Governança de Dados Sensíveis: Um Estudo de Caso sobre a Arquitetura Integrada Baseada no Five Safes Framework, SCIF e Programa de Privacidade e Segurança da Informação (PPSI)

1st José Marivaldo de Aragão Silva
Nucleo de Segurança da Informação
CIDACS — FIOCRUZ
Salvador, Bahia, Brazil
jose.msilva@fiocruz.br

2st Eliseu Silva Torres
Nucleo de Segurança da Informação
CIDACS — FIOCRUZ
Salvador, Bahia, Brazil
eliseu.torres@fiocruz.br

3nd Prof. Dr. Luiz Antonio Ribeiro Junior
Assistant Professor (Institute of Physics)
Universidade de Brasília (UnB)
Distrito Federal, Brasília, Brazil
ribeirojr@unb.br

***Resumo*—O crescente uso de dados sensíveis em pesquisas em saúde demanda mecanismos robustos de governança, segurança da informação e proteção da privacidade. Este estudo tem como objetivo analisar a implementação integrada do *Five Safes Framework* (FSF), do *Sensitive Compartmented Information Facility* (SCIF) e do Programa de Privacidade e Segurança da Informação (PPSI) na governança de dados sensíveis. A pesquisa adota uma abordagem qualitativa, caracterizada como estudo de caso, fundamentada em pesquisa documental e apoiada por uma revisão de literatura estruturada, tendo como contexto de análise o Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS). A revisão de literatura estruturada permitiu identificar e consolidar os fundamentos teóricos e as boas práticas relacionadas à governança de dados sensíveis, à segurança da informação e à proteção da privacidade, fornecendo o embasamento conceitual para a análise da implementação integrada do FSF, da SCIF e do PPSI no contexto do estudo de caso. Os resultados evidenciam que a implementação integrada desses componentes fortalece os controles de acesso, a proteção dos ambientes de processamento e análise, a gestão de riscos, a conformidade regulatória e a governança dos dados sensíveis. Conclui-se que a adoção coordenada desses componentes contribui para um ambiente de pesquisa mais seguro, transparente e alinhado às boas práticas de governança de dados, oferecendo subsídios para instituições que lidam com informações sensíveis em saúde e pesquisa.**

***Index Terms*—Governança de Dados; Dados Sensíveis; FSF; SCIF; PPSI; CIDACS; Segurança da Informação.**

I. Introdução

A evolução das tecnologias digitais aplicadas à saúde e pesquisa tem impulsionado novas possibilidades para coleta, integração e análise de grandes volumes de informações provenientes de diferentes fontes institucionais. Esses ecossistemas de dados têm desempenhado papel fundamental no avanço da pesquisa epidemiológica, na identificação de padrões populacionais e no desenvolvimento de estratégias baseadas em evidências para o aprimoramento das políticas públicas de saúde. Entretanto, a crescente complexidade desses ambientes amplia a necessidade de estruturas de governança capazes de assegurar que o uso dessas informações ocorra de forma segura, ética e responsável.

No contexto da pesquisa em saúde, dados provenientes de registros administrativos, sistemas de informação governamentais e bases populacionais frequentemente apresentam características sensíveis, exigindo mecanismos específicos para proteção da privacidade dos indivíduos e mitigação de riscos associados ao acesso, processamento, compartilhamento e possível reidentificação. Dessa forma, a governança de dados sensíveis passa a incorporar não apenas aspectos tecnológicos, mas também dimensões organizacionais, regulatórias e operacionais relacionadas ao gerenciamento do ciclo de vida das informações [1], [2].

O *Centro de Integração de Dados e Conhecimentos para Saúde* (CIDACS) [3], vinculado à Fundação Oswaldo Cruz (Fiocruz Bahia), constitui uma infraestrutura de pesquisa dedicada à integração e análise de grandes volumes de dados administrativos e de saúde, com o objetivo de apoiar estudos epidemiológicos e subsidiar a formulação e avaliação de políticas públicas. Na literatura, o CIDACS é reconhecido pelo desenvolvimento de estratégias avançadas de *record linkage*, voltadas à integração segura de diferentes fontes de dados, preservando requisitos de segurança da informação, privacidade e confidencialidade. Entre suas principais contribuições destaca-se o sistema *CIDACS Record Linkage* (CIDACS-RL), uma solução computacional que combina técnicas determinísticas e probabilísticas, estratégias de indexação e algoritmos de pontuação para realizar a vinculação de registros em bases contendo milhões de indivíduos, alcançando elevados níveis de acurácia, escalabilidade e desempenho [4].

Para garantir que o acesso e o processamento dessas informações ocorram de forma segura e em conformidade com requisitos éticos e regulatórios, ambientes de pesquisa em saúde e pesquisa demandam arquiteturas integradas de governança de dados, segurança da informação e privacidade. Nesse contexto, o CIDACS apresenta mecanismos alinhados aos princípios do *Five Safes Framework* (FSF) , ambientes se-

guros para processamento e análise de informações sensíveis, representados pela abordagem *Sensitive Compartmented Information Facility* (SCIF), e iniciativas relacionadas ao Programa de Privacidade e Segurança da Informação (PPSI). Essa abordagem combina princípios de governança de acesso aos dados, controles institucionais e mecanismos técnicos de proteção, fortalecendo a confidencialidade, integridade e disponibilidade das informações sensíveis. [5]–[7]

A pesquisa adota uma abordagem qualitativa, caracterizada como estudo de caso, fundamentada em pesquisa documental e apoiada por uma revisão de literatura estruturada. A revisão permitiu identificar os principais conceitos, frameworks e práticas relacionados à governança de dados sensíveis, segurança da informação e privacidade, enquanto a pesquisa documental possibilitou analisar políticas, procedimentos e mecanismos institucionais associados ao contexto estudado.

Nesse contexto, este artigo busca contribuir para a discussão sobre arquiteturas integradas de proteção de dados sensíveis, evidenciando como diferentes abordagens de governança, segurança e privacidade podem ser combinadas para fortalecer ecossistemas de segurança. A análise considera que a integração entre frameworks de governança, ambientes seguros, como SCIF,salvaguardas e controles técnicos constitui um elemento estratégico para promover confidencialidade, integridade, disponibilidade e conformidade regulatória no tratamento de informações sensíveis.

O restante deste artigo está organizado da seguinte forma: a Seção **II** apresenta a fundamentação teórica sobre governança de dados, privacidade,segurança da informação e as principais normas e frameworks associados; a Seção **III** descreve o contexto do CIDACS e sua infraestrutura de integração de dados; a Seção **IV** apresenta a análise da implementação integrada do *Five Safes Framework*, da *Secure Compartmented Information Facility* (SCIF) e do Programa de Privacidade e Segurança da Informação (PPSI); a Seção **V** discute os resultados obtidos; e, por fim, a Seção **VI** apresenta as considerações finais e as perspectivas para trabalhos futuros.

II. Fundamentação Teórica

A. Governança de Dados

A governança de dados pode ser definida como o conjunto de políticas, papéis, processos, padrões e métricas que asseguram o uso eficaz e eficiente da informação, permitindo que uma organização alcance seus objetivos [8]. Segundo o *Data Management Body of Knowledge* (DAMA-DMBOK), a governança de dados constitui uma função transversal responsável por orientar as demais disciplinas de gestão de dados, incluindo qualidade, arquitetura, segurança, integração e metadados, por meio de mecanismos de tomada de decisão, definição de responsabilidades (*accountability*), conformidade e controle [8]. Dessa forma, a governança estabelece diretrizes para garantir que os ativos de dados sejam gerenciados de forma consistente, segura e alinhada aos objetivos estratégicos da organização.

