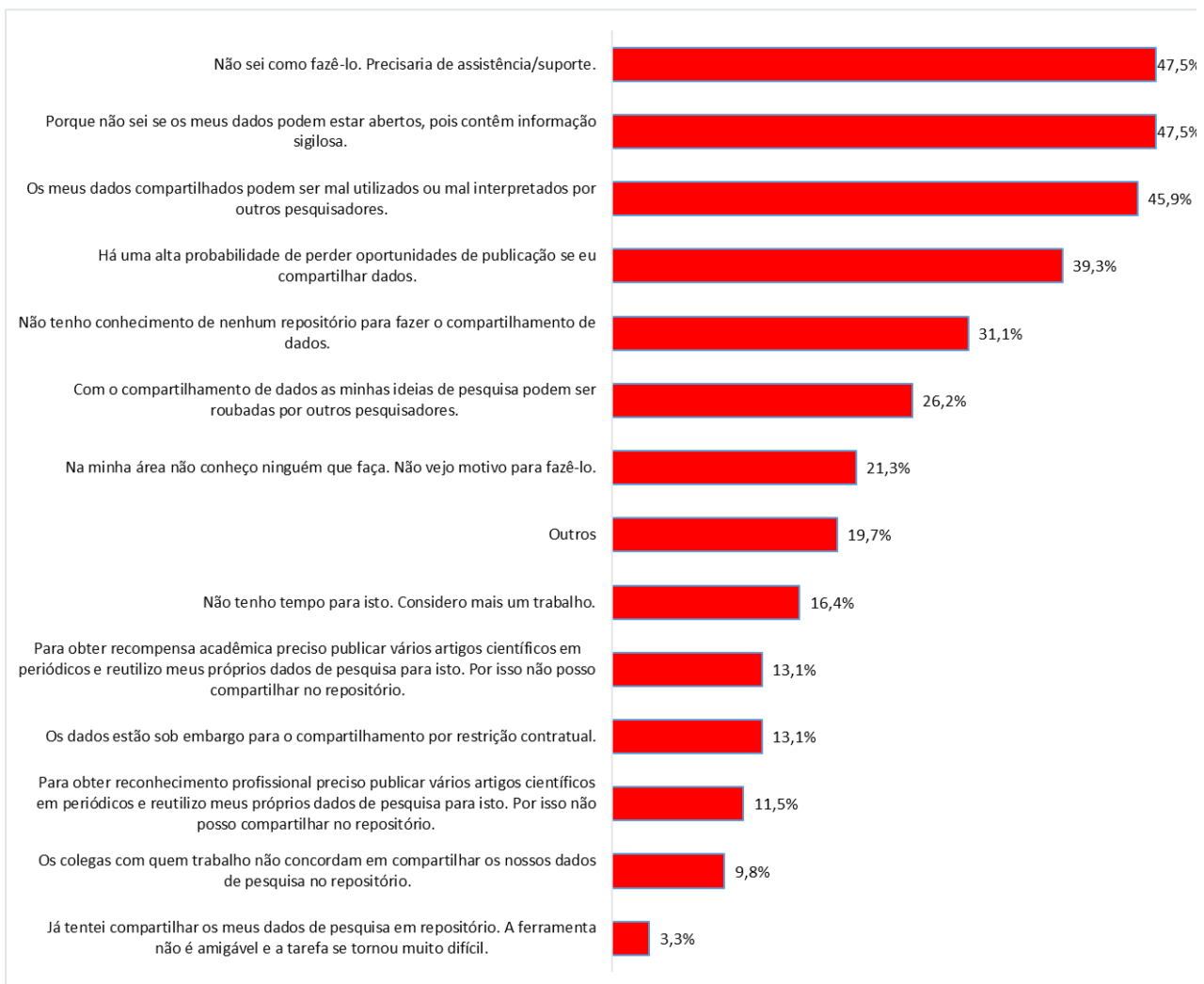
Gráfico 2 – Motivações para abertura dos dados de pesquisa



Para conhecer as principais barreiras percebidas pelos pesquisadores no compartilhamento e abertura de dados, solicitaram-se as principais razões para não compartilhar os dados de pesquisa em plataformas de dados. Do total de respondentes, 47% não compartilham por não saber como fazê-lo, e pela dúvida se os dados podem estar abertos, pois contêm informação sigilosa; 45% consideram que os dados compartilhados podem ser mal utilizados ou mal interpretados por outros pesquisadores; 39% acreditam ter alta probabilidade de perder

oportunidades de publicação se abrir os dados; 31% não conhecem de nenhum repositório para fazer o compartilhamento; 26% têm receio de que, ao fazer a abertura de dados, as ideias de pesquisa possam ser roubadas por outros pesquisadores; 21% observam que na sua área não conhecem ninguém que faça isto, portanto não veem motivo para fazê-lo; 9% declaram que os colegas com quem trabalham não concordam em compartilhar os dados de pesquisa no repositório.

