



PainelPedBR: Um Painel Analítico da Oferta de Alta Complexidade Pediátrica no Brasil

Trabalho de Conclusão de Curso

Karinna Lima Santos

Pablo Vieira Florentino
Orientador

Marcus Vinicius Teixeira Navarro
Co-orientador

Instituto Federal da Bahia – IFBA
Curso de Análise e Desenvolvimento de Sistemas
Campus Salvador

Salvador, Bahia, Brasil
Julho 2026

Resumo

A distribuição territorial da oferta de alta complexidade pediátrica no Brasil permanece inacessível para a maior parte dos cidadãos e de difícil consulta mesmo para gestores e pesquisadores, apesar da existência de bases de dados públicas que registram sistematicamente informações sobre estabelecimentos de saúde, procedimentos assistenciais e dados populacionais. Este trabalho descreve o desenvolvimento do PainelPedBR, uma aplicação analítica interativa de código aberto que integra o Cadastro Nacional de Estabelecimentos de Saúde (CNES), o Sistema de Gerenciamento da Tabela de Procedimentos do SUS (SIGTAP) e o Censo Demográfico 2022 do IBGE em um ambiente único de consulta, acessível publicamente sem exigência de conhecimento técnico especializado. O desenvolvimento seguiu o paradigma da Design Science Research com abordagem iterativa e validação especializada dos critérios metodológicos junto a pesquisador da área de saúde pública. Como o CNES não dispõe de campo direto de classificação por nível de complexidade, a identificação da oferta de alta complexidade pediátrica foi construída indiretamente a partir do SIGTAP, que classifica cada procedimento por complexidade e faixa etária. Foram classificados 1.195 procedimentos em três categorias — exclusivamente pediátrico, com nomenclatura pediátrica explícita e de faixa etária compartilhada — dos quais derivaram critérios de seleção aplicados em duas camadas hierárquicas: habilitação formal e serviço declarado. Um pipeline ETL modular gerou quinze produtos de dados pré-processados que alimentam seis módulos funcionais: Panorama Nacional, Desertos e Cobertura, Recursos Humanos, Procedimentos, Mapa de Estabelecimentos e Perfil do Estabelecimento. O escopo analítico resultante abrange 86.209 estabelecimentos em 5.546 municípios brasileiros, com indicadores normalizados pela população infantil de zero a dezenove anos. Os resultados demonstram que a integração metodologicamente fundamentada de bases públicas de saúde, quando apresentada em interface acessível, pode transformar um patrimônio informacional técnico em instrumento efetivo de cidadania e de apoio à gestão da saúde pediátrica no Brasil.

Palavras-chave: saúde pediátrica; alta complexidade; visualização de dados em saúde; Shiny; CNES; SIGTAP; engenharia de dados.

Abstract

The territorial distribution of high-complexity pediatric healthcare supply in Brazil remains inaccessible to most citizens and difficult to consult even for managers and researchers, despite the existence of public databases that systematically record information on health facilities, care procedures, and population data. This work describes the development of PainelPedBR, an open-source interactive analytical application that integrates the National Registry of Health Establishments (CNES), the SUS Procedures Table Management System (SIGTAP), and the 2022 IBGE Demographic Census into a single publicly accessible consultation environment requiring no specialized technical knowledge. Development followed the Design Science Research paradigm with an iterative approach and expert validation of methodological criteria by a public health researcher. Since CNES lacks a direct complexity-level classification field, identification of high-complexity pediatric supply was constructed indirectly from SIGTAP, which classifies each procedure by complexity and age range. A total of 1,195 procedures were classified into three categories — exclusively pediatric, with explicit pediatric nomenclature, and shared age range — from which selection criteria were derived and applied in two hierarchical layers: formal accreditation and declared services. A modular ETL pipeline generated fifteen pre-processed data products feeding six functional modules: National Overview, Coverage Gaps, Human Resources, Procedures, Establishment Map, and Establishment Profile. The resulting analytical scope covers 86,209 establishments across 5,546 Brazilian municipalities, with indicators normalized by the child population aged zero to nineteen years. The results demonstrate that methodologically grounded integration of public health databases, when presented in an accessible interface, can transform a restricted technical heritage into an effective instrument of citizenship and support for pediatric healthcare management in Brazil.

Keywords: pediatric healthcare; high complexity; health data visualization; Shiny; CNES; SIGTAP; data engineering.

Lista de Figuras

Figura 1 — Arquitetura de Alto Nível da Solução	36
Figura 2 — Diagrama de componentes do PainelPedBR	46
Figura 3 — Interface do módulo Panorama Nacional	52
Figura 4 — Interface do módulo Sobre	54
Figura 5 — Interface do módulo Desertos e Cobertura	56
Figura 6 — Interface do módulo Recursos Humanos	57
Figura 7 — Interface do módulo Procedimentos	58
Figura 8 — Interface do módulo Mapa de Estabelecimentos (modo Explorar).....	59
Figura 9 — Interface do módulo Mapa de Estabelecimentos (modo Comparar).....	60
Figura 10 — Interface do módulo Perfil do Estabelecimento.....	61

Lista de Tabelas

Tabela 1 — Caracterização operacional das fontes de dados.....	23
Tabela 2 — Classificação tripartite de procedimentos pediátricos de alta complexidade.....	25
Tabela 3 — Modelo de duas camadas de seleção de estabelecimentos.....	27
Tabela 4 — Pacotes R utilizados no PainelPedBR	37
Tabela 5 — Organização dos scripts do pipeline ETL.....	38
Tabela 6 — Produtos de dados gerados (competência dezembro de 2024)	41
Tabela 7 — Arquivos de teste	49
Tabela 8 — Mapeamento das interações de navegação entre módulos do painel...	63

Sumário

Resumo.....	2
Abstract.....	4
Lista de Figuras.....	4
Lista de Tabelas	5
1 Visão Geral.....	6
1.1 Declaração do Problema.....	8
1.2 Objetivos.....	9
1.3 Trabalhos Relacionados.....	9
1.4 Proposta de Solução de Software	11
2 Fundamentação Teórica.....	11
2.1 Contexto Normativo e Institucional	12
2.2 Sistemas de Informação em Saúde.....	14
2.3 Fundamentos Técnicos	18
3 Metodologia	21
3.1 Abordagem de Desenvolvimento	22
3.2 Operacionalização das Fontes de Dados	22
3.3 Metodologia de Classificação de Procedimentos	24
3.4 Modelo de Duas Camadas de Seleção de Estabelecimentos	26
3.5 Limitações Metodológicas	27
4 Requisitos	29
4.1 Requisitos Funcionais.....	29
4.2 Requisitos Não-Funcionais.....	33
5 Design e Implementação.....	34
5.1 Arquitetura Geral da Solução	35
5.2 Stack Tecnológico	36
5.3 Pipeline ETL.....	37
5.4 Estrutura de Dados	43
5.5 Arquitetura de Componentes.....	44
5.6 Design da Interface	46
6 Testes de Software.....	48
6.1 Estratégia de Testes.....	48
6.2 Testes Unitários com testthat	49
6.3 Resultados e Cobertura.....	50
7 Implantação	50

7.1 Reprodutibilidade com renv.....	50
7.2 Publicação Online	50
8 Manual do Usuário.....	51
8.1 Requisitos de Acesso	51
8.2 Navegação Geral.....	51
8.3 Cruzamento de Informações.....	62
9 Considerações Finais	65
Referências	68
Apêndice A — Glossário, Siglas e Abreviações.....	71

1 Visão Geral

A atenção à saúde no SUS é organizada em três níveis de complexidade — atenção básica, média complexidade e alta complexidade —, estruturados de forma hierárquica e complementar para garantir a integralidade do cuidado (BRASIL, 1990). Essa arquitetura segue o princípio da regionalização: cada nível reúne ações com densidade tecnológica e escala de produção compatíveis com sua função no sistema, cabendo à Atenção Básica coordenar o fluxo do usuário entre eles (MENDES, 2011).

A atenção básica constitui a porta de entrada preferencial do SUS e o primeiro ponto de contato da população com os serviços de saúde. Por meio das Unidades Básicas de Saúde (UBS) e da Estratégia Saúde da Família (ESF), são ofertadas ações de promoção da saúde, prevenção de doenças, diagnóstico e tratamento de condições de alta prevalência e menor densidade tecnológica (BRASIL, 2012).

A média complexidade compreende ações e serviços que requerem profissionais especializados e recursos diagnósticos e terapêuticos de maior especificidade, sem demandar a infraestrutura de um hospital de alta tecnologia. Incluem-se neste nível consultas com especialistas, exames de imagem e laboratoriais de maior custo, procedimentos ambulatoriais especializados e internações clínico-cirúrgicas de menor porte. O acesso é mediado por centrais de regulação municipais e estaduais, que organizam os fluxos segundo a capacidade instalada de cada região de saúde (BRASIL, 2002).

A alta complexidade concentra os procedimentos de maior densidade tecnológica, custo elevado e exigência de infraestrutura especializada: oncologia, transplantes de órgãos e tecidos, cirurgia cardíaca, neurocirurgia, UTI neonatal e pediátrica, reabilitação de alta complexidade, entre outros. Esses procedimentos são normatizados pelo Ministério da Saúde e habilitados formalmente para cada estabelecimento que comprova capacidade técnica e estrutural para executá-los.

A distribuição dos recursos de alta complexidade não é apenas uma questão assistencial — é uma questão de planejamento territorial e de equidade em saúde. São esses recursos que determinam se um paciente com câncer, insuficiência renal ou cardiopatia congênita receberá tratamento no seu município, na sua região, ou precisará percorrer centenas de quilômetros para acessar o cuidado adequado.

Para a população em geral — e especialmente para famílias com crianças que necessitam de cuidados especializados —, conhecer a rede de alta complexidade disponível no seu território é uma forma de exercer o direito à informação em saúde. Saber se há uma UTI pediátrica, um centro de oncologia infantil ou um serviço de reabilitação acessível na sua cidade ou região permite que famílias tomem decisões mais informadas e, quando necessário, reivindiquem junto ao sistema público o cuidado que lhes é devido. A transparência sobre essa distribuição é, portanto, um instrumento de cidadania.

Em caráter complementar, gestores e formuladores de políticas também se beneficiam dessa visão territorial: ela subsidia decisões como pleitear novas habilitações junto ao Ministério da Saúde, dimensionar pactuações intermunicipais, identificar vazios assistenciais e orientar a alocação de recursos via emendas e transferências fundo a fundo — inclusive no âmbito de programas como o PNAISC (Política Nacional de Atenção Integral à Saúde da Criança) e das redes temáticas do SUS.

1.1 Declaração do Problema

O Sistema Único de Saúde mantém um conjunto extenso de bases de dados públicas que registram, de forma sistemática e continuamente atualizada, informações sobre a infraestrutura assistencial, os recursos humanos, os procedimentos autorizados e os dados populacionais do país. Essas bases constituem um patrimônio informacional de alto valor, sendo disponibilizadas de forma aberta pelo Ministério da Saúde e por outros órgãos federais (BRASIL, 2024).

Na prática, porém, esse patrimônio permanece inacessível para a maior parte da população. Uma família que precisa saber se há uma UTI pediátrica ou um centro de oncologia infantil acessível na sua cidade ou região não dispõe de nenhuma ferramenta de consulta direta, simples e integrada para responder a essa pergunta — embora a Lei nº 8.080/1990 garanta expressamente o direito à informação sobre os serviços de saúde disponíveis (BRASIL, 1990). As ferramentas de consulta pública existentes, como o TabNet, mantido pelo Departamento de Informática do SUS, permitem a extração de dados de forma tabular e isolada, sem suporte a cruzamentos entre bases distintas e sem recursos de visualização territorial integrada (BRASIL, 2024). Análises que demandam a combinação de informações provenientes de sistemas diferentes exigem extração manual, tratamento e integração por parte do analista, pressupondo domínio técnico em manipulação de dados que não pode ser assumido como dado entre os cidadãos em geral.

Esse cenário se torna especialmente crítico quando o objeto de análise é a oferta de alta complexidade pediátrica no território brasileiro. Trata-se de um segmento assistencial que se distingue dos demais pela exigência de infraestrutura especializada, habilitações formais específicas junto ao Ministério da Saúde e profissionais com formação diferenciada nas subespecialidades pediátricas. Agrava esse cenário o fato de que as próprias bases de dados não dispõem de campos que classifiquem diretamente os estabelecimentos por nível de complexidade assistencial — o que exige que essa identificação seja construída indiretamente, por meio do cruzamento entre a tabela de procedimentos e o cadastro de estabelecimentos. Compreender onde essa oferta se concentra no território — identificando quais estabelecimentos a realizam, em que quantidade, com qual perfil e em relação a qual população — requer justamente o tipo de análise integrada entre múltiplas bases que as ferramentas disponíveis não suportam de forma acessível.

A ausência de uma solução que consolide essas informações em um ambiente analítico integrado, de acesso direto e sem exigência de conhecimento técnico especializado, representa uma lacuna concreta sobretudo para famílias e cidadãos que precisam compreender onde a alta complexidade pediátrica é ofertada no Brasil. Em caráter complementar, essa lacuna também afeta

estudantes, pesquisadores, gestores e formuladores de políticas com interesse no tema. É essa lacuna que o PainelPedBR se propõe a contribuir para minimizar.

1.2 Objetivos

1.2.1 Objetivo Geral

Desenvolver um painel analítico interativo para visualização, identificação e localização da distribuição da oferta de alta complexidade pediátrica no Brasil, por meio da integração de bases de dados públicas oficiais, com o propósito de oferecer à população uma ferramenta de consulta acessível e sem exigência de conhecimento técnico especializado — e, complementarmente, subsidiar gestores, pesquisadores e formuladores de políticas públicas de saúde com informação integrada e territorialmente orientada.

1.2.2 Objetivos Específicos

1. Mapear, coletar e integrar dados públicos provenientes do Cadastro Nacional de Estabelecimentos de Saúde (CNES), do Sistema de Gerenciamento da Tabela de Procedimentos do SUS (SIGTAP) e do Censo Demográfico 2022 do Instituto Brasileiro de Geografia e Estatística (IBGE), aplicando critérios metodológicos explícitos para a identificação da oferta de alta complexidade pediátrica — incluindo a classificação indireta a partir da tabela de procedimentos do SIGTAP, dado que as bases disponíveis não dispõem de campo direto de nível de complexidade por estabelecimento.
2. Construir um pipeline de engenharia de dados modular para extração, limpeza, padronização, harmonização e integração das bases, gerando produtos de dados otimizados para consumo pela aplicação.
3. Definir a arquitetura lógica do modelo de dados, adotando o procedimento ofertado como critério de classificação da oferta de alta complexidade pediátrica e o estabelecimento de saúde como unidade de agregação e apresentação dos recursos, com município e Unidade Federativa como atributos de agrupamento territorial.
4. Desenvolver a aplicação em arquitetura modular, com separação clara entre as camadas de dados pré-processados, lógica de negócio e interface do usuário, organizada em módulos funcionais agrupados em três eixos temáticos: Análise (Panorama Nacional e Desertos e Cobertura), Recursos (Recursos Humanos e Procedimentos) e Exploração (Mapa de Estabelecimentos e Perfil do Estabelecimento).
5. Documentar tecnicamente todas as etapas do desenvolvimento, incluindo decisões metodológicas do pipeline de dados, critérios de classificação de procedimentos, modelo lógico de dados, arquitetura da aplicação e limitações identificadas.

1.3 Trabalhos Relacionados

A análise das ferramentas existentes para acesso e visualização de dados de saúde pública no Brasil permite situar o PainelPedBR no ecossistema disponível e evidenciar a lacuna que a aplicação se propõe a preencher. As iniciativas identificadas organizam-se em duas categorias: ferramentas oficiais de consulta pública mantidas pelo Ministério da Saúde e ferramentas de acesso programático desenvolvidas pela comunidade de pesquisa.

O TabNet é a principal ferramenta oficial de consulta às bases do DATASUS e constitui a referência central de acesso público aos dados do SUS (BRASIL, 2024). A ferramenta permite a construção de tabelas de contingência a partir de diferentes bases, com flexibilidade na seleção de variáveis e filtros. No entanto, sua limitação estrutural em relação ao PainelPedBR é que cada consulta se restringe a uma única fonte de dados. Não é possível, por meio do TabNet, cruzar informações sobre infraestrutura de estabelecimentos com a classificação normativa de procedimentos ou com dados populacionais em uma única operação. A ferramenta não oferece recorte por faixa etária atendida, não produz indicadores normalizados por população e não disponibiliza visualizações geoespaciais integradas. Para um cidadão sem formação técnica em análise de dados, essas limitações tornam a ferramenta inoperável como instrumento de consulta direta.

O portal de consulta do CNES, também mantido pelo DATASUS, permite a busca de estabelecimentos de saúde por Unidade Federativa, município e tipo de unidade, com acesso a algumas informações cadastrais da unidade selecionada (BRASIL, 2024). Sua limitação central, para os propósitos do PainelPedBR, reside justamente na ausência de critério de classificação por nível de complexidade assistencial ou por faixa etária atendida — o que confirma que a identificação da oferta de alta complexidade pediátrica não pode ser obtida diretamente do CNES, exigindo cruzamento com outras bases. O portal não integra dados de outras fontes e não oferece recursos de análise territorial comparativa ou de normalização por população.

No campo das ferramentas de acesso programático, o pacote microdatasus, desenvolvido para o ambiente R, oferece funções para extração e pré-processamento de microdados de diferentes sistemas de informação em saúde do DATASUS, incluindo o CNES (SALDANHA; BASTOS; BARCELLOS, 2019). Trata-se de uma contribuição relevante para a comunidade de pesquisadores, pois automatiza etapas de extração que demandam acesso manual às bases. No entanto, o pacote se posiciona como uma biblioteca de acesso a dados — não como um painel analítico para o usuário final. Ele viabiliza a extração das bases, mas não realiza integração entre fontes, não aplica critérios metodológicos de classificação de procedimentos e não entrega uma interface interativa de visualização.

A análise dessas iniciativas evidencia uma lacuna objetiva: não existe, atualmente, uma ferramenta que integre dados do CNES, da tabela de procedimentos do SIGTAP e do Censo Demográfico com critério metodológico explícito para a identificação da oferta de alta complexidade

pediátrica, entregando o resultado em uma interface interativa e acessível ao público em geral — com georreferenciamento, indicadores normalizados por população pediátrica e navegação progressiva do panorama nacional ao perfil do estabelecimento. É nesse espaço que o PainelPedBR se insere.

1.4 Proposta de Solução de Software

O PainelPedBR é uma aplicação analítica interativa de código aberto concebida para consolidar, em um ambiente único e acessível, informações sobre a distribuição da oferta de alta complexidade pediátrica no Brasil. A aplicação integra dados do Cadastro Nacional de Estabelecimentos de Saúde (CNES), do Sistema de Gerenciamento da Tabela de Procedimentos do SUS (SIGTAP) e do Censo Demográfico do Instituto Brasileiro de Geografia e Estatística (IBGE), respondendo a uma lacuna concreta: a inexistência de uma ferramenta que torne essa informação disponível de forma direta e acessível, sem exigir do usuário conhecimento especializado em manipulação de dados.

A solução abrange quatro dimensões complementares da análise assistencial. A dimensão da infraestrutura compreende as habilitações formais, os tipos de leito e o perfil detalhado de cada estabelecimento com oferta de alta complexidade pediátrica identificado. A dimensão territorial contempla a distribuição geográfica desses recursos em múltiplas escalas — do panorama nacional ao perfil de uma unidade específica — com suporte à comparação entre territórios e à identificação de lacunas de cobertura. A dimensão dos procedimentos abrange a distribuição da oferta de procedimentos pediátricos de alta complexidade por tipo, categoria normativa e território, permitindo identificar quais serviços existem e onde estão disponíveis. A dimensão dos recursos humanos abrange a distribuição dos profissionais pediátricos por especialidade e território, com indicadores normalizados pela população infantil de cada município.

Dois aspectos distinguem o PainelPedBR das ferramentas de consulta atualmente disponíveis. O primeiro diz respeito à base metodológica da identificação dos estabelecimentos: em vez de depender exclusivamente da autodeclaração cadastral, a classificação da oferta de alta complexidade pediátrica é derivada dos critérios normativos da tabela de procedimentos do SIGTAP, o que confere à análise rastreabilidade e consistência metodológica. O segundo diz respeito à forma como os resultados são expressos: os indicadores produzidos pela aplicação relacionam a oferta de serviços à população pediátrica de cada território, permitindo comparações que refletem a real capacidade de cobertura e não apenas a distribuição absoluta dos recursos.

O resultado é uma ferramenta que transforma dados dispersos em informação estruturada e navegável, colocando à disposição da população uma visão integrada da oferta de alta complexidade pediátrica que, até o momento, só seria acessível mediante extração e cruzamento manual de múltiplas bases públicas.

2 Fundamentação Teórica

O desenvolvimento do PainelPedBR articula conhecimentos de três domínios distintos, cuja compreensão é necessária para fundamentar as decisões de projeto tomadas ao longo do trabalho. O primeiro diz respeito ao contexto normativo e institucional que rege a organização da atenção à saúde no Brasil, em particular a estruturação da atenção pediátrica de alta complexidade no âmbito do SUS e os critérios que definem o recorte etário adotado. O segundo abrange os sistemas de informação em saúde que constituem as fontes de dados do projeto — sua natureza institucional, estrutura e papel dentro do ecossistema de dados públicos do SUS —, incluindo o marco legal que regula o acesso a essas informações. O terceiro compreende os fundamentos técnicos que orientam as escolhas de engenharia e desenvolvimento: os processos de integração de dados heterogêneos, os princípios de visualização de dados aplicados ao contexto da saúde pública e a arquitetura de aplicações analíticas interativas.

Esta seção está organizada em três partes correspondentes a esses domínios: Contexto Normativo e Institucional, Sistemas de Informação em Saúde e Fundamentos Técnicos.

2.1 Contexto Normativo e Institucional

A compreensão das decisões metodológicas que fundamentam o PainelPedBR exige, como ponto de partida, a caracterização do arcabouço normativo e institucional que organiza a atenção à saúde no Brasil. Esta seção apresenta os elementos desse contexto que são diretamente relevantes para o projeto: a estrutura do Sistema Único de Saúde e o modelo de organização da atenção de alta complexidade, o marco normativo específico da atenção à saúde da criança e, por fim, a discussão sobre o critério etário pediátrico adotado, cuja definição não é unívoca no ordenamento brasileiro e impõe ao projeto uma escolha metodológica explícita e justificada.

2.1.1 O Sistema Único de Saúde e a organização da atenção de alta complexidade

O Sistema Único de Saúde foi instituído pela Constituição Federal de 1988, que estabeleceu a saúde como direito de todos e dever do Estado, organizando o sistema sob os princípios de universalidade, integralidade e equidade (BRASIL, 1988, art. 196–200). A regulamentação infraconstitucional do SUS foi consolidada pela Lei n. 8.080, de 19 de setembro de 1990, que dispõe sobre as condições para a promoção, proteção e recuperação da saúde e define a organização e o funcionamento dos serviços em todo o território nacional (BRASIL, 1990). Desde sua criação, o SUS constitui um dos maiores sistemas públicos de saúde do mundo, com cobertura universal e estrutura descentralizada que articula os três níveis federativos (PAIM et al., 2011).

