



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz



CURSO DE ESPECIALIZAÇÃO EM INFORMAÇÃO CIENTÍFICA E TECNOLÓGICA EM SAÚDE

DESENVOLVIMENTO DE UMA PLATAFORMA DE GESTÃO DE DADOS EM SAÚDE A PARTIR DE PRONTUÁRIOS ELETRÔNICOS

por

MARIANA CECI DE FRANÇA E SILVA

Trabalho de Conclusão de Curso apresentado ao Instituto de Comunicação e Informação Científica e Tecnológica em Saúde da Fundação Oswaldo Cruz.

Orientador: Dr. Jânio Gustavo Barbosa

Rio de Janeiro, Novembro de 2022

RESUMO

O Instituto Santos Dumont (ISD) é uma Organização Social de nível federal que possui duas unidades na zona rural da cidade de Macaíba, no Rio Grande do Norte. No ano de 2021, 52.531 pessoas foram atendidas no ISD, nas áreas de saúde materno-infantil e saúde da pessoa com deficiência. Isso gera uma série de registros que precisam estar disponíveis para as equipes multiprofissionais, tanto para a atuação clínica como para a resolução de problemas, pesquisa e gestão. Apesar de já adotar o sistema de Prontuário Eletrônico do Paciente (PEP), o ISD ainda não dispõe de um fluxo visível para reuso de dados gerados nos atendimentos. O objetivo deste trabalho é desenvolver um plano de mapeamento dos fluxos de pesquisa e *open data* da instituição para embasar a criação de um sistema de gestão de dados abertos, que será formulado a partir de um conjunto de *scripts*, para que esses dados possam ser disponibilizados de forma anônima, respeitando o alto grau de sensibilidade inerente aos dados em saúde.

Palavras-chave: Dados abertos em saúde; Prontuário Eletrônico do Paciente (PEP); Infraestrutura da informação; reuso de dados

ABSTRACT

The Santos Dumont Institute (ISD, as Portuguese abbreviation) is a social organization at the federal level that has two units located in the rural area of the city of Macaíba, in Rio Grande do Norte. In 2021, 52,531 people were assisted at the ISD, in the areas of maternal and child health and the health of people with disabilities. This generates a series of records that need to be available to multidisciplinary teams, both for clinical practice and for problem solving, research and management. Despite already adopting the Electronic Patient Record (EPR) system, ISD still does not have a visible flow for the reuse of data generated in the consultations. The objective of this work is to develop a plan to map the institution's research flows and open data to support the creation of an open data management system, which will be formulated from a set of scripts, so that these data can be made available anonymously, respecting the high degree of sensitivity inherent in health data.

Keywords: Open health data; Electronic Patient Record (EPR); Information infrastructure; data reuse.

SUMÁRIO

1. INTRODUÇÃO	5
2. JUSTIFICATIVA	10
3. REFERENCIAL TEÓRICO	14
3.1. Dado, informação e conhecimento	14
3.2. Prontuário Eletrônico do Paciente (PEP) e a Lei Geral de Proteção de Dados	14
3.3. Anonimização dos dados e compartilhamento	15
3.4 Política de dados abertos	16
4. OBJETIVOS	18
4.1. OBJETIVO GERAL	18
4.2. OBJETIVOS ESPECÍFICOS	18
5. METODOLOGIA	19
6. RESULTADOS ESPERADOS	21
7. REFERÊNCIAS CONSULTADAS	22
8. CRONOGRAMA	24
9. ORÇAMENTO	25

1. INTRODUÇÃO

O Instituto Santos Dumont (ISD) é uma Organização Social de nível federal vinculada ao Ministério da Educação, localizada na zona rural do município de Macaíba, na região metropolitana de Natal, capital do Rio Grande do Norte. O ISD funciona a partir de duas unidades: o Centro de Educação e Pesquisa em Saúde Anita Garibaldi (Anita) e o Instituto Internacional de Neurociências Edmond e Lily Safra (IIN-ELS), nas quais são atendidos usuários do Sistema Único de Saúde (SUS) da 7ª Região de Saúde do RN nas áreas de saúde materno-infantil e saúde da pessoa com deficiência, e onde estão situados os programas de pós-graduação de Residência Multiprofissional no Cuidado à Saúde da Pessoa com Deficiência e Mestrado em Neuroengenharia, ambos pioneiros em suas respectivas áreas no Brasil.

Além disso, o ISD é habilitado desde 2021 como Centro Especializado em Reabilitação de nível 4 pelo Ministério da Saúde, e opera 8 linhas de cuidado especializadas que assistem em sua maioria à população da 7ª Região de Saúde, a mais populosa do RN, que engloba os municípios de Natal, Parnamirim, Macaíba, Extremoz e São Gonçalo do Amarante. Essa assistência se estende para todo o RN em algumas linhas de cuidado específicas, como a da epilepsia farmacorresistente, ou seja, aquela de difícil controle por medicamentos, e a de medicina fetal ambulatorial para anomalias fetais e infecções de transmissão vertical, aquelas que podem passar da mãe para o bebê durante a gestação.

