26º ENANCIB
Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação
A Amazônia na sociedade global em rede: assimetrias e potencialidades informacionais, tecnológicas e ambientais
Belém, Pará, 09 a 13 de novembro de 2026
Universidade Federal do Pará

ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO
ISSN 2177-3688

GT 8 – Dados, Informação e Tecnologia

PROPOSTA DE *WORKFLOW* CIENTÍFICO PARA A VIGILÂNCIA DA RESISTÊNCIA ANTIMICROBIANA

PROPOSAL FOR A SCIENTIFIC WORKFLOW FOR ANTIMICROBIAL RESISTANCE SURVEILLANCE

Fábio Bernardo da Silva – Fundação Osvaldo Cruz (FIOCRUZ)
Anderson Silva de Araujo – Fundação Oswaldo Cruz (FIOCRUZ)
Viviane Santos de Oliveira Veiga – Fundação Osvaldo Cruz (FIOCRUZ)
Patrícia Corrêa Henning – Universidade Federal do Estado do Rio de Janeiro (UNIRIO)

Modalidade: Trabalho Completo

Resumo: A crescente produção de dados ômicos, especialmente no contexto da resistência antimicrobiana, impõe desafios relacionados à sua organização, documentação, interoperabilidade, preservação e reúso. Embora os Princípios FAIR orientem a localização, o acesso, a interoperabilidade e a reutilização de dados científicos, sua aplicação requer escolhas concretas nem sempre adotadas. O estudo tem como objetivo propor um *workflow* científico para a gestão e a curadoria de dados ômicos relacionados à resistência antimicrobiana, fundamentado nos Princípios FAIR e operacionalizado por meio de um Perfil de Implementação FAIR. Trata-se de uma pesquisa qualitativa, aplicada, exploratória, descritiva e analítica, desenvolvida a partir de levantamento bibliográfico e documental, análise de *workflows* existentes e identificação de lacunas relativas à gestão de dados ômicos. Como resultado, propõe-se um *workflow* estruturado em sete etapas: planejamento; coleta; processamento e análise; curadoria e representação; armazenamento e preservação; compartilhamento e acesso; e reúso e geração de conhecimento. O Perfil de implementação FAIR atua transversalmente ao modelo, registrando recursos e decisões de implementação FAIR ao longo do ciclo de vida dos dados. Conclui-se que a integração entre FAIR, Perfil de implementação FAIR e *workflow* científico pode fortalecer a governança, a rastreabilidade, a interoperabilidade e o reúso responsável de dados ômicos aplicados à vigilância da resistência antimicrobiana. A proposta possui caráter conceitual e preliminar e deverá ser validada empiricamente em contextos reais de pesquisa e vigilância.

Palavras-chave: dados ômicos; resistência antimicrobiana; Princípios FAIR; Perfil de Implementação FAIR; *workflow* científico.

Abstract: The growing production of omics data, particularly in the context of antimicrobial resistance, poses challenges related to their organization, documentation, interoperability, preservation, and reuse. Although the FAIR Principles guide the findability, accessibility, interoperability, and reusability of scientific data, their implementation requires concrete choices that are not always adopted. This study aims to propose a scientific workflow for the management and curation of omics data related to antimicrobial resistance, grounded in the FAIR Principles and operationalized through a FAIR Implementation Profile. This qualitative, applied, exploratory, descriptive, and analytical study is based on a bibliographic and documentary review, an analysis of existing workflows, and the identification of gaps in omics data management. As a result, a workflow structured into seven stages is proposed: planning; data collection; processing and analysis; curation and representation; storage and preservation; sharing and access; and reuse and knowledge generation. The FAIR Implementation Profile acts transversally across the model by documenting FAIR-enabling resources and implementation decisions throughout the data life cycle. It is concluded that integrating FAIR, FAIR Implementation Profile, and scientific workflows can strengthen the governance, traceability, interoperability, and responsible reuse of omics data applied to antimicrobial resistance surveillance. The proposal is conceptual and preliminary and should be empirically validated in real-world research and surveillance settings.
Keywords: omics data; antimicrobial resistance; FAIR Principles; FAIR Implementation Profile; scientific workflow.

1 INTRODUÇÃO

A Ciência Aberta consolidou-se a partir do final da década de 1990, como um movimento orientado à ampliação da transparência, da colaboração e do acesso aos resultados da pesquisa científica. Esse movimento abrange práticas como o acesso aberto às publicações, o compartilhamento e o reúso de dados de pesquisa, a disponibilização de protocolos, códigos e parâmetros analíticos, bem como formas mais abertas de avaliação e colaboração científica. Tais práticas têm repercussões diretas sobre a reprodutibilidade, a confiabilidade e o impacto social do conhecimento científico (Fecher; Friesike, 2014; UNESCO, 2021).