A organização da atenção à saúde no SUS estrutura-se em níveis hierárquicos de complexidade crescente, compreendendo a atenção básica, a média complexidade e a alta complexidade. Esse modelo foi aprofundado com a concepção das Redes de Atenção à Saúde, definidas como arranjos organizativos de ações e serviços de saúde de diferentes densidades tecnológicas que, integradas por sistemas de apoio técnico, logístico e de gestão, buscam garantir a integralidade do cuidado (MENDES, 2011). A Portaria GM/MS n. 4.279, de 30 de dezembro de 2010, estabeleceu as diretrizes para a organização dessas redes no âmbito do SUS, consolidando o modelo de atenção em rede como referência para o planejamento e a prestação dos serviços (BRASIL, 2010).

A alta complexidade corresponde ao conjunto de procedimentos que, pela sua densidade tecnológica e pelo custo elevado, exige estrutura física especializada, equipamentos de alta tecnologia e profissionais com formação diferenciada (BRASIL, 2017). Para viabilizar a oferta desses serviços com segurança e qualidade, o SUS adota o mecanismo das habilitações: instrumentos formais pelos quais o Ministério da Saúde credencia um estabelecimento para a realização de determinados procedimentos, condicionando esse credenciamento à comprovação de requisitos estruturais mínimos — entre os quais a disponibilidade de leitos específicos, equipamentos e profissionais qualificados. Esse modelo de credenciamento seletivo implica que a distribuição dos serviços de alta complexidade pelo território nacional reflete não apenas a capacidade instalada das regiões, mas também a conformidade dos estabelecimentos com exigências normativas rigorosas, o que torna a análise dessa distribuição um objeto de interesse tanto para o planejamento assistencial quanto para a pesquisa em saúde pública.

2.1.2 Política Nacional de Atenção à Saúde da Criança

A atenção à saúde da criança no âmbito do SUS é orientada pela Política Nacional de Atenção à Saúde da Criança (PNAISC), instituída pela Portaria GM/MS n. 1.130, de 5 de agosto de 2015. A política estabelece diretrizes e objetivos para a organização da atenção integral à saúde da criança de zero a nove anos, com extensão de algumas ações ao adolescente, e define como eixos estratégicos a atenção humanizada ao recém-nascido, o aleitamento materno, a alimentação saudável, o desenvolvimento integral, a atenção às crianças em situação de vulnerabilidade e a vigilância e prevenção de óbitos infantis (BRASIL, 2015). A PNAISC representa o principal marco normativo federal que orienta a organização dos serviços pediátricos no SUS e fundamenta a relevância da análise da rede de atenção voltada a esse segmento populacional.

No que diz respeito à alta complexidade pediátrica especificamente, a política reconhece a necessidade de garantir acesso a serviços especializados para condições de saúde que demandam cuidado intensivo, procedimentos de alta densidade tecnológica e acompanhamento multiprofissional. Serviços como Unidades de Terapia Intensiva Neonatal e Pediátrica, centros de oncologia infantil, unidades de transplante pediátrico e serviços de cardiologia intervencionista congênita constituem componentes críticos da rede de atenção à saúde da criança, cuja organização e distribuição territorial têm impacto direto sobre o acesso e os desfechos clínicos dessa população (VICTORA et al., 2011). A

literatura evidencia que a concentração geográfica desses serviços nos grandes centros urbanos impõe barreiras de acesso às populações residentes em regiões menos assistidas, o que reforça a importância de instrumentos que permitam monitorar e analisar essa distribuição de forma sistemática (MACINKO; HARRIS, 2015).

A PNAISC articula-se com outras políticas e redes temáticas do SUS que incidem sobre a atenção pediátrica de alta complexidade, entre as quais a Rede Cegonha — instituída pela Portaria GM/MS n. 1.459, de 24 de junho de 2011 —, voltada à atenção ao parto, ao nascimento e ao desenvolvimento infantil, com ênfase na qualificação das Unidades de Terapia Intensiva Neonatal (BRASIL, 2011). Esse conjunto normativo constitui o referencial institucional que sustenta a definição do escopo assistencial do PainelPedBR e justifica o recorte de alta complexidade adotado no projeto.

2.1.3 Delimitação da Faixa Etária Pediátrica

A definição da faixa etária pediátrica não é unívoca no ordenamento brasileiro. Diferentes instrumentos normativos adotam limites distintos, o que impõe ao projeto uma decisão metodológica explícita e justificada, tratada na seção 3.3.

O Estatuto da Criança e do Adolescente, instituído pela Lei n. 8.069, de 13 de julho de 1990, define criança como a pessoa com até doze anos de idade incompletos e adolescente como aquela entre doze e dezoito anos, estabelecendo um limite superior de dezessete anos, onze meses e vinte e nove dias para o conjunto dessa população (BRASIL, 1990). No âmbito dos serviços de pediatria do SUS, a Portaria GM/MS n. 1.130/2015 adota como referência para atendimento em serviços pediátricos o limite de catorze anos, onze meses e vinte e nove dias, sem prejuízo de ajustes por norma local (BRASIL, 2015).

A Organização Mundial da Saúde adota definição mais abrangente, compreendendo o período de zero a dezenove anos, onze meses e vinte e nove dias como faixa pediátrica (OMS, 2023). Esse critério é referendado pela Sociedade Brasileira de Pediatria como parâmetro para a prática clínica pediátrica no Brasil (SBP, 2020). A justificativa clínica para esse limite mais amplo reside no fato de que diversas condições de saúde de início na infância — como cardiopatias congênitas, doenças onco-hematológicas e erros inatos do metabolismo — frequentemente demandam acompanhamento especializado em serviços pediátricos além dos catorze anos, e a interrupção abrupta do cuidado ao atingir esse limite comprometeria a continuidade assistencial (FERRI et al., 2011).

A pluralidade de referências normativas existentes evidencia que a delimitação da faixa etária pediátrica é uma decisão com implicações diretas sobre o escopo de qualquer análise que tome essa população como objeto, demandando escolha fundamentada e transparente por parte do projeto.

2.2 Sistemas de Informação em Saúde

A construção do PainelPedBR apoia-se em três bases de dados públicas oficiais cujo entendimento é indispensável para a compreensão das decisões de projeto. Esta seção apresenta, inicialmente, o marco legal que regula o acesso a dados governamentais no Brasil e no qual essas bases se inserem. Em seguida, descreve cada uma das fontes sob a perspectiva de sua natureza institucional, estrutura e papel dentro do ecossistema de informação em saúde do SUS. Os detalhes operacionais sobre versões utilizadas, métodos de acesso e decisões de transformação aplicadas sobre cada base são tratados na seção de Metodologia.

2.2.1 Dados governamentais abertos no contexto brasileiro

A disponibilidade das bases de dados públicas utilizadas no PainelPedBR insere-se em um movimento mais amplo de abertura de dados governamentais que se consolidou no Brasil ao longo das últimas décadas. O marco legal central desse movimento é a Lei n. 12.527, de 18 de novembro de 2011, conhecida como Lei de Acesso à Informação (LAI), que regulamentou o direito de acesso às informações públicas previsto na Constituição Federal e estabeleceu como diretriz a disponibilização de informações em formatos abertos, estruturados e legíveis por máquina (BRASIL, 2011). A LAI impõe aos órgãos públicos o dever de divulgação proativa de informações de interesse coletivo, independentemente de solicitação, o que fundamenta a obrigatoriedade de publicação periódica das bases do DATASUS e do IBGE.

O conceito de dados governamentais abertos vai além da simples disponibilização de informações: preconiza que os dados públicos sejam liberados em formatos não proprietários, processáveis por máquina e sem restrições de uso, de forma a permitir que a sociedade civil, pesquisadores e desenvolvedores possam utilizá-los para produzir conhecimento e aplicações de interesse público (ISOTANI; BITTENCOURT, 2015). O Brasil integra a Open Government Partnership desde 2011, tendo desenvolvido sucessivos planos de ação nacionais voltados à ampliação da transparência e ao fortalecimento da política de dados abertos (OGP, 2024). No âmbito federal, o Decreto n. 8.777, de 11 de maio de 2016, instituiu a Política de Dados Abertos do Poder Executivo Federal, estabelecendo diretrizes para a abertura progressiva das bases de dados dos órgãos e entidades federais (BRASIL, 2016).

No setor da saúde especificamente, o DATASUS constitui a principal infraestrutura pública de dados, reunindo e disponibilizando as bases dos sistemas de informação em saúde oriunda do SUS. A política de dados abertos do Ministério da Saúde tem ampliado progressivamente a disponibilidade dessas bases em formatos estruturados e de acesso programático, o que viabiliza iniciativas como o PainelPedBR, que dependem da integração automatizada de múltiplas fontes (AVELINO; POMPEU; FONSECA, 2021). A qualidade e a utilidade dessas bases para fins analíticos, no entanto, dependem não apenas de sua disponibilização, mas também da consistência do preenchimento e da regularidade das atualizações realizadas pelos gestores responsáveis em cada nível federativo.

2.2.2 Cadastro Nacional de Estabelecimentos de Saúde (CNES)

O Cadastro Nacional de Estabelecimentos de Saúde é o sistema de informação oficial para o cadastramento de estabelecimentos de saúde no Brasil, gerenciado pelo Departamento de Informática do SUS. Instituído pela Portaria MS/SAS n. 511, de 29 de dezembro de 2000, o CNES tem como finalidade registrar as características de infraestrutura, recursos humanos, equipamentos, serviços especializados e habilitações de cada unidade de saúde cadastrada no território nacional, independentemente de sua vinculação ao SUS (BRASIL, 2000). O sistema constitui a principal fonte oficial de informações sobre a capacidade instalada da rede assistencial brasileira e serve de base para o planejamento, a regulação e a avaliação dos serviços de saúde pelos gestores nos três níveis federativos.

O CNES abrange a totalidade dos estabelecimentos de saúde em funcionamento no país, compreendendo unidades ambulatoriais, hospitalares, de apoio diagnóstico e terapêutico, de atenção básica e de urgência e emergência, entre outras categorias. As informações registradas organizam-se em diferentes domínios: dados cadastrais e de localização geográfica do estabelecimento, tipo de gestão e natureza jurídica, leitos disponíveis por tipo e destinação, equipamentos instalados, habilitações concedidas pelo Ministério da Saúde, serviços especializados declarados e vínculos de profissionais de saúde (BRASIL, 2000). Essa estrutura multidimensional torna o CNES uma fonte singular para análises que demandam a caracterização simultânea de diferentes dimensões da capacidade assistencial de um estabelecimento.

A base é atualizada mensalmente pelos gestores municipais, estaduais e federais, o que a torna uma fonte dinâmica e sensível a alterações na rede assistencial ao longo do tempo. A disponibilidade de coordenadas geográficas para os estabelecimentos cadastrados viabiliza análises territoriais e a construção de visualizações cartográficas da distribuição dos serviços. Estudos que utilizam o CNES como fonte de dados destacam tanto seu potencial analítico quanto a necessidade de atenção à qualidade do preenchimento, que depende da regularidade e precisão das atualizações realizadas pelos gestores locais (CARVALHO; EDUARDO, 1998; LIMA et al., 2009).

No contexto do PainelPedBR, o CNES constitui a fonte primária de informações sobre os estabelecimentos que compõem a rede de alta complexidade pediátrica, fornecendo os dados de infraestrutura, habilitações, leitos, equipamentos e profissionais que alimentam os módulos analíticos da aplicação. Os detalhes sobre a versão utilizada, o método de acesso e as decisões de transformação aplicadas sobre essa base são tratados na seção de Metodologia.

2.2.3 Sistema de Gerenciamento da Tabela de Procedimentos, Medicamentos e OPM do SUS (SIGTAP)

O Sistema de Gerenciamento da Tabela de Procedimentos, Medicamentos, Órteses, Próteses e Materiais Especiais do SUS é o sistema oficial que gerencia a tabela de procedimentos do SUS, constituindo a referência normativa central para o financiamento, a regulação e a auditoria dos serviços de saúde prestados no âmbito do sistema público. Instituído pela Portaria GM/MS n. 321, de 8 de

fevereiro de 2007, o SIGTAP unificou as tabelas anteriores do Sistema de Informações Ambulatoriais e do Sistema de Informações Hospitalares, consolidando em uma única base a classificação oficial de todos os procedimentos autorizados no SUS (BRASIL, 2007).

Cada procedimento registrado no SIGTAP possui um conjunto de atributos normativos estruturados que descrevem suas características clínicas, administrativas e estruturais. Entre esses atributos, destacam-se o nível de complexidade assistencial — com os valores Atenção Básica, Média Complexidade e Alta Complexidade —, a faixa etária elegível para a realização do procedimento expressa em meses, as habilitações que o estabelecimento deve possuir para executá-lo, os serviços especializados requeridos, os tipos de leito associados e as ocupações profissionais necessárias (BRASIL, 2007; BRASIL, 2017). Essa estrutura multidimensional torna o SIGTAP singular entre as bases públicas do SUS: é a única fonte que define, de forma normativa e estruturada, os requisitos assistenciais e estruturais associados a cada procedimento, incluindo a faixa etária para a qual cada um é tecnicamente indicado.

A tabela é atualizada mensalmente pelo Ministério da Saúde, com cada versão denominada competência, e disponibilizada publicamente para consulta e download. As atualizações incorporam a inclusão de novos procedimentos, alterações nos atributos de procedimentos existentes e exclusões, refletindo as revisões normativas da política assistencial do SUS ao longo do tempo. Essa periodicidade de atualização é relevante para projetos analíticos que utilizam o SIGTAP como referência, pois implica que os critérios normativos podem ser alterados entre competências (BRASIL, 2017).

No contexto do PainelPedBR, o SIGTAP ocupa posição central na arquitetura metodológica do projeto: é a partir de seus atributos normativos que se derivam os critérios de identificação dos procedimentos de alta complexidade pediátrica e, por consequência, os requisitos estruturais utilizados para selecionar os estabelecimentos que compõem a rede analisada. Essa lógica e seus desdobramentos operacionais são detalhados na seção de Metodologia.

2.2.4 Instituto Brasileiro de Geografia e Estatística — Censo 2022

O Instituto Brasileiro de Geografia e Estatística é o órgão federal responsável pela produção e disseminação de informações estatísticas e geográficas sobre o Brasil, constituindo a principal referência nacional para dados populacionais, socioeconômicos e territoriais (IBGE, 2023). O Censo Demográfico, realizado periodicamente pelo IBGE, representa o levantamento mais abrangente sobre a população brasileira, cobrindo a totalidade dos domicílios e residentes no território nacional. O Censo de 2022, cujos resultados foram divulgados a partir de 2023, é o mais recente disponível e constitui a fonte de dados populacionais adotada no PainelPedBR.

Os dados censitários oferecem informações populacionais desagregadas por município, faixa etária, sexo e outras variáveis socioeconômicas, com cobertura de todos os 5.570 municípios brasileiros. Essa granularidade territorial e demográfica é indispensável para a construção de indicadores normalizados por população, que permitem comparações entre territórios de portes

distintos de forma metodologicamente adequada (IBGE, 2023). Sem a normalização por população, a análise da distribuição de recursos de saúde limitar-se-ia a contagens absolutas, que favorecem sistematicamente os territórios mais populosos e obscurecem as disparidades de cobertura per capita.

No contexto do PainelPedBR, os dados do Censo 2022 são utilizados especificamente para a construção de indicadores de cobertura da rede pediátrica de alta complexidade em relação à população infantil de cada território. A definição da faixa etária pediátrica adotada para o recorte populacional — e sua compatibilização com os grupos etários disponíveis na estrutura do Censo — é tratada na seção de Metodologia.

2.3 Fundamentos Técnicos

O desenvolvimento do PainelPedBR mobiliza conhecimentos de três áreas técnicas cujos fundamentos conceituais são apresentados nesta seção. A primeira é a engenharia de dados, com foco nos processos de extração, transformação e carga que viabilizam a integração de fontes heterogêneas em um modelo analítico unificado. A segunda é a visualização de dados aplicada ao contexto da saúde pública, área que orienta as escolhas de representação visual e de design de interface da aplicação. A terceira é a arquitetura de aplicações analíticas interativas em *R/Shiny*, com ênfase no modelo reativo e no padrão modular que estruturam o desenvolvimento da ferramenta.

2.3.1 Engenharia de dados: ETL e integração de fontes heterogêneas

A engenharia de dados compreende o conjunto de práticas, processos e tecnologias voltados à coleta, ao tratamento e à disponibilização de dados para fins analíticos. No contexto de projetos que dependem de múltiplas fontes de dados, o processo central é o ETL — acrônimo para *Extract, Transform, Load* —, que descreve o fluxo de extração de dados de sistemas de origem, aplicação de transformações de qualidade e conformidade, e carga dos dados resultantes em um destino analítico (KIMBALL; CASERTA, 2004). Vassiliadis (2009), em revisão abrangente da literatura sobre o tema, estima que processos ETL consomem entre 60% e 80% do esforço total em projetos de integração de dados, evidenciando sua centralidade no desenvolvimento de soluções analíticas.

A fase de extração envolve a coleta de dados brutos a partir de sistemas de origem que frequentemente diferem entre si em formato, protocolo de acesso, periodicidade de atualização e estrutura de dados. No contexto de bases governamentais brasileiras, essa heterogeneidade é particularmente pronunciada: diferentes sistemas adotam formatos de arquivo distintos, *encodings* variados, convenções de nomenclatura próprias e interfaces de acesso que vão de portais de download manual a serviços de API estruturados (SALDANHA; BASTOS; BARCELLOS, 2019). A robustez da camada de extração é determinante para a confiabilidade de todo o pipeline, pois falhas ou inconsistências nessa fase propagam-se para as etapas subsequentes.

A fase de transformação compreende as operações de limpeza, padronização, harmonização e enriquecimento dos dados extraídos. Limpeza refere-se à identificação e ao tratamento de valores

ausentes, duplicados ou inconsistentes. Padronização envolve a conversão de representações distintas para um formato comum — como a normalização de *encodings* de texto ou a uniformização de formatos de data. Harmonização diz respeito ao alinhamento semântico entre campos de fontes distintas que representam o mesmo conceito com nomenclaturas ou codificações diferentes. Enriquecimento consiste na derivação de novos atributos a partir dos dados existentes, como o cálculo de indicadores normalizados por população (KIMBALL; ROSS, 2013). A qualidade dos dados resultantes dessa fase determina diretamente a confiabilidade das análises produzidas pela aplicação.

A fase de carga consiste na persistência dos dados transformados em uma estrutura otimizada para o consumo analítico. Em projetos baseados em R/Shiny, uma estratégia comum é a geração de produtos de dados pré-processados em formatos de leitura eficiente, que são carregados em memória na inicialização da aplicação, eliminando a necessidade de transformações em tempo de execução e garantindo tempos de resposta adequados para a interatividade da interface (WICKHAM, 2021). Além da eficiência no acesso aos dados, outro requisito relevante em pipelines de dados é a garantia de que sua reexecução produza sempre os mesmos resultados. Essa propriedade, denominada idempotência, assegura que o processo possa ser repetido — seja para incorporar uma nova competência dos dados de origem, seja para corrigir uma falha intermediária — sem que os produtos gerados sejam alterados de forma inesperada, o que contribui diretamente para a reprodutibilidade e a confiabilidade do processo (HUMMER et al., 2013).

2.3.2 Visualização de dados aplicada à saúde pública

A visualização de dados é a representação gráfica de informações com o objetivo de facilitar a compreensão, a análise e a comunicação de padrões e relações contidas nos dados (TUFTE, 2001). No contexto da saúde pública, onde decisões de planejamento e alocação de recursos dependem da capacidade de interpretar volumes expressivos de dados heterogêneos, a visualização eficaz desempenha papel instrumental: transforma dados brutos em informação analítica acessível a diferentes perfis de usuários, incluindo gestores que não possuem formação técnica em análise quantitativa (FEW, 2013).

Os fundamentos da visualização eficaz de dados foram sistematizados por Tufte (2001), que estabeleceu princípios como a maximização da razão entre dados e tinta — priorizando representações que transmitam o máximo de informação com o mínimo de elementos gráficos desnecessários — e a integridade gráfica, que exige que as representações visuais preservem as proporções e relações presentes nos dados originais. Cairo (2016) amplia essa perspectiva ao tratar a visualização como uma forma de comunicação que deve equilibrar precisão analítica e acessibilidade ao público-alvo, alertando para os riscos de distorção interpretativa introduzidos por escolhas inadequadas de escalas, cores e tipos de gráficos.

No campo específico dos painéis analíticos interativos, Shneiderman (1996) propôs o princípio que se tornou referência no design de interfaces de exploração de dados: visão geral primeiro, zoom e

filtragem, detalhes sob demanda. Esse princípio orienta a organização hierárquica da informação em aplicações analíticas, permitindo que o usuário parta de uma visão panorâmica e navegue progressivamente para níveis crescentes de granularidade conforme seu interesse analítico. Few (2013) complementa essa abordagem com diretrizes específicas para o design de painéis de monitoramento, enfatizando a importância de contextualizar valores absolutos com referências comparativas — médias, percentis ou benchmarks — para que o usuário possa avaliar a significância do que observa.

A dimensão geoespacial é particularmente relevante na visualização de dados de saúde pública, uma vez que a distribuição territorial dos serviços e recursos é frequentemente o objeto central de análise. Cromley e McLafferty (2012) demonstram que a visualização geoespacial permite identificar padrões de concentração, vazios assistenciais e relações entre a localização dos serviços e as características das populações que deles dependem. Entre as representações cartográficas mais utilizadas nesse contexto, destacam-se os mapas coropléticos — que comunicam valores de variáveis por meio da coloração diferenciada de unidades territoriais, sendo amplamente empregados para comparar indicadores normalizados por população entre territórios de portes distintos — e os mapas de pontos, adequados para representar a localização de estabelecimentos individuais e sua distribuição no espaço geográfico.

A interatividade é um atributo central das aplicações analíticas modernas, distinguindo-as dos relatórios estáticos tradicionais. Filtros dinâmicos, seleção de subconjuntos de dados e navegação progressiva ampliam significativamente a capacidade interacional e exploratória do usuário, permitindo que cada análise seja personalizada conforme o recorte territorial, temporal ou temático de interesse (SHNEIDERMAN, 1996). No contexto do PainelPedBR, esses princípios orientam diretamente as decisões de design da interface, desde a organização modular com navegação progressiva até a escolha dos tipos de visualização empregados em cada módulo analítico.