A partir do tripé ensino, pesquisa e extensão, a instituição funciona como um grande centro de educação para as profissões da saúde e engenharias, onde as pesquisas desenvolvidas estão diretamente vinculadas às principais demandas dos usuários, e por onde passam estagiários de graduação, residentes, pesquisadores de programas de pós-graduação e profissionais em atividades de educação permanente em saúde.

A instituição recebe estagiários de graduação oriundos dos cursos de Fisioterapia, Medicina, Pedagogia, Psicologia e Serviço Social dos diversos campi da Universidade Federal do Rio Grande do Norte (UFRN); residentes externos vinculados ao Hospital Infantil Varela Santiago, em Natal; Escola Multicampi de

Ciências Médicas (EMCM/UFRN), em Caicó; Hospital Universitário Onofre Lopes (HUOL/UFRN); Hospital Universitário Ana Bezerra (HUAB/UFRN), em Santa Cruz; Maternidade Escola Januário Cicco (MEJC/UFRN) também em Natal; Universidade Federal de São Paulo (Unifesp/SP) e Hospital Geral Roberto Santos (Salvador/BA).

Além disso, possui projetos de pesquisa multicêntricos em parceria com instituições como o Instituto Nacional da Mulher, Criança e Adolescente Fernandes Figueira, da Fundação Oswaldo Cruz (IFF/Fiocruz) e colaborações com pesquisadores de instituições nacionais e internacionais, como a École Polytechnique de Lausanne (EPFL), da Suíça. Essas parcerias resultam em um fluxo constante de pesquisadores em trânsito entre o ISD e suas respectivas instituições, compartilhando dados e realizando colaborações científicas para o desenvolvimento de novos protocolos de tratamento, novos desenvolvimentos tecnológicos e dispositivos de saúde voltados para as principais necessidades dos usuários atendidos pelo ISD.

Apenas no ano de 2021, o ISD atendeu 52.531 usuários, que são acompanhados por equipes multiprofissionais e estudantes vinculados a programas de graduação e pós-graduação. Esses usuários têm seus dados registrados no sistema de Prontuário Eletrônico do Paciente (PEP), a partir do *software* TOTVS.

O prontuário é um documento único, que reúne um conjunto de registros, sinais e imagens registradas sobre a saúde do paciente e a assistência prestada a ele. Ele é o documento que permite a comunicação entre membros de uma equipe multiprofissional. O PEP, em contraste com o sistema não informatizado, permite que mais dados sejam armazenadas por mais tempo e, quando processados, os dados existentes nesses prontuários permitem a identificação de tendências sanitárias em determinadas populações, informação de grande utilidade não apenas para os profissionais da saúde e pesquisadores, mas também para os gestores da saúde (GONÇALVES, J.P.P *et al*, 2013).

Apesar do grande volume de dados gerados a partir desses pacientes e de já adotar um sistema eletrônico, o ISD ainda não dispõe de um fluxo visível para reuso dos dados de atendimento, principalmente para as pesquisas científicas desenvolvidas na instituição.

A importância e utilização de dados para a tomada de decisões e análise de situações não é exclusividade de nossos tempos, sejam elas, em pesquisas públicas, no mercado privado ou em ações governamentais. Nas últimas décadas, no entanto, graças ao aumento da disponibilidade de dados possibilitada pelo meio digital e sua difusão, a sociedade experimenta um fenômeno que vai da escassez à extrema abundância desses dados (Sayão e Lopes, 2020).

Por dado, entendemos uma sequência de símbolos quantificados ou quantificáveis, que por consequência podem ser armazenados em um computador, processados por ele e que podem ou não se transformar em informação (Setzer, 2001). Os novos sistemas de computadores e redes possibilitam a integração e exploração de conjuntos massivos de dados, que revelam relações e padrões que muitas vezes passariam despercebidos ou necessitam de um extenso tempo de trabalho para serem identificados e transformados em informação e conhecimento.

Um exemplo concreto da possibilidade de utilização de dados para traçar estratégias na área de saúde está no Projeto Barriguda, criado em 2015 pela instituição, em parceria com a comunidade quilombola de Capoeiras, a maior e mais antiga do Rio Grande do Norte.

Até o ano de 2014, a eclâmpsia, uma condição que pode provocar convulsões durante a gravidez, atingia em média 20% das mulheres da comunidade de Capoeiras, provocando um coeficiente de mortalidade materna de 223 óbitos para cada 100 mil nascidos vivos. A partir do mapeamento dessa realidade a partir dos dados de saúde e dos relatos das mulheres da comunidade quilombola de Capoeiras, o ISD foi capaz de esquematizar e colocar em prática o Projeto Barriguda, que leva à comunidade o pré-natal e o acompanhamento profissional, e que foi capaz de zerar a quantidade de mortes maternas por eclâmpsia na localidade.