No entanto, a simples disponibilização e abertura dos dados não é suficiente para garantir seu reúso efetivo. É necessário que os dados sejam devidamente descritos, localizáveis, acessíveis sob condições claras, interoperáveis com outros sistemas e passíveis de reúso. Com essa finalidade, foram propostos os Princípios FAIR, acrônimo para *Findable*, *Accessible, Interoperable* e *Reusable*, isto é, localizável, acessível, interoperável e reutilizável (Wilkinson *et al.*, 2016).

A necessidade de práticas FAIR torna-se ainda mais evidente no campo das ciências ômicas. Tecnologias de alto rendimento, como o sequenciamento de nova geração, em inglês *Next-Generation Sequencing* (NGS), permitem a produção rápida e massiva de dados genômicos, transcriptômicos, proteômicos e metabolômicos. Embora essas tecnologias ampliem significativamente a capacidade de compreensão dos sistemas biológicos, também produzem conjuntos de dados complexos, heterogêneos e fortemente dependentes de metadados contextuais para sua interpretação, reprodução e reutilização (Hasin; Saria; Mason, 2017; Slatko; Gardner; Ausubel, 2018).

Os dados genômicos possuem relevância para a investigação de doenças, a identificação de variantes genéticas e o desenvolvimento de novas abordagens terapêuticas (Collins; Venter, 2001; Manolio *et al.*, 2013). No campo da resistência antimicrobiana (RAM), esses dados são estratégicos para identificar genes e mecanismos de resistência, monitorar a disseminação de clones bacterianos e apoiar a vigilância de patógenos emergentes. Essa importância é ampliada pela abordagem *One Health*, que reconhece a interdependência entre saúde humana, saúde animal e ambiente na compreensão e no enfrentamento da RAM (Laxminarayan *et al.*, 2013; WHO, 2019).

Diversas infraestruturas internacionais, como o *National Center for Biotechnology Information* (NCBI), o *European Nucleotide Archive* (ENA) e o *Comprehensive Antibiotic Resistance Database* (CARD), contribuem para ampliar a disponibilidade de dados genômicos e informações sobre resistência antimicrobiana (Alcock *et al.*, 2023; Amid *et al.*, 2020; Hayes *et al.*, 2024). Entretanto, a diversidade de formatos, padrões de metadados, vocabulários controlados, políticas de acesso e práticas de documentação ainda limita a integração e o reúso efetivo de dados produzidos por diferentes instituições, países e domínios de pesquisa.

Nesse cenário, os Princípios FAIR oferecem diretrizes gerais para a gestão de dados, mas não definem, por si sós, quais padrões, identificadores, ontologias, repositórios ou tecnologias devem ser adotados em cada comunidade científica. Essa lacuna pode ser enfrentada por meio do Perfil de Implementação FAIR, em inglês *FAIR Implementation Profile* (FIP), instrumento voltado à documentação estruturada das escolhas de implementação FAIR de uma comunidade de prática. O FIP registra os Recursos de Habilitação FAIR, tais como identificadores persistentes, esquemas de metadados, ontologias, vocabulários controlados, licenças, protocolos de acesso e repositórios adotados por uma determinada comunidade (Schultes *et al.*, 2020).

Por outro lado, a gestão de dados ômicos requer a organização explícita das etapas de planejamento, coleta, processamento, análise, curadoria, preservação, compartilhamento e reúso. O *workflow* científico permite estruturar essas etapas, documentar decisões analíticas e técnicas e favorecer a rastreabilidade, a transparência e a reprodutibilidade dos processos de pesquisa.

Diante desse contexto, assume-se que a articulação entre dados ômicos, Princípios FAIR, Perfil de Implementação FAIR (FIP) e *workflow* científico pode contribuir para a padronização, a interoperabilidade, a governança e o reúso responsável dos dados, especialmente em contextos de saúde pública e vigilância genômica.

Este estudo tem como objetivo geral propor um *workflow* científico voltado à gestão e à curadoria de dados ômicos relacionados à resistência antimicrobiana, fundamentado nos Princípios FAIR e articulado às decisões de implementação registradas em um FIP, como estratégia para favorecer a interoperabilidade, a governança e o reúso responsável desses dados no âmbito da saúde pública. A partir desse objetivo geral, definem-se os seguintes objetivos específicos:

a) Mapear *workflows* científicos, padrões de metadados, vocabulários controlados, ontologias, repositórios, bases de dados e tecnologias de implementação FAIR utilizados em pesquisas com dados ômicos, considerando sua aplicabilidade ao domínio da resistência antimicrobiana;
b) Analisar *workflows* científicos baseados em dados ômicos aplicados à resistência antimicrobiana à luz dos Princípios FAIR, identificando lacunas, riscos, requisitos e oportunidades de aprimoramento relacionados à gestão, à curadoria, à interoperabilidade e ao reúso dos dados;
c) Propor, estruturar e descrever um workflow científico para a gestão e a curadoria de dados ômicos relacionados à resistência antimicrobiana, articulando as etapas do ciclo de vida dos dados aos recursos e às decisões de implementação FAIR documentados no FIP.

