

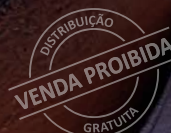
MINISTÉRIO DA SAÚDE


DEMOGRAFIA
da enfermagem
no Brasil 2025

Fundamentos Metodológicos e
Questões Éticas e Jurídicas
Relacionadas
ao Uso de
Bancos de
Dados

VOLUME 2

Brasília – DF
2025


DISTRIBUIÇÃO
VENDA PROIBIDA
GRATUITA

MINISTÉRIO DA SAÚDE
Secretaria de Gestão do Trabalho e da Educação na Saúde
Departamento de Gestão e Regulação do Trabalho na Saúde

Demografia da
Enfermagem no Brasil 2025

VOLUME 2

Fundamentos Metodológicos e Questões Éticas e Jurídicas Relacionadas ao Uso de Bancos de Dados

Brasília – DF
2025





Esta obra é disponibilizada nos termos da Licença Creative Commons – Atribuição – Não Comercial – Compartilhamento pela mesma licença 4.0 Internacional. É permitida a reprodução parcial ou total desta obra, desde que citada a fonte.

A coleção institucional do Ministério da Saúde pode ser acessada, na íntegra, na Biblioteca Virtual em Saúde do Ministério da Saúde: bvsm.sau.gov.br.

1ª edição – 2025 – versão eletrônica

Elaboração, distribuição e informações:

MINISTÉRIO DA SAÚDE

Secretaria de Gestão do Trabalho e da Educação na Saúde

Departamento de Gestão e Regulação do Trabalho na Saúde

Esplanada dos Ministérios, bloco G

CEP: 70058-900 – Brasília/DF

Site: www.sau.gov.br/sgtes

E-mail: sgtes@sau.gov.br

Ministro de Estado da Saúde:

Alexandre Rocha Santos Padilha

Secretário de Gestão do Trabalho e da Educação na Saúde:

Felipe Proença de Oliveira

Colaboração:

Adriana Fortaleza Rocha da Silva – DEGES/SGTES/MS

Aldiney José Doreto – ESPP-SESA

Alice Mariz Porciuncula – FENF/UERJ

Ana Carolina Tavares Vieira – FENF/UERJ

Ana Lygia Pires Melaragno – CORA

André Almeida de Moura – EEUSP/USP

Bianca Freire Ferreira – FSJ

Brendo de Araujo Gomes – PBV/UFRJ

Bruna Drummond Silveira – Fiocruz

Bruno Guimarães de Almeida – DEGERTS/SGTES/MS

Cristiane Helena Gallasch – FENF/UERJ

Cristiane Scolari Gosch – Opas/OMS

Daniela Moraes Morelli – Global Research Nurses para a América Latina e o Caribe

Daniela Paula Polessa – NCE e UERJ

Décio Santiago da Silva Jr. – FEN/UERJ

Deusineth de Oliveira Ferreira da Costa – PPGEN/UERJ

Elen Cristiane Gandra – ENF/UFMG

Ellen Marcia Peres – Cofen

Fabiano Saldanha Gomes de Oliveira – IMS/UERJ

Fernando Canto Michelotti – SGTES/MS

Fernando de Oliveira Santoro – ESDI/UERJ

Francisco Carlos Félix Lana – EEUSP/UFMG

Gerson Luiz Marinho – EEAN/UFRJ

Gustavo de Oliveira Tavares – DENF/UFRN

Gustavo Hoff – DEGERTS/MS

Handerson Silva Santos – EEUFBA

Helena Maria Scherlowski Leal David – FENF/UERJ

Heloísa Helena Oliveira da Silva – Cofen

Hercules de Oliveira Carmo – EEUSP/USP

Hugo Pinto de Almeida – FENF/UERJ

Janaina Sampaio Guerra Oliveira – SGTES/MS

João André Tavares Álvares da Silva – ESP-MG

Jonas Sâmí Albuquerque de Oliveira – DENF/UFRN

Jorge de Abreu Soares – CEFET/RJ

Kênia Lara da Silva – ENF/UFMG

Leila Senna Maia – IMS/UERJ

Leonardo Dresch Eberhardt – FENF/UNICAMP

Lidiane da Silveira Gouvea Toledo – CIDACS/Fiocruz

Livia Angeli Silva – EEUFBA

Lucas Henrique Lobato de Araujo – UFMG

Magda Guimarães de Araujo Faria – FENF/UERJ

Maria Claudia da Silva Vater – NUBEA/UFRJ

Mário Ribeiro Alves – FENF/UERJ

Mario Roberto Dal Poz – IMS/UERJ

Mercedes de Oliveira Neto – FENF/UERJ

Miguel Ângelo Caixeta Rodrigues – UFJF

Mônica Diniz Durães – Opas/OMS

Neyson Pinheiro Freire – Cofen

Paulo Cabral Filho – LNCC

Pedro Luiz Caixeta Rodrigues – IMS/UERJ

Rafaela Siqueira Costa Schreck – ENF/UFMG

Raquel Nunes da Silva Veiga – MS

Ricardo de Mattos Russo Rafael – FENF/UERJ

Rodrigo Winter Afonso – UNESA

Rogério Bittencourt de Miranda – FENF/UERJ

Samuel da Silva Guedes – DENF/UFRN

Sílvia Dias Passos – EEUFBA

Tatiana Rodrigues de Araujo Eleuterio – FENF/UERJ

Vinicius Layter Xavier – IMS/UERJ

Wanesca Caroline Pereira – DENF/UFRN

Washington Leite Junger – IMS/UERJ

Coordenação da equipe de pesquisa:

Décio Santiago Jr. – FEN/UERJ

Fabiano Gomes – IMS/UERJ

Helena Leal David – FENF/UERJ

Kenia Lara da Silva – ENF/UFMG

Mario Roberto Dal Poz – IMS/UERJ

Coordenação de Comunicação:

Juliana Lima – SGTES/MS

Priscilla Leonel – SGTES/MS

Supervisão editorial:

Priscila Tuy – Núcleo de Revisão Ascom/MS

Projeto gráfico, diagramação, ilustração da capa e arte final:

Eduardo Grisoni – SGTES/MS

Revisão textual:

Priscila Tuy – Núcleo de Revisão Ascom/MS

Normalização:

Daniela Ferreira Barros da Silva – Editora MS/CGDI

Impresso no Brasil / Printed in Brazil

Ficha Catalográfica

Brasil. Ministério da Saúde. Secretaria de Gestão do Trabalho e da Educação na Saúde. Departamento de Gestão e Regulação do Trabalho na Saúde.

Fundamentos metodológicos e questões éticas e jurídicas relacionadas ao uso de bancos de dados [recurso eletrônico] / Ministério da Saúde, Secretaria de Gestão do Trabalho e da Educação na Saúde, Departamento de Gestão e Regulação do Trabalho na Saúde. – Brasília : Ministério da Saúde, 2025.

142 p. : il. – (Demografia da Enfermagem no Brasil 2025 ; v. 2)

Modo de Acesso: World Wide Web: https://bvsm.sau.gov.br/bvs/publicacoes/demografia_enfermagem_mercado_trabalho_v2.pdf

ISBN 978-65-5993-928-2

1. Enfermagem. 2. Demografia. 3. Ética. I. Título. II. Série.

CDU 614.253.5:314.02

Catálogo na fonte – Bibliotecária: Daniela Ferreira Barros da Silva – CRB 1/2686 – Editora MS/CGDI – OS 2025/0535

Título para indexação:

Methodological foundations and ethical and legal issues related to the use of databases

Sumário

Apresentação	5
1 Análise dos problemas éticos e legais da utilização de bancos de dados	7
1.1 Contexto	9
1.1.1 Hipóteses de tratamento de dados: fins de pesquisa e legítimo interesse	11
1.1.2 O papel da ANPD na aplicação da LGPD	11
1.2 Metodologia	12
1.3 Resultados e discussão	13
1.3.1 Pesquisa bibliográfica	14
1.3.2 Pesquisa documental	19
1.4 Reflexões finais	34
2 Descrição e análise de bases de dados	37
2.1 Seleção das fontes de dados	39
2.2 Bases selecionadas	41
2.3 Avaliação da estrutura disponibilizada do conjunto de dados	43
2.4 Avaliação da qualidade dos dados	44
2.5 Produção do dicionário de dados de cada base de dados	45
2.6 Carregamento dos dados nos servidores seguros do LNCC	46
2.7 Produção de informação para o consumo dos grupos temáticos de pesquisa	49
2.8 Possibilidade de futura integração entre as bases de dados	50
2.9 Reflexões finais	50
3 Um estudo da modelagem da Força de Trabalho baseado na dinâmica de sistemas	53
3.1 Metodologia	54
3.2 Modelagem tradicional do mercado de trabalho	55
3.2.1 Setor educacional	58
3.2.2 Mercado de trabalho	60
3.2.3 Complexidades do mercado de trabalho em saúde	64
3.3 Modelagem dinâmica do mercado de trabalho	66
3.3.1 Articulação do problema	66
3.3.2 Hipótese dinâmica e formulação do modelo	68
3.4 Reflexões finais	79
Referências	81
Apêndices	89
Apêndice A – Descrição das Bases	89
Apêndice B – Modelo estrutural estático do gerador de demografias	116
Apêndice C – O modelo comportamental do gerador de demografias	135



Apresentação

O primeiro volume desta obra, *Demografia da Enfermagem no Brasil 2025 – Volume 1: Demografia e Mercado de Trabalho em Enfermagem no Brasil*, estabeleceu um panorama abrangente da força de trabalho em enfermagem no Brasil. Nele, destacamos a relevância crucial da enfermagem para os Objetivos de Desenvolvimento Sustentável e os desafios impostos pela pandemia da covid-19, que evidenciaram a necessidade de sistemas de saúde mais resilientes e o fortalecimento dos Recursos Humanos em Saúde (RHS). Trazemos também a expansão da formação, as assimetrias de distribuição e as condições de trabalho, sublinhando a urgência de dados atualizados e análises aprofundadas para um planejamento eficaz.


Este segundo volume dá continuidade à abordagem da força de trabalho da profissão, trata dos pilares metodológicos e considerações subjacentes que sustentam a análise da demografia e do mercado de trabalho em enfermagem. Enquanto o *Volume 1* apresentou os resultados e cenários, o *Volume 2* detalha o “como” essa compreensão foi construída, abordando aspectos cruciais de governança de dados, infraestrutura tecnológica e modelagem complexa. Ele expande o conhecimento previamente adquirido, fornecendo a base para a robustez das análises apresentadas.

Os capítulos que compõem este volume são:

- **Capítulo 1: Análise dos problemas éticos e legais da utilização de bancos de dados** – Explora os desafios éticos e legais do uso de dados em pesquisa científica, com foco na LGPD, e discute alternativas para acesso e tratamento de informações sensíveis.
- **Capítulo 2: Descrição e análise de bases de dados** – Detalha a metodologia de construção de um Data Warehouse para pesquisa em enfermagem, abordando seleção, avaliação, tratamento e armazenamento seguro de 14 bases de dados cruciais.
- **Capítulo 3: Um estudo da modelagem da força de trabalho baseado na dinâmica de sistemas** – Modela a força de trabalho em enfermagem no Brasil usando dinâmica de sistemas, analisando interações entre variáveis educacionais, demográficas e econômicas que afetam oferta e demanda.

Ao desmistificar a complexidade por trás da coleta, tratamento e interpretação de grandes quantidades de dados, o texto aqui apresentado complementa o Volume 1. São oferecidas, para pesquisadores, gestores e formuladores de políticas públicas, ferramentas e conhecimentos necessários para compreender o cenário atual e desenvolver estratégias baseadas em evidências robustas, garantindo sustentabilidade e equidade dos serviços de saúde no Brasil. Acreditamos que este produto consolidará ainda mais a base para futuras investigações e intervenções no campo da enfermagem.





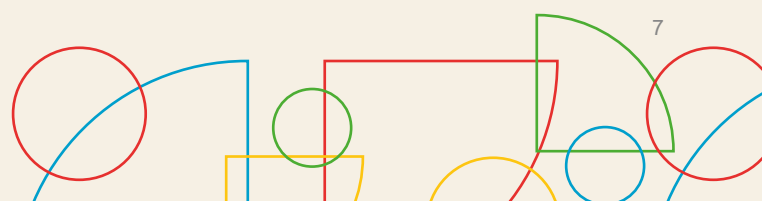
1 Análise dos problemas éticos e legais da utilização de bancos de dados

Este capítulo traz a análise dos desafios éticos e legais da utilização de bancos de dados no âmbito do projeto *Demografia e Mercado de Trabalho em Enfermagem*.

De um modo geral, a pesquisa científica faz uso de diferentes tipos de dados, que podem ser pessoais, públicos ou sensíveis. A coleta de dados primários tem sido regulada por conselhos de ética, enquanto o uso de dados secundários recebeu pouca regulamentação até a Lei n.º 13.709/2018, conhecida como Lei Geral de Proteção de Dados Pessoais (LGPD), que trouxe diversas mudanças sobre seu gerenciamento. Desde sua promulgação, a LGPD afetou diretamente o acesso e manejo dos dados, influenciando as pesquisas científicas. A lei sugere alternativas para esses casos, ainda pouco utilizadas.

Vale registrar que quando se trata de pesquisas na área da saúde, o Brasil contribui com o esforço global para proteger dados pessoais sensíveis tanto na pesquisa acadêmica quanto no contexto jurídico, por meio de Normas e Resoluções do Conselho Nacional de Saúde (CNS) e da LGPD. Quando os dados coletados são classificados como pessoais, é necessário garantir os direitos dos participantes da pesquisa, seja por meio da avaliação prévia de um Conselho de Ética em Pesquisa (CEP), seja pela restrição de acesso ou uso de um banco de dados obtido para outra finalidade. Juridicamente, o cenário se complica quando os dados são sensíveis além de pessoais. Nesse contexto é essencial definir o que são esses dados e o que os distingue.

Conforme a Lei n.º 13.709/2018, dado pessoal é qualquer informação que identifique ou possa identificar uma pessoa. Dados sensíveis abrangem detalhes sobre origem racial, crenças religiosas, opiniões políticas, associação sindical, saúde, vida sexual, dados genéticos ou biométricos. No Brasil, a proteção dos dados pessoais é



garantida pela Constituição Federal de 1988 e está ligada aos direitos à privacidade e à intimidade: “são invioláveis a intimidade, a vida privada, a honra e a imagem das pessoas, assegurado o direito a indenização pelo dano material ou moral decorrente de sua violação” (Brasil, 1988, art. 5º, inc. X).

A LGPD regulamenta essa proteção, destacando o respeito à privacidade, autodeterminação informativa e inviolabilidade da intimidade, honra e imagem (art. 1º). Inspirada no Regulamento Geral sobre a Proteção de Dados (GDPR) da União Europeia, parte da LGPD entrou em vigor no Brasil em 18 de setembro de 2020, enquanto as disposições sobre sanções administrativas começaram a valer em agosto de 2021, durante a pandemia da covid-19. A LGPD alterou profundamente a forma como os dados pessoais são tratados no País, introduzindo novos conceitos e definindo as partes envolvidas e os tipos de dados. A lei confere aos titulares dos dados direitos ao longo do processo de tratamento de suas informações, além de estabelecer requisitos de transparência.

Este estudo visa responder: *quais são as principais implicações éticas e legais do manejo de dados em pesquisas científicas na área da saúde após a implementação da LGPD?* Como resposta preliminar, é importante destacar que, devido ao curto período de vigência da lei, iniciada em 2021 por conta da pandemia da covid-19, existem poucas decisões judiciais sobre o assunto. Tanto o poder judiciário quanto os órgãos reguladores têm restringido ou negado pedidos para o tratamento de dados pessoais em pesquisas acadêmicas, conforme estudo da Autoridade Nacional de Proteção de Dados (ANPD) o que pode dificultar o progresso das pesquisas e criar desafios para a conformidade das práticas acadêmicas com a LGPD (ANPD, 2022). No entanto, vale lembrar que nas pesquisas de saúde, existem regulamentações específicas estabelecidas pela Comissão Nacional de Ética em Pesquisa (Conep) que têm como objetivo garantir a proteção dos participantes por meio da análise criteriosa e aprovação prévia dos protocolos dos estudos.

Contudo, nota-se uma aparente falta de interesse tanto dos órgãos reguladores quanto do judiciário em tratar desse assunto com decisões fundamentadas na razoabilidade e na legalidade, mesmo diante das leis e resoluções vigentes. Portanto, este capítulo tem como propósito identificar e examinar as principais questões éticas, jurídicas e legais que influenciam a pesquisa sobre Demografia e Mercado de Trabalho na área de Enfermagem, tais como: verificar as alterações trazidas pela LGPD na gestão de dados em estudos de saúde; examinar as decisões judiciais relacionadas à proteção de dados pessoais e à aplicação da LGPD, e identificar leis posteriores à LGPD que regulamentam pesquisas no campo da saúde.

1.1 Contexto

A LGPD estabelece os princípios e diretrizes para o tratamento de informações pessoais no Brasil. A lei visa proteger a privacidade dos indivíduos ao mesmo tempo que permite o uso responsável e ético das informações coletadas, assegurando que as atividades de pesquisa possam continuar a contribuir significativamente para avanços médicos e científicos. Compreender os conceitos fundamentais dessa legislação é crucial para garantir que sua implementação não restrinja indevidamente o acesso a dados utilizados em pesquisas científicas, principalmente aquelas focadas na área da saúde.

O que é tratamento de dados?

De acordo com o art. 5º, inciso X, da LGPD, “tratamento de dados” é definido como qualquer ação que envolva um dado pessoal durante sua manipulação. Isso inclui uma ampla gama de atividades, tais como: coleta (obtenção de dados); produção (criação ou geração de dados); recepção (recebimento de dados provenientes de outra fonte); classificação (organização de dados em categorias); utilização (aplicação dos dados para um propósito específico); acesso (capacidade de visualizar ou recuperar os dados); reprodução (cópia dos dados); transmissão (envio de dados a outro local ou entidade); distribuição (divulgação dos dados para um grupo mais amplo); processamento (manipulação técnica dos dados); arquivamento (guardar os dados de maneira organizada para futuro uso); armazenamento (retenção dos dados por um período); eliminação (remoção definitiva dos dados); avaliação ou controle da informação (análise dos dados para verificar qualidade ou finalidade); modificação (alteração dos dados); comunicação (transmissão de informações derivadas dos dados); transferência (movimentação dos dados entre diferentes entidades ou locais); difusão (disseminação dos dados para um público maior); e extração (retirada de dados de uma fonte).

Dentro do escopo da LGPD, o tratamento dos dados pessoais pode ser executado por dois principais agentes: o controlador e o operador. A lei define o controlador como a pessoa física ou jurídica, pública ou privada, que tem a autoridade para tomar decisões sobre como os dados pessoais serão tratados. Ou seja, é o controlador quem determina as finalidades e os meios do tratamento de dados. Na área da Administração Pública, o controlador é representado pela pessoa jurídica do órgão ou entidade pública que está sujeita à LGPD. Este controlador é tipicamente a autoridade responsável pelas decisões sobre o tratamento dos dados dentro daquele órgão ou entidade.

Por outro lado, o operador é definido como a pessoa física ou jurídica, pública ou privada, que realiza o tratamento de dados pessoais em nome do controlador. Isso pode incluir funcionários públicos que executam essas funções, bem como outras organizações que, embora sejam distintas do controlador, realizam atividades de tratamento de dados na base de contratos ou acordos semelhantes.

De acordo com a LGPD, o tratamento de dados pessoais só será autorizado se previsto nas hipóteses do art. 7º da lei, que são:

1. Consentimento
2. Cumprimento de obrigação legal ou regulatória
3. Execução de políticas públicas
4. Realização de estudos por órgão de pesquisa
5. Execução ou criação de contrato
6. Exercício regular de direitos
7. Proteção da vida
8. Tutela da saúde
9. Legítimo interesse
10. Proteção do crédito

A pesquisa focaliza as hipóteses três, quatro e nove do art. 7º da LGPD, que se refere à implementação de políticas públicas, realização de estudos por entidades de pesquisa e interesse legítimo dos controladores. A legislação destaca a importância de proteger a privacidade dos titulares de dados, recomendando a anonimização para evitar qualquer possibilidade de associação direta ou indireta com o titular. Na prática, isso significa que os dados devem ser processados de forma a garantir que não possam ser usados para identificar indivíduos específicos.

Em casos de estudos relacionados à saúde pública, a LGPD estipula que os dados devem ser tratados exclusivamente pela entidade de pesquisa responsável pelo estudo. Isso inclui seguir rigorosos protocolos de segurança para proteger a integridade e confidencialidade das informações. Além disso, essas entidades precisam justificar

o uso dos dados e demonstrar que o processamento é essencial para alcançar os objetivos do estudo. As diretrizes da LGPD buscam equilibrar a necessidade de utilizar dados para avanços científicos e tecnológicos com a proteção dos direitos fundamentais dos indivíduos, promovendo um ambiente de confiança e responsabilidade na gestão de informações pessoais.

1.1.1 Hipóteses de tratamento de dados: fins de pesquisa e legítimo interesse

O conceito de legítimo interesse é uma das justificativas legais mais amplas e adaptáveis estipuladas pela LGPD. Essa legislação permite que dados pessoais sejam processados para favorecer os interesses do controlador ou de terceiros, contanto que isso não contrarie os direitos essenciais do titular dos dados. Conforme delineado no art. 10, qualquer tratamento de dados deve visar objetivos legítimos fundamentados em situações específicas (Brasil, 2018a). É importante notar que dados pessoais sensíveis não podem ser tratados com base no legítimo interesse. Além disso, se o processamento de dados se baseia nesta justificativa, a ANPD pode requerer ao responsável pelo tratamento um relatório demonstrando o impacto na proteção dos dados.

As hipóteses que fundamentam o tratamento de dados pelo legítimo interesse estão elencadas nos arts. 7 e 10 da LGPD e são:

- i. compatibilidade com o ordenamento jurídico;
- ii. lastro em situações concretas;
- iii. vinculação a finalidades legítimas, específicas e explícitas.

1.1.2 O papel da ANPD na aplicação da LGPD

A Autoridade Nacional de Proteção de Dados Pessoais (ANPD) é o órgão responsável por supervisionar e aplicar a LGPD no Brasil. A ANPD atua como intermediária entre a sociedade e o governo, desempenhando um papel crucial na proteção dos dados pessoais. Uma de suas funções é esclarecer dúvidas da população sobre a LGPD, além de coletar sugestões que possam aprimorar a legislação e receber denúncias de irregularidades relacionadas ao uso indevido de dados pessoais.

Para garantir a segurança e privacidade dos dados pessoais, a ANPD orienta tanto entidades governamentais quanto empresas privadas sobre as melhores práticas no manuseio e processamento desses dados. Essa orientação visa assegurar que todas essas instituições estejam em conformidade com a LGPD, prevenindo violações e garantindo que os direitos dos titulares dos dados sejam respeitados.

Vinculada ao Ministério da Justiça, a ANPD é também responsável pela elaboração de normas e diretrizes específicas que complementam e detalham os dispositivos da LGPD. Esse trabalho inclui a criação de regulamentações técnicas e procedimentos operacionais que devem ser seguidos por todas as organizações que lidam com dados pessoais. Além disso, a ANPD realiza monitoramentos regulares para verificar se as entidades estão cumprindo as normativas estabelecidas.

Outro aspecto importante do trabalho da ANPD é a educação e conscientização do público sobre a importância da proteção de dados pessoais. A autoridade promove campanhas informativas e educativas, visando aumentar o entendimento da população sobre seus direitos e a forma de proteger suas informações pessoais. Nesse esforço, a ANPD busca alinhar-se a regulamentações internacionais, como o Regulamento Geral de Proteção de Dados (GDPR) da União Europeia, promovendo uma harmonização global das práticas de proteção de dados. Em resumo, a ANPD desempenha um papel fundamental na proteção dos dados pessoais no Brasil, garantindo que a LGPD seja aplicada de maneira eficiente e eficaz, protegendo os direitos dos cidadãos e promovendo uma cultura de respeito à privacidade e segurança da informação.

Até dezembro de 2024, não existia um acordo definitivo sobre como os dados pessoais devem ser tratados de acordo com a LGPD. A questão não é suficientemente debatida no meio acadêmico e é pouco abordada pelo judiciário.

1.2 Metodologia

Para abordar a questão *Quais são as principais implicações éticas e legais no manuseio de dados em pesquisas científicas na área de saúde após a implementação da LGPD?*, foi realizada uma análise de documentos e revisão de literatura. Foram investigadas as questões éticas relacionadas ao uso de dados depois da implementação da LGPD e identificadas áreas em que o conhecimento ainda é insuficiente. O estudo ocorreu entre 5 de julho e 20 de outubro de 2024. O método incluiu várias etapas: definição da pergunta de pesquisa; planejamento dos termos de busca e seleção das

bases de dados; realização da pesquisa conforme o protocolo definido; organização dos dados coletados; e síntese dos resultados obtidos.

Foram utilizados os seguintes termos de busca: “proteção de dados”; “LGPD”; “pesquisa científica”; “área de saúde”; “prejuízos”; “restrições” e “limitações”. Para combinação dos termos de busca, foram utilizados os operadores booleanos: “OR”, “AND” e “NOT”, este último com a finalidade de excluir resultados relacionados às palavras que não tivessem relação com a pesquisa. Durante a pesquisa, foram consultadas bases de dados que incluíam trabalhos das áreas de saúde, tecnologia da informação e direito. Os critérios para inclusão e exclusão foram os seguintes: (1) Focar na produção acadêmica a partir de 2018, ano em que a LGPD foi formulada; e (2) Selecionar trabalhos relacionados com a LGPD, principalmente aqueles na área da saúde.

Na revisão documental, foi adotado como critério a seleção de textos que tratassem da temática em âmbito nacional a partir do advento da LGPD. Os documentos foram levantados nos sites de órgãos governamentais, tais como o Conselho da Justiça Federal (<https://www.cjf.jus.br/publico/lgpd/index.html>) e a Autoridade Nacional de Proteção de Dados (<https://www.gov.br/anpd/pt-br>) e Portal da Legislação (<https://www4.planalto.gov.br/legislacao>).

Ao mesmo tempo, foi conduzida uma investigação nos Tribunais de Justiça do Sudeste para coletar decisões judiciais relacionadas ao tratamento de dados em pesquisas científicas, principalmente na área da saúde. A pesquisa considerou como critérios as sentenças dadas desde 2020, baseadas na LGPD. As decisões dos tribunais de justiça do Sudeste foram extraídas dos seguintes sites: Tribunal de Justiça do Espírito Santo: <https://www.tjes.jus.br/>; Tribunal de Justiça de Minas Gerais: <https://www.tjmg.jus.br/>; Tribunal de Justiça do Rio de Janeiro: <http://www.tjrj.jus.br/>; e Tribunal de Justiça de São Paulo: <https://www.tjsp.jus.br/>.

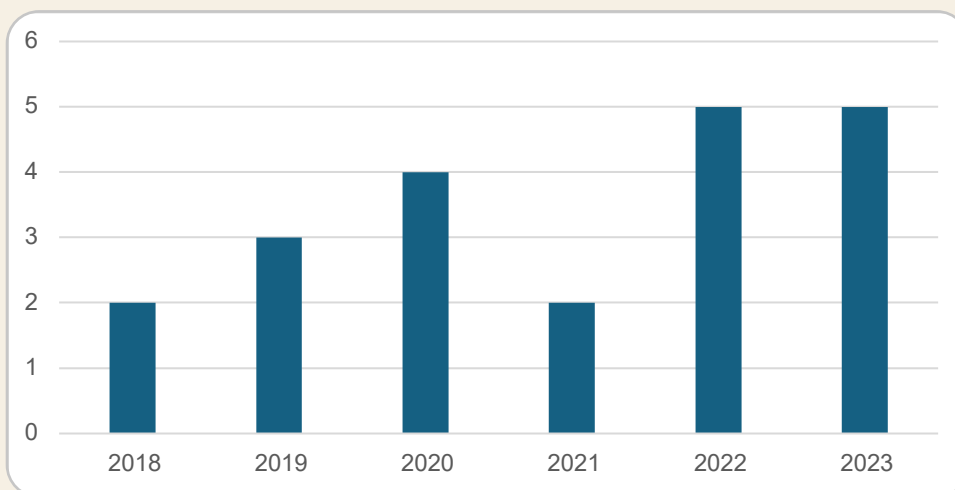
1.3 Resultados e discussão

Os resultados serão apresentados de duas maneiras distintas. Primeiramente, será realizada uma análise da produção bibliográfica, destacando sua origem e a distribuição dos estudos. Em seguida, os dados organizados serão apresentados considerando-se cinco categorias de análise: 1- Publicações relacionadas à regulamentação pela Autoridade Nacional de Proteção de Dados (ANPD); 2- Regulamentações no Brasil; 3- Legislação e decisões judiciais; 4- A posição dos tribunais; e 5- Decisões de órgãos administrativos.

1.3.1 Pesquisa bibliográfica

A pesquisa encontrou 21 títulos, publicados em 3 idiomas e 16 revistas diferentes: 17 artigos em português (81%), 3 artigos em inglês (14%) e 1 artigo em espanhol (5%). As revistas são originárias de 4 países: Brasil, com 17 artigos (81%); Portugal, com 2 artigos (9%); Espanha, com 1 artigo (5%); e Suécia, com 1 artigo (5%). A análise dos títulos por ano de publicação indicou que 2022 e 2023 tiveram o maior número de publicações, com cinco artigos cada. A distribuição anual das publicações pode ser vista na Figura 1.

Figura 1 – Produção acadêmica por ano de publicação



Fonte: elaboração própria.

Ressalta-se que apesar da LGPD ter sido promulgada em 18 de setembro de 2020, as sanções administrativas começaram a valer em 1º de agosto de 2021, o que explica o aumento da produção nos anos seguintes. Em 2018, dois artigos abordaram tanto a Lei Geral de Proteção de Dados (GDPR) da União Europeia quanto a legislação brasileira, despertando simultaneamente o interesse dos acadêmicos portugueses.

A produção brasileira sobre o tema começou em 2019 com dois artigos. Nos anos seguintes, houve quatro artigos em 2020, dois em 2021, quatro em 2022 e cinco em 2023. Quadro 1 apresenta a distribuição da produção intelectual dos países, número de artigos e anos de publicação. Até outubro de 2024, não foram encontrados resultados para 2024 nas bases pesquisadas.

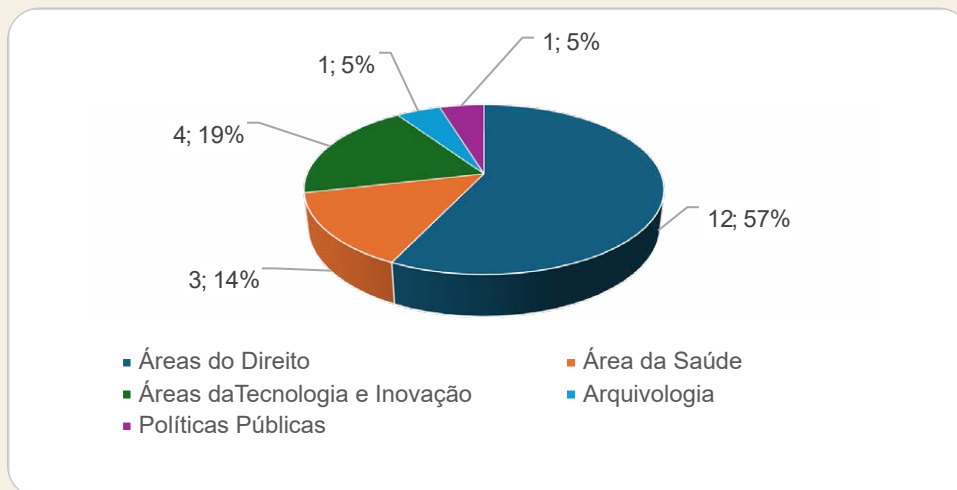
Quadro 1 – Número de artigos anuais por países de publicação

Ano	País	N. ° de títulos
2018	Portugal	2
2019	Suécia	1
	Brasil	2
2020	Brasil	4
2021	Brasil	2
2022	Brasil	4
	Espanha	1
2023	Brasil	5

Fonte: elaboração própria.

É importante notar que as publicações abrangem várias áreas do conhecimento, como apresentado na Figura 2.

Figura 2 – Distribuição da produção por áreas de conhecimento



Fonte: elaboração própria.

A ilustração revela que a maior parte das produções se concentra nas áreas de direito (57%), depois tecnologia e inovação (19%), saúde (14%), políticas públicas (1%) e arquivologia (1%).

O Quadro 2 traz a lista completa dos 23 títulos, especificando os autores, ano, idioma, periódicos e países onde foram publicados.

Quadro 2 – Produção bibliográfica sobre a LGPD selecionada no estudo

Base	Autor	Título	Ano	Periódico	Idioma	País
Scopus	Luis Douglas Binda Filho <i>et al.</i>	A Lei Geral de Proteção de Dados e suas implicações na saúde: as Avaliações de Impacto no tratamento de dados no âmbito clínico-hospitalar	2022	Revista de Bioética y Derecho	Português	Espanha
	Anabela Cristina Hirata <i>et al.</i>	A Proteção Jurídica dos Dados Pessoais Sensíveis Relativos à Saúde	2022	Revista Jurídica UNICURITIBA	Português	Brasil
	Nathan Lea	Como é que o regulamento geral da proteção de dados irá afetar a Saúde?	2018	Acta Med Portuguesa -Revista Científica da Ordem dos Médicos	Inglês	Portugal
	Carlo Botrugno	Cibersegurança, privacidade e proteção de dados de saúde na estratégia digital da União Europeia	2022	Revista de Estudos Constitucionais, Hermenêutica e Teoria do Direito	Inglês	Brasil
	Elda Coelho de Azevedo Bussinguer <i>et al.</i>	Governança da Privacidade e as Organizações Sociais de Saúde: Uma Análise da Lei Geral de Proteção de Dados	2022	Revista Jurídica UNICURITIBA	Português	Brasil
	Isabel Maria Lopes <i>et al.</i>	Aplicabilidade do regulamento geral sobre proteção de dados em clínicas de saúde	2018	Revista Ibérica de Sistemas e Tecnologias de Informação	Espanhol	Portugal
	Daniel de Araujo Dourado <i>et al.</i>	A regulação da inteligência artificial na saúde no Brasil começa com a Lei Geral de Proteção de Dados Pessoais	2022	Revista de Saúde Pública	Português	Brasil
	Mário Cesar Botelho <i>et al.</i>	A Aplicação da Lei Geral de Proteção de Dados na Saúde	2020	Revista Direito Sanitário	Português	Brasil
	Cristina Mendes Santos <i>et al.</i>	A guided internet-delivered individually-tailored ACT-influenced cognitive behavioural intervention to improve psychosocial outcomes in breast cancer survivors (iNOVBC): Study protocol	2019	Internet Interventions	Inglês	Suécia
Diadorim	Gabrielle Bezerra Sales Sarlet <i>et al.</i>	O consentimento informado e a proteção de dados pessoais de saúde na internet: uma análise das experiências legislativas de Portugal e do Brasil para a proteção integral da pessoa humana	2019	Civilistica.com	Português	Brasil
	Manuel David Masseno <i>et al.</i>	A segurança na proteção de dados: entre o GDPR europeu e a LGPD brasileira	2020	Revista. do Cejur: Prestação Jurisdicional	Português	Brasil/SC

continua

conclusão

Base	Autor	Título	Ano	Periódico	Idioma	País
Google Scholar	Bethânia de A. Almeida <i>et al.</i>	Compartilhamento e abertura de dados na Ciência Aberta e a pesquisa em saúde pública na Lei Geral de Proteção de Dados	2021	Revista Bahia Análise & Dados	Português	Brasil/BA
Google Scholar	Paula Cotrim de Abrantes <i>et al.</i>	Dados sensíveis de pesquisa relacionados à saúde numa perspectiva da Lei Geral de Proteção de Dados Pessoais	2023	Revista Conhecimento em Ação	Português	Brasil/RJ
Crossref	Suélynn Mattos de Aragão <i>et al.</i>	Lei Geral de Proteção de Dados: desafio do Sistema Único de Saúde	2020	Revista Eletrônica de Comunicação, Informação e Inovação em Saúde	Português	Brasil/RJ
Google Scholar	Paula Cotrim de Abrantes <i>et al.</i>	Dados sensíveis de pesquisa relacionados à saúde numa perspectiva da Lei Geral de Proteção de Dados Pessoais	2023	Revista Conhecimento em Ação	Português	Brasil/RJ
Crossref	Suélynn Mattos de Aragão <i>et al.</i>	Lei Geral de Proteção de Dados: desafio do Sistema Único de Saúde	2020	Revista Eletrônica de Comunicação, Informação e Inovação em Saúde	Português	Brasil/RJ
Diadorim	Lenora de Beaurepaire da Silva Schwaitzer	LGPD E ACERVOS HISTÓRICOS: impactos e perspectivas	2020	Archeion Online	Português	Brasil/PB
Google Scholar	Camilla Ellen Aragão Costa <i>et al.</i>	Lei Geral De Proteção de Dados Aplicada à Pesquisa Científica	2023	Revista de Pesquisa e Educação Jurídica	Português	Brasil/SC
Latindex	Marcos Antonio Madeira de Mattos Martins <i>et al.</i>	Monitoramento e Compartilhamento de Dados: a Suposta Antinomia entre o Poder de Vigilância Sanitária e a Lei Geral de Proteção de Dados	2023	Revista de Ciências Jurídicas e Empresariais	Português	Brasil/SP
Google Scholar	Fabricio Vasconcelos Gomes <i>et al.</i>	Proteção de dados e instituições de ensino: o que fazer com dados de alunos?	2023	Revista Brasileira de Políticas Públicas	Português	Brasil/DF
Latindex	José Maurício de Carvalho <i>et al.</i>	A Proteção dos Dados Sensíveis na Pesquisa Científica, Miguel Reale e a Individualidade da Pessoa	2023	Revista De Propriedade Intelectual - Direito Contemporâneo E Constituição	Português	Brasil/ SE
Google Scholar	Edith Maria Barbosa Ramos <i>et al.</i>	Questões éticas e perspectiva jurídica da proteção de dados	2021	Cadernos Ibero-Americanos de Direito Sanitário	Português	Brasil/DF
Google Scholar	Gabrielle Bezerra Sales Sarlet <i>et al.</i>	Questões tecnológicas, éticas e normativas da proteção de dados pessoais na área da saúde em um contexto de big data	2019	Revista Direitos Fundamentais & Justiça	Português	Brasil/MG

Fonte: elaboração própria.

Não foram localizadas dissertações ou teses de doutorado que tratem diretamente da temática desta pesquisa. Considerando que a LGPD entrou em vigor em 2020, complementada em 2021, há uma possibilidade de que o tema esteja atualmente sendo investigado. Os estudos publicados até o momento concentram-se em entender a dinâmica da LGPD, mas não exploram de maneira aprofundada suas consequências práticas e legais. Em linhas gerais, esses trabalhos mostram como as práticas seriam modificadas pela LGPD, porém não detalham de forma concreta os impactos no tratamento de dados na pesquisa em saúde.

1.3.2 Pesquisa documental

A pesquisa documental será apresentada sob cinco perspectivas de análises como informadas na sequência.

1.3.2.1 Publicações da ANPD

Como já explicitado, a ANPD elabora normas que regulamentam a lei e orienta os agentes sobre a proteção de dados pessoais. No repositório da autoridade (<https://www.gov.br/anpd/pt-br/documentos-e-publicacoes>), foram encontradas diversas publicações, como cartilhas, estudos e notas técnicas sobre o assunto, sendo a maioria documentos técnicos datada de 2023 e 2024.

