

Estado da publicação: O preprint não foi publicado em outro meio.

Indicadores para o SINAEP: Aderência dos dados abertos governamentais às dimensões avaliativas

Anderson Luiz Oliveira, Denise Maria Martins, Myllena Araujo Alves Pereira

<https://doi.org/10.1590/SciELOPreprints.17217>

Submetido em: 2026-08-19

Postado em: 2026-08-20 (versão 1)

(AAAA-MM-DD)

A moderação deste preprint recebeu o(s) endosso(s) de:

- Rosália Prados (ORCID: <https://orcid.org/0000-0003-2138-8422>)

ARTIGO

INDICADORES PARA O SINAEP: ADERÊNCIA DOS DADOS ABERTOS GOVERNAMENTAIS ÀS DIMENSÕES AVALIATIVAS

ANDERSON LUIZ DE OLIVEIRA¹

ORCID: <https://orcid.org/0009-0009-6084-6853>

<anderson.oliveira@cps.sp.gov.br>

DENISE MARIA MARTINS²

ORCID: <https://orcid.org/0000-0003-2956-0573>

<denise.martins@cps.sp.gov.br>

MYLLENA ARAUJO ALVES PEREIRA³

ORCID: <https://orcid.org/0009-0005-7162-451X>

<myllena.pereira@aluno.cps.sp.gov.br>

¹ Centro Estadual de Educação Tecnológica Paula Souza. São Paulo, SP, Brasil.

² Centro Estadual de Educação Tecnológica Paula Souza. São Paulo, SP, Brasil.

³ Centro Estadual de Educação Tecnológica Paula Souza. São Paulo, SP, Brasil.

RESUMO: O Decreto nº 12.603/2025 e a Portaria MEC nº 524/2026 instituíram o SINAEP^T, sistema nacional de avaliação da Educação Profissional e Tecnológica (EPT), com cinco dimensões avaliativas: condições de oferta; fluxo e rendimento; avaliação diagnóstica de competências e habilidades; articulação com as demandas do mundo do trabalho; e acompanhamento dos egressos. Como sua operacionalização depende de matrizes de indicadores ainda em construção pelo INEP, este estudo tem como objetivo compreender, a partir da articulação entre essas dimensões e o modelo CIPP (Contexto, Insumo, Processo e Produto), o grau e a natureza da aderência dos indicadores disponíveis nas bases de dados abertos governamentais a essas dimensões. Trata-se de pesquisa documental qualitativa, com análise de conteúdo categorial dedutiva sobre bases educacionais (INEP e MEC) e do mundo do trabalho (MTE e IBGE). Cada indicador foi situado na dimensão correspondente pela grade SINAEP^T × CIPP e avaliado, de forma dicotômica, pelos eixos de adequação conceitual, disponibilidade informacional e viabilidade técnica. Os 21 indicadores revelam aderência decrescente: plena nas condições de oferta e frágil na articulação com o trabalho e no acompanhamento de egressos. As lacunas não são de ausência de dados brutos, mas de integração entre bases e de significado dos dados, tornando-a prioritária para a implementação do sistema.

Palavras-chave: educação profissional e tecnológica, avaliação educacional, SINAEP^T, dados abertos governamentais, pesquisa documental qualitativa.

INDICATORS FOR SINAEP: ALIGNMENT OF OPEN GOVERNMENT DATA WITH EVALUATIVE DIMENSIONS

ABSTRACT: Decree No. 12,603/2025 and MEC Ordinance No. 524/2026 established SINAEP, the national evaluation system for Vocational and Technological Education (EPT), with five evaluative dimensions: supply conditions; flow and performance; diagnostic assessment of skills and abilities; alignment with labor market demands; and graduate tracking. As its operationalization depends on indicator matrices currently under development by INEP, this study aims to understand—based on the articulation between these dimensions and the CIPP model (Context, Input, Process, and Product)—the degree and nature of adherence of indicators available in open government databases to these dimensions. This is qualitative documentary research using deductive categorical content analysis on educational databases (INEP and MEC) and labor market databases (MTE and IBGE). Each indicator was assigned to its corresponding dimension via the SINAEP $\times$ CIPP grid and evaluated dichotomously across three axes: conceptual adequacy, informational availability, and technical feasibility. The 21 indicators reveal decreasing adherence: full in supply conditions, and weak in labor market alignment and graduate tracking. The gaps do not stem from a lack of raw data, but from database integration and data meaning, making integration a priority for the system's implementation.

Keywords: vocational and technological education, educational evaluation, SINAEP, open government data, qualitative documentary research.

INDICADORES PARA EL SINAEP: ADECUACIÓN DE LOS DATOS ABIERTOS GUBERNAMENTALES A LAS DIMENSIONES EVALUATIVAS

RESUMEN: El Decreto N.º 12.603/2025 y la Ordenanza MEC N.º 524/2026 instituyeron el SINAEP, sistema nacional de evaluación de la Educación Profesional y Tecnológica (EPT), con cinco dimensiones evaluativas: condiciones de oferta; flujo y rendimiento; evaluación diagnóstica de competencias y habilidades; articulación con las demandas del mundo laboral; y seguimiento de egresados. Como su operacionalización depende de matrices de indicadores aún en construcción por parte del INEP, este estudio busca comprender, mediante la articulación entre estas dimensiones y el modelo CIPP (Contexto, Insumo, Proceso y Producto), el grado y la naturaleza de la adherencia de los indicadores disponibles en las bases de datos abiertas gubernamentales a dichas dimensiones. Se trata de una investigación documental cualitativa, con análisis de contenido categorial deductivo sobre bases educativas (INEP y MEC) y del mundo laboral (MTE e IBGE). Cada indicador se situó en la dimensión correspondiente mediante la matriz SINAEP $\times$ CIPP y se evaluó, de forma dicotómica, según los ejes de adecuación conceptual, disponibilidad informacional y viabilidad técnica. Los 21 indicadores revelan una adherencia decreciente: plena en las condiciones de oferta, y frágil en la articulación con el trabajo y el seguimiento de egresados. Las brechas no radican en la ausencia de datos brutos, sino en la integración entre bases y el significado de estos, prioridad para la implementación del sistema.

Palabras clave: educación profesional y tecnológica, evaluación educativa, SINAEP, datos abiertos gubernamentales, investigación documental cualitativa.

INTRODUÇÃO

A Educação Profissional e Tecnológica (EPT) ocupa posição estratégica no cenário educacional brasileiro, articulando formação para o trabalho, elevação da escolaridade e desenvolvimento dos territórios em que se realiza. Nos últimos anos, a modalidade passou por expansão expressiva, com crescimento contínuo das matrículas registrado pelo Censo Escolar, impulsionado sobretudo pela integração entre ensino médio e formação técnica, e esse crescimento ampliou a necessidade de instrumentos capazes de aferir a qualidade da oferta, historicamente ausentes na modalidade.

Diferentemente da educação superior, avaliada desde 2004 pelo Sistema Nacional de Avaliação da Educação Superior (SINAES), e da educação básica, acompanhada por instrumentos consolidados como o IDEB, a EPT não dispôs, por décadas, de um sistema nacional de avaliação próprio. Essa lacuna começou a ser preenchida com o Decreto nº 12.603, de 28 de agosto de 2025, que instituiu a Política Nacional de Educação Profissional e Tecnológica (PNEPT) e criou o Sistema Nacional de Avaliação da Educação Profissional e Tecnológica (SINAEPT). A operacionalização do sistema foi regulamentada pela Portaria MEC nº 524, de 2 de junho de 2026, que definiu cinco dimensões avaliativas, detalhadas na fundamentação teórica, e atribuiu ao Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) a competência de regulamentar as matrizes de referência, os indicadores e os procedimentos de coleta, com implementação gradual prevista para até dois anos.

A construção dessas matrizes coloca uma questão que antecede a definição dos indicadores: saber se o sistema brasileiro de dados já produz a informação necessária para alimentá-los. A avaliação da EPT depende de informações que extrapolam o sistema educacional, abrangendo dados sobre o mundo do trabalho, a inserção profissional dos egressos e as demandas produtivas dos territórios, dispersos entre diferentes órgãos e bases públicas, com graus variados de acessibilidade e de compatibilidade entre si. Enquanto os esforços já em curso no INEP se concentram no desenvolvimento de novos instrumentos de avaliação direta, sobretudo para a mensuração de competências e habilidades dos estudantes (INEP, 2025), há espaço para um exame complementar, voltado ao aproveitamento da informação já existente nas bases de dados abertos governamentais.