A Justificativa de elaboração desse workflow fundamenta-se na necessidade de integrar, em uma estrutura explícita e sistematizada, as diferentes etapas envolvidas na gestão e na curadoria de dados ômicos aos Princípios FAIR e às escolhas de implementação registradas no FIP do domínio. Embora a literatura apresente diferentes workflows voltados

ao processamento e à análise de dados ômicos, persistem desafios relacionados à sua gestão e curadoria, especialmente no que se refere à heterogeneidade de formatos, padrões de metadados, vocabulários controlados, mecanismos de identificação, políticas de acesso, práticas de documentação e condições de armazenamento, preservação e compartilhamento. Esses aspectos podem limitar a interoperabilidade e o reúso efetivo e responsável dos dados.

Sob essa ótica, a proposta não decorre da ausência de *workflows* científicos ou bioinformáticos, mas da necessidade de estabelecer uma articulação explícita entre as diferentes etapas do ciclo de vida dos dados: planejamento; coleta; processamento e análise; curadoria e representação; armazenamento e preservação; compartilhamento e acesso; reúso e geração de conhecimento, além das decisões necessárias à implementação dos Princípios FAIR.

Nessa arquitetura, o FIP assume uma função transversal, ao permitir que as escolhas de implementação FAIR sejam explicitadas, documentadas e relacionadas às diferentes etapas do *workflow*. Essas escolhas podem envolver, entre outros elementos, padrões de metadados, vocabulários controlados, ontologias, identificadores persistentes, formatos de dados, repositórios, protocolos e demais recursos técnicos e semânticos necessários para tornar os dados progressivamente mais localizáveis, acessíveis, interoperáveis e reutilizáveis.

A contribuição específica, portanto, não objetiva substituir *workflows* bioinformáticos existentes, mas em acrescentar uma camada integrada de gestão e curadoria capaz de relacionar cada etapa do ciclo de vida dos dados às respectivas decisões de implementação dos principios FAIR registradas no FIP. Essa articulação permite estabelecer maior rastreabilidade entre requisitos, recursos e decisões de implementação, tornando mais explícitos os componentes técnicos, semânticos, informacionais, organizacionais e institucionais necessários à interoperabilidade, à governança e ao reúso responsável dos dados ômicos no domínio da resistência antimicrobiana.

Além desta introdução, o artigo apresenta o referencial teórico que fundamenta a relação entre dados ômicos, Princípios FAIR, FIP e *workflows* científicos, seguido do procedimento metodológico adotado para o levantamento bibliográfico e documental. Posteriormente, são analisados os *workflows* científicos identificados e discutidas as principais lacunas relacionadas à gestão e à curadoria de dados ômicos. Em seguida, apresenta-se o *workflow* proposto, estruturado em sete etapas do ciclo de gestão e curadoria dos dados, bem como a atuação transversal do FIP na explicitação e no registro dos recursos e das decisões de

implementação dos principios FAIR. Por fim, são apresentados os resultados e discurssões, seguido das considerações finais decorrentes da integração entre Princípios FAIR, FIP e *workflow* científico no contexto dos dados ômicos relacionados à RAM no âmbito da saúde pública.

2 REFERENCIAL TEÓRICO

A resistência antimicrobiana constitui uma das principais ameaças contemporâneas à saúde pública mundial. Embora a resistência seja um fenômeno evolutivo natural, o uso inadequado e excessivo de antimicrobianos na saúde humana, animal e na produção de alimentos favorece a seleção e a disseminação de microrganismos resistentes (Guimarães; Momesso; Pupo, 2010; WHO, 2019).

Os impactos da RAM ultrapassam a falha terapêutica individual e alcançam os sistemas de saúde, as cadeias produtivas, a segurança alimentar e a economia. Por essa razão, seu enfrentamento exige estratégias integradas de prevenção, monitoramento, vigilância e controle, em conformidade com a abordagem *One Health* (Laxminarayan *et al.*, 2013; O'Neill, 2016).

Nesse contexto, os dados ômicos desempenham papel relevante ao permitirem a caracterização de diferentes níveis moleculares dos sistemas biológicos, incluindo genomas, transcriptomas, proteomas, metabolomas e epigenomas. Em particular, os dados genômicos permitem identificar genes de resistência, investigar relações filogenéticas entre microrganismos, acompanhar surtos e analisar padrões de disseminação de patógenos em diferentes populações, territórios e ambientes.

Entretanto, esses dados apresentam elevado volume, complexidade estrutural, heterogeneidade de formatos e dependência de informações contextuais. Para que possam ser interpretados e reutilizados adequadamente, devem estar associados a metadados sobre origem das amostras, condições experimentais, instrumentos empregados, parâmetros de processamento, versões de *software*, métodos analíticos, critérios de qualidade e aspectos éticos ou legais relacionados ao seu acesso e uso (Hasin; Saria; Mason, 2017; Lefever *et al.*, 2009).