Gráfico 3 – Principais razões para não compartilhar dados de pesquisa em plataformas de dados



Outra análise realizada foi quanto ao tipo de dados utilizados pelos pesquisadores: 52% usam dados pessoais e dados pessoais sensíveis; 47% indicaram usar dados sem sigilos (ostensivos); 47% usam dados potencialmente sigilosos, porém ainda não submetidos à análise de sua confidencialidade ou classificação; 43% usam dados com parcerias interinstitucionais; e 27% usam dados que geraram ou que podem gerar patentes.

Os pesquisadores também informaram qual o volume total do conjunto de dados em um projeto típico: 26% responderam que é maior que 1GB e menor que 10GB; 12% maior que 10GB e menor que 100GB; 12% menor que 1GB; 8% maior que 100GB e menor que 1TB; e 30% não sabem o volume.

Por último, verificou-se ainda que os colaboradores da pesquisa gostariam que a instituição fornecesse alguns serviços para garantir o acesso aos dados no futuro, como serviço de armazenamento e preservação (73%) e sistema de backup de arquivos (72%), apoio técnico na elaboração do plano de gestão de dados (67%), serviço de compartilhamento e acesso de dados (52%), formação e consultoria na gestão de dados (50%), informação sobre repositório temático (44%) e identificação de normas e esquemas disciplinares de descrição de dados (36%).

Verifica-se que os pesquisadores estão interessados em serviços institucionais que apoiem as boas práticas na gestão dos dados em todo o ciclo de vida da pesquisa.