Diante desse contexto, a questão que orienta este estudo é: em que medida as bases de dados abertos governamentais brasileiras oferecem indicadores aderentes às dimensões avaliativas estabelecidas pelo SINAEPT? Para respondê-la, o objetivo geral é compreender, a partir da articulação entre essas cinco dimensões e o modelo CIPP (Contexto, Insumo, Processo e Produto), o grau e a natureza da aderência dos indicadores disponíveis a essas dimensões.

O estudo contribui em dois planos que se articulam. No plano prático, oferece ao processo de regulamentação do SINAEPT um diagnóstico compreensivo da base informacional disponível para cada uma das cinco dimensões avaliativas, identificando onde o sistema de dados já sustenta indicadores, onde persistem lacunas e, sobretudo, qual é a natureza dessas lacunas, informação que interessa diretamente à definição das matrizes em curso no INEP e à priorização dos investimentos em integração de bases. No plano metodológico, demonstra um percurso de diagnóstico informacional reproduzível com recursos acessíveis a qualquer instituição pública, baseado na consulta direta a portais oficiais de dados abertos, o que o torna adequado às condições de instituições com recursos limitados e replicável à medida que o marco regulatório evoluir.

FUNDAMENTAÇÃO TEÓRICA

O referencial que fundamenta esta análise articula o marco normativo do SINAEPT e a especificidade avaliativa da EPT; o modelo CIPP e a tradição dos modelos lógicos como estrutura analítica; a governança da informação e os dados abertos governamentais como infraestrutura da avaliação; e o risco da avaliação por proxy, que orienta a questão central do estudo.

O SINAEPT e a especificidade avaliativa da EPT

A literatura da área já apontava a ausência de um sistema nacional de avaliação como lacuna histórica da EPT: Allain, Gruber e Wollinger (2020) observavam que a educação profissional não dispunha de sistema próprio, o que dificultava a formulação de políticas públicas baseadas em evidências. Mais recentemente, o próprio INEP tem desenvolvido trabalho metodológico voltado à operacionalização do SINAEPT, incluindo a elaboração de fundamentação teórico-metodológica para as matrizes de avaliação e a realização de pesquisas-piloto voltadas à avaliação direta de habilidades e do desempenho de estudantes concluintes (INEP, 2025). O presente estudo situa-se de forma complementar a esse esforço: enquanto as iniciativas do INEP se concentram na criação de novos instrumentos, este trabalho examina em que medida os dados já existentes no sistema brasileiro de dados abertos são aderentes às dimensões que o SINAEPT estabelece.

A EPT organiza-se em torno de princípios que a distinguem das demais etapas do sistema educacional brasileiro: a formação integral do trabalhador, que articula as dimensões técnica, científica, cultural e política do processo formativo (Frigotto; Ciavatta; Ramos, 2012); a politecnia, voltada ao domínio dos fundamentos científicos das diferentes técnicas que caracterizam o processo de trabalho produtivo, e não ao adestramento em uma ocupação específica (Saviani, 2003); e a indissociabilidade entre teoria e prática, que impede a separação entre o conhecimento produzido em sala de aula e a experiência no mundo do trabalho (Frigotto; Ciavatta; Ramos, 2012). Esses princípios situam a EPT em diálogo permanente com os arranjos produtivos dos territórios onde se realiza.

O próprio Decreto nº 12.603/2025 já enumerava, em seu art. 19, as dimensões a serem consideradas pelo sistema (Brasil, 2025). A Portaria MEC nº 524/2026 as detalhou em cinco dimensões avaliativas (Brasil, 2026, art. 5º), com remissão expressa ao art. 19 do Decreto e correspondência integral de conteúdo entre os dois atos; a única variação substantiva é a explicitação, pela Portaria, do âmbito regional na dimensão de articulação com o mundo do trabalho, que o Decreto referia apenas aos âmbitos local e nacional. São elas: (I) as condições institucionais de oferta, compreendendo a organização didático-pedagógica dos cursos, o corpo docente e técnico e a infraestrutura e demais recursos disponíveis; (II) as estatísticas de oferta, fluxo e rendimento educacional, com foco na permanência e na conclusão dos cursos; (III) a avaliação diagnóstica dos conhecimentos, das competências e das habilidades práticas desenvolvidas; (IV) a articulação da oferta de EPT com as demandas do mundo do trabalho nos âmbitos local, regional e nacional; e (V) o acompanhamento da inserção dos egressos no mundo do trabalho e da continuidade dos estudos.

Lidas à luz dos princípios que fundamentam a modalidade, essas dimensões correspondem a perguntas sobre a especificidade formativa da EPT: as dimensões I e III questionam se as condições institucionais e os processos formativos efetivamente viabilizam a politecnia e a indissociabilidade entre

teoria e prática; as dimensões IV e V perguntam se essa formação se traduz em articulação real com os arranjos produtivos e em inserção profissional dos egressos (Brasil, 2025, art. 4º, V). Ao longo deste estudo, as cinco dimensões do SINAEP^T constituem o organizador primário da análise, e o modelo CIPP, apresentado a seguir, funciona como lente teórica que as ordena em uma cadeia avaliativa.

Modelos de avaliação institucional: CIPP e os modelos lógicos

Entre os modelos disponíveis para estruturar uma avaliação institucional multidimensional, o modelo CIPP (Contexto, Insumo, Processo e Produto), formalizado por Stufflebeam (1971, 2003), oferece uma arquitetura especialmente adequada às cinco dimensões do SINAEP^T. Concebido como instrumento de apoio à tomada de decisão em programas educacionais, o CIPP propõe que a avaliação examine, de forma articulada, as necessidades e o contexto que justificam um programa (Contexto), os recursos mobilizados para sua execução (Insumo), a forma como as atividades planejadas são realizadas (Processo) e os resultados produzidos (Produto). Essa lógica não é exclusiva do CIPP: na pesquisa sobre eficácia escolar, Scheerens (1990) propôs modelo semelhante para orientar a construção de indicadores educacionais, posteriormente consolidado por Scheerens e Bosker (1997) e adotado, com variações, por sistemas internacionais de indicadores.

A escolha do CIPP como lente de leitura das dimensões do SINAEP^T tem uma razão técnica específica: a dimensão IV, de articulação com as demandas do mundo do trabalho, comporta duas leituras temporais distintas, a das demandas que antecedem e orientam a oferta (leitura *ex ante*, correspondente ao Contexto) e a da articulação que resulta da formação realizada (leitura *ex post*, correspondente ao Produto). O CIPP fornece o vocabulário que torna essa dupla leitura explícita, evitando que informações de naturezas distintas sejam tratadas como equivalentes. Cabe reconhecer, contudo, uma limitação: o modelo pressupõe uma sequência relativamente linear entre contexto, insumo, processo e produto, ao passo que a EPT envolve retroalimentação contínua entre formação e mundo do trabalho; o desdobramento da dimensão IV em dois momentos atenua parcialmente essa linearidade.

Uma segunda tradição, com origem nos modelos lógicos de avaliação de programas (W. K. Kellogg Foundation, 2004; Wholey, 1979), descreve a cadeia causal de um programa como sequência de insumos, atividades, produtos, resultados e impactos, distinguindo com precisão o resultado, entendido como o efeito direto sobre os beneficiários, do impacto, entendido como a transformação mais ampla e de longo prazo nos territórios. Essa distinção é adotada neste estudo para precisar o sentido da categoria Produto do CIPP, interpretada como o resultado da formação sobre os egressos.

Governança da informação e dados abertos governamentais

A digitalização das atividades governamentais transformou a produção e a circulação de informações públicas, e os dados abertos governamentais passaram a ocupar posição central nas agendas de transparência e modernização administrativa. Bonina e Eaton (2020), em análise comparativa de plataformas latino-americanas de dados abertos, identificam que a sustentabilidade dessas iniciativas depende da governança institucional, da padronização técnica e do engajamento de múltiplos atores. Gao, Janssen e Zhang (2023) argumentam que a simples expansão quantitativa da disponibilidade de dados não basta: é necessária uma transformação estrutural das infraestruturas informacionais, que inclu

processos de organização, contextualização e governança capazes de viabilizar o uso estratégico dos dados.

