

Maturidade para o Impacto (MPI): uma proposta de transformação avaliativa para programas de pós-graduação profissionais

Maturity for Impact (MPI): a proposal for an evaluative transformation of professional graduate programs

Andrea Viviana Waichman ¹, Eduardo Meireles ², Genildo Cavalcante Ferreira-Júnior³

¹ Universidade Federal do Amazonas, Manaus, AM, Brasil

² Universidade do Estado de Minas Gerais, Belo Horizonte, MG, Brasil

³Instituto Federal do Acre, Rio Branco, AC, Brasil

Resumo

Este artigo apresenta uma arquitetura completa para a transformação do modelo de avaliação dos Programas de Pós-Graduação Profissionais (PPs) no Brasil, com foco no conceito de Maturidade para o Impacto (MPI) e no Índice de Maturidade para o Impacto (IMI). Fundamentando-se na necessidade de calibrar indicadores de qualidade universal com a missão aplicada e orientada a problemas dos PPs, a proposta distingue-se da avaliação tradicional voltada para programas acadêmicos. O modelo organiza-se em três quesitos principais: – Governança e Conexão Estratégica; Desenvolvimento de Soluções e Formação de Talentos; e Impacto e Transformação do Território – e opera por meio do IMI, um instrumento que agrega seis critérios ponderados para avaliar a maturidade dos Produtos Técnico-Tecnológicos (PTTs). O artigo detalha a estrutura proposta, sua operacionalização, cronograma de implementação e estratégias de monitoramento e comunicação, posicionando-se como um complemento para a avaliação quadrienal da Capes que tem o propósito de alinhar incentivos para a geração de valor público.

Palavras-chave: Avaliação de Pós-Graduação; Mestrado Profissional; Impacto Socioeconômico.

Área Tecnológica: Prospecção Tecnológica de Assuntos Específicos.

Abstract

This article presents a comprehensive framework for transforming the evaluation model of Professional Graduate Programs (PPs) in Brazil, grounded in the concept of Maturity for Impact (MPI) and the Maturity for Impact Index (IMI). Based on the need to calibrate universal quality indicators with the applied, problem-oriented mission of PPs, this proposal differs from traditional evaluation models designed for academic programs. The framework is organized into three main criteria – Governance and Strategic Alignment; Solutions Development and Talent Training; and Impact and Territorial Transformation – and operates through the IMI, an instrument that aggregates six weighted criteria to assess the maturity of Technical-Technological Products (PTTs). The article details the structure of the proposed model, its operationalization, implementation timeline, and monitoring and communication strategies, positioning itself as a complementary mechanism to Capes' quadrennial evaluation, aimed at aligning incentives with the generation of public value.

Keywords: Graduate Evaluation; Professional Master's Degree; Socioeconomic Impact.

1 Introdução

A pós-graduação profissional brasileira ocupa um lugar estratégico na política de ciência, tecnologia e inovação ao articular conhecimento científico, experiência prática e demandas de setores produtivos, governamentais e sociais. Diferentemente dos programas acadêmicos, sua missão é orientada por problemas, pela formação de profissionais e pela entrega de soluções em formatos diversos, como protocolos, metodologias, *softwares*, processos, materiais educativos, normas técnicas, modelos de gestão e tecnologias sociais. Essa diversidade amplia o alcance social da pós-graduação, mas tensiona modelos tradicionais de avaliação, historicamente estruturados para captar produção bibliográfica e impacto acadêmico.

Nesse sentido, a avaliação quadrienal da Capes organiza-se por áreas, combinando dimensões de formação, produção e impactos, com ajustes de pesos e critérios por área para refletir especificidades de cada campo de conhecimento. O desafio, já reconhecido na própria análise de fichas, é equilibrar indicadores de qualidade universal com a missão distinta de programas acadêmicos e profissionais, missão que, no caso dos Programas de Pós-Graduação Profissionais (PPs), é aplicada e orientada a problemas (validação em contexto real, adoção por terceiros, valor público). Em diversos PPs, persiste a influência de métricas concebidas sobretudo para programas acadêmicos, como contagens de publicações e citações, foco excessivo em periódicos e baixa valorização de demonstrações em contexto real, validações por usuários e adoção em escala. Essa assimetria pode gerar incentivos desalinhados com a missão dos PPs, que é essencialmente profissionalizante, orientada por necessidades de setores específicos e pela geração de valor público. Ao mesmo tempo, o desafio institucional é legítimo: como calibrar diferenças entre modalidades e áreas, preservando padrões universais de qualidade, comparabilidade mínima e rastreabilidade.