Gráfico 4 – Proteção de dados

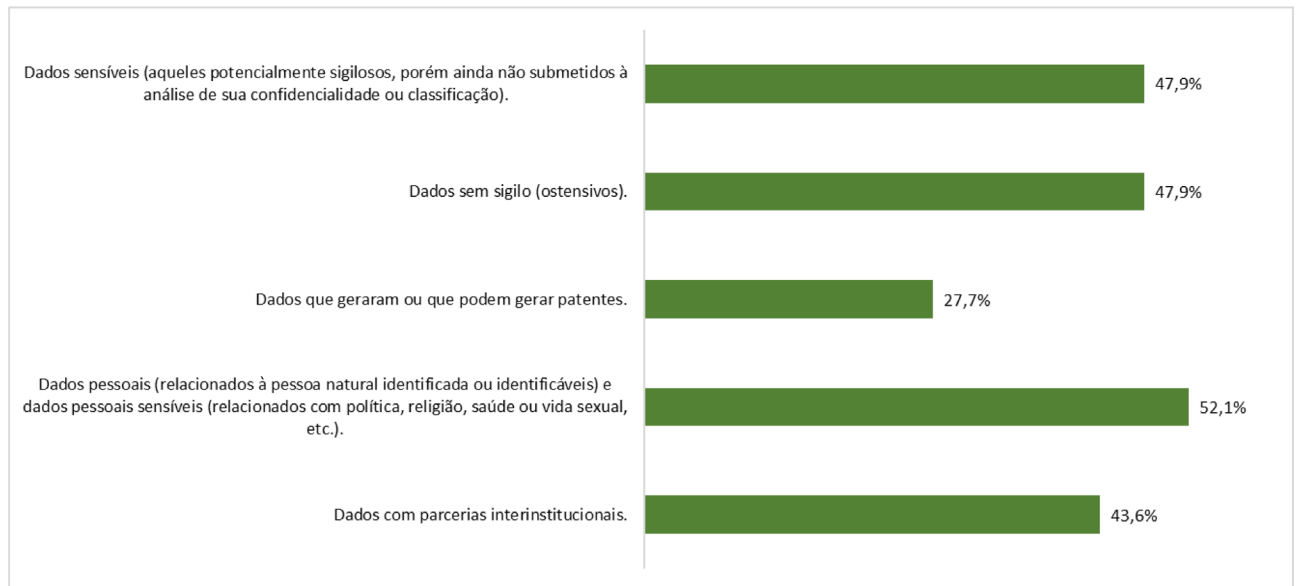
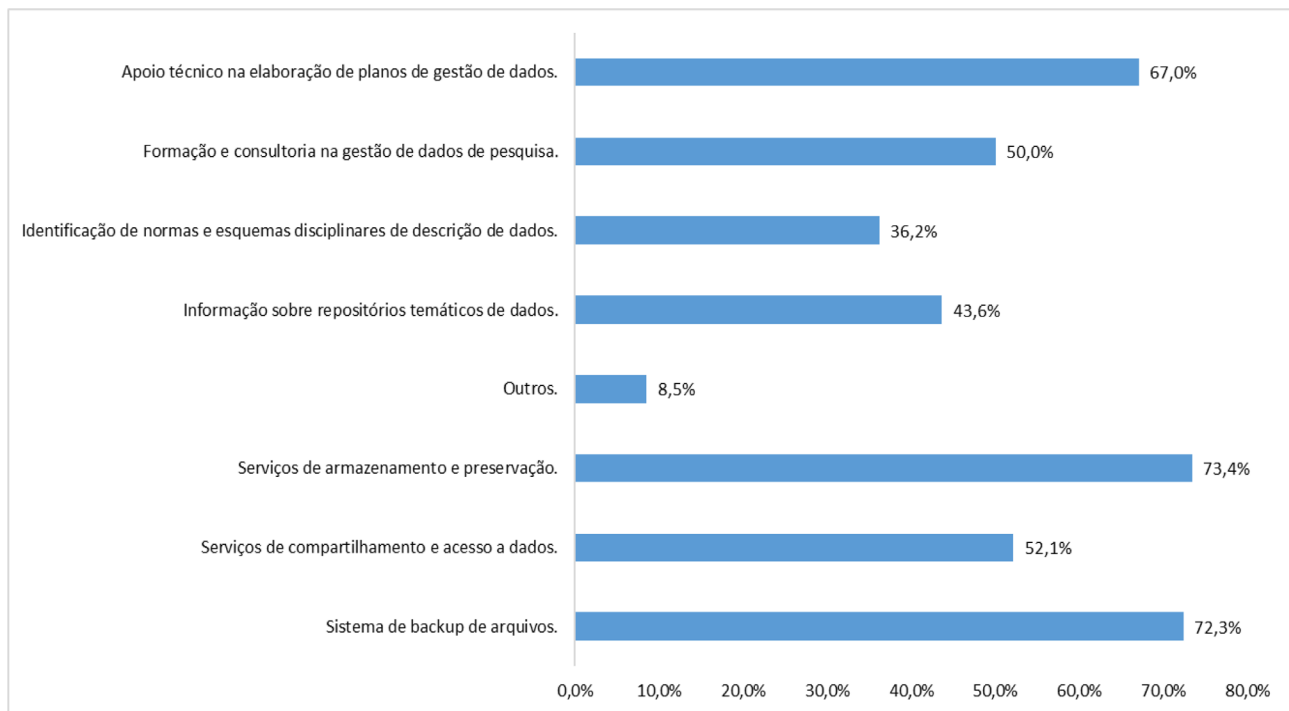


Gráfico 5 – Necessidades dos pesquisadores



## CONCLUSÕES

Os resultados da pesquisa aqui expostos mostram os principais anseios do pesquisador. Nesse contexto, o plano de gestão de dados, considerado uma tendência de forma global, é um instrumento fundamental para a gestão de dados de pesquisa e tem se tornado um requisito para o financiamento em várias agências de fomento, como H2020, WellCome Trust, Bill e Melinda Gates, NIH, e no Brasil, mais recentemente, se tornou requisito em alguns editais da Fundação de Amparo à Pesquisa de São Paulo (Fapesp).

Apesar da existência desses requisitos, os resultados da pesquisa demonstraram que até o momento os pesquisadores da instituição têm pouca familiaridade com o instrumento. Apenas 20% dos respondentes já fizeram algum tipo de documento com algum elemento característico de um plano de gestão de dados, como por exemplo, política de preservação de dados, de gestão de registros, de compartilhamento e/ou de estratégia de eliminação de dados.