Para a EPT, a questão assume contornos particulares, pois a modalidade depende de informações que transcendem o sistema educacional: dados sobre o mercado de trabalho, os arranjos produtivos locais e as demandas setoriais são insumos essenciais para indicadores que articulem formação e trabalho, exigência que decorre diretamente dos princípios de indissociabilidade entre teoria e prática e de politecnia. O Brasil dispõe de um conjunto amplo de bases públicas potencialmente relevantes para essa finalidade, detalhadas no percurso metodológico. No caso específico da Rede Federal, a Plataforma Nilo Peçanha foi instituída em resposta a questões de ordem técnica, teórica e metodológica que comprometiam a confiabilidade das estatísticas produzidas a partir do SISTEC, processo que envolveu a revisão semântica das categorias utilizadas e a remodelagem dos indicadores acadêmicos e de gestão (Brasil, 2020). O episódio ilustra, no interior da própria EPT, a centralidade da governança da informação.

O uso conjunto dessas bases depende, ainda, de interoperabilidade. Segundo a OCDE (2023), a interoperabilidade em bases públicas de educação compreende quatro dimensões (semântica, técnica, organizacional e legal), e sua ausência torna o cruzamento de dados propenso a erros e custoso. Para a EPT, essa interoperabilidade significaria, por exemplo, cruzar os dados do Censo Escolar, que identifica cursos técnicos por eixo tecnológico, com os da RAIS, que classifica ocupações pelo Código Brasileiro de Ocupações (CBO). Três circunstâncias descritas nesta literatura são particularmente relevantes para a análise que se segue, pois oferecem as bases teóricas das explicações possíveis para eventuais lacunas informacionais: a exigência de interoperabilidade semântica e de identificadores comuns entre bases (OCDE, 2023), a dependência histórica das infraestruturas de dados em relação às finalidades para as quais foram criadas (Gao; Janssen; Zhang, 2023) e a fragmentação da governança entre múltiplos entes e redes (Bonina; Eaton, 2020), que no caso brasileiro se expressa na distribuição da oferta de EPT entre redes federal, estaduais e privadas, cada uma com sistemas de coleta próprios.

Estatísticas públicas, indicadores e o problema da avaliação por *proxy*

Conforme a distinção proposta por Jannuzzi (2017), dados são registros elementares de observações; estatísticas públicas resultam da organização desses registros em descrições quantitativas; indicadores resultam da seleção e interpretação de medidas à luz de um referencial teórico; e índices agregam múltiplos indicadores em medidas sintéticas. Essa diferenciação é relevante porque a construção de um indicador envolve escolhas teóricas e normativas: não existe indicador neutro. Um dos riscos centrais apontados pelo autor é tratar estatísticas públicas como substitutas diretas de indicadores sociais, sem a devida fundamentação conceitual. Quando isso ocorre, registros concebidos para uma finalidade são reinterpretados para responder a questões para as quais não foram projetados, configurando a avaliação por *proxy*, isto é, o uso de medidas indiretas para representar fenômenos cuja observação direta é mais complexa.

Esse problema assume contornos particulares na EPT. A produção de estatísticas educacionais no Brasil foi historicamente estruturada para atender à educação básica e superior, de modo que indicadores amplamente disponíveis, como o desempenho em avaliações padronizadas, o fluxo escolar e a distorção idade-série, foram concebidos para representar fenômenos próprios dessas etapas, não da EPT. Bowers e Choi (2023) observam que a ausência de padrões de governança de dados educacionais, na linha dos princípios FAIR (dados encontráveis, acessíveis, interoperáveis e reutilizáveis),

compromete a equidade no reuso desses dados, privilegiando o que é facilmente mensurável em detrimento do que seria mais relevante avaliar. Schildkamp, Lai e Earl (2013) acrescentam que a interpretação de dados fora de seu contexto original de produção distorce não apenas o diagnóstico, mas também as decisões de política dele derivadas.

Essa literatura fornece o fundamento teórico dos três eixos interpretativos operacionalizados no percurso metodológico: a exigência de correspondência entre a medida e o construto avaliativo fundamenta o eixo da adequação conceitual (Jannuzzi, 2017); a exigência de que os dados sejam encontráveis e acessíveis em formatos reutilizáveis fundamenta o eixo da disponibilidade informacional (Bowers; Choi, 2023); e a exigência de granularidade, cobertura e interoperabilidade entre bases fundamenta o eixo da viabilidade técnica (OCDE, 2023). É dessa mesma literatura, portanto, que deriva a tipologia de lacunas utilizada na análise, na qual cada lacuna é tipificada pelo eixo que o indicador não satisfaz.

MÉTODOS E PROCEDIMENTOS

O desenho metodológico desta pesquisa fundamenta-se na articulação entre o marco normativo do SINAEP e o modelo CIPP, estruturando-se em um fluxo operacional que compreende três etapas sucessivas de identificação exploratória, verificação documental e análise categorial.

Delineamento da pesquisa

Este estudo caracteriza-se como pesquisa aplicada, de natureza documental e abordagem qualitativa, cujo delineamento combina duas tradições metodológicas complementares. Da análise documental (Cellard, 2008; Lüdke; André, 2013), o estudo toma o tratamento de documentos como fontes que exigem exame crítico de seu contexto de produção: as bases de dados abertos, seus dicionários de variáveis, suas notas metodológicas e os atos normativos do SINAEP são tratados como documentos, produzidos por instituições específicas, com finalidades específicas, cuja adequação a um novo uso, a avaliação da EPT, não pode ser presumida. Na análise de conteúdo, o estudo adota a abordagem categorial dedutiva proposta por Mayring (2014), caracterizada pela definição a priori das categorias analíticas: o corpus, composto exclusivamente por documentos oficiais e bases de dados públicas de acesso aberto, sem qualquer coleta de dados junto a pessoas, foi analisado a partir de uma grade normativo-teórica construída com base na articulação entre as cinco dimensões avaliativas do SINAEP e as quatro categorias do modelo CIPP, que orientou a classificação e a interpretação de cada unidade de registro.

Em conformidade com a Política de Integridade na Atividade Científica do CNPq (CNPq, 2026), declara-se o uso de ferramentas de Inteligência Artificial generativa nesta pesquisa, com a especificação da ferramenta, da fase e da finalidade: na etapa de identificação exploratória, utilizou-se modelo de linguagem de grande porte de código aberto executado localmente, descrito adiante; na fase de elaboração do manuscrito, utilizou-se o assistente Claude (Anthropic) como apoio à redação, à revisão e à organização do texto. Em todos os casos, a concepção da pesquisa, a definição do modelo analítico, a categorização dos indicadores e a interpretação dos resultados são de responsabilidade integral dos autores, que verificaram a procedência das informações e referências apresentadas.

Unidade de análise e corpus documental

A unidade de análise é o indicador potencial: cada variável que, em princípio, poderia compor um indicador de uma das cinco dimensões do SINAEP¹ e que seja passível de ser derivada, integral ou parcialmente, das bases de dados abertos consultadas. Na terminologia da análise de conteúdo, o indicador potencial constitui a unidade de registro do estudo. Não constituem unidades de análise as instituições de ensino, os cursos ou os estudantes: o foco recai sobre a informação disponível sobre eles.

O corpus é constituído pelos atos normativos que instituem e regulamentam o SINAEP¹, pela documentação oficial das bases de dados abertos governamentais, isto é, dicionários de variáveis, notas metodológicas, guias de referência e páginas de disseminação nos portais oficiais, e pelas próprias estruturas de dados disponibilizadas (tabelas, microdados e extratores). As bases consultadas, sintetizadas no Quadro 1, abrangem tanto o sistema educacional (INEP e MEC) quanto o mundo do trabalho (Ministério do Trabalho e Emprego e IBGE), abrangência exigida pelas dimensões IV e V do SINAEP¹.

Quadro 1 - Bases de dados governamentais consultadas.