Os Princípios FAIR foram propostos para aprimorar a capacidade de localização, acesso, integração e reutilização de dados científicos por pessoas e máquinas (Wilkinson *et al.*, 2016). Diferentemente de abordagens centradas exclusivamente na abertura dos dados,

FAIR não significa necessariamente que todos os dados devem ser irrestritos. Dados podem ser FAIR mesmo quando submetidos a controles de acesso, desde que seus metadados sejam localizáveis, suas condições de acesso sejam explícitas e seus procedimentos de solicitação estejam documentados.

No domínio ômico, a adoção dos Princípios FAIR é especialmente importante devido à diversidade de plataformas analíticas, formatos de arquivo, escalas de dados e contextos institucionais envolvidos. A aplicação desses princípios pode contribuir para que conjuntos de dados gerados em diferentes estudos sejam comparados, combinados e reutilizados de maneira responsável, fortalecendo a pesquisa colaborativa e a vigilância em saúde.

Embora os Princípios FAIR indiquem o que deve ser buscado na gestão de dados, sua implementação depende de escolhas concretas feitas por comunidades científicas e institucionais. O Perfil de Implementação FAIR (FIP) foi proposto como um instrumento para documentar essas escolhas de maneira estruturada (Schultes *et al.*, 2020).

O FIP descreve quais recursos uma comunidade adota para implementar os subprincípios FAIR. Esses recursos podem incluir identificadores persistentes, padrões de metadados, vocabulários controlados, ontologias, formatos de dados, protocolos de acesso, licenças, repositórios e ferramentas de avaliação. Dessa forma, o FIP explicita o "como" da implementação FAIR.

No contexto dos dados ômicos aplicados à RAM, um FIP pode apoiar a seleção e a documentação de recursos como DOI, ORCID*, BioProject, BioSample, MIxS, DataCite Metadata Schema, NCBI Taxonomy, Sequence Ontology*, repositórios institucionais, entre outros. Sua adoção pode favorecer o alinhamento entre equipes de laboratório, bioinformatas, curadores de dados, pesquisadores, gestores e instituições responsáveis pela vigilância genômica.

O workflow científico pode ser entendido como a organização estruturada de etapas, atividades, ferramentas, dados de entrada, produtos intermediários e resultados de um processo de pesquisa. Em estudos ômicos, os *workflows* geralmente envolvem planejamento experimental, geração de dados, controle de qualidade, processamento bioinformático, análise, interpretação, documentação, preservação e compartilhamento.

A formalização desses fluxos contribui para a transparência metodológica, a reprodutibilidade dos resultados e a rastreabilidade das decisões tomadas ao longo da pesquisa. Quando articulados aos Princípios FAIR, os *workflows* deixam de ser apenas

sequências técnicas de processamento e passam a incorporar requisitos de gestão, curadoria, metadados, proveniência, licenciamento, preservação e acesso aos dados.

Assim, a integração entre *workflow* científico e FIP permite associar cada etapa do ciclo de vida dos dados às escolhas técnicas e semânticas necessárias para sua implementação FAIR. Essa integração constitui a base conceitual da proposta desenvolvida neste estudo.

3 PROCEDIMENTOS METODOLÓGICOS

A pesquisa adota abordagem qualitativa, de natureza aplicada, com procedimentos exploratórios, descritivos e analíticos. Seu foco consiste em propor um *workflow* científico para a gestão e a curadoria de dados ômicos relacionados à resistência antimicrobiana.

O estudo foi estruturado em três etapas complementares, organizadas de modo a relacionar o mapeamento dos recursos, a análise comparativa dos workflows existentes e a elaboração da proposta.

Na primeira etapa, realizou-se levantamento bibliográfico e documental sobre dados ômicos, resistência antimicrobiana, Princípios FAIR, Perfil de Implementação FAIR (FIP) e workflows científicos. As buscas principais foram realizadas entre 1º e 30 de junho de 2025, abrangendo publicações do período de 2014 a 2024, e recuperaram 2.765 registros: PubMed (327), Scopus (1.536), Web of Science (797), LILACS (5) e BRAPCI (100). Além disso, uma busca exploratória realizada entre 3 e 10 de março de 2025 reuniu 208 documentos. Foram previstos como critérios de exclusão: registros duplicados, documentos sem texto completo, textos opinativos ou editoriais e estudos fora do escopo. As informações foram organizadas em matriz de extração contemplando tipo de recurso, finalidade, domínio ômico, padrões de metadados, vocabulários e ontologias, identificadores, repositórios e bases de dados, protocolos de acesso, licenças e relação com os subprincípios FAIR.