O projeto, que em 2017 venceu o prêmio de Laboratório de Inovação sobre a Participação Social na Atenção Integral à Saúde da Mulher, concedido pela Organização Pan-americana de Saúde (OPAS/OMS), mostra como a utilização de dados pode permitir avaliar de forma mais minuciosa os problemas locais e buscar estratégias e soluções que possam efetivamente solucioná-los. Por analisar um grupo específico (mulheres grávidas) em uma comunidade de pouco

mais de 1.300 habitantes, foi possível fazer o levantamento manual desses registros e suas principais correlações.

Entretanto, no campo da ciência, a nova realidade da curva de disponibilidade de dados reverbera na criação de novas metodologias, que impulsionam o progresso das pesquisas. Em um mundo onde a disponibilidade de dados é crescente, localizar, acessar e manter esses dados muitas vezes não padronizados é um desafio. Para poderem ser utilizados, os dados precisam seguir uma série de critérios fundamentais: devem ser acessíveis, inteligíveis, avaliáveis e utilizáveis pelas partes interessadas.

"Acessíveis", pois entende-se que os dados devem estar localizados de tal forma que podem ser rapidamente encontrados de uma forma passível de uso. "Inteligíveis" e "avaliáveis", porque precisam ser apresentadas de forma abrangente para aqueles que desejam se debruçar sobre eles para fazer um julgamento ou avaliação daquilo que foi comunicado a partir deles, incluindo dados concomitantes que possam influenciar na confiança do público. "Utilizáveis" diante da necessidade de serem apresentados em um formato que o dado possa ser reutilizado para diferentes propósitos (The Royal Society, 2012).

No caso das pesquisas em saúde, por exemplo, a disponibilização desses conteúdos de forma padronizada pode permitir diversas análises em grupos amostrais amplos, que demandam uma grande quantidade de tempo e recursos humanos para serem feitos de forma manual.

A utilização desses dados de pesquisa em saúde tem como contrapartida exigências rigorosas em relação à sua gestão, que precisa considerar o alto grau de sensibilidade dos registros, respeitando o anonimato dos pacientes, seus familiares e afins, de acordo com a resolução 1.632/2002 do Conselho Federal de Medicina, que aprovou a utilização dos PEPs e dispõe sobre as regras de controle de uso.

O presente projeto nasceu a partir da observação da possibilidade de criação de um sistema como esses para o Instituto Santos Dumont (ISD), instituição de ensino e pesquisa em que o atendimento aos usuários do SUS está vinculado às práticas de ensino, pesquisa e extensão, e onde o uso do Prontuário Eletrônico do Paciente (PEP) já é uma realidade, porém ainda sem fluxo definido para acesso aberto de pesquisadores e profissionais.

O projeto tem como objetivo criar um sistema de gestão de dados a partir de prontuários eletrônicos, que considere os aspectos citados: acessibilidade, inteligibilidade, possibilidade de avaliação e utilização pelas partes, no caso, pesquisadores.

Atualmente, não há um sistema fixo para compartilhamento de dados entre os pesquisadores, que precisam entrar em contato com os membros da equipe multiprofissional de cada linha de cuidado a fim de obter os dados necessários para o desenvolvimento dos estudos, o que muitas vezes é prejudicado pelas diferentes estruturas de cada linha de cuidado (algumas que se expandem para todo RN, outras que estão focadas na 7ª Região de Saúde, por exemplo, o que impacta no volume de registros de cada uma).

Isso resulta na necessidade de novas coletas e buscas de registros que já haviam sido previamente coletados, mas que não podem ser facilmente resgatados, acessados ou compartilhados.

A intenção aqui é compreender como funcionam os fluxos de pesquisa a fim de permitir a disponibilização dos dados em uma plataforma a partir de métricas pré-definidas, como por exemplo, números de atendimentos, procedimentos clínicos, aspectos sociais, epidemiológicos e uso dos recursos de saúde disponíveis. Espera-se que a criação de um sistema de gestão de dados abertos, com critérios bem definidos para permitir o acesso, amplie as possibilidades de análise para pesquisas e estudos iniciados na instituição, além de favorecer os estudos multicêntricos, a revisão de estudos científicos e ampliar as possibilidades de colaboração entre profissionais de diferentes campos que estejam pesquisando a respeito dos temas abarcados pelas linhas de atuação do ISD.

2. JUSTIFICATIVA

O compartilhamento de dados é uma etapa essencial no subsídio de pesquisas, elaboração e adoção de políticas públicas na área de saúde. (Mota e Babêto, 2004). Com o surgimento de plataformas computacionais para análise de grande volumes de dados, ficou cada vez mais comum ver aplicações de métodos preditivos em dados de saúde. O objetivo dessas análises, segundo Mota e Babêto (2004, p.3) é fornecer melhorias para o próprio sistema de saúde, bem como propor tratamentos que atendam às características individuais do paciente.