2.3.3 Arquitetura de aplicações analíticas interativas (Shiny)

O *Shiny* é um framework para o desenvolvimento de aplicações web interativas na linguagem R, que permite construir interfaces analíticas sofisticadas sem exigir conhecimento profundo de tecnologias web como HTML, CSS e JavaScript (CHANG et al., 2024). Desde sua introdução pela Posit — anteriormente RStudio —, o *Shiny* consolidou-se como a principal plataforma para a construção de painéis analíticos no ecossistema R, sendo amplamente adotado em pesquisa científica, setor público e organizações de saúde que utilizam R como linguagem de análise (WICKHAM, 2021).

A arquitetura de uma aplicação *Shiny* fundamenta-se no paradigma da programação reativa, no qual mudanças em entradas — como seleções de filtros ou interações com gráficos — propagam-se automaticamente para as saídas que delas dependem, sem que o desenvolvedor precise gerenciar manualmente a sincronização entre componentes (WICKHAM, 2021). Esse modelo contrasta com a programação imperativa tradicional, em que o fluxo de execução é controlado explicitamente pelo código. Na programação reativa, o framework mantém um grafo de dependências entre entradas e

saídas e garante que apenas os componentes afetados por uma alteração sejam recalculados, o que contribui para a eficiência da aplicação.

Uma aplicação *Shiny* é composta por dois elementos principais: a interface do usuário, que define o *layout* e os componentes visuais da aplicação, e o servidor, que contém a lógica de processamento dos dados e de geração das saídas. Em aplicações de maior complexidade, a organização do código em módulos é a abordagem recomendada: cada módulo encapsula a interface e a lógica de servidor de uma funcionalidade específica, com seu próprio espaço de nomes, o que reduz o acoplamento entre componentes e facilita a manutenção e a evolução da aplicação ao longo do tempo (WICKHAM, 2021). Esse padrão modular é especialmente relevante em painéis analíticos com múltiplas seções funcionais independentes, como é o caso do PainelPedBR.

Do ponto de vista do desempenho, aplicações *Shiny* que operam sobre volumes expressivos de dados beneficiam-se do pré-processamento dos dados fora do ciclo reativo. A estratégia de gerar produtos de dados otimizados na fase de ETL — carregados em memória na inicialização da aplicação — elimina a necessidade de transformações custosas em tempo de execução e garante tempos de resposta adequados para a interatividade da interface, independentemente do volume de dados subjacente (WICKHAM, 2021). A hospedagem de aplicações *Shiny* pode ser realizada por meio da plataforma shinyapps.io, que oferece implantação gerenciada sem necessidade de configuração de infraestrutura própria, tornando-a adequada para projetos acadêmicos e de pequeno porte que demandem acesso público à aplicação.

3 Metodologia

O PainelPedBR é, simultaneamente, um produto de *software* e uma contribuição de pesquisa. Essa dualidade determina a natureza da abordagem metodológica adotada: não se trata de uma pesquisa puramente descritiva ou explicativa, mas de uma investigação orientada à construção de um artefato tecnológico que, por si só, constitui a resposta ao problema identificado. Esse tipo de pesquisa encontra fundamento no paradigma da *Design Science Research*, proposto por Hevner et al. (2004) como referencial para estudos em sistemas de informação que produzem artefatos inovadores como contribuição ao conhecimento, em contraposição a estudos que apenas descrevem ou explicam fenômenos existentes.

O desenvolvimento seguiu uma abordagem iterativa e incremental, organizada em ciclos nos quais decisões de projeto eram implementadas, avaliadas e refinadas antes do avanço para a etapa seguinte. Dois pontos de validação especializada estruturaram esse ciclo: a revisão metodológica dos critérios de classificação de procedimentos junto ao Prof. Marcus Vinicius Teixeira Navarro, especialista em saúde pública do LABPROSAUD/IFBA, e o acompanhamento contínuo do Prof. Pablo Vieira Florentino, orientador do trabalho. Essa estrutura de validação garantiu que as decisões metodológicas tivessem respaldo tanto no domínio da saúde pública quanto nos objetivos técnicos do projeto.

Esta seção detalha a sequência de fases e passos ao longo do desenvolvimento: a operacionalização das fontes de dados, a metodologia de classificação dos procedimentos pediátricos de alta complexidade, o modelo de seleção dos estabelecimentos que compõem a rede analisada e as limitações metodológicas reconhecidas.

3.1 Abordagem de Desenvolvimento

A natureza do projeto — integração de bases públicas com estruturas, formatos e lógicas heterogêneas, combinada com a necessidade de validação especializada dos critérios metodológicos em um domínio fora da competência estrita da engenharia de dados — tornou inadequada qualquer abordagem que pressuponha requisitos completamente estáveis desde o início. Metodologias tradicionais de desenvolvimento em cascata, que separam rigidamente as fases de levantamento de requisitos, projeto e implementação, pressupõem estabilidade de escopo que raramente se verifica em projetos analíticos com múltiplas fontes de dados governamentais (SOMMERVILLE, 2016). Optou-se, portanto, por uma abordagem iterativa e incremental, na qual o trabalho foi organizado em ciclos que produziam incrementos funcionais do sistema, permitindo validação contínua e ajuste de curso antes do avanço para as etapas seguintes.

O desenvolvimento organizou-se em duas grandes frentes sequenciais. A primeira compreende o processo de engenharia de dados, responsável pela extração, transformação e geração dos produtos de dados que alimentam a aplicação. A segunda compreende a aplicação Shiny, construída sobre os produtos de dados gerados pela primeira frente. Essa separação explícita entre as camadas de dados e de interface é uma prática recomendada para aplicações analíticas em R, pois garante que a aplicação opere exclusivamente sobre dados pré-processados carregados em memória, sem dependência de transformações em tempo de execução (WICKHAM, 2021).

Uma decisão transversal ao desenvolvimento foi a separação entre critérios metodológicos e código de implementação. Os critérios que definem o escopo analítico do projeto — classificações, limiares etários, listas de códigos relevantes — foram centralizados em arquivos de configuração no formato YAML, lidos dinamicamente pelo pipeline em cada execução. Essa abordagem facilita a auditoria e a revisão dos critérios por especialistas sem exigir leitura de código — como nas validações realizadas com o Prof. Navarro — e o pipeline exporta automaticamente as listas de códigos derivadas em formato YAML como artefato de rastreabilidade, permitindo verificar exatamente quais procedimentos, habilitações e especialidades compõem o escopo da análise. A reprodutibilidade do ambiente de execução é assegurada pelo `renv`, que registra as versões exatas de todos os pacotes utilizados, e pelo versionamento do código e dos arquivos de configuração no GitHub.

3.2 Operacionalização das Fontes de Dados

O projeto utiliza três fontes de dados públicas oficiais. A fim de garantir consistência entre as fontes e estabilidade do recorte analítico, optou-se por fixar a competência de dezembro de 2024 como ponto de corte para o CNES e o SIGTAP — competência esta já consolidada e sem alterações retroativas esperadas no período de desenvolvimento do projeto. Os dados demográficos e socioeconômicos provêm do IBGE e estão sujeitos às defasagens estruturais de publicação características dessa fonte: a variável de população refere-se ao Censo Demográfico de 2022, cuja divulgação completa ocorreu em 2023, enquanto o indicador de PIB per capita corresponde ao ano de 2023, que representa o exercício mais recente disponível nas Contas Regionais do IBGE no momento da coleta. A Tabela 1 apresenta uma síntese das características operacionais de cada fonte.

Tabela 1 — Caracterização operacional das fontes de dados

Fonte	Recorte	Formato	Método de Acesso
CNES	Dez/2024	CSV (<i>encoding latin1</i> , delimitador ;)	Download HTTPS — DATASUS
SIGTAP	Dez/2024	Posicional de largura fixa com arquivos de layout separados	Download FTP — DATASUS
IBGE (Censo 2022)	População: Censo 2022 PIB per capita (2023)	JSON via API	API SIDRA — IBGE

Fonte: elaboração própria.

O CNES é acessado por meio de download direto via HTTPS do DATASUS, em um arquivo ZIP consolidado que contém múltiplas tabelas em formato CSV (BRASIL, 2024). A análise exploratória dos arquivos identificou duas características estruturais não explicitadas na documentação pública disponível, que exigem tratamento explícito na etapa de extração: a codificação latin1 e o delimitador ponto-e-vírgula, cuja não configuração correta produz corrupção de campos textuais com acentuação. A mesma análise revelou a coexistência de dois identificadores de estabelecimento com papéis distintos. O CO_CNES, de sete dígitos, corresponde ao identificador público do estabelecimento, exposto ao usuário final e utilizado como chave primária nos produtos de dados da aplicação; o CO_UNIDADE, de trinta e um dígitos, opera como chave técnica de integração entre as tabelas internas da base, sendo referenciado como chave estrangeira nas tabelas de habilitações, serviços, leitos e carga horária. A não observância dessa distinção inviabiliza as junções entre tabelas e produz resultados incorretos sem nenhuma mensagem de erro explícita.

O SIGTAP é acessado via FTP do DATASUS. Seus arquivos estão estruturados em formato posicional de largura fixa, acompanhados de arquivos de layout separados que descrevem a posição

inicial, o tamanho e o tipo de cada campo. Esse formato exige a implementação de um parser dedicado no pipeline, que lê as especificações de layout antes de processar os arquivos de dados. As tabelas consumidas abrangem os procedimentos e seus atributos de complexidade e faixa etária, além das tabelas de relacionamento que vinculam cada procedimento às habilitações, serviços especializados, tipos de leito e ocupações profissionais requeridos para sua execução (BRASIL, 2024).

Os dados populacionais e econômicos do IBGE são obtidos por meio da API SIDRA. A população residente desagregada por município e faixa etária provém do Censo Demográfico 2022 (tabela 9514), enquanto o PIB per capita municipal é extraído da tabela 5938, com recorte de 2023 — ano mais recente disponível para essa série. Ambas as consultas retornam JSON e são processadas com lógica de retry configurável. A malha municipal georreferenciada, utilizada para a renderização dos mapas da aplicação, é obtida por meio do pacote R *geobr* (PEREIRA; GONÇALVES, 2024), que disponibiliza os polígonos municipais oficiais do IBGE reprojatados para WGS84.

A integração das três fontes impõe um desafio técnico que merece registro explícito: as bases adotam codificações de município incompatíveis entre si. O CNES utiliza o código municipal de seis dígitos, sem o dígito verificador, enquanto o IBGE adota o código completo de sete dígitos. A harmonização é realizada de forma determinística no pipeline: o código de 6 dígitos utilizado pelo CNES corresponde exatamente aos 6 primeiros dígitos do código de 7 dígitos do IBGE, sendo o sétimo dígito o verificador. A conversão é aplicada programaticamente em cada execução do pipeline, sem necessidade de tabela de correspondência persistida.

3.3 Metodologia de Classificação de Procedimentos

A identificação dos estabelecimentos com oferta de alta complexidade pediátrica pressupõe uma definição operacional de o que constitui, no contexto do SUS, um serviço de alta complexidade voltado à pediatria. A abordagem mais imediata seria partir diretamente do CNES e verificar o que cada estabelecimento declara possuir. No entanto, a análise da estrutura do CNES revelou uma limitação fundamental para esse propósito: a base registra o que o estabelecimento possui — leitos, equipamentos, serviços declarados, habilitações formais — mas não classifica esses recursos por nível de complexidade assistencial nem por faixa etária atendida. Não existe, no CNES, nenhum campo que indique diretamente se um estabelecimento é de alta complexidade ou se seus serviços atendem à população pediátrica. Qualquer inferência sem ancoragem normativa introduziria subjetividade metodológica incompatível com os objetivos do projeto.

Essa limitação redirecionou a análise para o SIGTAP, que é, nas bases públicas do SUS, a única fonte que contém, de forma estruturada e com respaldo normativo, tanto a classificação de complexidade quanto a faixa etária dos procedimentos (BRASIL, 2007). Cada procedimento registrado no SIGTAP possui atributos explícitos de complexidade assistencial e de limites etários mínimo e máximo, expressos em meses. Esses atributos são definidos normativamente pelo Ministério da Saúde e reafirmados a cada competência mensal da tabela. Essa constatação definiu o princípio metodológico

central do projeto: em vez de partir dos estabelecimentos, parte-se dos procedimentos — e a partir deles, derivam-se quais estabelecimentos estão aptos a realizá-los. Essa inversão de lógica, de uma abordagem centrada no estabelecimento para uma abordagem descendente centrada em procedimentos, é a decisão metodológica mais fundamental do projeto.

3.3.1 Critério etário adotado

Conforme discutido na seção 2.1.3, a definição de faixa etária pediátrica não é unívoca no ordenamento brasileiro. O projeto adota o critério da Organização Mundial da Saúde, que delimita a faixa pediátrica em duzentos e trinta e nove meses — equivalente a zero a dezenove anos, onze meses e vinte e nove dias (OMS, 2023). Essa escolha fundamenta-se no argumento de ser o critério mais abrangente, garantindo que nenhum procedimento de alta complexidade com indicação pediátrica seja excluído do escopo analítico.

3.3.2 Classificação tripartite de procedimentos

A aplicação do critério etário sobre os procedimentos de alta complexidade do SIGTAP revelou que uma abordagem binária — incluir ou excluir — não consegue lidar adequadamente com a diversidade de situações presentes na tabela. Alguns procedimentos têm a faixa etária máxima restrita ao intervalo pediátrico, sendo aplicáveis exclusivamente a crianças. Outros não possuem restrição etária registrada, mas seu próprio nome declara o contexto pediátrico. Há ainda procedimentos cuja faixa etária começa na infância mas se estende além dela, sendo aplicáveis tanto a crianças quanto a adultos. Uma abordagem binária produziria subcobertura — pela exclusão de procedimentos clinicamente relevantes — ou sobre cobertura — pela inclusão de procedimentos sem relação com a população pediátrica. A solução adotada foi reconhecer que essas três situações correspondem a três graus distintos de evidência do caráter pediátrico do procedimento e tratá-las como categorias separadas, conforme apresentado na Tabela 2.

Tabela 2 — Classificação tripartite de procedimentos pediátricos de alta complexidade

Categoria	Denominação	Critério operacional
CAT1	Exclusivamente pediátrico	Idade máxima registrada no SIGTAP $\leq$ 239 meses. O procedimento não pode ser realizado em adultos pela própria definição normativa da tabela.
CAT2	Sem restrição etária, com nomenclatura pediátrica explícita	Sem restrição etária registrada, mas o título do procedimento contém ao menos um dos seguintes termos: <i>pediátrico</i> , <i>neonatal</i> , <i>infantil</i> , <i>recém-nascido</i> , <i>lactente</i> , <i>congenito</i> ou <i>premature</i> .

CAT3	Sem restrição etária, com início na faixa pediátrica	Idade mínima ≤ 239 meses e idade máxima > 239 meses. O procedimento é aplicável a crianças, mas não é exclusivo delas.
------	--	--

Fonte: elaboração própria.

A avaliação segue ordem de prioridade CAT1 > CAT2 > CAT3, garantindo que procedimentos com restrição etária formal sejam classificados pela evidência mais forte disponível. A aplicação desses critérios sobre a competência de dezembro de 2024 do SIGTAP resultou em sessenta e seis procedimentos CAT1, dois procedimentos CAT2 e mil, cento e vinte e sete procedimentos CAT3, totalizando mil, cento e noventa e cinco procedimentos no escopo analítico do projeto.

3.3.3 Derivação dos critérios estruturais

Os procedimentos classificados constituem a base para a derivação dos critérios estruturais utilizados na seleção de estabelecimentos. A partir das tabelas de relacionamento do SIGTAP, foram extraídas as habilitações, os serviços especializados e as ocupações profissionais associados a pelo menos um procedimento do escopo — resultando em oitenta e duas habilitações, sessenta e nove pares de serviço e classificação e cento e trinta e cinco ocupações profissionais. Os tipos de leito pediátrico relevantes — quinze ao total, distribuídos entre leitos pediátricos gerais, UTI pediátrica, UTI neonatal e unidades de cuidados intermediários pediátricos — foram definidos com base na codificação oficial do CNES e incorporados ao pipeline como critério estrutural complementar. Esses conjuntos de critérios, armazenados nos arquivos de configuração YAML do projeto, são utilizados nas três camadas de seleção de estabelecimentos descritas na seção seguinte.

3.4 Modelo de Duas Camadas de Seleção de Estabelecimentos

Com o conjunto de procedimentos classificados e os critérios estruturais derivados — habilitações e serviços especializados — o passo seguinte foi identificar, no CNES, quais estabelecimentos possuem capacidade de realizar esses procedimentos. O CNES não registra procedimentos diretamente por estabelecimento; a ponte entre as duas bases é feita por meio dos recursos estruturais que o SIGTAP exige para a execução de cada procedimento.

A análise das tabelas de relacionamento do SIGTAP revelou que nem todos os procedimentos do escopo analítico exigem habilitação formal: dos 1.195 procedimentos classificados, 284 não possuem nenhuma habilitação vinculada à tabela normativa. Esses procedimentos são definidos no SIGTAP apenas por meio de serviços especializados requeridos, o que torna a habilitação formal um critério insuficiente para capturar a integralidade do escopo. Uma abordagem que utilizasse

exclusivamente esse critério produziria subcobertura estrutural, excluindo estabelecimentos aptos a realizar procedimentos para os quais o próprio ordenamento normativo não exige credenciamento formal.

Para contornar essa limitação, adotou-se um modelo de duas camadas hierárquicas de seleção, aplicadas em ordem decrescente de rigor normativo, conforme descrito na Tabela 3.

Tabela 3 — Modelo de duas camadas de seleção de estabelecimentos

Camada	Mecanismo	Critério	Força evidencial
C1	Habilitação formal	Estabelecimento possui habilitação junto ao Ministério da Saúde compatível com ao menos um procedimento do escopo	Alta — exige credenciamento explícito e verificado pelo gestor do SUS
C2	Serviço declarado	Estabelecimento declara serviço especializado compatível com ao menos um procedimento do escopo, sem habilitação formal correspondente	Média — cobre procedimentos sem exigência de habilitação formal e estabelecimentos com registro de habilitação incompleto ou desatualizado

Fonte: elaboração própria.

A camada C1 é a de maior rigor normativo: a habilitação formal representa um credenciamento explícito, concedido pelo gestor do SUS mediante verificação dos requisitos estruturais do estabelecimento. A camada C2 atende a duas situações complementares: procedimentos que, por definição normativa no SIGTAP, não exigem habilitação — sendo identificáveis apenas pelo serviço especializado requerido —, e estabelecimentos que, embora satisfaçam os critérios para habilitação, não mantêm o registro formal atualizado no sistema.

As camadas não são mutuamente exclusivas: um estabelecimento pode satisfazer simultaneamente C1 e C2. Cada estabelecimento selecionado carrega o registro da combinação de camadas que determinou sua inclusão — com os valores possíveis C1, C2 e C1+C2 —, o que garante rastreabilidade metodológica completa e permite ao usuário do painel identificar, para qualquer estabelecimento, o tipo de evidência que fundamenta sua presença no escopo analítico.

A aplicação desse modelo sobre a competência de dezembro de 2024 do CNES resultou na seleção de 86.209 estabelecimentos, distribuídos entre 3.853 estabelecimentos no produto final com habilitação formal registrada (C1 ou C1+C2) e os demais incorporados pela camada de serviço declarado (C2), com cobertura de 5.546 dos 5.570 municípios brasileiros.

3.5 Limitações Metodológicas

Toda metodologia envolve escolhas que implicam concessões. Esta seção registra as limitações reconhecidas do modelo adotado, com o objetivo de orientar a interpretação dos resultados produzidos pelo PainelPedBR e subsidiar eventuais refinamentos futuros.

A primeira limitação diz respeito à natureza autodeclaratória do CNES. A base é alimentada pelos próprios estabelecimentos e gestores municipais e estaduais, o que torna sua qualidade dependente da regularidade e da precisão das atualizações realizadas em cada nível federativo. Informações sobre leitos, habilitações, serviços e vínculos profissionais podem estar desatualizadas, incompletas ou incorretas em relação à realidade operacional do estabelecimento no momento da análise (LIMA et al., 2009). Os dados devem ser interpretados como um retrato administrativo do período de referência adotado — dezembro de 2024 — e não como um inventário operacional verificado in loco.

A segunda limitação diz respeito à precisão geoespacial de parte dos estabelecimentos. O pipeline de geocodificação opera em três etapas em ordem decrescente de precisão: aproveitamento das coordenadas originais do CNES, geocodificação por endereço via Nominatim e, como recurso de contingência, atribuição do centroide municipal. Estabelecimentos que chegam ao terceiro nível têm sua localização fixada no ponto central do município, o que introduz imprecisão em qualquer análise que dependa de distâncias ou de proximidade entre unidades. Esses casos são sinalizados ao usuário na interface da aplicação, mas devem ser considerados na interpretação de análises espaciais baseadas em localização pontual.

A terceira limitação refere-se ao critério linguístico utilizado na identificação dos procedimentos CAT2. A ausência de restrição etária registrada em parte dos procedimentos de alta complexidade obriga a identificação com base nos termos presentes no nome oficial do procedimento no SIGTAP. Esse critério é sensível ao vocabulário adotado pelo Ministério da Saúde na nomenclatura dos procedimentos: procedimentos de alta complexidade com indicação pediátrica, mas cuja descrição não contenha nenhum dos termos de busca definidos, serão excluídos do escopo analítico. A abordagem adotada é deliberadamente conservadora, aceitando a possibilidade de falsos negativos em favor da reprodutibilidade e da rastreabilidade dos critérios, mas essa opção implica que o conjunto de procedimentos identificados pode não ser exaustivo.

A quarta limitação diz respeito à ausência de relação formal entre os equipamentos registrados no CNES e os procedimentos de alta complexidade pediátrica. Ao contrário das habilitações e dos serviços especializados, para os quais o SIGTAP mantém tabelas de relacionamento estruturadas com os procedimentos, não existe na base uma relação normativa que vincule tipos de equipamento a procedimentos ou a faixas etárias. Os dados de equipamentos do CNES não foram utilizados no modelo, pois sua incorporação sem ancoragem normativa introduziria subjetividade metodológica incompatível com os princípios do projeto.

A quinta limitação é inerente à classificação tripartite adotada e diz respeito à amplitude do escopo produzido pela categoria CAT3. Dos 1.195 procedimentos classificados, 1.127 — equivalente

a 94,3% do total — pertencem à categoria de faixa etária mista, sendo aplicáveis tanto a crianças quanto a adultos. Como consequência, o escopo analítico resultante apresenta duas fontes de heterogeneidade esperada: uma parcela dos estabelecimentos incluídos pode ter sua produção predominantemente voltada a pacientes adultos, sendo selecionada em razão de procedimentos que, embora cubram a faixa pediátrica, não são exclusivos dela; e uma parcela pode corresponder a estabelecimentos de atenção básica ou de menor complexidade que, por registrarem a oferta de ao menos um procedimento classificado no escopo, satisfazem o critério de seleção adotado sem que seu perfil assistencial dominante seja de alta complexidade. Essa é a principal concessão da abordagem adotada: ao priorizar a sensibilidade — garantindo que nenhum serviço com indicação pediátrica seja excluído —, aceita-se o risco de inclusão de estabelecimentos cujo perfil assistencial predominante não é de alta complexidade pediátrica. Os dados do painel devem ser interpretados à luz dessa característica do modelo, e a coluna de camada de evidência (C1, C2 ou C1+C2) disponível para cada estabelecimento oferece ao usuário um indicador do grau de rigor normativo que fundamenta cada inclusão.