Foram identificadas 21 cartilhas relacionadas ao tratamento de dados e à aplicação da LGPD no Brasil, das quais, para esta pesquisa, selecionamos as seguintes:

- Guia Orientativo – Tratamento de dados pessoais para fins acadêmicos e para a realização de estudos e pesquisas – junho 2023
- Guia Orientativo – Tratamento de dados pessoais pelo Poder Público – junho 2023
- Guia Orientativo – Hipóteses legais de tratamento de dados pessoais legítimo interesse – fevereiro 2024
- O Glossário de Proteção de Dados Pessoais e Privacidade – maio 2024

O portal do governo federal disponibiliza importantes estudos técnicos fornecidos pela ANPD. Com base nos critérios da pesquisa, esses estudos podem ser acessados em <https://www.gov.br/anpd/pt-br/documentos-e-publicacoes>:

- Estudo Técnico – Anonimização de dados na LGPD: uma visão de processo baseado em risco e técnicas computacionais

- Estudo Técnico – Anonimização de dados na LGPD: Análise Jurídica
- Estudo Técnico – Estudos de Casos sobre Anonimização de Dados na LGPD
- Texto para discussão nº 1/2022 – A LGPD e o tratamento de dados pessoais para fins acadêmicos e para a realização de estudos por órgão de pesquisa

No mesmo site, foram encontradas 14 notas técnicas que regulam o tratamento de dados de acordo com os princípios da LGPD. Dentre estas, foram selecionadas aquelas que estão relacionadas ao tema desta pesquisa:

- Nota Técnica n.º 22/2024/FIS/CGF/ANPD – Orientações aos servidores da Coordenação-Geral de Fiscalização (CGF) sobre a análise de publicidade dos documentos que instruem os processos de sua competência em atenção à Lei de Acesso à Informação
- Nota Técnica n.º 19/2023/CGF/ANPD – Atividade de Monitoramento
- Nota Técnica n.º 12/2023/CGF/ANPD – Avaliação dos Relatórios de Impacto à Proteção de Dados elaborados pelo Inep para fins adequação da divulgação dos microdados do censo escolar e do Enem à Lei n.º 13.709, de 14 de agosto de 2018 — Lei Geral de Proteção de Dados Pessoais (LGPD)
- Nota Técnica n.º 4/2023/CGTP/ANPD – Manifestação técnica da Coordenação-Geral de Tecnologia e Pesquisa acerca do alinhamento de conformidade com a LGPD e sua aplicação no varejo farmacêutico
- Nota Técnica n.º 46/2022/CGF/ANPD – Manifestação técnica da Coordenação-Geral de Fiscalização acerca da divulgação dos microdados do Enem e de censos escolares pelo Inep à luz da Lei n.º 13.709, de 14 de agosto de 2018 — Lei Geral de Proteção de Dados Pessoais (LGPD)
- Nota Técnica n.º 20/2022/CGN/ANPD – Proposta de realização de Tomada de Subsídios para regulamentação de transferência internacional de dados pessoais, nos termos dos arts. 33 e 35 da LGPD da Lei n.º 13.709, de 14 de agosto de 2018. É acompanhada do Aviso de Tomada de Subsídio e da Lista de Perguntas da Tomada de Subsídio

1.3.2.2 Legislação e decisões judiciais

Desde que a LGPD entrou em vigor, outras legislações foram criadas para complementar e esclarecer pontos controversos sobre o tratamento de dados. No site do Conselho da Justiça Federal <https://www.cjf.jus.br/publico/lgpd/index.html>, estão disponíveis as seguintes legislações:

- Lei n.º 13.709/2018 - Lei Geral de Proteção de Dados Pessoais

- MP 1.124 - Alterou a Lei n.º 13.709, de 14 de agosto de 2018
- Lei n.º 13.853/2019 - Alterou a Lei n.º 13.709, de 14 de agosto de 2018, para dispor sobre a proteção de dados pessoais e para criar a Autoridade Nacional de Proteção de Dados e dá outras providências
- Lei n.º 14.010/2020 - Adiou de 1º de janeiro de 2021 para 1º de agosto de 2021 a vigência das sanções que a Autoridade Nacional de Proteção de Dados (ANPD), ainda pendente de instalação, pode aplicar nos órgãos, entidades e empresas que lidam com o tratamento de dados
- Portaria n.º 212 de 15/10/2020 - Instituiu um Grupo de Trabalho destinado à elaboração de estudos e de propostas votadas à adequação dos tribunais à Lei Geral de Proteção de Dados
- Portaria n.º 213 de 15/10/2020 - Institui Comitê Gestor da Lei Geral de Proteção de Dados Pessoais
- Resolução n.º 334/2020, do Conselho Nacional de Justiça - Instituiu o Comitê Consultivo de Dados Abertos e Proteção de Dados no âmbito do Poder Judiciário
- Recomendação n.º 73/2020, do Conselho Nacional de Justiça - Orientou os órgãos do Poder Judiciário brasileiro quanto à adoção de medidas preparatórias e ações iniciais para a adequação às disposições contidas na LGPD
- Resolução n.º 363/2021, do Conselho Nacional de Justiça - Estabeleceu medidas para o processo de adequação à LGPD a serem adotadas pelos tribunais
- EC n.º 115/2022 - Emenda Constitucional que altera a Constituição Federal para incluir a proteção de dados pessoais entre os direitos e garantias fundamentais
- Lei n.º 14.460/2022 - Transforma a Autoridade Nacional de Proteção de Dados (ANPD) em autarquia de natureza especial
- Resolução STJ/GP n.º 5/2023 - Estabelece as diretrizes para a Política de Proteção de Dados Pessoais do Superior Tribunal de Justiça
- Resolução n.º 738/24 do Conselho Nacional de Saúde. Dispõe sobre uso de bancos de dados com finalidade de pesquisa científica envolvendo seres humanos

A investigação sobre decisões judiciais foi conduzida nos Tribunais de Justiça da região sudeste, abrangendo sentenças desde 2020 com base na Lei Geral de Proteção

de Dados referentes ao tratamento de informações em pesquisas científicas e na área da saúde. Constatou-se uma escassez de decisões e publicações sobre este assunto, devido ao curto período de vigência da Lei e à pandemia, o que afetou os resultados obtidos, essenciais para enfrentar os desafios atuais no acesso a dados.

1.3.2.3 Regulamentação no Mundo

A preocupação com a proteção de dados começou a ganhar relevância de forma global a partir da década de 1970. Foi em Hessen, na Alemanha, que surgiu a primeira lei específica sobre o tema. Atualmente, diversos países regulam a manipulação de dados pessoais por meio de legislações específicas. Na União Europeia, por exemplo, todos os membros aplicam o Regulamento Geral de Proteção de Dados (GDPR), abrangendo países como Alemanha, França, Itália.

Aproximadamente 19% das nações, incluindo a Índia, ainda não estabeleceram leis específicas para lidar com essa questão, conforme informações da Conferência das Nações Unidas sobre Comércio e Desenvolvimento (UNCTAD). Nos Estados Unidos, a regulamentação se dá de maneira diferenciada entre os estados, com atos como o *Colorado Privacy Act* (CPA), o *California Consumer Privacy Act* (CCPA) e o *Utah Consumer Privacy Act* (UCPA).

- Cenário mundial

Neste tópico serão apresentados os países que possuem legislação específica sobre o tema e ao final algumas considerações sobre essas legislações importantes para a construção dos objetivos desta pesquisa conforme demonstrado no Quadro 3.

Quadro 3 – Países com legislações específicas sobre a temática, seus objetivos e limites

Continente	País	Legislação	Data legislação	Objetivos	Limites	Observações
Américas	Argentina	25.326/00	04/10/2000	Proteger a privacidade de dados pessoais e garantir o acesso individual a informações armazenadas em bancos de dados		
Américas	Brasil	13.709/18	13/08/2018	Estabelecer regras de tratamento de dados pessoais, para proteger a liberdade e privacidade dos titulares	Restringiu o tratamento de dados aos casos previstos em lei; trouxe o instituto do consentimento	A lei entrou em vigor integralmente apenas em 1 de agosto de 2021
Américas	Canadá	S.C. 2000, c. 5	abr/00	Responsabilidade; Identificação da Finalidade; Consentimento; Limitação da Obtenção das Informações e Uso; Divulgação e Retenção	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	
Américas	Colômbia	1.266/08 1.273/09 1.581/12 1.377/13	31/12/2008 07/6/2009 17/10/2012 27/06/2013	Consentimento; Diretrizes sobre crimes cibernéticos; Direito dos titulares de determinar a coleta e o armazenamento dos dados; Privacidade de dados comerciais e financeiros	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	
Américas	Estados Unidos	California Consumer Privacy Act	01/01/2020	Direito de saber: os consumidores devem ser informados sobre quais informações pessoais uma organização coleta sobre eles e como essas informações são usadas. O direito de excluir: com algumas exceções, os consumidores podem excluir as informações coletadas sobre eles. O direito de recusar: os consumidores podem evitar a venda de suas informações a terceiros. O direito de não discriminação: uma organização não pode tratar os usuários que exercem seus direitos de forma diferente	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	A lei é de 2018, mas só entrou em vigor em 2020
Américas	México	Ley Federal de Protección de Datos Personales en Posesión de los Particulares	06/07/2020	Estabelecer procedimentos fundamentais para garantir o direito de todas as pessoas de proteger suas informações pessoais na posse de outros	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	
Ásia	China	GB/T 35273-2020	01/10/2020	Estabelecer diretrizes sobre transparência, direitos do titular e consentimento	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	

continua

continuação

Continente	País	Legislação	Data legislação	Objetivos	Limites	Observações
Ásia	Indonésia	Regulação n.º 82 de 2012 Lei n.º 19 de 2016 Regulação n.º 20 de 2016 Lei n.º 11 de 2018 Lei n.º 27 de 2022	17/12/2022	Processamento de dados pessoais de acordo com as bases legais; Consentimento do titular dos dados; Necessidade contratual e interesse legítimo; Medidas de segurança adequadas para proteger os dados pessoais; Direitos dos titulares dos dados, de acesso, correção e exclusão dos seus dados pessoais; Transferência de dados pessoais para fora da Indonésia	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	Lei de Proteção de Dados Pessoais (Lei PDP) é a Lei 27 de 2022. Lei de Informação e Transações Eletrônicas é Lei nº 11 de 2018. Das demais legislações complementam a Lei nº 11 de 2018
Ásia	Índia	Lei 21 de 2000 Regras de Tecnologia da Informação de 2011 Lei de Proteção de Dados Pessoais 2019 (revogada)	Não foi encontrada tal informação	Regras específicas sobre como proteger dados pessoais; Requerimentos para garantir a privacidade de dados; Regras voltadas para a coleta e o tratamento de dados pessoais nos setores bancário e de saúde	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	De acordo com a Lei de Privacidade de 2019, empresas teriam que pedir uma solicitação explícita para utilizar ou armazenar os dados dos cidadãos. Em 03 de agosto de 2022 a lei foi revogada
Ásia	Filipinas	10173/12	08/09/2012	Garantir o direito dos titulares de saber a finalidade da coleta de dados, quem fará o manuseio e quem terá o acesso aos dados coletados	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	Conhecida como Lei de Privacidade de Dados a Legislação foi elaborada em 2012, mas entrou em vigor em 2016
Ásia	Japão	Act on the Protection of Personal Information – APPI de 2003 APPI de 2017	15/05/2003	Consentimento: Os operadores devem obter o consentimento adequado dos titulares de dados antes de coletar, usar ou divulgar suas informações pessoais; Propósito Específico: Os operadores só podem usar informações pessoais para os fins específicos para os quais foram coletadas; Segurança: Os operadores devem implementar medidas técnicas e organizacionais para proteger as informações pessoais contra acesso não autorizado, perda, alteração ou divulgação	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	APPI de 2017 é a Emenda da APPI de 2003. Altera o tratamento de dados pessoais por empresas públicas e privadas, e traz diretrizes para proteger os titulares.
Ásia	Malásia	Personal Data Protection Act	02/06/2010	Coleta, o uso e a divulgação de dados pessoais; Direitos aos titulares dos dados, incluindo direitos de acesso, correção, desistência de consentimento e objeção a determinados tipos de processamento	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	
Europa	Alemanha	Bundesdatenschutzgesetz – BDSG	2017	Estabelecer que órgãos públicos e privados devem contratar um profissional responsável por privacidade de dados; Estabelecer regras claras para avaliações de score de crédito; Criar diretrizes específicas para o tratamento de dados de funcionários	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	

continua

conclusão

Continente	País	Legislação	Data legislação	Objetivos	Limites	Observações
Europa	Dinamarca	Não foi encontrada a lei, apenas trabalhos sobre ela	2018	Estabelecer regra para: Processamento de dados; Divulgação de informações pessoais; Consentimento; Transferência de dados; Estabelecimento de multas para casos de violação	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	Complementa e reforça a GDPR de 2018 (<i>General Data Protection Regulation</i>) Legislação de proteção de dados da União Europeia
Europa	Grécia	2.472/1997 3.471/2006 3.917/2011	1997 2006 2011	Criar princípios sobre: transparência; tratamento de dados com propósito e responsabilização dos agentes	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	Ainda não há legislação em vigor Adequada à GDPR de 2018
Europa	França	2016-1321	2016	Estabeleceu as seguintes diretrizes: Direito do usuário saber quem é o agente de tratamento e direito do usuário saber o propósito da coleta desses dados	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	Complementa e reforça a GDPR de 2018 (<i>General Data Protection Regulation</i>) Legislação de proteção de dados da União Europeia
Europa	Finlândia	Lei 759 de 2004 Lei 917/2014 HE 9/2018	2004 2014 2018	Definir os seguintes direitos aos titulares: Acessar os seus dados pessoais; Solicitar correções, anonimização, bloqueio ou eliminação dos seus dados; Solicitar a transferência dos seus dados pessoais a terceiros; Saber com quais entidades os seus dados são compartilhados; Apresentar oposição ao tratamento dos seus dados	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	A legislação HE 9/18 revoga a legislação de 199 e complementa e reforça a GDPR de 2018 (<i>General Data Protection Regulation</i>) Legislação de proteção de dados da União Europeia
Europa	União Europeia	General Data Protection Regulation - GDPR	25/05/2018	Estabelece princípios sobre transparência; licitude; minimização de dados; responsabilidade; integridade	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	É a lei mais completa e complexa sobre o assunto. A legislação europeia exige que os países da União Europeia atualizem ou ratifiquem em suas leis próprias para que estejam alinhadas às regras da GDPR
Oceania	Austrália	Privacy Act 1988	04/11/1988	Regulamentar o uso e divulgação de dados; os direitos do titular dos dados; a manutenção da qualidade dos dados e a transparência e anonimidade		
Oceania	Nova Zelândia	Privacy Act	01/12/2020	Estabelecer princípios para a coleta, uso e gerenciamento de informações pessoais	Restringiu o tratamento de dados aos casos previstos em lei. Trouxe o instituto do consentimento	Revogou a Lei de Privacidade de 1993

Fonte: elaboração própria.

Ao analisar as legislações que fazem parte do cenário mundial acerca da proteção de dados, verifica-se que as principais diferenças entre as leis de proteção de dados mencionadas no documento são:

Base Legal: cada país tem uma base legal diferente para a coleta e tratamento de dados. Por exemplo, a GDPR da União Europeia tem seis bases legais, enquanto a PIPEDA do Canadá opera sobre dez princípios de boas práticas.

Consentimento: a exigência de consentimento do usuário varia. Na Argentina, a coleta de dados só pode ser feita com o consentimento do usuário, enquanto nos Estados Unidos, algumas leis permitem a coleta de dados sem consentimento em certos setores.

Direitos do Titular: os direitos concedidos aos titulares dos dados também diferem. A GDPR garante direitos como acesso, correção, exclusão e portabilidade dos dados, enquanto a Lei de Privacidade da Califórnia (CCPA) garante direitos de notificação, acesso, opção de coleta e acesso igualitário a serviços.

Aplicação: a aplicação das leis pode ser nacional ou setorial. Nos Estados Unidos, existem leis estaduais e setoriais, enquanto a GDPR se aplica a todos os países da União Europeia e a qualquer empresa que processe dados de cidadãos da UE.

Multas e Penalidades: as multas e penalidades por violação das leis variam. A GDPR estabelece multas significativas para violações, enquanto outras leis podem ter penalidades menos severas.

Transferência Internacional de Dados: as regras para transferência internacional de dados também diferem. A GDPR tem diretrizes rigorosas para transferências internacionais, enquanto a PIPEDA do Canadá permite transferências internacionais e inter-regionais de dados pessoais com base em princípios de boas práticas.

Ainda sobre o tema, destaca-se que as leis de proteção de dados ao redor do mundo têm recebido diversas críticas, tanto por especialistas quanto por controladores e usuários. Muitos controladores, criticam a complexidade e o custo de conformidade com essas leis.

A implementação de sistemas e processos para garantir a conformidade pode ser onerosa e exigir recursos significativos. Algumas leis, como a GDPR, são criticadas por sua ambiguidade em certos aspectos, o que pode levar a diferentes interpretações e criar incertezas para os controladores sobre como exatamente cumprir os requisitos legais.

Há preocupações de que regulamentações rigorosas possam limitar a inovação, especialmente em áreas como inteligência artificial e *big data*, dificultando o desenvolvimento de novas tecnologias e serviços. A aplicação das leis pode ser desigual, com alguns controladores enfrentando penalidades severas enquanto outros escapam de sanções, criando um ambiente de concorrência desleal.

Embora os direitos dos titulares de dados sejam uma parte importante dessas leis, há críticas sobre a dificuldade de implementação prática desses direitos, com controladores enfrentando desafios para atender a solicitações de acesso, correção e exclusão de dados de forma eficiente. As regras rigorosas para transferência internacional de dados, especialmente sob a GDPR, são criticadas por dificultar operações globais, exigindo que controladores naveguem por um labirinto de requisitos legais para transferir dados entre países. As multas significativas por violação das leis, como as previstas pela GDPR, são vistas como excessivas por alguns controladores, que temem que penalidades severas possam prejudicar financeiramente aqueles que cometem erros inadvertidos.

Algumas críticas apontam para um conflito entre privacidade e segurança, no qual a proteção rigorosa de dados pode, em alguns casos, dificultar a implementação de medidas de segurança eficazes. Essas diferenças refletem as diversas abordagens e prioridades de cada país em relação à proteção de dados e privacidade.

1.3.2.4 Regulamentação no Brasil

A proteção de dados no Brasil teve início com o artigo 5º da Constituição Federal de 1988, que protege a inviolabilidade da intimidade, vida privada, honra e imagem das pessoas, garantindo indenização em casos de violação desses direitos. Antes da criação da LGPD, não existia uma lei específica que tratasse da proteção de dados pessoais de forma abrangente. O assunto era abordado por várias legislações dispersas, como o Código de Defesa do Consumidor, a Lei do Cadastro Positivo, a Lei de Acesso à Informação e o Marco Civil da Internet (Lei nº 12.965 e Decreto nº 8.771/16).

A LGPD, que foi fortemente influenciada pela GDPR (Regulamento Geral sobre a Proteção de Dados da União Europeia), trouxe avanços significativos para a regulamentação da privacidade e proteção de dados no país. Com sua implementação, outras legislações foram editadas para complementar e esclarecer o tratamento de dados pessoais, estabelecendo diretrizes claras e consistentes. As principais leis que atualmente regulam essa área no Brasil incluem não apenas a LGPD, mas também outras normativas e regulamentos que visam assegurar a transparência e segurança no tratamento dos dados.

- Lei n.º 13.709/2018 - Lei Geral de Proteção de Dados Pessoais

A legislação vigente no Brasil sobre o tratamento de dados pessoais visa proteger os direitos fundamentais de liberdade, privacidade e o livre desenvolvimento da personalidade da pessoa natural. Seu foco é criar um cenário de segurança jurídica, padronizando regulamentos e práticas para promover a proteção dos dados pessoais de todos os cidadãos no País, em conformidade com os parâmetros internacionais.

A lei foi aprovada com os seguintes vetos:

- Autoridade Nacional de Proteção de Dados

O presidente da República na ocasião, Michel Temer, vetou a criação da Autoridade Nacional de Proteção de Dados depois de consultar os Ministérios da Justiça; da Fazenda; da Transparência, Fiscalização e Controladoria-Geral da União; do Planejamento, Desenvolvimento e Gestão; da Segurança Pública; da Ciência, Tecnologia, Inovações e Comunicações; e o Banco Central. A ANPD seria vinculada ao Ministério da Justiça e teria função de órgão regulador para fiscalizar as normas da nova lei, além de aplicar sanções.

Na razão para o veto, o presidente Temer alegou que a criação da ANPD implicaria inconstitucionalidade do processo legislativo por trazer vício de iniciativa (a criação teria que partir do Executivo Federal). Mas o presidente já sinalizou que concorda no mérito com a criação do órgão, e que enviará um projeto de lei para tal.

Foi vetada ainda a implantação do Conselho Nacional de Proteção de Dados Pessoais e da Privacidade, que auxiliaria a ANPD. A justificativa para o veto também alegou vício de iniciativa.

- Dados compartilhados com o poder público

Outro mecanismo vetado era o que obrigava a publicidade de dados pessoais compartilhados com entidades de direito público. Temer alegou que “a publicidade irrestrita da comunicação ou do uso compartilhado de dados pessoais entre órgãos e entidades de direito público, imposta pelo dispositivo, pode tornar inviável o exercício regular de algumas ações públicas como as de fiscalização, controle e polícia administrativa.” (Brasil, 2018b, art. 28).

- Suspensão ou proibição de banco de dados

Foi vetada a suspensão ou proibição de atividades de bancos de dados devido a violações à nova lei, visando evitar insegurança e consequências negativas nas operações financeiras.

- Lei n.º 13.853/2019

Esta lei alterou a Lei n.º 13.709/2018 (LGPD), estabelecendo diretrizes sobre a proteção de dados pessoais e instituindo a Autoridade Nacional de Proteção de Dados (ANPD). Essas modificações ocorreram antes da entrada em vigor da LGPD.

- Lei n.º 14.010/2020
- O art. 20 dessa lei adiou a aplicação das penalidades administrativas e multas previstas na LGPD para 1º de agosto de 2021, devido à pandemia da covid-19
- Portaria n.º 212, de 15 de outubro de 2020

Estabelece um Grupo de Trabalho encarregado de realizar estudos e sugerir propostas para adaptar os tribunais aos requisitos da Lei Geral de Proteção de Dados. Esse documento foi revogado pela Portaria n.º 362 em 14 de outubro de 2022, após o término de sua vigência.

- Portaria n.º 213, de 15 de outubro de 2020
- Cria o Comitê Gestor da Lei Geral de Proteção de Dados Pessoais (CGLGPD) no âmbito do Conselho Nacional de Justiça, com a missão de coordenar e supervisionar a aplicação da lei nos órgãos judiciais.
- Resolução n.º 334/2020, do Conselho Nacional de Justiça
- Institui o Comitê Consultivo de Dados Abertos e Proteção de Dados no Poder Judiciário, especificando suas responsabilidades, que incluem fornecer orientações e desenvolver políticas relacionadas à proteção de dados e à transparência dos dados abertos.
- Recomendação n.º 73/2020, do Conselho Nacional de Justiça
- Insta os órgãos do Poder Judiciário a tomarem medidas iniciais e adotarem preparativos para se adequarem às exigências da Lei Geral de Proteção de Dados Pessoais. Essa recomendação foi posteriormente modificada pela Recomendação n.º 89 em 24 de fevereiro de 2021.
- Resolução n.º 363/2021, do Conselho Nacional de Justiça
- Estabelece diretrizes e providências para a adaptação dos tribunais às

normas estipuladas pela Lei Geral de Proteção de Dados Pessoais, visando assegurar conformidade legal.

- MP 1.124/22 - Altera a Lei n.º 13.709, de 14 de agosto de 2018
- Medida Provisória que transforma a Autoridade Nacional de Proteção de Dados em uma autarquia de natureza especial e altera a estrutura dos cargos comissionados. Essa medida foi convertida na Lei n.º 14.460 em 25 de outubro de 2022.
- Emenda Constitucional n.º 115/2022
- Modifica a Constituição Federal para incluir a proteção de dados pessoais como um dos direitos e garantias fundamentais, conferindo à União competência exclusiva para legislar sobre proteção e tratamento de dados pessoais.
- Lei n.º 14.460/2022
- Resultante da Medida Provisória n.º 1.124/22, transforma a Autoridade Nacional de Proteção de Dados (ANPD) em uma autarquia de natureza especial e ajusta a composição dos cargos comissionados.
- Resolução STJ/GP n.º 5/2023
- Define as diretrizes para a implementação da Política de Proteção de Dados Pessoais no Superior Tribunal de Justiça (STJ), abrangendo medidas de segurança e privacidade.
- Resolução n.º 738/24 do Conselho Nacional de Saúde

Aprovada em 1º de fevereiro de 2024, essa resolução regula o uso de bancos de dados para fins de pesquisas científicas na área da saúde, especialmente aquelas envolvendo seres humanos.

Vale registrar que a nova regulamentação do Conselho Nacional de Saúde (CNS) estabelece diretrizes abrangentes para o uso de bancos de dados em pesquisas. Ela define claramente as responsabilidades tanto dos controladores quanto dos operadores desses bancos, além de listar os direitos fundamentais dos participantes das pesquisas. Entre os princípios básicos destacados estão a autodeterminação informativa, que permite aos indivíduos controlarem suas próprias informações, bem como a garantia de liberdade e privacidade dos participantes. A regulamentação orienta sobre como constituir e utilizar bancos de dados, sempre focando na proteção da dignidade e dos direitos dos indivíduos envolvidos.

De acordo com essa regulamentação, é permitido o tratamento de dados pessoais para fins de pesquisa científica por instituições autorizadas. No entanto, enfatiza-se a importância de garantir a anonimização dos dados sempre que possível, para proteger a identidade dos participantes. Os responsáveis pela gestão dos bancos de dados podem incluir patrocinadores das pesquisas, pesquisadores diretamente envolvidos no estudo, ou gestores de biobancos. Estes devem seguir rigorosamente os princípios estabelecidos pela Resolução do CNS, assegurando que todas as práticas estejam em conformidade com as normas vigentes.

A LGPD foi desenvolvida através de amplos debates e revisões posteriores, incluindo a criação da ANPD. Com o objetivo de adaptar a lei às necessidades práticas, foram introduzidas novas regulamentações, como a inclusão expressa da proteção de dados nos direitos fundamentais e ajustes necessários no sistema judiciário. Apesar desses avanços, identificou-se a necessidade de legislações complementares para cobrir lacunas deixadas pela LGPD. Um exemplo disso é a Resolução n.º 738/24 do CNS, que busca abordar especificamente o uso de bancos de dados em pesquisas científicas envolvendo seres humanos, oferecendo um *framework* legal mais robusto para essas atividades.

1.3.2.5 A posição dos tribunais

Nesta seção, será analisado como o poder judiciário se comporta em relação à proteção dos dados pessoais e à aplicação da LGPD. A investigação mostrou que a maioria das ações judiciais referentes ao tratamento de dados pessoais são iniciadas pelos próprios titulares desses dados, principalmente em situações em que houve divulgação sem seu consentimento. Os tribunais frequentemente determinam a remoção das informações divulgadas indevidamente e impõem penalidades aos responsáveis pela infração, geralmente envolvendo compensações por danos morais, atendendo assim às solicitações dos titulares dos dados.

A pesquisa indicou que esses processos judiciais são uma resposta direta a violações da privacidade e da segurança dos dados pessoais. Os tribunais têm tomado medidas rigorosas para garantir que as disposições da LGPD sejam cumpridas, enfatizando a importância do consentimento explícito dos titulares para qualquer tipo de uso ou divulgação de seus dados pessoais. Além de ordenar a retirada das informações, as decisões judiciais também costumam incluir sanções financeiras significativas como forma de compensar os danos sofridos pelos titulares dos dados e de desencorajar futuras violações.

- Ações Judiciais que mencionam artigos da LGPD em suas decisões

Até o momento, as ações judiciais que mencionam a LGPD em suas decisões têm os seguintes temas:

Direito Civil e Direito do Consumidor

- Assédio de empresas especializadas em negociação de dívidas a partir do uso de seus dados
- Pedido de indenização por danos relativos a fraudes resultantes de falhas de instituições financeiras na proteção de dados
- Decisões automatizadas

Direito do Trabalho

- A Justiça do Trabalho tem condenado empresas a indenizar funcionários — e até revertidos demissões por justa causa — por tratamento inadequado dos dados pessoais e violação à privacidade dos empregados
- Decisões automatizadas

Decisões automatizadas consistem em selecionar, classificar, aprovar ou rejeitar, atribuir notas, calcular riscos ou probabilidades com base em dados pessoais através de regras, cálculos, algoritmos, análises estatísticas, inteligência artificial ou aprendizado de máquina. O art. 20 da LGPD assegura o direito de revisão quando a decisão é tomada de forma automatizada.

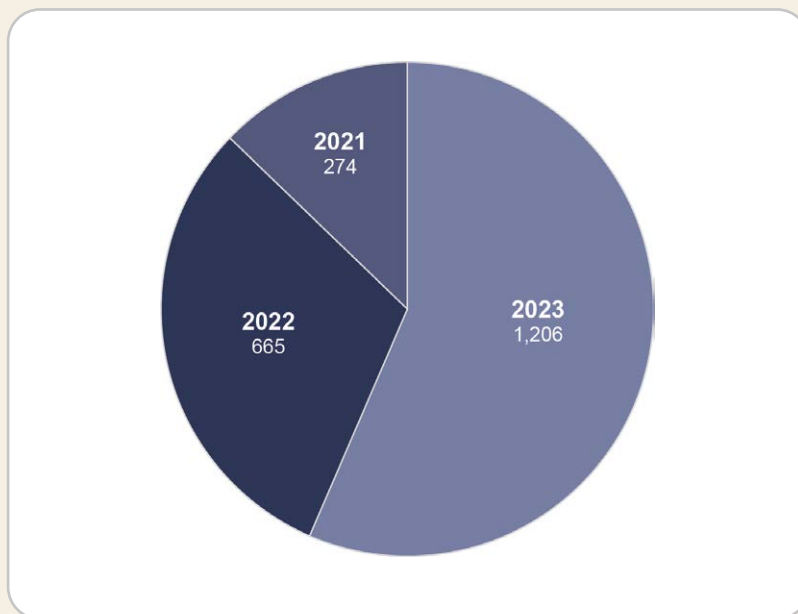
Decisões judiciais sobre acesso a dados em pesquisas acadêmicas

O Poder Judiciário

[...] ‘vem se manifestando pelo indeferimento de pedidos realizados por pessoa natural para o tratamento de dados pessoais para fins de pesquisa acadêmica’. O mesmo órgão reconhece que esse posicionamento pode inviabilizar a realização de trabalhos acadêmicos, ‘razão pela qual se busca alternativas legais, alicerçadas pelo órgão responsável por zelar, implementar e fiscalizar o cumprimento da LGPD em todo o território nacional’ (ANPD, 2022, p. 5).

A Figura 3 demonstra o aumento do número de processos judiciais que mencionam a LGPD.

Figura 3 – Aumento no número de processos judiciais que mencionam da LGPD em seus fundamentos



Fonte: Mendes; Fujimoto (2024).

- Principais temas das ações judiciais relacionadas à LGPD

Em ações judiciais movidas por empresas para proteger direitos garantidos pela LGPD, os temas abordados são diferentes daqueles nas ações iniciadas por indivíduos, como discutido anteriormente. Nessas situações, embora o consentimento seja mencionado, ele não está entre os principais tópicos, que incluem segurança, exclusão de dados, finalidade da divulgação e tratamento de informações.

A Figura 4 apresenta os temas das ações judiciais que mais mencionam a LGPD em seus fundamentos:

Figura 4 – Temas das ações judiciais que mais mencionam a LGPD em seus fundamentos

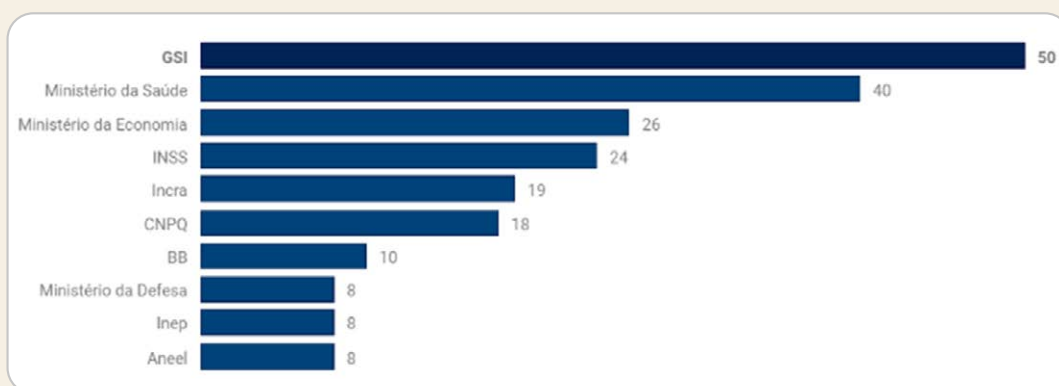


Fonte: Mendes; Fujimoto (2024).

1.3.2.6 Decisões de órgãos administrativos

Atualmente, os órgãos de controle que mais negam pedidos de acesso à dados são os representados na Figura 5.

Figura 5 – Os órgãos de controle que mais negam pedidos de acesso à dados



Fonte: elaboração própria, a partir de dados da Transparência Brasil.

Quando o motivo para o acesso aos dados tem por justificativas fins de pesquisa, as universidades públicas no Brasil vêm adotando o posicionamento de negar o acesso. A fim de ilustrar o aqui dito, destaca-se o texto para discussão n.º 1/2022 da ANPD sobre

A LGPD e o tratamento de dados pessoais para fins acadêmicos e para a realização de estudos por órgão de pesquisa. O estudo cita a decisão de uma universidade federal que adotou, por cautela, a postura de negar pedidos de acesso à informação para fins de pesquisa,

‘por considerar a inexistência de ato normativo regulamentador’ que disponha sobre a aplicabilidade do art. 7º, IV e do art. 11, II, c, da LGPD. Ainda segundo a mesma instituição, ‘o que se verifica é que a LGPD, ao tempo que desejou não impedir o desenvolvimento de pesquisas no país, desejou preservar os dados pessoais [...]. Entretanto, consideramos que tais questionamentos precisam de uma melhor orientação [...] sobre o procedimento a se realizar com relação aos pedidos de acesso à informação de dados pessoais e/ou sensíveis para fins de pesquisas’ (ANPD, 2022, p. 5).

Em regra, os órgãos controladores têm se utilizado do mesmo fundamento para negar acesso à dados para fins de pesquisa.

1.4 Reflexões finais

Durante a elaboração deste trabalho, confirmamos a hipótese inicial. A Lei Geral de Proteção de Dados (LGPD) alterou significativamente o acesso e o tratamento de dados pessoais no Brasil, mas ainda é um tema pouco explorado por cientistas e tribunais. A falta de estudos pode ser atribuída, em parte, ao fato de a lei ter entrado em vigor durante a pandemia da covid-19, o que retardou tanto sua aplicação quanto a pesquisa, principalmente na área da saúde, onde o tratamento de dados é crucial.

Desde a implementação da LGPD, várias legislações complementares foram propostas para preencher suas lacunas, contudo, nenhuma delas conseguiu resolver de maneira eficaz as questões éticas envolvidas. Isso se deve à natureza subjetiva da ética para os órgãos de controle e judiciais, o que torna desafiador o acesso aos dados nesse contexto e a formulação de recursos eficazes para contestar decisões negativas.

As ações judiciais relacionadas à LGPD geralmente são movidas principalmente por titulares de dados cujo tratamento foi realizado sem consentimento, concentrando-se nas áreas de direito do consumidor e justiça do trabalho. As poucas produções acadêmicas existentes sobre a LGPD tendem a explicar a legislação sem abordar seus efeitos práticos de forma aprofundada. Além disso, as negativas de acesso aos dados por parte dos órgãos controladores são comuns e frequentemente justificadas pela “ausência de legislação específica,” apesar de a LGPD tratar explicitamente do tema.

A ausência de decisões não implica necessariamente que o acesso ou tratamento de dados não esteja sendo negado, mas sim sugere que há uma lacuna no entendimento do tema que limita os direitos dos pesquisadores. Uma investigação detalhada sobre como a pandemia afetou a aplicação da LGPD seria altamente relevante, dado que a lei teve pouco debate nos primeiros anos de vigência e seu impacto completo ainda não foi plenamente compreendido.

Portanto, torna-se evidente a necessidade urgente de estudos mais profundos sobre a LGPD por parte de cientistas e do Poder Judiciário. À medida que a lei oferece diversas possibilidades de acesso a dados, essas oportunidades são frequentemente negadas devido à falta de conhecimento técnico e compreensão das suas implicações práticas. Assim, fomentar pesquisas e debates sobre a LGPD contribuirá significativamente para uma aplicação mais justa e eficaz da lei.



2 Descrição e análise de bases de dados

O Brasil possui uma das maiores forças de trabalho em enfermagem do mundo, com mais de três milhões de profissionais (Cofen, 2024). Esse contingente representa cerca de 70% dos trabalhadores da saúde do País (Mendes *et al.*, 2022), evidenciando sua relevância na elaboração de políticas públicas na área. Nesse contexto, aspectos como desigualdades salariais, variações regionais e a expansão de cursos de formação intensificam a necessidade de estudos aprofundados sobre as transformações do mercado de trabalho e suas implicações para a área da saúde (Salgado *et al.*, 2022; Inep, 2024b).

Embora o Brasil tenha demonstrado avanços importantes no registro dos trabalhadores do setor da saúde, os estudos na área ainda são escassos. Entre os existentes destacam-se Scheffer (2023), Machado (2012) e Nigenda *et al.* (2011). Para fazer frente à demanda por informações sobre o tema supramencionado, a criação de um repositório de dados com a arquitetura Data Lake, para análise integrada de bases de dados sobre o trabalho em saúde surge como estratégia fundamental para consolidar informações dispersas e permitir estudos mais robustos sobre demografia, mercado e condições de trabalho dos profissionais de enfermagem.

Este capítulo apresenta, de forma objetiva, o processo de construção de um repositório, detalhando as etapas de coleta, integração, tratamento e disponibilização dos dados que subsidiarão análises futuras sobre o setor. O projeto exigiu, entre outros esforços, a organização de dados provenientes de diferentes fontes, sejam de acesso público e aberto, ou de acesso restrito, cujos dados foram liberados após assinatura de termos de cooperação.

No Brasil, essas fontes são disponibilizadas por entidades públicas e privadas, e podem contribuir significativamente para a análise da situação e dinâmica dos mercados de trabalho e da educação em enfermagem em um período determinado, possibilitando também análises temporais. Essas bases de dados contêm informações valiosas sobre características demográficas, educacionais e ocupacionais, incluindo rendimentos, afastamentos, local de atuação, entre outras variáveis relevantes para a análise do mercado de trabalho em saúde (RIPSA, 2008).

O levantamento aqui apresentado incluiu a identificação das principais bases de dados governamentais de interesse à área da saúde, abrangendo coletas relacionadas a pesquisas sobre estabelecimentos de saúde, registros administrativos, registros censitários e gestão em saúde. Cabe ressaltar que o presente estudo não tem a pretensão de ser exaustivo, mas sim de apresentar a metodologia utilizada para a criação de um *Data Lake* destinado ao suporte de pesquisas integrantes do projeto de *Demografia e Mercado de Trabalho em Enfermagem*.

O foco está na identificação e descrição de fatores relacionados a estudos demográficos e à análise do mercado de trabalho em enfermagem. A fim de atingir esse macro objetivo, uma importante demanda técnica envolveu os processos de seleção, avaliação, armazenamento e manutenção dessas bases de dados, da forma mais completa e segura possível. Dessa forma, importa informar que, o projeto de *Demografia e Mercado de Trabalho em Enfermagem*, contou com assessoria jurídica especificamente em observância a Lei Geral de Proteção de Dados Pessoais (LGPD) de 2018.

E é nesse contexto que são registradas as etapas de construção e utilização do compêndio de dados, estruturados em cinco segmentos: (1) introdução, (2) metodologia; (3) descrição das bases de dados utilizadas, (4) as considerações finais e (5) Referências. A metodologia inclui o elenco das bases escolhidas. O tópico seguinte apresenta detalhadamente as características de cada base, organizadas da seguinte forma: (1) descrição, (2) origem dos dados, (3) estrutura, (4) restrições pela LGPD, (5) possibilidades de investigação e (6) limitações. Cada descrição inclui ainda links para o dicionário de dados¹ padronizado e para a análise exploratória elaboradas pelo projeto.

¹Um dicionário de dados é uma coleção de informações que detalha o significado, as características e o uso dos elementos de dados em um sistema, banco de dados ou projeto de pesquisa. Ele funciona como um glossário para dados, fornecendo o nome do registro, descrições, tipos, tamanho, restrições etc. garantindo a interpretação consistente e a precisão das informações por diferentes utilizadores. Por exemplo, se há necessidade de registrar o nome de um cliente, existirá uma definição de como este registro será feito no banco de dados. Esta definição dirá que o registro se chamará "nome_completo", uma descrição, "Nome completo do cliente", o tipo, "texto", o tamanho, "100" e uma restrição, "Não pode ser nulo".

2.1 Seleção das fontes de dados

A seleção das bases teve como critério primário o potencial para responder a 19 questões norteadoras do projeto:

1. Qual é a distribuição estatística e geográfica do número de vagas disponíveis anualmente para educação técnica em enfermagem no País?
2. Quais são as modalidades, padrões curriculares e relação docente/aluno dos cursos técnicos em enfermagem no País?
3. Quais são os mecanismos de credenciamento para instituições de ensino técnico em enfermagem no País?
4. Quais são as despesas per capita no investimento em formação técnica em enfermagem no País?
5. Qual é a distribuição estatística e geográfica das instituições de ensino em enfermagem credenciadas (graduação e curso técnico) no País?
6. Qual é a distribuição estatística e geográfica do número de vagas disponíveis anualmente para educação superior em enfermagem no País?
7. Qual é a distribuição estatística do número de vagas para educação superior em enfermagem no País?
8. Qual relação docente/aluno dos cursos de graduação em enfermagem no País?
9. Quais são as modalidades, padrões curriculares dos cursos de graduação em enfermagem no País?
10. Quais são os mecanismos de credenciamento para instituições de ensino em graduação em enfermagem no País?
11. Quais são as despesas per capita no investimento em graduação e técnico em enfermagem no setor de educação privado do País?
12. Qual é o perfil demográfico dos profissionais de enfermagem?

13. Como é a evolução do estoque/densidade de profissionais de enfermagem segundo cada categoria?
14. Qual é a distribuição geográfica dos profissionais de enfermagem ativos no período?
15. Os profissionais de enfermagem ativos têm se inserido no mercado de trabalho na mesma Unidade da Federação ou Município em que se formaram?
16. Qual é a distribuição dos profissionais de enfermagem por instituição pública e privada e por tipo de unidade segundo nível de atenção à saúde?
17. Qual é a projeção do estoque de profissionais de enfermagem?
18. Quantos enfermeiros especialistas ativos estão cadastrados na base de dados?
19. Qual é a proporção e distribuição geográfica de enfermeiros estrangeiros no País?

A partir dessas questões, a coordenação do projeto elegeu de forma empírica 14 bases de dados, conforme explicitadas no Quadro 4. Os repositórios de dados utilizados serão abordados ao longo deste capítulo. As bases estão detalhadas no Apêndice A. Os Apêndices B e C trazem aspectos mais técnicos da metodologia, utilizando um modelo de classes da *Unified Modeling Language* (UML).

Quadro 4 – Bases de dados selecionadas para pesquisa

Base de dados	Sigla	Responsabilidade
Cadastro Geral de Empregados e Desempregados	Caged	Ministério do Trabalho e Emprego (MTE)
Cadastro Nacional de Estabelecimentos de Saúde	CNES	Ministério da Saúde (MS)
Censo Demográfico	-	Instituto Brasileiro de Geografia e Estatística (IBGE)
Censo da Educação Básica	-	Instituto Nacional de Estudos e Pesquisas Educacionais (Inep)
Censo da Educação Superior	-	Instituto Nacional de Estudos e Pesquisas Educacionais (Inep)
e-Gestor Atenção Básica	-	Ministério da Saúde (MS)

continua

conclusão

Exame Nacional de Desempenho dos Estudantes	Enade	Ministério da Educação (MEC)
Instituto Nacional de Seguridade Social	INSS	Ministério da Previdência Social (MPS)
Pesquisa Nacional por amostra de domicílios contínua	PNAD contínua	Instituto Brasileiro de Geografia e Estatística (IBGE)
Pesquisa Nacional de Saúde	PNS	Instituto Brasileiro de Geografia e Estatística (IBGE)
Registros administrativos do Conselho Federal de Enfermagem	-	Conselho Federal de Enfermagem (Cofen)
Relação Anual de Informações Sociais	Rais	Ministério do Trabalho e Emprego MTE
Sistema de Informações de Mortalidade	SIM	Ministério da Saúde (MS)
Sistema de Seleção Unificada	Sisu	Ministério da Educação (MEC)

Fonte: elaboração própria.