KHATRI E BROWN(2010) propõem um modelo de governança de dados estruturado em cinco domínios de decisão: princípios de dados, qualidade de dados, metadados, acesso aos dados e ciclo de vida dos dados. Esse modelo é amplamente utilizado como referência na literatura por estabelecer uma estrutura de tomada de decisão para os ativos de dados, em analogia aos mecanismos de governança de Tecnologia da Informação (TI) [9].

No contexto da governança corporativa de TI, a norma ISO/IEC 38500 estabelece princípios para o uso eficaz, eficiente e aceitável da tecnologia da informação nas organizações, fornecendo diretrizes para que a alta administração avalie, direcione e monitore o uso da TI [10]. Complementarmente, o framework COBIT 2019 amplia essa perspectiva ao definir objetivos de governança e gestão (*Governance and Management Objectives*), conectando estratégia organizacional, gestão de riscos, conformidade e controles relacionados aos ativos de informação e aos dados, sendo amplamente empregado como referência em iniciativas de governança de dados [11].

B. Privacidade de Dados

A privacidade, no contexto informacional, é fundamentada no conceito de *Privacy by Design* (PbD), proposto por CAVOUKIAN (2009), que estabelece sete princípios destinados à incorporação da proteção de dados pessoais desde a concepção de sistemas, processos e tecnologias. Entre esses princípios destacam-se a proteção proativa (e não reativa), a privacidade como configuração padrão (*privacy as the default setting*), a integração da privacidade ao projeto (*privacy embedded into design*) e a garantia da segurança durante todo o ciclo de vida da informação [12].

Sob a perspectiva acadêmica, SOLOVE(2006) desenvolveu uma taxonomia da privacidade que classifica os riscos informacionais em quatro categorias principais: coleta de informações, processamento de informações, disseminação de informações e invasões à esfera privada. Essa taxonomia tornou-se uma referência para a identificação e análise dos riscos associados ao tratamento de dados pessoais em diferentes contextos organizacionais e tecnológicos [13].

No âmbito regulatório, a Lei Geral de Proteção de Dados Pessoais (LGPD) estabelece princípios como finalidade, adequação, necessidade, livre acesso, qualidade dos dados, transparência, segurança, prevenção e responsabilização (*accountability*), regulamentando o tratamento de dados pessoais no Brasil [14]. A legislação brasileira foi fortemente influenciada pelo Regulamento Geral sobre a Proteção de Dados da União Europeia (*General Data Protection Regulation* – GDPR), considerado a principal referência internacional em proteção de dados ao introduzir mecanismos como o *Data Protection Impact Assessment* (DPIA), a proteção de dados desde a concepção (*Data Protection by Design and by Default*) e a figura do Encarregado pela Proteção de Dados (*Data Protection Officer(DPO)* [15].

Complementando esse arcabouço, a norma ISO/IEC 27701 estende os requisitos dos Sistemas de Gestão de Segurança da Informação (SGSI), definidos na ISO/IEC 27001, para

contemplar um Sistema de Gestão de Informações de Privacidade (*Privacy Information Management System (PIMS)*, fornecendo controles específicos para organizações que atuam como controladoras ou operadoras de dados pessoais [16]. Em uma abordagem baseada em gestão de riscos, o *NIST Privacy Framework* estrutura a governança da privacidade em três componentes principais (*Core*, *Profiles* e *Implementation Tiers*) permitindo que organizações identifiquem, avaliem e gerenciem riscos relacionados à privacidade de maneira integrada às estratégias de governança e segurança da informação [17].

No âmbito brasileiro, o Programa de Privacidade e Segurança da Informação (PPSI) estabelece um conjunto estruturado de diretrizes, políticas, processos e controles destinados à proteção de dados sensíveis ao longo de todo o seu ciclo de vida. Em uma abordagem semelhante à proposta pelo *Center for Internet Security (CIS) Controls*, o PPSI organiza controles técnicos, administrativos e organizacionais voltados à gestão de riscos, controle de acesso, classificação da informação, monitoramento, resposta a incidentes, conformidade regulatória e capacitação dos usuários, promovendo uma estratégia integrada de governança de dados, segurança da informação e privacidade [6], [18]. Dessa forma, o programa operacionaliza os requisitos previstos na Lei Geral de Proteção de Dados Pessoais (LGPD), nas normas da família ISO/IEC 27000 e nas políticas institucionais da Fiocruz, estabelecendo um modelo contínuo de melhoria da maturidade em segurança e privacidade da informação.

C. Segurança da Informação

A segurança da informação é tradicionalmente fundamentada na tríade Confidencialidade, Integridade e Disponibilidade (CID), que estabelece os princípios essenciais para a proteção dos ativos de informação. A confidencialidade assegura que as informações sejam acessadas apenas por entidades autorizadas; a integridade garante a exatidão, consistência e completude dos dados ao longo de seu ciclo de vida; e a disponibilidade assegura que informações e serviços permaneçam acessíveis sempre que necessários. Esses princípios constituem a base conceitual para o desenvolvimento de políticas, controles e mecanismos de proteção em ambientes computacionais [19].

No contexto normativo internacional, a família ISO/IEC 27000 representa o principal conjunto de normas para gestão da segurança da informação. A ISO/IEC 27001 estabelece os requisitos para implementação, manutenção e melhoria contínua de um Sistema de Gestão de Segurança da Informação (SGSI), baseado em uma abordagem sistemática de gestão de riscos [20]. Complementarmente, a ISO/IEC 27002 fornece diretrizes para seleção, implementação e gerenciamento de controles de segurança, organizados em quatro grandes categorias: controles organizacionais, controles de pessoas, controles físicos e controles tecnológicos, constituindo uma referência para a implementação das medidas previstas na ISO/IEC 27001 [21].

No âmbito das estruturas de referência para cibersegurança, o *NIST Cybersecurity Framework* (CSF) estabelece uma abordagem baseada em riscos para o gerenciamento da segurança cibernética, estruturada nas funções *Govern*, *Identify*, *Protect*, *Detect*, *Respond* e *Recover*. Essas funções permitem integrar governança, gestão de riscos e controles operacionais em organizações de diferentes portes e setores [22]. Em complemento, a publicação especial NIST SP 800-53 apresenta um catálogo abrangente de controles de segurança e privacidade aplicáveis a sistemas de informação e organizações, enquanto a NIST SP 800-37 define o *Risk Management Framework* (RMF), estabelecendo um processo contínuo para categorização, seleção, implementação, avaliação, autorização e monitoramento dos controles de segurança [23], [24].

No contexto operacional, o *Center for Internet Security* (CIS) disponibiliza os *CIS Critical Security Controls*, um conjunto priorizado de salvaguardas técnicas e administrativas destinadas à mitigação das ameaças cibernéticas mais relevantes. Organizados em dezoito domínios de controle, os CIS Controls fornecem uma abordagem prescritiva para implementação de controles de segurança, sendo amplamente utilizados como referência para programas de maturidade em cibersegurança e para a operacionalização dos requisitos estabelecidos por frameworks como a ISO/IEC 27001 e o NIST CSF [25].

D. Five Safes Framework

O *Five Safes Framework* (FSF) foi desenvolvido como uma abordagem sociotécnica para apoiar o acesso seguro e responsável a dados sensíveis, especialmente em ambientes de pesquisa que utilizam informações confidenciais. Diferentemente de modelos tradicionais de segurança baseados exclusivamente em restrições técnicas, o Five Safes considera que a proteção dos dados depende da interação entre aspectos humanos, organizacionais, tecnológicos e contextuais.

Segundo Desai et al. [26], o Five Safes tem sido adotado internacionalmente por organizações responsáveis por disponibilizar dados para pesquisa, permitindo equilibrar a necessidade de acesso científico com requisitos de privacidade e segurança. Dessa forma, o framework apresenta uma abordagem baseada em risco, na qual diferentes camadas de proteção são combinadas para permitir o uso seguro de dados sensíveis.