A adoção do PEP, regulamentada em 2002, traz diversas vantagens, como:

- Acesso rápido;
- Disponibilidade remota;
- Uso simultâneo por várias pessoas;
- Legibilidade absoluta (não se perderá tempo decifrando os "hieróglifos" dos médicos, enfermeiros e atendentes);
- Redução do espaço de armazenamento;
- Maior confiabilidade;
- Segurança e confidencialidade dos dados por meio da utilização de senhas digitais;
- Extinção das pilhas de papéis passíveis de deterioração e consequente inutilização dos dados diante de más condições de acondicionamento.

Nesse sentido, o PEP, representante maior da chamada documentação sanitária, se configura como uma importante fonte de informação em saúde que pode ser utilizada nessas análises e previsões, já compartilhadas entre os membros da equipe multiprofissional.

Importante salientar que o prontuário do paciente é um documento composto de vários tipos de textos: descritivos, narrativos, dissertativos, redigidos por profissionais de várias áreas da saúde e, por vezes, diferentes organizações (Miranda e Bentes Pinto, 2005). Esse é um dos motivos pelos quais, na maior parte das vezes, os dados contidos nesses documentos se apresentam de forma não estruturada e são armazenados de maneira difusa, com uma grande

variedade terminológica não necessariamente padronizada dificultando assim a recuperação e identificação de informação no contexto dos dados em saúde.

Por se tratar de uma Organização Social, o ISD tem o compromisso de prestar contas à entidade pública a qual está vinculado, no caso, o Ministério da Educação, cumprindo uma série de metas pactuadas entre as instituições, que englobam atividades de pesquisa, ensino e extensão. Essa avaliação é feita semestralmente, com as metas mensuradas a partir de indicadores, que avaliam aspectos como quantidade de publicações por membro do corpo docente e discente, horas cumpridas para estágios de graduação e residência médica e multiprofissional, e ações de educação permanente em saúde.

Com o objetivo de facilitar a aquisição dos dados necessários para elaborar o relatório semestral, o Instituto buscou criar um sistema de gestão capaz de gerar visualizações de dados referentes a esses indicadores, como o número de publicações feitas por professor pesquisador, o total de estudantes de graduação e pós-graduação que passaram pela instituição e o total de procedimentos e atendimentos realizados.

Esse sistema, no entanto, se restringe aos indicadores pactuados junto ao Ministério da Educação, e seu acesso está restrito aos gestores e responsáveis pela elaboração do documento de prestação de contas, sem acesso à registros por linha de cuidado que traga registros passíveis de utilização em pesquisas na área de saúde.

A criação e a manutenção desse tipo de base de dados, necessita de conhecimento especializado e muitas vezes uma equipe de tecnologia da informação dedicada. Nesse sentido, a proposta desse projeto é fornecer as diretrizes básicas para mapear os fluxos de pesquisa da instituição, permitindo disponibilizar uma fluxo que favoreça a criação de um sistema mínimo que se conecte a um banco de dados e disponibilize um painel de dados com medidas básicas para acompanhamento das métricas de atendimento, para compartilhamento responsável e anonimizado de dados em saúde, bem como, acompanhamento do panorama de desenvolvimento das ações realizadas pelo ISD.

Por estar na região mais populosa do estado do Rio Grande do Norte, e organizar suas linhas de atuação com o objetivo de cobrir vazios assistenciais existentes no estado, os dados coletados a partir do PEP no ISD se mostram

relevantes para a condução de pesquisas, estudos e predições que podem ter impacto não apenas na criação de novos protocolos clínicos e dispositivos tecnológicos, mas na própria gestão do SUS.

Um exemplo disso são os estudos conduzidos em parceria com o IFF/Fiocruz, voltados à investigação das condições de bexiga e intestino neurogênicos em crianças com Microcefalia causada em decorrência do vírus Zika, que permitiu identificar condições passíveis de reabilitação nas crianças acompanhadas pelo Instituto em sua linha de cuidado especializada, com impacto direto na qualidade de vida das famílias. Essa pesquisa faz parte de um estudo multicêntrico, conduzido também em Pernambuco e na Bahia, e é um exemplo da potencialidade do compartilhamento de dados na área de saúde.

Nesse sentido, o sistema de gestão de dados é uma tecnologia que permite criar efeitos de sinergia, como apontam Mota e Babêto (2005), reduzindo custos e erros de sistema, otimizando recursos, proporcionando treinamento contínuo por toda a cadeia de compartilhamento de dado, permitindo também a geração de predições epidemiológicas que podem ter impacto direto em estudos científicos e na gestão em saúde.

A criação desse sistema não pode acontecer de forma aleatória: segundo a resolução 1.639/2002 do Conselho Federal de Medicina (CONSELHO FEDERAL DE MEDICINA, 2002), o uso de sistemas informatizados para guarda e manuseio de PEP devem preencher requisitos de integridade da informação e qualidade do serviço; cópia de segurança; banco de dados; privacidade e confiabilidade; autenticação e auditoria, que devem andar lado a lado dos princípios de acessibilidade, ininteligibilidade, possibilidade de avaliação e utilização pelas partes apontados pela Royal Society como essenciais para sua eficácia.