O processo de seleção não se revelou exaustivo ou exclusivo. Os conjuntos de dados disponibilizados em cada uma das bases selecionadas foram acessados de forma concorrente, segundo critérios de seleção definidos *a priori*. À medida que os resultados de cada base se consolidavam, outras bases de dados e/ou plataformas foram avaliadas e estudadas seguindo a mesma dinâmica com o propósito de subsidiar o escopo do projeto.

2.2 Bases selecionadas

Aqui, são apresentadas uma descrição detalhada e uma análise crítica das bases de dados públicas e privadas relevantes para a força de trabalho em enfermagem no Brasil, destacando suas principais características, aplicações práticas e limitações. O objetivo é fornecer uma visão estruturada dessas fontes, de forma a permitir que pesquisadores, gestores e formuladores de políticas públicas compreendam como utilizá-las de maneira eficiente e informada em estudos e iniciativas relacionadas à área. Os Quadros 5, 6 e 7 sintetizam o detalhamento das informações considerando três macro dimensões: relevância para o estudo, periodicidade dos dados e características das bases.

Quadro 5 – Relevância para o estudo da força de trabalho em enfermagem

N.º	Base	Relevância para o estudo da força de trabalho em enfermagem		
		Perfil demográfico	Condições de trabalho	Educação
1	Caged	Sim	Sim	Sim
2	Censo Ed. Básica	Sim	Não	Sim
3	Censo Ed. Superior	Sim	Não	Sim
4	Censo Demográfico	Sim	Não	Não
5	CNES	Sim	Sim	Não
6	Cofen	Sim	Não	Não
7	e-Gestor	Não analisado	Não analisado	Não analisado
8	INSS	Sim	Sim	Não
9	PNAD Contínua	Sim	Sim	Sim
10	Enade	Não analisado	Não analisado	Não analisado
11	PNS	Não analisado	Não analisado	Não analisado
12	Rais	Sim	Sim	Sim
13	SIM	Não	Sim	Não
14	Sisu	Não analisado	Não analisado	Não analisado

Fonte: elaboração própria.

Quadro 6 – Dimensão temporal

N.	Base	Dimensão temporal		
		Periodicidade	Início	Fim
1	Caged	Mensal	jan/2007	set/2024
2	Censo Ed Básica	Anual	1995	2023
3	Censo Ed Superior	Anual	2009	2023
4	Censo Demográfico	Anual	2000	2022
5	CNES	Mensal	jul/2007	set/2024
6	Cofen	Diário	jul/1973	fev/2024
7	e-Gestor	Não analisado	Não analisado	Não analisado
8	INSS	Mensal	jan/2019	set/2024
9	PNAD Contínua	Trimestral	1T/2012	3T/2024
10	Enade	Não analisado	Não analisado	Não analisado
11	PNS	Não analisado	Não analisado	Não analisado
12	Rais	Anual	1985	2022
13	SIM	Mensal	1979	2024
14	Sisu	Não analisado	Não analisado	Não analisado

Fonte: elaboração própria.

Quadro 7 – Características das bases

N.	Base	Características das Bases		
		Acesso aos dados	Dados agregados	Menor granularidade local
1	Caged	Público	Não	Município
2	Censo Ed Básica	Público	Sim	Município
3	Censo Ed Superior	Público	Sim	Município
4	Censo Demográfico	Público	Sim	Município
5	CNES	Público	Não	Município
6	Cofen	Privado	Não	Bairro
7	e-Gestor	Não analisado	Não analisado	Não analisado
8	INSS	Público	Não	Município
9	PNAD Contínua	Público	Sim	Município
10	Enade	Não analisado	Não analisado	Não analisado
11	PNS	Não analisado	Não analisado	Não analisado
12	Rais	Público	Não	Município
13	SIM	Público	Não	Município
14	Sisu	Não analisado	Não analisado	Não analisado

Fonte: elaboração própria.

Nota: A PNAD contínua apresenta granularidade para municípios como capitais e metrópoles. A PNAD também apresenta alguns resultados com disponibilidade trimestral.

2.3 Avaliação da estrutura disponibilizada do conjunto de dados

Foi feita uma avaliação técnica para verificação de quais locais os conjuntos de dados poderiam ser acessados. Dependendo da base de dados pretendida, existem dois possíveis contextos de acesso. Se os dados em tela forem oferecidos de forma pública, é necessário saber se existe alguma restrição de acesso por parte dos guardiões dos dados.

No entanto, existem acessos mais simples, como, por exemplo, a plataforma de atendimento para acesso às bases da Relação Anual de Informações Sociais (Rais) e do Cadastro Geral de Empregados e Desempregados (Caged), via portal <http://gov.br>. Porém, existem bases como o Cadastro Nacional de Estabelecimentos de Saúde (CNES), oficialmente disponível no Portal CNES, do DataSUS, que pode

ser encontrada em diversas outras fontes (por acesso via protocolo FTP do próprio site do DataSUS, ou outros hospedeiros). Em cada site, há uma estrutura de CNES diferenciada, com subconjunto de atributos diferentes. Ocorre também frequentemente a existência de dados ausentes, conflitantes (um mesmo profissional sendo, por exemplo, do sexo masculino, feminino, indeterminado e nulo), valores fora da faixa válida, e outras anomalias que fragilizam sobremaneira a confiança no conjunto de dados. Entretanto, não há como prescindir de uma base como o CNES em pesquisas de demografia.

2.4 Avaliação da qualidade dos dados

Com o propósito de verificar a natureza dos dados obtidos, a etapa seguinte de avaliação das bases de dados acessadas é a sua qualidade. Assim, procede-se uma análise exploratória dos dados, em que se avalia a integridade e qualidade dos dados conforme ilustrado na Figura 6.

Figura 6 – Estrutura do relatório de análise exploratória dos dados

A	B	C	D	E	F	G	H	I	J	K	L	M
	Atributo	Cntde Preenchida	Cntde Null	% Preenchido	% Null	Valores Distintos	Erros	Em branco	Outliers	Outs		
5	5	7.762.247	29.977	99,61%	0,39%	32.195	4		4.281			
6	6	7.788.124	0	100,00%	0,00%	22.732	1.291.131	73.490				abaixo de 01/01/1925 ou acim
7	7	7.788.124	0	100,00%	0,00%	233	18.772	78.502				
8	8	7.788.124	0	100,00%	0,00%	118	4	32.152				
9	9	7.788.124	0	100,00%	0,00%	82	215.454	7.588.042				
10	10	7.788.124	0	100,00%	0,00%	9	309	4.854.991				
11	11	7.788.124	0	100,00%	0,00%	513.471	28	9.040				
12	12	7.788.124	0	100,00%	0,00%	689.989	965.746	1.374.339				
13	13	7.788.124	0	100,00%	0,00%	576	53.884	7.019.131				
14	14	7.788.124	0	100,00%	0,00%	143.293	198	114.182				
15	15	7.788.124	0	100,00%	0,00%	9.874	1.519.458	9.301				
16	16	7.788.124	0	100,00%	0,00%	79	1.712	6.586				
17	17	7.788.124	4.201	99,95%	0,05%	8	0	4.201				
18	18	7.783.847	4.277	99,95%	0,05%	13	11.566	4.277				
19	19	7.788.124	0	100,00%	0,00%	4	1	4.200				
20	20	7.788.124	0	100,00%	0,00%	7	4	4.277				
21	21	7.788.124	0	100,00%	0,00%	3.118	8.612	5.881.276				
22	22	7.783.843	4.281	99,95%	0,05%	8.687	4	4.281	112.415			abaixo de 01/01/1901 ou acim
23	23	7.788.124	0	100,00%	0,00%	29	1	7				
24	24	7.788.063	61	100,00%	0,00%	6	4	57				
25	25	7.788.124	0	100,00%	0,00%	2.018.501	0	279.558				
26	26	7.788.124	0	100,00%	0,00%	3.441	250.888	1.251.021				
27	27	7.788.124	0	100,00%	0,00%	20.016	752.843	1.245.676				
28	28	6.977.817	810.307	89,60%	10,40%	15.811	0	810.307	8.030			abaixo de 12/01/1973 ou acim
29	29	7.788.124	0	100,00%	0,00%	3	0	61				
30	30	6.656	7.781.468	0,09%	99,91%	814	0	7.781.468				
31	31	7.788.124	0	100,00%	0,00%	2.085.505	-323.195	872.045				
32	32	2.827.480	4.905.604	36,24%	63,76%	13.540	0	4.905.604				
33	33	7.788.124	0	100,00%	0,00%	9.023	-323.195	872.045				
34	34	7.788.124	0	100,00%	0,00%	21.675	300.337	895.309				
35	35	7.788.124	0	100,00%	0,00%	18.828	0	2.162.500				
36	36	2.173.635	5.614.589	27,91%	72,09%	12.806	0	5.614.589	50.556			abaixo de 12/01/1973 ou acim
37	37	7.788.033	91	100,00%	0,00%	15.521	0	91	26.296			abaixo de 12/01/1973 ou acim
38	38	7.788.024	100	100,00%	0,00%	10	1	100				não tem a opção 4 (parteira)
39	39	7.787.918	206	100,00%	0,00%	13	0	206				
40	40	7.788.124	0	100,00%	0,00%	4	1.172.127	88				

Fonte: elaboração própria.

Nesta figura se pode ver como está o conteúdo de cada campo. Na coluna “B” está o nome do conteúdo analisado enquanto as demais colunas mostram análises como a quantidade de registros preenchidos (coluna “C”), a quantidade de registro vazios (coluna “D”) etc. Esta informação permite que os pesquisadores possam entender a

qualidade de suas conclusões. Como exemplo, se no campo “idade” houver 10% de registros deixados em branco as conclusões a partir destes dados devem ser melhores que se houver 60% de registros deixados em branco.

Da mesma maneira que para os dicionários de dados, os resultados das análises exploratórias implementadas estão disponíveis em links no final da descrição de cada base de dados.

2.5 Produção do dicionário de dados de cada base de dados

Uma das primeiras providências que devem ser adotadas quando da proposta de utilização de um conjunto de dados é o conhecimento de sua estrutura. Assim, procuramos, como ação primária, acessar os metadados² das bases de dados candidatas à futura incorporação ao projeto. Entretanto, dada a diversidade de fontes de dados, temos as mais variadas ofertas: desde nenhuma (em que precisamos “descobrir”, observando os dados, o que cada atributo significa), até páginas na internet com dicionários de dados razoavelmente estruturados. Há ainda situações nas quais um mesmo conjunto de dados, por ser disponibilizado em vários locais, possuem catálogos diferenciados.

Com o objetivo de mitigar a heterogeneidade anteriormente informada, foi elaborada uma estrutura universal de dicionário de dados a ser útil a todas as bases, conforme proposta da coordenação do projeto. Essa estrutura, que busca contemplar a semântica máxima existente entre as bases de dados utilizadas, está ilustrada na Figura 7.

²Metadados são dados sobre outros dados (ou metainformação), como título, autor, data de criação, local e tamanho de um arquivo de computador. Estes metadados fornecem contexto e descrevem um recurso para facilitar sua organização, busca e gerenciamento.

Figura 7 – Exemplo da estrutura do dicionário de dados estendido

Cadastro_IES				
N	Nome da Variável	Descrição da Variável	Tipo	Domínio
1	NU_ANO_CENSO	Ano de referência do Centro de Educação Superior	Num	
2	NO_REGIÃO_IES	Nome da região geográfica da sede administrativa ou reitoria da IES	Char	
3	CO_REGIÃO_IES	Código da região geográfica da sede administrativa ou reitoria da IES	Num	
4	NO_UF_IES	Nome da Unidade da Federação da sede administrativa ou reitoria da IES	Char	
5	CO_UF_IES	Código da Unidade da Federação da sede administrativa ou reitoria da IES	Num	
6	NO_MUNICÍPIO_IES	Nome do Município da sede administrativa ou reitoria da IES	Char	
7	CO_MUNICÍPIO_IES	Código do Município da sede administrativa ou reitoria da IES	Num	
8	NO_CAPITAL_IES	Nome da cidade da sede administrativa ou reitoria da IES	Char	
9	CO_CAPITAL_IES	Código da cidade da sede administrativa ou reitoria da IES	Num	
10	NO_MESORREGIÃO_IES	Nome da Mesorregião da sede administrativa ou reitoria da IES	Char	
11	CO_MESORREGIÃO_IES	Código da Mesorregião da sede administrativa ou reitoria da IES	Num	
12	NO_MICRORREGIÃO_IES	Nome da Microrregião da sede administrativa ou reitoria da IES	Char	
13	CO_MICRORREGIÃO_IES	Código da Microrregião da sede administrativa ou reitoria da IES	Num	
14	TP_ORGANIZACAO_ACADEMICA	Tipo de Organização Acadêmica da IES	Num	1. Universidade 2. Centro Universitário 3. Faculdade 4. Instituto Federal de Educação, Ciência e Tecnologia 5. Centro Federal de Educação Tecnológica
15	TP_CATEGORIA_ADMINISTRATIVA	Tipo de Categoria Administrativa da IES	Num	1. Pública Federal 2. Pública Estadual 3. Pública Municipal 4. Privada sem fins lucrativos 5. Privada sem fins lucrativos 6. Privada - Particular em sentido amplo 7. Especial 8. Privada comunitária 9. Privada profissional
16	NO_MANTENEDORA	Nome da mantenedora da IES	Char	
17	CO_MANTENEDORA	Código único de identificação da mantenedora da IES	Num	
18	NO_IES	Nome da IES	Char	
19	CO_IES	Código único de identificação da IES	Num	
20	NO_ENDERECO_IES	Endereço da sede administrativa/reitoria da IES	Char	
21	CO_ENDERECO_IES	Código do endereço	Num	
22	NO_CAMPUS_ENTRADA_ENDERECO_IES	Componente do endereço	Char	
23	NO_BAIRRO_IES	Bairro	Char	
24	NO_CEP_IES	CEP	Num	
25	NO_TEC_TOTAL	Quantidade de Funcionários Nôcio administrativos	Num	

Fonte: elaboração própria.

É importante observar que houve a preocupação em registrar eventuais alterações estruturais ao longo do tempo nos esquemas dos conjuntos de dados. Infelizmente, não é incomum atributos serem descontinuados, criados, ou, o mais grave: continuarem semanticamente existindo, mas com nomes diferentes (exemplo: *id-profissional* em um ano, *cod-profissional* no ano seguinte). Tais mudanças, se não observadas com muita atenção, podem gerar significativos prejuízos aos pesquisadores e suas análises e os produtos da pesquisa.

Todos os dicionários de dados elaborados estão disponibilizados em links registrados ao fim da descrição de cada base de dados.

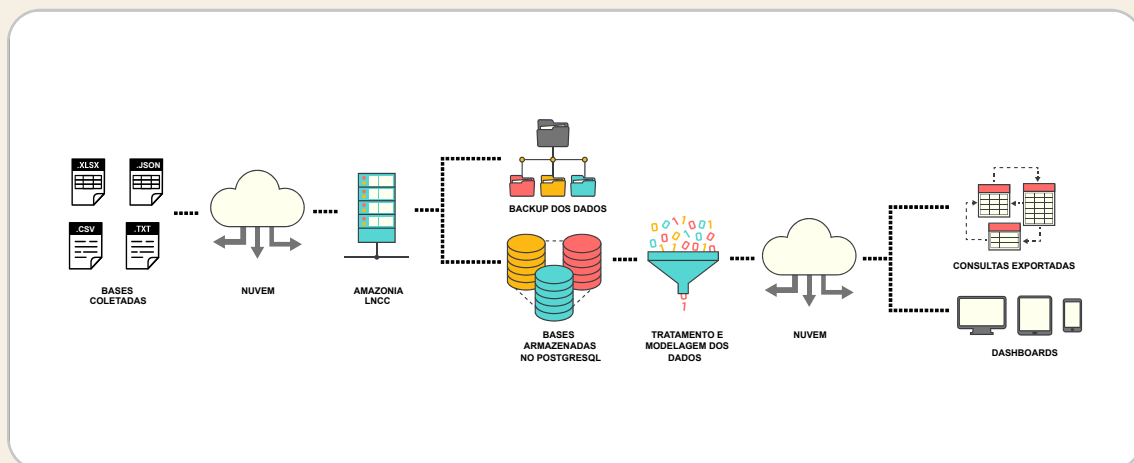
2.6 Carregamento dos dados nos servidores seguros do LNCC

As metas de produção de informação e conhecimento almejadas pelo projeto suscitam o armazenamento, tratamento, guarda e atualização de dados, processos por muitas vezes críticos, que podem evidenciar característica sensíveis e/ou que permitam identificação. A despeito de a maioria das bases de dados acessadas pelo projeto consistir em dados públicos, há uma responsabilidade social inerente ao projeto no

seu manejo. Existem situações nas quais os dados são obtidos por meio de parcerias institucionais, viabilizadas por assinatura de convênio institucional - fato que torna mais crítica a segurança demandada.

Nesse contexto, o projeto estabeleceu uma parceria com o Laboratório Nacional de Computação Científica (LNCC), instituto de pesquisa vinculado ao Ministério da Ciência, Tecnologia e Inovação (MCTIC), localizado na cidade de Petrópolis/RJ. Nessa parceria institucional, o referido laboratório disponibiliza ao projeto, sem ônus, um servidor – que consiste em um aparato computacional de um servidor *multicore*, ligado a um dispositivo de armazenamento (*storage*) com capacidade da ordem de *terabytes*. A arquitetura de tratamento e armazenamento de dados está demonstrada na Figura 8.

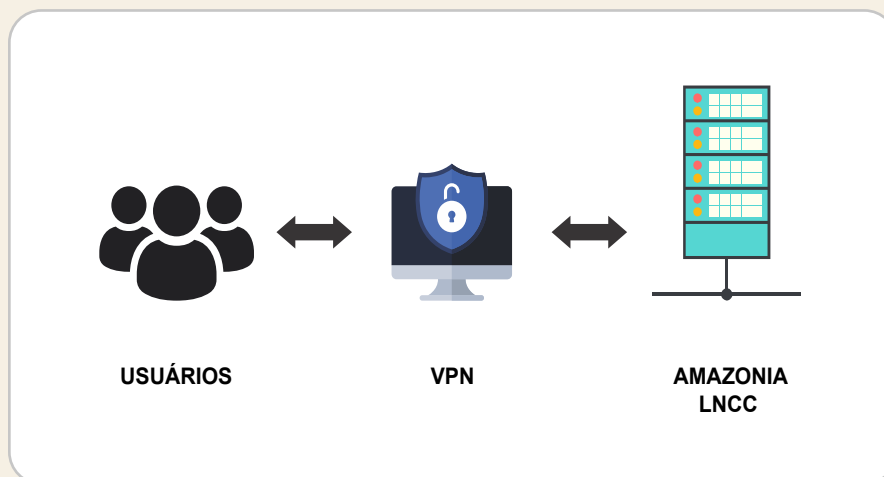
Figura 8 – Arquitetura de tratamento e armazenamento de dados do projeto, e sua relação com a infraestrutura do LNC



Fonte: elaboração própria.

Esse ambiente de alta capacidade de armazenamento de dados, parte deles sensíveis, deve garantir segurança tanto física quanto lógica. O LNCC possui certificação ISO 27001, norma internacional de requisitos para segurança da informação, que o gabarita a hospedar bases contendo dados sensíveis. O acesso aos dados é restrito aos analistas autorizados pela coordenação do projeto. Estes, por sua vez, estão sujeitos às normas estabelecidas por respectivos termos de confidencialidade, no manuseio desses dados. O acesso dos analistas é feito de forma presencial no laboratório ou remotamente, através de canais seguros de comunicação do tipo VPN (*Virtual Private Network*) conforme demonstrado na Figura 9.

Figura 9 – Esquema de comunicação entre o projeto e o LNCC



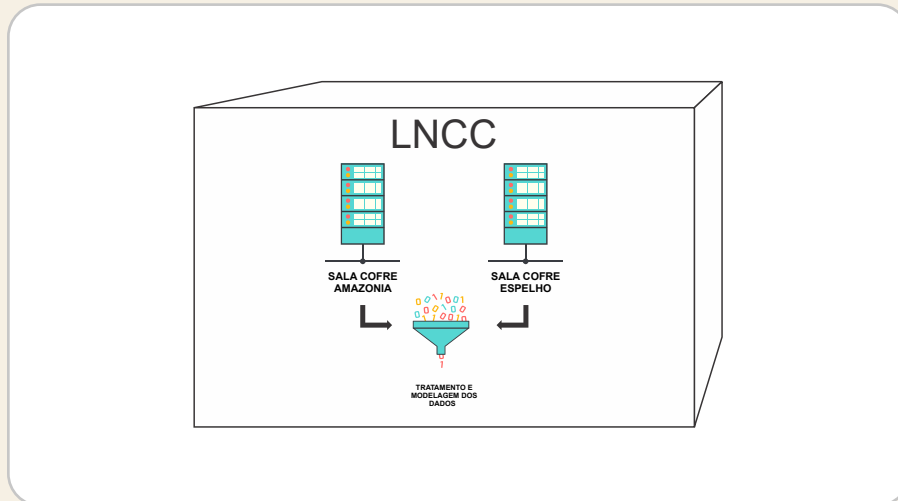
Fonte: elaboração própria.

O servidor armazena as bases de dados e possui um espelho funcionando como cópia de segurança. O ambiente físico onde essas máquinas estão hospedadas é equipado com controle de climatização e proteção contra surtos eletromagnéticos e incêndio.

O tratamento e modelagem dos dados requerem processamento de alto desempenho, em que muitos processadores atuam em paralelo viabilizando o processamento dessas bases em tempo hábil. O LNCC fornece esse serviço para o projeto.

Com conexões de internet de 10Gb/s, o laboratório oferece aos analistas do projeto acesso remoto eficiente às bases disponibilizadas, que são respectivamente carregadas, tratadas e modeladas por eles. A Figura 10 apresenta um esquema simplificado da infraestrutura de armazenamento no ambiente seguro AMAZÔNIA hospedada no LNCC. Apenas analistas autorizados do projeto gerenciam. As requisições enviadas aos analistas são retornadas, sob a forma de consultas exportadas (bases tratadas, anonimizadas) e/ou *dashboards* (painéis de visualizações consolidadas das bases).

Figura 10 – Infraestrutura de armazenamento



Fonte: elaboração própria.

2.7 Produção de informação para o consumo dos grupos temáticos de pesquisa

Com os dados carregados e disponibilizados na estrutura computacional do projeto, foi possível disponibilizar duas formas de produção de informação e conhecimento: a produção de *dashboards* com o cruzamento de atributos de cada base de dados, ou o atendimento de demandas específicas (*ad-hoc*) dos grupos de pesquisa temáticos associados ao projeto. No primeiro caso, os painéis são desenvolvidos utilizando a ferramenta *Microsoft Power BI*, associada ao sistema gerenciador de bancos de dados *PostgreSQL*. Diversas possibilidades de cruzamentos são oferecidas no *dashboard*. Com isso, seus usuários podem produzir várias informações de interesse sem o intermédio de terceiros.

Entretanto, demandas específicas não atendidas pelos painéis são passíveis de solicitação via coordenação do projeto, desde que com propósito definido. Este tipo de solicitação é o resultado de reuniões síncronas, seguidas ou não de discussão em fóruns assíncronos.

O projeto também produziu uma metodologia de estudo para a demografia e mercado de trabalho da enfermagem. Essa metodologia foi descrita em *Unified Modeling Language* (UML) e generalizada como ferramenta de estudo de qualquer demografia. Os Apêndices B e C apresentam o modelo teórico dessa generalização, que servirá

como uma importante ferramenta metodológica, no estudo de demografias da força de trabalho em saúde.

2.8 Possibilidade de futura integração entre as bases de dados

Um dos maiores desafios enfrentados por quem trabalha com dados é a possibilidade de cruzamento entre bases de dados. Do ponto de vista técnico, para que essa junção aconteça, cada conjunto de dados precisa conter um identificador da pessoa física ou do posto de trabalho, que seja comum a todas as bases. Por exemplo: entre o conjunto de atributos do CNES, o número do Cartão Nacional de Saúde (CNS) consta como um identificador. No entanto, esse número não consta na base de dados Rais, outro conjunto de dados importante aos objetivos do projeto.

É consenso que, no estágio tecnológico e legal atual, a única possibilidade técnica de integração é a disponibilização do número do Cadastro de Pessoas Físicas (CPF) da Receita Federal. Apesar de não ser considerado pela LGPD como um dado pessoal sensível, observa-se um cuidado na sua divulgação, dada a sua capacidade de identificação derivada de diversos outros dados da pessoa física. Entretanto, considerando os propósitos do projeto, assim como a relevância institucional das entidades envolvidas, busca-se a disponibilização, por meio de convênios, desse identificador no maior número de bases possível.

2.9 Reflexões finais

Apresentamos o processo de identificação, seleção e organização das bases de dados consideradas essenciais para uma análise abrangente da demografia e do mercado de trabalho em enfermagem no Brasil. Ao longo do texto, evidenciou-se a relevância de cada uma das etapas executadas como suporte para subsidiar decisões estratégicas, fortalecer políticas públicas baseadas em evidências e aprimorar a compreensão da força de trabalho em saúde.

Dada a relevância desse grupo de trabalhadores, que constitui a maioria da força de trabalho em saúde no Brasil, compreender sua dinâmica ocupacional é fundamental para o planejamento estratégico e a formulação de políticas públicas que assegurem a sustentabilidade do sistema de saúde e a equidade no acesso aos serviços.

O estudo evidenciou inconsistências nos registros e a ausência de segmentações específicas que limitam a profundidade das análises e que podem comprometer o embasamento de políticas públicas. Dados mais confiáveis e robustos são indispensáveis não apenas para diagnósticos mais precisos, mas também para a formulação de estratégias de gestão baseadas em evidências.

A partir das possibilidades apresentadas, destaca-se a oportunidade de utilizar essas informações para responder a questões cruciais sobre a formação, distribuição e sustentabilidade da força de trabalho em enfermagem. Por exemplo, análises mais detalhadas sobre a relação entre oferta educacional e demanda de mercado podem direcionar iniciativas para reduzir desigualdades regionais. Além disso, investigações sobre a mobilidade dos profissionais e a influência das condições de trabalho na retenção da força de trabalho podem fundamentar políticas que promovam maior estabilidade e valorização da categoria.

O estudo também destacou a importância de consolidar e integrar as bases de dados selecionadas, visando a criação de um ambiente seguro e robusto para análise integrada das informações sobre a força de trabalho em enfermagem. Essa integração é fundamental para o monitoramento contínuo de tendências, padrões e transformações no mercado de trabalho, permitindo análises mais abrangentes sobre a distribuição, formação e condições de trabalho dos trabalhadores da enfermagem. Ademais, a adoção de metodologias inovadoras e o fortalecimento de parcerias institucionais, como exemplificado na cooperação com o LNCC, ampliam a capacidade de armazenamento seguro, análise exploratória e disponibilização de dashboards customizados, facilitando respostas mais precisas e efetivas às necessidades do setor.

Assim, o presente estudo pretende ser um marco inicial para futuras iniciativas que busquem não apenas compreender, mas também transformar o mercado de trabalho em enfermagem no Brasil. Essas transformações são fundamentais para a sustentabilidade do sistema de saúde e para a melhoria contínua da qualidade e equidade nos serviços prestados à população.



3 Um estudo da modelagem da Força de Trabalho baseado na dinâmica de sistemas

Com sua ampla extensão territorial e diversidade socioeconômica, o Brasil se apoia fortemente nos profissionais de enfermagem para garantir o acesso à saúde em vários graus de complexidade, desde a assistência básica até os atendimentos especializados em instituições hospitalares. No entanto, a administração eficaz dessa mão de obra se depara com desafios como a escassez de profissionais em determinadas áreas, a alta rotatividade, as condições de trabalho contingentes e a formação deficiente para atender às demandas do mercado de trabalho.

As enfermeiras são essenciais para cumprir a promessa de 'não deixar ninguém para trás' e para o esforço global de alcançar os Objetivos de Desenvolvimento Sustentável (ODS). Elas contribuem de forma central para as metas nacionais e globais relacionadas a uma série de prioridades de saúde, incluindo cobertura universal de saúde, saúde mental e doenças não transmissíveis, preparação e resposta a emergências, segurança do paciente e a oferta de cuidados integrados e centrados nas pessoas (OMS, 2020a, p. 2, tradução nossa).

Neste contexto, entender como se relacionam as variáveis que influenciam o sistema é importante para escolher quais dados levantar e analisar, tanto para compreender os eventos, quanto para planejar políticas públicas. Modelos tradicionais de relacionamento entre as variáveis, como o proposto por Sousa *et al.* (2013), têm sido bem-sucedidos em expressar as macroestruturas do sistema de saúde. Todavia, falham em identificar e expressar as micro relações. Sterman (2000) destaca que uma modelagem baseada na dinâmica de sistemas facilita a compreensão de como as variáveis interagem e evoluem ao longo do tempo, permitindo a análise de políticas e estratégias em sistemas complexos.

A modelagem por meio da dinâmica de sistemas mostra-se como um instrumento sólido para examinar os fluxos e estoques que formam a força de trabalho de enfermagem, possibilitando entender como as variáveis estruturais se relacionam ao longo do tempo. A dinâmica dos sistemas se distingue por sua habilidade de captar a interdependência e a não linearidade dessas variáveis, como o efeito das políticas educacionais, inovação e condições laborais sobre o *turnover* de trabalhadores do setor. Além disso, possibilita a análise das interações que se estabelecem entre a disponibilidade de trabalhadores, a qualidade do serviço fornecido e a satisfação dos próprios trabalhadores.

Segundo Barreto, Carreira e Marcon (2015), a demanda por serviços de saúde, impulsionada pelo envelhecimento da população e pelo aumento da prevalência de doenças crônicas, intensifica a pressão sobre o sistema de saúde, exigindo um planejamento estratégico mais acurado para garantir a presença de profissionais competentes. Simultaneamente, as situações econômicas e políticas podem variar consideravelmente a atratividade da profissão de enfermagem, seja através do efeito nas oportunidades de aprendizado ou nas condições de trabalho e treinamento. Nessa perspectiva, o uso da dinâmica de sistemas possibilita a simulação de cenários de longo prazo e a análise de diversas políticas públicas, identificando áreas de alavancagem que podem aprimorar a distribuição de recursos humanos no campo da enfermagem.

O objetivo deste estudo é modelar a força do trabalho de enfermagem no Brasil por meio da metodologia de dinâmica de sistemas, a fim de examinar como os variados fatores – educacionais, econômicos, demográficos e políticos – influenciam a disponibilidade e a procura por profissionais nesta área. O modelo desenvolvido permitirá a exploração de cenários futuros que possam contribuir para a criação de políticas públicas mais precisas, garantindo a sustentabilidade e o equilíbrio entre a oferta de enfermeiros e as demandas de ascensão da população.

3.1 Metodologia

Este trabalho foi desenvolvido a partir de duas metodologias combinadas: descrição do modelo estático desenvolvido por Souza *et al.* (2013), seguida pelo desenvolvimento do modelo dinâmico desdobrando as variáveis conforme necessário para evidenciar as interações.

Para descrever essas interações foi utilizada a metodologia da Dinâmica de Sistemas. Segundo Sterman (2000), trata-se de uma metodologia eficaz para entender e modelar sistemas complexos, como o mercado de trabalho na área da saúde.

Esta abordagem permite registrar interações dinâmicas entre os elementos de um sistema, identificando relações causais, *loops* de feedback e *delays* que influenciam o comportamento do sistema ao longo do tempo, auxiliando na identificação de como políticas ou ações podem provocar efeitos não pretendidos devido às interconexões e aos comportamentos não lineares do sistema.

De acordo com Sterman (2000) o processo de modelagem via Dinâmica de Sistemas segue seis etapas fundamentais: 1) Articulação do problema; 2) Formulação de um modelo dinâmico; 3) Formulação de um modelo simulável; 4) Teste do modelo; 5) Formulação e avaliação de políticas; e 6) Implementação. Essas fases possibilitam entender e simular a dinâmica de sistemas complexos, auxiliando na tomada de decisões estratégicas e na sugestão de políticas efetivas para intervenções no sistema simulado. Neste trabalho serão abordados os passos 1 e 2.

3.2 Modelagem tradicional do mercado de trabalho

Em um setor tão dinâmico e vital como o da saúde, modelos bem desenvolvidos servem como ferramentas estratégicas para melhorar a alocação de recursos humanos e garantir que as demandas atuais e futuras da população sejam atendidas. Muitos são os modelos que buscam descrever a macroestrutura do mercado de trabalho em saúde, todavia, o modelo desenvolvido por Sousa *et al.* (2013) tem sido, atualmente, o mais utilizado, inclusive pela Organização Mundial de Saúde em seus estudos.

Um mercado de trabalho pode ser definido como o mecanismo que facilita a satisfação de uma demanda por serviços de trabalho por aqueles que desejam ou são capazes de fornecer esses serviços. A dinâmica entre o número e o tipo de empregos oferecidos no mercado e o número e o tipo de trabalhadores de saúde disponíveis e dispostos a aceitar uma vaga oferecida é fundamental para determinar o nível de emprego e a configuração do mercado de trabalho, assim como seus resultados (OMS, 2021, p. xxii, tradução nossa).

Essa força de trabalho, de acordo com o Health Labour Market Analysis Guidebook (OMS, 2021), tem papel central em todo sistema de saúde e é crucial para o atingimento da cobertura universal de saúde. O setor de saúde e assistência social é um segmento em expansão acelerada em diversos países; por exemplo, aproximadamente 1 em cada 10 postos de trabalho nos países da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) se encontram neste segmento. Contudo, vários países também enfrentam grandes desafios na força de trabalho da saúde, tais como a escassez de trabalhadores, uma combinação inapropriada de competências, uma

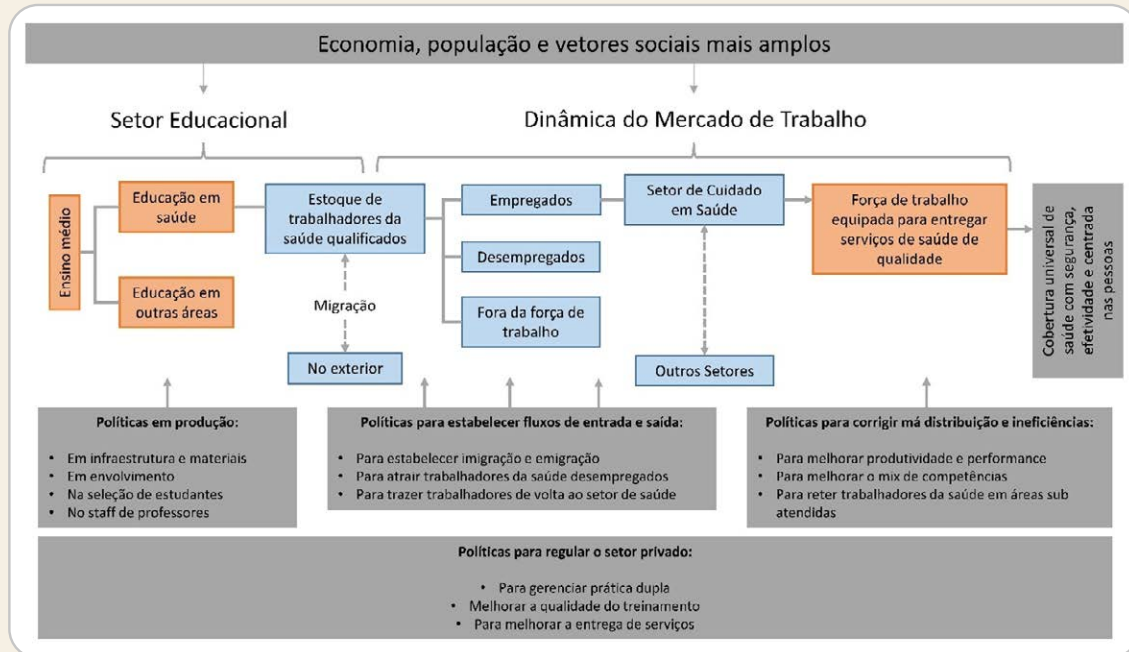
distribuição geográfica desigual de profissionais e deficiências em suas habilidades, motivação e desempenho.

Esses desafios, ainda segundo o Health Labour Market Analysis Guidebook (OMS, 2021), somados à escassez prevista de trabalhadores da saúde até 2030, principalmente em países de renda baixa e média baixa, travam o avanço rumo à cobertura universal de saúde, à segurança na saúde global e aos Objetivos de Desenvolvimento Sustentável relacionados à saúde. Nesse cenário, é crucial construir uma força de trabalho de saúde adequada ao propósito e pronta para a prática.

Para Sousa *et al.* (2013) tais problemas não se restringem apenas às nações de renda baixa e média baixa; muitos países de alta renda provavelmente enfrentarão uma grave falta de profissionais de saúde devido aos cortes no orçamento destinados aos serviços sociais causados pela crise econômica mundial. A população envelhecida exerce maior pressão sobre os sistemas de saúde, elevando a necessidade de cuidados médicos. Além disso, o constante movimento da mão de obra, como o crescimento do fluxo de trabalhadores de um país em desenvolvimento para outro, constitui um obstáculo para os mercados globais de trabalho na área da saúde.

Para abordar os desafios descritos e atingir a cobertura universal de saúde, os países terão que desenvolver políticas efetivas para otimizar o fornecimento de profissionais de saúde. Isso só pode ser alcançado por meio de um planejamento abrangente da força de trabalho em saúde com base em uma análise aprofundada do mercado de trabalho em saúde para entender as forças motrizes que afetam a oferta e a demanda de força de trabalho, tanto dentro dos países quanto em nível global (Sousa *et al.*, 2013, p. 892, tradução nossa).

Figura 11 – Modelo de Mercado de Trabalho para a Cobertura Universal de Saúde



Fonte: adaptado de Sousa *et al.* (2013).

A Figura 11 apresenta um *framework* do mercado de trabalho em saúde, dividido basicamente em duas partes: 1) Setor Educacional e 2) Dinâmicas do Mercado de trabalho, oferecendo uma visão abrangente das inter-relações entre a formação de profissionais, as políticas de alocação e retenção, e os desafios para atender à demanda por serviços de saúde. “Ambos são influenciados pela economia, pelo ambiente político, pela demografia da população e por outros fatores mais amplos da sociedade” (OMS, 2021, p. xxii, tradução nossa).

A imagem ilustra como o “Setor Educacional” e a “Dinâmica do Mercado” se relacionam para criar um fluxo constante de Recursos Humanos em Saúde (RHS). Ela ressalta, por exemplo, a importância das estratégias de inscrição e escolha de alunos, além das políticas de capacitação, fundamentais para garantir um número adequado de profissionais capacitados. Adicionalmente, a análise mostra a relevância das políticas de alocação para combater a má distribuição e as ineficiências, discutindo táticas para atrair trabalhadores desempregados e reter profissionais em regiões isoladas.

A Figura 11 nos possibilita uma visão unificada das diversas políticas e estratégias implementadas no setor de saúde, simplificando a compreensão dos processos e das possíveis ações para corrigir falhas estruturais. Combinadas, essas informações

constituem um alicerce analítico para debater como diversas intervenções podem afetar o mercado de trabalho na área da saúde, auxiliando na ampliação da cobertura dos serviços de saúde e aprimoramento da qualidade dos atendimentos prestados.

3.2.1 Setor educacional

A interação do setor educacional sugere que o planejamento e a administração das vagas nos programas de capacitação são fundamentais para prevenir tanto a falta quanto o excesso de profissionais em determinados setores. Um equilíbrio na distribuição de especialidades e níveis de formação garante que o sistema de saúde tenha uma força de trabalho adequada para atender à demanda. Assim, políticas de educação alinhadas às demandas do mercado de trabalho auxiliam na eficácia e sustentabilidade do sistema de saúde, garantindo que os serviços sejam de alta qualidade e acessíveis ao público.

A oferta e a demanda por educação de trabalhadores da saúde constituem um mercado que tem um forte impacto no mercado de trabalho da saúde e, conseqüentemente, no desempenho do próprio sistema de serviços de saúde. A oferta corresponde ao número de vagas em programas de educação básica e de treinamento em especialidades (OMS, 2021, p. 103, tradução nossa).

Segundo o *Health Labour Market Analysis Guidebook* (OMS, 2021), a oferta de vagas em programas educacionais de saúde é determinada por dois fatores principais: investimentos em infraestrutura, remuneração e condições de trabalho para docentes, e regulamentações por parte dos órgãos governamentais e conselhos profissionais que definem normas para a educação e, em alguns contextos, limitam o número de admissões ao ensino superior. Esses elementos impactam a quantidade de profissionais formados e a adequação da formação às demandas do setor de saúde.

Sousa *et al.* (2013) mostram que a situação econômica do país e suas prioridades influenciam diretamente a disponibilidade de vagas enquanto sindicatos e associações científicas frequentemente pressionam para restringir a entrada no mercado, com o objetivo de proteger o valor de suas profissões. Além disso, para o *Health Labour Market Analysis Guidebook* (OMS, 2021), *stakeholders*, tais como associações de pacientes e organizações do setor privado, exercem influência na busca por maior acesso a profissionais da saúde ou pela oportunidade de inserção no mercado educacional. A abertura de vagas para alunos estrangeiros, que geralmente pagam taxas mais elevadas, e as políticas de pressão para atender às necessidades de *stakeholders* influentes também têm um papel crucial na determinação da oferta educacional nesse setor.

Portanto a oferta de vagas para a formação de profissionais de saúde é um processo complexo, que vai além da simples resposta às necessidades do sistema de saúde e do interesse público. Ela é moldada por elementos econômicos e de pressão de *stakeholders*, que buscam impactar a quantidade de postos de trabalho e as condições de acesso ao mercado de trabalho. Dessa forma, uma avaliação das políticas de educação para profissionais de saúde deve levar em conta esses múltiplos interesses, que podem, em determinadas situações, colidir com as necessidades da população e com a igualdade na alocação de profissionais de saúde.

A demanda é impulsionada tanto por estudantes que buscam qualificação na área da saúde quanto por uma demanda secundária estimulada por governos que aumentam ou limitam o número de vagas de treinamento e quaisquer políticas associadas, como bolsas de estudo. Assim como a oferta, espera-se que a demanda por educação de profissionais de saúde esteja alinhada com a demanda por esses trabalhadores, derivada da demanda por serviços de saúde (OMS, 2011).