No cenário da pesquisa em saúde, o Five Safes apresenta elevada relevância devido à necessidade de conciliar a utilização de grandes bases populacionais com a proteção dos direitos dos indivíduos. Ambientes que processam dados clínicos, epidemiológicos e administrativos necessitam de mecanismos capazes de controlar quem acessa os dados, em quais condições o processamento ocorre e quais informações podem ser disponibilizadas como resultado das análises.

Já no contexto brasileiro, os princípios do *Five Safes Framework* encontram aderência às práticas adotadas pelo Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS), cuja infraestrutura foi concebida para viabilizar o acesso seguro a grandes bases de dados sensíveis utilizadas

TABLE I
DIMENSÕES DO *Five Safes Framework*.

Dimensão	Descrição
Safe Projects	Avalia se o projeto é apropriado, legal, ético e capaz de produzir benefícios científicos ou sociais, em conformidade com requisitos institucionais e regulatórios.
Safe People	Verifica se os usuários possuem capacitação, credenciamento e autorização para acessar e utilizar os dados de forma responsável, respeitando políticas de segurança e privacidade.
Safe Data	Analisa o nível de sensibilidade dos dados e os mecanismos empregados para reduzir riscos de identificação ou reidentificação, incluindo anonimização, pseudonimização e minimização de dados.
Safe Settings	Refere-se ao ambiente físico e computacional utilizado para armazenamento, processamento e análise dos dados, contemplando controles de acesso, monitoramento, auditoria e mecanismos de isolamento e proteção.
Safe Outputs	Avalia os resultados produzidos pelas análises, assegurando que as informações divulgadas não permitam a identificação direta ou indireta dos indivíduos presentes nas bases de dados.

em pesquisas populacionais. A combinação de processos de governança de dados, mecanismos de controle de acesso, técnicas de desidentificação, ambientes computacionais seguros e procedimentos de avaliação dos resultados demonstra uma abordagem alinhada às cinco dimensões .Dessa forma, o CIDACS representa um ambiente de referência para a aplicação integrada desses princípios, conciliando a produção de conhecimento científico com a proteção da privacidade dos indivíduos, a conformidade com requisitos éticos e legais e a mitigação dos riscos associados ao tratamento de dados sensíveis. Nesse cenário, o presente trabalho propõe o fortalecimento da dimensão *Safe Settings* por meio da adoção de uma arquitetura inspirada no conceito de *Sensitive Compartmented Information Facility* (SCIF), integrada ao Programa de Privacidade e Segurança da Informação (PPSI) pela salvaguarda , ampliando os mecanismos de proteção já existentes na infraestrutura do CIDACS.

E. Programa de Privacidade e Segurança da Informação (PPSI)

Programas de Privacidade e Segurança da Informação (PPSI) constituem estruturas institucionais destinadas ao estabelecimento de diretrizes, responsabilidades, processos e mecanismos de controle voltados à proteção dos ativos informacionais e ao gerenciamento dos riscos associados à segurança da informação e à privacidade dos dados. Diferentemente de abordagens baseadas em controles isolados ou medidas predominantemente técnicas, um PPSI propõe uma visão integrada, articulando dimensões estratégicas, operacionais, normativas e regulatórias, de modo a promover uma abordagem sistemática para a governança da informação.

No Brasil, os PPSI têm sido adotados como instrumentos de fortalecimento da governança de segurança da informação em organizações públicas e privadas, estabelecendo uma conexão entre requisitos institucionais, princípios de gestão de riscos, proteção de dados pessoais e conformidade regulatória [6]. Sua implementação demanda a definição de políticas e normas internas, mecanismos de gestão de identidade e acesso, processos de classificação e tratamento da informação, planos de resposta a incidentes, programas de conscientização e capacitação, avaliação contínua de riscos e monitoramento da efetividade dos controles de segurança implementados.

Em ambientes de pesquisa em saúde, a adoção de programas estruturados de privacidade e segurança da informação assume papel estratégico em razão da criticidade e sensibilidade dos dados processados. Dados relacionados à saúde apresentam elevado potencial de impacto social, ético e institucional quando submetidos a acessos indevidos, alterações não autorizadas, perdas de integridade ou usos incompatíveis com sua finalidade original. Nesse cenário, mecanismos robustos de governança tornam-se fundamentais para assegurar que as atividades de coleta, armazenamento, processamento, compartilhamento e disponibilização dos dados estejam alinhadas aos princípios de confidencialidade, integridade, disponibilidade, transparência e responsabilização.

A implementação de um Programa de Privacidade e Segurança da Informação (PPSI) requer a adoção de um conjunto estruturado de controles capazes de estabelecer mecanismos institucionais, técnicos e operacionais voltados à proteção dos ativos informacionais. Esses controles devem considerar não apenas aspectos tecnológicos, mas também fatores humanos, processos organizacionais, requisitos regulatórios e princípios de governança, possibilitando uma abordagem integrada para gerenciamento de riscos relacionados à segurança e privacidade dos dados.

Entretanto, considerando o escopo deste artigo, não será realizada uma descrição detalhada dos 25 controles que compõem um PPSI. A análise será direcionada aos controles que apresentam maior aderência ao contexto da pesquisa, especialmente aqueles relacionados aos desafios de governança e proteção de dados sensíveis no Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS). Dessa forma, serão priorizados os controles associados à governança da informação, gestão de ativos, classificação e proteção de dados, controle de acesso, gestão de riscos, monitoramento, rastreabilidade e conformidade regulatória.

Essa delimitação permite estabelecer uma relação direta entre os controles selecionados do PPSI, os princípios do Five Safes Framework e os requisitos de um Secure Data Environment, ou Sensitive Compartmented Information Facility (SCIF), considerando a necessidade de criação de ambientes controlados para acesso, processamento e análise de dados sensíveis em pesquisas na área da saúde. Assim, os controles analisados representam aqueles que contribuem diretamente para o fortalecimento da governança sociotécnica, da proteção da privacidade e da segurança operacional dos ambientes de pesquisa baseados em dados.

F. Sensitive Compartmented Information Facility (SCIF)

A *Sensitive Compartmented Information Facility* (SCIF) representa um ambiente físico e tecnológico projetado para proteger o processamento, armazenamento e análise de informações sensíveis por meio de controles rigorosos de segurança física, tecnológica e administrativa [27]. Originalmente desenvolvida no contexto de ambientes governamentais e de inteligência, a abordagem SCIF estabelece requisitos específicos para restringir acessos não autorizados, reduzir riscos de comprometimento da informação e garantir que atividades envolvendo dados críticos sejam realizadas em ambientes controlados e devidamente protegidos.

Uma SCIF combina mecanismos de segurança física, controles tecnológicos, procedimentos operacionais e restrições administrativas, constituindo uma arquitetura de proteção em camadas. Entre os principais elementos associados a esses ambientes destacam-se o controle rigoroso de acesso físico e lógico, autenticação e autorização de usuários, segregação de ambientes computacionais, monitoramento contínuo das atividades, auditoria, restrições quanto ao uso de dispositivos externos e aplicação de políticas específicas para manipulação de informações classificadas ou sensíveis [27], [28]. Esses mecanismos têm como objetivo garantir que somente indivíduos autorizados possam acessar, processar ou interagir com informações protegidas, reduzindo a possibilidade de exposição, alteração indevida ou extração não autorizada dos dados.