A sexta limitação refere-se à heterogeneidade temporal das fontes utilizadas. O escopo analítico do projeto é definido a partir da competência de dezembro de 2024 do CNES e do SIGTAP, e qualquer alteração posterior na tabela de procedimentos ou nos registros cadastrais não é refletida automaticamente nos produtos de dados gerados — a atualização do escopo requer a reexecução do pipeline sobre uma nova competência. Adicionalmente, os dados populacionais provêm do Censo Demográfico de 2022 e os indicadores de PIB per capita municipal referem-se ao ano de 2023, edição mais recente disponível no IBGE à época da execução do pipeline. Análises que combinem dados de oferta assistencial com indicadores populacionais ou socioeconômicos envolvem, portanto, recortes temporais distintos, o que deve ser considerado na interpretação de comparações entre capacidade instalada e condições territoriais.

4 Requisitos

O levantamento de requisitos do PainelPedBR partiu das perguntas que a aplicação deve ser capaz de responder. Cada módulo foi concebido para responder a uma pergunta distinta sobre a oferta de alta complexidade pediátrica no Brasil, e os requisitos funcionais derivam diretamente dessas perguntas. Essa abordagem é coerente com a natureza do sistema: uma ferramenta de consulta e visualização cujo valor está em tornar acessível, de forma direta e sem exigência de conhecimento técnico especializado, informação antes fragmentada em múltiplas bases de dados — permitindo que famílias e cidadãos compreendam o que está disponível no seu território e, complementarmente, que gestores e pesquisadores identifiquem lacunas na distribuição dos recursos.

Os requisitos funcionais são organizados por módulo e precedidos pela pergunta analítica que orienta cada um. Os requisitos não-funcionais são organizados segundo os atributos de qualidade definidos pela norma ISO/IEC 25010 (ISO/IEC, 2011).

4.1 Requisitos Funcionais

Os requisitos funcionais são apresentados por módulo. Cada módulo é precedido pela pergunta analítica central que orienta seu escopo e justifica os requisitos que o compõem.

Módulo 1 — Panorama Nacional

Pergunta central: *como está distribuída a oferta de alta complexidade pediátrica no Brasil?*

- [RF01] A aplicação deve apresentar um mapa coroplético nacional com métricas de cobertura por Unidade Federativa — incluindo leitos pediátricos e de UTI por dez mil crianças —, permitindo a identificação visual imediata de disparidades regionais na capacidade instalada.
- [RF02] A aplicação deve exibir indicadores agregados nacionais — número total de estabelecimentos no escopo, pediatras por cem mil crianças, leitos de UTI pediátrica por cem mil crianças e municípios identificados como desertos pediátricos —, com possibilidade de refinamento por Unidade Federativa, camada de inclusão e tipo de estabelecimento.
- [RF03] A aplicação deve apresentar a distribuição dos estabelecimentos por região com indicadores normalizados por cem mil crianças e linha de referência nacional, e uma classificação dos vinte municípios com maior proporção de leitos de UTI pediátrica por população infantil.
- [RF04] A aplicação deve apresentar uma tabela resumo por Unidade Federativa com indicadores absolutos e normalizados — estabelecimentos, municípios cobertos, leitos pediátricos e indicadores de cobertura por população infantil —, com possibilidade de download em formato CSV.

Módulo 2 — Desertos e Cobertura

Pergunta central: *quais municípios estão sendo deixados para trás?*

- [RF05] A aplicação deve exibir indicadores de lacuna de cobertura — municípios sem pediatra, crianças residentes em deserto pediátrico total, distância mediana ao município com pediatra mais próximo, percentual da população infantil sem acesso a UTI pediátrica, percentual de municípios sem UTI pediátrica e crianças em cobertura crítica —, com possibilidade de filtro por região e Unidade Federativa.
- [RF06] A aplicação deve apresentar um mapa coroplético de municípios classificados por categoria de cobertura pediátrica, permitindo a identificação geográfica das lacunas territoriais de acesso.
- [RF07] A aplicação deve apresentar um gráfico de barras empilhadas com a distribuição percentual dos municípios por categoria de cobertura pediátrica — deserto total, cobertura

crítica, abaixo da média e cobertura adequada — desagregado por região geográfica, permitindo a comparação entre regiões da proporção de municípios em cada nível de cobertura.

- [RF08] A aplicação deve apresentar um gráfico de dispersão relacionando PIB per capita municipal à cobertura pediátrica, com demarcação de quadrantes para identificação de municípios em situação de dupla vulnerabilidade socioeconômica e assistencial.
- [RF09] A aplicação deve apresentar a classificação dos municípios com maior número absoluto de crianças residentes em deserto pediátrico.

Módulo 3 — Recursos Humanos

Pergunta central: *como estão distribuídos os profissionais pediátricos pelo território?*

- [RF10] A aplicação deve exibir indicadores do total de profissionais pediátricos únicos vinculados aos estabelecimentos do escopo e contagens desagregadas por especialidade — pediatras gerais, neonatologistas, intensivistas pediátricos, oncologistas pediátricos e cirurgiões pediátricos —, com possibilidade de refinamento por região, Unidade Federativa e município.
- [RF11] A aplicação deve apresentar um mapa coroplético de profissionais pediátricos por cem mil crianças desagregado por município, com filtro de especialidade independente, permitindo a identificação geográfica de lacunas na oferta de cada tipo de profissional.
- [RF12] A aplicação deve apresentar a classificação das dez Unidades Federativas com maior número de profissionais pediátricos, filtrável por especialidade, e um gráfico de barras empilhadas com a composição percentual de especialidades por região geográfica.
- [RF13] A aplicação deve disponibilizar uma tabela de profissionais por município e especialidade, com a taxa por cem mil crianças, buscável e ordenável, filtrada pelos critérios geográficos selecionados no painel.

Módulo 4 — Procedimentos

Pergunta central: *quais procedimentos pediátricos de alta complexidade existem e onde estão disponíveis?*

- [RF14] A aplicação deve exibir indicadores da oferta de procedimentos no recorte selecionado — total de procedimentos únicos, estabelecimentos ofertantes, número de Unidades Federativas com oferta de procedimentos CAT1 e categoria de classificação dominante —, com possibilidade de filtro por região, Unidade Federativa e categoria de procedimento (CAT1, CAT2 ou CAT3).

- [RF15] A aplicação deve apresentar um treemap hierárquico dos procedimentos classificados, organizado por grupo, subgrupo e procedimento SIGTAP, com tamanho proporcional ao número de estabelecimentos ofertantes, permitindo a navegação visual pela estrutura da tabela de procedimentos de acordo com os filtros ativos.
- [RF16] A aplicação deve disponibilizar um seletor de procedimento com busca textual que, ao receber uma seleção, exiba as informações descritivas do procedimento e um mapa pontual dos estabelecimentos que o ofertam, permitindo a localização geográfica da oferta para procedimentos individuais.
- [RF17] A aplicação deve apresentar um mapa de calor dos vinte e cinco procedimentos mais ofertados por Unidade Federativa, com cor proporcional ao número de estabelecimentos ofertantes, atualizado de acordo com os filtros ativos no painel.

Módulo 5 — Mapa de Estabelecimentos

Pergunta central: *o que existe em um lugar específico e como ele se compara a outros territórios?*

- [RF18] O módulo deve operar em dois modos alternáveis — Explorar e Comparar — compartilhando um painel de filtros com nove critérios: região, Unidade Federativa, município, tipo de estabelecimento, tipo de gestão, presença de UTI pediátrica, presença de UTI neonatal, habilitação formal e camada de inclusão metodológica.
- [RF19] No modo Explorar, a aplicação deve exibir seis indicadores do recorte filtrado — total de estabelecimentos, leitos pediátricos gerais, leitos de UTI pediátrica e leitos de UTI neonatal com respectivas taxas por cem mil crianças, percentual de municípios cobertos e percentual de estabelecimentos com habilitação formal (C1).
- [RF20] No modo Explorar, a aplicação deve apresentar uma tabela de estabelecimentos com download em CSV e um mapa de localização com agrupamento de pontos; o clique em um ponto do mapa ou em uma linha da tabela deve acionar a navegação ao Módulo 6 com o estabelecimento selecionado.
- [RF21] No modo Comparar, a aplicação deve permitir a seleção de dois a quatro territórios — por Unidade Federativa ou por município — e apresentar um gráfico de indicadores normalizados e uma tabela comparativa lado a lado para os territórios selecionados.

Módulo 6 — Perfil do Estabelecimento

Pergunta central: *quais são os recursos disponíveis em um estabelecimento específico e por qual critério ele integra o escopo analítico?*

- [RF22] A aplicação deve permitir a seleção de um estabelecimento por meio de busca textual livre por nome, código CNES ou município, ou por navegação a partir do Módulo 5; ao

selecionar um estabelecimento, deve exibir badges indicando a presença de UTI Pediátrica, UTI Neonatal, habilitações de Oncologia e Transplante, e a camada de inclusão metodológica (C1, C2 ou C1+C2).

- [RF23] A aplicação deve apresentar os dados cadastrais do estabelecimento — nome fantasia, razão social quando distinta, código CNES, município e Unidade Federativa, tipo de unidade, tipo de gestão e natureza jurídica — e uma seção de rastreabilidade informando a camada de inclusão metodológica e o método de geocodificação utilizado.
- [RF24] A aplicação deve exibir um mapa de localização do estabelecimento selecionado com a demarcação dos demais estabelecimentos do escopo analítico situados no raio de cem quilômetros, permitindo a contextualização geográfica da unidade em relação à oferta regional de alta complexidade pediátrica.
- [RF25] A aplicação deve disponibilizar um painel de recursos organizado em quatro abas: Leitos, com gráfico de barras dos tipos de leito pediátrico declarados; Habilitações, com tabela de habilitações vigentes por código, nome, categoria pediátrica e grupo; Profissionais, com tabela de vínculos profissionais do estabelecimento; e Procedimentos, com contagem e tabela dos procedimentos pediátricos de alta complexidade disponíveis com seus códigos SIGTAP e categorias de classificação.

4.2 Requisitos Não-Funcionais

Desempenho

- [RNF01] Os filtros interativos da aplicação devem responder em até dois segundos sob condições normais de uso, garantindo experiência de navegação fluida sem recarga de página.
- [RNF02] A aplicação deve carregar os produtos de dados em memória na inicialização, eliminando a necessidade de transformações em tempo de execução e assegurando tempos de resposta estáveis independentemente do volume de dados subjacente.

Usabilidade

- [RNF03] A interface deve ser acessível a qualquer usuário, independentemente de formação técnica em análise de dados ou programação — incluindo famílias e cidadãos que buscam informações sobre serviços de saúde —, sendo igualmente adequada a gestores, pesquisadores e formuladores de políticas.
- [RNF04] A navegação entre os módulos deve seguir uma lógica progressiva e intuitiva — do panorama nacional ao perfil do estabelecimento —, de forma que o fluxo de uso não exija instruções prévias.

- [RNF05] Os indicadores apresentados devem ser acompanhados de descrições contextuais que expliquem sua interpretação, evitando ambiguidade na leitura dos resultados.

Portabilidade

- [RNF06] A aplicação deve ser acessível por meio de navegador web, sem necessidade de instalação de software adicional pelo usuário.

Manutenibilidade

- [RNF07] Os critérios metodológicos que definem o escopo analítico — classificações, limiares etários e listas de códigos relevantes — devem estar centralizados em arquivos de configuração YAML, separados do código de transformação, de forma que alterações metodológicas possam ser incorporadas sem modificação da lógica de implementação.
- [RNF08] A arquitetura da aplicação deve seguir o padrão modular do framework Shiny, com cada módulo encapsulando sua interface e lógica de servidor de forma independente, facilitando a manutenção e a evolução isolada de cada funcionalidade.
 - **Critério de verificação:** o isolamento é garantido estruturalmente pelo mecanismo `NS()` + `moduleServer()` do *Shiny*, que impede o compartilhamento de IDs e estado de sessão entre módulos, e pelos contratos explícitos de comunicação via parâmetros reativos em *server.R*.

Reprodutibilidade

- [RNF09] O ambiente de execução do projeto deve ser gerenciado pelo *renv*, com registro das versões exatas de todos os pacotes utilizados, garantindo que o pipeline de dados e a aplicação sejam executáveis em qualquer máquina com o mesmo resultado.
 - Critério de verificação: o isolamento é garantido estruturalmente pelo mecanismo `NS()` + `moduleServer()` do *Shiny*, que impede o compartilhamento de IDs e estado de sessão entre módulos, e pelos contratos explícitos de comunicação via parâmetros reativos em *server.R*.
- [RNF10] O código-fonte e os arquivos de configuração do projeto devem ser disponibilizados em repositório público no GitHub, permitindo a reprodução do pipeline de dados e da aplicação e a auditoria dos critérios metodológicos implementados.

5 Design e Implementação

As decisões de design do PainelPedBR foram orientadas por três princípios que permeiam toda a arquitetura do sistema. O primeiro é a separação explícita entre as camadas de dados e de interface:

o pipeline de engenharia de dados e a aplicação Shiny são componentes independentes que se comunicam exclusivamente por meio de produtos de dados pré-processados, eliminando qualquer acoplamento em tempo de execução entre as duas camadas. O segundo é a desnormalização orientada ao desempenho: os produtos de dados são gerados na fase de carga do pipeline já na estrutura que a aplicação necessita, eliminando transformações em tempo de execução e garantindo tempos de resposta adequados à experiência interativa da aplicação (conforme requisito RNF-02). O terceiro é a auditabilidade metodológica: todos os critérios que definem o escopo analítico do projeto são externalizados em arquivos de configuração versionados, separados do código de implementação, de forma que as decisões metodológicas sejam rastreáveis e alteráveis sem modificação do código de transformação.

Esta seção documenta as decisões de design que resultam desses princípios, organizadas em cinco perspectivas: a arquitetura geral do sistema, o pipeline ETL, a estrutura de dados, o diagrama de componentes e o design da interface da aplicação.

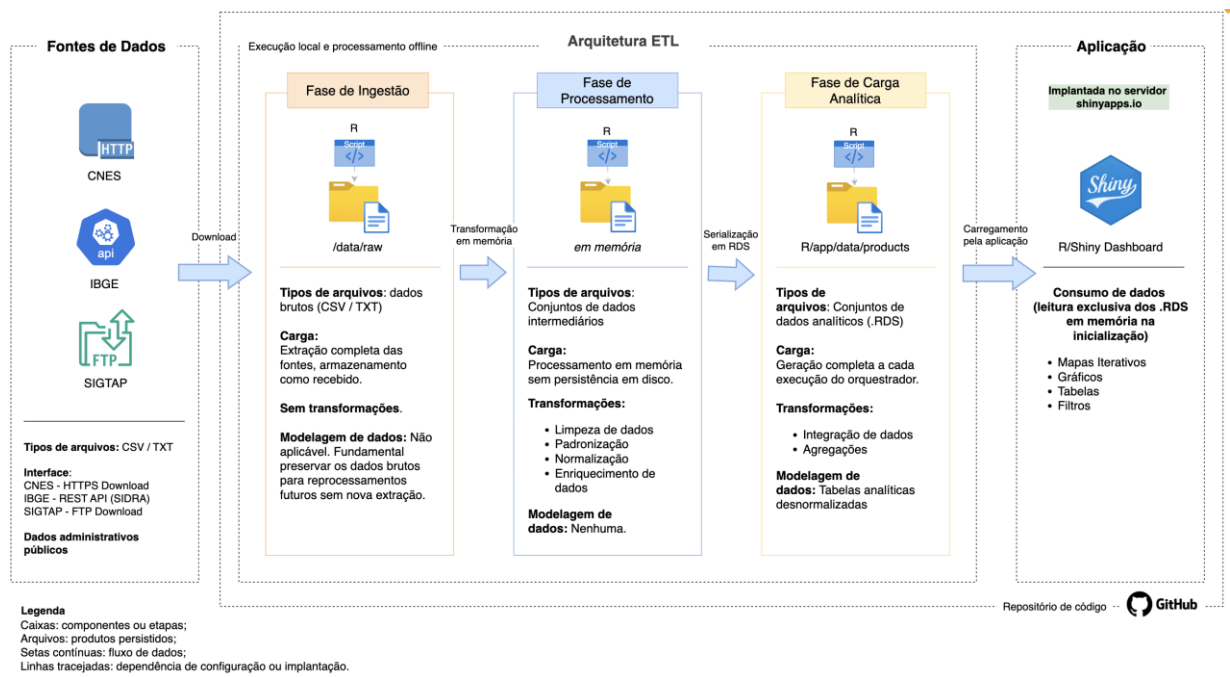
5.1 Arquitetura Geral da Solução

O PainelPedBR é composto por dois componentes com responsabilidades e ciclos de execução distintos: o pipeline de engenharia de dados e a aplicação Shiny. O pipeline é responsável pela extração, transformação e geração dos produtos analíticos a partir das fontes de dados oficiais. A aplicação é responsável pela apresentação interativa desses produtos ao usuário final. Os dois componentes comunicam-se exclusivamente por meio de arquivos de dados serializados no formato .rds, gerados pelo pipeline e consumidos pela aplicação na inicialização — sem dependência em tempo de execução entre os dois.

Essa separação é uma decisão arquitetural central do projeto. Ela implica que a atualização dos dados — pela reexecução do pipeline sobre uma nova competência das bases — é independente do ciclo de vida da aplicação, e que a aplicação opera sempre sobre dados pré-processados e estáveis, sem transformações em tempo de execução. Os critérios metodológicos são centralizados em arquivos de configuração YAML utilizados pelo pipeline, garantindo que as decisões analíticas estejam consolidadas em um único ponto e reflitam-se nos produtos consumidos pela aplicação.

A Figura 1 apresenta a arquitetura de alto nível do sistema, ilustrando o fluxo de dados desde as fontes oficiais até o consumo pela aplicação Shiny, com destaque para as três fases que compõem o pipeline e as características de cada uma.

Figura 1 — Arquitetura de Alto Nível da Solução



Fonte: elaboração própria.

Essa separação entre componentes também se reflete nos ambientes de execução: o pipeline é executado localmente no ambiente de desenvolvimento, enquanto a aplicação é hospedada na plataforma shinyapps.io, acessível publicamente por meio de navegador web sem necessidade de infraestrutura própria.

5.2 Stack Tecnológico

A escolha tecnológica central do PainelPedBR é o R, linguagem de programação estatística de código aberto amplamente adotada na pesquisa em saúde pública e epidemiologia (R CORE TEAM, 2024). A opção por R como linguagem única para todas as etapas do projeto — da extração e tratamento dos dados à construção da interface interativa — foi motivada por três fatores. Primeiro, R dispõe de um ecossistema consolidado de pacotes para as operações requeridas pelo projeto, eliminando a necessidade de múltiplas linguagens ou ferramentas externas. Segundo, a concentração de todo o fluxo de trabalho em uma única linguagem simplifica a manutenção do código e garante a reprodutibilidade integral do projeto. Terceiro, a integração nativa com o framework Shiny permite que o mesmo ambiente utilizado para o tratamento de dados seja empregado na construção da interface de visualização (CHANG et al., 2024).

Os pacotes utilizados no projeto estão organizados na Tabela 4, agrupados por camada funcional, com indicação da versão empregada e da finalidade de cada um no contexto do PainelPedBR.

Tabela 4 — Pacotes R utilizados no PainelPedBR

Pacote	Versão	Camada	Finalidade
<i>shiny</i>	1.13.0	Aplicação	Framework para construção da interface interativa
<i>shinyWidgets</i>	0.9.1	Aplicação	Componentes de entrada avançados (filtros com busca, botões de ação)
<i>shinybusy</i>	0.3.3	Aplicação	Indicador visual de carregamento durante processamento reativo
<i>dplyr</i>	1.2.0	ETL	Manipulação e transformação de dados tabulares
<i>readr</i>	2.2.0	ETL	Leitura de arquivos de texto com diferentes encodings, delimitadores e largura fixa
<i>stringr</i>	1.6.0	ETL	Padronização e limpeza de campos textuais
<i>stringi</i>	1.8.7	ETL	Operações avançadas de codificação e normalização de texto
<i>purrr</i>	1.2.1	ETL	Iteração funcional sobre listas e exportação de artefatos de auditoria
<i>tidyr</i>	1.3.2	ETL	Reestruturação de dados e tratamento de valores ausentes
<i>janitor</i>	2.2.1	ETL	Normalização de nomes de colunas para snake_case
<i>curl</i>	7.0.0	ETL	Acesso às bases do DATASUS com configuração de parâmetros de conexão
<i>yaml</i>	2.3.12	ETL	Leitura dos arquivos de configuração metodológica
<i>logger</i>	0.4.1	ETL	Registro estruturado de eventos do pipeline em arquivo e console
<i>pointblank</i>	0.12.3	ETL	Validação automatizada da integridade estrutural dos produtos de dados
<i>sf</i>	1.1-0	Visualização	Manipulação e renderização de dados geoespaciais
<i>leaflet</i>	2.2.3	Visualização	Mapas interativos de pontos e coropléticos
<i>plotly</i>	4.12.0	Visualização	Construção de gráficos interativos com suporte a eventos
<i>reactable</i>	0.4.5	Visualização	Tabelas analíticas com filtragem, ordenação e destaque condicional
<i>scales</i>	1.4.0	Visualização	Formatação de valores numéricos para exibição
<i>htmlwidgets</i>	1.6.4	Visualização	Interface de binding entre bibliotecas JavaScript de visualização e o ambiente R/Shiny
<i>renv</i>	1.1.8	Reprodutibilidade	Gestão de dependências e reprodutibilidade do ambiente

Fonte: elaboração própria.

A gestão de dependências é realizada pelo *renv*, que registra as versões exatas de todos os pacotes em um arquivo de bloqueio (*lockfile*), garantindo que o ambiente de execução seja reproduzível em qualquer máquina (USHEY; WICKHAM, 2024). O versionamento do código e dos arquivos de configuração é gerenciado pelo Git, com repositório público hospedado no GitHub. A publicação da aplicação para acesso externo é realizada na plataforma *shinyapps.io*, que oferece hospedagem gerenciada para aplicações *R/Shiny* sem necessidade de infraestrutura própria.

5.3 Pipeline ETL

O pipeline de engenharia de dados do PainelPedBR é implementado em R e organizado em três fases sequenciais — Extração, Transformação e Carga — com nove scripts modulares, cada um com responsabilidade lógica bem delimitada. Um orquestrador principal (`orchestrator_etl.R`) encadeia as três fases e controla o fluxo de execução do pipeline como um todo, permitindo que cada fase seja reexecutada de forma isolada quando necessário. A Tabela 5 apresenta a organização dos scripts por fase e sua descrição funcional.