BASE/PORTAL	ÓRGÃO RESPONSÁVEL	TIPO DE INFORMAÇÃO	PERIODICIDADE	VIA DE ACESSO ABERTA
Censo Escolar da Educação Básica	INEP/MEC	Matrículas, infraestrutura, corpo docente e gestão escolar, com identificação de etapas e modalidades (incluindo cursos técnicos/EPT)	Anual	Microdados e sinopses no portal do INEP
Indicadores Educacionais	INEP/MEC	Adequação da formação docente, esforço docente, complexidade de gestão, regularidade do corpo docente, distorção idade-série, taxas de rendimento	Anual	Planilhas no portal do INEP
IDEB	INEP/MEC	Indicador de qualidade que combina fluxo escolar e proficiência (SAEB)	Bienal	Planilhas no portal do INEP
ENEM (microdados)	INEP/MEC	Desempenho e perfil socioeconômico dos participantes	Anual	Microdados no portal do INEP
Plataforma Nilo Peçanha (PNP)	MEC/SETEC	Estatísticas oficiais e indicadores de gestão da Rede Federal de EPT	Anual	Módulo Extrator e portal de dados abertos do MEC
Novo CAGED	MTE	Admissões e desligamentos formais, por município, ocupação (CBO) e setor	Mensal	Microdados via PDET/MTE
RAIS	MTE	Vínculos formais de emprego por ocupação (CBO), setor e município	Anual	Microdados via PDET/MTE
PNAD Contínua	IBGE	Taxas de desocupação e de participação na força de trabalho, com desagregação por UF	Mensal/Trimestral	API SIDRA e interface web do SIDRA

PNAD Contínua, módulo Educação	IBGE	Frequência e conclusão de educação profissional, por sexo, cor/raça e UF	Anual	API SIDRA e interface web do SIDRA
CEMPRE (Cadastro Central de Empresas)	IBGE	Empresas, unidades locais, pessoal ocupado e salários, por atividade econômica (CNAE) e território	Anual	API SIDRA e interface web do SIDRA
Censo da Educação Superior	INEP/MEC	Matrículas e ingressantes no ensino superior	Anual	Microdados no portal do INEP

Fonte: Elaborado pelos autores (2026).

As consultas à documentação e às estruturas de dados das bases relacionadas no Quadro 1 foram realizadas entre 4 de abril e 15 de julho de 2026. O registro pormenorizado de cada consulta (edição ou ano-base da fonte aberta e data de acesso) integra a planilha de categorização disponibilizada no repositório do estudo. O levantamento restringiu-se a dados disponibilizados em formato estruturado e processável por máquina, não abrangendo informações publicadas apenas em formatos não estruturados, como relatórios em PDF ou painéis interativos sem opção de exportação; as implicações dessa delimitação são retomadas nas considerações finais. Cabe registrar que o Censo Escolar, principal fonte dos indicadores identificados, é de declaração obrigatória para todos os estabelecimentos de educação básica do país, o que confere cobertura censitária à oferta de EPT nele registrada.

Categorias analíticas, eixos interpretativos e tipologias

A grade de categorias definidas a priori resulta da articulação entre as cinco dimensões do SINAEP e as quatro categorias do CIPP (Quadro 2). Nela, as dimensões do SINAEP constituem o organizador primário, por serem o objeto normativo cuja base informacional se examina, e as categorias do CIPP funcionam como lente teórica que as ordena na cadeia avaliativa; a dimensão IV é desdobrada em duas leituras, *ex ante* (Contexto) e *ex post* (Produto), conforme fundamentado na revisão teórica.

Quadro 2 - Articulação entre as dimensões do SINAEP e as categorias do modelo CIPP.

DIMENSÃO DO SINAEP (PORTARIA Nº 524/2026, ART. 5º)	CATEGORIA CIPP	FOCO DA ARTICULAÇÃO
I. Condições institucionais de oferta	Insumo	Organização didático-pedagógica, corpo docente e técnico, infraestrutura e demais recursos
II. Estatísticas de oferta, fluxo e rendimento	Processo	Permanência e conclusão dos cursos
III. Avaliação diagnóstica de conhecimentos, competências e habilidades	Processo	Desenvolvimento de competências e habilidades práticas nos cursos
IV. Articulação com as demandas do mundo do trabalho	Contexto (leitura <i>ex ante</i>) e Produto (leitura <i>ex post</i>)	Demandas que orientam a oferta; articulação que resulta da formação
V. Acompanhamento da inserção dos egressos e da continuidade dos estudos	Produto/Resultado	Trajetória dos egressos no mundo do trabalho e nos estudos

Fonte: Elaborado pelos autores (2026). Nota: os indicadores alocados a cada dimensão são examinados, sem exceção, pelos três eixos interpretativos operacionalizados no Quadro 3.

Para os fins deste estudo, entende-se por aderência a propriedade de um indicador potencial satisfazer, simultaneamente, três requisitos, operacionalizados em três eixos interpretativos: a adequação conceitual, a disponibilidade informacional e a viabilidade técnica. Cumpre distinguir: os eixos não se confundem com as dimensões avaliativas do SINAEP; as dimensões definem o que se avalia, e os eixos definem como se julga a aptidão de cada indicador para servi-las. O Quadro 3 sintetiza a operacionalização de cada eixo, explicitando a passagem entre seu fundamento teórico, apresentado na revisão de literatura, e sua aplicação como critério de avaliação: para cada eixo, uma pergunta operacional, a evidência documental que a responde e a regra de decisão correspondente.

Quadro 3 - Operacionalização dos eixos interpretativos.

EIXO	FUNDAMENTO TEÓRICO	PERGUNTA OPERACIONAL	EVIDÊNCIA CONSULTADA	REGRAS DE DECISÃO (DICOTÔMICA)
Adequação conceitual	Correspondência entre medida e construto; risco da avaliação por <i>proxy</i> (JANNUZZI, 2017)	A informação mensura diretamente o construto do foco da dimensão (Quadro 2) ou apenas fenômeno correlato?	Definição oficial da variável no dicionário ou na nota metodológica, confrontada com o texto da dimensão (Brasil, 2026, art. 5º)	Satisfeito quando a variável expressa diretamente o construto; não satisfeito quando constitui aproximação (<i>proxy</i>)
Disponibilidade informacional	Princípios FAIR de encontrabilidade e acessibilidade (Bowers; Choi, 2023)	Os dados estão publicados em formato estruturado e processável por máquina, localizável no portal oficial?	Página de disseminação do órgão produtor e arquivos abertos (microdados, tabelas, extratores)	Satisfeito quando o acesso estruturado é confirmado na fonte; não satisfeito quando inexistente ou restrito a formatos não estruturados
Viabilidade técnica	Dimensões da interoperabilidade em bases educacionais (OCDE, 2023)	A granularidade, a cobertura, a periodicidade e as possibilidades de cruzamento permitem calcular o indicador para a EPT?	Estrutura dos arquivos abertos (colunas, filtros e desagregações) e identificadores documentados	Satisfeito quando o cálculo é possível para a EPT com os dados atuais; não satisfeito quando falta desagregação, cobertura ou vínculo entre bases

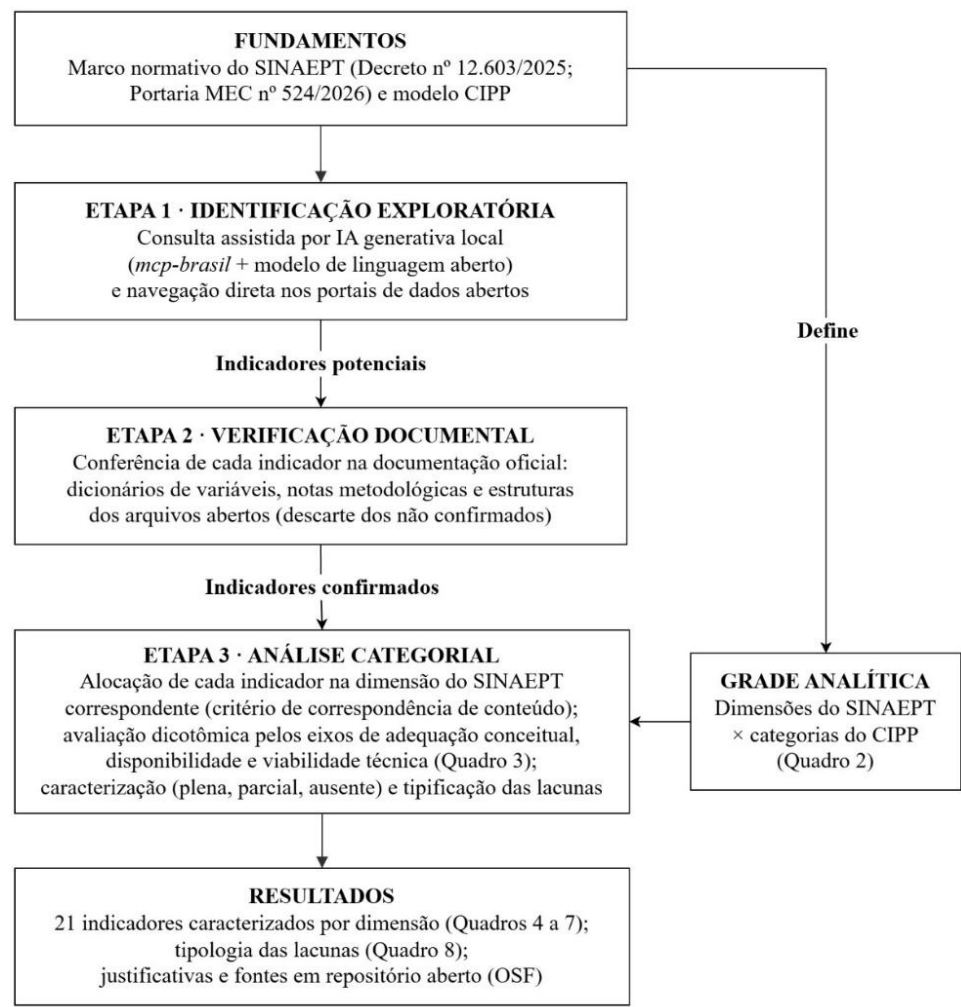
Fonte: Elaborado pelos autores (2026), com base em Jannuzzi (2017), Bowers e Choi (2023) e OCDE (2023).