Embora tenha origem em cenários relacionados à segurança nacional e proteção de informações classificadas, o conceito de ambientes seguros e segregados apresenta aplicabilidade crescente em outros domínios que demandam elevados níveis de proteção da informação, incluindo centros de pesquisa, instituições de saúde, organizações governamentais e ambientes destinados à análise de grandes bases de dados populacionais. Nesses contextos, os princípios associados ao modelo SCIF contribuem para minimizar riscos relacionados ao acesso indevido, vazamento de informações, comprometimento da confidencialidade e violação da privacidade dos indivíduos participantes de pesquisas.

No contexto da governança de dados sensíveis em saúde, a SCIF pode ser compreendida como uma camada operacional destinada à implementação da dimensão *Safe Settings* estabelecida pelo *Five Safes Framework*. Enquanto o Five Safes fornece um modelo de avaliação baseado em risco para determinar se um ambiente apresenta condições adequadas para acesso e uso de dados sensíveis, a SCIF fornece os mecanismos físicos, tecnológicos e procedimentais necessários para tornar esse ambiente confiável, controlado e auditável [26]–[28]. Dessa forma,esta dimensão estabelece os requisitos de segurança do ambiente, enquanto a arquitetura SCIF materializa esses requisitos por meio de controles concretos de proteção.Esses mecanismos incluem controle de acesso, segregação de ambientes e monitoramento contínuo para proteção das informações sensíveis.

G. Governança de Dados Sensíveis em Pesquisa e Saúde

Após a análise dos conceitos apresentados, observa-se que os dados de saúde ocupam uma posição singular na interseção entre governança de dados, privacidade e segurança da informação, pois são classificados como dados pessoais sensíveis tanto LGPD quanto pelo GDPR. Em razão de seu elevado potencial de impacto sobre os direitos e liberdades dos indivíduos, esses dados demandam mecanismos específicos de governança, bases legais apropriadas para seu tratamento e controles técnicos, administrativos e organizacionais capazes de reduzir riscos relacionados ao acesso não autorizado, vazamento, reidentificação e uso indevido das informações. No cenário internacional, destaca-se ainda a Health Insurance Portability and Accountability Act (HIPAA) [29], especialmente por meio da HIPAA Security Rule, que estabelece salvaguardas administrativas, físicas e técnicas para garantir a confidencialidade, a integridade e a disponibilidade das informações eletrônicas de saúde (ePHI - electronic Protected Health Information), constituindo uma das principais referências mundiais para a proteção de dados em ambientes de saúde.

Nesse contexto, os princípios do Privacy by Design assumem papel fundamental ao estabelecer que a proteção da privacidade deve ser ser incorporada desde a concepção dos processos, sistemas e infraestruturas responsáveis pelo tratamento dos dados. No cenário brasileiro, o Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS), vinculado à Fundação Oswaldo Cruz (Fiocruz Bahia), constitui uma infraestrutura de pesquisa que incorpora práticas de governança de dados, privacidade e segurança da informação em suas atividades de integração, armazenamento, processamento e disponibilização controlada de grandes bases populacionais em saúde. Sua atuação está estruturada por processos institucionais que contemplam governança, avaliação de projetos, controle de acesso, definição de responsabilidades, proteção da infraestrutura computacional, monitoramento e auditoria, apresentando forte alinhamento aos princípios do Five Safes Framework nas dimensões citadas neste artigo.

Além disso, o CIDACS adota ambientes seguros para o processamento e análise de dados sensíveis, utilizando controles institucionais e tecnológicos que convergem com os princípios da Secure Compartmented Information Facility (SCIF), especialmente quanto à segregação de ambientes, restrição de acesso, rastreabilidade das operações e proteção da infraestrutura. Nesse contexto, a integração entre os mecanismos já existentes no CIDACS, os princípios do Five Safes Framework, os requisitos inspirados em SCIF e as diretrizes do Programa de Privacidade e Segurança da Informação (PPSI) representa uma evolução da maturidade institucional, fortalecendo a governança, a gestão de riscos, a conformidade regulatória e a proteção dos dados sensíveis. Assim, a arquitetura proposta neste trabalho busca sistematizar e ampliar os controles já existentes, estabelecendo uma abordagem integrada que combina governança, processos e controles técnicos para promover o uso seguro e responsável de dados em pesquisas em saúde.

III. CIDACS: Contexto Institucional e Infraestrutura para Pesquisa com Dados Sensíveis

Esta seção apresenta o contexto institucional do Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS), destacando sua atuação como uma das principais infraestruturas brasileiras para integração, processamento e análise de grandes bases de dados em saúde. São abordados os principais projetos científicos desenvolvidos pelo centro, abrangendo pesquisas em doenças infecciosas, doenças crônicas, saúde materno-infantil, vigilância epidemiológica, avaliação de políticas públicas e determinantes sociais da saúde, fundamentadas na integração de registros administrativos e bases de dados de saúde. Essas iniciativas utilizam grandes volumes de dados populacionais para produzir evidências científicas que subsidiam a formulação, a implementação e a avaliação de políticas públicas em saúde.

A. Contexto Institucional do CIDACS

O Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS), vinculado ao Instituto Gonçalo Moniz da Fundação Oswaldo Cruz (Fiocruz Bahia), é uma infraestrutura nacional de pesquisa dedicada à integração, vinculação e análise de grandes volumes de dados administrativos, sociais e de saúde. Desde sua criação, em 2016, o centro desenvolve pesquisas voltadas ao estudo dos determinantes sociais da saúde e à avaliação de políticas públicas, reunindo equipes multidisciplinares e uma infraestrutura computacional de alto desempenho para o processar dados sensíveis [30].

As pesquisas conduzidas pelo centro baseiam-se na integração de registros administrativos e sistemas nacionais de informação em saúde, viabilizando a construção de grandes coortes populacionais e estudos longitudinais. Entre suas principais iniciativas destacam-se a *Coorte de 100 Milhões de Brasileiros*, estudos sobre doenças infecciosas e crônicas, saúde materno-infantil, vigilância epidemiológica, determinantes sociais da saúde e avaliação de políticas públicas, além do desenvolvimento do CIDACS-RL para vinculação de registros em larga escala [31], [32].Entre as principais fontes de dados utilizadas encontram-se registros provenientes do Cadastro Único para Programas Sociais (CadÚnico), do Sistema de Informações sobre Mortalidade (SIM), do Sistema de Informações sobre Nascidos Vivos (SINASC), do Sistema de Informação de Agravos de Notificação (SINAN), do Sistema de Informações Hospitalares do Sistema Único de Saúde (SIH/SUS), além de diversas outras bases nacionais disponibilizadas pelo Departamento de Informática do Sistema Único de Saúde (DATASUS). A integração dessas diferentes fontes possibilita acompanhar indivíduos ao longo do tempo, permitindo a realização de estudos sobre doenças infecciosas, doenças crônicas, saúde materno-infantil, vigilância epidemiológica, determinantes sociais da saúde e avaliação de programas e políticas públicas.

Devido ao tratamento de grandes volumes de dados sensíveis, o CIDACS adota uma estrutura de Sala Segura que dados, tecnologia da informação e segurança da informação. Nesse contexto, destacam-se o Núcleo de Segurança da Informação (NSI), responsável pela definição de políticas de segurança, gestão de riscos, controle de acesso, tratamento de incidentes, monitoramento e auditoria. Todas as etapas do processo são protocoladas e registradas, e os documentos e termos preservados para rastreamento (Cadeia de Custódia dos dados). O acesso a Sala Segura do Cidacs somente é autorizado pelo NSI a membros específicos da equipe, que não podem utilizar celular, pen drive ou qualquer outro dispositivo eletrônico, havendo ainda controle, monitoramento e auditoria de entrada e saída de quaisquer ativos pela equipe de Segurança da Informação.A atuação integrada dessas estruturas assegura que todo o ciclo de vida dos dados ocorra em conformidade com requisitos éticos, legais e técnicos de proteção da informação, fornecendo a base institucional para a implementação de mecanismos avançados de governança e segurança de dados [33].