Na segunda etapa, foram selecionados e analisados quatro *workflows* (*Galaxy*, *Nextflow*, *nf-core* e *WorkflowHub*), com base em critérios de representatividade, grau de adoção comunitária e complementaridade funcional no ecossistema de dados ômicos e bioinformática. A amostra contemplou plataformas de execução web e linguagens de orquestração analítica (*Galaxy* e *Nextflow*), iniciativas de padronização de pipelines (*nf-core*) e repositórios de registro e descoberta FAIR (*WorkflowHub*), excluindo-se ferramentas proprietárias ou de uso estritamente local sem impacto no compartilhamento comunitário. A análise comparativa investigou escopo, padronização de metadados, identificadores

persistentes, proveniência, licenciamento, alinhamento aos Princípios FAIR e lacunas na gestão contínua do ciclo de vida dos dados.

Na terceira etapa, as lacunas identificadas foram sintetizadas e associadas às sete etapas do ciclo de gestão e curadoria dos dados. Para cada etapa, foram relacionados requisitos de gestão, curadoria, documentação, interoperabilidade, acesso, preservação e reúso, bem como os tipos de recursos passíveis de registro em um FIP. A inclusão dos componentes no *workflow* foi orientada pela correspondência entre as lacunas observadas e os recursos de implementação FAIR necessários para enfrentá-las.

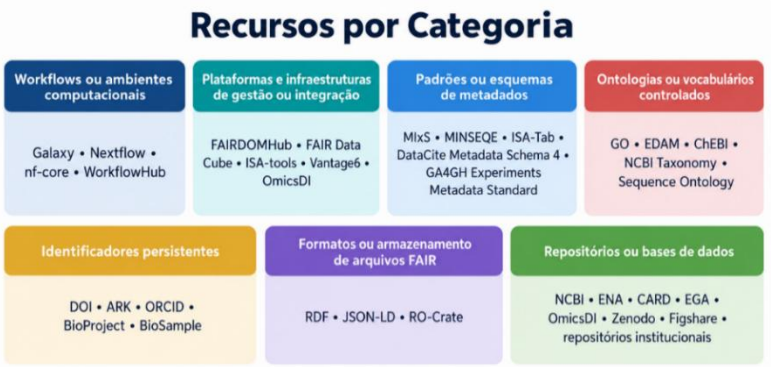
A proposta resultante possui caráter conceitual e preliminar. Nesta etapa do estudo, não foi realizada validação empírica do *workflow* com pesquisadores, laboratórios ou serviços de vigilância. Essa validação é prevista como etapa subsequente, numa oficina de cocriação com todos os *stakeholders*, a fim de avaliar aplicabilidade, usabilidade e capacidade de adaptação do modelo a diferentes contextos institucionais e aplicações na área de dados ômicos.

Declaração de uso de inteligência artificial generativa: Utilizou-se o ChatGPT, da OpenAI, para revisão ortográfica e geração da representação visual apresentada nas Figuras 1 e 2. O conteúdo, as categorias, os recursos e os fluxos representados foram definidos pelos autores. A imagem gerada foi revisada, corrigida e validada pelos autores, que assumem integral responsabilidade por sua precisão e adequação.

4 RESULTADOS E DISCUSSÃO

O levantamento bibliográfico e documental permitiu selecionar, até esta versão, 34 recursos distribuídos em sete categorias, conforme a figura 01, a seguir.

Figura 1 - Categorização dos recursos selecionados



Fonte: elaboração própria (2026), com uso do ChatGPT/OpenAI para geração da imagem, a partir de conteúdo e instruções fornecidos pelos autores. Imagem revisada e validada pelos autores.

A seleção desses recursos foi orientada pela sua capacidade de estruturar processos computacionais, registrar a proveniência, assegurar a representação semântica e favorecer a reprodutibilidade, o compartilhamento e a preservação sustentável dos dados ao longo das sete etapas do ciclo de vida, viabilizando seu registro transversal em um Perfil de Implementação FAIR (FIP).

Os resultados também evidenciam que a implementação FAIR em dados ômicos depende da articulação entre decisões técnicas, informacionais e institucionais. Nesse sentido, a elaboração de um FIP para dados ômicos mostra-se adequada como instrumento para registrar e tornar explícitas as escolhas de implementação FAIR adotadas ao longo do ciclo de vida dos dados. Ao documentar os Recursos de Habilitação FAIR, o FIP contribui para reduzir ambiguidades, orientar equipes multidisciplinares e favorecer a consistência entre diferentes estudos, laboratórios e instituições.