A avaliação de cada indicador foi, portanto, dicotômica em cada eixo, satisfeito ou não satisfeito, respondida exclusivamente com base na evidência documental indicada no Quadro 3 e registrada, com a respectiva justificativa, na planilha de categorização do repositório do estudo. A caracterização da aderência deriva diretamente dessa avaliação: plena quando os três eixos são satisfeitos, parcial quando dois o são e ausente quando apenas um ou nenhum o é. Essas categorias constituem uma tipologia interpretativa, e não uma escala de mensuração, e os eixos não foram hierarquizados entre si, por representarem requisitos de naturezas distintas e não substituíveis. Para os indicadores parciais, a lacuna é tipificada pelo eixo único não satisfeito; para os ausentes, pelo eixo cuja não satisfação é determinante para a inviabilidade do indicador. Os quadros de resultados explicitam, para cada indicador, o veredito de cada eixo e a caracterização de aderência dele resultante.

Etapas e procedimentos

O exame do material seguiu três etapas, representadas na Figura 1.

Figura 1 - Fluxo metodológico da pesquisa.



Fonte: Elaborado pelos autores (2026).

Etapa 1 - Identificação exploratória: A identificação inicial dos indicadores potenciais foi realizada entre abril e junho de 2026, com apoio de um instrumento de consulta assistida por Inteligência Artificial generativa executado integralmente em ambiente local: o servidor *mcp-brasil* (versão 0.14.0; protocolo *Model Context Protocol*, MCP; Anthropic, 2024), conectado ao modelo de linguagem de código aberto Qwen2.5-7B-Instruct (quantização Q4_K_M), executado via LM Studio (versão 0.4.16). A opção por ferramentas locais, gratuitas e de código aberto reduz barreiras de custo à reprodução do percurso e assegura que nenhum dado seja processado por servidores externos. A identificação exploratória foi complementada pela navegação direta nos catálogos dos portais oficiais de dados abertos (INEP, MEC, dados.gov.br, PDET/MTE e SIDRA/IBGE).

Etapa 2 - Verificação documental sistemática: Como modelos de linguagem podem produzir respostas textuais que não correspondem ao conteúdo efetivo das fontes, situação observada em testes preliminares, nenhum indicador foi caracterizado a partir do retorno do instrumento exploratório. Cada indicador potencial foi verificado diretamente na documentação oficial de sua base de origem: a página de disseminação no portal do órgão produtor, o dicionário de variáveis ou nota metodológica correspondente e, quando aplicável, a estrutura dos arquivos abertos (colunas, filtros e níveis de desagregação disponíveis). Indicadores cuja existência ou estrutura não pôde ser confirmada na fonte oficial foram descartados.

Etapa 3 - Análise categorial: Cada indicador verificado foi alocado na dimensão do SINAEP correspondente mediante um critério de correspondência de conteúdo: comparou-se o construto informacional do indicador (o que a variável efetivamente mensura, conforme sua documentação oficial) com o foco de cada dimensão descrito no Quadro 2, registrando-se a justificativa da alocação; indicadores relativos às demandas do mundo do trabalho foram alocados à leitura *ex ante* da dimensão IV, e indicadores relativos a resultados da formação, à leitura *ex post* da dimensão IV ou à dimensão V. Em seguida, cada indicador foi examinado segundo os três eixos interpretativos e caracterizado como de aderência plena, parcial ou ausente, com registro da justificativa interpretativa de cada juízo e tipificação da lacuna, quando existente. Os resultados dessa análise são apresentados na sequência das dimensões do SINAEP, com o veredito de cada eixo explicitado por indicador nos respectivos quadros e com a discussão de cada dimensão percorrendo os três eixos à luz da fundamentação teórica. Para conferir confiabilidade ao percurso interpretativo, o estudo mobilizou o registro sistemático das justificativas, ancoradas na documentação oficial verificada, e a auditabilidade pública do percurso: o registro das fontes consultadas estão disponíveis em <https://doi.org/10.17605/OSF.IO/KVAPC>

RESULTADOS E DISCUSSÃO

Os 21 indicadores potenciais caracterizados são discutidos a seguir, organizados conforme a sequência das dimensões do SINAEP, com um quadro síntese para cada etapa da análise. Em cada dimensão, os resultados são interpretados à luz da fundamentação teórica; ao final, uma tipologia consolidada sintetiza a natureza das lacunas identificadas.

Dimensão I: condições institucionais de oferta

Quadro 4 - Indicadores potenciais da dimensão I.

DIM. SINAEP	CIPP	BASE DE ORIGEM	#	INDICADOR POTENCIAL	AC	DI	VT	ADERÊNCIA	LACUNA
I	Insumo	Censo Escolar	1	Evolução da matrícula de EPT por rede	S	S	S	Plena	—
I	Insumo	Censo Escolar	2	Matrícula de EPT por dependência administrativa	S	S	S	Plena	—
I	Insumo	Censo Escolar	3	Matrícula de EPT por etapa/modalidade	S	S	S	Plena	—
I	Insumo	Censo Escolar	4	Proporção de matrícula vocacional no ensino médio regular, por UF	S	S	S	Plena	—
I	Insumo	Censo Escolar	5	Matrícula de EPT técnica por eixo tecnológico	S	S	S	Plena	—
I	Insumo	Indicadores Educacionais (INEP)	6	Adequação da Formação Docente	S	S	N	Parcial	Viabilidade técnica
I	Insumo	Indicadores Educacionais (INEP)	7	Esforço Docente	S	S	N	Parcial	Viabilidade técnica
I	Insumo	Indicadores Educacionais (INEP)	8	Complexidade de Gestão da Escola	S	S	N	Parcial	Viabilidade técnica

I	Insumo	Indicadores Educacionais (INEP)	9	Regularidade do Corpo Docente	S	S	N	Parcial	Viabilidade técnica
---	--------	---------------------------------	---	-------------------------------	---	---	---	---------	---------------------

Fonte: Elaborado pelos autores (2026). Nota: dimensão I: condições institucionais de oferta. AC: adequação conceitual; DI: disponibilidade informacional; VT: viabilidade técnica; S: eixo satisfeito; N: eixo não satisfeito, conforme os critérios do Quadro 3

A dimensão I é a mais bem servida pelo sistema de dados abertos: cinco dos nove indicadores apresentam aderência plena, todos correspondentes a estatísticas específicas de educação profissional do Censo Escolar, informação censitária, específica da modalidade e disponível em formato estruturado. Os quatro Indicadores Educacionais do INEP relativos a corpo docente e gestão, conceitualmente pertinentes à dimensão, foram caracterizados como parciais por lacuna de viabilidade técnica: a verificação de suas estruturas de disseminação (ano-base 2025) confirmou a ausência de desagregação por modalidade, e a busca sistemática pelos termos "profissional" e "técnico" nos blocos de variáveis dessas planilhas não retornou ocorrências, de modo que os docentes de disciplinas técnicas permanecem agregados às categorias gerais do ensino médio ou da educação básica. O padrão confirma, para a EPT, a observação de Bowers e Choi (2023) de que as infraestruturas de dados educacionais privilegiam o que é facilmente mensurável: a oferta é visível; as condições específicas da modalidade, não.

Dimensões II e III: fluxo, rendimento e avaliação diagnóstica

Quadro 5 - Indicadores potenciais das dimensões II e III.