Atualmente, o compartilhamento de dados entre pesquisadores internos e externos se dá a partir da assinatura de um convênio formal entre o ISD e a instituição. A aprovação e condução de novos estudos também deve ser aprovada pelo Comitê de Ética em Pesquisa (CEP) da instituição.

O zelo pelos aspectos éticos envolvendo o uso dos PEPs em pesquisas científicas deverá se manter no sistema de gestão de dados abertos, com a necessidade de aprovação da solicitação, análise da justificativa para acesso aos dados e diferenciação de níveis de acesso por tipo de pesquisador ou estudo a ser desenvolvido.

A principal inovação desta plataforma, além de condensar em um só espaço, de forma padronizada, os diferentes dados, protocolos e regras de acesso, está principalmente na possibilidade de reuso dos dados já existentes para novas análises e estudos. No caso de pesquisas relacionadas à bexiga e intestino neurogênicos, por exemplo, registros referentes ao exame de urodinâmica podem ser importantes para agregar em estudos multicêntricos que tratem desta temática.

Além disso, a criação do sistema deverá eliminar dois problemas atuais, que são a necessidade de contactar diretamente os responsáveis por cada linha de cuidado para fazer a busca manual dos dados que se pretende utilizar em determinada pesquisa científica, o que demanda um tempo considerável dos profissionais e, muitas vezes, resulta no atraso da execução de estudos científicos, além da impossibilidade de fornecer determinadas informações diante da falta de clareza sobre quais dados são passíveis de fornecimento ou não para cada solicitante. Com um fluxo bem definido, essas dúvidas seriam facilmente respondidas.

3. REFERENCIAL TEÓRICO

Para o desenvolvimento deste trabalho, será importante discutir algumas dimensões conceituais importantes, que serão necessárias para sua execução. É o caso dos conceitos de dado, Prontuário Eletrônico do Paciente, anonimização de dados em saúde e a Lei Geral de Proteção de Dados. Esses conceitos são relevantes para o projeto na medida em que demonstram as faces necessárias para a composição de fluxos e protocolos que podem se tornar modelos de funcionalidade e construção de bancos de dados, abordando aspectos como modelagem, funcionalidade e aspectos legais que devem ser considerados na construção de um sistema de pesquisa que envolva dados de prontuários eletrônicos.

● 3.1. Dado, informação e conhecimento

A primeira diferenciação é entre dado, informação e conhecimento. Entendemos o dado como o elemento bruto quantificável, desprovido de significação imediata, que podem ser armazenados e processados (Semidão, 2013). Quantificável, por sua vez, significa algo que pode ser contado de forma matemática ou estatística e posteriormente reproduzido. Ele é, necessariamente, puramente sintático, o que significa que podem ser descritos através de representações formais e estruturais, além de serem passíveis de armazenamento e processamento em um computador. (Setzer, 2001).

Se o dado é visto como o "átomo" ou elemento bruto, a informação remete ao processo relacionado ao conhecimento e à comunicação. Trata-se de uma abstração informal que, ainda segundo Setzer (2001), não pode ser formalizada ou quantificada através de uma teoria lógica ou matemática. A informação pode ser uma propriedade interior de um indivíduo, ou recebida por ele e, portanto, não pode ser processada por um computador. Enquanto o dado é puramente sintático, a informação possui uma carga semântica.

As palavras "sífilis", "mulher" e "grávida", por exemplo, são dados. Ao organizar essas palavras ("A mulher grávida tem sífilis") e contextualizá-las dentro de um serviço de saúde materno-infantil, esses dados podem assumir uma carga

de significado para o profissional da saúde, que vai encaminhar a mulher grávida que apresenta diagnóstico para sífilis para o tratamento adequado.

O conhecimento, por sua vez, diz respeito a uma abstração interior e pessoal, localizado puramente no âmbito subjetivo e descrito somente a partir da informação ou dado. Esses conceitos são relevantes na medida que este estudo se baseia na perspectiva da Ciência da Informação, especificamente, a gestão da informação e do conhecimento, um dos núcleos teóricos desta área, "que pode ser definido como o tipo de abordagem de problemas de informação e conhecimento que atua desde uma perspectiva de aperfeiçoamento de cabedais de informação e conhecimento para fins organizacionais" (Semidão, 2013, p. 7).

3.2. Prontuário Eletrônico do Paciente (PEP) e a Lei Geral de Proteção de Dados

O Prontuário Eletrônico do Paciente (PEP) terá atenção especial neste projeto por se tratar do documento que reúne todas as principais informações a serem organizadas no sistema de gestão de dados. Considerado o representante máximo da documentação sanitária, o prontuário é um documento único que dispõe de um conjunto de registros gerados a partir de fatos, acontecimentos e situações sobre a saúde de um indivíduo, assim como a assistência a ele prestada. (GONÇALVES, J.P.P et al, 2013). Seja em sua forma física ou eletrônica, o prontuário é regido por códigos deontológicos das diversas profissões da saúde, por se tratar do documento mais importante de registro do paciente. (Fernandes, 2021).