B. Infraestrutura de Integração de Dados em Saúde

A infraestrutura de dados do Centro foi concebida para suportar a integração, o armazenamento, o processamento e a disponibilização segura de grandes volumes de dados administrativos, sociais e de saúde utilizados em pesquisas científicas. Seu fluxo operacional contempla todas as etapas do ciclo de vida da informação, incluindo a obtenção dos dados junto aos órgãos provedores, curadoria, validação, padronização, harmonização, vinculação de registros (*Record Linkage*), anonimização, armazenamento e disponibilização controlada para projetos de pesquisa aprovados. Essa arquitetura possibilita a construção de bases longitudinais destinadas ao desenvolvimento de estudos epidemiológicos e à avaliação de políticas públicas, preservando simultaneamente a qualidade, a traíde CIA, e a rastreabilidade das informações [34].

Em conformidade com LGPD, o tratamento de dados pessoais , observa os princípios da finalidade, necessidade, segurança, transparência e responsabilização, utilizando bases legais específicas para pesquisa em saúde pública. Os dados são recebidos por canais seguros, submetidos a procedimentos de verificação de integridade, catalogação e controle de qualidade antes de serem incorporados ao ambiente de processamento. O acesso às bases é restrito exclusivamente a profissionais autorizados, mediante autenticação, segregação de ambientes computacionais, monitoramento contínuo, registros de auditoria e aplicação do princípio do menor privilégio. Após os processos de integração e vinculação dos registros, os conjuntos de dados disponibilizados aos pesquisadores passam por procedimentos de anonimização ou desidentificação, reduzindo significativamente os riscos de identificação dos titulares dos dados e reforçando a proteção da privacidade [35].

IV. Análise da Implementação Integrada

Fluxo operacional do tratamento de dados no CIDACS, contemplando as etapas de obtenção, curadoria digital, integração de registros, processamento, anonimização e disponibilização controlada para pesquisa científica como demonstrado na Figura 1.

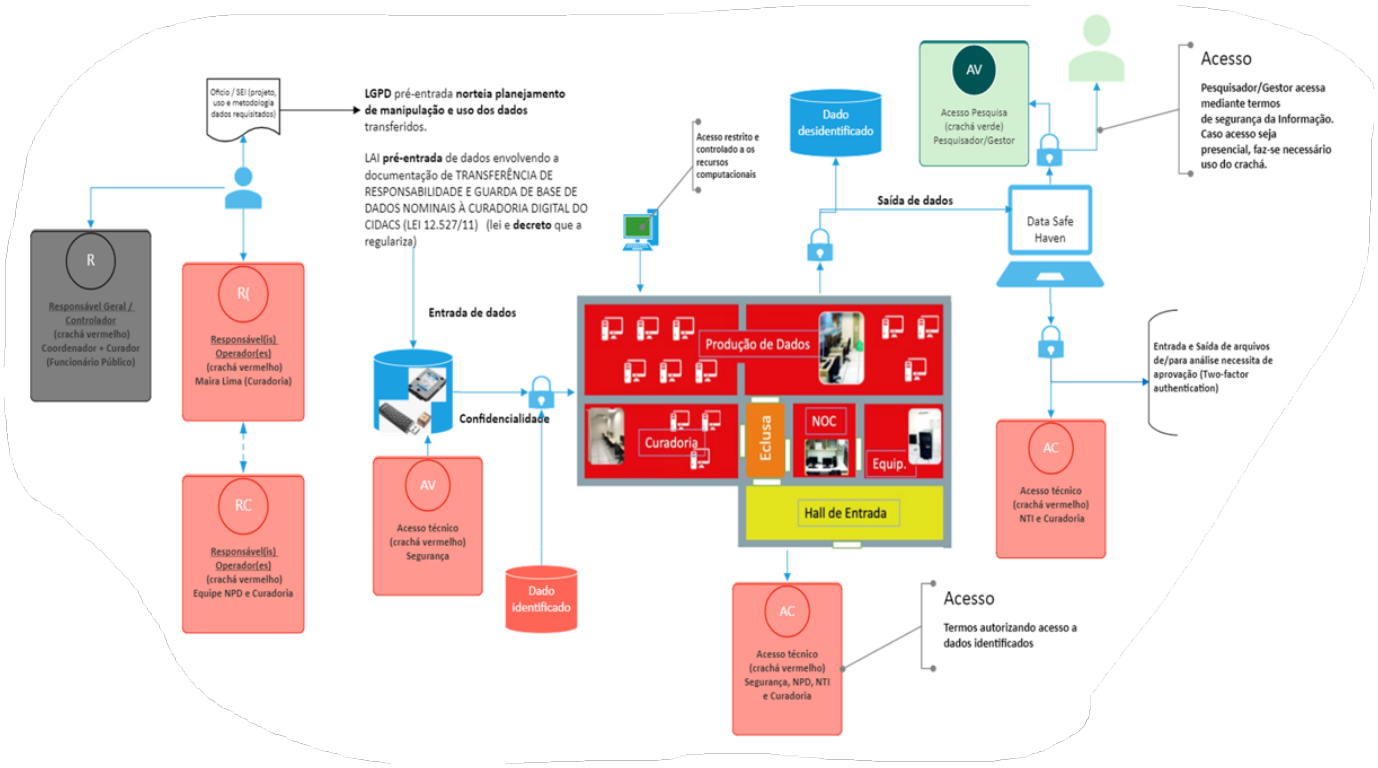

Fig. 1. Fluxo operacional do tratamento de dados no CIDACS. [34].

A infraestrutura tecnológica é complementada por recursos computacionais de alto desempenho destinados ao processamento de grandes volumes de dados e um por mecanismos de segurança , como exemplo a Sala Segura (SCIF do Cidacs) voltados à proteção da infraestrutura, dos ativos informacionais e dos ambientes de pesquisa. Esses mecanismos incluem controle de acesso físico e lógico, segregação entre ambientes de processamento, monitoramento das atividades realizadas, registro de eventos, auditoria, gestão de identidades e proteção dos ativos computacionais. Todas as etapas do processo são protocoladas e registradas, e os documentos e termos preservados para rastreamento. O acesso a Sala Segura somente é autorizado a membros específicos da equipe, que não podem utilizar qualquer outro dispositivo eletrônico, havendo ainda controle, monitoramento e auditoria de entrada e saída de quaisquer ativos pela equipe de Segurança da Informação. O acesso a este ambiente requer uma atuação integrada entre as diferentes áreas responsáveis pela governança, infraestrutura tecnológica e uso científico das informações. Nesse contexto, são estabelecidos mecanismos de articulação entre as lideranças do Núcleo de Segurança da Informação (NSI), do Núcleo de Tecnologia da Informação (NTI) e da Plataforma de Dados, por meio de reuniões periódicas de alinhamento técnico e estratégico. Esses encontros têm como objetivo avaliar requisitos de segurança, acompanhar a evolução dos ambientes computacionais, discutir demandas relacionadas aos projetos, analisar riscos identificados e definir ações nos processos de proteção da informação demonstrados na Tabela II

TABLE II
RESPONSABILIDADES DAS LIDERANÇAS NA ARQUITETURA DE PROTEÇÃO DE DADOS DO CIDACS.