DIM. SINAEP	CIPP	BASE DE ORIGEM	#	INDICADOR POTENCIAL	AC	DI	VT	ADERÊNCIA	LACUNA
II	Processo	IDEB (INEP)	10	IDEB do Ensino Médio	S	S	S	Plena	—
II	Processo	Indicadores Educacionais (INEP)	11	Taxa de Distorção Idade-Série	S	S	N	Parcial	Viabilidade técnica
II	Processo	Indicadores Educacionais (INEP)	12	Taxas de Rendimento Escolar	S	S	N	Parcial	Viabilidade técnica
III	Processo	Microdados do ENEM (INEP)	13	Desempenho em competências (<i>proxy</i>)	N	S	S	Parcial	Adequação conceitual

Fonte: Elaborado pelos autores (2026). Nota: dimensão II: estatísticas de oferta, fluxo e rendimento educacional; dimensão III: avaliação diagnóstica de conhecimentos, competências e habilidades práticas. AC: adequação conceitual; DI: disponibilidade informacional; VT: viabilidade técnica; S: eixo satisfeito; N: eixo não satisfeito, conforme os critérios do Quadro 3.

Nas dimensões de processo, apenas o IDEB apresenta aderência plena. Os demais indicadores limitam-se por motivos de natureza distinta, e a diferença é analiticamente relevante. A Taxa de Distorção Idade-Série e as Taxas de Rendimento Escolar apresentam lacuna de viabilidade técnica: a única desagregação que tangencia a EPT em suas estruturas é residual, restrita às colunas de quarta série do ensino médio (e, nas taxas de rendimento, de turmas não seriadas), que captam apenas os cursos técnicos integrados de quatro anos e deixam invisíveis os cursos de três anos e os subsequentes. Já os microdados do ENEM, utilizados como *proxy* de avaliação de competências para a dimensão III,

apresentam lacuna de adequação conceitual: mensuram desempenho acadêmico geral, não as habilidades práticas que a dimensão prevê, configurando exatamente o uso de estatística pública como substituta de indicador que Jannuzzi (2017) adverte exigir fundamentação conceitual. No primeiro caso, o dado é adequado, mas sua forma de divulgação o inviabiliza; no segundo, o dado é disponível e utilizável, mas não expressa o construto.

Dimensão IV, leitura *ex ante*: as demandas do mundo do trabalho

Quadro 6 - Indicadores potenciais da dimensão IV em sua leitura *ex ante*.

DIM. SINAEPT	CIPP	BASE DE ORIGEM	#	INDICADOR POTENCIAL	AC	DI	VT	ADERÊNCIA	LACUNA
IV (<i>ex ante</i>)	Contexto	PNAD Contínua (IBGE/SIDRA)	14	Taxa de desocupação e de participação na força de trabalho	N	S	S	Parcial	Adequação conceitual
IV (<i>ex ante</i>)	Contexto	PNAD Contínua (IBGE/SIDRA)	15	Condições da força de trabalho com desagregação territorial (UF)	N	S	S	Parcial	Adequação conceitual
IV (<i>ex ante</i>)	Contexto	PNAD Contínua, módulo Educação (IBGE/SIDRA)	16	Frequência e conclusão de educação profissional na população, por sexo, cor/raça e UF	N	S	S	Parcial	Adequação conceitual
IV (<i>ex ante</i>)	Contexto	CEMPRE (IBGE/SIDRA)	17	Estrutura produtiva territorial: empresas, pessoal ocupado e salários por CNAE e território	N	S	S	Parcial	Adequação conceitual

Fonte: Elaborado pelos autores (2026). Nota: dimensão IV: articulação da oferta de EPT com as demandas do mundo do trabalho, aqui em sua leitura *ex ante*. AC: adequação conceitual; DI: disponibilidade informacional; VT: viabilidade técnica; S: eixo satisfeito; N: eixo não satisfeito, conforme os critérios do Quadro 3.

A leitura *ex ante* da dimensão IV dispõe de aproximações múltiplas, porém nenhuma que expresse diretamente seu construto: os quatro indicadores foram caracterizados como parciais, todos por lacuna de adequação conceitual. As taxas de desocupação e de participação, mesmo com desagregação por unidade federativa, constituem aproximações macroeconômicas sem correspondência direta com as demandas setoriais por formação. O módulo Educação da PNAD Contínua descreve o alcance da educação profissional na população (quem a frequenta e quem a conclui, por sexo, cor/raça e UF), informação valiosa para caracterizar a demanda social atendida, mas que não expressa as demandas do mundo do trabalho; seu enquadramento nesta leitura, e não nas dimensões de processo, decorre da natureza da fonte, uma pesquisa amostral domiciliar externa às instituições de ensino, que retrata o

cenário demográfico-territorial da qualificação, não a operação dos cursos. O CEMPRE é o candidato conceitualmente mais próximo, ao descrever a estrutura produtiva por atividade econômica e território, aproximando-se da noção de arranjos produtivos locais; ainda assim, descreve a estrutura instalada, não a demanda por formação que dela decorreria. O conjunto ilustra o mecanismo da avaliação por *proxy* discutido a partir de Jannuzzi (2017): a dimensão que deveria orientar a oferta segundo os princípios da politécnica e da articulação territorial (Saviani, 2003; Frigotto; Ciavatta; Ramos, 2012) é justamente aquela em que só há medidas de fenômenos correlatos.

Dimensões IV, leitura *ex post*, e V: articulação realizada e trajetória dos egressos

Quadro 7 - Indicadores potenciais da dimensão IV em sua leitura *ex post* e da dimensão V.

DIM. SINAEP	CIPP	BASE DE ORIGEM	#	INDICADOR POTENCIAL	AC	DI	VT	ADERÊNCIA	LACUNA
IV (<i>ex post</i>)	Produto/Resultado	Novo CAGED (PDET/MTE)	18	Saldo e movimentação de empregos formais, por município e ocupação (CBO)	N	S	S	Parcial	Adequação conceitual
V	Produto/Resultado	Plataforma Nilo Peçanha (SETEC/MEC)	19	Indicadores de gestão e trajetória da Rede Federal de EPT	S	S	N	Parcial	Viabilidade técnica
IV-V	Produto/Resultado	RAIS (PDET/MTE)	20	Inserção de egressos por ocupação (CBO) e curso	N	S	N	Ausente	Viabilidade técnica
V	Produto/Resultado	Censo da Educação Superior (INEP)	21	Continuidade dos estudos de egressos no ensino superior	N	S	N	Ausente	Viabilidade técnica

Fonte: Elaborado pelos autores (2026). Nota: dimensão IV: articulação da oferta de EPT com as demandas do mundo do trabalho, aqui em sua leitura *ex post*; dimensão V: acompanhamento da inserção dos egressos no mundo do trabalho e da continuidade dos estudos. AC: adequação conceitual; DI: disponibilidade informacional; VT: viabilidade técnica; S: eixo satisfeito; N: eixo não satisfeito, conforme os critérios do Quadro 3.

Nas dimensões que expressam o resultado da formação, nenhum indicador alcançou aderência plena, e é aqui que a fragmentação informacional se mostra mais severa. Os indicadores de gestão da Rede Federal, disponíveis de forma estruturada na Plataforma Nilo Peçanha, foram caracterizados como parciais por lacuna de viabilidade técnica: os próprios metadados da base declaram abrangência restrita às unidades da Rede Federal, que responde por parcela minoritária da matrícula nacional de EPT, o que impede sua generalização como indicador do sistema; a PNP ilustra simultaneamente a viabilidade de um sistema estatístico específico para a EPT e o limite de sua fragmentação, ao não alcançar as demais redes, expressão concreta da governança fragmentada descrita por Bonina e Eaton (2020). O saldo de empregos formais do Novo CAGED, disponível em microdados

desagregados por município e ocupação, limita-se conceitualmente: descreve a dinâmica geral do emprego, não a trajetória dos egressos.

Os dois indicadores que expressariam diretamente o construto das dimensões, a inserção de egressos por ocupação e curso e a continuidade dos estudos no ensino superior, foram caracterizados como ausentes. A verificação dos leiautes oficiais tornou essa impossibilidade concreta: no leiaute público da RAIS não há identificadores pessoais, e a escolaridade do trabalhador é registrada em doze categorias genéricas, nas quais o egresso técnico é indistinguível sob o rótulo de ensino médio completo, sem qualquer campo que registre o curso ou o eixo tecnológico de formação; no caso dos censos do INEP, os identificadores pseudonimizados de estudantes utilizados pela educação básica e pela educação superior não são publicamente vinculáveis entre si. O indicador é concebível, as bases existem, mas o cálculo é impossível com os dados abertos atuais, materializando a ausência das dimensões semântica e técnica da interoperabilidade descritas pela OCDE (2023).

Tipologia das lacunas e o padrão de conjunto

Quadro 8 - Tipologia das lacunas identificadas entre os quinze indicadores parciais ou ausentes.