No Brasil, as regras acerca da digitalização, guarda, armazenamento e manuseio dos PEP estão dispostas na Lei 13.787 de 27 de dezembro de 2018, vinculada à Lei Geral de Proteção de Dados (LGPD), lei 13.709 de 14 de agosto de 2018 e devem respeitar também as normas da Autoridade Nacional de Proteção de Dados (ANPD), criada pelo decreto 10.474/2020.

Os dados disponíveis no PEP são de natureza individual e sensível, tendo em vista sua relação direta com a saúde e intimidade do paciente. Todas as instituições públicas ou privadas responsáveis pelo tratamento de dados pessoais devem ter suas atividades ajustadas conforme as regras e princípios da LGPD. As

instituições devem prever formas e estruturas seguras para armazenamento e acesso ao PEP, seja pelo paciente ou por outras pessoas autorizadas, salvo registros que possam comprometer ou colocar em situação de vulnerabilidade o paciente, como questões relacionadas à violência, abuso e saúde mental.

Tais sistemas devem ser certificados para garantir a segurança e o sigilo dos dados, além de estabelecer medidas para prevenir acessos não autorizados ou ilícitos que podem provocar a perda, destruição, alteração, comunicação ou qualquer outra forma de tratamento indevido do dado, segundo a LGPD.

3.3. Anonimização dos dados e compartilhamento

A utilização dos registros do PEP para pesquisas se encaixa no princípio da integralidade em saúde do SUS, expresso na Lei 8080/1990, que entende a prática como parte do conjunto articulado e contínuo de ações de caráter preventivo e curativo individuais e coletivos. A LGPD determina, ainda, que quando há solicitação de dados por alguém, a pessoa em questão deve justificar quanto à finalidade da solicitação (CEJAM, O Paciente e a LGPD).

Uma das estratégias que podem ser adotadas para utilização de dados oriundos de prontuários em pesquisas é a anonimização, que consiste na retirada de aspectos identificadores do paciente. A anonimização consiste na remoção irreversível de registros que liguem o indivíduo e seus registros médicos, a um grau em que seria impossível restabelecer a conexão entre um e outro (Kushida; Clete A., et al., 2012). Um exemplo prático, por exemplo, está em pacientes de idade muito avançada, acima dos 90 anos, ou em uma pessoa com uma condição rara, que necessitam da utilização de estratégias que não permitissem a identificação dos indivíduos em relação à seus dados de saúde.

- **3.4 Política de dados abertos**

Regulamentada pelo Decreto 8.777 de 2016, o Decreto 9.903 de 2019 e a Resolução nº 3 do Comitê Gestor da Infraestrutura Nacional de Dados Abertos (INDA), a Política Nacional de Dados Abertos define as regras para disponibilização de dados abertos no âmbito da administração pública brasileira.

Essa política é operacionalizada pelo Plano de Dados Abertos, "documento orientador para as ações de implementação e promoção de abertura de dados de cada órgão ou entidade da administração pública federal, obedecidos os padrões mínimos de qualidade, de forma a facilitar o entendimento e a reutilização das informações" (BRASIL, 2016).

A política tem como objetivo contribuir para aumentar a transparência, permitindo melhor controle social das ações governamentais. As normativas federais devem servir de base para orientar outras organizações na preparação para adequação às diretrizes da LGPD, que prevê o direito de acesso aos dados gerados pelas organizações públicas e privadas por parte dos titulares dos dados.

De acordo com a Resolução nº 3 de 13 de outubro de 2017, que dispõe sobre as normas para elaboração e publicação do Plano de Dados Abertos, as bases de dados a serem disponibilizadas devem ser justificadas e priorizadas pelo grau de relevância para o cidadão, o estímulo ao controle social, a obrigatoriedade legal ou compromisso assumido de disponibilização daquele dado, seu papel estratégico em projetos do governo, a possibilidade do dado de demonstrar resultados diretos e efetivos dos serviços públicos disponibilizados ao cidadão pelo Estado, a sua capacidade de fomento ao desenvolvimento sustentável e a possibilidade de fomento a negócios da sociedade.

Já a Lei nº 12.527 de 18 de novembro de 2011, que regula o acesso aos registros, determina-se que as plataformas onde estão disponibilizados os dados devem conter: ferramentas de pesquisa de conteúdo que permitam o acesso aos registros de forma objetiva, transparente, clara e com linguagem de fácil compreensão, a possibilidade de gravação dos relatórios em diversos formatos que permitam a fácil análise dos dados, como as planilhas, o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por computadores, a divulgação detalhada dos formatos utilizados para estruturar a informação, a atualização constante do sistema, a garantia de autenticidade e integridade, um local que permita a comunicação com o órgão ou entidade detentora da plataforma e as medidas necessárias para acessibilidade do conteúdo para pessoas com deficiência.