Lideranças	Principais responsabilidades
Plataforma de Dados	Coordenação da governança científica dos dados, definição dos fluxos de integração, organização e disponibilização das informações para pesquisa. Responsável pelo alinhamento entre os requisitos dos projetos científicos, qualidade dos dados, processos de curadoria, integração de registros (*Record Linkage*) e uso adequado das bases populacionais.
Núcleo de Segurança da Informação (NSI)	Coordenação das ações de segurança da informação, proteção dos ambientes computacionais, implementação de controles de acesso, monitoramento contínuo, auditoria, gestão de riscos, análise de eventos de segurança, resposta a incidentes e garantia da confidencialidade, integridade, disponibilidade e rastreabilidade das informações.
Atuação integrada	Articulação entre governança de dados e segurança da informação, garantindo que o acesso e processamento das informações ocorram em ambientes controlados, alinhados aos princípios do *Five Safes Framework*, aos requisitos arquiteturais inspirados em SCIF e aos controles institucionais do PPSI.

O monitoramento contínuo conduzido pelas lideranças apresenta papel fundamental nessa arquitetura, permitindo identificar comportamentos anômalos, validar a conformidade dos acessos autorizados, acompanhar eventos de segurança e responder de forma tempestiva a possíveis incidentes.

V. Resultados e Discussões

Os resultados obtidos indicam que a integração entre o *Five Safes Framework*, o Programa de Privacidade e Segurança da Informação (PPSI) e os princípios das *Sensitive Compartmented Information Facilities* (SCIFs) proporciona um modelo robusto de governança para proteção de dados sensíveis no contexto do Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS).

Observa-se que o PPSI desempenha o papel de estrutura normativa e estratégica, estabelecendo políticas institucionais, responsabilidades, requisitos regulatórios e diretrizes para a gestão da segurança da informação. Em complemento, o Five Safes Framework operacionaliza essas diretrizes por meio de controles voltados à autorização de projetos, qualificação dos usuários, proteção dos dados, segurança dos ambientes e validação dos resultados produzidos. Os princípios das SCIFs reforçam essa arquitetura ao incorporar mecanismos avançados de segurança física, lógica e operacional, ampliando a proteção dos ambientes que processam informações de elevada sensibilidade.

A análise demonstra que a adoção integrada desses três modelos contribui significativamente para o fortalecimento da governança de dados, destacando-se os seguintes benefícios:

- redução dos riscos de reidentificação e de vazamento de informações sensíveis;
- fortalecimento dos mecanismos de autenticação, autorização e controle de acesso baseados no princípio do menor privilégio (*Least Privilege*) e no *Need-to-Know*;
- aumento da conformidade com a Lei Geral de Proteção de Dados (LGPD) e demais requisitos regulatórios relacionados à privacidade e à segurança da informação;
- maior rastreabilidade das operações por meio de auditoria, monitoramento contínuo e registro de eventos;
- incremento da confiança institucional no compartilhamento e uso de dados para pesquisa científica, favorecendo a colaboração entre pesquisadores e instituições;
- fortalecimento da resiliência organizacional frente a ameaças internas, ataques cibernéticos e tentativas de acesso não autorizado.

Apesar dos benefícios observados, a implementação dessa arquitetura integrada apresenta desafios relevantes. Entre eles destacam-se a necessidade de investimentos em infraestrutura tecnológica e física, a adoção de processos contínuos de capacitação dos usuários, a integração entre mecanismos de governança e controles operacionais, além da manutenção permanente da conformidade com requisitos legais e normativos. No caso da adoção de controles inspirados em SCIFs, aspectos relacionados ao custo, à complexidade operacional e à adaptação desses ambientes ao contexto da pesquisa científica também representam fatores que devem ser considerados durante o planejamento e a implementação da solução.

Esse desafio adquire caráter estratégico no contexto brasileiro, uma vez que a Sociedade Brasileira para o Progresso da Ciência (SBPC) recomenda a criação de uma Plataforma Nacional Soberana de Dados em Saúde, destacando que os dados de aproximadamente 170 milhões de usuários do Sistema Único de Saúde (SUS) constituem um patrimônio nacional e uma questão de segurança. O documento ainda aponta o Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS) como referência para a integração de bases nacionais destinadas à pesquisa, aos estudos epidemiológicos e à vigilância em saúde [36]. Essa recomendação evidencia que a proteção de grandes repositórios de dados em saúde deve ser tratada não apenas sob a perspectiva da privacidade, mas também como um requisito estratégico para a soberania e a segurança da informação.

Embora frameworks de governança, como o *Five Safes Framework*, estabeleçam diretrizes para o acesso seguro aos dados, esses modelos concentram-se predominantemente em aspectos administrativos, éticos e operacionais, não contemplando de forma aprofundada requisitos avançados de segurança física, proteção da infraestrutura e mecanismos de contrainteligência empregados em ambientes destinados ao processamento de informações altamente sensíveis.

Por outro lado, as *Sensitive Compartmented Information Facilities* (SCIFs) representam um modelo consolidado para proteção de informações críticas, integrando controles físicos, tecnológicos e procedimentais capazes de mitigar ameaças relacionadas à espionagem, ao acesso não autorizado e ao comprometimento de informações sensíveis. Entretanto, a adoção integral desse modelo em ambientes de pesquisa científica apresenta limitações relacionadas à sua complexidade, aos elevados custos de implementação e às diferenças existentes entre os requisitos de proteção de informações classificadas e aqueles aplicáveis aos dados científicos e de saúde.

Embora o Programa de Privacidade e Segurança da Informação (PPSI) estabeleça um conjunto estruturado de controles institucionais para proteção de dados e ativos informacionais, sua abordagem está predominantemente direcionada à implementação de práticas de governança, gestão de riscos, segurança da informação e privacidade no contexto organizacional.

De maneira geral, a análise apresentada na tabela III evidencia que a integração entre o *PPSI*, o *FSF* e os princípios das *SCIFs* estabelece uma arquitetura multicamadas de segurança (*Defense in Depth*), integrando governança, proteção física, segurança cibernética e privacidade de dados. Essa abordagem fortalece o controle de acesso, a rastreabilidade e a gestão de riscos no tratamento seguro de informações sensíveis em ambientes de pesquisa.

Diferentemente de modelos como o *FSF*, que avalia a adequação do acesso e uso dos dados sob uma perspectiva baseada em princípios, e do conceito SCIF, que enfatiza a proteção física e arquitetural de ambientes altamente restritos, o PPSI apresenta uma abordagem operacional baseada em controles, processos e responsabilidades institucionais. Entretanto, sua aplicação isolada pode não contemplar integralmente aspectos específicos relacionados à avaliação contextual do uso dos dados e à construção de ambientes fisicamente segregados para processamento de informações sensíveis.

TABLE III
INTEGRAÇÃO ENTRE *Five Safes Framework*, SCIF E OS CONTROLES DO PPSI 2.0.

Five Safes	**Questão associada**	**Controles PPSI 2.0 relacionados**	**Alinhamento conceitual**
Safe Projects	O uso dos dados é apropriado, legal, ético e justificável?	Controle 0: Estruturação básica para governança; Controle 19: Registro das operações de tratamento; Controle 23: Análise das operações de tratamento; Controle 25: Princípios da LGPD	Relaciona-se à governança dos dados, definição de finalidade, conformidade regulatória, avaliação das operações de tratamento e legitimidade do uso das informações.
Safe People	Os usuários podem ser considerados confiáveis para utilizar os dados adequadamente?	Controle 5: Gestão de contas; Controle 6: Gestão de acesso; Controle 14: Conscientização e treinamento; Controle 21: Encarregado e direitos dos titulares	Relaciona-se à identificação, autenticação, autorização, capacitação e responsabilização dos usuários envolvidos no tratamento dos dados.
Safe Data	Existe risco de divulgação indevida nos próprios dados?	Controle 3: Proteção de dados; Controle 19: Registro das operações de tratamento; Controle 20: Ações de prevenção; Controle 25: Princípios da LGPD	Relaciona-se à proteção dos dados sensíveis, minimização de riscos, rastreabilidade das operações e redução de possibilidades de exposição ou reidentificação.
Safe Settings	O ambiente limita acessos não autorizados ou erros operacionais?	Controle 4: Configuração segura; Controle 6: Gestão de acesso; Controle 8: Gestão de registros de auditoria; Controle 12: Gestão de infraestrutura de rede; Controle 13: Monitoramento e defesa de rede; Controle 18: Testes de intrusão	Relaciona-se diretamente aos princípios arquiteturais inspirados em SCIF, envolvendo segregação de ambientes, proteção da infraestrutura, monitoramento contínuo, auditoria e controle rigoroso de acesso.
Safe Outputs	Os resultados produzidos evitam divulgação indevida?	Controle 8: Gestão de registros de auditoria; Controle 16: Segurança de aplicações; Controle 24: Compartilhamento e transferência internacional; Controle 25: Princípios da LGPD	Relaciona-se à avaliação dos resultados gerados, controle de compartilhamento das informações e prevenção da divulgação indireta de dados sensíveis.