Quadro 1 - Comparação funcional dos workflows

Recurso	**Função principal**	**Etapas apoiadas**	**Contribuição FAIR**	**Lacuna que justifica a proposta integrada**
Galaxy	Ambiente web para análises reprodutíveis	Processamento /análise; compartilhamento	Histórico, ferramentas, parâmetros e *workflows*	Não resolve sozinho PGD, preservação institucional, licenças e escolhas FIP.
Nextflow	Linguagem e mecanismo de execução de pipelines	Processamento /análise	Portabilidade, versionamento do *pipeline* e execução em diferentes infraestruturas	Depende de práticas externas para metadados, repositório, identificador persistente, licenciamento e curadoria.
nf-core	Coleção comunitária de *pipelines Nextflow* padronizados	Processamento /análise e controle de qualidade	Padronização, revisão comunitária, documentação e reprodutibilidade	Foco predominante no *pipeline*; gestão integral do ciclo de vida precisa ser complementada.
WorkflowHub	Registro e descoberta de *workflows* científicos	Compartilhamento, acesso e reuso	Publicação, descrição e descoberta de *workflows*	Não substitui execução, PGD, curadoria dos dados nem definição institucional de acesso.

Fonte: elaboração própria (2026).

A partir dos resultados já identificados no estudo, as principais lacunas podem ser agrupadas em cinco dimensões: (i) qualidade e completude dos metadados; (ii) adoção inconsistente de vocabulários controlados e ontologias; (iii) documentação insuficiente da proveniência; (iv) indefinição ou heterogeneidade das licenças; e (v) condições de acesso e reúso nem sempre explicitadas. Essas lacunas orientaram a estruturação do *workflow* proposto: as questões de metadados, vocabulários e ontologias foram associados sobretudo à etapa de curadoria e representação; a proveniência foi articulada às etapas de processamento, análise e curadoria; e as licenças e condições de acesso e reúso foram incorporadas às etapas de compartilhamento, acesso e geração de conhecimento. O FIP atua transversalmente, registrando as escolhas de implementação FAIR correspondentes.

Quadro 2 - Critérios de análise derivados de FAIR, FIP e ciclo de vida dos dados

Dimensão	**Critério verificável**	**Exemplos de evidência**	**Relação FAIR[1]**	**Etapa do workflow**
Metadados	Completude, consistência e uso de padrões	MIxS, MINSEQE, ISA-Tab, *DataCite*	F2, I1, I3, R1	Curadoria e representação
Semântica	Uso de ontologias e vocabulários controlados	GO, EDAM, ChEBI, NCBI *Taxonomy, Sequence Ontology*	I1, I2, I3	Curadoria e representação
Identificação	Uso de PID para dados, pessoas, projetos e amostras	DOI, ARK, ORCID, *BioProject, BioSample*	F1, F3	Planejamento; coleta; armazenamento
Proveniência e reprodutibilidade	Registro de versões, parâmetros, software e histórico	*Galaxy, Nextflow, nf-core, RO-Crate*	R1.2, R1.3	Processamento e análise
Acesso e licenciamento	Condições claras, protocolos e licenças	EGA, protocolos autenticados, licenças explícitas	A1, A1.1, A1.2, R1.1	Compartilhamento e acesso
Preservação e depósito	Repositório, formatos sustentáveis e política de preservação	NCBI, ENA, *Zenodo*, *Figshare,* repositórios institucionais	A2, R1.3	Armazenamento e preservação

Fonte: elaboração própria (2026).

A proposta de *workflow* apresentada na Figura 1 foi estruturado em sete etapas: planejamento; coleta; processamento e análise; curadoria e representação; armazenamento e preservação; compartilhamento e acesso; e reúso e geração de conhecimento. Em todas essas etapas, os Princípios FAIR orientam a gestão dos dados, enquanto o FIP atua como mecanismo de documentação, padronização e rastreabilidade das escolhas realizadas.

Figura 2 - Proposta de *workflow* científico para dados ômicos em resistência antimicrobiana



Fonte: elaboração própria (2026), com uso do ChatGPT/OpenAI para geração da imagem, a partir de conteúdo e instruções fornecidos pelos autores. Imagem revisada e validada pelos autores.

O FIP é apresentado como elemento transversal ao workflow, pois documenta as decisões relativas a padrões, identificadores, metadados, ontologias, vocabulários controlados, formatos, repositórios, protocolos de acesso e licenças. Dessa forma, a figura não deve ser entendida apenas como um fluxo técnico de análise de dados, mas como uma estrutura integrada de gestão e curadoria FAIR para dados ômicos em resistência antimicrobiana.

Como cenário ilustrativo de aplicação, em um projeto de vigilância genômica de microrganismos resistentes, o *workflow* pode orientar desde o planejamento da coleta e a definição dos metadados mínimos até o processamento bioinformático, a documentação da

proveniência, a escolha de vocabulários e ontologias, a seleção do repositório e a definição das condições de acesso e reúso.