A Lei dispõe, ainda, sobre os pedidos de acesso, que podem ser feitos por qualquer meio legítimo, contanto que possuam a identificação do requerente e a especificação da informação requerida, determinando um prazo não superior a 20

dias para o atendimento da solicitação, com possibilidade de prorrogação por mais 10 dias mediante uma justificativa expressa do órgão.

Essas resoluções são de importante compreensão na medida que permitem delinear a operacionalização e implementação da política de dados abertos dentro das demais organizações, permitindo compreender melhor quais atribuições dizem respeito aos operadores e controladores, que possuem a responsabilidade na operação de tratamento dos dados pessoais. (LAPIN, 2021).

4. OBJETIVOS

4.1. OBJETIVO GERAL

O presente trabalho tem como objetivo desenvolver e elaborar protocolos de mapeamento do fluxo de pesquisa e compartilhamento de dados abertos do ISD como base das diretrizes de desenvolvimento de sistemas de compartilhamento de dados de pesquisa.

4.2. OBJETIVOS ESPECÍFICOS

- Mapear os aspectos éticos e legais envolvendo o uso do PEP em bases de dados para pesquisa científica;
- Analisar o uso do PEP no Instituto Santos Dumont classificando as principais características para a disponibilização dos dados de forma anonimizada, dentro dos novos protocolos da Lei Geral de Proteção de Dados;
- Desenvolver protocolos de informação com o fluxo de dados abertos que favoreça ao desenvolvimento de um sistema com base em uma perspectiva multidisciplinar, com a contribuição de áreas como Ciência da Informação, Ciência de Dados, Comunicação Social e Tecnologia da Informação.

5. METODOLOGIA

O delineamento do presente estudo se deu em torno da proposta de criação de uma plataforma, por meio de um conjunto de *scripts*, recurso de programação que automatiza tarefas que seriam executadas por operadores humanos, para disponibilizar, de forma anônima, dados oriundos dos PEPs do Instituto Santos Dumont.

O estudo se configura como uma pesquisa aplicada, cujas primeiras etapas devem consistir na identificação dos fluxos de pesquisa e negócios da instituição, a fim de determinar as etapas de registro, aquisição, classificação, guarda e compartilhamento dos dados de pesquisa entre instituições e pesquisadores vinculados ao ISD. Em seguida, será feito o levantamento das variáveis existentes no PEP e as conexões viáveis com o banco de dados a ser disponibilizado.

A identificação dos fluxos existentes no ISD será fundamental para determinar os diferentes graus e tipos de acesso aos dados coletados, tendo em vista a natureza sensível dos registros em saúde. A partir dessa definição, serão estabelecidos os protocolos e requisições para os diferentes tipos de acesso a diferentes dados, como por exemplo: engajamento do paciente na reabilitação, dados de natureza socioeconômica, idade, gênero, município de origem etc.

A definição dessas variáveis se configura como uma etapa fundamental para o prosseguimento das fases subsequentes, tendo em vista que a definição das conexões existentes deverá levar em conta os aspectos éticos e legais sobre os dados para garantir sua anonimização de forma eficiente. Em seguida, projeta-se a criação do banco de dados a ser exibido no formato de painel, a definição das métricas e *scripts* de consulta, implementação do painel e disponibilização do código e documentação.

As etapas do projeto são pensadas a partir de uma perspectiva multiprofissional e interdisciplinar, no qual profissionais de diversos campos como Tecnologia da Informação, Ciência da Computação, Ciência da Informação e Comunicação Social devem trabalhar em conjunto para fazer as definições necessárias e a correta implementação do banco de dados.

As métricas pré-definidas serão baseadas em números de atendimentos e procedimentos clínicos, características sociais e epidemiológicas da população

atendida, uso dos recursos de saúde disponíveis e engajamento do paciente e família.

A definição do escopo de quais métricas serão implementadas ainda vai ser discutido no início do projeto.

As estratégias para anonimização dos dados serão baseadas nas técnicas descritas na literatura (Kushida; Clete A., *et al.*, 2012) além de não ser importado pelo sistema os dados que levem a identificação de pacientes (CPF, endereço, idade se elevada etc).

Diante de tudo isso, temos um plano de trabalho que se divide em três partes fundamentais: a primeira consiste na identificação de quais informações atualmente são preenchidas no PEP ou deixam de ser preenchidas em cada linha de cuidado. Em seguida, serão feitas entrevistas com os profissionais de referência dessas linhas de cuidado para compreender por que isso acontece. Após essa etapa, é hora de discutir com os pesquisadores quais são as principais informações que costumam ser solicitadas às linhas de cuidado para os estudos, e identificar se eles se encontram entre os dados que já costumam ser preenchidos ou não. Se não, será necessário discutir com os profissionais estratégias para que isso passe a ser preenchido nas consultas.