Por outro lado, a adoção do PPSI apresenta vantagens significativas ao fornecer mecanismos práticos para operacionalização dos requisitos de privacidade e segurança da informação, incluindo gestão de ativos, proteção de dados, controle de acesso, registros de auditoria, monitoramento, gestão de incidentes, conscientização e conformidade regulatória. No contexto de ambientes de pesquisa em saúde, esses controles contribuem para transformar princípios abstratos de proteção em procedimentos institucionais verificáveis e mensuráveis. Dessa forma, enquanto o *FSF* contribui para avaliar se determinado acesso aos dados é seguro e justificável, e o conceito *SCIF* fornece referências para construção de ambientes protegidos, o PPSI atua como mecanismo de implementação e sustentação contínua dos controles necessários para garantir a segurança, privacidade e governança dos dados sensíveis.

Vale salientar que a implementação da arquitetura de proteção de dados está associada a três ambientes especificos: a *sala segura*, o *datacenter* e o *ambiente corporativo*. Todos esses ambientes são submetidos a mecanismos contínuos de controle e monitoramento, que incluem sistemas de videomonitoramento, controle de acesso , autenticação de usuários, registros de atividades e procedimentos operacionais de segurança. Cada ambiente possui requisitos específicos de proteção, contribuindo de forma integrada para garantir a conformidade da traíde CIA e rastreabilidade das informações sensíveis ao longo de todo o ciclo de vida dos dados e ainda relacionando os ambientes, todo acesso é controlado por mecanismos físicos e lógicos, incluindo controle de entrada, videomonitoramento, autenticação, registros de acesso e procedimentos específicos para autorização e acompanhamento das atividades realizadas. Essa segmentação permite aplicar diferentes camadas de proteção, garantindo que o acesso às informações sensíveis ocorra de forma restrita, rastreável e compatível com os princípios mencionados neste artigo.Como exemplo, o acesso ao *datacenter* é realizado mediante um fluxo formal de autorização, no qual a solicitação deve ser previamente registrada 72 horas antes por meio de chamado institucional, contendo a justificativa, o responsável e a finalidade do acesso. A atividade é acompanhada por um membro do NSI, garantindo acessibilidade,o cumprimento dos procedimentos estabelecidos e a rastreabilidade das ações realizadas no ambiente. Esse controle reforça os princípios de segregação de funções, menor privilégio (*Least Privilege*) e *Need-to-Know*, e reduzindo riscos associados a acessos indevidos ou atividades não autorizadas em ambientes críticos. [37] Além do acompanhamento operacional, todas as atividades executadas no ambiente são registradas por meio de mecanismos de auditoria, permitindo a identificação do usuário, do horário de acesso, das operações realizadas e dos recursos utilizados. Esses registros possibilitam a realização de auditorias periódicas, investigações de incidentes de segurança e a verificação da conformidade com as políticas institucionais e requisitos regulatórios. Dessa forma, o monitoramento contínuo e de auditoria, promove transparência, responsabilização e capacidade de resposta de eventos de segurança ou desvios de conformidade.

VI. Considerações Finais

Este estudo demonstrou que a integração proposta contribui para uma visão sistêmica da proteção de dados sensíveis em ambientes de pesquisa, permitindo alinhar requisitos de governança, segurança e confiabilidade ao longo dos processos institucionais. Os resultados evidenciam que a adoção de mecanismos complementares favorece uma abordagem equilibrada entre a necessidade de utilização dos dados para geração de conhecimento científico e a preservação dos direitos dos titulares, estabelecendo uma base estruturada para ambientes de pesquisa mais seguros, auditáveis e resilientes.

No contexto do Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS), a integração do FSF, SCIF e PPSI possibilita equilibrar o compartilhamento de dados para fins científicos com elevados níveis de privacidade, confidencialidade e conformidade regulatória. A adoção de uma arquitetura multicamadas baseada no conceito de *Defense in Depth* fortalece o controle de acesso, reduz riscos de reidentificação e vazamento de informações, amplia a rastreabilidade das operações e aumenta a confiança institucional na utilização de dados para pesquisa.

Além disso, a integração entre esses modelos evidenciam que princípios de proteção originalmente aplicados a ambientes de informações críticas podem ser contextualizados para cenários de pesquisa em saúde, nos quais grandes volumes de dados sensíveis demandam elevados níveis de confiabilidade e proteção. Essa abordagem contribui para elevar a maturidade institucional em segurança da informação, fortalecendo processos de governança, gestão de riscos, controle de acesso e preservação da confiabilidade dos dados utilizados na produção científica.

Como trabalhos futuros, propõe-se ampliar as salvaguardas associadas ao uso de Inteligência Artificial (IA) em ambientes de pesquisa com dados sensíveis, considerando os novos desafios relacionados à privacidade, explicabilidade, governança dos modelos e possíveis riscos de reidentificação. Além disso, recomenda-se a incorporação de mecanismos avançados de inteligência de ameaças (*Threat Intelligence*), capazes de apoiar a identificação preventiva de vulnerabilidades, comportamentos anômalos e potenciais vetores de ataque direcionados aos ambientes de processamento de dados.

Também se pretende investigar a integração de técnicas emergentes, como automação de respostas a incidentes, análise comportamental de usuários, modelos de detecção baseados em aprendizado de máquina e abordagens de segurança adaptativa, com o objetivo de aprimorar a capacidade de monitoramento e resposta institucional. Essas evoluções permitirão fortalecer a arquitetura proposta, mantendo o equilíbrio entre acesso científico, inovação tecnológica e proteção dos direitos dos titulares dos dados.

Acknowledgment

Os autores agradecem à Universidade de Brasilia e ao LabRisk(Laboratorio de riscos, privacidade e segurança cibernética) e ao Centro de Integração de Dados e Conhecimentos para Saúde (Cidacs/Fiocruz Bahia)

References

[1] V. Khatri and C. V. Brown, "Designing Data Governance," *Communications of the ACM*, vol. 53, no. 1, pp. 148–152, 2010, doi: 10.1145/1629175.1629210.

[2] World Health Organization, "Ethics and Governance of Artificial Intelligence for Health," World Health Organization, Geneva, Switzerland, 2021.

[3] CIDACS. *Centro de Integração de Dados e Conhecimentos para Saúde*. Disponível em: https://cidacs.bahia.fiocruz.br/. Acesso em: 20 abr. 2026.

[4] G. C. G. Barbosa, M. S. Ali, B. Araújo, S. Reis, S. Sena, M. Y. Ichihara, J. M. Pescarini, R. L. Fiaccone, L. D. Amorim, R. Pita, M. L. Barreto, et al., "CIDACS-RL: A Novel Indexing Search and Scoring-Based Record Linkage System for Huge Datasets with High Accuracy and Scalability," *BMC Medical Informatics and Decision Making*, vol. 20, no. 1, p. 289, 2020. doi: 10.1186/s12911-020-01285-w.