Quadro 3 - Cenário operacional de aplicação à vigilância genômica de RAM

Etapa	Ação prática	Evidência esperada	Registro no FIP
Planejamento	Definir PGD, responsabilidades, requisitos éticos e dados mínimos	PGD; protocolo; responsáveis; escolha preliminar de padrões	Padrões de metadados, Identificadores persistentes, políticas e licenças
Coleta	Registrar origem, amostra, ambiente e condições experimentais	*BioSample*; MIxS; identificador da amostra; cadeia de custódia	Identificadores persistentes de amostra, esquema de metadados e vocabulários
Processamento e análise	Executar controle de qualidade e pipeline reprodutível	*Galaxy/Nextflow/nf-*core; versões; parâmetros; logs	Formatos, *software*, proveniência e empacotamento
Curadoria e representação	Normalizar metadados e relações semânticas	ISA-Tab; GO; EDAM; ChEBI; NCBI *Taxonomy; Sequence Ontology*	Esquemas, ontologias e vocabulários
Armazenamento e preservação	Selecionar ambientes e formatos adequados	NCBI/ENA/EGA; repositório institucional; política de preservação	Repositório, formatos, Identificadores persistentes e política de retenção
Compartilhamento e acesso	Definir abertura, restrição, protocolo e licença	Metadados públicos; acesso controlado quando necessário; licença explícita	Protocolos, autenticação, autorização e licença
Reúso e conhecimento	Permitir descoberta, citação, integração e nova análise	DOI; *OmicsDI; WorkflowHub; RO-Crate*; documentação	Identificadores persistentes, metadados ricos, proveniência e condições de reuso

Fonte: elaboração própria (2026).

Já o FIP funcionaria como registro transversal das escolhas adotadas em cada etapa, permitindo comparar práticas entre laboratórios e identificar pontos de convergência ou divergência na implementação FAIR. Esse cenário é apresentado apenas para demonstrar o potencial de aplicação do modelo e não constitui validação empírica da proposta.

5 CONSIDERAÇÕES FINAIS

Do ponto de vista teórico, o estudo traz contribuições para a Ciência da Informação ao aproximar os conceitos de organização, representação, curadoria, interoperabilidade e governança de dados científicos das necessidades concretas da pesquisa ômica aplicada à saúde pública. A integração entre FAIR, FIP e workflow amplia a compreensão de que a qualidade dos dados depende não apenas de seu processamento analítico, mas também de sua documentação, contextualização, preservação e capacidade de reúso.

Em termos práticos, a proposta pode apoiar equipes multidisciplinares na definição de práticas mais consistentes para a gestão de dados ômicos. O workflow pode ser adaptado por laboratórios, instituições de pesquisa, serviços de vigilância genômica e iniciativas de saúde pública interessadas em estruturar processos de gestão de dados alinhados aos Princípios FAIR.

Como limitação, reconhece-se que o workflow científico possui caráter preliminar e conceitual, tendo sido elaborado com base em levantamento bibliográfico e documental e na análise de workflows e recursos existentes. Embora a proposta sistematize etapas, recursos e decisões relacionadas à gestão e à curadoria de dados ômicos, sua utilização em contextos reais de pesquisa e vigilância ainda precisa ser validada junto à comunidade científica e aos profissionais do domínio, por meio de oficinas de cocriação com profissionais envolvidos na vigilância da resistência antimicrobiana na área de dados ômicos. Esta validação está prevista na próxima fase da pesquisa.

Essa etapa deverá combinar a aplicação piloto em um projeto real e avaliação estruturada de critérios como clareza, completude, usabilidade, aderência aos Princípios FAIR, capacidade de documentar recursos por meio do FIP e adaptabilidade a diferentes contextos institucionais.

Os resultados dessa validação poderão subsidiar a revisão e o refinamento do modelo, permitindo avaliar de forma mais robusta sua aplicabilidade, usabilidade e capacidade de adaptação a diferentes cenários de pesquisa em saúde.

Por fim, a integração entre Princípios FAIR, Perfil de Implementação FAIR e *workflow* científico oferece uma abordagem promissora para fortalecer a gestão, a curadoria e a governança de dados ômicos relacionados à resistência antimicrobiana. Ao tornar explícitas as escolhas técnicas, semânticas e institucionais necessárias ao ciclo de vida dos dados, a

proposta contribui para ampliar sua interoperabilidade, rastreabilidade e potencial de reúso na produção de conhecimento científico e no enfrentamento da RAM.