Ainda segundo o *Health Labour Market Analysis Guidebook* (OMS, 2021), a demanda por educação na área da saúde é influenciada também por fatores como prestígio profissional, altruísmo, tradição familiar, barreiras de acesso e o retorno financeiro e social esperado. Profissões de alto status, como a medicina, atraem mais candidatos, levando muitos a buscarem formação no exterior se não conseguem vaga em seus países. Aspectos como a satisfação pessoal em cuidar de outros e condições de mercado favoráveis, como bons salários, também impulsionam a escolha pela área da saúde, demonstrando que a decisão envolve tanto fatores econômicos e sociais quanto valores pessoais.

Esses elementos evidenciam que a decisão de seguir uma carreira na área da saúde é intrincada e multifacetada, englobando uma mistura de aspirações pessoais, estímulos financeiros e pressão da sociedade. Portanto, a procura por formação nessa área não é motivada apenas pelas necessidades do sistema de saúde, mas também por fatores de prestígio, contentamento pessoal e tradição familiar, que ajudam a definir o perfil dos profissionais em potencial e impactam diretamente a dinâmica do mercado de trabalho na área da saúde.

3.2.2 Mercado de trabalho

3.2.2.1 Oferta

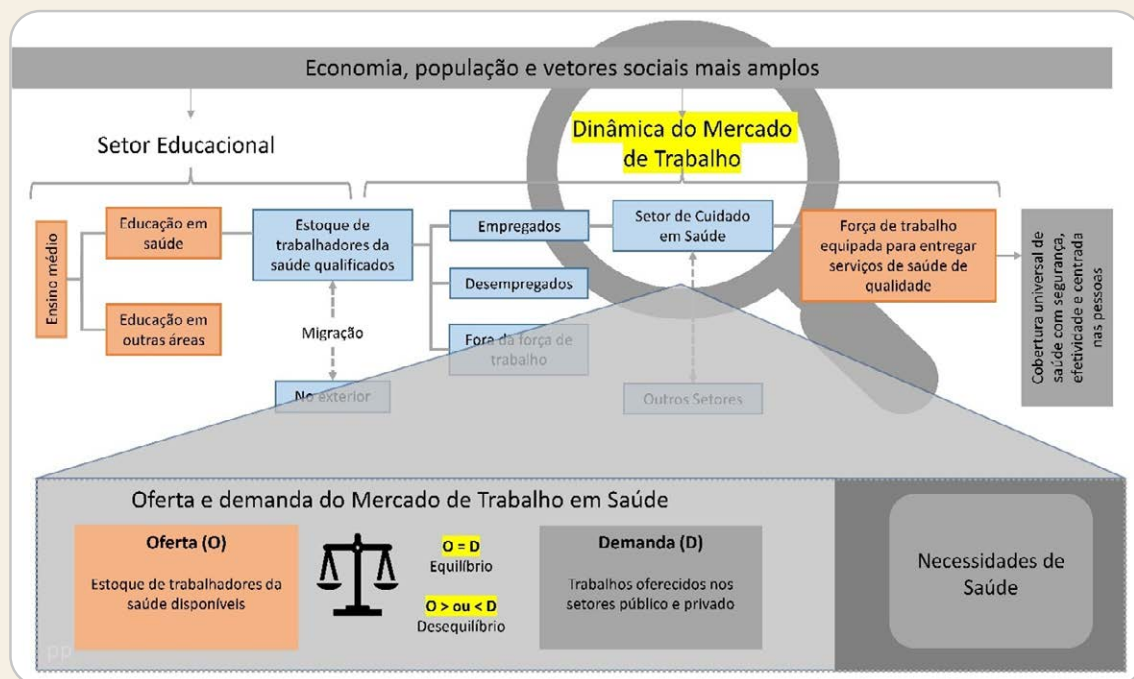
De acordo com o *Health Labour Market Analysis Guidebook* (OMS, 2021), os mercados de trabalho são vistos como equilibrados quando a oferta de trabalho é igual à procura. Ou seja, no ponto de equilíbrio salarial, a quantidade de profissionais de saúde prontos para trabalhar será igual à quantidade de postos de trabalho disponíveis. Quando o mercado de trabalho está em equilíbrio, não se nota nem excesso nem falta de trabalhadores.

Contudo, se esse equilíbrio não for atingido, podem ocorrer distorções, temporárias ou estruturais, resultado de uma incompatibilidade contínua entre a demanda por profissionais e sua oferta no mercado. Esta dinâmica destaca a necessidade de políticas eficientes para harmonizar essas variações e prevenir efeitos adversos no sistema.

A oferta de trabalhadores da saúde representa a quantidade de trabalhadores da saúde que estão dispostos a trabalhar com salários vigentes nas instituições de saúde de um país, incluindo hospitais, clínicas, farmácias, centros de diagnóstico, postos avançados de trabalhadores comunitários e centros de treinamento. No contexto do mercado de trabalho da saúde, o ponto de partida para entender a oferta é o conjunto de trabalhadores da saúde qualificados disponíveis (OMS, 2021).

As observações feitas acima estão representadas na Figura 12 a seguir:

Figura 12 – Dinâmica do Mercado de Trabalho dos Profissionais da Saúde (Oferta)



Fonte: adaptado de Sousa *et al.* (2013).

Para o *Health Labour Market Analysis Guidebook* (OMS, 2021), a oferta de trabalhadores de saúde é um componente essencial e dinâmico do mercado de trabalho em saúde, refletindo as interações entre fatores educacionais, econômicos e institucionais. Todavia, garantir um equilíbrio sustentável entre a formação e a distribuição de profissionais requer não apenas investimentos em infraestrutura educacional, mas também políticas estratégicas que abordem desigualdades geográficas, incentivem a retenção e promovam condições de trabalho dignas.

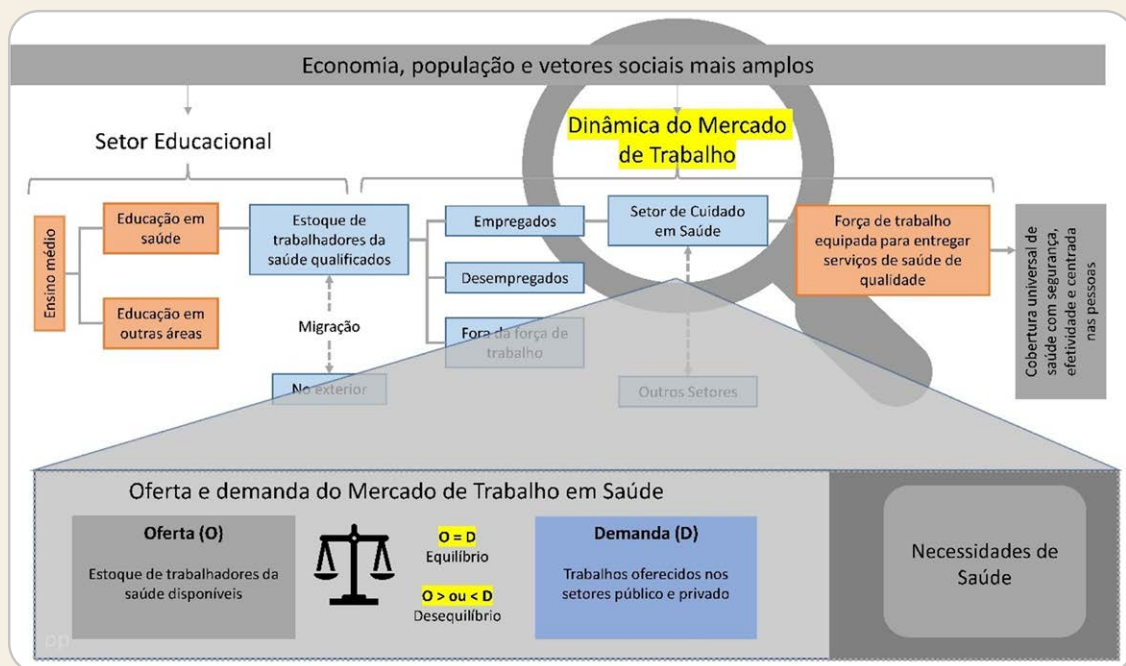
Dito isto, medir e monitorar a oferta por meio de indicadores claros e modelos previsionais é crucial para antecipar desafios e alinhar a força de trabalho às necessidades do sistema de saúde. Ao implementar essas estratégias, é possível fortalecer a capacidade do sistema de saúde em atender de forma equitativa e efetiva às demandas da população, reduzindo desigualdades e ampliando o acesso a serviços essenciais.

3.2.2.2 Demanda

Partindo para o outro prato da balança, “analisar a demanda por profissionais de saúde é essencial para entender melhor o mercado de trabalho em saúde, em particular a interação entre a oferta e demanda por profissionais de saúde, e para informar as políticas de força de trabalho em saúde” (OMS, 2021, p. 144, tradução nossa).

Ainda segundo o *Health Labour Market Analysis Guidebook* (OMS, 2021), a demanda por profissionais de saúde reflete a capacidade e a disposição de pagamento (do governo, setor privado, atores internacionais e outros) pela prestação de serviços de saúde, o que, por sua vez, impulsiona a necessidade de contratação de profissionais em unidades de saúde públicas ou privadas e outras partes do sistema, incluindo trabalhadores autônomos. Em outras palavras, a demanda por trabalho em saúde corresponde ao número de “vagas” para esses profissionais, considerando uma determinada faixa salarial e nível de produtividade.

Figura 13 – Dinâmica do Mercado de Trabalho dos Profissionais da Saúde (Demanda)



Fonte: adaptado de Sousa *et al.* (2013).

Essa dinâmica ressalta a relevância de aspectos econômicos e estruturais na determinação da demanda de profissionais da saúde. Por exemplo, durante períodos de recessão econômica, a capacidade de financiamento público/privado pode diminuir,

afetando a geração de empregos no ramo. Porém, em situações de grande necessidade, como em épocas de pandemias, a demanda por profissionais tende a aumentar, mesmo que momentaneamente, o que pode resultar em desequilíbrios consideráveis entre a oferta e a demanda. Portanto, é fundamental entender esses elementos para formular políticas que alinhem as necessidades de saúde da população aos recursos disponíveis.

Além de questões econômicas e estruturais, é importante destacar outros fatores ligados ao contexto social, demográfico etc.

[...] fatores contextuais, como aspectos sociodemográficos e epidemiológicos, influenciam diretamente a demanda por profissionais de saúde. O envelhecimento populacional, especialmente em países de alta renda e cada vez mais em países de média e baixa renda, aumenta a necessidade de serviços e profissionais relacionados, como enfermeiros para cuidados domiciliares. A demanda também é moldada pelo ambiente político, refletindo as prioridades dos tomadores de decisão no setor de saúde (OMS, 2021, p. 149, tradução nossa).

Para o *Health Labour Market Analysis Guidebook* (OMS, 2021), a demanda também é impactada pelo ambiente político, especialmente pelas prioridades determinadas pelos tomadores de decisão e outros *stakeholders* em relação ao setor de saúde e às políticas associadas, assim como pelos avanços científicos e tecnológicos, e sua aplicação na área da saúde, que estão mudando tanto o conteúdo quanto a organização dos serviços e seu desempenho. Outro ponto-chave

[...] para obter um quadro completo dos fatores que influenciam a demanda é essencial conhecer a perspectiva dos empregadores. Os empregadores não são um grupo homogêneo: eles podem ser públicos ou privados, com ou sem fins lucrativos, formais ou informais, e suas preferências podem variar consideravelmente. No setor da saúde, eles enfrentam grandes restrições, como um mercado de trabalho altamente regulamentado e grupos profissionais fortes, e uma demanda por serviços de saúde em rápida mudança (OMS, 2021, p. 151, tradução nossa).

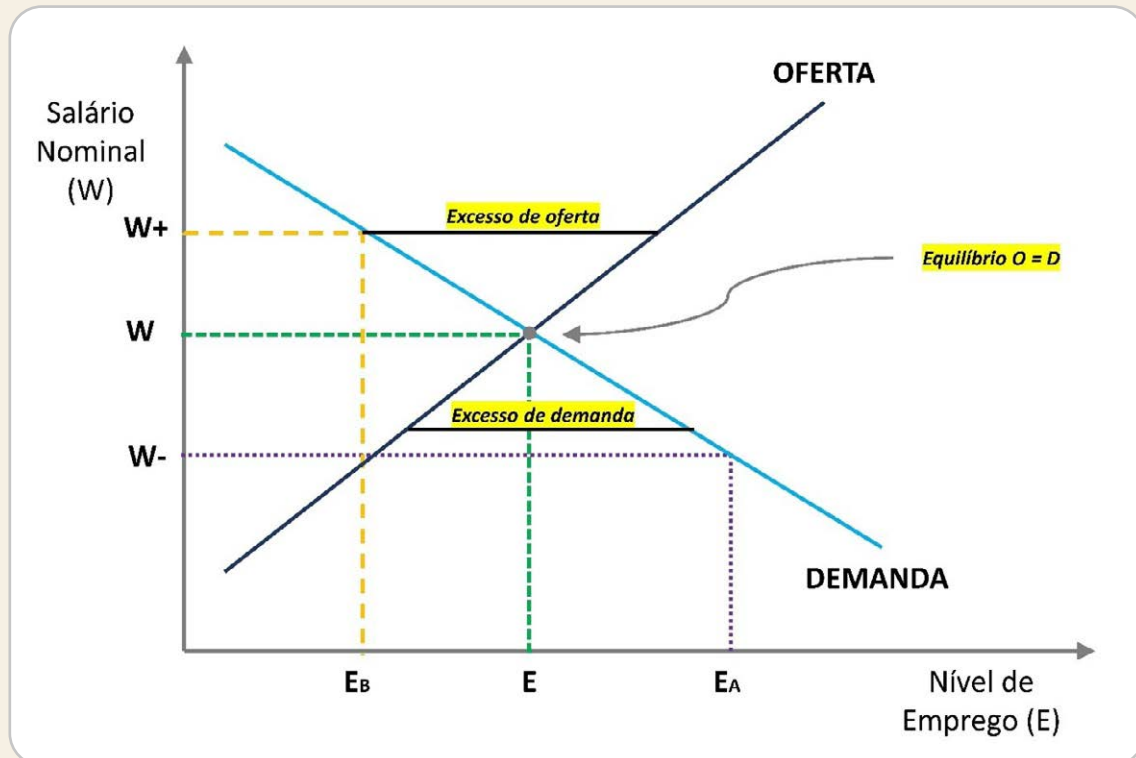
3.2.3 Complexidades do mercado de trabalho em saúde

O desalinhamento entre a oferta e a demanda no setor de saúde representa um dos maiores obstáculos para a efetividade do setor. Esse desequilíbrio pode se apresentar de várias maneiras, como o excesso de profissionais em certos campos e a escassez em outros, refletindo tanto limitações no planejamento da mão de obra quanto mudanças rápidas nas demandas do sistema de saúde.

Abordagens tradicionais para desenvolver estratégias de recursos humanos no setor de saúde têm se baseado principalmente no planejamento do lado da oferta, que estima as necessidades de força de trabalho em saúde e propõe aumentar (ou diminuir) as capacidades educacionais para estreitar a lacuna entre o número existente de trabalhadores de saúde e o número necessário. Esse foco na produção de trabalhadores de saúde negligencia outros aspectos importantes da dinâmica do mercado de trabalho, como mudanças na demanda, relações do empregador com trabalhadores e sindicatos profissionais, ou preferências de trabalho dos trabalhadores (OMS, 2021, p. xxxiv, tradução nossa).

Segundo o *Health Labour Market Analysis Guidebook* (OMS, 2021), o relacionamento entre oferta e demanda dos profissionais de saúde é fundamental para os resultados do mercado de trabalho no setor. Esse equilíbrio é influenciado tanto pelo setor educacional, que impacta a oferta de trabalhadores, quanto pelas condições do setor de saúde, que determinam a criação de empregos. O termo “desalinhamento” é usado para descrever as diferentes formas de desarmonia entre a oferta de trabalhadores e a demanda por vagas. A Figura 14 ilustra a interseção entre as curvas de oferta de trabalho (S) e demanda de trabalho (D) em um mercado competitivo.

Figura 14 – Curva de Oferta (S) e Demanda (D)



Fonte: Health Labour Market Analysis Guidebook (2021).

O equilíbrio ocorre quando a oferta é igual à demanda, gerando o salário competitivo w^* e o emprego E^* . O salário w^* é o salário de equilíbrio de mercado. No nível salarial $w+$, muitos profissionais de saúde competirão pelos poucos empregos disponíveis; conseqüentemente, haverá um excesso de oferta de profissionais de saúde. Em contraste, no salário $w-$, o número de empregos disponíveis para profissionais de saúde é maior do que o número de profissionais de saúde dispostos a trabalhar naquele nível salarial, resultando em uma escassez de profissionais de saúde (OMS, 2021, p. 205, tradução nossa).

Segundo o *Health Labour Market Analysis Guidebook* (OMS, 2021), os desalinhamentos no mercado de trabalho em saúde geram efeitos indesejáveis no desempenho dos sistemas de saúde, impactando pacientes, trabalhadores e empregadores. Esses efeitos são categorizados em três áreas principais: acesso aos cuidados, desempenho dos serviços de saúde (em termos de produtividade e qualidade) e bem-estar dos profissionais de saúde.

Esses desalinhamentos são de diversos tipos, mas podem ser categorizados basicamente em três grupos, de acordo com o *Health Labour Market Analysis Guidebook* (OMS, 2021): podem ocorrer de forma quantitativa, quando o número de profissionais disponíveis não corresponder às vagas existentes, ou de habilidades; quando houver incompatibilidade entre as qualificações dos trabalhadores e os requisitos dos cargos; e podem ser consequência de discriminação no mercado de trabalho, como por gênero ou raça, ou de outros tipos de preconceito no ambiente de trabalho.

3.3 Modelagem dinâmica do mercado de trabalho

No item anterior foi descrita a modelagem tradicional do mercado de trabalho desenvolvida por Sousa *et al.* (2013). Esta modelagem é a mais comumente aceita e utilizada atualmente na academia, pelos profissionais da área e pela Organização Mundial da Saúde em alguns estudos. O modelo possui uma perspectiva analítica e macroestrutural, ressaltando os componentes principais do sistema e suas ligações de forma hierárquica e linear. Ela distingue claramente os principais estoques e fluxos, tais como o setor educacional, a mão de obra disponível e os mecanismos de políticas públicas que estruturam o sistema, sem entrar em detalhes sobre as interações dinâmicas que ocorrem entre os seus componentes ou os ciclos de feedback que os ligam.

Esta omissão no tratamento das interações dinâmicas entre componentes e entre os ciclos de *feedback* impede um entendimento mais aprofundado das interações que ocorrem no mercado de trabalho em saúde. A relevância desses processos dinâmicos e dos ciclos de *feedback* reside no fato de que eles moldam o comportamento do sistema ao longo do tempo, impactando diretamente sua resiliência, eficácia e habilidade de se ajustar a mudanças externas e internas.

Assim, nesta etapa não se descreve um novo modelo, mas as interações e conexões dinâmicas que ocorrem entre os componentes do modelo tradicional, bem como os ciclos de *feedback* construídos por essas interações, com o intuito de alcançar uma compreensão mais profunda sobre o funcionamento do mercado de trabalho em saúde e assim possibilitar aos tomadores de decisão uma visão mais holística e integrada.

3.3.1 Articulação do Problema

Para a OECD e diversos autores (Leerapan *et al.*, 2021; OCDE, 2021a; Sousa *et al.*, 2013; Stermann, 2000), altas taxas de aposentadoria, emigração para outros países, distribuição desigual de enfermeiros e formação insuficiente de novos profissionais são

grandes desafios enfrentados pela força de trabalho em enfermagem no Brasil. Esses elementos provocam um sensível desequilíbrio entre oferta e demanda. Paralelamente a isso, o aumento da expectativa de vida e o respectivo aumento de doenças crônicas elevam a necessidade de atenção especializada. Em virtude destes pontos, nesta pesquisa, busca-se compreender os fluxos de entrada e saída de profissionais de enfermagem, considerando os principais fatores que impactam a força de trabalho e orientar políticas públicas para atenuar a falta de profissionais e garantir a sustentabilidade do sistema de saúde brasileiro.

Segundo Schlosser *et al.* (2022), houve no Brasil, ao longo dos últimos 20 anos, um aumento significativo no número de profissionais de enfermagem, passando de 3,41 por mil habitantes em 2001 para 11,65 em 2021, um aumento de aproximadamente 341%. Esse crescimento se deu em virtude do aumento no número de matrículas nos cursos de graduação e colaborou para o rejuvenescimento desta força de trabalho.

Mesmo em face disso, segundo Souza *et al.* (2021a), as taxas de aposentadoria estão crescendo, especialmente entre os profissionais com menor qualificação, que enfrentam condições adversas de trabalho para complementar a renda devido às aposentadorias de valor baixo. Tais autores também enfatizam que o Brasil lida com desafios relacionados à migração de enfermeiros para países com melhores condições de trabalho e remuneração, o que intensifica a falta de profissionais na área. A disparidade na alocação de profissionais, com grande concentração no Sudeste e escassez em regiões isoladas do Norte e Nordeste, é uma questão persistente e amplamente debatida nas últimas duas décadas.

Conforme Schlosser *et al.* (2022) e Souza *et al.* (2021a), espera-se que a necessidade de profissionais de enfermagem aumente ainda mais nos próximos 20 anos, impulsionada pelo envelhecimento da população e pelo crescimento das doenças crônicas. Além disso, a taxa de aposentadoria da atual força de trabalho, que ainda é jovem, deverá aumentar consideravelmente até 2030, exercendo uma pressão extra sobre a disponibilidade de profissionais. A Organização Mundial da Saúde (OMS) e outras organizações globais alertam que essa tendência global de envelhecimento da força de trabalho será crucial, especialmente em nações de renda baixa e média, como o Brasil, a menos que sejam empregadas estratégias efetivas de retenção e treinamento de novos profissionais.

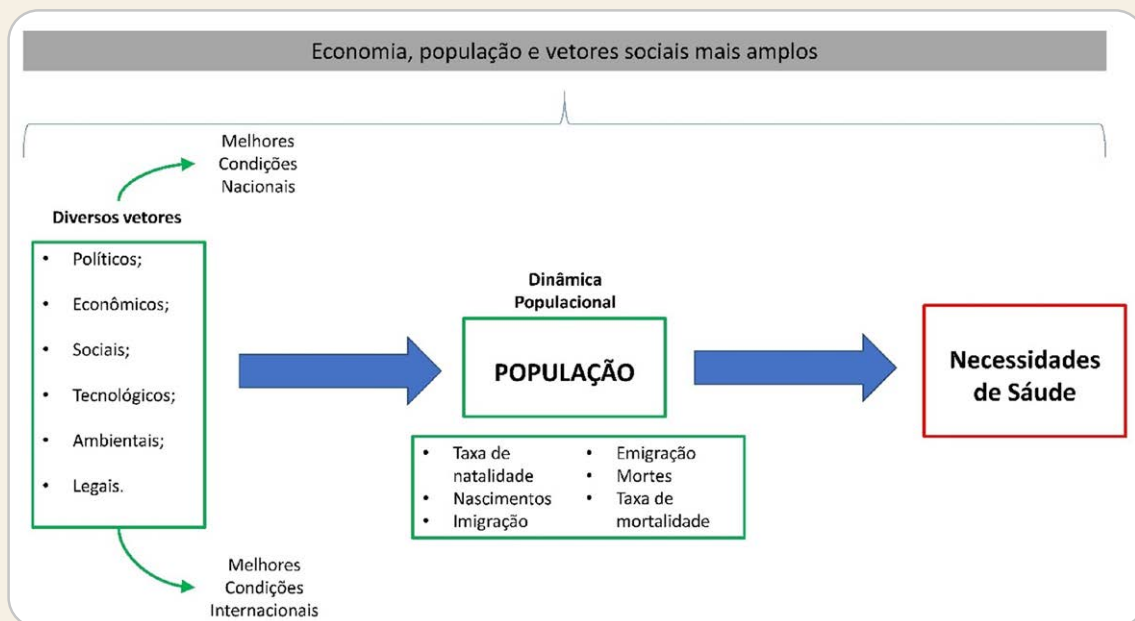
Tais pesquisas sugerem que o período de 20 anos é adequado tanto para compreender as dinâmicas históricas que formaram a situação presente quanto para projetar o futuro e formular políticas públicas adequadas para atenuar a escassez prevista e garantir a sustentabilidade do sistema de saúde.

3.3.2 Hipótese dinâmica e formulação do modelo

A hipótese dinâmica inicial propõe uma conexão entre a entrada de novos profissionais de enfermagem através da graduação, a taxa de imigração, a taxa de aposentadorias e taxa de emigração, particularmente em regiões isoladas, que afetam o estoque disponível de recursos humanos. Afetando estas variáveis se elenca um conjunto de fatores agora explicitados, a saber: políticos, econômicos, socioculturais, tecnológicos, ambientais e legais. Este conjunto de fatores, conhecidos como PESTAL, afetam as escolhas dos indivíduos no que tange as decisões que afetam as variáveis acima.

Outros fatores, como a disparidade na alocação de profissionais, o aumento da expectativa de vida e o aumento das enfermidades crônicas, pressionam o sistema de saúde a ponto de resultar no referido déficit crescente de profissionais capacitados.

Figura 15 – Construindo o Modelo Dinâmico a partir do Modelo Tradicional



Fonte: autoria própria.

O mercado de trabalho em saúde é um dos diversos mercados de trabalho. Como os demais, ele está inserido no todo do sistema social, interagindo direta ou indiretamente com diversos setores. Nesta interação, o conjunto de fatores PESTAL

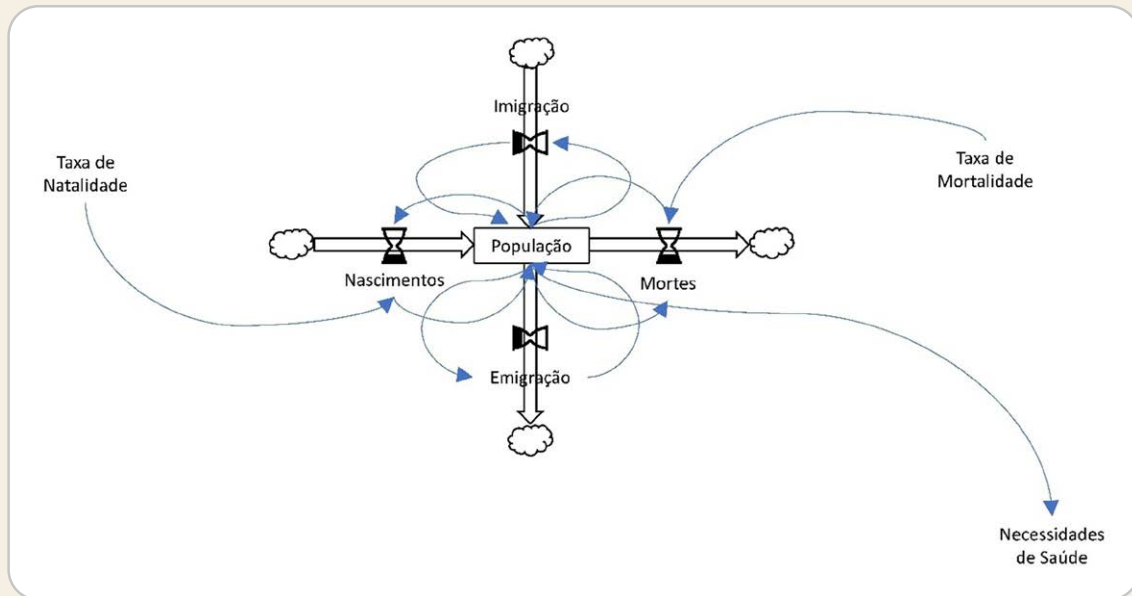
passa a influenciar e a impactar o funcionamento de sua dinâmica, especialmente no que se refere à população. Conforme a Figura 15, esses fatores influenciam e impactam diretamente duas importantes variáveis: “Melhores Condições Nacionais” e “Melhores Condições Internacionais”.

Na Figura 15, temos ao centro a variável população, que representa o total de pessoas em uma região ou país, e é a principal impactada pelas outras variáveis deste setor do modelo. No quadro logo abaixo de população temos a variável taxa de natalidade que representa o número de nascimentos por unidade de população em um determinado período e que influencia diretamente a variável nascimentos, que se refere à quantidade de nascimentos ocorridos numa população em dado período. O número de nascidos impacta positivamente a população. Em contrapartida, quanto maior a população maior o número de nascimentos. Encerrando as variáveis que influenciam positivamente a população, temos o fluxo de imigração, que representa o número de pessoas que saem de um país para viver no Brasil. A seguir, temos as variáveis que impactam negativamente a população. A primeira delas é o fluxo de emigração, que representa a quantidade de pessoas que saem do Brasil para viver noutro país. Em seguida temos o fluxo de mortes, que diminuem o quantitativo populacional, e são reforçados pela taxa de mortalidade, que é a medida de quantas mortes ocorrem numa dada população num determinado período, e pelo aumento populacional. Assim, essas são as variáveis que representam a dinâmica populacional básica.

Já “Necessidades de Saúde” representam uma variável crucial em nosso modelo, refletindo a procura global por serviços de saúde. Esta procura é impactada diretamente pelas variáveis demográficas relatadas anteriormente. Com o crescimento e o envelhecimento da população, espera-se que as exigências por cuidados de saúde aumentem, especialmente no que diz respeito a cuidados especializados. A inclusão desta variável é crucial para vincular o segmento demográfico do modelo à força de trabalho, possibilitando uma avaliação mais detalhada de como as alterações na população impactam diretamente a oferta e demanda por profissionais de enfermagem, ajudando a projetar futuras necessidades de recursos humanos, fundamentando políticas públicas de adequação.

Na figura 16, é possível observar o acréscimo da variável “Necessidades de Saúde” à parte do modelo descrito anteriormente.

Figura 16 – Representação via Dinâmica de Sistemas do Modelo Tradicional da Figura 15

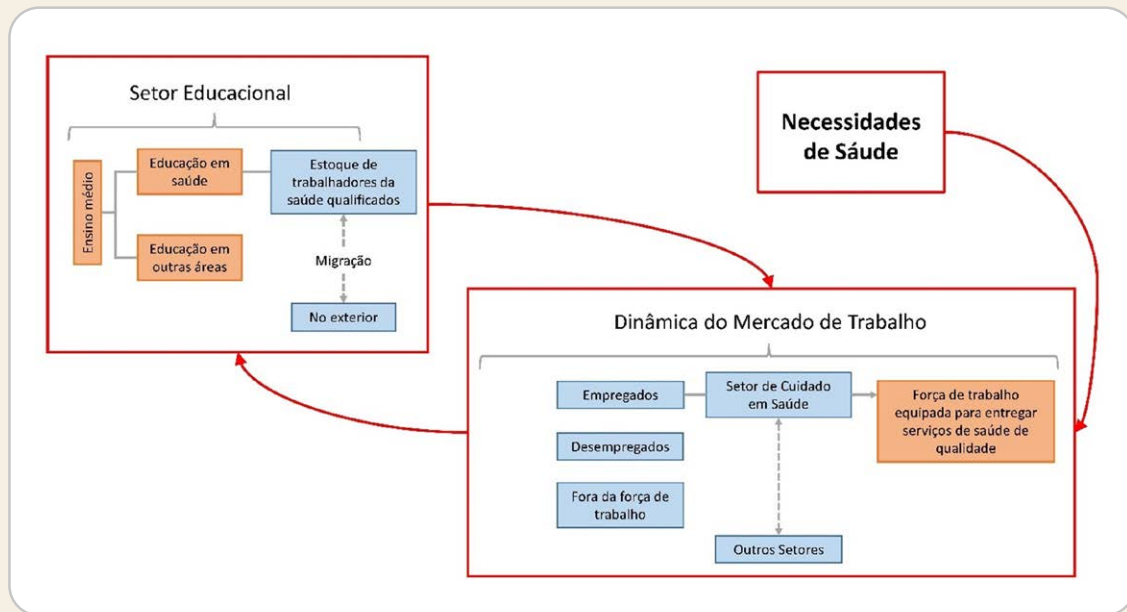


Fonte: elaboração própria.

As influências entre as variáveis estão representadas no modelo pelas setas azuis com sinais positivos e negativos. Na dinâmica de sistemas, segundo Sterman (2000), os sinais positivos (+) em uma relação causal indicam que há uma relação direta entre as variáveis conectadas. Isso significa que um aumento em uma variável resulta em um aumento na outra, e uma diminuição em uma leva a uma diminuição na outra, caracterizando uma mudança na mesma direção entre as variáveis. Já os sinais negativos (-) em uma relação causal indicam que há uma relação inversa ou de equilíbrio entre as variáveis conectadas. Isso significa que, quando uma variável aumenta, outra diminui, e vice-versa. Esse tipo de relação é caracterizado por uma mudança em direções opostas entre as variações. Os sinais positivos e negativos não representam algo bom ou ruim, apenas indicam se as variáveis caminham na mesma direção ou em direções opostas.

Fica então entendido que com o aumento da população via nascimentos ou imigração há o aumento por necessidades de saúde. O contrário também ocorre, a diminuição da população via número de mortes ou emigração causa uma queda por necessidades de saúde. Como dito anteriormente, “Necessidades de Saúde” conecta o setor populacional do modelo ao setor da força de trabalho.

Figura 17 – Necessidades de Saúde e sua Conexão com os Dois Setores do Modelo Tradicional



Fonte: elaboração própria.

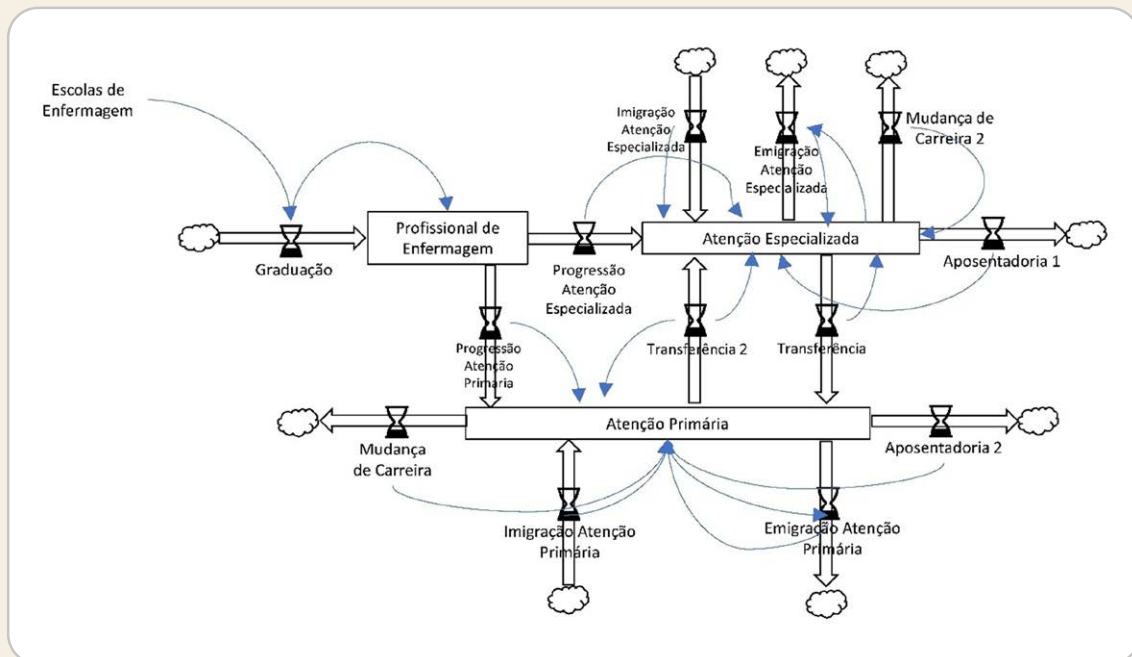
As “Necessidades de Saúde” da população exercem influência tanto no setor educacional quanto no mercado de trabalho em saúde, impactando a formação e a distribuição dos profissionais de saúde. Essa relação é complexa e dinâmica, refletindo as demandas emergentes e as condições de saúde da sociedade. No setor educacional, “Necessidades de Saúde” determinam indiretamente a oferta de cursos e programas voltados para a formação de profissionais capacitados a atender às demandas específicas da população. No que tange ao mercado de trabalho em saúde, as necessidades de saúde da população influenciam a demanda por profissionais em diversas especialidades e regiões. Conforme destacado por Campos *et al.* (2021), há uma determinação recíproca entre o mercado de trabalho e a educação, onde o mercado impacta a educação e vice-versa, sendo essas determinações mútuas objeto de reflexões. Esta conexão criada por “Necessidades de Saúde” ocorre através de duas novas variáveis, “Atenção Primária” e “Atenção Especializada”, no entendimento de que quanto maior a necessidade de saúde maior a procura pelos serviços da atenção primária e da atenção especializada.

Segundo Andrade *et al.* (2021), a “Atenção Primária” é a porta de entrada para o sistema de saúde, possuindo um papel fundamental nos cuidados preventivos, diagnósticos e terapêuticos básicos e fornecendo serviços contínuos e completos para

promoção da saúde, prevenção de doenças e tratamento de condições agudas e crônicas. Outra característica fundamental da Atenção Primária é sua habilidade de diminuir a necessidade de Atenção Especializada. Quando os serviços de saúde são eficientes, diversas condições de saúde podem ser tratadas ou controladas antecipadamente, prevenindo a evolução para situações mais sérias que requerem atenção especial. Uma forte atenção primária auxilia na sustentabilidade do sistema de saúde a longo prazo, enquanto qualquer insuficiência nesse nível leva a um excesso de trabalho, elevação dos custos e variações na qualidade do atendimento em outros níveis.

Por sua vez, a “Atenção Especializada” abrange serviços de saúde mais complexos, tais como tratamentos em hospitais, cirurgias e o gerenciamento de condições crônicas. No modelo, ela pode ser aplicada como uma variável diretamente relacionada às Necessidades de Saúde, especialmente com o crescimento da expectativa de vida e a incidência de doenças crônicas na população, que demandam assistência especializada. Da mesma forma que a “Atenção Primária”, a “Atenção Especializada” também é afetada por variáveis como “Aposentadoria”, “Emigração” e “Imigração”, que impactam a disponibilidade de profissionais para suprir as necessidades desses serviços mais complexos.

Figura 18 – Diagrama de Estoque e Fluxo do setor de Força de Trabalho do Modelo



Fonte: elaboração própria.

No modelo, a Atenção Primária e a Atenção Especializada têm funções interconectadas, espelhando vários graus de complexidade nos serviços de saúde, ambos essenciais para o funcionamento adequado do sistema. Variáveis como “Progressão na Atenção Primária” e “Mudança de carreira” influenciam diretamente a disponibilidade de profissionais. Os enfermeiros podem ser promovidos, mudar de área ou se aposentar, o que afeta a quantidade de profissionais à disposição no nível primário de assistência. O mesmo ocorre para a “Atenção Especializada”, pois inferimos que existe uma constante transferência, fluxo, de pacientes e profissionais de um nível de atenção para outro.

Ambos os níveis de atenção são afetados pelas variáveis “Aposentadoria” e “Emigração” de profissionais. A menor disponibilidade de profissionais diminui a qualidade de atendimento, especialmente na “Atenção Especializada”, que requer uma equipe com formação mais capacitada. Esses movimentos de emigração são intensificados pelas condições socioeconômicas internacionais mais desenvolvidas, que atraem profissionais para fora do país. Contudo, o avanço na carreira, que envolve uma mudança de profissionais da atenção primária para a especializada, assim como “Melhores Condições Socioeconômicas Nacionais” podem amenizar esse efeito em parte, dependendo da efetividade das políticas de manutenção e formação.

Portanto, a “Atenção Primária” e a “Atenção Especializada” estão intrinsecamente ligadas, cada uma afetando a outra através de fluxos de profissionais e pacientes. Essas interações são fundamentais para garantir a eficácia e a continuidade do serviço de saúde, ressaltando a relevância de políticas que reforçam ambos os níveis e garantem a distribuição de profissionais para suprir a demanda em expansão. Assim, podemos perceber, através da Figura 13, que estas duas variáveis dependem de uma intrincada rede de conexões interdependentes, da qual podemos destacar: “Escolas de Enfermagem”, “Graduação” e “Profissional de Enfermagem”.

As instituições de ensino de enfermagem são a base para a formação de novos profissionais. Segundo o Perfil da Enfermagem no Brasil (Machado, 2017), houve um aumento significativo no número de escolas de enfermagem. No entanto, esse avanço foi acompanhado por desafios ligados à qualidade do ensino, com diversos cursos oferecidos por entidades privadas sem um controle de qualidade adequado. Esses temas suscitam dúvidas sobre a capacidade dos recém-formados em lidar com as intrincadas exigências do setor de saúde.

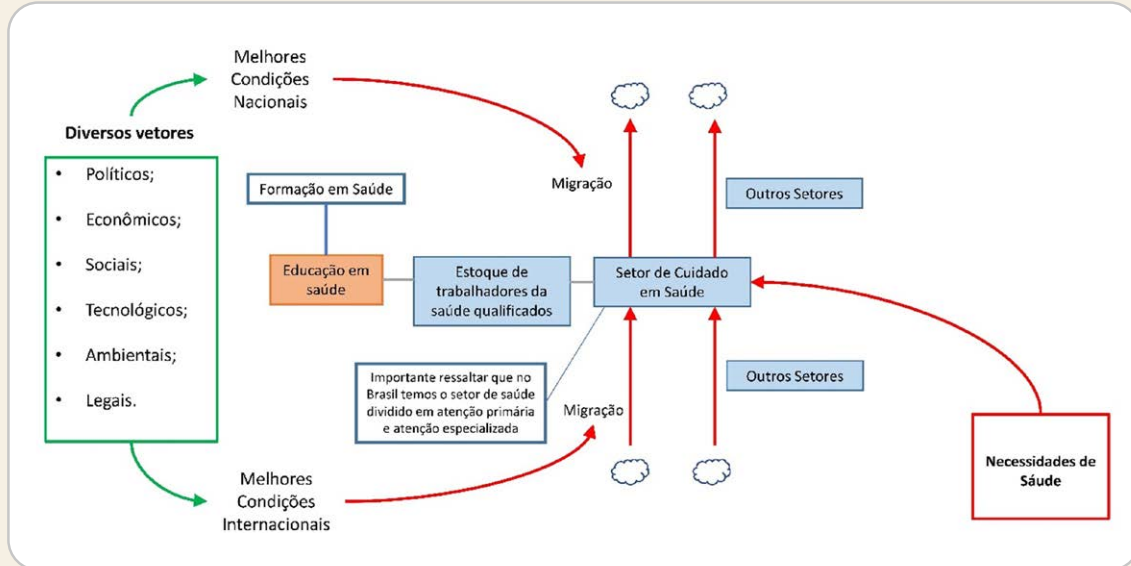
O processo de graduação em enfermagem é o que capacita os alunos a se tornarem profissionais prontos para ingressar no mercado de trabalho. No Brasil, os cursos de enfermagem, geralmente com duração de quatro a cinco anos, trazem uma mescla de teoria e prática, com destaque para projetos supervisionados que são fundamentais para a formação dos profissionais em potencial. Machado (2017), ressalta que, mesmo com o aumento no número de graduados, a capacidade do mercado de trabalho em absorver esses profissionais é limitada, principalmente em áreas periféricas onde a disponibilidade de empregos é restrita.