TIPO DE LACUNA	CARACTERIZAÇÃO	OCORRÊNCIA	IMPLICAÇÕES
Adequação conceitual	Dado disponível e tecnicamente viável, mas que representa apenas <i>proxy</i> do construto avaliativo	Dimensão III (um), dimensão IV <i>ex ante</i> (quatro) e dimensão IV <i>ex post</i> (um), em um total de seis indicadores	Os indicadores mais acessíveis medem fenômenos correlatos, não os construtos das dimensões
Viabilidade técnica	Dado disponível e conceitualmente pertinente, mas com desagregação, cobertura ou interoperabilidade insuficientes	Dimensão I (quatro), dimensão II (dois) e dimensões IV–V (três), em um total de nove indicadores	A EPT permanece invisível no interior de estatísticas gerais, e as bases não se vinculam entre si
Disponibilidade	Dado não acessível em formato estruturado	Nenhum indicador, após a verificação documental nas fontes oficiais	O problema informacional do SINAEP não é de ausência de dados brutos, mas de integração e de significado

Fonte: Elaborado pelos autores (2026).

Dos 21 indicadores, seis apresentam aderência plena, treze aderência parcial e dois estão ausentes, em distribuição que configura um gradiente decrescente ao longo das dimensões: a dimensão I concentra toda a aderência plena de base censitária, as dimensões II e III sustentam-se parcialmente e as dimensões IV e V, que expressariam a especificidade da EPT como formação articulada ao trabalho, são as mais frágeis. A constatação mais significativa do Quadro 8, contudo, é negativa: nenhuma lacuna remanescente é de disponibilidade. Na etapa exploratória, três indicadores haviam sido tidos como indisponíveis por não terem sido localizados pelo instrumento de consulta assistida por IA; a verificação documental nas fontes oficiais reverteu esse juízo, pois os dados da Plataforma Nilo Peçanha estão disponíveis de forma estruturada e os microdados da RAIS e do Censo da Educação Superior são públicos. O que parecia ausência era, em parte, invisibilidade ao instrumento: os dados existiam, mas não eram encontráveis pelas vias de acesso programático consultadas, o que materializa a dimensão da encontrabilidade dos princípios FAIR (Bowers; Choi, 2023) e ilustra o argumento de Gao, Janssen e

Zhang (2023) de que a expansão da disponibilidade de dados não substitui a transformação estrutural das infraestruturas informacionais.

Desse padrão decorrem três proposições interpretativas sobre a persistência das lacunas, ancoradas nas bases teóricas apresentadas na revisão de literatura e que este estudo, de natureza documental, registra como agenda a investigar, e não como conclusões. A primeira, de ordem técnica, é que a integração entre bases educacionais e do mundo do trabalho exige interoperabilidade semântica, isto é, um mapeamento entre eixos tecnológicos e ocupações que hoje não existe, e identificadores comuns compatíveis com a proteção de dados pessoais (OCDE, 2023). A segunda, de ordem histórico-institucional, é que a produção de estatísticas educacionais foi estruturada para a educação básica e superior, de modo que o acompanhamento de egressos da EPT nunca constituiu prioridade dos sistemas de informação existentes (GAO; JANSSEN; ZHANG, 2023). A terceira, de ordem federativa, é que a distribuição da oferta entre redes federal, estaduais e privadas, cada uma com sistemas e categorias próprios, dificulta a constituição de um fluxo nacional padronizado de informações (BONINA; EATON, 2020).

CONSIDERAÇÕES FINAIS

Este estudo teve como objetivo compreender, a partir da articulação entre as cinco dimensões avaliativas do SINAEP e o modelo CIPP, o grau e a natureza da aderência dos indicadores disponíveis nas bases de dados abertos governamentais brasileiras a essas dimensões. A análise dos 21 indicadores potenciais respondeu a esse objetivo em dois níveis. Quanto ao grau, a aderência distribui-se em gradiente decrescente ao longo das dimensões: o sistema brasileiro de dados abertos descreve bem o que a EPT oferta (dimensão I), descreve de forma limitada como essa oferta se realiza (dimensões II e III) e quase não descreve o que a antecede (as demandas do mundo do trabalho) nem o que dela resulta (a trajetória dos egressos). Quanto à natureza, as lacunas identificadas não são de ausência de dados brutos, e sim de integração, quando a informação existe mas está diluída em estatísticas gerais ou fragmentada em bases que não dialogam, e de significado, quando a informação disponível constitui apenas aproximação do construto avaliativo.

A principal contribuição do estudo consiste em oferecer um diagnóstico sistemático da aderência dos indicadores disponíveis às dimensões do SINAEP, qualificando as lacunas informacionais da EPT como problemas predominantemente relacionados à interoperabilidade entre bases e à adequação conceitual dos dados. Esse diagnóstico desloca o debate da mera produção de novas informações para a articulação e o uso mais consistente das bases existentes, além de evidenciar os riscos de se privilegiar, na avaliação, indicadores facilmente mensuráveis, mas pouco representativos dos construtos avaliativos. No plano metodológico, o estudo propõe um percurso de diagnóstico informacional estruturado por uma grade analítica definida a priori a partir da articulação entre as dimensões do SINAEP e as categorias do modelo CIPP. Esse percurso combina a identificação exploratória apoiada por Inteligência Artificial generativa local com a verificação documental sistemática em fontes oficiais, configurando um procedimento auditável, de baixo custo e potencialmente replicável por outras instituições e em diferentes contextos regulatórios.

Dessas conclusões decorrem recomendações diretas ao processo de regulamentação em curso: que a definição das matrizes reconheça explicitamente a assimetria informacional entre as dimensões, evitando que a maior facilidade de mensuração da oferta conduza a um sistema que avalie o

que é mensurável em vez do que é relevante; que se priorize a construção de mecanismos de integração entre as bases educacionais e as do mundo do trabalho, cujos requisitos o diagnóstico identifica (o mapeamento semântico entre eixos tecnológicos e ocupações e um vínculo de trajetórias compatível com a proteção de dados pessoais), ainda que a avaliação de soluções técnicas para atendê-los extrapole o escopo deste estudo; e que a interoperabilidade entre os sistemas das diferentes redes de oferta seja considerada desde o início, sob pena de se produzirem indicadores nacionais não comparáveis.

Os resultados devem ser interpretados à luz das limitações do estudo. O corpus restringiu-se a dados disponibilizados em formato estruturado, excluindo fontes relevantes sem essa forma de acesso, como sistemas estaduais de informação educacional. A verificação da viabilidade técnica baseou-se na documentação oficial e na estrutura divulgada de cada base, sem o processamento exaustivo dos microdados, e não contemplou aspectos relacionados à qualidade das informações, como subnotificações e defasagens de atualização. Além disso, os critérios de classificação dos indicadores foram fundamentados em julgamento analítico apoiado na documentação oficial e, embora registrados de forma auditável no repositório da pesquisa, não estão isentos de subjetividade, podendo ser reinterpretados em futuras aplicações da grade analítica. Essas limitações orientam a agenda de pesquisas futuras, que inclui o cálculo empírico dos indicadores classificados como plenamente aderentes, a investigação das proposições interpretativas junto aos gestores dos órgãos produtores de dados, a replicação do diagnóstico após a divulgação das matrizes oficiais do SINAEP pelo INEP e a incorporação de fontes não estruturadas ao corpus.

REFERÊNCIAS

ALLAIN, Olivier; GRUBER, Cristhian; WOLLINGER, Pedro Renato. *O que avaliar em educação profissional? Princípios epistemológicos da formação de trabalhadores*. In: MORAES, Gustavo Henrique; ALBUQUERQUE, Ana Elizabeth Maia de; SANTOS, Robson dos; SILVA, Susiane de Santana Moreira Oliveira da (org.). *Avaliação da educação profissional e tecnológica: um campo em construção*. Brasília, DF: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2020. p. 19-42.

ANTHROPIC. *Introducing the Model Context Protocol*. [S. l.]: Anthropic, 2024. Disponível em: <https://www.anthropic.com/news/model-context-protocol>. Acesso em: 12/05/2026.

BONINA, Carla; EATON, Ben. Cultivating open government data platform ecosystems through governance: Lessons from Buenos Aires, Mexico City and Montevideo. *Government Information Quarterly*, v. 37, n. 3, art. 101479, 2020. <https://doi.org/10.1016/j.giq.2020.101479>

BOWERS, Alex J.; CHOI, Yujin. Building school data equity, infrastructure, and capacity through FAIR data standards: Findable, accessible, interoperable, and reusable. *Educational Researcher*, v. 52, n. 7, p. 450–458, 2023.