A literatura recente sobre a avaliação da pós-graduação brasileira tem sido consistente ao apontar que o modelo Capes, tal como operacionalizado, tende a concentrar peso excessivo em métricas bibliométricas e de contagem, como estratificação de periódicos e volume de publicações, o que desloca o eixo da avaliação para aquilo que é mais facilmente mensurável, nem sempre para aquilo que é mais significativo em termos de impacto e relevância social (Barata, 2019, McManus *et al.*, 2022). Esse desenho não é neutro: ao transformar indicadores em alvo, ele passa a induzir condutas e reconfigurar prioridades institucionais, criando um ambiente em que o desempenho é confundido com produtividade e a melhoria de programas pode ser substituída pelo esforço de “adequação” aos parâmetros avaliativos (Patrus; Dantas; Shigaki, 2015).

As críticas mais frequentes se organizam em torno de três problemas estruturais. O primeiro é o produtivismo, associado à pressão por resultados de curto prazo, com sobrecarga docente e incentivo à priorização de quantidade em detrimento de densidade intelectual, fenômeno discutido em estudos de diferentes áreas e experiências institucionais. O segundo é a padronização, quando métricas de “tamanhos único” desconsideram culturas de publicação e formatos legítimos de produção, penalizando especialmente campos em que livros, produtos técnicos, artísticos e culturais são centrais, como enfatizam McManus, Baeta Neves e Carvalho (2023) ao tratarem da produção técnica e artística na pós-graduação. O terceiro é a reprodução de assimetrias institucionais e regionais, na medida em que critérios uniformes tendem a favorecer programas já consolidados e territórios com maior infraestrutura, aprofundando desigualdades e tensionando o papel público da pós-graduação (Oliveira; Rothen, 2022).

No conjunto, os autores convergem mais para uma agenda de recalibração do que para a rejeição da avaliação. McManus *et al.* (2022) defendem maior sensibilidade para impacto e contexto, integração e transparência dos dados e cronogramas mais claros, enquanto outros trabalhos apontam a necessidade de uma inflexão menos punitiva e mais formativa. Nessa direção, aparecem propostas como fortalecer componentes qualitativos e de julgamento por pares (Perez, 2025), incorporar métricas que capturem melhor a função formativa, como indicadores de treinamento e coautoria discente-docente (Gheno *et al.*, 2019), e reduzir o peso exclusivo de métricas bibliométricas, de modo a reequilibrar qualidade, diversidade de áreas e relevância social como dimensões avaliáveis

Esta proposta posiciona-se exatamente nesse ponto, o de preservar a qualidade e recalibrar a leitura de impacto para os PPGs Profissionais, evitando a importação acrítica de métricas concebidas para programas acadêmicos (exemplo: ênfase desproporcional em periódicos e citações) e valorizando transferência, adoção e alcance de soluções (serviços, processos, produtos, normas, dados, entre outros) na sociedade e no mercado. Na construção dos indicadores, foram adotados princípios de uso responsável de métricas (DORA, 2012; Hicks *et al.*, 2015; Santos, 2020), que recomendam diversidade de modelos de avaliação para respeitar missões distintas: (i) combinar indicadores com evidências auditáveis e narrativas contextualizadas; (ii) promover transparência de pesos, regras e algoritmos; (iii) evitar proxies simplistas (exemplo: fator de impacto) como fim em si; e (iv) proteger contra efeitos comportamentais (a “medida virar meta”). Assim, a sistemática proposta posiciona-se como um complemento para a avaliação quadrienal da Capes, explicando desde o início como a combinação de quesitos, pesos e fichas por área pode ser lida à luz da missão dos Programas Profissionais (PPs).