O “Profissional de Enfermagem” desempenha um papel crucial na manutenção e melhoria da saúde pública no Brasil, particularmente no Sistema Único de Saúde (SUS). Contudo, é um desafio constante manter esses profissionais no mercado de trabalho. De acordo com Silva e Machado (2019), diversos profissionais lidam com condições de trabalho adversas, tais como excesso de trabalho, altos níveis de pressão e condições precárias, o que resulta em um alto índice de rotatividade e emigração para países que proporcionem melhores condições de trabalho e renda. Esses elementos limitam a capacidade do sistema de saúde em atender às demandas populacionais em ascensão, particularmente em um contexto de envelhecimento da população e crescimento das doenças crônicas.

A capacidade do Brasil de formar, manter e desenvolver sua força de trabalho na área da saúde é diretamente influenciada pelas variáveis escolas de enfermagem, graduação e profissionais da área. A ampliação da oferta de cursos de enfermagem requer ações de controle de qualidade e políticas governamentais que promovam a retenção e a distribuição equânime desses profissionais, com o objetivo de suprir as necessidades de saúde em todas as áreas do país.

Na Figura 19, é possível verificar a integração das Figuras 15 e Figura 17, anteriormente apresentadas, consolidando as relações entre o setor educacional, o mercado de trabalho em saúde e as necessidades de saúde. Essa integração evidencia a complexidade do sistema, unindo a perspectiva macroestrutural e analítica do modelo tradicional com a abordagem dinâmica, que considera as interações, ciclos de feedback e influências externas.

Figura 19 – Diagrama Geral do Modelo Tradicional Mostrando suas Conexões Causais



Fonte: elaboração própria.

O setor educacional, representado pela formação em saúde, conecta-se diretamente ao “pool de trabalhadores qualificados”, que funciona como o estoque central de profissionais disponíveis para o setor de saúde. Esses trabalhadores, por sua vez, são alocados para atender às demandas do sistema de saúde, que, no Brasil, se divide em atenção primária e atenção especializada. As necessidades de saúde atuam como um direcionador estratégico, impactando tanto o setor educacional, ao demandar a formação de novos profissionais, quanto o mercado de trabalho, ao regular a oferta e a distribuição de trabalhadores.

A influência das necessidades de saúde sobre o mercado de trabalho e o setor educacional também é destacada por Campos *et al.* (2021), que apontam a necessidade de ajustes contínuos na formação de profissionais para atender às mudanças nas condições epidemiológicas e demográficas. Nesse contexto, a integração dos elementos na figura proporciona uma base para análises mais abrangentes e para a construção de cenários preditivos, permitindo aos tomadores de decisão identificarem pontos de alavancagem para intervenções eficazes.

Com relação à variável “Melhores Condições Socioeconômicas Nacionais”, Jacinto, Tejada e Sousa (2010) apontam que, em períodos de crescimento econômico, caracterizados pelo aumento do emprego e da renda, a saúde da população tende a melhorar, levando a uma diminuição nas taxas de mortalidade e à expansão do

acesso aos serviços de saúde. Isso ocorre devido à melhoria das condições de vida e à disponibilidade de serviços, tais como saneamento básico e educação, que impactam diretamente a saúde.

Albuquerque *et al.* (2017) indicam que a expansão do PIB per capita e o aumento da renda familiar no Brasil entre 2000 e 2016 contribuíram para o desenvolvimento socioeconômico, o que, conseqüentemente, influenciou positivamente a disponibilidade de serviços de saúde em certas áreas. No entanto, esse avanço foi desigual, com maior concentração de desenvolvimento e disponibilidade de serviços nas regiões Sudeste e Sul, enquanto as regiões Norte e Nordeste ainda lidam com obstáculos socioeconômicos que restringem o acesso à saúde. Ainda segundo Albuquerque *et al.* (2017), a ampliação do acesso à educação e a regulamentação do mercado trabalhista impactam a qualidade da saúde pública. Um grau de educação mais alto tem uma tendência de melhoria da saúde da população, ao mesmo tempo que promove a capacitação de profissionais para a área de enfermagem e saúde pública.

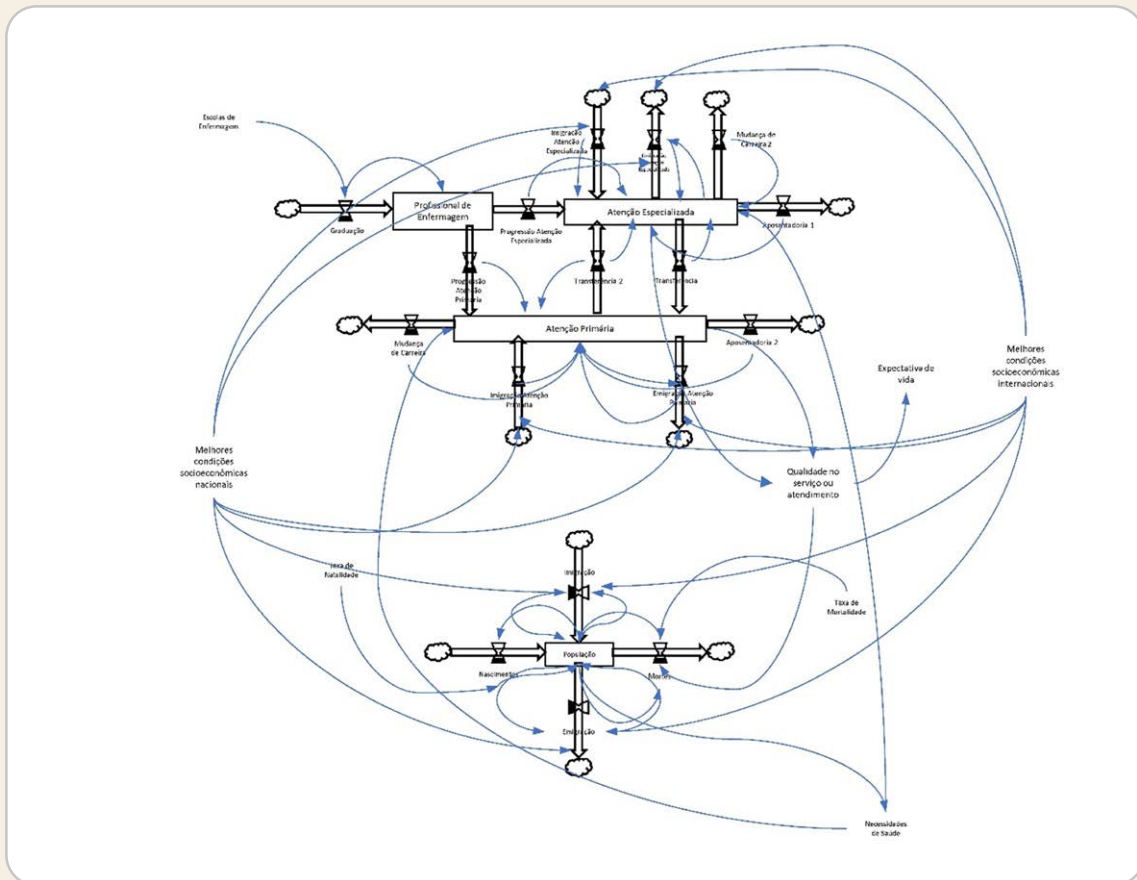
Segundo a OCDE (2021b), o trabalho formal está intrinsecamente ligado à estabilidade financeira das famílias, ampliando o acesso aos planos de saúde e, conseqüentemente, aliviando o Sistema Único de Saúde (SUS) para que possa prestar um atendimento mais eficaz às populações mais vulneráveis. Assim, a variável “Melhores Condições Socioeconômicas Nacionais” impactam tanto a procura por assistência médica quanto a disponibilidade de profissionais encontrados para satisfazer as demandas de uma população em crescimento e com aumento na expectativa de vida.

No cenário da saúde mundial, a variável “Melhores Condições Socioeconômicas Internacionais” exerce uma influência significativa na força de trabalho de enfermagem e na qualidade dos sistemas de saúde de vários países, incluindo o Brasil. A globalização, juntamente com uma economia global mais sólida, afeta diretamente a oferta de profissionais e a demanda por serviços de saúde. “Melhores Condições Socioeconômicas” em nações desenvolvidas atraem profissionais de saúde de países em desenvolvimento, como o Brasil, proporcionando melhores resultados laborais, condições de trabalho e chances de progressão profissional. Esta migração de profissionais da saúde pode resultar numa falta de enfermeiros e médicos em países menos desenvolvidos, incluindo a oferta de serviços de saúde, especialmente nas regiões mais desfavorecidas.

Segundo a OCDE (2021b), melhores condições socioeconômicas internacionais impactam diretamente os investimentos na área da saúde, seja na infraestrutura ou na formação de profissionais. As economias mais robustas costumam investir mais em

saúde pública, o que pode ser um exemplo para o Brasil na tentativa de fortalecer o SUS e aprimorar a eficácia de seus serviços. Na Figura 20 temos o modelo tradicional proposto por Sousa *et al.* (2013), apresentado pela abordagem da Dinâmica de Sistemas.

Figura 20 – Diagrama de Estoque e Fluxo da Força de Trabalho em Enfermagem via Dinâmica de Sistemas



Fonte: elaboração própria.

Para Buss (2007), a desigualdade no desenvolvimento socioeconômico global e as normas desiguais do comércio mundial também impactam os países em desenvolvimento, intensificando a pressão sobre seus sistemas de saúde e criando obstáculos para manter uma mão de obra engajada e motivada. Portanto, compreender e atenuar o efeito das condições socioeconômicas globais é crucial para a criação de políticas públicas que busquem garantir a sustentabilidade dos sistemas de saúde no Brasil e em outras nações em desenvolvimento.

Finalizando a explicação temos duas outras variáveis restantes: “Qualidade no Serviço ou Atendimento” e “Expectativa de Vida”. A “Qualidade no Serviço ou Atendimento” é um fator crucial para avaliar o rendimento dos sistemas de saúde. Ela demonstra não só o desempenho dos profissionais de saúde, mas também a adequação da infraestrutura e o acesso a tratamentos e assistência de alta complexidade. A excelência na assistência está intrinsecamente relacionada à capacitação e acessibilidade de profissionais competentes, como enfermeiros, além das condições de trabalho desses profissionais. Pesquisas apontam que, em contextos de maior pressão laboral, como excesso de pacientes e escassez de recursos, a qualidade do serviço tende a cair, afetando diretamente os resultados de saúde e a satisfação dos pacientes (Machado, 2017). Ademais, estratégias voltadas para aprimorar a qualidade do serviço frequentemente incluem o treinamento constante dos profissionais de saúde e a melhoria das condições laborais, elementos que impactam diretamente na retenção e motivação dos enfermeiros.

A “Expectativa de Vida” é um indicador que espelha o estado geral de saúde de uma comunidade. O aumento na expectativa de vida, muitas vezes impulsionado por avanços nas condições socioeconômicas e no acesso a serviços de saúde, leva a um envelhecimento da população. Este fenômeno afeta diretamente a demanda por serviços de saúde, já que pessoas mais velhas precisam de assistência médica constante, principalmente em casos de doenças crônicas e condições ligadas ao envelhecimento. O aumento da expectativa de vida apresenta desafios para o sistema de saúde, que precisa estar preparado para oferecer assistência adequada a uma população que vive mais e tem necessidades de saúde mais complexas (Silva; Machado, 2019). Simultaneamente, o crescimento da expectativa de vida requer uma preparação adequada dos profissionais de saúde, particularmente dos enfermeiros, para lidar com essa demanda, intensificando a necessidade de planejamento e distribuição de recursos no setor.

Essas duas variáveis estão interconectadas: com o aumento da expectativa de vida, a procura por serviços de saúde de alta qualidade aumenta, obrigando o sistema de saúde a aprimorar continuamente a qualidade dos serviços ou atendimentos para assegurar o bem-estar da população mais idosa e vulnerável. Dessa forma, o modelo apresenta uma síntese dinâmica e sistêmica, proporcionando uma visão integrada do mercado de trabalho em saúde, essencial para análises preditivas e para a formulação de políticas públicas que promovam a eficiência, a equidade e a sustentabilidade do sistema.

3.4 Reflexões finais

A utilização da dinâmica de sistemas na modelagem da força de trabalho em enfermagem no Brasil proporciona uma visão integrada das interdependências entre variáveis educacionais, demográficas, econômicas e profissionais. O modelo desenvolvido permite a identificação dos principais desafios enfrentados pelo sistema, como o envelhecimento populacional, as desigualdades regionais na distribuição de profissionais, as taxas de aposentadoria e migração, e as fragilidades no setor educacional, que comprometem a sustentabilidade da oferta de enfermeiros no país.

Embora este estudo não tenha avançado para a simulação de cenários, a abordagem proposta fornece uma base sólida para reflexões estratégicas e construção de políticas públicas mais assertivas. Recomenda-se a implementação de políticas voltadas para a retenção regional de profissionais, o fortalecimento da formação educacional com foco em qualidade e alinhamento às demandas do mercado, além da valorização profissional por meio de melhores condições de trabalho e progressão na carreira.

A composição desse modelo destaca a importância de uma abordagem sistêmica e integrada para enfrentar os desafios do mercado de trabalho em enfermagem. Ao fornecer uma representação clara das relações causais e dos ciclos de *feedback*, o estudo abre caminho para análises futuras, incluindo simulações, que poderão servir de base para intervenções específicas e avaliação de impactos. Dessa forma, a modelagem proposta se torna uma importante ferramenta para o planejamento estratégico, garantindo a eficiência e a sustentabilidade do mercado de trabalho.



Referências

ALBUQUERQUE, M. V. de *et al.* Desigualdades regionais na saúde: mudanças observadas no Brasil de 2000 a 2016. **Ciência & Saúde Coletiva**, Rio de Janeiro, v. 22, p. 1055-1064, 2017. DOI: <https://doi.org/10.1590/1413-81232017224.26862016>. Disponível em: <https://www.scielo.br/j/csc/a/mnpHNBCXdptWTzt64rx5GSn/?format=html&lang=pt>. Acesso em: 23 out. 2025.

ANDRADE, M. V. *et al.* Transition to universal primary health care coverage in Brazil: analysis of uptake and expansion patterns of Brazil's Family Health Strategy (1998-2012). **PloS One**, San Francisco, v. 16, n. 5, e0251764, 2021. DOI: <https://doi.org/10.1371/journal.pone.0251764>. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0201723>. Acesso em: 23 out. 2025.

AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS (Brasil). **Estudo técnico**: a LGPD e o tratamento de dados pessoais para fins acadêmicos e para a realização de estudos por órgão de pesquisa. Brasília, DF: ANPD, 2022. (Texto para discussão, n. 1/2022). Disponível em: https://www.gov.br/anpd/pt-br/assuntos/noticias/sei_00261-000810_2022_17.pdf. Acesso em: 8 ago. 2024.

AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS (Brasil). **Guia orientativo**: tratamento de dados pessoais para fins acadêmicos e para a realização de estudos e pesquisas. Brasília, DF: ANPD, 2023. Disponível em: <https://www.gov.br/anpd/pt-br/documentos-e-publicacoes/documentos-de-publicacoes/web-guia-anpd-tratamento-de-dados-para-fins-academicos.pdf>. Acesso em: 10 ago. 2024.

BARRETO, M. da S.; CARREIRA, L.; MARCON, S. S. Envelhecimento populacional e doenças crônicas: reflexões sobre os desafios para o Sistema de Saúde Pública. **Revista Kairós-Gerontologia**, São Paulo, v. 18, n. 1, p. 325-339, 2015. Disponível em: <https://revistas.pucsp.br/index.php/kairos/article/view/26092>. Acesso em: 23 out. 2025.

BEZERRA, E. **Princípios de análise e projeto de sistemas com UML**. 3. ed. Rio de Janeiro: GEN LTC, 2015.

BOTELHO, M. C.; CAMARGO, E. P. do A. A aplicação da Lei Geral de Proteção de Dados na saúde. **Revista de Direito Sanitário**, São Paulo, v. 21, e0021, 2021. DOI: <https://doi.org/10.11606/issn.2316-9044.rdisan.2021.168023>.

Disponível em: <https://revistas.usp.br/rdisan/article/view/168023/178494>.
Acesso em: 10 jul. 2024.

BOTRUGNO, C. Cybersecurity, privacy, and health data protection in the digital strategy of the European Union. **Revista de Estudos Constitucionais, Hermenêutica e Teoria do Direito**, São Leopoldo, v. 14, n. 3, p. 300-312, set./dez. 2022. DOI: <https://doi.org/10.4013/rechtd.2022.143.01>. Disponível em: <https://revistas.unisinos.br/index.php/RECHTD/article/view/26455>. Acesso em: 3 jul. 2024.

BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República, 1988. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 24 out. 2025.

BRASIL. Presidência da República. **Lei n.º 12.527, de 18 de novembro de 2011**. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. Brasília, DF: PR, 2011. Disponível em: <https://www.gov.br/saude/pt-br/acesso-a-informacao/banco-de-precos/legislacao/lei-no-12-527-de-18-de-novembro-de-2011.pdf/view>. Acesso em: 22 out. 2025.

BRASIL. Presidência da República. **Lei n.º 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Brasília, DF: PR, 2018a. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 22 out. 2025.

BRASIL. Presidência da República. **Mensagem n.º 451, de 14 de agosto de 2018**. Brasília, DF: PR, 2018b. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/Msg/VEP/VEP-451.htm. Acesso em: 22 out. 2025.

BUSS, P. M. Globalização, pobreza e saúde. **Ciência & Saúde Coletiva**, São Paulo, v. 12, p. 1575-1589, 2007. DOI: <https://doi.org/10.1590/S1413-81232007000600019>. Disponível em: <https://www.scielo.br/j/csc/a/VLr6csVhCdhmLyh5tqwQnkn/?lang=pt>. Acesso em: 23 out. 2025.

BUSSINGUER, E. C. A.; ZAGANELLI, J. C. Governança da privacidade e as organizações sociais de saúde: uma análise da Lei Geral de Proteção de Dados. **Revista Jurídica**, Curitiba, v. 4, n. 71, p. 855-869, 2022. Disponível em: <https://revista.unicuritiba.edu.br/index.php/RevJur/article/view/4873/pdf>. Acesso em: 10 jul. 2024.



CAMPOS, F. E. *et al.* **Profissões e mercado de trabalho em saúde: perspectivas para o futuro.** Rio de Janeiro: Fiocruz, 2021. (Textos para Discussão, n. 64). Disponível em: https://saudeamanha.fiocruz.br/wp-content/uploads/2021/07/Campos-et-al_Profissoes_e_mercado_de_trabalho_em_saude_TD_64.pdf. Acesso em: 22 out. 2025.

CASSEL, P. E.; PETEROSI, H. G. Considerações sobre o impacto da Lei Geral de Proteção de Dados na pesquisa. *In: SIMPÓSIO DOS PROGRAMAS DE MESTRADO PROFISSIONAL*, 15., 2020, São Paulo. **Anais** [...]. São Paulo: CPS, 2020. Disponível em: <http://www.pos.cps.sp.gov.br/files/artigo/file/980/807203593ea41491fb2cd91d3496d59a.pdf>. Acesso em: 10 jul. 2024.

COLUSSI, F. A. M.; SANTOS, T. L. V. dos. Novas tecnologias e liberdade de expressão na pesquisa científica: uma análise sobre a proteção de dados genéticos e de saúde. **Revista de Biodireito e Direito dos Animais**, Florianópolis, v. 4, n. 2, p. 1-21, jul./dez. 2018. Disponível em: <https://www.indexlaw.org/index.php/revistarbda/article/view/4690>. Acesso em: 7 jul. 2024.

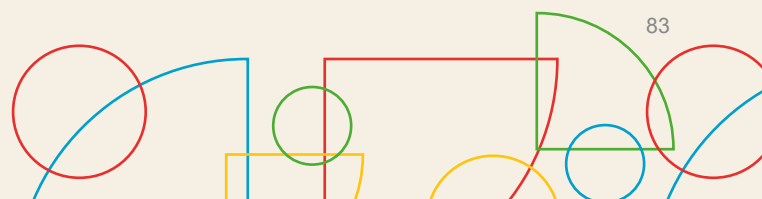
CONSELHO FEDERAL DE ENFERMAGEM (Brasil). **Enfermagem em números.** Brasília, DF: Cofen, 2024. Disponível em: <https://www.cofen.gov.br/enfermagem-em-numeros/>. Acesso em: 28 nov. 2024.

CONSELHO NACIONAL DE SAÚDE (Brasil). **Resolução n.º 466, de 12 de dezembro de 2012.** Brasília, DF: CNS, 2012. Disponível em: https://bvsms.saude.gov.br/bvs/saudelegis/cns/2013/res0466_12_12_2012.html. Acesso em: 22 out. 2025.

CONSELHO NACIONAL DE SAÚDE (Brasil). **Resolução n.º 738, de 1º de fevereiro de 2024.** Dispõe sobre o uso de bancos de dados com finalidade de pesquisa científica envolvendo seres humanos. Brasília, DF: CNS, 2024. Disponível em: <https://www.gov.br/conselho-nacional-de-saude/pt-br/atos-normativos/resolucoes/2024/resolucao-no-738.pdf/view>. Acesso em: 22 out. 2025.

DONEDA, D.; BARRETO, M. L.; ALMEIDA, B. de A. Uso e proteção de dados pessoais na pesquisa científica. **Direito Público**, Brasília, DF, v. 16, n. 90, 2019. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/3895>. Acesso em: 10 set. 2024.

DOURADO, D. de A.; AITH, F. M. A. A regulação da inteligência artificial na saúde no Brasil começa com a Lei Geral de Proteção de Dados Pessoais. **Revista de Saúde Pública**, São Paulo, v. 56, p. 80, 2022. Disponível em: <https://www.scielo.br/j/rsp/a/k38jGvJdbQSYN4MpzGZpfXw/?format=pdf&lang=pt>. Acesso em: 10 set. 2024.



FACHIN, Z.; HIRATA, A. C. A proteção jurídica dos dados pessoais sensíveis relativos à saúde. **Revista Jurídica**, Curitiba, v. 3, n. 70, p. 1-23, 2022. Disponível em: <https://revista.unicuritiba.edu.br/index.php/RevJur/article/view/5969/371373914>. Acesso em: 10 set. 2024.

FRAZÃO, A. Big data e aspectos concorrenciais do tratamento de dados pessoais. In: BIONI, B. *et al.* (coord.). **Tratado de proteção de dados pessoais**. Rio de Janeiro: Forense, 2021. p. 535-552.

INEP. **Censo Escolar**: resultados. Brasília, DF: Inep, 2024a. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-escolar/resultados>. Acesso em: 16 nov. 2024.

INEP. **Censo da Educação Superior**: microdados do Censo da Educação Superior. Brasília, DF: Inep, 2024b. Disponível em: <https://www.gov.br/inep/pt-br/area-de-atuacao/dados-abertos/microdados/censo-da-educacao-superior>. Acesso em: 16 nov. 2024.

JACINTO, P. de A.; TEJADA, C. A. O.; SOUSA, T. R. V. de. Efeitos das condições macroeconômicas sobre a saúde no Brasil. **Revista de Saúde Pública**, São Paulo, v. 44, n. 2, p. 310-317, abr. 2010. DOI: <https://doi.org/10.1590/S0034-89102010000200011>. Disponível em: <https://www.scielo.br/j/rsp/a/VB66cKWbMY35XmLstfVrVFL/?lang=pt>. Acesso em: 24 out. 2025.

LASKOWSKI-JONES, L.; CASTNER, J. The great resignation, newly licensed nurse transition shock, and emergency nursing. **Journal of Emergency Nursing**, St. Louis, v. 48, n. 3, p. 236-242, May 2022.

LEA, N. C. Como é que o regulamento geral da proteção de dados irá afetar a saúde? **Acta Médica Portuguesa**, Lisboa, v. 31, n. 7-8, p. 363-365, jul./ago. 2018. Disponível em: <https://www.actamedicaportuguesa.com/revista/index.php/amp/article/view/10881/5467>. Acesso em: 10 out. 2024.

LEERAPAN, B. *et al.* System dynamics modelling of health workforce planning to address future challenges of Thailand's Universal Health Coverage. **Human Resources for Health**, Londres, v. 19, p. 1-16, 2021. DOI: <https://doi.org/10.1186/s12960-021-00572-5>. Disponível em: <https://human-resources-health.biomedcentral.com/articles/10.1186/s12960-021-00572-5>. Acesso em: 22 out. 2025.

LOPES, I. M.; OLIVEIRA, P. Aplicabilidade do regulamento geral sobre proteção de dados em clínicas de saúde. **Revista Ibérica de Sistemas e Tecnologias de Informação**, Lisboa, n. E15, p. 118-129, 2018. Disponível em: <http://hdl.handle.net/10198/17773>. Acesso em: 1º ago. 2024.

MACHADO, M. (coord.). **Perfil da enfermagem no Brasil**: relatório final. Rio de Janeiro: Cofen; Fiocruz, 2017. Disponível em: <https://www.cofen.gov.br/perfilenfermagem/pdfs/relatoriofinal.pdf>. Acesso em: 24 out. 2025.

MACHADO, M. H. **Perfil da enfermagem no Brasil**. Belém: Cofen; Fiocruz, 2012. Disponível em: https://www.cofen.gov.br/wp-content/uploads/2012/03/26_06_12_palestra_mariahelenamacado.pdf. Acesso em: 22 out. 2025.

MARINHO, G. L.; QUEIROZ, M. E. V. de. Cobertura populacional de enfermeiros no Brasil: estimativas com base em diferentes fontes de dados. **Trabalho, Educação e Saúde**, Rio de Janeiro, v. 21, e00916202, 3 fev. 2023. DOI: <https://doi.org/10.1590/1981-7746-ojs916>. Disponível em: <https://www.scielo.br/j/tes/a/YHp7xfrKdXhV3HbwwmJC9Ty/?format=html&lang=pt>. Acesso em: 24 out. 2025.

MASSENO, M. D.; MARTINS, G. M.; FALEIROS JÚNIOR, J. L. de M. A segurança na proteção de dados: entre o GDPR europeu e a LGPD brasileira. **Revista do CEJUR**: Prestação Jurisdicional, Florianópolis, v. 8, n. 1, e346, jan./dez. 2020. DOI: <https://doi.org/10.21902/rctjsc.v8i1.346>. Disponível em: <https://revistadocejur.tjsc.jus.br/cejur/article/view/346/181>. Acesso em: 10 ago. 2024.

MENDES, L. S.; FUJIMOTO, M. (org.). **Relatório do painel LGPD nos tribunais**. São Paulo: Cedis-IDP; Jusbrasil, 2024. Disponível em: <https://static.jusbr.com/painel-lgpd/edicoes/relatorio-do-painel-lgpd-nos-tribunais-2023.pdf>. Acesso em: 10 ago. 2024.

MENDES, M. *et al.* Força de trabalho de enfermagem: cenário e tendências. **Revista de Enfermagem da UFSM**, Santa Maria, v. 12, e11, mar. 2022. DOI: <https://doi.org/10.5902/2179769267928>. Disponível em: <https://periodicos.ufsm.br/reufsm/article/view/67928>. Acesso em: 22 out. 2025.

MENDES-SANTOS, C. *et al.* A guided internet-delivered individually-tailored ACT-influenced cognitive behavioural intervention to improve psychosocial outcomes in breast cancer survivors (iNOVBC): study protocol. **Internet Interventions**, Amsterdam, v. 17, p. 100236, Sept. 2019. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2214782918300794>. Acesso em: 3 ago. 2024.

NIGENDA, G. G. *et al.* Towards the construction of health workforce metrics for Latin America and the Caribbean. **Human Resources for Health**, Londres, v. 9, p. 24, 2011. DOI: <https://doi.org/10.1186/1478-4491-9-24>. Disponível em: <https://human-resources-health.biomedcentral.com/articles/10.1186/1478-4491-9-24>. Acesso em: 24 out. 2025.

OBJECT MANAGEMENT GROUP. **UML – Unified Modeling Language**. Milford: OMG, Dec. 2017. Version 2.5.1. Disponível em: <https://www.omg.org/spec/UML/2.5.1>. Acesso em: 12 maio 2025.

ORGANIZAÇÃO DE COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Estudo da OCDE da Atenção Primária à Saúde no Brasil**. Paris: OCDE, 2021a. Disponível em: https://www.oecd.org/pt/publications/estudo-da-ocde-da-atencao-primaria-a-saude-no-brasil_9bf007f4-pt.html. Acesso em: 22 out. 2025.

ORGANIZAÇÃO DE COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Estudos da OCDE sobre os Sistemas de Saúde: Brasil 2021**. Paris: OCDE, 2021b. Disponível em: https://www.oecd.org/pt/publications/estudos-da-ocde-sobre-os-sistemas-de-saude-brasil-2021_f2b7ee85-pt.html. Acesso em: 22 out. 2025.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. **Health labour market analysis guidebook**. Geneva: OMS, 2021. Disponível em: <https://iris.who.int/bitstream/handle/10665/348069/9789240035546-eng.pdf>. Acesso em: 22 out. 2025.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. **Situación de la enfermería en el mundo 2020**: resumen de orientación. [Geneva]: OMS, 2020a. Disponível em: <https://www.paho.org/es/documentos/situacion-enfermeria-mundo-2020-resumen-orientacion>. Acesso em: 22 out. 2025.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. **State of the world's nursing 2020**: investing in education, jobs and leadership. Geneva: WHO, 2020b. Disponível em: <https://iris.who.int/handle/10665/331677>. Acesso em: 22 out. 2025.

ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE. **La situación de la enfermería en la Región de las Américas**: informe final del Foro Regional para el Avance de la Enfermería en América Latina: 8 y 9 de septiembre del 2022. Washington, D.C.: OPAS, 2023. Disponível em: <https://iris.paho.org/handle/10665.2/57220>. Acesso em: 24 out. 2025.

PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO. **Presença da LGPD com relevância nas decisões judiciais quase dobraram no Brasil em um ano**. Brasília, DF: PNUD, 21 dez. 2023. Disponível em: <https://www.undp.org/pt/brazil/news/presenca-da-lgpd-com-relevancia-nas-decisoes-judiciais-quase-dobraram-no-brasil-em-um-ano>. Acesso em: 10 ago. 2024.

RAMOS, E. M. B. *et al.* Questões éticas e perspectiva jurídica da proteção de dados. **Cadernos Ibero-Americanos de Direito Sanitário**, Brasília, DF, v. 10, n. 3, p. 172-190, jul./set. 2021. DOI: <https://doi.org/10.17566/ciads.v10i3.796>. Disponível em: <https://www.cadernos.prodisa.fiocruz.br/index.php/cadernos/article/view/796>. Acesso em: 19 set. 2024.

REDE INTERAGENCIAL DE INFORMAÇÕES PARA A SAÚDE. **Indicadores básicos para a saúde no Brasil: conceitos e aplicações**. 2. ed. Brasília, DF: Opas, 2008. Disponível em: <https://www.ripsa.org.br/publicacoes/indicadores-basicos-para-a-saude-no-brasil-livro-verde/>. Acesso em: 22 out. 2025.

RIBEIRO, A. T. **Os impactos da Lei Geral de Proteção de Dados na pesquisa clínica: aspectos contratuais**. 2022. Trabalho de Conclusão de Curso (Especialização em Direito Contratual) – Faculdade de Direito, Pontifícia Universidade Católica de São Paulo, São Paulo, 2022. Disponível em: <https://repositorio.pucsp.br/jspui/handle/handle/31470>. Acesso em: 19 set. 2024.

SALGADO, M. P. C. *et al.* **Nudos críticos sobre la desigualdad de género**. Buenos Aires: CLACSO, 2022. (Becas de investigación).

SANTOS, B. M. P. dos *et al.* Perfil e essencialidade da enfermagem no contexto da pandemia da COVID-19. **Ciência & Saúde Coletiva**, Rio de Janeiro, v. 28, p. 2785-2796, out. 2023. DOI: <https://doi.org/10.1590/1413-812320232810.09772023>. Disponível em: <https://www.scielo.br/j/csc/a/pKyr8tjqxVDbjC8XV7zDjNj/>. Acesso em: 22 out. 2025.

SARLET, G. B. S.; CALDEIRA, C. O consentimento informado e a proteção de dados pessoais de saúde na internet: uma análise das experiências legislativas de Portugal e do Brasil para a proteção integral da pessoa humana. **Civilistica.com**, Rio de Janeiro, v. 8, n. 1, p. 1-27, 2019. Disponível em: <https://civilistica.emnuvens.com.br/redc/article/view/411>. Acesso em: 10 ago. 2024.

SARLET, G. B. S.; MOLINARO, C. A. Questões tecnológicas, éticas e normativas da proteção de dados pessoais na área da saúde em um contexto de big data. **Revista Brasileira de Direitos Fundamentais & Justiça**, Porto Alegre, v. 13, n. 41, p. 183-212, 2020. DOI: <https://doi.org/10.30899/dfj.v13i41.811>. Disponível em: <https://dfj.emnuvens.com.br/dfj/article/view/811>. Acesso em: 10 ago. 2024.

SCHEFFER, M. (coord.). **Demografia médica no Brasil 2023**. São Paulo: FMUSP; AMB, 2023. Disponível em: https://amb.org.br/wp-content/uploads/2023/02/DemografiaMedica2023_8fev-1.pdf. Acesso em: 22 out. 2025.

SCHLOSSER, F. *et al.* Strategies for talent engagement and retention of Brazilian nursing professionals. **Revista Brasileira de Enfermagem**, São Paulo, v. 75, n. 6, e2022v75n6refl, 2022. DOI: <https://doi.org/10.1590/0034-7167-2022v75n6refl>. Disponível em: <https://www.scielo.br/j/reben/a/PDDsPQh6M7rPBNxTyMCgHCB/?lang=en>. Acesso em: 22 out. 2025.

SILVA, M. C. N. da; MACHADO, M. H. Sistema de saúde e trabalho: desafios para a enfermagem no Brasil. **Ciência & Saúde Coletiva**, Rio de Janeiro, v. 25, n. 1, p. 7-13, jan. 2019. DOI: <https://doi.org/10.1590/1413-81232020251.27572019>. Disponível em: <https://www.scielo.br/j/csc/a/wqFyYK4y49f8WZPmkvrwVsQ/?lang=pt>. Acesso em: 22 out. 2025.

SOUSA, A. *et al.* A comprehensive health labour market framework for universal health coverage. **Bulletin of the World Health Organization**, Geneva, v. 91, p. 892-894, 2013. DOI: <https://doi.org/10.2471/BLT.13.118927>. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC3853957/pdf/BLT.13.118927.pdf>. Acesso em: 22 out. 2025.

SOUZA, H. S. de *et al.* The Brazilian nursing workforce faced with the international trends: an analysis in the International Year of Nursing. **Physis: Revista de Saúde Coletiva**, São Paulo, v. 31, n. 1, e310111, 2021a. DOI: <https://doi.org/10.1590/S0103-73312021310111>. Disponível em: <https://www.scielo.br/j/physis/a/9LWQsVhYPbjYLkm8fsKqncf/?lang=en>. Acesso em: 22 out. 2025.

SOUZA, P. F. de *et al.* Condições laborais e a pandemia por covid-19: considerações dos gestores em enfermagem. *In*: SOARES, S. S. S. *et al.* (org.). **O complexo mundo do trabalho em saúde e as implicações para enfermagem**. Belo Horizonte: Synapse Editora, 2021b. p. 99-109.

STERMAN, J. **Business dynamics**: system thinking and modeling for a complex world. Boston: Irwin/McGraw-Hill, 2000. v. 1.

TRANSPARÊNCIA BRASIL. [**Portal Transparência Brasil**]. [S. l.: s. n.], [202-?]. Disponível em: <https://www.transparencia.org.br/>. Acesso em: 29 out. 2025.

ZAGANELLI, M. V.; BLINDA FILHO, D. L. A Lei Geral de Proteção de Dados e suas implicações na saúde: as avaliações de impacto no tratamento de dados no âmbito clínico-hospitalar. **Revista de Bioética y Derecho**, Barcelona, n. 54, p. 215-232, 2022. DOI 10.1344/rbd2021.54.36005. Disponível em: <https://scielo.isciii.es/pdf/bioetica/n54/1886-5887-bioetica-54-215.pdf>. Acesso em: 1º ago. 2024.

Apêndices

Apêndice A – Descrição das Bases

Cadastro Geral de Empregados e Desempregados (Caged)

O Caged é uma das principais fontes de informação sobre o mercado de trabalho formal com abrangência nacional e periodicidade mensal. Criado com o objetivo de monitorar e fiscalizar as admissões e dispensas de trabalhadores regidos pela Consolidação das Leis do Trabalho (CLT), o Caged desempenha um papel crucial na análise de dinâmicas empregatícias no País. No contexto da força de trabalho em enfermagem, o Caged permite a identificação de padrões de crescimento ou retração dos postos de trabalho, bem como análises sobre a estabilidade, a distribuição regional e a demanda por profissionais do setor.

Origem dos dados: os dados são disponibilizados de forma aberta, sem necessidade de cadastro, no endereço: <https://www.gov.br/trabalho-e-emprego/pt-br/assuntos/estatisticas-trabalho/microdados-rais-e-caged>.

Para baixar as tabelas, faz-se necessário conectar-se ao servidor FTP no endereço: <ftp://ftp.mtps.gov.br/pdet/microdados/>.

Estrutura: o Caged disponibiliza dados a partir de janeiro de 2007. Essa base, até o momento da produção deste texto (dezembro de 2024), continua sendo mantida. Desde o início de vigência do cadastro, as tabelas apresentaram alterações em seus campos, bem como a base principal foi descrita com dois nomes, a saber: Caged (Jan/07 a Dez/2019) e Novo Caged (Jan/20 aos dias atuais).

Principais Dimensões: município; UF; código da ocupação; grau de instrução; idade; raça/cor; sexo; competência da movimentação; horas contratadas; tipo de movimento (admissão/desligamento); tipo de deficiência; e salário mensal.

Restrições pela LGPD: dados descaracterizados.

Possibilidades de Investigação: os dados disponibilizados pelo Caged oferecem diversas possibilidades de investigação para a análise da força de trabalho em enfermagem no Brasil. A partir das informações de admissão e desligamento, é viável realizar análises longitudinais sobre a evolução do mercado de trabalho formal, identificando tendências

de crescimento ou retração do emprego. Além disso, a segmentação dos dados por características como idade, sexo, raça/cor e grau de instrução possibilita investigar questões relacionadas à diversidade e inclusão no mercado. A análise regional permite compreender como os profissionais de enfermagem estão distribuídos entre estados e municípios, identificando áreas com maior deficiência de recursos humanos e possíveis desigualdades na oferta de serviços de saúde. A investigação das movimentações laborais também possibilita explorar a estabilidade e a rotatividade desses profissionais, gerando *insights* para o desenvolvimento de políticas de retenção e a identificação de fatores associados à vulnerabilidade ocupacional na enfermagem.

Limitações: a base de dados apresenta algumas limitações importantes. Primeiramente, ela não registra contratações de trabalhadores informais e autônomos, nem de profissionais que prestam serviços de forma contínua e regular, mas que são remunerados por meio de pessoa jurídica (PJ). Além disso, a base não fornece o quantitativo total da força de trabalho nos períodos analisados, limitando-se a registrar apenas as entradas e saídas dos postos de trabalho.

Outro aspecto relevante é a mudança significativa no número de atributos ao longo de três períodos distintos: entre 2010 e 2018, de 2018 a 2020 e a partir de 2020, com o advento do Novo Caged. Essas alterações dificultam a análise longitudinal de algumas variáveis, já que determinados atributos não estão presentes em todas as versões da base ao longo do tempo. Um exemplo disso é a variável “categoria”, introduzida pelo Novo Caged, que descreve o tipo de contrato do trabalhador com base na CLT. Embora essa variável tenha ampliado significativamente as possibilidades de análise do mercado de trabalho formal, ela se encontra disponível apenas a partir de 2020, impossibilitando uma análise consistente desse atributo em anos anteriores.

Link para o dicionário de dados: <https://abrir.link/AloKW>

Link para a análise exploratória: <https://abrir.link/EcQLD>

Cadastro Nacional de Estabelecimentos de Saúde (CNES)

O CNES é o sistema oficial de governo responsável por registrar informações detalhadas sobre todos os estabelecimentos de saúde do País, independentemente de sua natureza jurídica ou de sua vinculação ao Sistema Único de Saúde (SUS).

Administrado pelo Ministério da Saúde, o CNES tem como principais objetivos cadastrar e manter atualizadas as informações sobre os estabelecimentos de saúde,

incluindo dados sobre infraestrutura, recursos físicos, trabalhadores e os serviços ofertados. Dessa forma, o sistema reflete a realidade da capacidade instalada e da força de trabalho dos serviços de saúde no Brasil, abrangendo tanto unidades públicas quanto privadas, sejam conveniadas ao SUS ou não.

As informações contidas no CNES são integradas a outros sistemas de informação em saúde, permitindo que gestores, pesquisadores e profissionais possam acessar e utilizar esses dados para diversas finalidades. Além disso, o sistema fornece informações acessíveis ao público sobre a localização, a disponibilidade e o funcionamento dos serviços de saúde, possibilitando aos gestores e à sociedade melhor conhecimento sobre a oferta assistencial em cada território.

No contexto da força de trabalho em enfermagem, o CNES é fonte essencial para entender a distribuição dos profissionais de saúde, permitindo análises sobre a quantidade e as características dos trabalhadores por estabelecimentos, municípios, estados e regiões. Essas informações são fundamentais para apoiar o planejamento e a alocação de recursos, contribuindo para a formulação de políticas públicas voltadas à melhoria da cobertura e da equidade no acesso aos serviços de saúde.