[5] UK Data Service, "Five Safes: Designing Data Access Systems," UK Data Service, University of Essex, 2016. Available: https://ukdataservice.ac.uk/help/secure-lab/five-safes/

[6] Gabinete de Segurança Institucional da Presidência da República (GSI/PR), "Programa de Privacidade e Segurança da Informação (PPSI)," Brasília, DF, Brasil, 2023. Available: https://www.gov.br/gsi/pt-br/assuntos/seguranca-da-informacao

[7] COMMITTEE ON NATIONAL SECURITY SYSTEMS (CNSS). Intelligence Community Standard (ICS) 700-1: Facilities for Sensitive Compartmented Information. Fort Meade, MD: CNSS, 2021.

[8] DAMA INTERNATIONAL. *The DAMA Guide to the Data Management Body of Knowledge (DAMA-DMBOK2)*. 2. ed. Basking Ridge, NJ: Technics Publications, 2017.

[9] KHATRI, V.; BROWN, C. V. Designing data governance. *Communications of the ACM*, v. 53, n. 1, p. 148–152, 2010.

[10] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). *ISO/IEC 38500:2015 – Information technology – Governance of IT for the organization*. Geneva: ISO, 2015.

[11] ISACA. *COBIT 2019 Framework: Governance and Management Objectives*. Schaumburg, IL: ISACA, 2019.

[12] CAVOUKIAN, A. *Privacy by Design: The 7 Foundational Principles*. Toronto: Information and Privacy Commissioner of Ontario, 2009.

[13] SOLOVE, D. J. A taxonomy of privacy. *University of Pennsylvania Law Review*, v. 154, n. 3, p. 477–560, 2006.

[14] BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Diário Oficial da União, Brasília, DF, 15 ago. 2018.

[15] UNIÃO EUROPEIA. Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho, de 27 de abril de 2016. Regulamento Geral sobre a Proteção de Dados (General Data Protection Regulation – GDPR). Jornal Oficial da União Europeia, L119, p. 1–88, 2016.

[16] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). *ISO/IEC 27701:2019 – Security techniques – Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management*. Geneva: ISO, 2019.

[17] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST). *NIST Privacy Framework: A Tool for Improving Privacy through Enterprise Risk Management*. Version 1.0. Gaithersburg, MD: NIST, 2020.

[18] Center for Internet Security (CIS), "CIS Critical Security Controls Version 8," Center for Internet Security, East Greenbush, NY, USA, 2021. Available: https://www.cisecurity.org/controls

[19] WHITMAN, M. E.; MATTORD, H. J. *Principles of Information Security*. 7. ed. Boston: Cengage Learning, 2021.

[20] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). *ISO/IEC 27001:2022 – Information security, cybersecurity and privacy protection – Information security management systems – Requirements*. Geneva: ISO, 2022.

[21] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). *ISO/IEC 27002:2022 – Information security, cybersecurity and privacy protection – Information security controls*. Geneva: ISO, 2022.

[22] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST). *Cybersecurity Framework (CSF) 2.0*. Gaithersburg, MD: NIST, 2024.

[23] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST). *Security and Privacy Controls for Information Systems and Organizations*. NIST Special Publication 800-53 Revision 5. Gaithersburg, MD: NIST, 2020.

[24] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST). *Risk Management Framework for Information Systems and Organizations*. NIST Special Publication 800-37 Revision 2. Gaithersburg, MD: NIST, 2018.

[25] National Institute of Standards and Technology (NIST), "Cybersecurity Framework (CSF) 2.0," Gaithersburg, MD, USA: National Institute of Standards and Technology, 2024. doi: 10.6028/NIST.CSWP.29.

[26] Desai, T.; Ritchie, F.; Welpton, R. *The Five Safes: Designing Data Access for Research*. Economics Working Paper Series, University of the West of England, 2016.

[27] OFFICE OF THE DIRECTOR OF NATIONAL INTELLIGENCE (ODNI). *Intelligence Community Directive (ICD) 705: Sensitive Compartmented Information Facilities*. Washington, DC: Office of the Director of National Intelligence, 2010.

[28] OFFICE OF THE DIRECTOR OF NATIONAL INTELLIGENCE (ODNI). *Intelligence Community Standard (ICS) 705-01: Physical and Technical Security Standards for Sensitive Compartmented Information Facilities*. Washington, DC: Office of the Director of National Intelligence, 2010.

[29] U.S. Department of Health and Human Services. *HIPAA Security Rule*. Health Information Privacy. Disponível em: https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/index.html. Acesso em: 26 jul. 2026.

[30] CENTRO DE INTEGRAÇÃO DE DADOS E CONHECIMENTOS PARA SAÚDE (CIDACS). *Quem Somos*. Salvador: Fundação Oswaldo Cruz (Fiocruz Bahia). Disponível em: https://cidacs.bahia.fiocruz.br/sobre/quem-somos/. Acesso em: 26 jul. 2026.

[31] J. M. Pescarini, E. Williamson, R. L. Fiaccone, L. D. Amorim, M. Y. Ichihara, E. S. Paixão, A. L. S. Barbosa, R. Pita, M. L. Barreto, et al., "The 100 Million Brazilian Cohort: A Resource for Epidemiological and Operational Research," *International Journal of Epidemiology*, vol. 47, no. 4, p. e27, 2018. doi: 10.1093/ije/dyy137.

[32] D. Almeida, D. Gorender, M. Y. Ichihara, S. Sena, L. Menezes, G. C. G. Barbosa, R. L. Fiaccone, E. S. Paixão, R. Pita, and M. L. Barreto, "Examining the Quality of Record Linkage Process Using Nationwide Brazilian Administrative Databases to Build a Large Birth Cohort," *BMC Medical Informatics and Decision Making*, vol. 20, no. 1, p. 173, 2020. doi: 10.1186/s12911-020-01192-0.

[33] CENTRO DE INTEGRAÇÃO DE DADOS E CONHECIMENTOS PARA SAÚDE (CIDACS). *Como Funcionamos*. Salvador: Fundação Oswaldo Cruz (Fiocruz Bahia). Disponível em: https://cidacs.bahia.fiocruz.br/sobre/como_funcionamos/. Acesso em: 26 jul. 2026.

[34] CENTRO DE INTEGRAÇÃO DE DADOS E CONHECIMENTOS PARA SAÚDE (CIDACS). *Como funcionamos*. Disponível em: https://cidacs.bahia.fiocruz.br/sobre/como_funcionamos/. Acesso em: 26 jul. 2026.

[35] CENTRO DE INTEGRAÇÃO DE DADOS E CONHECIMENTOS PARA SAÚDE (CIDACS). *O CIDACS e a LGPD*. Disponível em: https://cidacs.bahia.fiocruz.br/o-cidacs-e-a-lgpd/. Acesso em: 26 jul. 2026.

[36] SOCIEDADE BRASILEIRA PARA O PROGRESSO DA CIÊNCIA (SBPC). *Brasil que Queremos: Ciência, Democracia e Soberania*. Observatório Eleições 2026, São Paulo: SBPC, 2026. Disponível em: https://portal.sbpcnet.org.br/observatorio-eleicoes2026/assets/pdf/cadernos_2026.pdf. Acesso em: 24 jul. 2026. p. 61.

[37] Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS), "Datacenter", Fundação Oswaldo Cruz (Fiocruz Bahia), 2026. [Online]. Available: https://cidacs.bahia.fiocruz.br/datacenter/. Accessed: Jul. 26, 2026.