A proposta preserva indicadores de qualidade universal e recalibra sua interpretação para valorizar transferência, adoção e valor público. Em vez de proxies simplistas, propõe-se triangulação entre métricas, evidências auditáveis e narrativa contextual, com pesos ajustáveis por área e trilhas de auditoria para cada evidência. A agenda internacional de reforma da avaliação científica tem criticado o uso acrítico de indicadores simplificados. O *Leiden Manifesto*, por exemplo, recomenda que a avaliação apoie o julgamento qualitativo por especialistas, contextualize métricas, considere diferenças entre campos e seja transparente e auditável. A Declaração de São Francisco (DORA, 2012) reforça a necessidade de valorizar a diversidade de produtos de pesquisa e evitar a substituição da avaliação de mérito por métricas de periódicos. No Brasil, iniciativas como o Projeto Métricas da Universidade de São Paulo (USP) consolidaram discussões sobre uso responsável de indicadores e seus efeitos sobre a governança acadêmica.

Propôs-se o conceito de Maturidade para o Impacto (MPI) para avaliar, ao longo do tempo, em que medida o programa reúne processos, dados e parcerias capazes de gerar e de sustentar impacto. Avaliar a “maturidade para o impacto (MPI)” é a situação almejada, deslocando o eixo avaliativo para a maturidade das soluções criadas e para sua validação externa, com ênfase em resultados aplicados e valor público gerado. O conceito criado se assemelha à ideia de Societal Readiness Levels (SRL) (IFD, 2019; Bernstein *et al.*, 2022), mas mensura objetivos diferentes. Enquanto o SRL mensura quanto o ambiente social (atores, instituições, regras, valores e arranjos de uso) está pronto para adotar, aceitar e sustentar uma solução, o MPI pergunta se a solução criada reúne qualidade, validação e evidências suficientes para produzir valor público de forma consistente. Enquanto um foca na “disposição do mundo” para receber a inovação, o outro foca na “capacidade da inovação” de gerar efeitos materiais e verificáveis no mundo. O conceito de Maturidade para o Impacto (MPI) sugere um processo de desenvolvimento e aprimoramento que culmina na capacidade de gerar valor. Uma metodologia, um protocolo ou uma obra de arte também “amadurecem”. Não é medir a maturidade por si só, mas sim a maturidade com o objetivo de gerar impacto. O termo “impacto” é amplo o suficiente para englobar todas as dimensões de valor que a pós-graduação profissional busca (social, econômico, cultural, ambiental, científico etc.), sem se restringir ao viés econômico ou tecnológico.

Assim, a avaliação pautada pelo Quadro de Maturidade para o Impacto (MPI) e pelo Índice de Maturidade para o Impacto (IMI) busca valorizar a robustez das soluções desenvolvidas, sua validação externa e a capacidade de transformação territorial e setorial, em contraste com a preponderância de métricas tradicionalmente voltadas para a produção acadêmica pura. Com essa mudança, espera-se promover PTTs que realmente mudem práticas,

processos decisórios e políticas; formar egressos preparados para resolver desafios complexos; e estruturar redes com empresas, entes públicos e organizações sociais, assegurando a difusão de soluções que funcionam em situações reais.

Diante desse cenário, este artigo tem como objetivo apresentar uma proposta de sistemática avaliativa específica para Programas de Pós-Graduação Profissionais, centrada no conceito de Maturidade para o Impacto (MPI) e operacionalizada por meio do Índice de Maturidade para o Impacto (IMI). O MPI interpreta a trajetória do programa e sua capacidade de produzir, sustentar e comunicar impactos; o IMI qualifica a maturidade de impacto de cada produto técnico e tecnológico (PTT) e, por agregação, do portfólio do programa. Busca-se descrever os fundamentos, princípios orientadores, critérios e indicadores que compõem o modelo, demonstrando como essa abordagem pode aprimorar a mensuração da qualidade e dos resultados aplicados gerados pelos Programas Profissionais, além de contribuir com evidências para o alinhamento entre a missão dessas modalidades e seus mecanismos de avaliação institucional de modo a transformar a avaliação em instrumento de melhoria contínua.

2 Metodologia

A elaboração da proposta de avaliação partiu da identificação da necessidade de transformação do modelo atualmente adotado pela Capes para os Programas de Pós-Graduação Profissionais (PPs), no qual predominam métricas e indicadores orientados para a produção acadêmica tradicional. Assim, foram analisadas as diferenças entre a situação presente, caracterizada pela subvalorização de produtos técnico-tecnológicos e impacto aplicado, e a situação almejada, que demanda uma sistemática alinhada à natureza profissionalizante desses programas. O estudo é de natureza metodológica e propositiva, orientado para o desenho de um modelo de avaliação centrado em evidências. A construção do MPI/IMI seguiu três eixos complementares.