Origem dos dados: os dados são disponibilizados de forma aberta, sem necessidade de cadastro, no endereço <https://CNES.datasus.gov.br/pages/profissionais/extracao.jsp>

Estrutura: O CNES disponibiliza, dentre outros dados por estabelecimento e por profissionais, com periodicidade mensal, desde julho de 2007, sem ocorrência de alterações na estrutura dos seus campos.

Principais Dimensões:

Tabela de Estabelecimentos: Período (mês/ano); Município; Estado; Bairro; CEP; CNES do Estabelecimento.

Tabela de Profissionais: Período (mês/ano); Município; Estado; CNES do estabelecimento; descrição da profissão; Sexo; Natureza Jurídica do estabelecimento; Gestão (Federal, Estadual, Municipal); Pertence ao SUS; Carga horária trabalhada).

Restrições pela LGPD: dados descaracterizados.

Possibilidades de Investigação: o CNES oferece amplas possibilidades de investigação sobre a força de trabalho em saúde. Com os dados disponíveis, é possível realizar análises detalhadas da distribuição geográfica dos profissionais de enfermagem,

permitindo a identificação de possíveis desigualdades na oferta de recursos humanos em saúde entre diferentes estados e municípios.

A análise da carga horária dos profissionais vinculados a estabelecimentos de saúde possibilita uma compreensão mais profunda sobre o perfil de trabalho dos enfermeiros, incluindo aspectos como tempo de dedicação e variações regionais. Esses dados ajudam a identificar áreas com sobrecarga de trabalho e a detectar padrões de superexploração laboral.

Outro aspecto relevante refere-se aos tipos de vínculos dos profissionais com diferentes tipos de gestão (federal, estadual e municipal) e à natureza jurídica dos estabelecimentos (públicos ou privados). Essas informações são essenciais para compreender o papel de cada setor na contratação de profissionais de enfermagem e identificar lacunas significativas. Além disso, a vinculação dos estabelecimentos ao SUS permite analisar a participação dos profissionais de enfermagem na prestação de serviços públicos, destacando regiões com maior necessidade de expansão dessa força de trabalho para assegurar um atendimento adequado à população.

A estrutura detalhada dos dados também permite explorar a capacidade instalada dos estabelecimentos, sendo fundamental para correlacionar as disponibilidades de infraestrutura e de recursos humanos. Essas informações auxiliam gestores e formuladores de políticas na alocação eficiente de recursos e no planejamento de intervenções que promovam maior equidade na cobertura dos serviços de saúde, especialmente em áreas mais vulneráveis.

Limitações: a primeira grande dificuldade em relação ao CNES foi escolher entre os vários conjuntos de dados (*datasets*) publicamente disponibilizados. Diferentes acessos oferecem *datasets* distintos entre si, não quanto aos atributos, mas quanto aos dados em si mesmos. A base disponibilizada no site do DataSUS apresenta uma série de inconsistências, como dados ausentes, alternância no registro de sexo para a mesma pessoa (variando entre masculino e feminino ao longo do tempo) e alterações no tipo de administração dos estabelecimentos de saúde, que oscilam entre as esferas municipal, federal e estadual, diversas vezes em um mesmo período. Além disso, no momento do acesso, foi possível obter apenas as bases referentes aos profissionais, uma vez que os dados das dimensões de estabelecimentos estavam ausentes.

Diante dessas limitações, o projeto buscou outras fontes. Um segundo servidor DataSUS e o ElastiCNES ofereceram outros conjuntos de dados que ainda estão em análise. Entretanto é possível registrar que o portal ElastiCNES organiza de forma mais

enxuta e estruturada as dimensões de profissionais e estabelecimentos, superando algumas das dificuldades encontradas na base do DataSUS. Apesar disto, o portal ElastiCNES apresenta dois pontos negativos: (1) a dificuldade da extração de um grande volume de dados; e (2) o desconhecimento sobre o procedimento para extração e tratamento desses dados. Estes dois conjuntos de dados, entretanto, ainda estão em análise.

Censo Demográfico

Descrição: o Censo Demográfico é a pesquisa estatística mais ampla e detalhada realizada no País, cuja elaboração, execução e divulgação de resultados é de responsabilidade do Instituto Brasileiro de Geografia e Estatística (IBGE). Seu principal objetivo é oferecer um panorama preciso das características sociodemográficas e das condições de vida dos brasileiros, fornecendo subsídios para a formulação de políticas públicas em diferentes esferas governamentais, bem como para orientar decisões estratégicas de investimento da iniciativa privada.

Por meio do Censo, são coletados dados sobre a distribuição da população, composição etária, escolaridade, condições de moradia e acesso a serviços, abrangendo desde áreas urbanas densamente povoadas até localidades rurais mais isoladas. Esses dados representam uma fonte confiável e sistemática de informações sobre a situação de vida nos municípios, possibilitando um diagnóstico mais qualificado das disparidades regionais e das especificidades locais.

Para a análise da força de trabalho em enfermagem, o Censo Demográfico desempenha um papel fundamental ao possibilitar o cruzamento de informações sobre distribuição populacional e indicadores socioeconômicos com a oferta de serviços de saúde. Esses dados permitem identificar áreas com maiores carências de profissionais de saúde, analisar a adequação da oferta de serviços de enfermagem às necessidades locais e orientar políticas públicas voltadas à promoção da equidade no acesso a cuidados de saúde, especialmente em regiões com maiores necessidades.

Origem dos dados: os dados são disponibilizados de forma aberta, sem necessidade de cadastro, no endereço: <https://sidra.ibge.gov.br/pesquisa/censo-demografico/demografico-2022/universo-populacao-e-domicilios-situacao-urbana-ou-rural>.

Estrutura: o projeto utilizou os dados de acesso público disponibilizados pelo IBGE, dos Censos de 2000, 2010 e 2022. A observação dos dados tabulados originais desses

três censos revelou além da criação de novas planilhas, a ocorrência de alterações em campos de tabelas já existentes de forma a adicionar ou excluir colunas ao longo dos anos.

Principais Dimensões: município; ano; características da população; características das residências.

Restrições pela LGPD: dados descaracterizados e agregados por dimensão.

Possibilidades de Investigação: uma das principais contribuições do censo é a possibilidade de analisar a distribuição demográfica e socioeconômica da população em conjunto com a oferta de serviços de saúde, permitindo identificar regiões com maior carência de profissionais de enfermagem. Essa abordagem auxilia na compreensão da relação entre a densidade populacional e a presença de serviços de saúde, orientando políticas de alocação de recursos humanos que promovam uma distribuição mais equitativa.

Outra possibilidade de investigação envolve a análise das características socioeconômicas das regiões, como o nível de escolaridade e as condições de moradia, fatores associados à demanda por cuidados de saúde e que influenciam o planejamento da oferta de profissionais de enfermagem. Com base nesses dados, é possível identificar áreas mais vulneráveis e direcionar esforços para melhorar o acesso e a qualidade dos serviços de enfermagem.

O Censo também possibilita investigar o perfil das comunidades atendidas, permitindo comparações entre áreas urbanas e rurais e identificando diferenças nas necessidades de cuidados de saúde nesses contextos. Embora a frequência da coleta de dados seja uma limitação, uma vez que possui periodicidade original decenal, o conjunto de informações obtido pode subsidiar o desenvolvimento de políticas públicas e estratégias para reduzir as desigualdades no acesso aos cuidados de saúde em diferentes regiões do Brasil.

Limitações: a base apresenta as características do universo populacional em um determinado ano, tornando o processo de análise de evolução por período dependente de cenários estatísticos. Essa conexão se torna ainda mais evidente diante das dificuldades enfrentadas na publicação do último Censo Demográfico, originalmente previsto para 2020, e realizado apenas em 2022, com atrasos significativos na disponibilização dos dados.

Esses atrasos impactam diretamente a utilização do censo como ferramenta de análise em estudos longitudinais, principalmente em áreas como a demografia da enfermagem, que exige dados atualizados e detalhados para avaliar a distribuição e as tendências populacionais. A falta de informações recentes e completas dificultou a compreensão de mudanças recentes no perfil demográfico e socioeconômico das regiões, comprometendo a precisão de projeções e estratégias baseadas nesses dados.

Além disso, sua periodicidade limita sua capacidade de capturar mudanças mais dinâmicas, como fluxos migratórios, transformações no mercado de trabalho e variações na composição da força de trabalho em saúde. Esses fatores tornam necessária sua integração com outras fontes de dados mais frequentes, como a PNAD Contínua, para complementar lacunas e fornecer um panorama mais detalhado e atualizado.

Link para o dicionário de dados: <https://abrir.link/jVBHC>

Link para a análise exploratória: <https://abrir.link/rDmmq>

Censo da Educação Básica (Censo Escolar)

Descrição: o Censo da Educação Básica, ou Censo Escolar, é a pesquisa estatística educacional realizada em âmbito nacional, considerada essencial para o entendimento da estrutura e dinâmica da educação básica no Brasil. Coordenado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), sua execução ocorre em parceria com secretarias estaduais e municipais de educação, além de contar com a participação de escolas públicas e privadas de todo o território nacional.

A pesquisa é realizada anualmente e abrange múltiplas modalidades e etapas educacionais, incluindo educação infantil, ensino fundamental, ensino médio, educação especial, educação de jovens e adultos (EJA) e educação profissional e tecnológica. Essas informações são fundamentais para o monitoramento do sistema educacional brasileiro e para a avaliação da efetividade das políticas públicas implementadas.

No contexto da força de trabalho em enfermagem, o Censo Escolar fornece, a partir de 2023, dados estratégicos sobre a formação de técnicos em enfermagem. A análise dessas informações possibilita identificar áreas com maior ou menor oferta de cursos, além de traçar um panorama do perfil educacional dos futuros profissionais de saúde.

Origem dos dados: os dados são disponibilizados de forma aberta, sem necessidade de cadastro, no endereço: <https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-escolar/resultados>.

Estrutura: o Censo da Educação Básica tem seus dados disponibilizados de 1995 até 2023, considerando o momento em que este estudo está sendo escrito. Durante esse tempo, as tabelas apresentaram alterações em seus campos, sendo adicionadas e excluídas colunas.

Principais Dimensões: ano; região; município; UF; escola oferece ensino técnico; quantidade de matrículas do ensino médio; quantidade de docentes do ensino médio.

Restrições pela LGPD: dados descaracterizados e agregados por dimensão.

Possibilidades de Investigação: os dados do Censo da Educação Básica possibilitam uma ampla gama de investigações relacionadas à formação de técnicos em enfermagem e à sua distribuição pelo País. Com as informações sobre a oferta de ensino técnico e a quantidade de matrículas no ensino médio, é possível mapear a disponibilidade de cursos técnicos em enfermagem em diferentes regiões, estados e municípios, identificando áreas com oferta insuficiente para atender às necessidades locais de saúde. A análise dessas informações também contribui para compreender a relação entre a formação técnica e as demandas do mercado de trabalho, permitindo avaliar a capacidade regional de atender às exigências do setor de saúde.

Limitações: até o ano de 2022, a base não continha informações sobre o tipo de formação técnica, o que tornava inviável filtrar principalmente as escolas que oferecem cursos técnicos de enfermagem. Além disso, também não são fornecidos dados sobre a quantidade de egressos por ano, informação crucial para a análise do fluxo desses profissionais ingressando no mercado de trabalho.

A ausência desses dados impacta diretamente a capacidade de compreender a relação entre este tipo de formação educacional e as demandas do mercado de trabalho. Sem informações detalhadas sobre as escolas técnicas de enfermagem, não é possível identificar disparidades regionais na oferta desses cursos. Da mesma forma, a inexistência de dados sobre os egressos impede uma avaliação precisa sobre o número de profissionais efetivamente disponíveis para atender às necessidades do setor de saúde em diferentes regiões e dificulta o monitoramento de tendências no setor educacional e sua relação com o mercado de trabalho, como a taxa de absorção de recém-formados em posições formais e informais.

Essa lacuna limita a possibilidade de análises estratégicas que poderiam direcionar investimentos para regiões com menor oferta educacional ou maior demanda por profissionais.

O impacto se torna ainda mais crítico em um contexto de mudanças demográficas e aumento das demandas por serviços de saúde, em que é essencial alinhar a formação profissional à evolução das necessidades populacionais.

Link para o dicionário de dados: <https://abrir.link/pSPUk>

Link para a análise exploratória: <https://abrir.link/zuwRX>

Censo da Educação Superior

Descrição: o Censo da Educação Superior constitui uma das principais bases de dados para a análise do ensino superior no Brasil, sendo realizado anualmente pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep).

O levantamento abrange o conjunto de instituições cadastradas no sistema e-MEC e coleta informações sobre diversos elementos estruturais e operacionais, como instituições, locais de oferta, cursos, vagas disponíveis, corpo docente e perfil dos estudantes matriculados em cursos de graduação. Esses dados oferecem uma visão abrangente sobre a dimensão e a distribuição do ensino superior no Brasil, permitindo análises comparativas ao longo do tempo.

No campo da enfermagem, o Censo da Educação Superior é uma fonte valiosa para compreender a formação de profissionais desse nível de ensino. Ele possibilita identificar a localização e a capacidade dos cursos de graduação em enfermagem, avaliar o número de ingressantes e concluintes, bem como traçar um perfil dos cursos com formação presencial ou à distância. Analisando o censo, também é possível planejar políticas educacionais para a adequação da formação acadêmica às demandas de saúde pública e privada no Brasil.

Origem dos dados: os dados são disponibilizados de forma aberta, sem necessidade de cadastro, no endereço: <https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-da-educacao-superior/resultados>

Estrutura: o Censo da Educação Superior tem seus dados disponibilizados no período de 2009 a 2023, considerando o momento da produção deste texto. Durante esse tempo, as tabelas apresentaram alterações em seus campos, acrescentando e excluindo colunas ao longo dos anos.

Principais Dimensões:

Tabela de Instituições de Ensino (IES): ano; região; município; UF; IES na capital; modalidade de ensino; organização acadêmica; categoria administrativa; características do corpo técnico; características do corpo docente; infraestrutura da IES.

Tabela de Curso: ano; região; município; UF; área geral do curso; área específica do curso; área detalhada do curso; modalidade de ensino; nível acadêmico; curso gratuito; número e características das vagas; número e características dos ingressantes; número e características do matriculados; número e características do concluintes.

Restrições pela LGPD: dados descaracterizados e agregados por dimensão.

Possibilidades de Investigação: o Censo da Educação Superior oferece amplas oportunidades para investigações relacionadas à formação de enfermeiros no Brasil. Com os dados sobre a distribuição e a capacidade dos cursos de graduação em enfermagem, é possível analisar a concentração geográfica dos cursos, identificando desigualdades regionais e áreas com maior ou menor oferta de formação superior para essa categoria profissional. Além disso, as informações sobre modalidades de ensino (presencial e a distância) e características dos ingressantes e concluintes permitem avaliar tendências de crescimento na educação a distância e o impacto dessas modalidades na qualificação dos profissionais.

O censo também possibilita a exploração da relação entre a quantidade de vagas ofertadas e a demanda de estudantes, subsidiando a formulação de políticas educacionais que busquem ajustar a oferta à demanda por serviços de saúde. Por fim, o Censo da Educação Superior permite a análise das características dos concluintes e sua relação com a dinâmica do mercado de trabalho, contribuindo para identificar se o número de graduados é suficiente para atender às demandas de saúde pública e privada nas diferentes regiões do Brasil. Embora a ausência de dados financeiros e de informações detalhadas sobre os planos pedagógicos, limite algumas análises, as informações disponíveis são fundamentais para a formulação de estratégias voltadas à qualificação da força de trabalho em enfermagem e ao fortalecimento do sistema de saúde.

Limitações: a base das IES não está segmentada por curso, o que impossibilita a análise do corpo técnico e de docentes de forma individual. Além disso, não há dados financeiros das IES e dos cursos, dificultando o planejamento estratégico e a formulação de políticas de investimento. Também não estão disponíveis informações sobre os planos pedagógicos, o que limita a compreensão da estrutura e da qualidade dos programas ofertados.

A ausência de segmentação por curso impede a identificação de características específicas das equipes técnicas e docentes em áreas como enfermagem, dificultando análises detalhadas sobre o perfil e a qualificação dos profissionais diretamente envolvidos na formação acadêmica. Isso é especialmente prejudicial em áreas críticas como a saúde, em que o alinhamento entre o corpo docente e as demandas do mercado de trabalho é essencial para garantir uma formação que responda às necessidades da sociedade.

A falta de dados financeiros também compromete a capacidade de planejar investimentos de forma eficiente. Sem informações detalhadas, torna-se impossível avaliar a sustentabilidade econômica dos cursos e das IES, bem como identificar desigualdades regionais no financiamento do ensino superior. Essa lacuna dificulta a formulação de políticas públicas que garantam a equidade na alocação de recursos e no acesso à educação superior de qualidade.

Essas limitações impactam diretamente o planejamento e a execução de estratégias que visem a melhoria da formação superior em áreas essenciais, como enfermagem. Em um cenário de crescente demanda por profissionais de saúde qualificados, a ausência dessas informações fragiliza a capacidade de gestores e formuladores de políticas de promoverem iniciativas baseadas em evidências para aprimorar a formação, a alocação de recursos e a qualidade do ensino superior no Brasil.

Link para o dicionário de dados: <https://abrir.link/TPKBx>

Conselho Federal de Enfermagem (Cofen)

Descrição: a base de dados do Conselho Federal de Enfermagem (Cofen) reúne informações sobre o registro dos profissionais, compondo um retrato abrangente da categoria no Brasil. Essa base inclui dados demográficos e profissionais de enfermagem (enfermeiros, técnicos e auxiliares de enfermagem), como idade, sexo, local de registro e outras características essenciais para a análise do perfil da força de trabalho na enfermagem.

Os dados disponibilizados pelo Cofen permitem não apenas monitorar a quantidade de profissionais em atividade, mas também compreender sua distribuição regional e características específicas, sendo uma importante ferramenta para a tomada de decisão e o planejamento de políticas públicas voltadas ao fortalecimento da enfermagem. Essa base é fundamental para avaliar a capacidade instalada em termos

de recursos humanos, bem como para identificar possíveis carências de profissionais de enfermagem em diferentes regiões do Brasil.

Além disso, o Cofen contribui para análises sobre o crescimento do setor, acompanhando a evolução no número de registros ao longo do tempo. Essas informações são cruciais para compreender a dinâmica de entrada e saída de profissionais no mercado de trabalho e, assim, orientar iniciativas que promovam a sustentabilidade da força de trabalho e a qualidade dos serviços de saúde prestados.

Origem dos dados: os dados foram disponibilizados de forma reservada, por meio de um convênio entre a UERJ e o Cofen.

Estrutura: o Cofen disponibiliza seus dados por movimentação do profissional, ou seja, cada nova ação (cadastro ou alterações) gera um novo registro na base. Assim, a periodicidade é diária, com dados disponíveis desde a fundação do Conselho, em 1973, até fevereiro de 2024, considerando o momento da produção deste texto. Durante esse período, não houve alteração na estrutura dos campos.

Principais Dimensões: categoria do profissional; sexo; idade; local de nascimento (município e país); status da inscrição; regional de registro; data de cadastro.

Restrições pela LGPD: dados descaracterizados.

Possibilidades de Investigação: a base de dados oferece diversas possibilidades de investigação sobre a força de trabalho em enfermagem no Brasil. Uma das principais análises possíveis diz respeito à distribuição dos profissionais por região, permitindo identificar disparidades geográficas na alocação de enfermeiros, especialmente entre diferentes estados e municípios. Essa análise possibilita mapear onde há escassez ou excesso de profissionais, facilitando a formulação de políticas públicas voltadas à redistribuição equitativa dos recursos humanos em saúde.

Além disso, os dados demográficos, como idade e sexo, possibilitam estudos sobre o perfil dos enfermeiros ativos no País, o que é essencial para compreender tendências de envelhecimento da força de trabalho, diversidade de gênero e entrada de novas gerações na profissão.

A análise do status de inscrição dos profissionais, como registros ativos, cancelados ou suspensos, permite avaliar a dinâmica de permanência dos enfermeiros no mercado, oferecendo subsídios para entender fatores que influenciam a desistência da profissão e propor estratégias de retenção.

Outra possibilidade de investigação envolve o acompanhamento das alterações no cadastro dos profissionais ao longo do tempo, o que pode ser utilizado para identificar tendências na migração dos enfermeiros entre regiões ou mudanças de categorias profissionais. Isso é particularmente relevante para avaliar a mobilidade dos profissionais e identificar os fatores que motivam a realocação, contribuindo para estratégias de atração e fixação de profissionais em áreas carentes.

Por fim, o histórico de registros do Cofen desde sua fundação até o presente possibilita análises longitudinais que permitem compreender a evolução do número de profissionais de enfermagem ao longo das décadas, além de fornecer um panorama sobre o crescimento da categoria em resposta a eventos críticos, como crises econômicas ou pandemias. Essas investigações são fundamentais para prever a demanda futura por profissionais e assegurar a sustentabilidade da força de trabalho em enfermagem no Brasil.

Limitações: apesar do esforço coordenado entre os Conselhos Regionais de Enfermagem (Corens) e o Cofen para padronizar campos, tratar inconsistências e realizar o pré-processamento adequado, a base do Cofen ainda apresenta diversos tipos de limitações. Entre as principais dificuldades, destaca-se a falta de padronização nos campos de resposta, que resulta em dados heterogêneos e pouco consistentes. A base também inclui registros muito antigos, como, por exemplo, de pessoas nascidas no início do século XX, o que pode não refletir com precisão o perfil atual da força de trabalho. A presença de campos de duplicados também é um problema recorrente, dificultando a análise precisa dos dados. Ademais, os dados refletem as inscrições realizadas nos conselhos regionais. Portanto, um mesmo profissional pode possuir mais de uma inscrição, seja na mesma categoria (inscrição secundária em outra regional) ou em categorias distintas (na mesma ou outra regional), sendo contabilizado na base mais de uma vez.

Outro ponto crítico é a concentração dos registros mais recentes: quase 50% do total de registros (7.778.067) refere-se ao período de 2010 em diante. Essa distribuição desigual dificulta a análise histórica do número de novos registros em períodos anteriores, podendo levar à subestimação de dados em décadas passadas. Essa limitação compromete a possibilidade de traçar séries históricas consistentes e detalhadas.

Apesar dessas limitações, a base possui um potencial significativo para estudos acadêmicos, considerando ser o único conjunto de dados confiável que cobre o universo

completo dos profissionais de enfermagem registrados no Brasil. Contudo, no estado atual, as análises realizadas estão sujeitas a vieses e inconsistências que afetam a confiabilidade dos resultados.

Link para o dicionário de dados: <https://abrir.link/BIUPP>

Link para a análise exploratória: <https://abrir.link/uwzpf>

e-Gestor Atenção Básica

Descrição: o e-Gestor Atenção Básica (e-Gestor AB) é uma plataforma online desenvolvida pelo Ministério da Saúde para centralizar o acesso aos sistemas e dados relacionados à atenção básica em saúde no Brasil. A ferramenta funciona como um hub de informações, permitindo que gestores estaduais e municipais planejem, acompanhem, e avaliem ações de saúde de forma integrada e organizada.

Embora não seja uma base de dados em si, o e-Gestor AB possibilita o acesso a informações estratégicas sobre equipes de saúde, indicadores de desempenho, cobertura populacional e recursos disponíveis. Essas funcionalidades facilitam o monitoramento das políticas de saúde e a tomada de decisões baseadas em evidências, tornando-se uma ferramenta essencial para a gestão e melhoria contínua da atenção básica, tanto em nível local quanto nacional.

Origem dos dados: foi realizada, no dia 31/01/2024, uma consulta ao endereço: <https://sisab.saude.gov.br/index.xhtml> com objetivo de extrair os dados no formato de tabelas. Os dados foram disponibilizados de forma aberta, sem necessidade de cadastro.

Estrutura: a plataforma apresentou tabelas com resultados agregados, mas não permitiu o download de microdados estruturados. Essa limitação foi determinante para a decisão de não utilizar esse recurso no âmbito do projeto, uma vez que a ausência de granularidade inviabilizou análises mais detalhadas.

Principais Dimensões: não analisado, considerando a impossibilidade de acessar microdados conforme descrito anteriormente.

Restrições pela LGPD: relatórios descaracterizados.

Possibilidades de Investigação: não analisado, devido à ausência de microdados estruturados, conforme o relato da seção de estrutura.

Limitações: não analisado, considerando a ausência de granularidade nos dados disponíveis.

Link para o dicionário de dados: <https://abrir.link/byiXy>

Link para a análise exploratória: <https://abrir.link/Dwcbe>

Exame Nacional de Desempenho dos Estudantes (Enade)

Descrição: o Exame Nacional de Desempenho dos Estudantes (Enade) é uma avaliação conduzida pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), com o objetivo de medir o desempenho acadêmico dos estudantes concluintes dos cursos de graduação no Brasil. O exame avalia o conhecimento dos estudantes em relação aos conteúdos específicos definidos pelas diretrizes curriculares de seus cursos, bem como competências e habilidades que contribuem para sua formação profissional e geral.

Com periodicidade anual, o Enade busca fornecer indicadores que contribuam para o aprimoramento da qualidade do ensino superior no País. No contexto da enfermagem, os dados do Enade permitem uma análise detalhada do desempenho dos estudantes, destacando áreas que necessitam de maior atenção nos programas de ensino. Essas informações são valiosas para instituições de ensino, gestores e formuladores de políticas públicas, auxiliando na promoção de melhorias contínuas na formação de futuros profissionais de saúde.

Origem dos dados: os dados são disponibilizados de forma aberta, sem necessidade de cadastro, no endereço: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enade>.

Estrutura: os dados do Enade abrangem o período de 2004 a 2022, considerando o momento da produção deste texto. Durante esse tempo, as tabelas passaram por alterações nos campos, com adições e exclusões de colunas ao longo dos anos, o que pode impactar a comparabilidade de algumas variáveis entre períodos.

Principais Dimensões: a base do Enade será incorporada futuramente ao *data lake* do projeto. Essa decisão foi tomada com base na necessidade de estruturar as principais bases de dados e relacioná-las às dimensões estratégicas do projeto, mesmo que elas não fossem usadas de imediato.

Restrições pela LGPD: os dados são descaracterizados em conformidade com a Lei Geral de Proteção de Dados (LGPD), garantindo a privacidade das informações pessoais dos estudantes.

Possibilidades de Investigação: não analisado, considerando o relato do item principais dimensões.

Limitações: não analisado, considerando o relato do item principais dimensões.

Link para o dicionário de dados: <https://abrir.link/mRKJe>

Link para a análise exploratória: <https://abrir.link/bouAS>

Instituto Nacional do Seguro Social (INSS)

Descrição: o Instituto Nacional do Seguro Social (INSS) é uma autarquia federal vinculada ao Ministério da Previdência Social, responsável pela gestão dos benefícios previdenciários no Brasil. O INSS tem como principal objetivo conceder e manter benefícios que garantem a proteção social dos trabalhadores e de seus dependentes, cobrindo situações de risco social, como doenças, invalidez, maternidade, morte e aposentadoria. A base de dados do INSS é uma fonte relevante de informações sobre a concessão de benefícios, como aposentadorias, pensões por morte e auxílios relacionados à incapacidade laboral.

No contexto da força de trabalho em enfermagem, os dados do INSS permitem análises valiosas sobre o perfil dos enfermeiros que solicitam benefícios previdenciários, como aposentadoria ou auxílio-doença, possibilitando compreender melhor as condições de trabalho e de saúde desses profissionais. Esse tipo de informação é fundamental para a formulação de políticas públicas voltadas à melhoria da saúde ocupacional, bem como ao desenvolvimento de estratégias para minimizar o desgaste e o afastamento de profissionais em função de problemas de saúde. A análise dos registros previdenciários também possibilita a avaliação da longevidade profissional e a identificação de padrões associados a doenças ocupacionais na área da enfermagem.

Origem dos dados: os dados são disponibilizados de forma aberta, sem necessidade de cadastro, no endereço: <https://dados.gov.br/dados/organizacoes/visualizar/instituto-nacional-do-seguro-social>.

Estrutura: os dados do INSS são disponibilizados em periodicidade mensal, organizados em quatro categorias: benefícios concedidos, emitidos, indeferidos e comunicações de

acidentes de trabalho. A abrangência temporal varia conforme o tipo de dado, sendo, para fins deste estudo, considerado o período entre 2019 a setembro de 2024.

Principais Dimensões:

Benefícios concedidos: Competência da concessão; Município; Estado; Tipo de aposentadoria; Classificação Internacional de Doenças; Data de nascimento; e Sexo.

Benefícios emitidos: Data de início de validade; Município de residência; Estado de residência; Município de pagamento; Estado de pagamento; Sexo; e e Valor líquido.

Benefícios indeferidos: Competência de indeferimento; Data do Indeferimento; Motivo indeferimento; Data Nascimento; Sexo; Município; e Estado.

Comunicação de acidentes de trabalho: Data Emissão Comunicação de acidente do trabalho; Data Acidente; Data Nascimento; Município do acidente; Município do empregador; Sexo; Natureza da lesão; Parte do corpo atingida; Indica óbito; e Código da ocupação do segurado.

Restrições pela LGPD: dados descaracterizados.

Possibilidades de Investigação: a base do INSS oferece diversas possibilidades de investigação relacionadas à população brasileira. A análise dos dados de benefícios concedidos e emitidos possibilita traçar um perfil da população que recorre à aposentadoria ou a auxílios por incapacidade, considerando variáveis como idade, sexo e motivo da concessão. Esses dados contribuem para identificar padrões de aposentadoria precoce e para levantar os principais fatores de afastamento dos solicitantes, fornecendo subsídios para a criação de políticas públicas voltadas à retenção de profissionais e à redução de afastamentos evitáveis.

Outro aspecto relevante de análise envolve as comunicações de acidentes de trabalho, que incluem informações sobre o tipo de lesão, a parte do corpo atingida e o município do acidente. Esses dados são especialmente importantes no contexto da enfermagem, pois permitem identificar os tipos mais comuns de acidentes que afetam os profissionais, bem como avaliar a distribuição desses eventos em diferentes regiões do país. Essa análise é essencial para identificar vulnerabilidades no ambiente de trabalho e desenvolver ações preventivas que promovam a segurança e a saúde dos profissionais da saúde.

Além disso, a base permite compreender características demográficas e ocupacionais, como idade, sexo e tipo de vínculo empregatício, influenciam a frequência e a gravidade dos incidentes. Com isso, é possível identificar quais segmentos da força de trabalho em enfermagem estão mais suscetíveis a riscos ocupacionais, contribuindo para o desenvolvimento de políticas específicas que reduzam o desgaste físico e mental dos enfermeiros e promovam condições de trabalho mais seguras e saudáveis.

Limitações: a base apresenta uma limitação importante: somente as comunicações de acidentes de trabalho incluem o código de ocupação do profissional. Isso impossibilita a realização de análises mais amplas sobre a força de trabalho em enfermagem nas demais tabelas do INSS, como aquelas relacionadas à concessão de benefícios ou auxílios. Essa restrição reduz o potencial de integração dos dados, dificultando a construção de um panorama completo sobre os impactos das condições laborais na saúde e no bem-estar dos profissionais de enfermagem.

A limitação referente à ausência do código de ocupação em tabelas além das comunicações de acidentes de trabalho é particularmente relevante para estudos sobre a força de trabalho em enfermagem. Essa ausência impede a identificação direta dos profissionais dessa área nas bases relacionadas a aposentadorias, auxílios por incapacidade ou outros benefícios, dificultando análises que conectem as condições laborais às necessidades de proteção social dessa categoria.

Link para o dicionário de dados: <https://abrir.link/ywyQG>

Link para a análise exploratória: <https://abrir.link/eDdAZ>

Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua)

Descrição: a Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua), conduzida pelo Instituto Brasileiro de Geografia e Estatística (IBGE), tem como objetivo principal produzir informações contínuas sobre a inserção da população no mercado de trabalho e suas características, como idade, sexo e nível de instrução. Além disso, permite o estudo do desenvolvimento socioeconômico do País através da produção de dados anuais sobre outras formas de trabalho, trabalho infantil, migração, entre outros temas, monitorando as flutuações e tendências da força de trabalho brasileira ao longo do tempo.

Implantada em caráter definitivo em 2012, a PNAD Contínua coleta dados em todo o território nacional, gerando indicadores trimestrais e anuais que permitem análises

detalhadas do mercado de trabalho e de temas socioeconômicos relevantes. A pesquisa utiliza uma amostra representativa de domicílios, abordando aspectos como emprego, renda, educação e acesso à tecnologia. Esses dados são essenciais para a elaboração de políticas públicas voltadas ao desenvolvimento socioeconômico do Brasil.

No contexto da força de trabalho em enfermagem, a PNAD Contínua permite análises sobre ocupação, posição na ocupação, existência de carteira de trabalho assinada, contribuição para a previdência, horas trabalhadas, desejo e disponibilidade para trabalhar mais horas, rendimentos e distribuição regional. Esses dados fornecem um panorama sobre a inserção dos profissionais de enfermagem no mercado formal e informal, contribuindo para entender a dinâmica ocupacional e identificar tendências ou mudanças no perfil desses profissionais ao longo do tempo.

Origem dos dados: os dados são disponibilizados de forma aberta, sem necessidade de cadastro, e podem ser acessados através do link: [IBGE, PNAD Contínua - Microdados](#).

Estrutura: os dados da PNAD Contínua são disponibilizados em diferentes periodicidades. Mensalmente, são apresentados indicadores relacionados à força de trabalho, disponíveis apenas no nível geográfico nacional. Trimestralmente, são disponibilizados indicadores relacionados à força de trabalho para todos os níveis de divulgação. Anualmente, a pesquisa aborda demais temas permanentes e indicadores complementares à força de trabalho.

A pesquisa contempla o período de 2012 a 2024, durante o qual ocorreram alterações nos campos das tabelas, como a inclusão e exclusão de variáveis. Os dados são divulgados em diversos níveis geográficos, abrangendo o Brasil, suas grandes regiões, as unidades da federação e vinte regiões metropolitanas que incluem municípios das capitais, como Manaus, Belém, Fortaleza, São Paulo, entre outras. Também são apresentados dados referentes aos municípios das capitais e à região integrada de desenvolvimento da Grande Teresina.

Principais Dimensões: entre as principais dimensões disponíveis na pesquisa estão o período, especificado por ano e trimestre, as características do trabalho, as características da escolaridade, as características da moradia, além das informações relacionadas aos estados e aos municípios das capitais.

Restrições pela LGPD: dados descaracterizados.

Possibilidades de Investigação: a PNAD Contínua oferece diversas possibilidades de análise relacionadas à força de trabalho em enfermagem. Primeiramente, a pesquisa permite comparar a inserção dos profissionais no mercado formal e informal, identificando vínculos trabalhistas regulares e benefícios, como previdência e férias remuneradas. Essa análise é fundamental para compreender a precarização do trabalho e as condições enfrentadas pelos enfermeiros no mercado laboral. Além disso, possibilita a análise da distribuição dos enfermeiros nas diferentes regiões do Brasil, cruzando variáveis como estado, região, escolaridade e rendimento médio. Com isso, é possível identificar disparidades regionais na remuneração e nas oportunidades de trabalho, o que orienta a alocação de recursos e o desenvolvimento de políticas voltadas à equidade no emprego.

Outro ponto relevante é a possibilidade de investigar a relação entre o nível de escolaridade dos profissionais e suas condições de trabalho, como rendimento e estabilidade no emprego. Essa análise oferece informações valiosas sobre a valorização dos profissionais com diferentes graus de qualificação e o impacto disso na sustentabilidade da força de trabalho. Além disso, a pesquisa possibilita a análise do impacto de mudanças econômicas e sociais ao longo dos anos no mercado de trabalho dos enfermeiros, como os efeitos de crises econômicas ou períodos de crescimento na empregabilidade e na remuneração desses profissionais.

Limitações: apesar dessas potencialidades, a PNAD Contínua possui algumas limitações que devem ser consideradas. Por ser uma pesquisa baseada em amostragem, os dados podem estar sujeitos a erros amostrais, o que exige cautela na interpretação dos resultados. Além disso, a precisão das informações depende da habilidade dos entrevistadores e da honestidade dos entrevistados, o que pode gerar inconsistências nos dados. Por fim, para o correto uso da base, é necessário o domínio de conhecimentos estatísticos e a utilização de softwares especializados para a análise de dados amostrais complexos.

Link para o dicionário de dados: <https://abrir.link/xLOxh>

Link para a análise exploratória: <https://abrir.link/mEuYj>

Pesquisa Nacional de Saúde (PNS)

Descrição: a Pesquisa Nacional de Saúde (PNS) é uma iniciativa conduzida pelo IBGE, em parceria com o Ministério da Saúde, destinada a coletar dados abrangentes sobre a saúde da população brasileira. Seu principal objetivo é fornecer informações

detalhadas sobre as condições de saúde, estilos de vida, acesso e uso de serviços de saúde, ações preventivas e continuidade dos cuidados, além de aspectos relacionados ao financiamento da assistência.

Realizada em âmbito nacional, a PNS utiliza uma amostra representativa de domicílios para investigar questões como prevalência de doenças crônicas, hábitos de vida, cobertura vacinal e utilização de serviços médicos. Esses dados são fundamentais para subsidiar decisões de políticas públicas e avaliar o desempenho do sistema de saúde, promovendo melhorias no atendimento e na qualidade dos serviços prestados.

Origem dos dados: os dados são disponibilizados de forma aberta, sem necessidade de cadastro, no endereço: <https://www.ibge.gov.br/estatisticas/sociais/saude/29540-2013-pesquisa-nacional-de-saude.html?edicao=9177&t=microdados>

Estrutura: a PNS tem seus dados disponibilizados entre 2013 e 2019, considerando o momento da produção deste texto. Comparando as duas bases, as tabelas apresentaram alterações em seus campos, adicionando e excluindo colunas.

Principais Dimensões: na data em que este estudo está sendo escrito, dezembro de 2024, os dados da PNS ainda não foram integrados ao *data lake*. A inserção ocorrerá futuramente no decorrer do projeto. Essa decisão foi tomada com base na necessidade de estruturar as principais bases de forma integrada e alinhada a todas as dimensões do projeto.

Restrições pela LGPD: dados descaracterizados.

Possibilidades de Investigação: não analisado, considerando o relato do item principais dimensões.

Limitações: As limitações da PNS também não foram analisadas nesta fase, devido à ausência de dados estruturados no *data lake*.

Link para o dicionário de dados: <https://abrir.link/HWTZE>

Link para a análise exploratória: <https://abrir.link/ircBZ>

Relação Anual de Informações Sociais (Rais)

Descrição: a Relação Anual de Informações Sociais (Rais) é uma base desenvolvida pelo governo brasileiro com o objetivo de fornecer dados abrangentes sobre a atividade

trabalhista no Brasil. Criada pelo Decreto n.º 76.900, de 1975, e regulamentada pelo Decreto n.º 10.854, de 2021, a Rais visa atender à necessidade de controle e monitoramento das relações de trabalho, além de contribuir para a formulação de políticas públicas voltadas ao mercado de trabalho e fornecer dados relevantes para a elaboração de estatísticas oficiais.

A Rais possuiu, durante seu período de publicação, natureza obrigatória e periodicidade anual, sendo preenchida por todos os estabelecimentos, tanto públicos quanto privados, que possuíssem vínculos empregatícios formais. Os dados coletados incluíam informações detalhadas sobre a força de trabalho, como ocupação, salário, jornada de trabalho, entre outros aspectos. Essas informações são estrategicamente relevantes, pois permitem monitorar o comportamento do mercado de trabalho, identificar tendências e acompanhar as mudanças na estrutura de emprego no Brasil.

No contexto da força de trabalho em enfermagem, a Rais foi uma fonte fundamental para entender o perfil dos profissionais que atuam em regime formal, analisando sua distribuição por região, faixa salarial e tipo de ocupação. A base de dados possibilita, ainda, identificar padrões de contratação e rotatividade, além de fornecer subsídios para avaliar a evolução do emprego na enfermagem ao longo do tempo, auxiliando na formulação de políticas que visem à valorização e retenção dos profissionais de saúde.

Origem dos dados: os dados são disponibilizados de forma aberta, sem necessidade de cadastro no endereço <https://www.gov.br/trabalho-e-emprego/pt-br/assuntos/estatisticas-trabalho/microdados-rais-e-caged>. Para baixar as tabelas, faz-se necessário logar no FTP: <ftp://ftp.mtpps.gov.br/pdet/microdados/>.

Estrutura: os dados da Rais são disponibilizados com periodicidade anual, abrangendo informações detalhadas sobre vínculos empregatícios formais em empresas públicas e privadas. A estrutura da base inclui registros individuais de vínculos de trabalho ativos em 31 de dezembro de cada ano, além de dados relacionados a admissões e desligamentos durante o período. Esses registros incluem uma ampla gama de atributos, como ocupação, faixa etária, sexo, raça/cor, nacionalidade, tipo de vínculo, tipo de movimento (admissão, desligamento, etc.), e remuneração mensal de dezembro.

A abrangência temporal dos dados inicia-se em 1976, ano da criação da Rais, até 2022, quando a base foi descontinuada e substituída pelo sistema eSocial. Apesar de sua descontinuação, os dados históricos da Rais continuam sendo uma boa fonte para análises do mercado de trabalho formal no Brasil. A estrutura dos campos manteve consistência em relação a alguns atributos comuns ao longo do tempo, como município

do trabalho, código de ocupação e faixa salarial, permitindo a construção de séries históricas para diversas análises. No entanto, alterações pontuais em certos períodos, como a inclusão ou exclusão de campos, podem impactar a comparabilidade de algumas variáveis entre anos.

Principais Dimensões: Ano; Mês admissão; Mês desligamento; Código da ocupação; Vínculo ativo 31/12; Faixa etária; Sexo; Raça/cor; Valor da remuneração de dezembro; Município do trabalho; Nacionalidade.

Restrições pela LGPD: dados descaracterizados.

Possibilidades de Investigação: a base de dados da Rais oferece uma ampla gama de possibilidades de investigação relacionadas à força de trabalho em enfermagem no Brasil. Com informações detalhadas sobre vínculos formais, é possível realizar análises que identificam a força de trabalho da enfermagem no mês de dezembro de cada ano, além de traçar o perfil dos vínculos ao longo do tempo. A análise dos dados de admissões e desligamentos contribui para uma melhor compreensão da dinâmica do mercado de trabalho formal na área da saúde, incluindo taxas de rotatividade e permanência.