Tabela 5 — Organização dos scripts do pipeline ETL

Fase	Script	Responsabilidade
Extração	<code>extract_cnes.R</code>	<i>Download</i> e armazenamento dos dados brutos do CNES via HTTPS
Extração	<code>extract_sigtap.R</code>	<i>Download</i> e <i>parsing</i> dos arquivos posicionais do SIGTAP via FTP
Extração	<code>extract_ibge.R</code>	Coleta dos dados populacionais via API SIDRA e da malha municipal georreferenciada
Transformação	<code>transform_sigtap.R</code>	Classificação tripartite dos procedimentos e derivação dos critérios estruturais
Transformação	<code>transform_cnes.R</code>	Limpeza, padronização e seleção de estabelecimentos nas duas camadas
Transformação	<code>transform_ibge.R</code>	Normalização dos dados populacionais e preparação das geometrias municipais
Transformação	<code>geocode.R</code>	Geocodificação por endereço via Nominatim com cache persistente e centroide municipal como contingência
Carga	<code>build_data_products.R</code>	Integração das entidades processadas das três fontes e geração dos produtos analíticos finais
Carga	<code>save_data_products.R</code>	Validação e serialização dos produtos em formato <code>.rds</code> para consumo pela aplicação

Fonte: elaboração própria.

5.3.1 Fase de Extração

A fase de extração é responsável pela coleta dos dados brutos a partir das três fontes oficiais e pelo seu armazenamento no diretório `data/raw/` sem nenhuma modificação. Cada script registra em log a data e a hora de extração, o tamanho dos arquivos obtidos e eventuais erros de conexão, garantindo rastreabilidade do ciclo de atualização dos dados. Os parâmetros de acesso a cada fonte — URLs, endereço do servidor FTP, identificadores da API SIDRA e competência de referência — são lidos dos arquivos de configuração YAML, sem nenhum valor fixo no código dos scripts. O pipeline implementa verificação de integridade dos arquivos já presentes em disco: nas reexecuções, o procedimento de download é suprimido quando o arquivo correspondente já existe no diretório de destino e seu tamanho em bytes é compatível com o esperado, o que reduz o tempo de execução nas competências subsequentes à coleta inicial. Esta fase encontra-se integralmente implementada e validada para a competência de dezembro de 2024.

5.3.2 Fase de Transformação

A fase de transformação aplica a lógica metodológica descrita na seção 3 sobre os dados brutos e produz entidades relacionais limpas e padronizadas. As transformações ocorrem inteiramente em memória, sem persistência de dados intermediários em disco — decisão justificada pela ausência de consumidores externos das estruturas intermediárias e pelo tempo de execução reduzido das transformações (KIMBALL; CASERTA, 2004).

5.3.2.1 `transform_sigtap.R`

Este script aplica a classificação tripartite de procedimentos, operação central da metodologia do projeto. A partir das tabelas SIGTAP, os procedimentos são classificados em três categorias mutuamente exclusivas: CAT1, composta por procedimentos exclusivamente pediátricos definidos por teto etário de até 239 meses no SIGTAP; CAT2, composta por procedimentos de faixa etária universal com nomenclatura explicitamente pediátrica; e CAT3, composta por procedimentos de faixa compartilhada, que abrangem populações pediátrica e adulta simultaneamente. A partir dessa classificação, o script deriva as listas de habilitações, serviços e ocupações profissionais que compõem o escopo analítico do projeto. Os critérios resultantes são exportados para `config/classificacoes_derivadas.yaml`, que opera como artefato de rastreabilidade — permitindo auditar, a qualquer momento, quais códigos fundamentam as decisões de seleção de estabelecimentos nas etapas subsequentes.

5.3.2.2 `transform_cnes.R`

Este script aplica as duas camadas de seleção de estabelecimentos sobre a base do CNES. A Camada C1 seleciona estabelecimentos por habilitação formal registrada no CNES, critério de maior rigor normativo. A Camada C2 seleciona por declaração de serviço, critério complementar

fundamentado nos serviços vinculados aos procedimentos do escopo analítico via SIGTAP. Além da seleção, o script executa duas operações de qualidade sobre os dados: a validação de coordenadas geográficas — identificando registros com latitude e longitude ausentes, zeradas ou fora dos limites territoriais do Brasil — e a padronização de campos textuais, que inclui normalização de encoding, remoção de espaços redundantes e uniformização de capitalização.

5.3.2.3 transform_ibge.R

Este script estrutura os dados populacionais por município e faixa etária a partir do Censo Demográfico 2022 e prepara as geometrias municipais para uso nos mapas da aplicação. A estruturação por faixa etária é necessária para o cálculo de indicadores normalizados por população pediátrica — denominador distinto da população total municipal — que compõem os produtos analíticos dos módulos de visualização. As geometrias são processadas e persistidas em formato compatível com o pacote sf, otimizando o tempo de carregamento na inicialização da aplicação Shiny.

5.3.2.4 geocode.R

Este script trata os estabelecimentos cujas coordenadas geográficas foram sinalizadas como ausentes ou inválidas pelo transform_cnes.R. A geocodificação segue uma estratégia em três etapas progressivas. Na primeira etapa, o script consulta o cache persistente em data/cache/, reutilizando coordenadas já resolvidas em execuções anteriores sem nenhuma chamada externa. Na segunda etapa, para os endereços não encontrados em cache, realiza chamadas à API Nominatim com limitação de taxa de 1.100 ms por requisição, respeitando os termos de uso do serviço e prevenindo bloqueios por excesso de requisições. Na terceira etapa, para os endereços não resolvidos pela API, aplica o centroide do município como coordenada aproximada, garantindo que nenhum estabelecimento fique sem representação geoespacial nos produtos finais. Essa abordagem incremental assegura que nenhuma coordenada já resolvida seja reprocessada em execuções subsequentes, propriedade alinhada ao princípio de idempotência que orienta o design de todo o pipeline.

5.3.3 Fase de Carga

A fase de carga integra as entidades produzidas pela fase de transformação em um modelo unificado e gera os produtos de dados consumidos pela aplicação Shiny. Os produtos são gerados com estrutura desnormalizada — campos de múltiplas entidades são consolidados em tabelas orientadas ao consumo analítico, eliminando a necessidade de junções em tempo de execução. Uma decisão transversal desta fase é a pré-computação de agregações: produtos como os resumos de recursos humanos e de oferta de procedimentos são calculados no ETL e persistidos como tabelas prontas para consumo, evitando que o Shiny precise carregar centenas de megabytes de vínculos individuais em tempo de execução. Antes da serialização, o produto (dp_estabelecimentos) é submetido a validação automatizada via pacote pointblank, que verifica a integridade estrutural e a consistência dos campos

críticos. A Tabela 6 apresenta os quinze produtos de dados gerados e seus volumes na competência de dezembro de 2024.

Tabela 6 — Produtos de dados gerados (competência dezembro de 2024)

Produto	Registros	Descrição
dp_estabelecimentos.rds	86.209	Um registro por estabelecimento no escopo analítico, com atributos de infraestrutura (leitos, habilitações, profissionais), localização georreferenciada, tipo de unidade, gestão e camada de inclusão metodológica
dp_leitos_por_estab.rds	48.315	Leitos do escopo pediátrico por estabelecimento e tipo, com quantidades totais e SUS e indicadores de categoria (pediátrico geral, UTI pediátrica, UTI neonatal, intermediário pediátrico)
dp_habilitacoes_por_estab.rds	10.539	Habilitações formais do escopo por estabelecimento, com código, nome, grupo e categoria pediátrica de referência
dp_profissionais_unicos.rds	2.218.104	Profissionais pediátricos distintos vinculados aos estabelecimentos do escopo, com carga horária total agregada e número de estabelecimentos vinculados – base para KPIs nacionais
dp_profissionais_vinculos.rds	3.059.188	Vínculos individuais profissional–estabelecimento com carga horária ambulatorial e hospitalar e indicadores de especialidade, base para distribuição territorial
dp_rh_por_municipio.rds	15.149	Profissionais pediátricos pré-agregados por município e especialidade, com taxa por 100 mil crianças. Substitui o carregamento de 350 MB de vínculos individuais em tempo de execução
dp_rh_por_uf.rds	160	Profissionais pediátricos distintos pré-agregados por Unidade Federativa e

		especialidade, sem dupla contagem de profissionais vinculados a estabelecimentos em municípios diferentes
dp_municipios_resumo.rds	5.570	Indicadores agregados para todos os municípios brasileiros (incluindo os sem cobertura identificada), com dados populacionais do Censo 2022 e indicador de pediatras por 100 mil crianças
dp_procedimentos_por_estab.rds	915.132	Procedimentos pediátricos de alta complexidade por estabelecimento, com categoria de classificação (CAT1/CAT2/CAT3), hierarquia SIGTAP e indicador de habilitação formal
dp_oferta_resumo.rds	28.929	Oferta pré-agregada por procedimento e Unidade Federativa, com número de estabelecimentos ofertantes. Alimenta treemap, mapa de calor, seletor de procedimentos e KPIs do Módulo de Procedimentos
dp_oferta_mapa.rds	2.545.808	Oferta por procedimento e estabelecimento, limitada a estabelecimentos com coordenadas georreferenciadas, para renderização do mapa pontual de oferta por procedimento individual
dp_vinculos_por_estab.rds	496.490	Vínculos profissionais agrupados por estabelecimento, CBO e tipo de vínculo, com contagem de profissionais e carga horária agregada (versão compacta para o Módulo Perfil do Estabelecimento)
dp_procs_slim.rds	2.545.808	Procedimentos com colunas mínimas necessárias para exibição e tipos compactos (factors) – versão otimizada para consumo pelo Módulo Perfil do Estabelecimento

dp_constantes_startup.rds	lista — 3 elementos	Constantes pré-computadas na inicialização: total de médicos pediatras, distância mediana nacional ao município com pediatra mais próximo e tabela de lookup com distâncias medianas por região e Unidade Federativa
malha_municipios_2022.rds	5.570	Malha municipal georreferenciada do IBGE 2022, simplificada à tolerância de 1 km para otimização de renderização, com polígonos em WGS84

Fonte: elaboração própria.

5.4 Estrutura de Dados

O PainelPedBR não utiliza banco de dados relacional. A camada de persistência do sistema é composta pelos quinze produtos de dados no formato binário nativo do R (.rds), apresentados na Seção 5.3.3, armazenados no diretório R/app/data/products/ e produzidos exclusivamente pela fase de carga do pipeline ETL. Optou-se por essa abordagem porque a aplicação é somente-leitura — nenhum usuário escreve dados — e, portanto, a complexidade de manutenção de um servidor de banco de dados relacional não se justifica. O formato .rds preserva os tipos de dados originais do R e dispensa qualquer camada de mapeamento objeto-relacional ou driver de conexão.

Definida a camada de persistência, a organização interna dos produtos seguiu um modelo desnormalizado, orientado ao padrão de acesso de cada módulo da aplicação. Em vez de um esquema normalizado com tabelas relacionadas por chaves estrangeiras, cada produto de dados foi pré-processado para responder a uma consulta específica (KIMBALL; ROSS, 2013). Esse princípio é especialmente relevante em aplicações Shiny hospedadas: como o servidor processa requisições de múltiplos usuários simultaneamente, operações computacionalmente custosas — junções entre grandes tabelas, agregações por múltiplas dimensões, cálculos de distância geográfica — realizadas em tempo de execução deteriorariam a experiência de todos os usuários concorrentes. A solução adotada foi deslocar essas operações para o ETL, executando-as uma única vez na fase de carga e persistindo os resultados prontos para consumo direto. Produtos como dp_rh_por_municipio e dp_rh_por_uf substituem o carregamento de 3 milhões de vínculos individuais por tabelas pré-agregadas por especialidade; dp_procs_slim e dp_oferta_resumo substituem o carregamento de 915 mil registros de procedimentos por estruturas otimizadas para cada padrão de acesso específico; dp_constantes_startup disponibiliza distâncias medianas entre municípios pré-calculadas, eliminando operações geoespaciais de alta complexidade computacional em runtime. O campo co_cnes — código nacional do estabelecimento no CNES — funciona como chave de ligação entre os produtos, permitindo que os módulos relacionem informações sem nenhuma junção em tempo de execução.

A estratégia de carregamento em memória é consequência direta da separação dos produtos em grupos funcionais distintos. Oito produtos são carregados na inicialização da sessão Shiny, em `global.R`: `dp_estabelecimentos`, `dp_municipios_resumo`, `dp_rh_por_municipio`, `dp_rh_por_uf`, `dp_oferta_resumo`, `dp_vinculos_por_estab`, `dp_constantes_startup` (desserializado como constantes individuais de domínio) e a malha municipal georreferenciada. Esses produtos alimentam os filtros globais, os indicadores dos módulos exibidos por padrão e as constantes de domínio da aplicação. Quatro produtos são carregados sob demanda pelos módulos temáticos, mediante acesso ao dicionário de caminhos *PATHS* definido em `global.R`: `dp_leitos_por_estab`, `dp_habilitacoes_por_estab`, `dp_oferta_mapa` e `dp_procs_slim`. Os três produtos restantes — `dp_profissionais_vinculos`, `dp_profissionais_unicos` e `dp_procedimentos_por_estab` — são persistidos pelo ETL como produtos intermediários auditáveis, mas não são carregados diretamente pela aplicação: suas informações chegam ao usuário por meio dos produtos pré-agregados e otimizados que os substituem em runtime. Essa separação impede que o tempo de inicialização da aplicação cresça proporcionalmente ao volume total de dados persistidos.

5.5 Arquitetura de Componentes

Para representar a organização interna do PainelPedBR, adotou-se o diagrama de componentes da Unified Modeling Language (UML), que descreve a estrutura de um sistema de software em termos de unidades funcionais com interfaces bem definidas e dependências explícitas (FOWLER, 2003). Essa notação é mais adequada ao projeto do que diagramas de classe, visto que o R não é uma linguagem orientada a objetos no sentido estrito: a estrutura do sistema é melhor descrita por módulos funcionais e seus contratos de dados do que por hierarquias de herança.

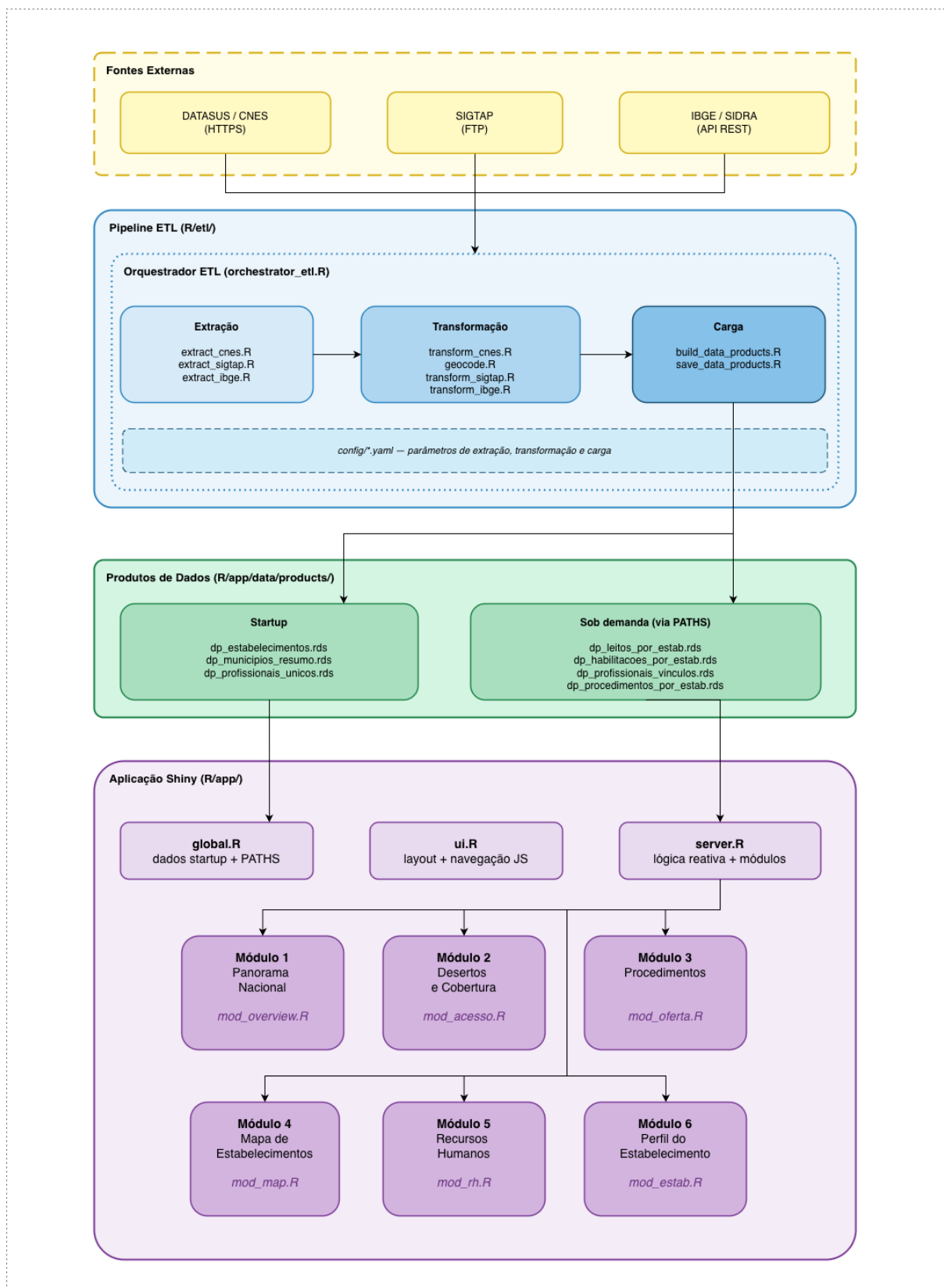
O sistema organiza-se em três camadas com responsabilidades distintas. A primeira é o pipeline ETL, composto pelos scripts de extração, transformação e carga localizados em `R/etl/`. A segunda é a camada de dados persistidos, representada pelos quinze arquivos `.rds` em `R/app/data/products/`. A terceira é a aplicação Shiny, em `R/app/`, que consome os produtos de dados e os apresenta ao usuário. Os arquivos `.rds` funcionam como a interface formal entre o pipeline e a aplicação: nenhuma das duas camadas depende da implementação interna da outra, o que permite executar o ETL de forma independente e substituir versões dos dados sem reinicializar o código da aplicação (BASS; CLEMENTS; KAZMAN, 2012).

Dentro da aplicação Shiny, três arquivos principais exercem papéis complementares. O `global.R` é responsável pelo carregamento dos dados de inicialização e pela definição das constantes compartilhadas entre os módulos, incluindo o dicionário de caminhos *PATHS*. O `ui.R` define a estrutura visual do painel — barra lateral, barra superior e as áreas de conteúdo —, acessando os dados de inicialização exclusivamente para a exibição de metadados estáticos da sessão. O `server.R` instancia os seis módulos temáticos e gerencia o estado reativo compartilhado entre eles: filtros geográficos (UF, região e município), módulo ativo e a seleção de estabelecimento, transmitida do módulo Mapa de Estabelecimentos para o módulo Perfil do Estabelecimento. Cada módulo encapsula sua própria função

de interface (`_ui`) e função de lógica reativa (`_server`), carregando sob demanda os produtos de dados de que necessita por meio do dicionário `PATHS`.

A Figura 2 apresenta o diagrama de componentes do sistema, com as relações de dependência entre as três camadas e os componentes internos da aplicação Shiny.

Figura 2 — Diagrama de componentes do PainelPedBR



Fonte: elaboração própria.

5.6 Design da Interface

Embora o Shiny elimine a necessidade de implementar em JavaScript a lógica reativa e a comunicação cliente-servidor da aplicação, extensões pontuais em CSS e JavaScript foram utilizadas para comportamentos de interface que ultrapassam os padrões nativos do framework — o layout flexbox customizado, a responsividade mobile e as funções de navegação entre módulos. Essas adições operam exclusivamente na camada de apresentação e não substituem o papel central do Shiny na arquitetura da aplicação.

5.6.1 Estrutura de *Layout*

O layout do PainelPedBR foi estruturado em duas regiões principais: uma barra lateral (sidebar) fixa à esquerda e uma área de conteúdo principal à direita. Essa divisão foi implementada por meio do modelo de caixas flexíveis (flexbox) do CSS (Cascading Style Sheets), sem dependência das classes de grade do framework Bootstrap (MEYER; WEYL, 2017). A decisão de não utilizar componentes de layout de pacotes externos foi tomada para garantir controle total sobre o comportamento visual em diferentes larguras de tela.

5.6.2 Organização dos Módulos

A interface dos módulos analíticos foi projetada a partir de um padrão estrutural comum: painel de filtros geográficos no topo, grade de indicadores-chave (KPIs) e visualização analítica principal. Essa padronização reduz a curva de aprendizado do usuário — o padrão de interação aprendido em um módulo aplica-se diretamente aos demais — e simplifica o desenvolvimento incremental, pois novos módulos podem ser construídos sobre a mesma estrutura sem alterações na arquitetura de navegação.

A visualização principal varia conforme o recorte analítico de cada módulo: mapas interativos nos módulos com dimensão territorial, gráficos analíticos nos módulos de oferta de procedimentos e distribuição de recursos humanos, e tabela navegável no módulo de exploração de estabelecimentos. O Módulo 6 — Perfil do Estabelecimento — constitui exceção ao padrão: ao invés de filtros geográficos, é acionado por busca textual ou por seleção direta no mapa, exibindo uma ficha consolidada com todos os dados disponíveis para o estabelecimento selecionado.

O Módulo 1 — Panorama Nacional — é a visão padrão exibida ao abrir o painel e o módulo de maior densidade informacional. Sua interface organiza-se em quatro camadas verticais: painel de filtros, grade de seis KPIs nacionais, mapa interativo com alternância entre visualização pontual e coroplética, e gráficos complementares com distribuição regional, funil de camadas metodológicas e classificação de municípios por leitos UTI pediátricos.

5.6.3 Navegação entre Módulos

A navegação entre os módulos do painel foi implementada sem recarga de página. Ao selecionar um módulo na sidebar, a função JavaScript `showModule(id)` oculta todos os contêineres de módulo e exibe apenas o correspondente à seleção, alterando também a classe CSS `active` no item de

navegação. A função notifica o servidor Shiny sobre o módulo ativo por meio de `Shiny.setInputValue()`, que registra o identificador do módulo como entrada reativa — permitindo que o servidor atualize o título exibido na topbar sem requisição HTTP adicional.

A navegação pode ser acionada também pelo servidor, por meio de um handler de mensagens customizado (`addCustomMessageHandler`). Esse mecanismo é utilizado, por exemplo, quando o usuário clica em um estabelecimento no Módulo 5 — Mapa de Estabelecimentos: o servidor envia uma mensagem ao cliente instruindo-o a exibir o Módulo 6 — Perfil do Estabelecimento — com o estabelecimento selecionado já carregado.