Conclui-se com os achados da pesquisa, até o momento, que a maioria dos pesquisadores da Fiocruz compreende a importância do compartilhamento e abertura de dados de pesquisa, principalmente para a transparência na conduta da pesquisa e otimização do avanço científico. Além disto, a percepção dos benefícios de compartilhar achados como meio de ampliar a acessibilidade, confiabilidade, visibilidade e o reconhecimento profissional também mostra que a Fiocruz possui um ambiente propício para se trabalhar a abertura de dados de pesquisa na cauda longa da ciência (*small science*).

Além disso, percebe-se que a desinformação é a principal barreira para o compartilhamento e abertura de dados na Fiocruz, talvez por desconhecimento da legislação vigente sobre a proteção de dados pessoais, ou seja, por não conhecer as plataformas de dados disponíveis para o compartilhamento.



A desinformação sobre as formas e normas de citação de dados e do uso dos identificadores persistentes (que facilitam a citação dos dados) também afasta o pesquisador do compartilhamento de dados. A má utilização ou interpretação dos dados também preocupa o pesquisador, mas ele desconhece a documentação que deve acompanhar os dados compartilhados, minimizando a utilização ou interpretação inadequada por outros pesquisadores.

Considerando esses anseios do pesquisador, a Fiocruz está elaborando uma proposta de Plano de Gestão de Dados coordenado pelo Instituto de Comunicação Científica e Tecnológica em Saúde (ICICT), em conformidade com o que os pesquisadores sinalizaram como necessidades descritas nos resultados dessa pesquisa.

O desconhecimento sobre o Plano de Gestão de Dados (PGD) é um fator importante que necessita ser considerado. É fundamental o apoio aos pesquisadores para compreender a relevância do PGD e as ferramentas disponíveis para elaborá-lo. Além desse apoio, é necessária também a disponibilização, clara, de infraestrutura que apoie a gestão de todo o ciclo de vida da pesquisa, bem como serviços que auxiliem os pesquisadores no uso das ferramentas disponíveis para o compartilhamento e abertura de dados de pesquisa na Fiocruz e no campo da saúde.

---

## REFERÊNCIAS

FUNDAÇÃO OSWALDO CRUZ. *Abertura de dados para pesquisa na Fiocruz: perspectivas de um novo paradigma da Ciência*. Rio de Janeiro, 18 jun. 2018. Disponível em: <https://portal.fiocruz.br/noticia/evento-marca-inicio-da-discussao-sobre-abertura-de-dados-para-pesquisa-na-fiocruz>. Acesso em: 28 jun. 2018.

HALL, H. Social exchange for knowledge exchange. In: *MANAGING KNOWLEDGE: CONVERSATIONS AND CRITIQUES*, 2001. *Presentation* [...]. Leicester: University of Leicester Management Centre, 2001.

MINISTÉRIO DA SAÚDE. Fundação Oswaldo Cruz. *Política de Acesso Aberto ao Conhecimento*. [Rio de Janeiro]: Fiocruz, 2014. Disponível em: [https://portal.fiocruz.br/sites/portal.fiocruz.br/files/documentos/portaria\\_-\\_politica\\_de\\_acesso\\_aberto\\_ao\\_conhecimento\\_na\\_fiocruz.pdf](https://portal.fiocruz.br/sites/portal.fiocruz.br/files/documentos/portaria_-_politica_de_acesso_aberto_ao_conhecimento_na_fiocruz.pdf). Acesso em: 02 set. 2019.

U.S. NATIONAL COMMITTEE FOR CODATA. Committee for a pilot study on database interfaces. *Bits of power: issues in Global Access to Scientific Data*. Washington, D.C.: National Academy Press, 1997. Disponível em: <http://site.ebrary.com/id/10041138>. Acesso em: 20 mar. 2017.

VEIGA, V. S. de O. *Percepção dos pesquisadores portugueses e brasileiros da área de Neurociências quanto ao compartilhamento de artigos científicos e dados de pesquisa no acesso aberto verde: custos, benefícios e fatores contextuais*. 2017. 294 f. Tese (Doutorado em Informação e Comunicação em Saúde) - Instituto de Comunicação e Informação Científica e Tecnológica em Saúde, Fundação Oswaldo Cruz, Rio de Janeiro, 2017.