Outro aspecto relevante a ser investigado é a distribuição geográfica dos enfermeiros em atividade, levando em consideração variáveis como município, região e faixa salarial. Essas informações permitem identificar possíveis desigualdades na oferta de profissionais entre diferentes regiões do país, além de avaliar se as localidades com maior demanda por serviços de saúde estão sendo adequadamente atendidas em termos de força de trabalho. Esses dados são fundamentais para auxiliar gestores e formuladores de políticas públicas no planejamento de intervenções que promovam uma distribuição mais equitativa de enfermeiros em regiões carentes.

Adicionalmente, a Rais possibilita a análise do perfil sociodemográfico dos enfermeiros em regime formal de trabalho, considerando variáveis como idade, sexo e raça/cor. Essa análise pode revelar disparidades na participação de diferentes grupos na força de trabalho, contribuindo para o desenvolvimento de políticas que promovam maior diversidade e inclusão na enfermagem. A relação entre remuneração e características demográficas também pode ser explorada, permitindo a identificação de possíveis desigualdades salariais, o que é essencial para a formulação de ações voltadas à valorização e à promoção da igualdade no setor de saúde.

Apesar base tenha sido descontinuada em 2022, os dados históricos fornecidos pela Rais continuam sendo valiosos para compreender as mudanças na composição e

no comportamento do mercado de trabalho formal em enfermagem ao longo do tempo. Essas informações possibilitam a construção de séries históricas que podem subsidiar o planejamento futuro e a tomada de decisões estratégicas voltadas ao fortalecimento da força de trabalho na saúde.

Limitações: a Rais não registra contratações de trabalhadores informais e autônomos, nem de profissionais que prestam serviços de forma contínua e regular, mas que são remunerados por meio de pessoa jurídica (PJ). Como seu escopo é restrito ao mercado formal, trabalhadores informais e autônomos são contemplados apenas em outras bases, como a PNAD Contínua (IBGE).

Outro ponto crítico é a ausência de um campo identificador para os profissionais, o que limita as análises ao nível dos postos de trabalho e impede um rastreamento mais detalhado dos vínculos ao longo do tempo. Além disso, entre 2010 e 2022, a Rais passou por alterações em sua estrutura, com mudanças nos campos disponibilizados ao longo dos anos, o que dificulta a comparabilidade de algumas variáveis entre períodos.

Especificamente entre 2010 e 2014, a base não apresenta dados de remuneração para meses específicos, limitando-se a registrar os valores referentes ao mês de dezembro. Essa ausência restringe a possibilidade de estudos detalhados sobre a remuneração mensal dos profissionais da saúde durante esse período, tornando necessário o uso de outras bases para análises mais completas.

Embora a Rais possua de 44 a 45 atributos comuns em todas as tabelas, o poder explicativo da base é impactado pela descontinuidade nos dados em 2022 e por sua substituição pelo sistema eSocial. O eSocial unifica informações trabalhistas, previdenciárias e fiscais, com intuito de reduzir redundâncias e inconsistências, mas ainda enfrenta desafios em sua implementação. A transição para o eSocial representa um avanço no potencial de análise de dados trabalhistas no Brasil, mas também introduz um período de ajuste que pode dificultar comparações diretas com os dados históricos da Rais.

Apesar dessas limitações, a Rais se destaca por responder a perguntas que não podem ser atendidas pelo Caged, como causas de afastamento, motivo de desligamento e nacionalidade do trabalhador. A padronização e robustez de seus registros a tornaram uma ferramenta essencial para análises históricas e estratégicas do mercado de trabalho formal no Brasil.

Link para o dicionário de dados: <https://abrir.link/EvjQI>

Link para a análise exploratória: <https://abrir.link/CGGRf>



Sistema de Informações sobre Mortalidade (SIM)

Descrição: o Sistema de Informações sobre Mortalidade (SIM) é um instrumento de vigilância epidemiológica desenvolvido pelo Ministério da Saúde com o propósito de coletar e organizar dados relativos aos óbitos ocorridos no território nacional. O sistema tem como base a Declaração de Óbito (DO), um documento padronizado nacionalmente para registro compulsório de informações sobre a causa e as circunstâncias do óbito. O SIM é responsável por fornecer dados consistentes e detalhados sobre mortalidade, sendo fundamentais para a formulação e avaliação de políticas públicas de saúde.

No contexto da força de trabalho em enfermagem, o SIM é uma ferramenta estratégica para compreender como condições de saúde e riscos ocupacionais afetam diretamente os profissionais dessa categoria. Seus dados permitem identificar a ocorrência de óbitos entre enfermeiros e correlacionar fatores como causas de morte, faixa etária, sexo, raça/cor e local de atuação. Essa análise fornece um panorama detalhado sobre os principais riscos enfrentados pela categoria, permitindo identificar padrões de mortalidade específicos e direcionar ações para mitigar esses riscos.

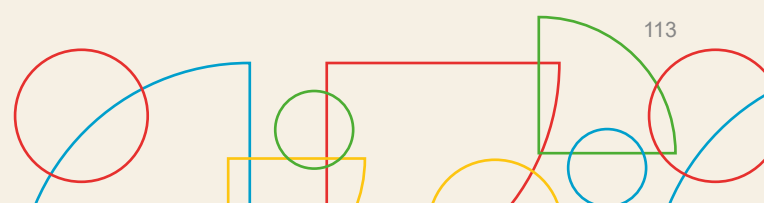
Origem dos dados: os dados são disponibilizados de forma aberta, sem necessidade de cadastro, em endereço <https://datasus.saude.gov.br/transferencia-de-arquivos/> e <https://svs.aids.gov.br/daent/centrais-de-conteudos/dados-abertos/sim/>.

Estrutura: o SIM apresenta dados por registro de ocorrência, organizados com base na data do óbito. O período coberto é entre 1979 e 2024, considerando o momento da produção deste texto. As informações incluem variáveis demográficas, geográficas e relacionadas às circunstâncias do óbito, permitindo análises detalhadas e abrangentes.

Principais Dimensões: Data do óbito; Município de nascimento; Data de Nascimento; Idade; Sexo; Raça/cor; Estado Civil; Escolaridade; Código de ocupação do trabalho; Município da ocorrência; e Tipo de óbito.

Restrições pela LGPD: dados descaracterizados.

Possibilidades de Investigação: a base de dados do SIM possibilita uma análise aprofundada sobre a mortalidade entre profissionais de enfermagem, oferecendo uma perspectiva valiosa para estudos de demografia dessa força de trabalho. A análise detalhada das causas de óbito pode identificar padrões específicos de mortalidade entre enfermeiros, considerando variáveis como idade, sexo, raça/cor e local de atuação. Esses dados são cruciais para compreender os fatores de risco que afetam a longevidade



dos profissionais e para mapear diferenças significativas na mortalidade entre grupos demográficos e regiões do País.

Adicionalmente, o SIM permite investigar o impacto das condições de trabalho e dos riscos ocupacionais na mortalidade dos enfermeiros. Por meio da análise das causas mais frequentes de morte na categoria, é possível correlacionar essas condições ao ambiente de trabalho e identificar fatores que contribuem para a vulnerabilidade dos profissionais, como jornadas extensas, exposição a doenças e falta de suporte adequado no ambiente de trabalho.

O sistema também oferece a possibilidade de cruzar variáveis sociodemográficas, como escolaridade e estado civil, com os dados de mortalidade, contribuindo para identificar perfis mais vulneráveis dentro da força de trabalho em enfermagem. Esses estudos podem ajudar a delinear políticas que promovam a proteção e a segurança ocupacional, além de orientar iniciativas para reduzir as desigualdades na saúde dos profissionais.

Outro aspecto relevante é a análise da distribuição geográfica dos óbitos, que permite identificar regiões com índices elevados de mortalidade entre enfermeiros. Essa investigação pode revelar fatores associados, como infraestrutura inadequada, escassez de recursos e condições de trabalho precárias, subsidiando intervenções estratégicas que busquem melhorar o ambiente de trabalho e a qualidade de vida dos profissionais.

Limitações: embora o SIM seja uma base robusta e abrangente, algumas limitações devem ser consideradas. A qualidade dos dados depende da completude e da precisão das informações registradas nas declarações de óbito, o que pode levar a subnotificações ou inconsistências. Além disso, a utilização da variável “código de ocupação do trabalho” para identificar enfermeiros exige um mapeamento cuidadoso, pois inconsistências na codificação podem impactar a precisão das análises.

Por fim, o SIM não inclui informações adicionais que poderiam enriquecer as análises no contexto da enfermagem, como detalhes sobre o vínculo empregatício ou o histórico ocupacional dos profissionais. Apesar dessas limitações, o sistema permanece como uma ferramenta essencial para avaliar a mortalidade e os riscos enfrentados pela força de trabalho em enfermagem, fornecendo subsídios valiosos para a formulação de políticas públicas e estratégias de saúde ocupacional.



Sistema de Seleção Unificada (Sisu)

Descrição: o Sistema de Seleção Unificada (Sisu) é uma plataforma digital gerida pelo Ministério da Educação (MEC) que centraliza a oferta de vagas em cursos de graduação de instituições públicas de ensino superior no Brasil. A seleção é baseada no desempenho dos candidatos no Exame Nacional do Ensino Médio (Enem), permitindo que os estudantes escolham suas opções de curso e instituição conforme suas notas. O Sisu funciona como uma ferramenta de democratização do acesso ao ensino superior, facilitando a alocação de vagas de forma transparente e acessível.

O sistema opera em edições regulares, regidas por editais publicados pelo MEC e pelas próprias instituições participantes. Esses editais detalham os critérios de seleção, prazos e procedimentos específicos, garantindo que as regras do processo sejam claras para todos os candidatos e instituições envolvidas.

Origem dos dados: os dados são disponibilizados de forma aberta, sem necessidade de cadastro no endereço: <https://dadosabertos.mec.gov.br/sisu>.

Estrutura: os dados do Sisu abrangem o período entre 2017 e 2022, considerando o momento da produção deste texto. Durante esse tempo, as tabelas passaram por alterações em seus campos, com adições e exclusões de colunas ao longo dos anos, o que pode impactar a comparabilidade de variáveis entre edições.

Principais Dimensões: a base do Sisu será incorporada ao *data lake* em momento futuro. Essa decisão foi tomada com base na necessidade de organizar as principais bases de dados de forma estruturada e relacionadas às dimensões estratégicas do projeto.

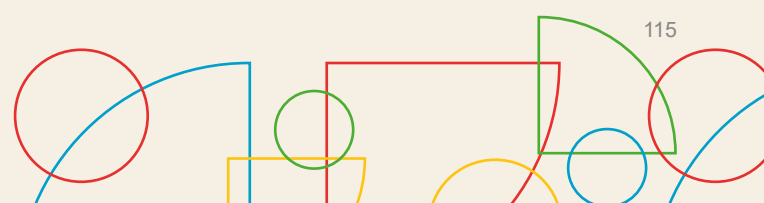
Restrições pela LGPD: dados descaracterizados.

Possibilidades de Investigação: não analisado, considerando o relato do item principais dimensões.

Limitações: não analisado, considerando o relato do item principais dimensões.

Link para o dicionário de dados: <https://abrir.link/IGKUA>

Link para a análise exploratória: <https://abrir.link/ltzpH>



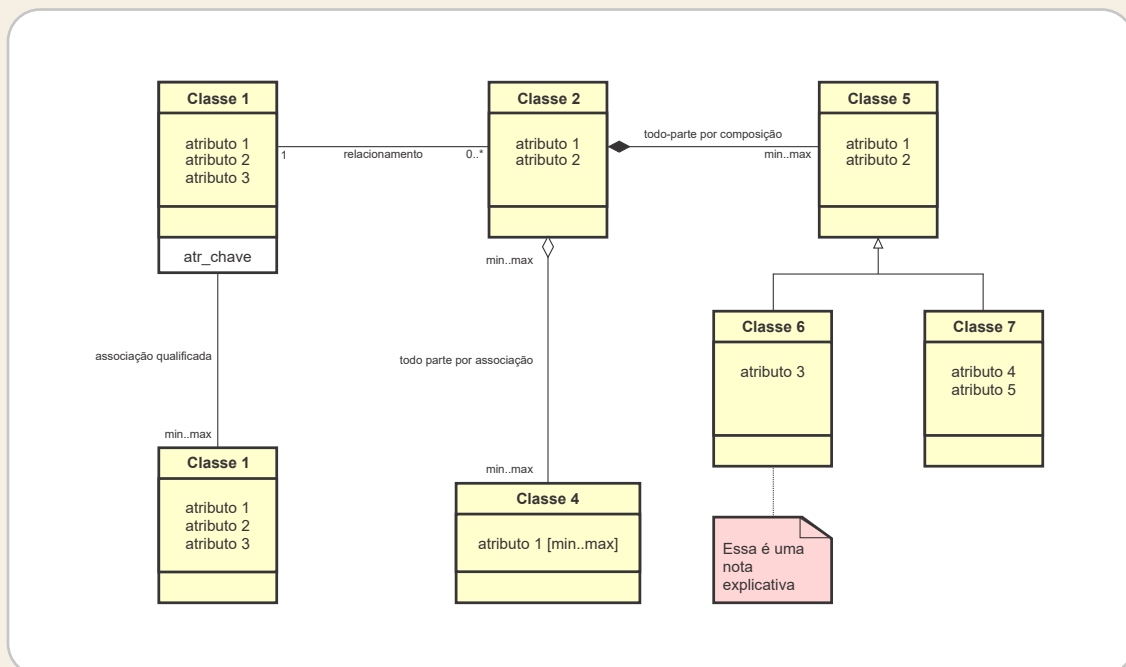
Apêndice B – Modelo estrutural estático do gerador de demografias

Com o objetivo de representar as estruturas e relacionamentos dos elementos de dados do gerador de demografias, adotamos como ferramenta de representação o modelo conceitual de classes da UML (*Unified Modeling Language*), em sua versão 2.5.1 (OBJECT MANAGEMENT GROUP, 2017; Bezerra, 2015).

Artefatos da UML utilizados na representação – Diagrama de classe conceitual

O diagrama de classe conceitual genérico, apresentado a seguir, traz consigo elementos gráficos disponíveis na UML 2.5.1 para o projeto conceitual de dados, com o objetivo de auxiliar o leitor que desconhece a notação na leitura do texto que se segue (Figura 1). É importante registrar que eles não esgotam o tema. A literatura sobre o assunto é vasta, com referências muito consistentes e consolidadas no tema (Bezerra, 2014; OBJECT MANAGEMENT GROUP, 2017).

Figura 1 – Diagrama de classe conceitual genérico



Fonte: elaboração própria.

Classes

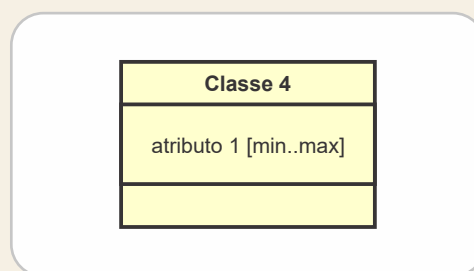
No diagrama apresentado, podemos observar a ocorrência de classes, nomeadas como “Classe 1”, “Classe 2”, até “Classe 5”. Cada classe representa a abstração de características (atributos) e/ou comportamentos (métodos) de objetos (ou instâncias) similares, bem como os relacionamentos (ou associações) entre, de forma geral, os objetos das classes. No caso em específico, como estamos focalizando a estrutura estática, não representaremos os comportamentos das classes envolvidas.

Atributos multivalorados

As características de objetos de uma classe são modeladas por meio de *atributos*, cada qual associado a um tipo de dados, simples ou complexo. Eles representam valores únicos para cada instância da classe – o que significa que cada atributo só pode armazenar nenhum (situação em que se atribui o valor NULO) ou no máximo um valor do tipo de dado relacionado. Esse tipo de atributo padrão é chamado *atributo atômico*.

Entretanto, a UML 2.5.1 permite a modelagem, no diagrama conceitual de classes, de atributos que podem armazenar, para cada objeto de uma classe, uma quantidade mínima e máxima de valores do tipo de dado associado àquele atributo. Essa situação caracteriza os chamados **atributos multivalorados**, cuja representação está representada na Figura 2.

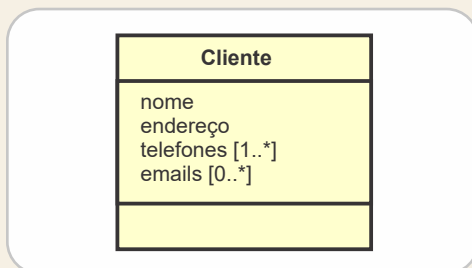
Figura 2 – “Classe 4” apresenta o atributo multivalorado “atributo 1”, com um número mínimo min e um número máximo max de ocorrências



Fonte: elaboração própria.

Em situações como essa, o “atributo 1” de um objeto da “Classe 4” pode armazenar vários valores. Um exemplo concreto dessa situação é o representado na Figura 3.

Figura 3 – Representação da classe “Cliente”



Fonte: elaboração própria.

A classe “Cliente” possui quatro atributos: dois simples (“nome” e “endereço”) e dois multivalorados (“telefones” e “emails”). O atributo “telefones” deve guardar, para cada objeto da classe “Cliente”, no mínimo um telefone (especificado na cardinalidade mínima da multiplicidade igual a um, representada dentro dos colchetes ao lado do nome desse atributo); no máximo, o objeto da classe “Cliente” pode ter um número indeterminado de telefones (cardinalidade máxima da multiplicidade representada pelo asterisco).

O atributo multivalorado “emails” tem semântica similar ao atributo “telefones”, à exceção da cardinalidade mínima de sua multiplicidade. No seu caso, um objeto da classe “Cliente” pode não ter nenhum e-mail associado.

Associações

Uma associação é representada, a princípio, por uma linha contínua ligando duas classes. Sua semântica pode ser mais bem compreendida se reduzida ao escopo da concretização (materialização) de um objeto de cada classe envolvida na associação. Assim, se a abstração “Classe 1” materializar-se no objeto “Objeto 1”, e “Objeto 2” representar a abstração “Classe 2”, então a associação “relacionamento” do esquema anteriormente representado pode associar “Objeto 1” a nenhum (0) ou vários (*) objetos similares a “Objeto 2” – e, por essa razão, temos a representação **0..*** ao lado da “Classe 2”, referindo-se aos objetos da “Classe 1”. Por outro lado, “Objeto 2” relaciona-se necessariamente no mínimo a um (1) e no máximo a um (1) objeto similar a “Objeto 1” da “Classe 1” (indicação **1** do lado da “Classe 1”, que também poderia ser apresentada como **1..1**). Essa combinação de mínimos e máximos de associações do objeto de uma classe a objetos de outra classe é denominada **multiplicidade**, com suas **cardinalidades mínima e máxima**. Quando a cardinalidade mínima é zero (0),

ela pode ser omitida, e indica que a associação dos objetos da classe a outros objetos da outra classe é *opcional*. Caso contrário (cardinalidade mínima maior ou igual a um), a associação é *obrigatória*.

Existem três casos especiais de relacionamentos que foram utilizados no esquema supracitado. Trata-se da **associação qualificada** e das **associações todo-parte por agregação e por composição**.

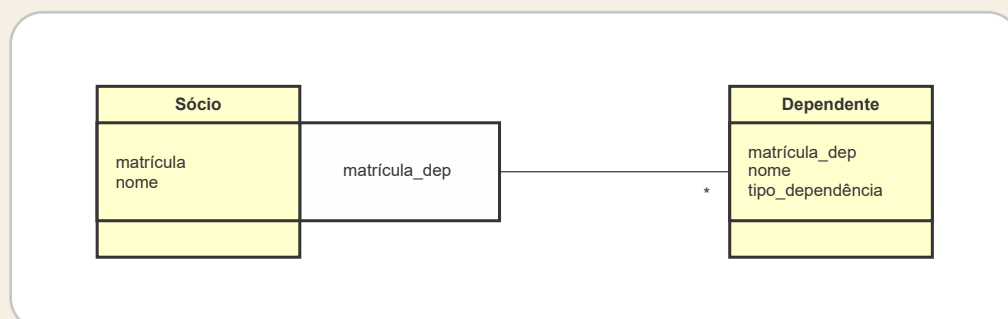
Associações qualificadas

Existem situações nas quais a instanciação de um objeto depende fundamentalmente da existência prévia de um outro objeto. A fim de capturar tal semântica, a UML representa tal ideia como associação qualificada, com o desenho de um retângulo adjacente à classe que abstrai os objetos que necessitam ser previamente materializados. Para fins práticos, e remetendo-nos à antiga representação, chamaremos essa classe de “*classe forte*”, e a outra, “*classe fraca*”. Ao lado da classe fraca, aparece a multiplicidade dos objetos da classe forte frente aos da classe fraca. No sentido oposto, a multiplicidade sempre é 1..1 (ou simplesmente 1), pois, se um dado objeto precisa de outro para existir, não há sentido em associá-lo a nenhum ou vários objetos.

Por fim, o atributo utilizado no retângulo adjacente à classe forte é dito “*atributo-chave*”. Ele é um identificador obtido no rol de atributos da classe fraca, e serve como um índice entre os objetos dessa classe.

Como exemplo, a Figura 4 representa um modelo conceitual de classes de um clube associativo, envolvendo sócios e seus dependentes.

Figura 4 – Exemplo do caso de uma associação qualificada



Fonte: elaboração própria.

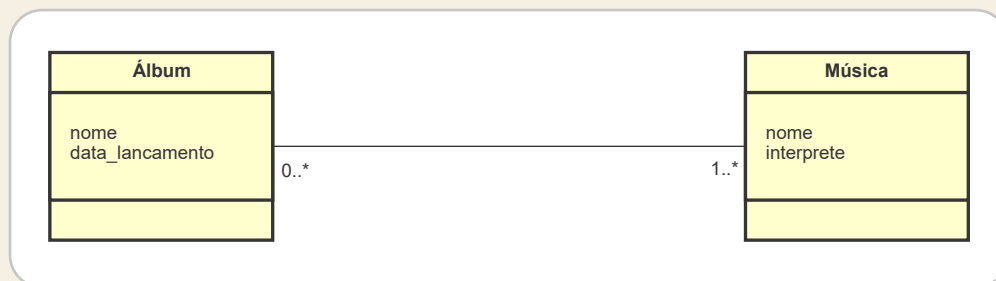
A leitura do modelo apresentado diz que: um objeto da classe “Sócio” pode estar associado a nenhum (mínimo zero, permissivamente omitido), e no máximo a vários objetos da classe “Dependente” (representado pelo símbolo *). Por outro lado, um objeto da classe “Dependente” precisa estar associado a exatamente um (mínimo um e máximo um, com representação de multiplicidade omitida) objeto da classe “Sócio”.

Observe que, para que possa existir um dependente, há a necessidade imperiosa de que exista no mínimo um sócio vinculado. No que tange ao máximo, também temos a limitação de um sócio, pois a dependência é de exatamente um (o vinculante). A decorrência dessa situação implica em que um dado objeto da classe “Dependente” só existe se houver um sócio associado a ele.

Relacionamentos todo-parte por associação

Por vezes, em uma associação entre objetos de classes, por vezes identificamos uma interessante particularidade: os objetos de uma classe, se reunidos forem, concretizam um objeto de outra classe. Porém, esses objetos, que são partes de um todo, possuem existência independente. Vejamos um exemplo na Figura 5.

Figura 5 – Relacionamento inicial entre objetos das classes “Álbum” e “Música”



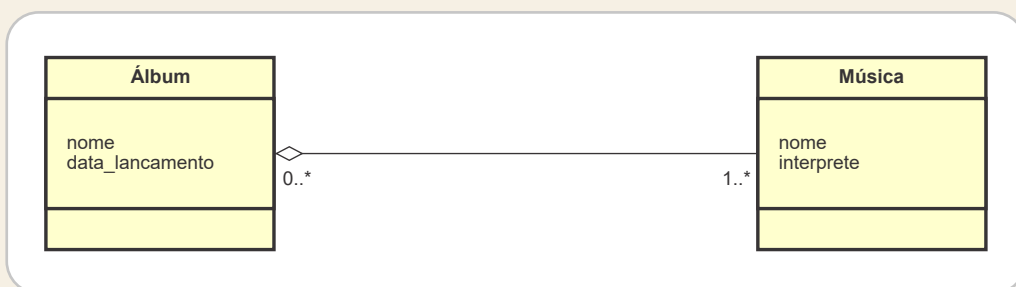
Fonte: elaboração própria.

Um objeto da classe “Álbum” é composto por, no mínimo um, e no máximo vários objetos da classe “Música”. Entretanto, um objeto da classe “Música”, que pode ser instanciado de forma independente com relação à classe “Álbum”, pode não estar associado a nenhum ou vários outros objetos daquela classe.

Observe que essa situação carece de semântica adicional para complementação plena do modelo. Isso acontece pois lhe falta a caracterização da relação de todo e parte anteriormente mencionada. E isso acontece por se tratar de uma situação caracterizada por um **relacionamento todo-parte por associação**: todo-parte por um conjunto de

objetos da classe “Música” (*classe parte*) materializarem um objeto da classe “Álbum” (*classe todo*); e por associação porque os objetos da classe “Música” são instanciados sem a necessidade da existência prévia da classe “Álbum”. A representação do relacionamento todo-parte por associação acontece pelo desenho de um losango ao lado da classe todo. Completando o modelo da Figura 10 temos a representação completa na Figura 6.

Figura 6 – Exemplo de relacionamento do tipo todo-parte por associação



Fonte: elaboração própria.

Relacionamentos todo-parte por composição

Considere a combinação das duas situações anteriormente apresentadas (associação qualificada e associação todo-parte por associação): um relacionamento do tipo todo-parte em que, para que um objeto da classe parte seja instanciado, dependem da existência prévia de um objeto da classe “todo”. Essa situação caracteriza o que se denomina **associação todo-parte por composição**. A Figura 7 nos traz um exemplo representativo.

Figura 7 – Exemplo de relacionamento todo-parte por composição



Fonte: elaborada pelos autores.

Um conjunto de objetos da classe parte “Item Nota Fiscal” realizam um objeto da classe todo “Nota Fiscal”. Adicionalmente, para que um objeto da classe “Item Nota Fiscal” possa ser criado, é necessário que ocorra a existência prévia de um objeto da classe “Nota Fiscal”.

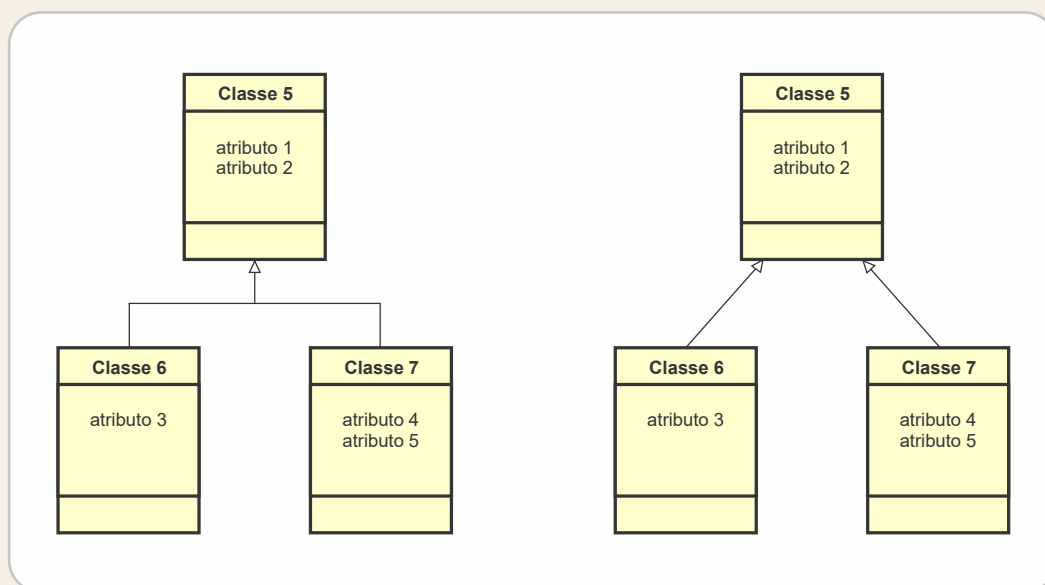
Herança

Todas as situações anteriormente apresentadas tratavam de associações de objetos de uma classe com instâncias, materializações de outra classe. Entretanto, no caso em tela, a **herança**, temos o único caso no diagrama de classes conceitual de um relacionamento entre classes.

Na herança (ou generalização/especialização), existe o conceito de *superclasse* (ou classe genérica, ou classe mãe), e *subclasses* (ou classes filhas, ou classes especializadas). O conceito central desse tipo de relacionamento é o de transmitir toda a semântica associada à superclasse para as subclasses – sejam os atributos, e/ou as associações, e/ou os comportamentos (métodos) a ela atrelados.

As possíveis formas de representação da herança no modelo de classes conceitual da UML 2.5.1 são ilustradas na Figura 8.

Figura 8 – Formas de representação da herança na UML

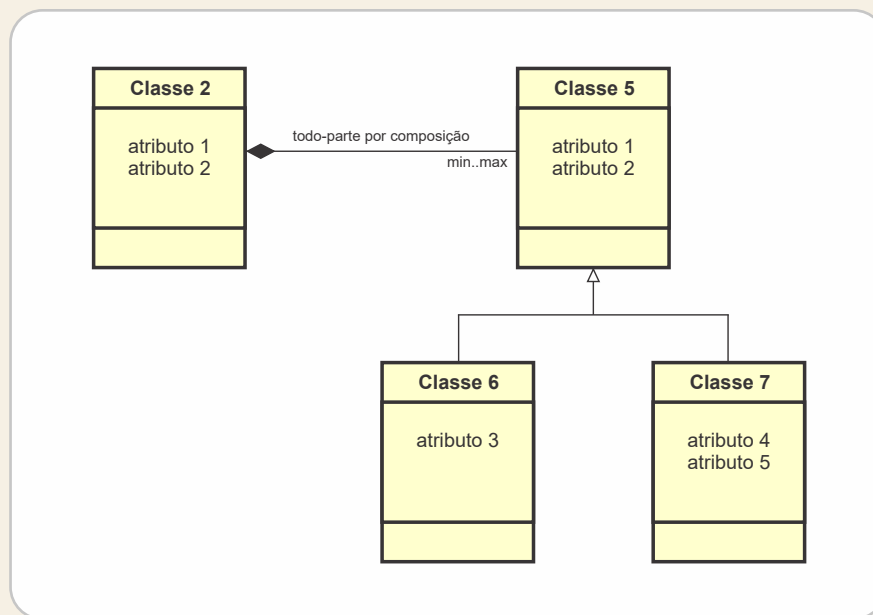


Fonte: elaboração própria.

No esquema apresentado nas Figuras 8, a “Classe 6” possui, de fato, três atributos: “atributo 1”, “atributo 2” e “atributo 3”; a “Classe 7” tem como atributos “atributo 1”, “atributo 2”, “atributo 4” e “atributo 5”, em um total de quatro atributos.

Na Figura 9, o relacionamento entre as classes “Classe 2” e “Classe 5” são transmitidas, por herança, para as classes “Classe 6” e “Classe 7”.

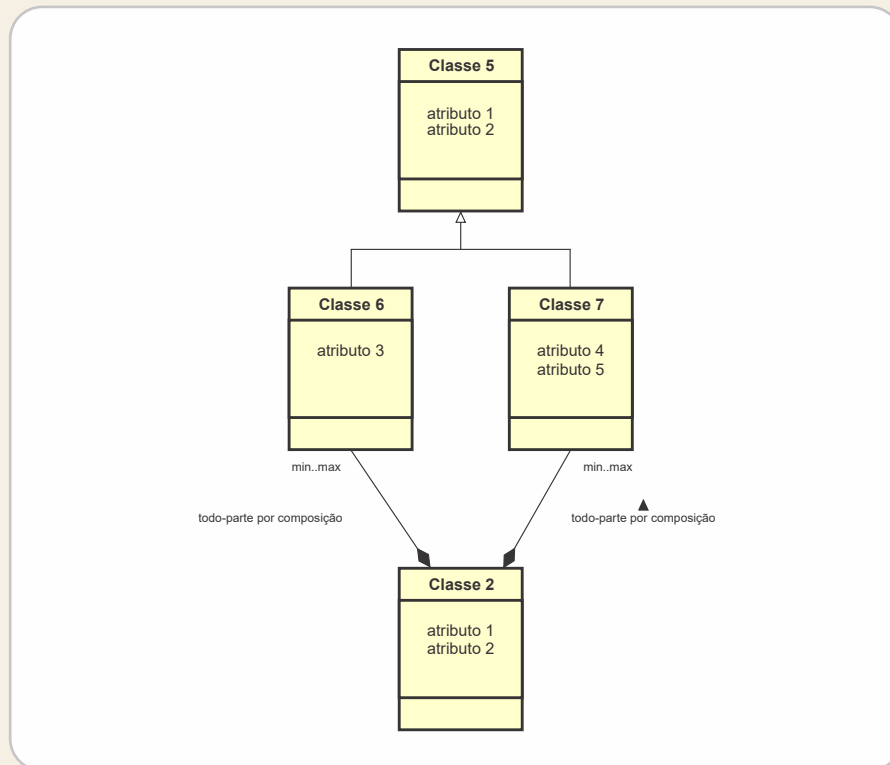
Figura 9 – Trecho do modelo conceitual de classe originalmente apresentado nesta seção



Fonte: elaboração própria.

Destarte, semanticamente, o modelo da Figura 10 equivale ao da Figura 9.

Figura 10 – Representação equivalente (e inadequada) ao modelo da Figura 12

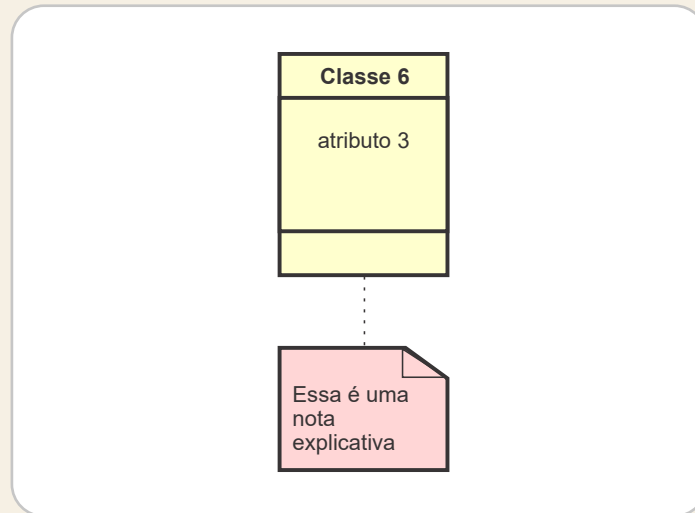


Fonte: elaboração própria.

Notas explicativas

As notas explicativas são elementos auxiliares da UML, não restritas ao diagrama de classes. Trata-se de metaelementos, utilizados para elucidar, documentar, justificar e/ou complementar o entendimento do modelo. Podem, por conseguinte, serem utilizadas em qualquer artefato disponível na UML 2.5.1, como apresentado na Figura 11. No diagrama de classes, as notas explicativas justificam inclusões e extensões, anotam requisitos funcionais ou não funcionais relacionados, explicitam valores de domínios de atributos, por exemplo. No diagrama de atividades, que será apresentado na sequência, explica de regras de negócios, condições em decisões, e exceções não modeladas graficamente.

Figura 11 – Exemplo de uma nota explicativa ligada à classe “Classe 6”



Fonte: elaboração própria.

Detalhamento do modelo

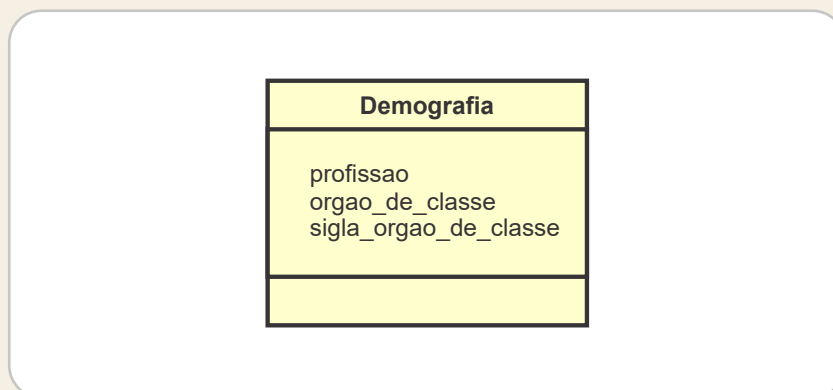
Preâmbulo

Com o conhecimento necessário para o pleno entendimento da semântica dos elementos gráficos utilizados no contexto desse projeto, no que tange à modelagem estrutural estática de seus elementos utilizando o diagrama de classes da UML 2.5.1 (OBJECT MANAGEMENT GROUP, 2017; BEZERRA, 2015), partiremos para a explanação dos elementos conceituais de dados de um gerador de demografias.

Construção conceitual do modelo

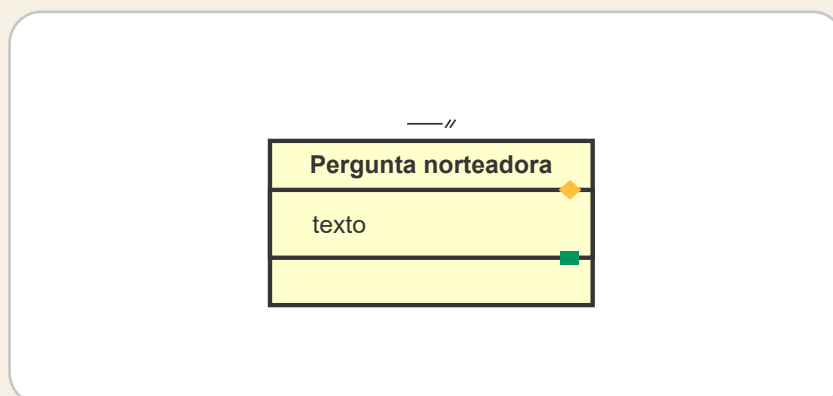
Caracterizam uma demografia dados tais como qual a profissão-alvo da análise, o seu respectivo órgão de classe nacional, bem como a sigla desse órgão, como apresentado na Figura 12.

Figura 12 – Descrição da classe “Demografia”



Fonte: elaboração própria.

Figura 13 – Descrição da classe “Pergunta norteadora”

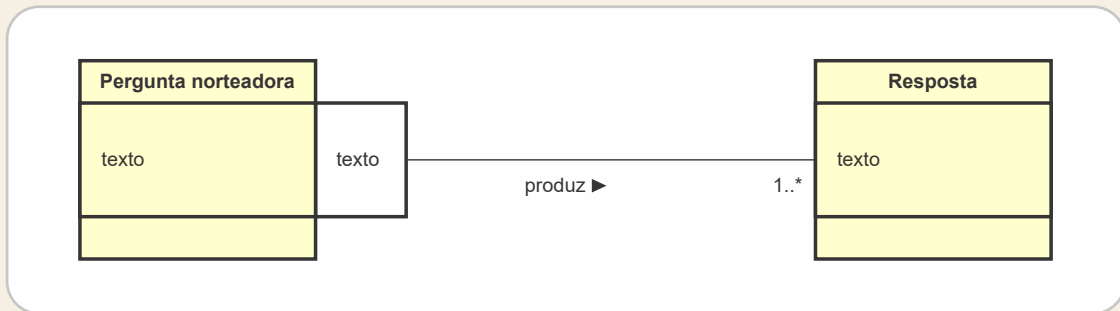


Fonte: elaboração própria.

Demografias iniciam-se pela reflexão, por parte de profissionais de uma área, de quais elementos desejam-se analisar. A partir daí, desdobram-se perguntas norteadoras (Figura 13), registradas apenas com o seu texto descritivo.

Essas perguntas norteadoras produzirão respostas, cujo registro somente existirá por conta da existência da pergunta norteadora. Uma pergunta norteadora deriva uma ou mais repostas. A Figura 14 representa a situação descrita.

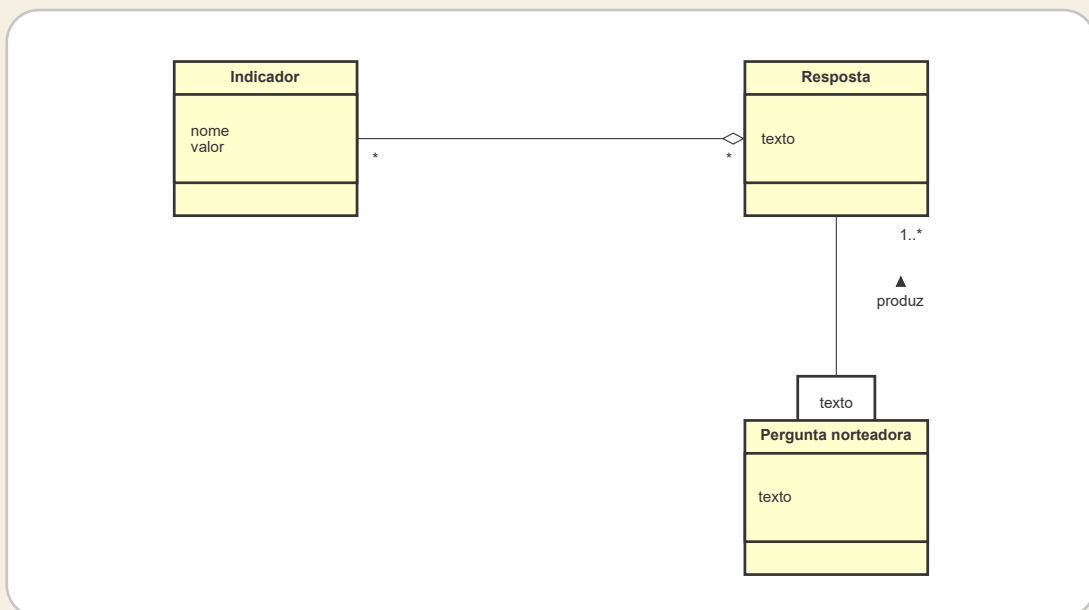
Figura 14 – Associação qualificada “produz” entre as classes “Pergunta norteadora” e “Resposta”



Fonte: elaboração própria.