Esse padrão de navegação, comum em aplicações de página única (Single Page Application), evita o custo de recarregar a página completa a cada troca de módulo. Os módulos já instanciados permanecem no DOM (Document Object Model) em estado oculto, preservando o estado dos filtros e dos gráficos renderizados durante a sessão.

5.6.4 Responsividade

A interface foi projetada para adaptar-se a diferentes larguras de tela por meio de media queries CSS (MARCOTTE, 2011). Quatro pontos de quebra (breakpoints) foram definidos: 1.280 pixels, para reduzir a grade de KPIs de seis para três colunas; 1.024 pixels, para empilhar verticalmente os gráficos dispostos lado a lado; 768 pixels, para ocultar a sidebar e exibi-la como painel deslizante acionado por um botão de alternância, reduzindo a grade de KPIs para duas colunas; e 480 pixels, para ajustar o tamanho de fonte e o espaçamento interno dos cartões em telas de dispositivos móveis pequenos. Em telas com largura inferior a 768 pixels, a área de conteúdo ocupa toda a largura disponível e a sidebar é controlada pelas funções JavaScript `toggleSidebar()` e `closeSidebar()`, com sobreposição semitransparente (overlay) para fechar o menu ao tocar fora dela.

6 Testes de Software

A verificação automatizada do PainelPedBR concentra-se na camada de transformação do pipeline ETL, onde reside a lógica de domínio mais crítica do projeto. Erros nessa camada propagam-se silenciosamente para os produtos de dados e, consequentemente, para todos os indicadores da aplicação — tornando-a a fronteira mais importante para verificação automatizada.

6.1 Estratégia de Testes

A suíte de testes utiliza o *framework testthat* (WICKHAM; BRYAN, 2023), padrão para testes unitários em R, e adota abordagem baseada em *fixtures* programáticas: todos os dados de entrada são construídos como *data.frame* diretamente nos arquivos de teste, sem leitura de arquivos externos e sem chamadas a APIs ou serviços de rede. Essa decisão garante que os testes sejam executáveis em qualquer ambiente, sem dependência da base de dados real ou de conexão com a internet, e que permaneçam estáveis independentemente de variações nas fontes oficiais.

O escopo cobre exclusivamente a fase de transformação. A fase de extração não é coberta por envolver *downloads* de fontes externas, incompatíveis com testes unitários isolados. A fase de carga é coberta indiretamente pela validação automatizada via *pointblank* executada em tempo de pipeline. A aplicação Shiny não possui testes automatizados nesta versão: a verificação do comportamento reativo e da interface demanda infraestrutura de automação de navegador — como os pacotes *shinytest2* ou *selenider* — cuja implementação constitui direção natural para trabalhos futuros. As funções utilitárias de I/O (*file_utils.R*) também não são cobertas: as funções de *download* dependem de rede e as demais encapsulam chamadas diretas ao sistema de arquivos sem lógica de domínio testável de forma isolada.

6.2 Testes Unitários com *testthat*

A suíte é composta por quatro arquivos com 54 casos de teste e 78 expectativas individuais, organizados por componente.

Tabela 7 — Arquivos de teste

Arquivo	Componente	Casos	Expectativas	O que verificam
test_transform_sigtap.R	classify_procedures()	15	20	Filtros de pré-seleção por complexidade e grupo SIGTAP; classificação CAT1, CAT2 e CAT3 com valores de fronteira; prioridade entre categorias; descarte de procedimentos sem categoria
test_transform_cnes.R	.validar_coordenadas() .derivar_campos_geograficos() .derivar_camada_inclusao()	15	15	Validação de coordenadas nos quatro limites geográficos do Brasil; derivação de UF e região a partir do código municipal; atribuição das camadas C1, C2 e C1+C2
test_transform_ibge.R	.transformar_populacao() .transformar_pib()	12	19	Filtro de nível municipal (NC=6); soma das quatro faixas etárias pediátricas; derivação de pop_total pelo código 100362; descarte de códigos com formato inválido; tratamento de entrada nula ou vazia
test_clean_utils.R	para_numerico() para_data_br() gerar_de_para_municipio()	12	24	Conversão de decimal com vírgula; tratamento de NA e strings inválidas; parsing de data no formato brasileiro; derivação do código CNES de 6 dígitos a partir do código IBGE de 7
Total		54	78	

Fonte: elaboração própria.

Os testes de *transform_sigtap.R* exercitam especialmente os valores de fronteira da classificação tripartite: o limiar de 239 meses como teto pediátrico, o valor 9999 como indicador de faixa etária sem restrição, e a detecção de terminologia pediátrica sem distinção de maiúsculas e minúsculas. Os testes de *transform_cnes.R* verificam os quatro limites geográficos do Brasil como condições de rejeição de coordenadas e as três combinações possíveis de camada de inclusão. Os testes de *transform_ibge.R* cobrem os dois ramos do tratamento de `pop_total` — presença e ausência do código 100362 — e os casos de entrada nula para o módulo de PIB.

6.3 Resultados e Cobertura

Todos os 54 casos de teste passam sem falhas ou avisos, executados com o comando `testthat::test_dir("tests/testthat/").` Nenhuma dependência de dados externos ou de conexão de rede é necessária para a execução.

A cobertura formal de linhas não foi mensurada com ferramenta automatizada. Qualitativamente, as funções com maior densidade de lógica de domínio — `classify_procedures()` e `.derivar_camada_inclusao()` — têm seus ramos principais exercitados, incluindo casos de fronteira e combinações de entrada que representam os cenários analíticos reais do projeto. Permanecem fora do escopo de testes automatizados: a geocodificação (*geocode.R*), a fase de extração, a fase de carga e a camada da aplicação *Shiny*. A função `.extrair_centroides()` em *transform_ibge.R* não é coberta por depender de objetos espaciais *sf* incompatíveis com a abordagem de fixtures simples adotada na suíte.

7 Implantação

O PainelPedBR é distribuído em dois ambientes com propósitos distintos: o ambiente local de desenvolvimento, onde o pipeline *ETL* é executado e os dados são reprocessados, e o ambiente de produção, onde a aplicação *Shiny* é servida publicamente. A reprodutibilidade do primeiro é garantida pelo *renv*; a acessibilidade do segundo, pelo *shinyapps.io*.

7.1 Reprodutibilidade com *renv*

A reprodutibilidade do ambiente de execução é gerenciada pelo pacote *renv* (USHEY; WICKHAM, 2024), que mantém uma biblioteca de pacotes isolada por projeto e registra as versões exatas de todas as dependências em um arquivo de bloqueio (*renv.lock*). O lockfile do PainelPedBR especifica a versão do *R* (4.5.1) e de 144 pacotes, todos provenientes do *CRAN*, com seus metadados de versionamento e hash de integridade.

Apenas o *renv.lock* é versionado no repositório *Git* — a biblioteca local (*renv/library/*) é excluída via *.gitignore*, evitando o versionamento de centenas de megabytes de binários compilados. Para recriar o ambiente em uma nova máquina, basta executar `renv::restore()` após clonar o repositório: o *renv* instala automaticamente todas as dependências nas versões registradas, sem interferência nos pacotes instalados globalmente no sistema.

7.2 Publicação Online

O PainelPedBR foi publicado na plataforma shinyapps.io e está disponível publicamente, sem necessidade de cadastro ou instalação de software, no endereço: <https://karinna-lima.shinyapps.io/PainelPedBR/>.

O acesso é compatível com navegadores modernos (Chrome, Firefox, Edge, Safari) em dispositivos desktop e móveis, conforme especificado no requisito não-funcional RNF06 (Seção 4.2). A aplicação é hospedada no plano gratuito do shinyapps.io, que oferece implantação gerenciada sem necessidade de infraestrutura própria. Esse plano impõe dois limites operacionais relevantes: 25 horas de uso ativo por mês, após as quais a aplicação fica indisponível até o início do ciclo seguinte, e 1 GB de memória RAM por instância. O segundo limite influenciou diretamente decisões de arquitetura documentadas na Seção 5.3.3: produtos de dados de grande volume foram substituídos por versões pré-agregadas no ETL, reduzindo a memória necessária em runtime. Para uso contínuo sem interrupção e com maior margem de memória, a migração para um plano pago da plataforma ou para infraestrutura própria — como um servidor Shiny Server — constitui direção natural para a evolução da implantação.

8 Manual do Usuário

Esta seção descreve os requisitos de acesso ao PainelPedBR e o funcionamento da interface de navegação. O painel é uma aplicação web interativa — não requer instalação de software e pode ser acessado diretamente pelo navegador.

8.1 Requisitos de Acesso

O acesso ao PainelPedBR não requer cadastro, instalação de software ou configuração prévia. Os únicos requisitos são conexão com a internet e um navegador moderno. A aplicação foi validada nos navegadores Chrome, Firefox, Edge e Safari, em suas versões correntes. O uso em dispositivos móveis é suportado — a interface adapta-se a telas menores por meio de layout responsivo — mas o uso em monitor desktop é recomendado em função do volume de informações exibidas simultaneamente.

A aplicação é hospedada no plano gratuito do shinyapps.io, que impõe um limite mensal de horas de uso ativo. Em caso de indisponibilidade temporária por esgotamento da cota, o acesso é restabelecido automaticamente no início do ciclo seguinte.

8.2 Navegação Geral

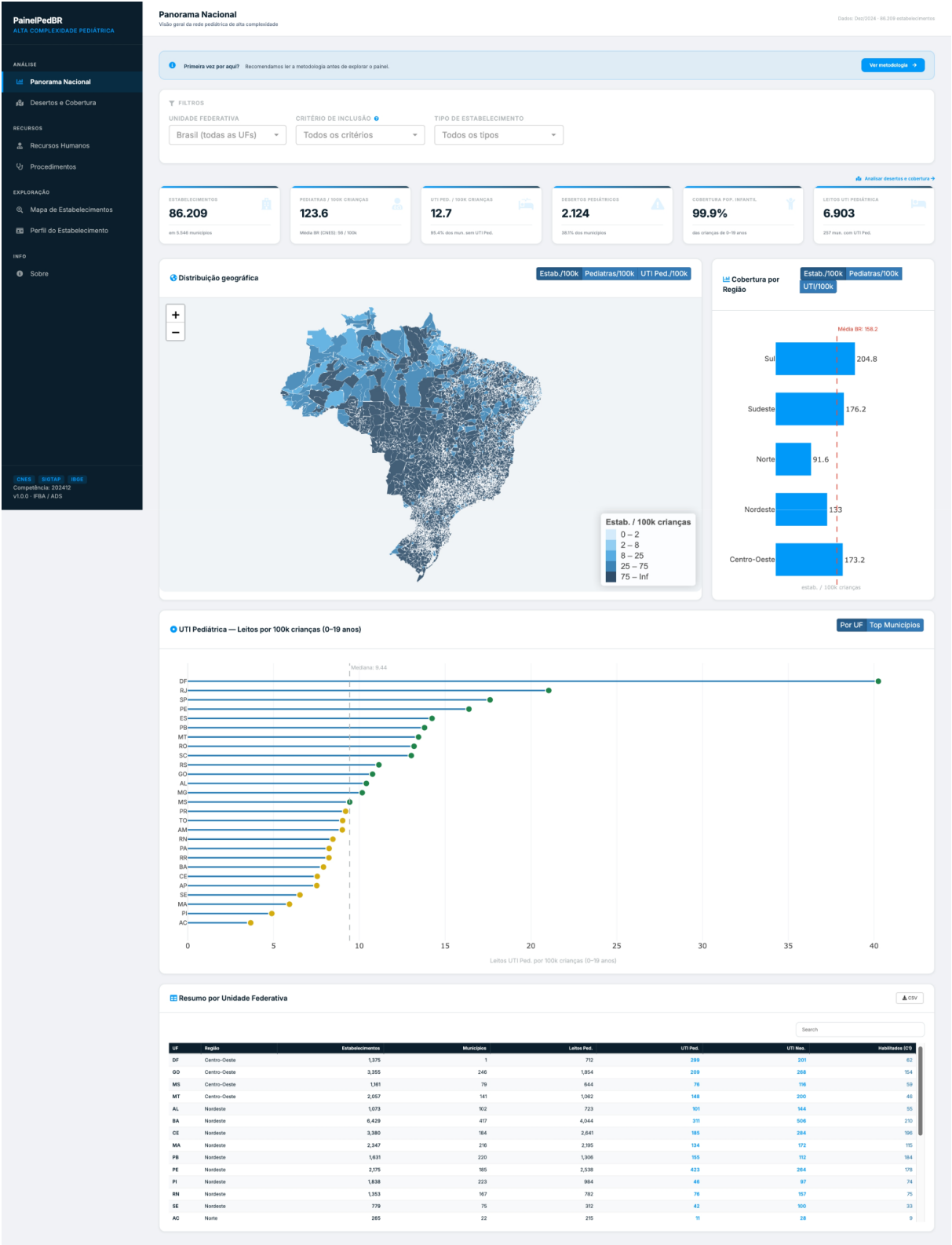
A barra lateral organiza as seções em quatro grupos temáticos: Análise (Panorama Nacional e Desertos e Cobertura), Recursos (Recursos Humanos e Procedimentos), Exploração (Mapa de Estabelecimentos e Perfil do Estabelecimento) e Info (Sobre). A navegação entre seções ocorre por

clique nos itens da barra lateral, sem recarregamento da página. A barra superior exibe o título da seção ativa. Em dispositivos móveis, a barra lateral é acionada pelo botão de menu no canto superior esquerdo.

8.2.1 Panorama Nacional

O Panorama Nacional é a seção exibida por padrão ao abrir o painel e apresenta uma visão agregada da rede pediátrica de alta complexidade no Brasil. O painel de filtros permite recortar a visualização por região geográfica, Unidade Federativa e camada metodológica (C1, C2 ou C1+C2). Os indicadores são atualizados reativamente a cada alteração de filtro.

Figura 3 — Interface do módulo Panorama Nacional



Fonte: elaboração própria.

A seção organiza-se em três blocos verticais. O bloco superior exibe seis indicadores-chave (KPIs): total de estabelecimentos, municípios cobertos, pediatras por 100 mil crianças, leitos UTI pediátrica por 10 mil crianças, municípios classificados como desertos pediátricos e cobertura da

população infantil. O bloco central apresenta o mapa interativo, com alternância entre visualização por pontos georreferenciados e mapa coroplético por município. O bloco inferior reúne três gráficos complementares: distribuição regional dos estabelecimentos, funil de camadas metodológicas C1/C2 e classificação dos dez municípios com maior número de leitos UTI pediátricos.

8.2.2 Sobre

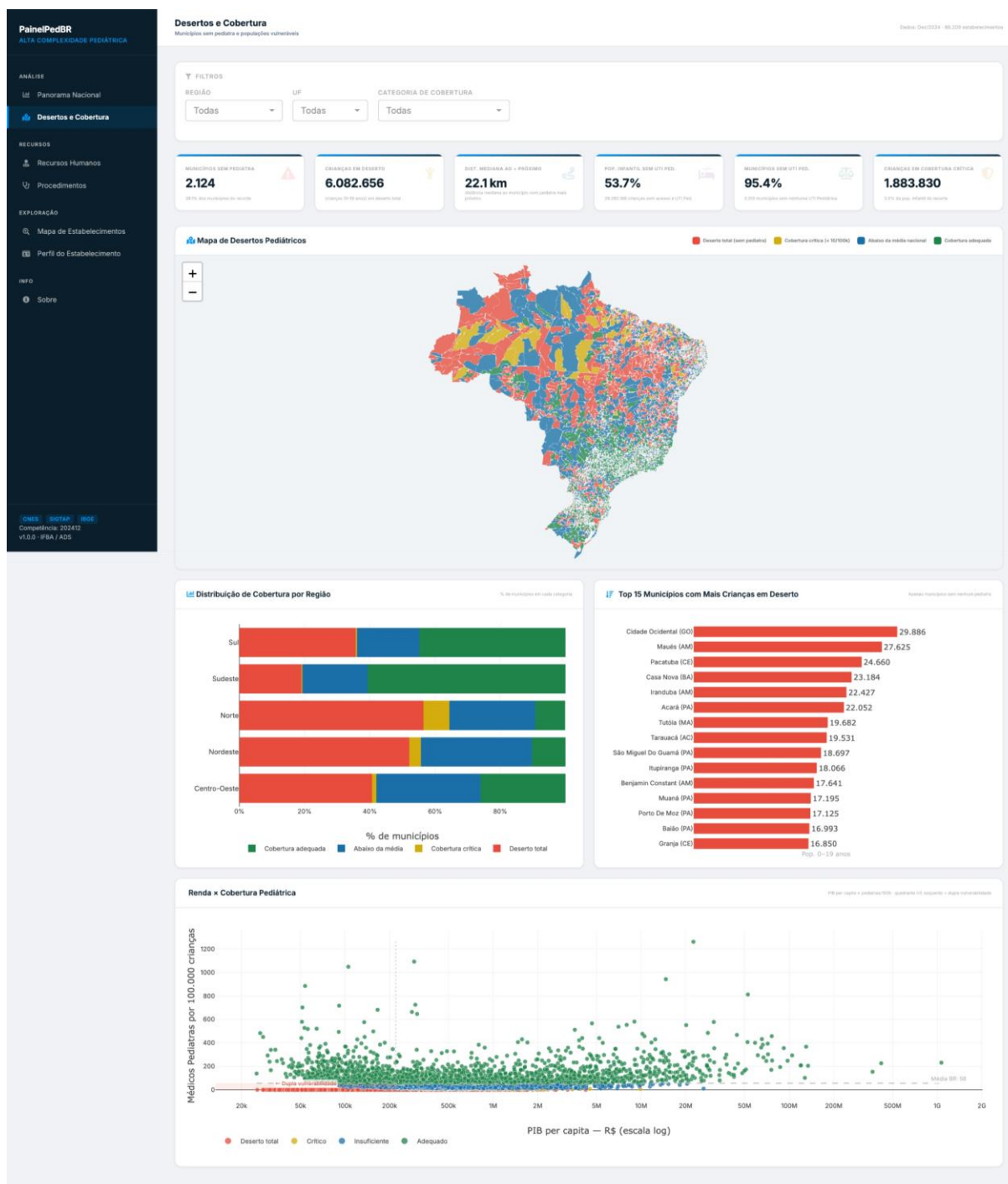
A seção "Sobre" reúne as informações institucionais e metodológicas do PainelPedBR em uma página estática de consulta. Não possui filtros ou elementos interativos.

A página organiza-se em seis blocos: um banner de apresentação com indicadores resumidos do painel (número de estabelecimentos, municípios cobertos, pediatras e competência dos dados); um guia de navegação com a descrição de cada seção do painel; a metodologia de seleção de estabelecimentos (camadas C1 e C2) e de classificação de procedimentos (CAT1, CAT2 e CAT3); as fontes de dados utilizadas (CNES, SIGTAP e IBGE/Censo 2022) com indicação da competência de referência; as limitações e considerações de uso; e a equipe do projeto com a sugestão de citação e link para o repositório público no GitHub.

8.2.3 Desertos e Cobertura

A seção Desertos e Cobertura responde à pergunta: "quais municípios estão sendo deixados para trás?". Os filtros de região, UF e categoria de cobertura permitem recortar a análise por território e por situação de acesso.

Figura 5 — Interface do módulo Desertos e Cobertura



Fonte: elaboração própria.

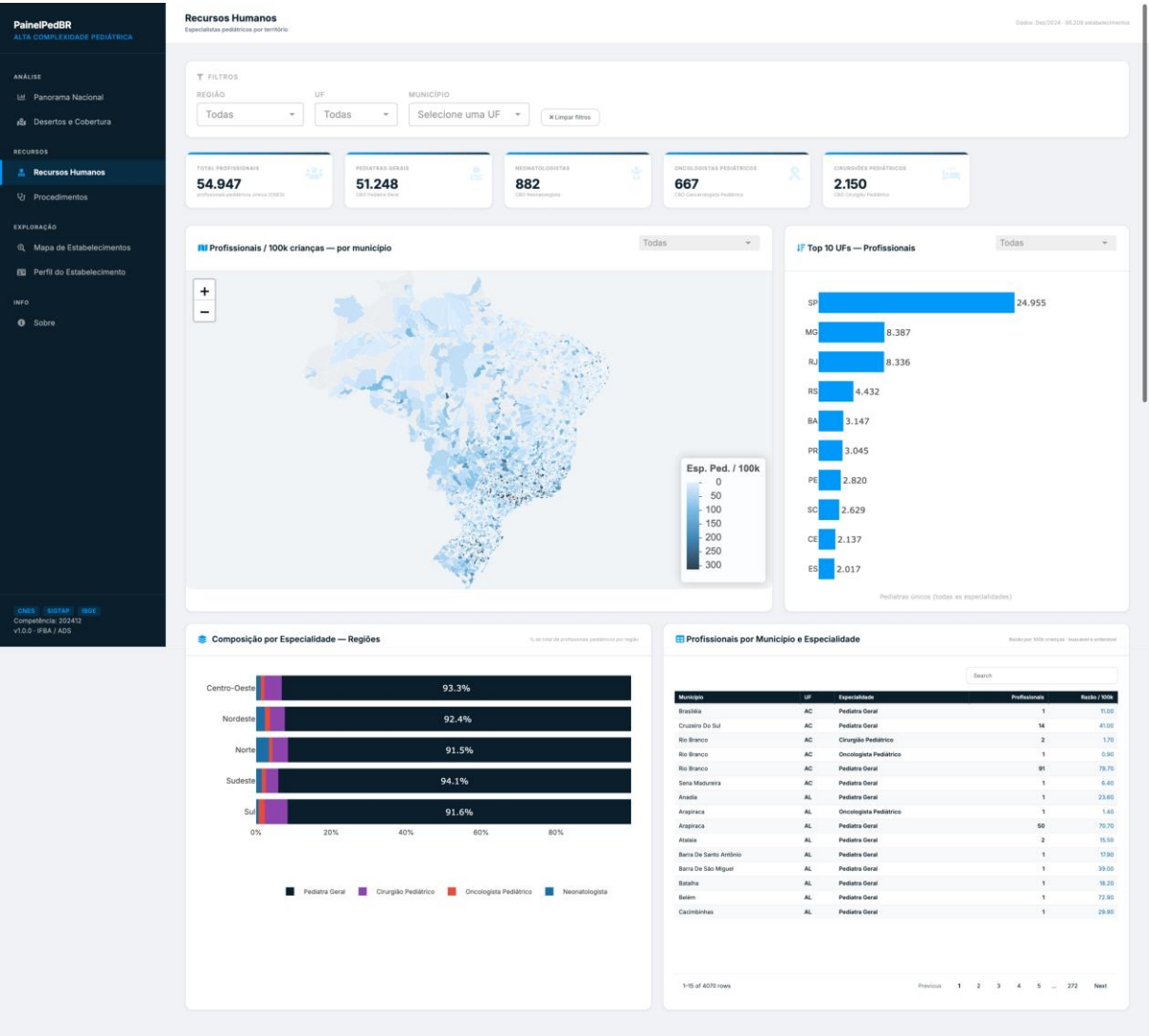
Seis indicadores sintetizam o cenário: municípios sem nenhum pediatra, crianças residentes em deserto pediátrico total, distância mediana ao município com pediatra mais próximo, percentual da população infantil sem acesso a UTI Pediátrica, percentual de municípios sem UTI Pediátrica e crianças em cobertura crítica. O mapa coroplético classifica os municípios em quatro categorias de cobertura. O gráfico de barras empilhadas compara a distribuição proporcional das categorias entre as regiões do país. O gráfico de dispersão relaciona PIB per capita e cobertura pediátrica por município, destacando o quadrante de dupla vulnerabilidade — baixa renda e baixa cobertura simultaneamente. A

classificação lista os quinze municípios com maior número absoluto de crianças em deserto total; clicar em um município no ranking navega automaticamente para o Mapa de Estabelecimentos com aquele território selecionado.