A segunda etapa diz respeito à definição de variáveis, padronização dos registros e como eles serão incorporados ao banco de dados. Por fim, a terceira etapa diz respeito à discussão sobre os diferentes níveis de acesso aos registros, que serão decididos em conjunto com os coordenadores de pesquisa e educação em saúde da instituição. Esse fluxo deverá ser bem desenhado, com cada escolha justificada a partir do grau de sensibilidade do dado.

6. RESULTADOS ESPERADOS

Com a execução desse estudo, espera-se ter uma visão clara dos fluxos de pesquisa do ISD, atualmente não delineados, de forma a poder disponibilizá-los de forma pública e organizada para aumentar a transparência da instituição e ampliar as possibilidades de parcerias para a execução de estudos tanto internos como multicêntricos.

Este mapeamento vai permitir estabelecer as diretrizes base para a criação do sistema de gestão de dados, em conformidade com as exigências da LGPD e do Plano Nacional de Dados Abertos, reafirmando o compromisso da instituição com a gestão responsável dentro e o fortalecimento do Sistema Único de Saúde (SUS) através da pesquisa científica, tendo em vista o princípio da integralidade da assistência em saúde.

Além disso, o estudo vai mapear as variáveis atualmente existentes nos PEPs e como elas podem ser de fato anonimizadas de tal forma que sua disponibilização seja possível em uma base de dados aberta, além de definir também os diferentes graus de acesso para cada tipo de solicitante.

Esse processo vai permitir automatizar tarefas que antes eram descentralizadas e de difícil execução prática, tendo em vista o elevado tempo demandado para sua execução e a necessidade de busca manual por registros ainda não organizados em um banco de dados de fácil acesso.

7. REFERÊNCIAS CONSULTADAS

BRASIL, Decreto nº 8.777, de 11 de maio de 2016. *Institui a Política de Dados Abertos do Poder Executivo federal*. Brasília: Governo Federal. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/D8777.htm
Acesso em 20 de novembro de 2022.

BRASIL. Decreto nº 7.724, de 16 de maio de 2012. *Regulamenta a Lei nº 12.527 de 18 de novembro de 2011*. Brasília: Governo Federal. 2012. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2012/Decreto/D7724.htm
Acesso em 20 de novembro de 2022.

BRASIL. Lei nº 12.527 de 18 de novembro de 2011. *Regula o acesso a informações*. Brasília: Governo Federal. 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2011/Lei/L12527.htm Acesso em 20 de novembro de 2022.

Gonçalves, João Paulo Pereira *et al.* "Prontuário Eletrônico: uma ferramenta que pode contribuir para a integração das Redes de Atenção à Saúde". Saúde em Debate, v. 37, 2013.

INSTITUTO SANTOS DUMONT. "Relatório anual do Contrato de Gestão 2021". Disponível em: <http://www.institutosantosdumont.org.br/relatorios-contrato-gestao/>

Kushida, Clete A., *et al.* "Strategies for de-identification and anonymization of electronic health record data for use in multicenter research studies". Medical care 50.Suppl (2012): S82.

LAPIN. "Controlador ou operador: quem sou eu? Cartilha sobre agentes de tratamento de dados pessoais". 2021.

Miranda, Nelson Julio de Oliveira; Pinto, Virginia Bentes. "Prontuário Eletrônico do Paciente: padronização e interoperabilidade". XVI Encontro Nacional de Pesquisa em Ciência da Informação, 2015.

Mota, Francisca R. L.; Babêto, H.S. *"Processamento e compartilhamento da informação em prontuários eletrônicos"*. Sociedade Brasileira de Informática em Saúde, Anais do IX Congresso Brasileiro de Informática em Saúde, 2004.

Sayão, Luis F.; Sales, Luana. *"Afiml, o que é dado de pesquisa?"*. Biblos: Revista do Instituto de Ciências Humanas e da Informação, v. 24, nº 02, 2020.

Setzer, Valdemar W. *"Os Meios Eletrônicos e a Educação: Uma Visão Alternativa."* São Paulo: Editora Escrituras, Coleção Ensaio Transversais Vol. 10, 2001.

Semidão, Rafael Aparecido Moron. *"Dados, Informação e Conhecimento: elementos de análise conceitual"*. DataGramZero - Revista de Informação, vol. 14, 2013.

Silva, João Carlos; Carvalho, Cedric Luiz. *"A sociedade da informação e do conhecimento: presente e futuro"*. Revista UFG, ano XI, nº 7, 2009.

Souza, A. D.; De Almeida, M. B. *Análise de dados clínicos textuais de Prontuários Eletrônicos do Paciente para integração com terminologias médicas padronizadas.*

THE ROYAL SOCIETY. *"Science as an open enterprise"*. London: The Royal Society Science Policy Centre, 2012. Disponível em: <https://royalsociety.org/topics-policy/projects/science-publicenterprise/report/>

9. ORÇAMENTO

Pagamento de bolsa de Iniciação Científica: R\$ 600 por mês/cada

Quantidade de bolsistas estimados para ingressar no projeto: 2

Tempo estimado: um ano

Total de investimento: R\$ 14.400,00