A Figura 15 mostra a seguinte situação: a resposta, por sua vez, é composta por nenhum ou mais indicadores. Todavia, um indicador pode colaborar na construção de nenhuma ou mais respostas. Com isso, reforça-se a validade da produção de indicadores.

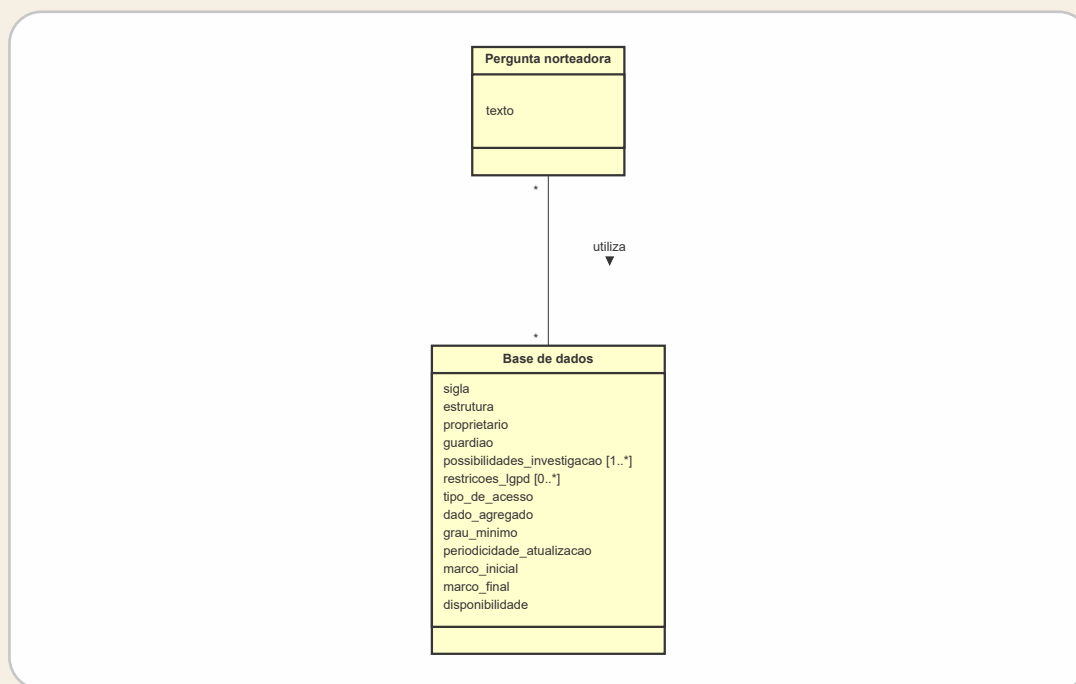
Figura 15 – Representação da associação todo-parte por agregação entre as classes “Indicador” e “Resposta”, e a associação qualificada entre as classes “Pergunta norteadora” e “Resposta”



Fonte: elaboração própria.

Cada pergunta norteadora pode ou não se valer do uso de bases de dados, que embasarão, em última análise, a produção dos indicadores utilizados pelas respostas. Cada base de dados é caracterizada por um nome, uma descrição, a fonte de origem dos dados, uma sigla, sua estrutura, proprietário dos dados, guardião (ou fiel depositário), as formas de investigação, as eventuais restrições relacionadas à Lei Geral de Proteção de Dados (LGPD – Lei n. 13.709, de 14 de agosto de 2018), o tipo de acesso (se público ou privado), se há dados agregados ou não (com opções de valores “sim”, “não” ou “não analisado”), a granularidade mínimo dos dados (se município, capital ou UF), a periodicidade de atualização (diária, semanal, mensal, trimestral, quadrimestral, semestral, anual ou não analisado), os marcos inicial e final (primeira e última datas de ocorrência de dados na base), e a forma de disponibilidade dos dados. A estrutura descrita é graficamente representada na Figura 16.

Figura 16 – Associação “utiliza” entre as classes “Pergunta norteadora” e “Base de dados”

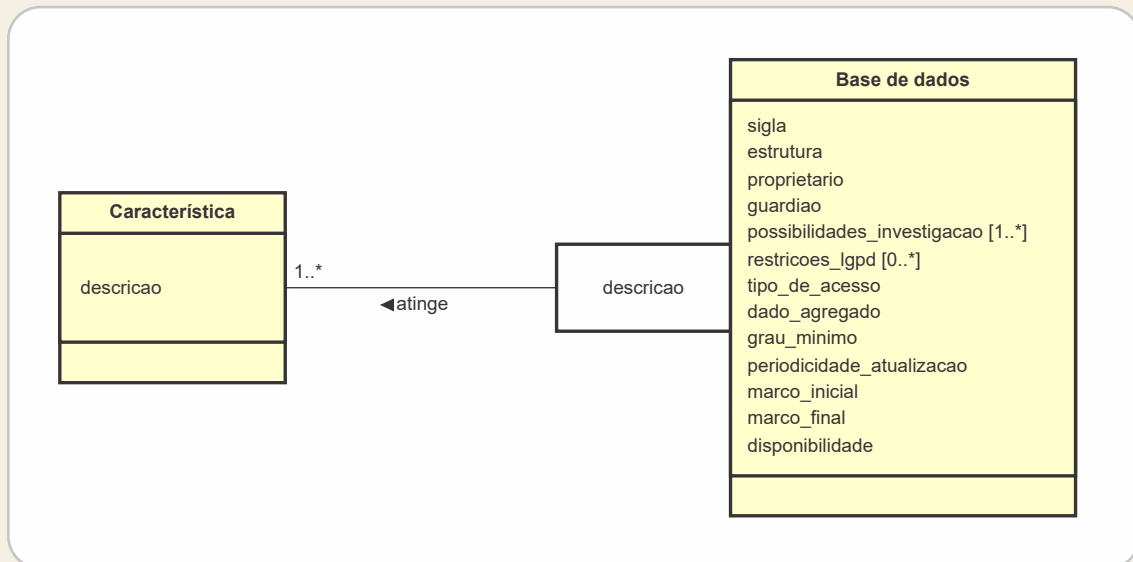


Fonte: elaboração própria.

As bases de dados possuem ao menos uma característica. Essas características descrevem, de forma objetiva, quais os propósitos de cada base de dados. Como exemplo, a base de dados do Cadastro Geral de Empregados e Desempregados (Caged), de responsabilidade do Ministério do Trabalho e Emprego (MTE), tem como característica evidenciar a entrada e saída de postos de trabalho dentro um mês, ao

longo dos anos. A Figura 17 apresenta a modelagem da classe “Característica”.

Figura 17 – Associação qualificada entre a classe fraca “Característica” e a classe forte “Base de dados”

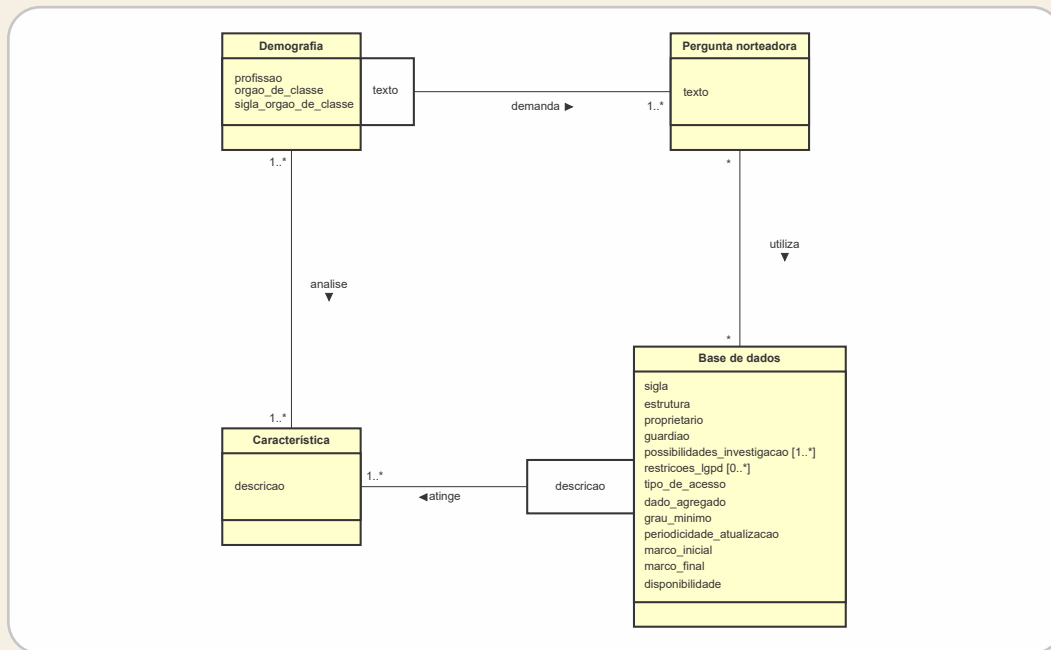


Fonte: elaboração própria.

Observe que a representação da classe “Característica” não pôde prescindir da classe “Base de dados”, pois um objeto da classe “Característica” necessita de um objeto da classe “Base de dados” para ser instanciado.

As características das bases de dados servem como elementos motivadores de demografias, pois elas inspiram a produção das perguntas norteadoras, conforme modelado na Figura 18. Um objeto da classe “Demografia” deve estar associado a, no mínimo, a um objeto da classe “Característica”, e, a, no máximo, vários objetos dessa última classe.

Figura 18 – Relacionamento “analisa” entre as classes “Característica” e “Demografia”

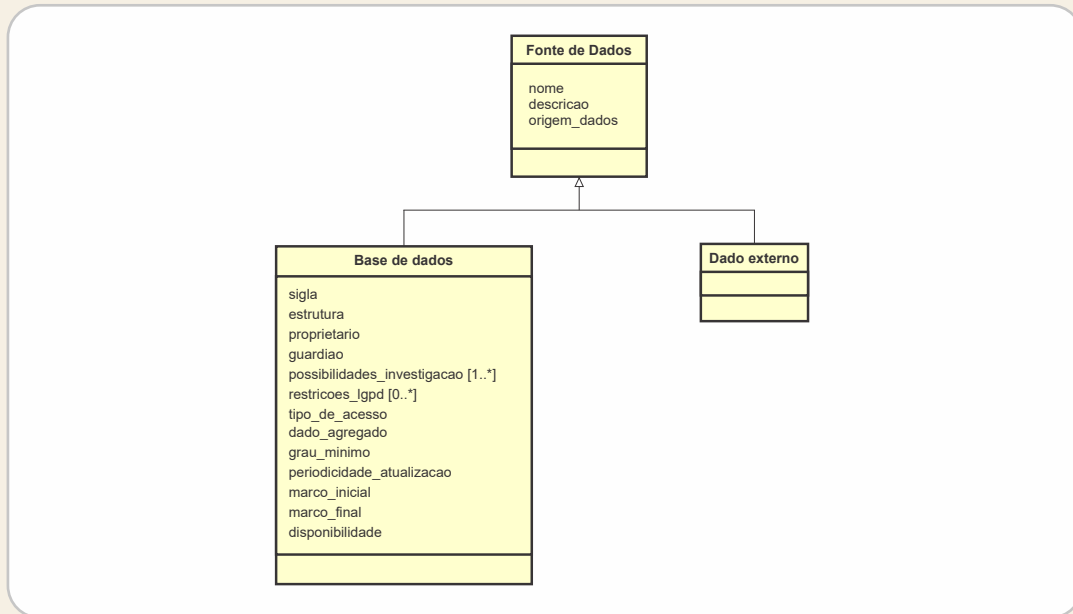


Fonte: elaboração própria.

É possível também consultar dados externos, que não necessariamente componham bases de dados estruturadas. Nesse caso, deseja-se conhecer apenas o nome da fonte, sua descrição, e a fonte de origem dos dados.

Considerando-se que tanto bases de dados como dados externos são, de forma genérica, fontes de dados, pode-se modelar estruturalmente tal relação por meio de uma herança, como a Figura 19 mostra.

Figura 19 – Herança envolvendo a superclasse “Fonte de Dados” e suas subclasses “Base de dados” e “Dado externo”



Fonte: elaboração própria.

Uma fonte de dados é composta por uma ou mais variáveis (ou atributos). Cada variável é identificada por um nome, descrição, o tipo do dado (inteiro, *string*, número real etc.), e em quais anos ela consta registrada na base de dados. Uma variável não existe se não for vinculada a uma fonte de dados. A coleção de todas as variáveis de uma fonte de dados a caracteriza plenamente. A Figura 20 ilustra tais conceitos.

Figura 20 – Relacionamento “apresenta” entre as classes “Fonte de Dados” e “Variável”



Fonte: elaboração própria.

A Figura 21 trata do relacionamento entre variáveis e regras de formação. Uma variável integra uma ou mais regras de formação, cada qual apresentando a sua estrutura. Uma coleção de variáveis de uma fonte de dados caracteriza plenamente uma ou mais regras de formação. Entretanto, a variável pode existir sem que a regra exista.

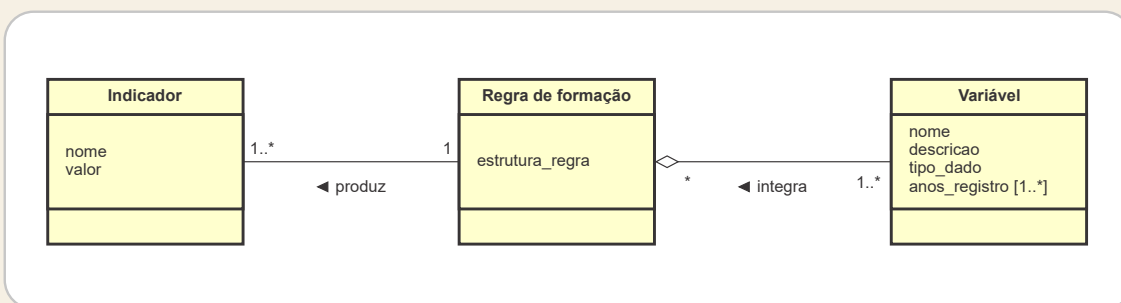
Figura 21 – Associação todo-parte por agregação “integra” entre as classes “Regra de formação” e “Variável”



Fonte: elaboração própria.

Considere a estrutura de uma regra como algo similar a uma fórmula, uma função, ou uma relação que, a partir de uma ou mais variáveis, produzirá um ou mais indicadores. Todavia, um indicador relaciona-se exatamente com uma única regra de formação, como a Figura 22 modela.

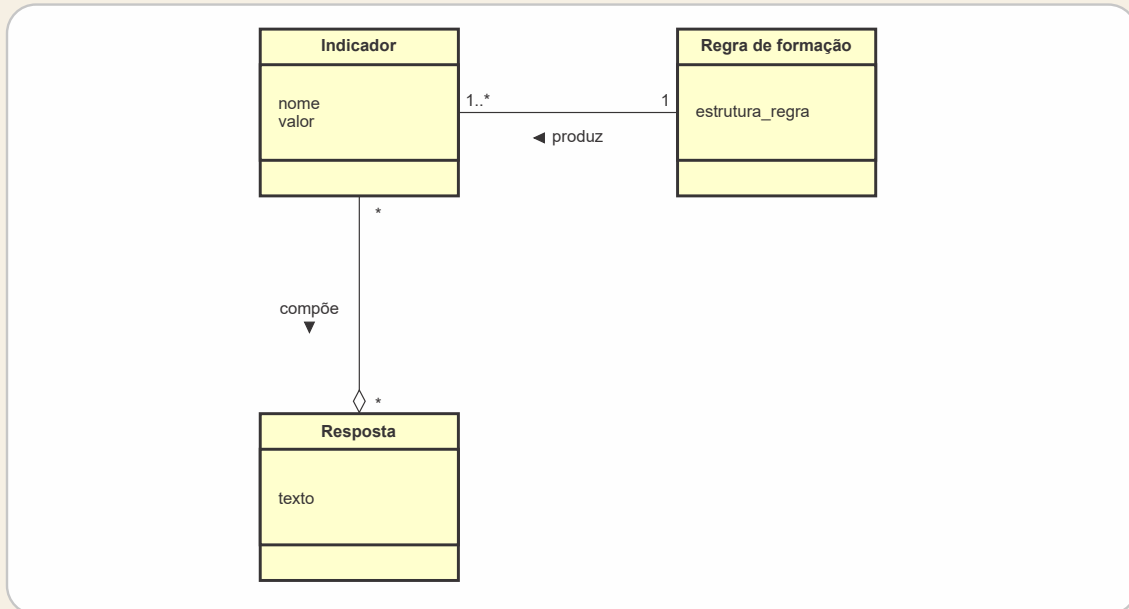
Figura 22 – Extensão do relacionamento entre variáveis e regras de formação, com a produção de indicadores



Fonte: elaboração própria.

Completando a integração do modelo, cada indicador compõe nenhuma ou mais respostas, e cada resposta pode ser composta por nenhum ou mais indicadores. Os indicadores independem das respostas. A Figura 23 ilustra o conceito.

Figura 23 – Associação entre os indicadores e a produção das respostas, modelada pelo relacionamento “compõe”

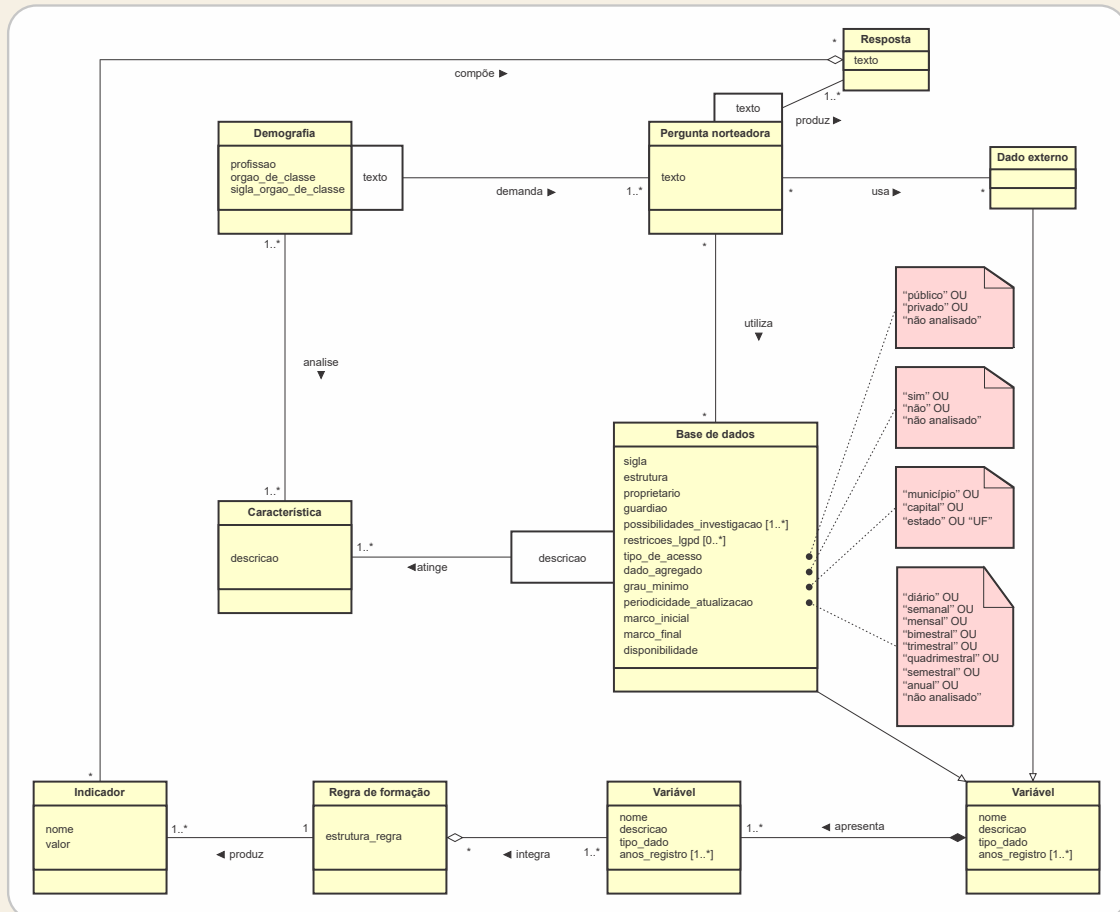


Fonte: elaboração própria.

Construção conceitual do modelo

Unindo todos os elementos conceituais apresentados, a Figura 24 apresenta a versão completa do modelo conceitual de dados, capturando todos os aspectos estruturais estáticos relacionados ao tema.

Figura 24 – Versão completa do diagrama conceitual de classes da UML do gerador de demografias



Fonte: elaboração própria.



Apêndice C – O modelo comportamental do gerador de demografias

Projetar as estruturas dos dados do sistema é tarefa imperiosa, porém não exclusiva. É de todo necessário que entendamos e modelemos ainda o comportamento do sistema, por possuir também característica dinâmica. Assim sendo, dentre as diversas ferramentas disponibilizadas pela UML 2.5.1, a escolha para representar o modelo comportamental do gerador de demografias foi o diagrama de atividades.

Tal escolha tomou como princípio sua simplicidade e facilidade de entendimento. Por guardar alguma semelhança com o esquema de fluxogramas, amplamente conhecido, ele é uma ferramenta usada para modelar fluxos de trabalho, processos de negócio ou lógicas de controle dentro de um sistema, enfatizando o fluxo de atividades e decisões.

Artefatos da UML utilizados na representação – Diagrama de atividades

Com o objetivo de apresentar os elementos do diagrama de atividades utilizados no contexto do gerador de demografias, apresentamos um diagrama de atividades genérico na Figura 30 e, na sequência (Quadro 5), o significado de cada elemento apresentado.

Cabe ressaltar que existem outros símbolos disponíveis na UML 2.5.1, com semânticas distintas, e que não foram apresentados neste texto. Referências tais como Bezerra (2015) e OBJECT MANAGEMENT GROUP (2017) contêm informações mais completas acerca do assunto.

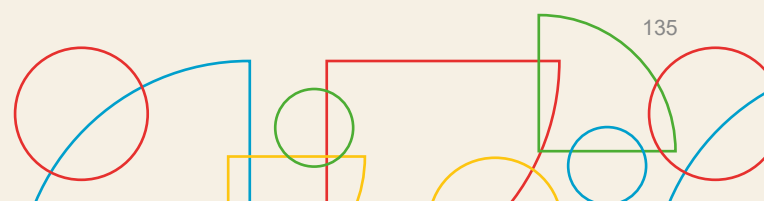
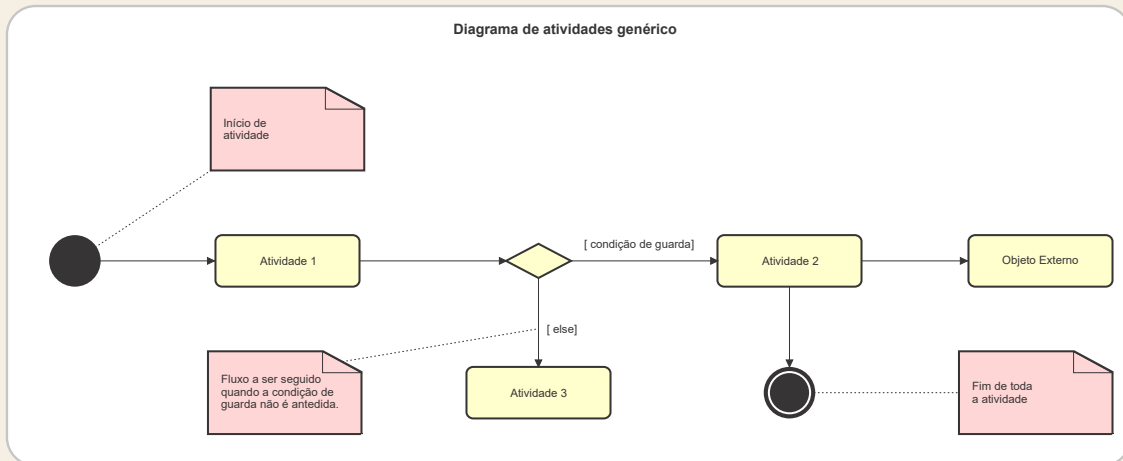



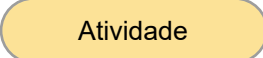


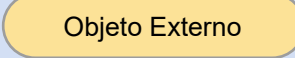

Figura 1 – Estrutura genérica de um diagrama de atividades com elementos gráficos utilizados no gerador de demografias



Fonte: elaboração própria.

O Quadro 1 indica o significado de cada um dos símbolos utilizados.

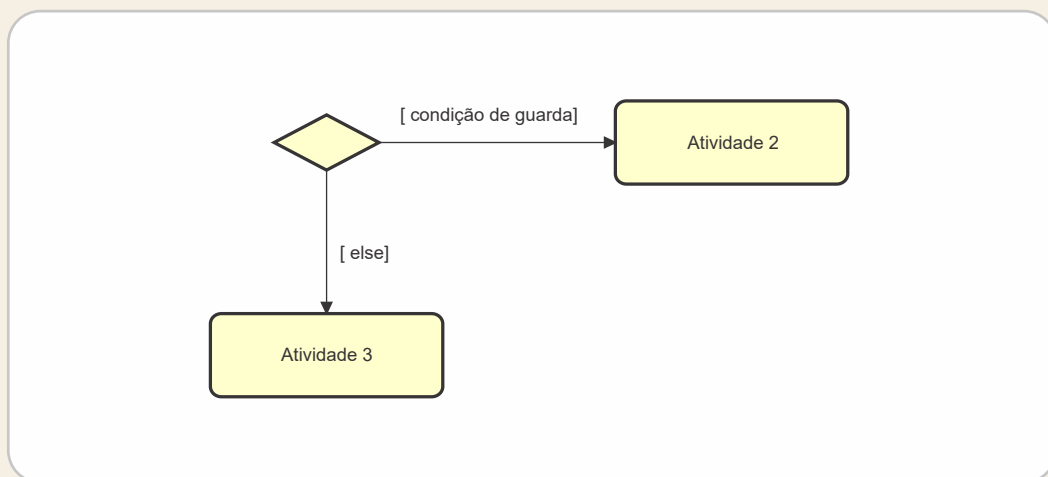
Quadro 1 – Símbolos do diagrama de atividades da UML 2.5.1 utilizados na modelagem comportamental do gerador de demografias

Símbolo	Significado
Nó inicial 	Indicador de início de fluxo
 Atividade	Ação ou tarefa a ser realizada
Fluxo de Controle 	Indicador da sequência de execução, do caminho a seguir
Nó de Decisão/Fusão 	Ponto de decisão de caminho, com a execução de uma condição de guarda
 Objeto Externo	Artefatos ou dados manipulados pelas ações
Nó Final de Atividade 	Indicador de término de fluxo, de término de todas as execuções em andamento dentro da atividade

Fonte: elaboração própria.

Especial atenção é necessária sobre o nó de decisão, com estrutura mais completa ilustrada na Figura 2. Seu funcionamento assemelha-se, em linguagens de programação, à instrução do tipo se <condição> então <ação 1> senão <ação 2>. No caso do diagrama de atividades, <condição> é chamada condição de guarda. Caso a condição de guarda seja verdadeira, seu fluxo de atividade adjacente é executado. Se for falsa, realiza-se a atividade do fluxo [else]. Ou seja, nunca as duas atividades (“Atividade 1” e “Atividade 2” serão executadas simultaneamente).

Figura 2 – Estrutura genérica de um nó de decisão



Fonte: elaboração própria.

Detalhamento do modelo

Preâmbulo

Produzir uma demografia envolve um conjunto de tarefas bem definidas. Algumas seguem um caminho obrigatório. Outras, dependem do atendimento de condições. Por esse motivo, o diagrama de atividades foi escolhido como o artefato da UML 2.5.1 a ser utilizado para modelar os aspectos comportamentais do gerador de demografias.

Detalhamento do fluxo de atividades do gerador de demografias

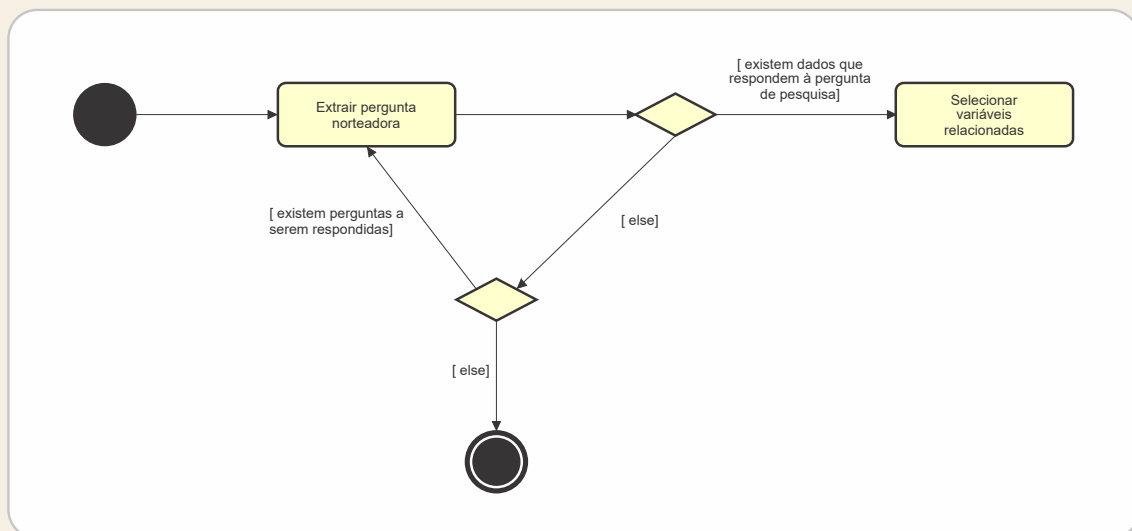
A Figura 3 apresenta o trecho inicial do gerador de demografias, que possui uma característica interessante. Observa-se que, após o nó inicial, a atividade “*Extrair pergunta norteadora*” é executada. Na sequência, a condição de guarda “existem dados que respondem à pergunta?” questiona se há a necessidade de utilização de algum dado para resolver aos desafios apresentados na pergunta em tela. Caso não seja utilizado

nenhum dado, o fluxo é desviado para a condição “else”, e, na sequência, outra condição de guarda entra em cena: “*existem perguntas a serem respondidas?*”. Caso existam, o fluxo retorna à primeira atividade (“*Extrair pergunta norteadora*”). Se a condição não for atendida, o fluxo segue a outra condição “else”, e toda a atividade termina.

Uma situação em que uma pergunta norteadora não utilize dados pode ser exemplificada quando há a necessidade de análise de documentos oficiais, normas ou legislações. Tal cenário ilustra a execução da condição da primeira condição “else”. Se tal quadro se repetir para todas as perguntas norteadoras, o gerador de demografias, escopo deste trabalho, em nada colabora com o processo, findando-se o fluxo de toda a atividade com a execução da segunda cláusula “else”.

Destacamos esse trecho para ressaltar o escopo do gerador de demografias. Seu propósito focaliza a produção de demografias que utilizam fontes de dados, sejam de bases de dados ou dados externos.

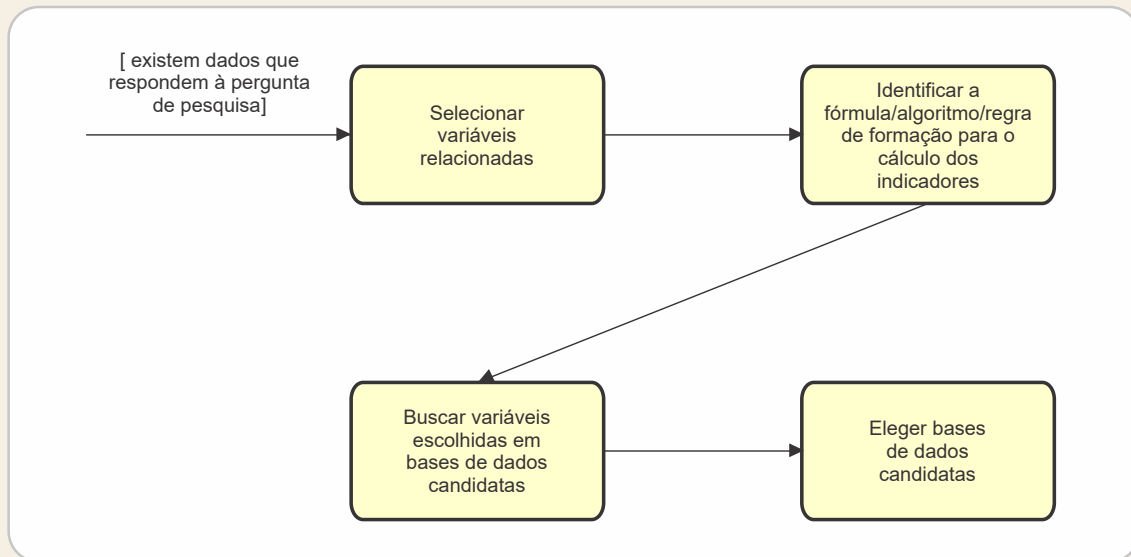
Figura 3 – Trecho inicial do fluxo de atividades do gerador de demografias



Fonte: elaboração própria.

O desafio seguinte ao atendimento da condição de guarda “*existem dados que respondem à pergunta*” é a seleção de variáveis que comporão a fórmula, algoritmo ou regra de formação necessária ao cálculo dos indicadores que serão produzidos, com o propósito de responder ao questionamento investigativo solicitado. Assim, será necessário identificar em qual(is) local(is) essa(s) variável(s) se encontra(m). Para isso, imperiosa é a tarefa de identificação desses elementos de dados nas bases de dados candidatas, que podem ser inicialmente (mas não obrigatoriamente) sugeridas pelos clientes contratantes de uma determinada demografia. A Figura 4 modela a situação descrita.

Figura 4 – Trecho do fluxo de atividades referente à produção e recuperação dos elementos de dados necessários para a resposta à pergunta de pesquisa



Fonte: elaboração própria.

Utilizaremos a partir de agora os termos “*variáveis*”, “*atributos*” e “*bases de dados*” no plural, ressaltando-se que é possível que aconteça uma única ocorrência de cada elemento desses em um determinado contexto.

Importante ressaltar, ainda no âmbito do trecho de fluxo supracitado, que existe uma relação de mapeamento entre as variáveis que responderão à pergunta de pesquisa e os atributos das bases de dados candidatas. Essas variáveis podem se refletir diretamente nos atributos de uma mesma base, ou serem compostas por atributos de bases diversas, ou mesmo variáveis derivadas, que necessitam de dados externos.

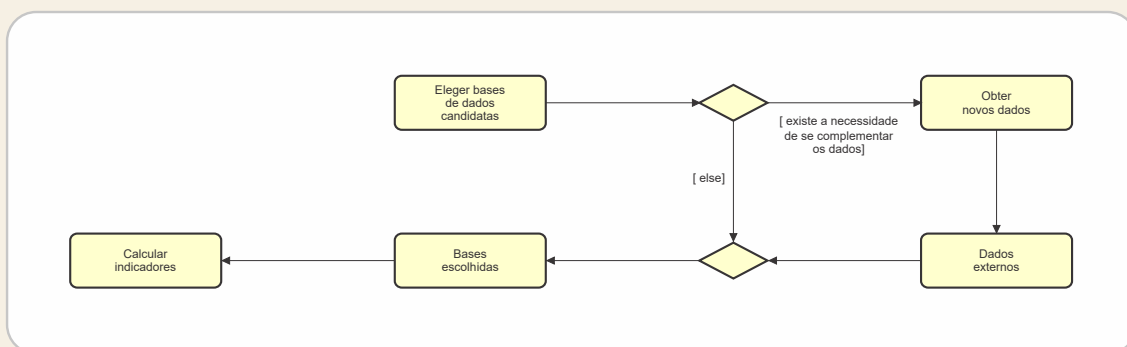
A escolha das bases a serem utilizadas é a última atividade desse fluxo. Ela pode, mas não necessariamente utiliza, como dito anteriormente, as bases de dados sugeridas pelo cliente contratante da demografia. Os especialistas técnicos e/ou pesquisadores da área podem eleger outras bases que considerem apropriadas, desde que justificadas as utilizações. É de todo conveniente também justificar o descarte de bases de dados sugeridas pelo cliente, em prol da completude de todo o processo.

Dando continuidade ao processo, a Figura 5 apresenta, a partir da eleição das bases de dados, o questionamento do acréscimo de dados externos para a completa resposta à pergunta de pesquisa. Assim, a condição de guarda “*existe a necessidade de se complementar os dados?*” é ativada. Se a resposta for afirmativa, a atividade “*Obter*

novos dados” é executada, e o objeto externo “*Dados externos*” é armazenado. Note que não se trata, na prática, de um único objeto. Referimo-nos, na verdade, a todos os dados externos, sejam de sites da internet, arquivos avulsos, ou quaisquer outras fontes necessárias à completação dos dados para responder ao questionamento proposto.

Se a resposta para a condição de guarda for negativa, significa que não há dados externos a serem buscados. Porém, em ambos os casos, bases de dados foram eleitas, e armazenadas no objeto externo “*Bases escolhidas*”. Da mesma forma, o objeto externo representa o conjunto de bases de dados apontadas como necessárias ao atendimento da demanda original.

Figura 5 – Trecho do fluxo que trata da eleição das fontes de dados, sejam bases de dados ou dados externos, e o posterior cálculo de indicadores

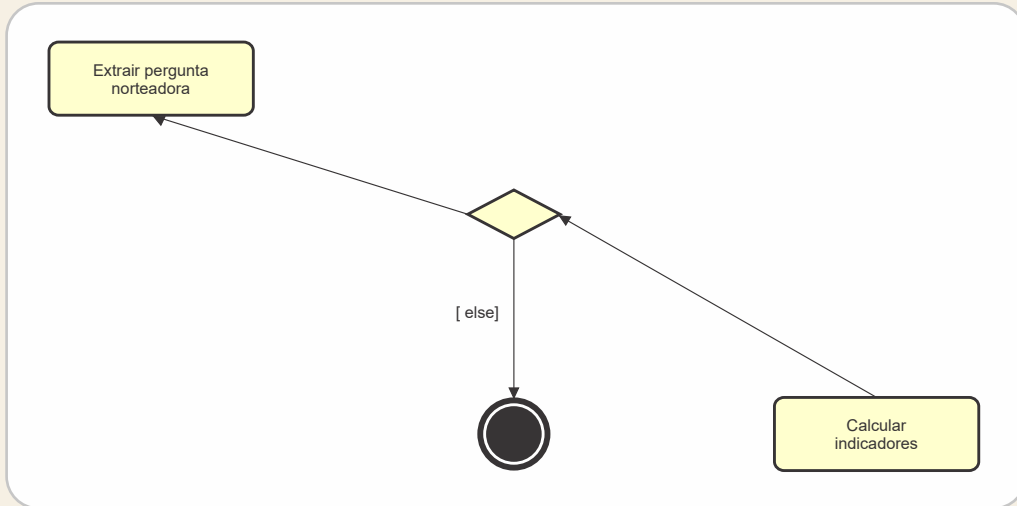


Fonte: elaboração própria.

De posse de todos os dados necessários, a última atividade do trecho, “*Calcular indicadores*”, é executada. Ela produzirá os insumos necessários para a produção da resposta para a pergunta norteadora.

Por fim, a Figura 6 ilustra o processo de término das etapas necessárias à produção da resposta da pergunta de pesquisa em estudo. Após a produção dos indicadores, realizada pela atividade “*Calcular indicadores*”, executa-se a condição de guarda “*existem perguntas a serem respondidas?*”. Em caso positivo, o processo reinicia com a execução da atividade “*Extraír pergunta norteadora*”. Caso contrário, toda a atividade se encerra.

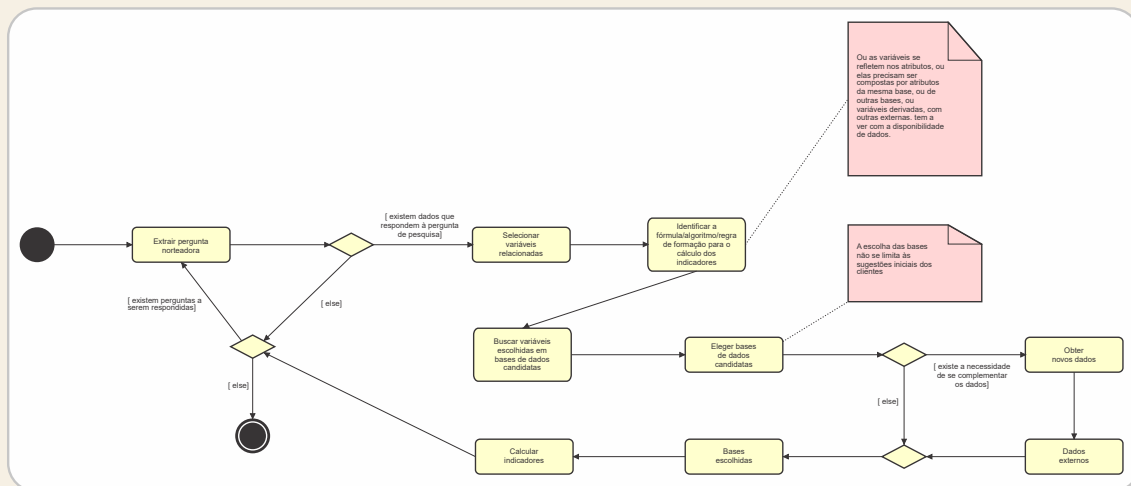
Figura 6 – Trecho final do fluxo de atividades do gerador de demografias



Fonte: elaboração própria.

A Figura 7 ilustra o diagrama de atividades completo do gerador de demografias, materializando o seu comportamento.

Figura 7 – Diagrama de atividades completo do gerador de demografias



Fonte: elaboração própria.

Enfim, o desenvolvimento de um projeto de construção de um gerador de demografias é ousado, porém instigante e recompensador. O estágio apresentado no presente texto revela o grau de maturidade em que o grupo integrante do projeto atualmente se encontra, após vários sucessos e insucessos. A pretensão é da utilização dessa ideia em demografias de diferentes naturezas, sejam profissões, ou quaisquer outros temas. O certo é que o trabalho não se encerra por aqui. Temos um projeto com muitos desafios por explorar.

Conte-nos o que pensa sobre esta publicação.
Clique aqui e responda a pesquisa.



Biblioteca Virtual em Saúde do Ministério da Saúde
bvsmms.saude.gov.br



MINISTÉRIO DA
SAÚDE