8.2.4 Recursos Humanos

A seção Recursos Humanos apresenta a distribuição dos profissionais de saúde vinculados ao escopo pediátrico por especialidade e território. Os filtros de região, UF e município permitem recortar a análise geograficamente.

Figura 6 — Interface do módulo Recursos Humanos



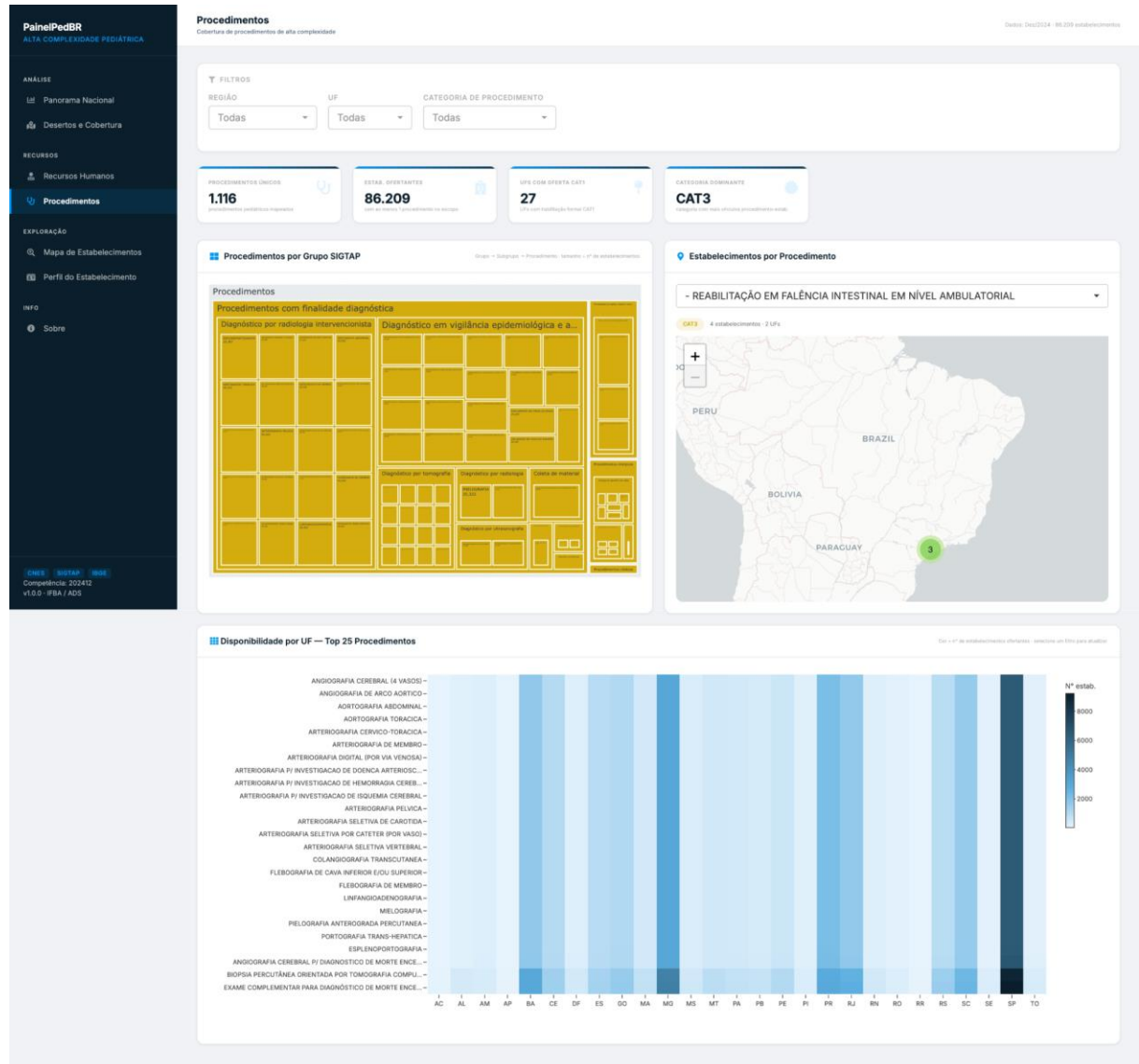
Fonte: elaboração própria.

Os profissionais são agrupados em cinco especialidades: Neonatologista, Oncologista Pediátrico, Cirurgião Pediátrico, Pediatra Geral e demais CBOs pediátricos. O mapa exibe a distribuição territorial e a tabela analítica lista os vínculos individuais com filtro por especialidade.

8.2.5 Procedimentos

A seção Procedimentos mapeia a oferta de procedimentos pediátricos de alta complexidade por estabelecimento e território. Os filtros permitem selecionar UF, categoria de classificação (CAT1, CAT2 ou CAT3) e grupo SIGTAP.

Figura 7 — Interface do módulo Procedimentos



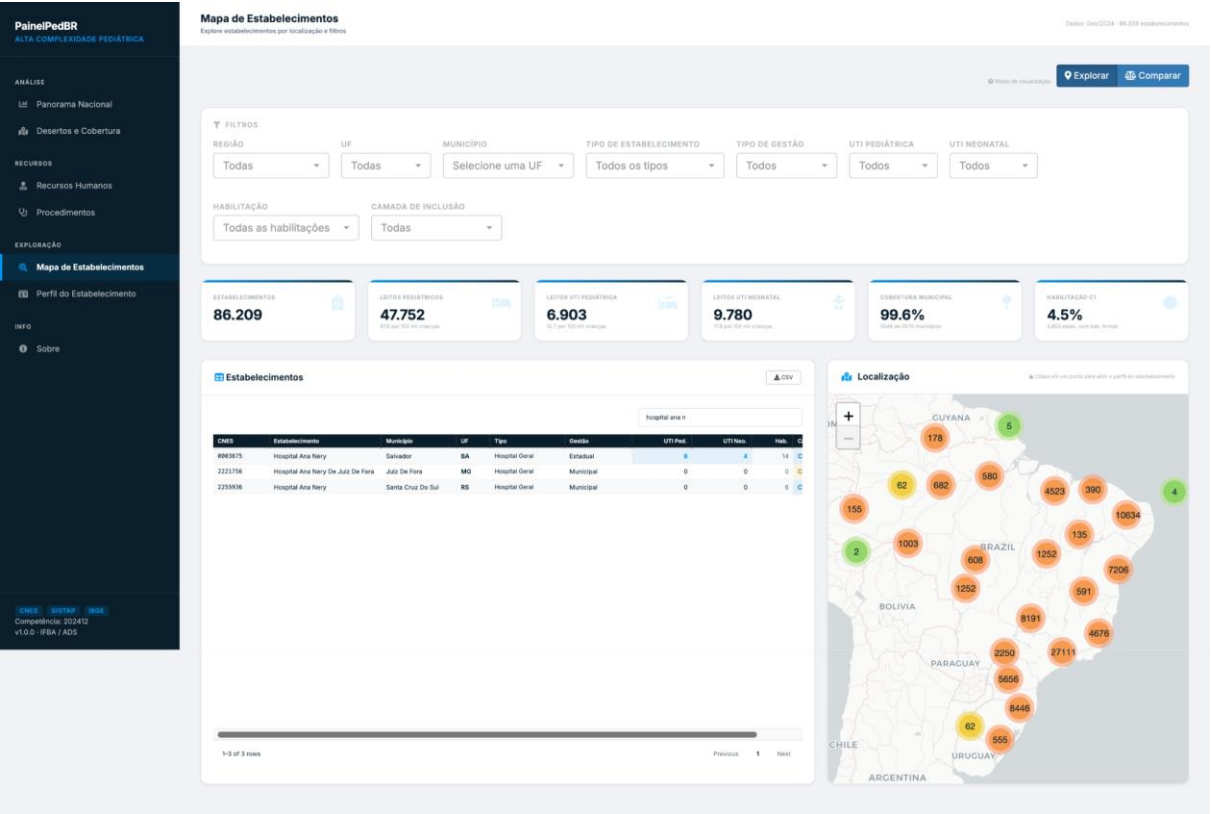
Fonte: elaboração própria.

O mapa de árvore hierárquico organiza os procedimentos por grupo e subgrupo SIGTAP, com tamanho proporcional ao número de estabelecimentos ofertantes. O mapa de calor cruza os vinte procedimentos mais ofertados com as Unidades Federativas. O mapa pontual filtra os estabelecimentos que ofertam o procedimento selecionado no mapa de árvore.

8.2.5 Mapa de Estabelecimentos

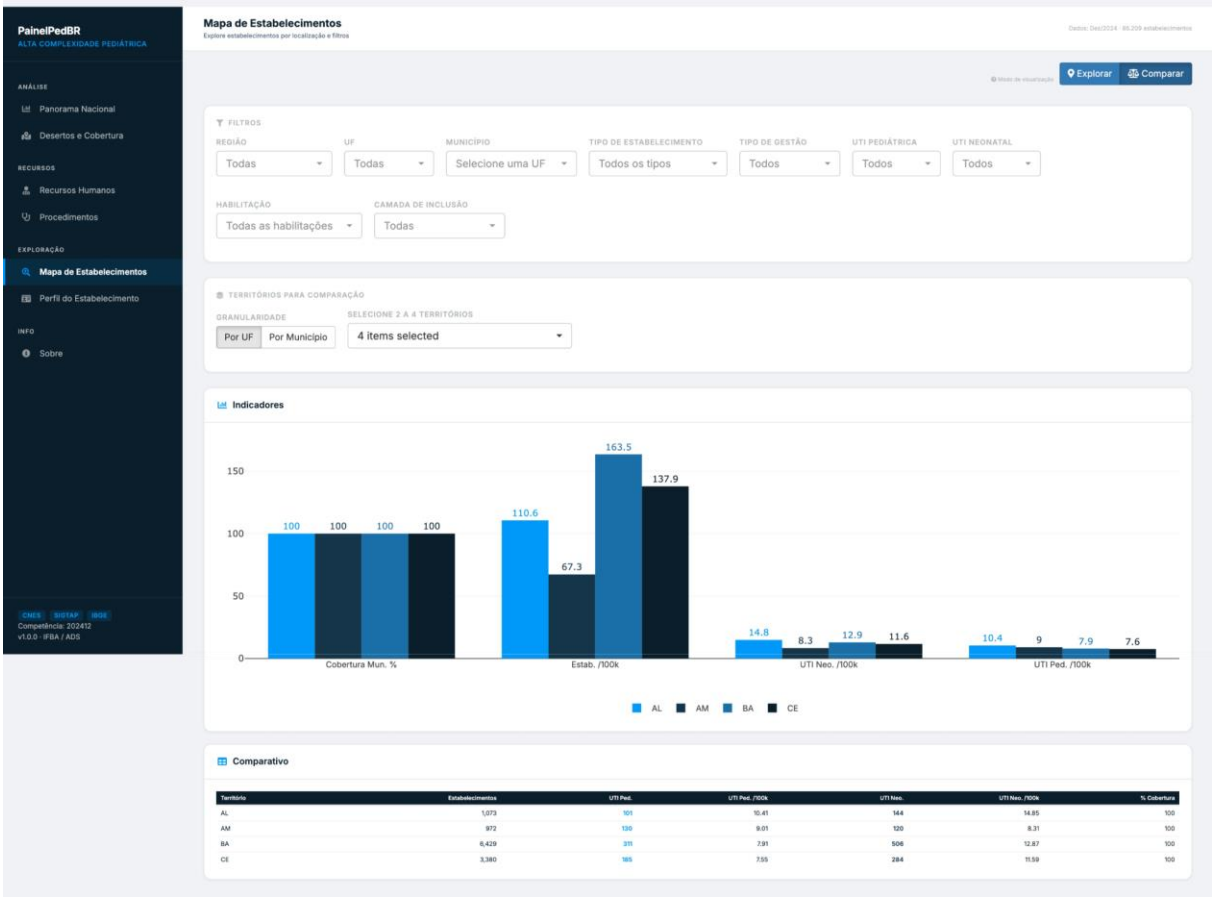
O Mapa de Estabelecimentos permite localizar e filtrar estabelecimentos do escopo analítico por região, UF, município, tipo e camada metodológica. A tabela analítica é o componente central da seção; o mapa atua como suporte visual à localização geográfica.

Figura 8 — Interface do módulo Mapa de Estabelecimentos (modo "Explorar")



Fonte: elaboração própria.

Figura 9 — Interface do módulo Mapa de Estabelecimentos (modo "Comparar")



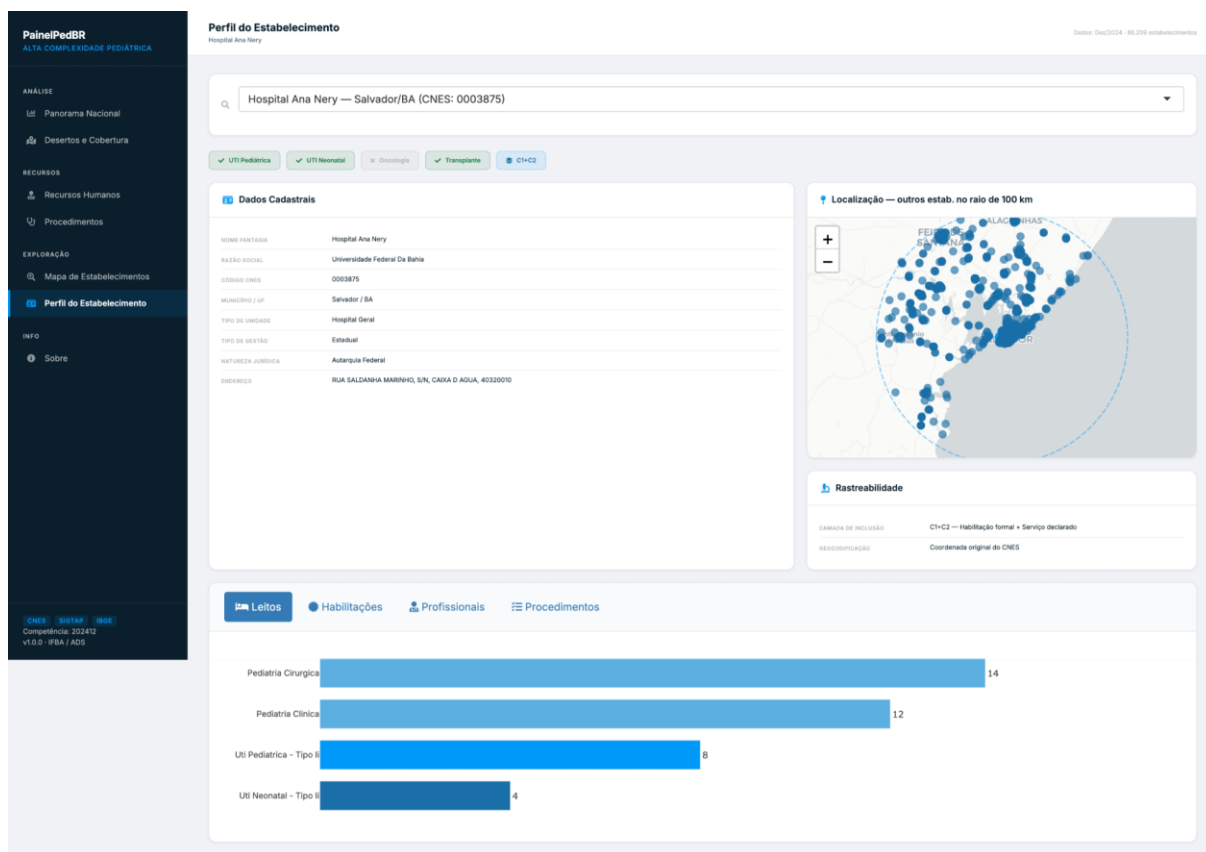
Fonte: elaboração própria.

A seção oferece dois modos de uso, selecionáveis no topo: Explorar, para filtrar e visualizar estabelecimentos individualmente, e "Comparar", para contrastar dois a quatro territórios em painel lado a lado. O clique em um marcador do mapa ou em uma linha da tabela abre automaticamente o Perfil do Estabelecimento correspondente.

8.2.7 Perfil do Estabelecimento

O Perfil do Estabelecimento exibe a ficha completa de um estabelecimento específico. O acesso pode ser feito por dois caminhos: busca textual pelo nome, código CNES ou município na barra de pesquisa da própria seção, ou clique em um estabelecimento no Mapa de Estabelecimentos.

Figura 10 — Interface do módulo Perfil do Estabelecimento



Fonte: elaboração própria.

A ficha consolida todas as informações disponíveis para o estabelecimento selecionado: dados cadastrais (endereço, tipo, camada de inclusão C1/C2), leitos por tipo, habilitações formais, vínculos profissionais e procedimentos ofertados. Todos esses dados são carregados sob demanda no momento da seleção.

8.3 Cruzamento de Informações

O PainelPedBR é projetado como um sistema integrado: cada seção responde a um ângulo específico da mesma questão — a disponibilidade de oferta pediátrica de alta complexidade —, e as informações de uma seção ganham profundidade quando combinadas com as demais. Esta seção descreve os mecanismos de integração disponíveis e os principais fluxos de análise que deles resultam.

8.3.1 Filtros geográficos compartilhados

A seleção de região ou UF é compartilhada entre as seções Panorama Nacional, Desertos e Cobertura, Procedimentos, Mapa de Estabelecimentos e Recursos Humanos. Ao navegar de uma seção para outra, o filtro geográfico ativo é preservado automaticamente, eliminando a necessidade de reaplicar seleções. Isso permite comparar a mesma unidade territorial, um estado, uma região, sob ângulos distintos sem interromper o recorte.

8.3.2 Atalhos de navegação direta

Além da navegação manual pela barra lateral, o painel oferece atalhos que conectam seções em pontos estratégicos da análise:

Tabela 8 — Mapeamento das interações de navegação entre módulos do painel

Origem	Elemento interativo	Destino
Panorama Nacional	Botão "Analisar desertos e cobertura"	Desertos e Cobertura
Panorama Nacional	Clique em uma UF no gráfico de UTI por UF	Desertos e Cobertura com a UF pré-selecionada
Panorama Nacional	Botão "Ver metodologia"	Sobre
Desertos e Cobertura	Clique em um município no <i>ranking</i> de crianças em deserto	Mapa de Estabelecimentos com o município pré-selecionado
Recursos Humanos	Clique em uma linha da tabela de profissionais	Mapa de Estabelecimentos com o município pré-selecionado
Mapa de Estabelecimentos	Clique em um marcador no mapa ou linha da tabela	Perfil do Estabelecimento

Fonte: elaboração própria.

8.3.3 Fluxos de análise integrada

Fluxo 1 — Diagnóstico estadual completo: "Como está o estado X?"

1. Em Panorama Nacional, selecionar a UF no filtro. Os indicadores, o mapa e o gráfico de cobertura por região atualizam para o recorte estadual. O gráfico de UTI por UF destaca a posição do estado em relação aos demais.
2. Clicar na barra da UF no gráfico de UTI navega diretamente para Desertos e Cobertura com aquela UF pré-selecionada. Os seis indicadores mostram quantos municípios não têm pediatria, quantas crianças estão em deserto, a distância mediana ao mais próximo e a proporção sem UTI Pediátrica. O gráfico de dispersão PIB per capita × cobertura revela se a baixa cobertura se concentra nos municípios de menor renda — o quadrante de dupla vulnerabilidade.
3. Navegar para Recursos Humanos — o filtro de UF persiste. Os indicadores detalham a composição por especialidade; o mapa coroplético permite alternar entre especialidades para identificar quais estão ausentes em determinadas microrregiões.
4. Navegar para Procedimentos — o filtro de UF persiste. O mapa de calor de disponibilidade por UF contextualiza quais categorias de procedimentos (CAT1, CAT2, CAT3) estão presentes no estado.

Fluxo 2 — Do deserto ao estabelecimento mais próximo: "Há alguma opção acessível?"

1. Em Desertos e Cobertura, identificar no ranking os municípios com maior número de crianças sem nenhum pediatra.
2. Clicar em um município no *ranking* navega para o Mapa de Estabelecimentos com aquele município pré-selecionado. A tabela e o mapa mostram os estabelecimentos do município e, se o município não tiver nenhum, os resultados podem ser ampliados removendo o filtro de município e mantendo apenas a UF, para visualizar os estabelecimentos mais próximos na região.
3. Clicar em qualquer estabelecimento na tabela ou no mapa abre o Perfil do Estabelecimento. O minimapa na ficha exibe todos os demais estabelecimentos do escopo no raio de 100 km, com a distância exata de cada um — permitindo avaliar o tempo de deslocamento para a população do município em deserto.

Fluxo 3 — Rastreamento de um procedimento específico: "Onde é possível realizar este procedimento?"

1. Em Procedimentos, selecionar a categoria de procedimento (CAT1, CAT2 ou CAT3) e navegar pelo mapa de árvore hierárquico (Grupo → Subgrupo → Procedimento) para localizar o procedimento de interesse.
2. Selecionar o procedimento no menu ao lado do mapa pontual. O mapa atualiza exibindo apenas os estabelecimentos que ofertam aquele procedimento, com indicação de quantos estão sem geolocalização.
3. Aplicar o filtro de UF ou Região para restringir a busca a um território. O mapa de calor abaixo mostra a disponibilidade comparada entre UFs para os 25 procedimentos mais ofertados, permitindo identificar estados com lacunas específicas.

Fluxo 4 — Investigação por tipo de profissional: "Onde há oncologistas pediátricos?"

1. Em Recursos Humanos, selecionar "Oncologista Pediátrico" no filtro de especialidade do mapa coroplético. O mapa atualiza mostrando a densidade de oncologistas por município, e o *ranking* exibe as 10 UFs com maior concentração.
2. Clicar em uma linha da tabela de profissionais navega para o Mapa de Estabelecimentos com aquele município selecionado, mostrando quais estabelecimentos estão cadastrados e seus dados de UTI Pediátrica.
3. Clicar em um estabelecimento na tabela ou no mapa abre o Perfil do Estabelecimento, onde a aba Profissionais detalha todos os CBOs pediátricos vinculados àquele estabelecimento específico.

Fluxo 5 — Comparação entre territórios: "Como se distribuem os indicadores entre territórios diferentes?"