REFERÊNCIAS

ALCOCK, Brian P. et al. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. **Nucleic Acids Research**, Oxford, v. 51, n. D1, p. D690-D699, 2023. Disponível em: https://pubmed.ncbi.nlm.nih.gov/36263822/. Acesso em: 10 nov. 2026

AMID, Clara; ALAKO, Blaise T. F.; KADHIRVELU, Vishnukumar Balavenkataraman; BURDETT, Tony *et al.* The European Nucleotide Archive in 2019**. Nucleic Acids Research**, Oxford, v. 48, n. D1, p. D70-D76, 2020. Disponível em: https://academic.oup.com/nar/article/48/D1/D70/5624970. Acesso em: 10 nov. 2026

COLLINS, Francis S.; GREEN, Eric D.; GUTTMACHER, Alan E.; GUYER, Mark S. A vision for the future of genomics research. **Nature**, London, v. 422, p. 835-847, 2003 . Disponível em: https://pubmed.ncbi.nlm.nih.gov/12695777/. Acesso em: 10 nov. 2026

FECHER, Benedikt; FRIESIKE, Sascha. Open science: one term, five schools of thought. *In*: BARTLING, Sönke; FRIESIKE, Sascha (ed.). **Opening science**: the evolving guide on how the Internet is changing research, collaboration and scholarly publishing. Cham: Springer, 2014. p. 17-47. Disponível em: https://link.springer.com/chapter/10.1007/978-3-319-00026-8_2. Acesso em: 10 nov. 2026

GUIMARÃES, Denise Oliveira; MOMESSO, Luciano da Silva; PUPO, Mônica Tallarico. Antibióticos: importância terapêutica e perspectivas para a descoberta e desenvolvimento de novos agentes**. Química Nova**, São Paulo, v. 33, n. 3, p. 667-679, 2010. Disponível em: https://www.scielo.br/j/qn/a/dhKT3h4ZxxvsQdkzyZ4VnpB/. Acesso em: 14 jun. 2026.

HASIN, Yehudit; SELDIN, Marcus; LUSIS, Aldons. Multi-omics approaches to disease. **Genome Biology**, London, v. 18, art. 83, 2017. Disponível em: https://link.springer.com/article/10.1186/s13059-017-1215-1. Acesso em: 10 nov. 2026

HAYES, C. Nelson et al. From omics to multi-omics: a review of advantages and tradeoffs. **Genes**, **Basel**, v. 15, n. 12, art. 1551, 2024. Disponível em: https://www.mdpi.com/2073-4425/15/12/1551. Acesso em: 10 nov. 2026

LAXMINARAYAN, Ramanan et al. Antibiotic resistance: the need for global solutions. The **Lancet Infectious Diseases**, London, v. 13, n. 12, p. 1057-1098, 2013. Disponível em: https://pubmed.ncbi.nlm.nih.gov/24252483/. Acesso em: 10 nov. 2026

LEFEVER, Steve et al. RDML: structured language and reporting guidelines for real-time quantitative PCR data**. Nucleic Acids Research**, Oxford, v. 37, n. 7, p. 2065-2069, 2009. Disponível em: https://academic.oup.com/nar/article/37/7/2065/1017981. Acesso em: 10 nov. 2026.

MANOLIO, Teri A.; CHISHOLM, Rex L.; OZENBERGER, Brad; RODEN, Dan M. *et al.* Implementing genomic medicine in the clinic: the future is here. **Genetics in Medicine**, New York, v. 15, n. 4, p. 258-267, 2013. Disponível em: https://doi.org/10.1038/gim.2012.157. Acesso em: 10 nov. 2026

O'NEILL, Jim. **Tackling drug-resistant infections globally**: final report and recommendations. London: Review on Antimicrobial Resistance, 2016. Disponível em: https://amr-review.org/. Acesso em: 17 jan. 2026.

SCHULTES, Erik et al. Reusable FAIR implementation profiles as accelerators of FAIR convergence. In: INTERNATIONAL CONFERENCE ON CONCEPTUAL MODELING, 39., 2020, [*s. l.*]. **Advances in conceptual modeling**. Cham: Springer, 2020. p. 138-147. Disponível em: https://dl.acm.org/doi/10.1007/978-3-030-65847-2_13. Acesso em: 10 nov. 2026

SLATKO, Barton E.; GARDNER, Andrew F.; AUSUBEL, Frederick M. Overview of next-generation sequencing technologies**. Current Protocols in Molecular Biology**, Hoboken, v. 122, n. 1, e59, 2018. Disponível em: https://doi.org/10.1002/cpmb.59. Acesso em: 10 nov. 2026

UNESCO**. Recomendação da UNESCO sobre ciência aberta**. Paris: UNESCO, 2021. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000379949_por. Acesso em: 17 jan. 2026.

WILKINSON, Mark D.; DUMONTIER, Michel; AALBERSBERG, IJsbrand Jan; APPLETON, Gabrielle *et al.* The FAIR guiding principles for scientific data management and stewardship. **Scientific Data**, London, v. 3, art. 160018, 2016. Disponível em: https://www.nature.com/articles/sdata201618. Acesso em: 10 nov. 2026

WORLD HEALTH ORGANIZATION. **No time to wait**: securing the future from drug-resistant infections. Report to the Secretary-General of the United Nations. Geneva: WHO, 2019. Disponível em: https://www.who.int/publications/i/item/no-time-to-wait-securing-the-future-from-drug-resistant-infections. Acesso em: 13 jun. 2025.