BRASIL. Decreto nº 12.603, de 28 de agosto de 2025. Institui a Política Nacional de Educação Profissional e Tecnológica e cria o Sistema Nacional de Avaliação da Educação Profissional e Tecnológica. *Diário Oficial da União*, Brasília, DF, 29 ago. 2025.

BRASIL. Ministério da Educação. Portaria MEC nº 524, de 2 de junho de 2026. Regulamenta o Sistema Nacional de Avaliação da Educação Profissional e Tecnológica (SINAEP). *Diário Oficial da União*, Brasília, DF, n. 103, Seção 1, p. 37–38, 3 jun. 2026.

BRASIL. Ministério da Educação. Secretaria de Educação Profissional e Tecnológica. *Plataforma Nilo Peçanha*: guia de referência metodológica: PNP 2020. Brasília, DF: MEC, 2020. Disponível em: https://dadosabertos.mec.gov.br/images/conteudo/pnp/2020/grm_pnp_2020.pdf. Acesso em: 08/04/2026.

CELLARD, André. *A análise documental*. In: POUPART, Jean; DESLAURIERS, Jean-Pierre; GROULX, Lionel-Henri; LAPERRIÈRE, Anne; MAYER, Robert; PIRES, Álvaro P. *A pesquisa qualitativa: enfoques epistemológicos e metodológicos*. Petrópolis: Vozes, 2008. p. 295-316.

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO. Portaria CNPq nº 2.664, de 6 de março de 2026. Institui a Política de Integridade na Atividade Científica do CNPq. *Diário Oficial da União*, Brasília, DF, Seção 1, p. 4, 11 mar. 2026.

FRIGOTTO, Gaudêncio; CIAVATTA, Maria; RAMOS, Marise (Orgs.). *Ensino médio integrado: concepção e contradições*. 3. ed. São Paulo: Cortez, 2012.

GAO, Yuzhuo; JANSSEN, Marijn; ZHANG, Cheng. Understanding the evolution of open government data research: Towards open data sustainability and smartness. *International Review of Administrative Sciences*, v. 89, n. 1, p. 59–75, 2023.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. *Fundamentação teórico-metodológica das matrizes de referência*: avaliação da educação profissional e tecnológica. Brasília, DF: INEP, 2025. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/avaliacao-da-educacao-profissional-e-tecnologica/matrizes-de-referencia>. Acesso em: 03/06/2026.

JANNUZZI, Paulo de Martino. *Indicadores sociais no Brasil*: conceitos, fontes de dados e aplicações. 6. ed. Campinas: Alínea, 2017.

LÜDKE, Menga; ANDRÉ, Marli E. D. A. *Pesquisa em educação*: abordagens qualitativas. 2. ed. São Paulo: EPU, 2013.

MAYRING, Philipp. *Qualitative content analysis*: Theoretical foundation, basic procedures and software solution. Klagenfurt, 2014. Disponível em: <https://nbn-resolving.org/urn:nbn:de:0168-ssolar-395173>. Acesso em: 25/04/2026.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *OECD digital education outlook 2023*: Towards an effective digital education ecosystem. Paris: OECD Publishing, 2023. <https://doi.org/10.1787/c74f03de-en>

SAVIANI, Dermeval. O choque teórico da politecnicidade. *Trabalho, Educação e Saúde*, v. 1, n. 1, p. 131–152, 2003. <https://doi.org/10.1590/S1981-77462003000100010>

SCHEERENS, Jaap. School effectiveness research and the development of process indicators of school functioning. *School Effectiveness and School Improvement*, v. 1, n. 1, p. 61–80, 1990. <https://doi.org/10.1080/0924345900010106>

SCHEERENS, Jaap; BOSKER, Roel J. *The foundations of educational effectiveness*. Oxford: Pergamon, 1997.

SCHILDKAMP, Kim; LAI, Mei Kuin; EARL, Lorna (Eds.). *Data-based decision making in education: Challenges and opportunities*. Dordrecht: Springer, 2013.

STUFFLEBEAM, Daniel L. *The CIPP model for evaluation*. In: KELLAGHAN, Thomas; STUFFLEBEAM, Daniel L. (Eds.). *International handbook of educational evaluation*. Dordrecht: Kluwer Academic Publishers, 2003. p. 31–62. https://doi.org/10.1007/978-94-010-0309-4_4

STUFFLEBEAM, Daniel L.; FOLEY, Walter J.; GEPHART, William J.; GUBA, Egon G.; HAMMOND, Robert L.; MERRIMAN, Howard O.; PROVUS, Malcolm M. *Educational evaluation and decision making*. Itasca: F. E. Peacock, 1971.

W. K. KELLOGG FOUNDATION. *Logic model development guide*. Battle Creek: W. K. Kellogg Foundation, 2004.

WHOLEY, Joseph S. *Evaluation: Promise and performance*. Washington, DC: The Urban Institute, 1979.

Submetido: 27/07/2026

Preprint: Não se aplica

Aprovado: XX/XX/20XX

Editor de seção:

DECLARAÇÃO SOBRE DISPONIBILIDADE DE DADOS

O artigo disponibiliza, no subtópico "Etapas e procedimentos", o link para um repositório anônimo que contém a cópia dos documentos analisados e o histórico das interações (chats) extraídas do software LM Studio. O material pode ser acessado em: <https://doi.org/10.17605/OSF.IO/KVAPC>

CONTRIBUIÇÃO DE AUTORIA

Anderson Luiz de Oliveira: Conceitualização, curadoria de dados, metodologia, análise formal, investigação, escrita do rascunho original e escrita e edição (*Conceptualization, Data curation, Methodology, Formal analysis, Investigation, Writing - original draft*).

Denise Maria Martins: Conceitualização, metodologia, supervisão, validação e revisão (*Conceptualization, Supervision, Validation, Methodology, Writing - review & editing*).

Myllena Araujo Alves Pereira: Conceitualização, escrita e rascunho original (*Conceptualization, Writing - review & editing*).

DECLARAÇÃO DE CONFLITO DE INTERESSE

Os autores declaram que não há conflito de interesse com o presente artigo.

DECLARAÇÃO DE USO DE INTELIGÊNCIA ARTIFICIAL

Para a elaboração desse manuscrito foi utilizada o recurso de Inteligência Artificial Claude (Anthropic) para a realização da(s) seguinte (s) atividade(s): apoio à redação, à revisão e à organização do texto. O (a) autor (a) declara que todo o material derivado de tal processo foi revisado e o (a) autor (a) assume total responsabilidade por todo o conteúdo do manuscrito.

Este preprint foi submetido sob as seguintes condições:

- Os autores declaram que os necessários Termos de Consentimento Livre e Esclarecido de participantes ou pacientes na pesquisa foram obtidos e estão descritos no manuscrito, quando aplicável.
- Os autores declaram que a elaboração do manuscrito seguiu as normas éticas de comunicação científica.
- Os autores declaram que estão cientes que são os únicos responsáveis pelo conteúdo do preprint e que o depósito no SciELO Preprints não significa nenhum compromisso de parte do SciELO, exceto sua preservação e disseminação.
- Os autores declaram que os dados, aplicativos e outros conteúdos subjacentes ao manuscrito estão referenciados.
- O manuscrito depositado está no formato PDF.
- Os autores declaram que a pesquisa que deu origem ao manuscrito seguiu as boas práticas éticas e que as necessárias aprovações de comitês de ética de pesquisa, quando aplicável, estão descritas no manuscrito.
- Os autores declaram que uma vez que um manuscrito é postado no servidor SciELO Preprints, o mesmo só poderá ser retirado mediante pedido à Secretaria Editorial do SciELO Preprints, que afixará um aviso de retratação no seu lugar.
- Os autores concordam que o manuscrito aprovado será disponibilizado sob licença [Creative Commons CC-BY](#).
- O autor submissor declara que as contribuições de todos os autores e declaração de conflito de interesses estão incluídas de maneira explícita e em seções específicas do manuscrito.
- Os autores declaram que o manuscrito não foi depositado e/ou disponibilizado previamente em outro servidor de preprints ou publicado em um periódico.
- Caso o manuscrito esteja em processo de avaliação ou sendo preparado para publicação mas ainda não publicado por um periódico, os autores declaram que receberam autorização do periódico para realizar este depósito.
- O autor submissor declara que todos os autores do manuscrito concordam com a submissão ao SciELO Preprints.