1. Em Panorama Nacional, a tabela resumo por UF (com download CSV) permite uma comparação inicial entre todos os estados — estabelecimentos, municípios cobertos, leitos pediátricos, UTI Pediátrica, UTI Neonatal e habilitações C1.
2. Em Mapa de Estabelecimentos, ativar o modo Comparar e selecionar de 2 a 4 UFs ou municípios. Para 2 territórios, o gráfico *de halteres* destaca as diferenças em três indicadores normalizados (cobertura municipal, leitos UTI Neonatal/100k e UTI Pediátrica/100k). Para 3 ou 4 territórios, o gráfico de barras agrupadas e a tabela comparativa apresentam os mesmos indicadores lado a lado.
3. Para aprofundar a comparação em recursos humanos, navegar para Recursos Humanos e alternar entre especialidades no *ranking* por UF, observando a posição relativa dos estados selecionados.

Fluxo 6 — Ficha completa de um estabelecimento: "O que este hospital oferece?"

1. Localizar o estabelecimento pelo Mapa de Estabelecimentos (filtros por UF, município, tipo, UTI Pediátrica ou habilitação) ou acessar diretamente o Perfil do Estabelecimento e buscar pelo nome, código CNES ou município.
2. A ficha consolida quatro dimensões em abas: Leitos (por tipo, com destaque para UTI Pediátrica e Neonatal), Habilitações (formais registradas no CNES, com categoria pediátrica de referência), Profissionais (vínculos por CBO, carga horária ambulatorial e hospitalar) e Procedimentos (mapeados por categoria CAT1/CAT2/CAT3 e grupo SIGTAP).
3. O minimapa integrado exibe todos os estabelecimentos do escopo no raio de 100 km, com distância em quilômetros — útil para avaliar se o estabelecimento é o único ponto de referência regional ou se há alternativas próximas.

9 Considerações Finais

No Brasil, os dados necessários para compreender a distribuição territorial dos serviços pediátricos especializados existem, são públicos e atualizados periodicamente, mas permanecem dispersos em sistemas independentes, com formatos heterogêneos e sem nenhum mecanismo de integração acessível ao público. Famílias com crianças que dependem de cuidados de alta complexidade não dispõem de qualquer forma direta de verificar se há uma unidade de terapia intensiva pediátrica, um serviço de oncologia infantil ou um centro de reabilitação a uma distância razoável de onde vivem. Gestores municipais, pesquisadores e formuladores de políticas públicas enfrentam obstáculo semelhante: obter uma visão integrada da oferta instalada exige extrair, tratar e cruzar manualmente ao menos três bases públicas independentes (o Cadastro Nacional de Estabelecimentos de Saúde, a Tabela de Procedimentos do Sistema Único de Saúde e os dados populacionais do Censo Demográfico de 2022), um processo que pressupõe domínio técnico que não pode ser assumido como dado entre os potenciais usuários dessas informações.

O PainelPedBR foi desenvolvido para eliminar esse obstáculo. Trata-se de uma aplicação web interativa e de acesso público irrestrito que integra as três bases mencionadas em um sistema analítico organizado em seis módulos temáticos, cobrindo as quatro dimensões definidas como objetivo do trabalho: infraestrutura assistencial (estabelecimentos, leitos pediátricos e de terapia intensiva), distribuição territorial (municípios sem cobertura especializada e desertos pediátricos), oferta de procedimentos (mapeamento dos procedimentos disponíveis por estabelecimento e território) e recursos humanos (distribuição de especialistas pediátricos por município e unidade federativa). A aplicação está publicada e disponível sem necessidade de instalação de *software* ou cadastro por parte do usuário.

As ferramentas públicas de consulta disponíveis na plataforma do Ministério da Saúde — como o TabNet e o e-Gestor — oferecem acesso à mesma base de dados, mas operam por exportação tabular de recortes predefinidos, sem integração entre fontes e sem recursos analíticos que tornem a informação interpretável sem conhecimento técnico especializado. O PainelPedBR se diferencia dessas ferramentas em dois aspectos fundamentais. O primeiro é a base metodológica da identificação dos estabelecimentos: em vez de depender exclusivamente do cadastro autodeclarado pelas unidades de saúde, o sistema identifica os estabelecimentos com oferta pediátrica de alta complexidade a partir dos critérios normativos da tabela de procedimentos — abordagem que confere rastreabilidade explícita ao escopo analítico e que pode ser auditada ou ajustada de forma independente, pois todos os critérios estão externalizados em arquivos de configuração separados do código de implementação. O segundo é a normalização dos indicadores pela população infantil de cada território: ao expressar a oferta em relação à população de zero a dezenove anos residente em cada município, o painel torna comparáveis territórios de tamanhos populacionais muito diferentes, substituindo contagens absolutas por indicadores de cobertura efetiva.

A escolha tecnológica central do projeto foi também a mais determinante: desenvolver todo o sistema — da extração e transformação dos dados à interface do usuário — em R, uma linguagem de programação consolidada em análise estatística e ciência de dados que, por meio do *framework Shiny* e de bibliotecas como *Leaflet* para mapas interativos, *Plotly* para gráficos e *reactable* para tabelas analíticas, permite construir aplicações web completas sem que o desenvolvedor precise dominar tecnologias de *front-end* separadas. Essa convergência não é apenas uma conveniência operacional: ela tornou possível que um único desenvolvedor projetasse, implementasse e publicasse um sistema de ponta a ponta de forma completamente autônoma. Sem ela, seria necessário integrar ao menos duas pilhas tecnológicas distintas — uma para análise de dados e outra para a interface *web* —, o que aumentaria consideravelmente a complexidade de manutenção e inviabilizaria o escopo do projeto nas condições em que foi realizado. A experiência do PainelPedBR reforça que, para projetos analíticos desenvolvidos de forma independente, essa integração representa uma referência difícil de abandonar: quem passa a desenvolver análise e interface em uma única linguagem dificilmente volta a estruturar projetos analíticos de outra forma. O gerenciamento de dependências por meio do sistema *renv*, que registra as versões exatas das 144 bibliotecas utilizadas em um arquivo versionado junto ao código, provou ser igualmente indispensável — tanto na publicação da aplicação quanto na reexecução do

pipeline após atualizações de sistema, situações em que incompatibilidades de versão são frequentes e de difícil diagnóstico sem esse controle. Também no plano da implantação, a publicação foi realizada na plataforma *shinyapps.io*, que oferece hospedagem gerenciada para aplicações *Shiny* sem necessidade de infraestrutura própria. O plano gratuito utilizado impõe um limite de memória de um *gigabyte* por instância — restrição que se revelou determinante para uma das decisões arquiteturais mais relevantes do projeto: em vez de carregar os conjuntos de dados brutos em memória e realizar agregações em tempo de execução, o *pipeline* de processamento gera produtos de dados pré-agregados que reduzem substancialmente o volume carregado pela aplicação no momento do acesso. O resultado é que a arquitetura do sistema foi moldada não apenas pela lógica analítica, mas também pelas condições reais de implantação — o que reforça a importância de considerar o ambiente de produção desde as fases iniciais do *design*. Para uso contínuo sem restrição de cota, a migração para um plano pago ou para infraestrutura própria com *Shiny Server* constitui a evolução natural da implantação.

A implementação revelou algo que a especificação inicial não antecipava. A adoção de testes automatizados para a lógica de classificação dos procedimentos — escritos antes de qualquer validação com dados reais — forçou a explicitação de todos os casos de fronteira das regras metodológicas. Situações de sobreposição entre categorias de classificação, que permaneceriam implícitas no código e poderiam contaminar silenciosamente os indicadores exibidos ao usuário, foram identificadas e resolvidas durante a escrita dos testes. A expectativa inicial era que os testes servissem para verificar o comportamento do código após sua escrita; o que a implementação demonstrou é que, neste contexto, escrevê-los foi parte do processo de especificação da própria metodologia — e não apenas de sua verificação. Isso sugere que, para sistemas analíticos cujos resultados dependem de classificações com múltiplos casos de fronteira, testes automatizados deveriam ser tratados como instrumento de *design* metodológico, não apenas de controle de qualidade.

Do ponto de vista da utilização, o painel foi projetado para operar de forma progressiva e integrada: o usuário parte de uma visão nacional agregada, restringe o território por região, estado ou município, e ao identificar um estabelecimento de interesse acessa sua ficha completa com leitos, habilitações formais, vínculos profissionais e procedimentos ofertados. As seções são conectadas por atalhos de navegação que preservam o filtro geográfico ativo, de modo que uma análise iniciada em um ângulo pode ser continuada em outro sem interromper o recorte territorial. Para a população em geral, isso significa poder verificar, sem intermediários técnicos, se há um serviço pediátrico especializado a uma distância acessível. Para gestores e pesquisadores, significa dispor de uma visão sistêmica que antes exigia processamento manual de grandes volumes de dados. Nesse sentido, o painel cumpre uma função que vai além da apresentação de informações: ele desloca a barreira de acesso à informação em saúde de um requisito técnico para uma decisão de navegação.

O desenvolvimento contou com a consultoria do Prof. Marcus Vinicius Teixeira Navarro, pesquisador do Laboratório de Pesquisa em Saúde Pública e Ambiente do Instituto Federal da Bahia, cuja contribuição foi determinante para o embasamento epidemiológico dos critérios de seleção de estabelecimentos e de classificação de procedimentos. Essa interlocução com a pesquisa aplicada em

saúde pública posiciona o trabalho além do desenvolvimento de *software*: os critérios implementados refletem discussões em curso no campo da epidemiologia pediátrica brasileira, e o painel está posicionado para servir como instrumento de suporte a investigações futuras sobre acesso e equidade na atenção pediátrica especializada.

Como continuidade natural, a direção mais relevante é a automatização da atualização periódica dos dados, hoje dependente de execução manual a cada nova competência das bases públicas. A incorporação de séries temporais permitiria analisar a evolução da distribuição dos recursos ao longo do tempo, e o cruzamento com dados de produção do sistema público de saúde confrontaria a oferta instalada com a utilização efetiva. No plano metodológico, a extensão do modelo de classificação para abranger os três níveis de complexidade assistencial — atenção básica, média e alta — permitiria delimitar o escopo da alta complexidade com maior precisão e oferecer ao usuário uma perspectiva comparativa entre níveis de atenção.

Para o campo da saúde pediátrica no Brasil, onde a invisibilidade territorial dos recursos é parte constitutiva do problema de equidade, uma ferramenta que transforma dados públicos dispersos em informação analítica navegável representa uma contribuição com impacto direto sobre quem mais precisa dela. O PainelPedBR demonstra que dados de saúde, quando integrados com rigor metodológico e apresentados em interface acessível, deixam de ser patrimônio técnico restrito a especialistas e passam a ser um recurso efetivamente disponível para qualquer cidadão que precise compreender onde a oferta de cuidados pediátricos especializados existe no seu território — e para quem, no exercício de políticas públicas ou de pesquisa, precise compreender onde ela falta e o que isso significa para a população infantil brasileira.

Referências

AVELINO, D. P. de; POMPEU, J. C. B.; FONSECA, I. F. da. **Democracia digital: mapeamento de experiências em dados abertos, governo digital e ouvidorias públicas**. Brasília: Instituto de Pesquisa Econômica Aplicada, 2021.

BASS, Len; CLEMENTS, Paul; KAZMAN, Rick. **Software architecture in practice**. 3. ed. Boston: Addison-Wesley, 2012.

BRASIL. **Constituição (1988)**. Constituição da República Federativa do Brasil. Brasília, DF: Senado Federal, 1988.

BRASIL. **Decreto nº 8.777, de 11 de maio de 2016**. Institui a Política de Dados Abertos do Poder Executivo Federal. Diário Oficial da União: seção 1, Brasília, DF, 12 maio 2016.

BRASIL. **Lei nº 8.069, de 13 de julho de 1990**. Dispõe sobre o Estatuto da Criança e do Adolescente. Diário Oficial da União: seção 1, Brasília, DF, 16 jul. 1990.

BRASIL. **Lei nº 8.080, de 19 de setembro de 1990**. Dispõe sobre as condições para a promoção, proteção e recuperação da saúde, a organização e o funcionamento dos serviços correspondentes. Diário Oficial da União: seção 1, Brasília, DF, 20 set. 1990.

BRASIL. **Lei nº 12.527, de 18 de novembro de 2011**. Lei de Acesso à Informação. Diário Oficial da União: seção 1, Brasília, DF, 18 nov. 2011.

BRASIL. Ministério da Saúde. **Política Nacional de Atenção Básica**. Brasília, DF: Ministério da Saúde, 2012.

BRASIL. Ministério da Saúde. **Portaria GM/MS nº 321, de 8 de fevereiro de 2007**. Institui a Tabela de Procedimentos, Medicamentos, Órteses, Próteses e Materiais Especiais do SUS. Diário Oficial da União: seção 1, Brasília, DF, 9 fev. 2007.

BRASIL. Ministério da Saúde. **Portaria GM/MS nº 373, de 27 de fevereiro de 2002**. Aprova a Norma Operacional da Assistência à Saúde – NOAS-SUS 01/2002. Diário Oficial da União: seção 1, Brasília, DF, 28 fev. 2002.

BRASIL. Ministério da Saúde. **Portaria GM/MS nº 1.130, de 5 de agosto de 2015**. Diário Oficial da União: seção 1, Brasília, DF, 6 ago. 2015.

BRASIL. Ministério da Saúde. **Portaria GM/MS nº 4.279, de 30 de dezembro de 2010**. Diário Oficial da União: seção 1, Brasília, DF, 31 dez. 2010.

BRASIL. Ministério da Saúde. **Portaria de Consolidação GM/MS nº 1, de 28 de setembro de 2017**. Diário Oficial da União: seção 1, Brasília, DF, 3 out. 2017.

BRASIL. Ministério da Saúde. Secretaria de Atenção à Saúde. **Portaria SAS/MS nº 511, de 29 de dezembro de 2000**. Diário Oficial da União: seção 1, Brasília, DF, 30 dez. 2000.

BRASIL. Ministério da Saúde. Departamento de Informática do SUS. **Cadastro Nacional de Estabelecimentos de Saúde (CNES)**. Brasília, DF, 2024. Disponível em: <http://cnes.datasus.gov.br>. Acesso em: 23 jun. 2026.

BRASIL. Ministério da Saúde. Departamento de Informática do SUS. **TabNet Win32 3.0**. Brasília, DF, 2024. Disponível em: <http://www2.datasus.gov.br/DATASUS/index.php?area=02>. Acesso em: 23 jun. 2026.

BRASIL. Ministério da Saúde. **Sistema de Gerenciamento da Tabela de Procedimentos, Medicamentos e OPM do SUS (SIGTAP)**. Brasília, DF, 2024. Disponível em: <http://sigtap.datasus.gov.br>. Acesso em: 23 jun. 2026.

CAIRO, Alberto. **The truthful art: data, charts, and maps for communication**. Berkeley: New Riders, 2016.

CARVALHO, A. O.; EDUARDO, M. B. de P. **Sistemas de informação em saúde para municípios**. São Paulo: Faculdade de Saúde Pública da USP, 1998.

CHANG, Winston et al. **shiny: web application framework for R**. R package version 1.9.1. 2024. Disponível em: <https://CRAN.R-project.org/package=shiny>. Acesso em: 23 jun. 2026.

CROMLEY, E. K.; McLAFFERTY, S. L. **GIS and public health**. 2. ed. New York: Guilford Press, 2012.

FERRI, R. L. et al. Transition from pediatric to adult care for young adults with chronic conditions. **Pediatrics**, Itasca, v. 128, n. 5, p. 1146–1152, 2011.

FEW, Stephen. **Information dashboard design: displaying data for at-a-glance monitoring**. 2. ed. Burlingame: Analytics Press, 2013.

FOWLER, Martin. **UML distilled: a brief guide to the standard object modeling language**. 3. ed. Boston: Addison-Wesley, 2003.

HEVNER, A. R. et al. Design science in information systems research. **MIS Quarterly**, Minneapolis, v. 28, n. 1, p. 75–105, 2004.

HUMMER, W. et al. A model-driven approach for document-oriented RESTful APIs. In: IEEE INTERNATIONAL CONFERENCE ON SERVICE-ORIENTED COMPUTING AND APPLICATIONS, 6., 2013. **Proceedings [...]**. Piscataway: IEEE, 2013. p. 247–254.

IBGE. **Censo Demográfico 2022: resultados do universo**. Rio de Janeiro: IBGE, 2023. Disponível em: <https://censo2022.ibge.gov.br>. Acesso em: 11 abr. 2026.

ISO/IEC. **ISO/IEC 25010:2011: systems and software engineering: systems and software quality requirements and evaluation (SQuaRE): system and software quality models**. Geneva: ISO, 2011.

ISO/IEC/IEEE. **ISO/IEC/IEEE 29148:2018: systems and software engineering: life cycle processes: requirements engineering**. Geneva: ISO, 2018.

ISOTANI, Seiji; BITTENCOURT, Ig Ibert. **Dados abertos conectados**. São Paulo: Novatec, 2015.

KIMBALL, Ralph; CASERTA, Joe. **The data warehouse ETL toolkit: practical techniques for extracting, cleaning, conforming, and delivering data**. Indianapolis: Wiley, 2004.

KIMBALL, Ralph; ROSS, Margy. **The data warehouse toolkit: the definitive guide to dimensional modeling**. 3. ed. Indianapolis: Wiley, 2013.

LIMA, C. R. A. et al. Revisão das dimensões de qualidade dos dados e métodos aplicados na avaliação dos sistemas de informação em saúde. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 25, n. 10, p. 2095–2109, 2009.

MARCOTTE, Ethan. **Responsive web design**. New York: A Book Apart, 2011.

MENDES, Eugênio Vilaça. **As redes de atenção à saúde**. 2. ed. Brasília, DF: Organização Pan-Americana da Saúde, 2011.

MEYER, Eric A.; WEYL, Estelle. **CSS: the definitive guide**. 4. ed. Sebastopol: O'Reilly Media, 2017.

OGP. **Brazil member page**. Washington, DC, 2024. Disponível em: <https://www.opengovpartnership.org/members/brazil>. Acesso em: 10 jun. 2026.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. **Adolescent health**. Geneva: WHO, 2023. Disponível em: <https://www.who.int/health-topics/adolescent-health>. Acesso em: 10 jun. 2026.

PAIM, J. et al. The Brazilian health system: history, advances, and challenges. **The Lancet**, London, v. 377, n. 9779, p. 1778–1797, 2011.

PEREIRA, Rafael H. M.; GONÇALVES, Caio Nogueira. **geobr: download official spatial data sets of Brazil**. R package version 1.9.1. 2024. Disponível em: <https://CRAN.R-project.org/package=geobr>. Acesso em: 14 jun. 2026.

R CORE TEAM. **R: a language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2024. Disponível em: <https://www.r-project.org>. Acesso em: 20 jun. 2026.

SALDANHA, R. de F.; BASTOS, R. R.; BARCELLOS, C. Microdatasus: pacote para download e pré-processamento de microdados do DataSUS no software R. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 35, n. 7, e00032419, 2019.

SHNEIDERMAN, Ben. The eyes have it: a task by data type taxonomy for information visualizations. In: IEEE SYMPOSIUM ON VISUAL LANGUAGES, 1996, Boulder. **Proceedings [...]**. Washington: IEEE, 1996. p. 336–343.

SOCIEDADE BRASILEIRA DE PEDIATRIA. **Guia prático de atualização: limites de idade para atendimento em pediatria**. Rio de Janeiro: SBP, 2020.

SOMMERVILLE, Ian. **Software engineering**. 10. ed. Harlow: Pearson, 2016.

TUFTE, Edward R. **The visual display of quantitative information**. 2. ed. Cheshire: Graphics Press, 2001.

USHEY, Kevin; WICKHAM, Hadley. **renv: project environments for R**. Version 1.1.8. 2024. Disponível em: <https://rstudio.github.io/renv/>. Acesso em: 23 jun. 2026.

VASSILIADIS, Panos. A survey of extract-transform-load technology. **International Journal of Data Warehousing and Mining**, Hershey, v. 5, n. 3, p. 1–27, 2009.

WICKHAM, Hadley. **Advanced R**. 2. ed. Boca Raton: Chapman and Hall/CRC, 2019.

WICKHAM, Hadley. **Mastering Shiny**. Sebastopol: O'Reilly Media, 2021.

Apêndice A — Glossário, Siglas e Abreviações

API	<i>Application Programming Interface</i> — interface de comunicação entre sistemas de software
CAT1 / CAT2 / CAT3	Categorias da classificação tripartite de procedimentos pediátricos adotada neste trabalho
C1 / C2	Camadas hierárquicas do modelo de seleção de estabelecimentos adotado neste trabalho
CNES	Cadastro Nacional de Estabelecimentos de Saúde
CSS	<i>Cascading Style Sheets</i> — linguagem de estilização de interfaces web
CSV	<i>Comma-Separated Values</i> — formato de arquivo de texto para dados tabulares
DATASUS	Departamento de Informática do SUS
DOM	<i>Document Object Model</i> — representação estruturada de documentos HTML em memória
ECA	Estatuto da Criança e do Adolescente (Lei n. 8.069/1990)
ETL	<i>Extract, Transform, Load</i> — processo de extração, transformação e carga de dados
FTP	<i>File Transfer Protocol</i> — protocolo de transferência de arquivos em redes de computadores
HTML	<i>HyperText Markup Language</i> — linguagem de marcação para estruturação de páginas web
HTTPS	<i>HyperText Transfer Protocol Secure</i> — extensão segura do protocolo HTTP
IBGE	Instituto Brasileiro de Geografia e Estatística
JSON	<i>JavaScript Object Notation</i> — formato leve de intercâmbio de dados estruturados

KPI	<i>Key Performance Indicator</i> — indicador-chave de desempenho
LAI	Lei de Acesso à Informação (Lei n. 12.527/2011)
OGP	<i>Open Government Partnership</i> — Parceria para Governo Aberto
OMS	Organização Mundial da Saúde
OPM	Órteses, Próteses e Materiais Especiais
PNAISC	Política Nacional de Atenção Integral à Saúde da Criança
RDS	Formato nativo de serialização de objetos da linguagem R (<i>R Data Serialization</i>)
SAS	Secretaria de Atenção à Saúde
SBP	Sociedade Brasileira de Pediatria
SIDRA	Sistema IBGE de Recuperação Automática
SIGTAP	Sistema de Gerenciamento da Tabela de Procedimentos, Medicamentos e OPM do SUS
SPA	<i>Single Page Application</i> — aplicação web de página única
SSL	<i>Secure Sockets Layer</i> — protocolo de segurança para comunicação em redes
SUS	Sistema Único de Saúde
UF	Unidade Federativa
UML	<i>Unified Modeling Language</i> — linguagem de modelagem unificada para sistemas de software
URL	<i>Uniform Resource Locator</i> — endereço de recurso em rede
UTI	Unidade de Terapia Intensiva

YAML *YAML Ain't Markup Language* — formato de serialização de dados legível por humanos, utilizado para arquivos de configuração

ZIP Formato de compressão e empacotamento de arquivos