Eixo 1: referenciais de métricas responsáveis e avaliação de pesquisa. Foram adotados como pilares o Leiden Manifesto e a DORA, por explicitarem princípios de contextualização, transparência, triangulação e prevenção de incentivos distorcidos. Esse eixo também incorporou as recomendações do relatório *The Metric Tide* (Wilsdon *et al.*, 2015), que sintetiza boas práticas para uso de métricas em avaliação e gestão da pesquisa.

Eixo 2: análise de instrumentos avaliativos de programas profissionais. A partir de uma leitura comparativa das fichas e descritores de avaliação das 51 áreas de avaliação da Capes, buscou-se mapear onde se concentram assimetrias e quais elementos de qualidade podem ser tratados como universais. O foco recaiu sobre a

lógica em quesitos e subitens, pesos e evidências exigidas, com atenção às dimensões mais diretamente relacionadas à entrega de produtos e resultados aplicados.

Eixo 3: síntese do modelo e desenho operacional. O MPI foi concebido como um constructo longitudinal de maturidade, enquanto o IMI foi estruturado como índice sintético por PTT. Foram definidos critérios e uma escala ordinal de cinco níveis, descritores ajustáveis por área e pesos parametrizáveis. Para garantir rastreabilidade, desenhou-se um dossiê digital de evidências por PTT e um painel de indicadores no nível do programa. Por fim, propôs-se um roteiro de implantação em fases, com piloto, calibragem interavaliadores, auditorias amostrais e revisão iterativa.

A proposta é apresentada como um modelo de referência, passível de adaptação por área e por programa, com ênfase na transparência sobre escolhas de pesos, descritores e critérios de validade de evidências. Tais princípios asseguram que a avaliação seja capaz de reconhecer diferentes trajetórias de desenvolvimento institucional e mensurar o potencial de transformação de cada programa. Dessa forma, estruturou-se uma sistemática que desloca a ênfase da avaliação para indicadores relacionados ao impacto, à inovação aplicada e à formação de profissionais capazes de atuar como agentes de mudança em seus contextos de prática.

A sistemática proposta tem como intuito mensurar a qualidade e o impacto dos Programas Profissionais com base em sua vocação para a aplicação do conhecimento e a geração de soluções tangíveis, distanciando-se do modelo predominante de avaliação acadêmica. Como princípios fundamentais da nova sistemática, destacam-se os seguinte:

- 1) Relevância e impacto para a sociedade/mercado: o foco principal deve ser a capacidade do programa de gerar soluções que respondam a problemas reais e que produzam valor (econômico, social, ambiental, tecnológico).
- 2) Orientação para a aplicação e a inovação: valorizar o desenvolvimento e a transferência de produtos, processos e metodologias que sejam inovadores e aplicáveis.
- 3) Conexão permanente com o setor produtivo/social: reconhecer e incentivar a interação orgânica e estratégica com empresas, setor público, terceiro setor e comunidades.
- 4) Formação de profissionais agentes de transformação: avaliar a capacidade do programa de formar egressos que se tornem líderes e inovadores em seus campos de atuação.
- 5) Evidência e validação externa: priorizar indicadores que possam ser comprovados por validações de terceiros, dados de mercado ou relatos de impacto com metodologia robusta.

3 Resultados e Discussão

A sistemática de avaliação propõe três macro quesitos: Q1 – Governança e Conexão Estratégica; Q2 – Desenvolvimento de Soluções e Formação de Talentos; e Q3 – Impacto e Transformação do Território. A lógica de avaliação passa a espelhar a lógica de gestão: o que foi ensinado, orientado, demonstrado, validado e transferido é o que será medido. O modelo proposto combina ajustes estruturais (governança e currículo), operacionais (processo de PTT, coleta e guarda de evidências, parcerias e validações) e de monitoramento (IMI e painéis de dados). A seguir, estão descritos os indicadores que compõem cada quesito de avaliação (Quadro 1).

Quadro 1 – Quesito, indicador e exemplos de evidências de cada indicador

QUESITO	INDICADOR	EXEMPLOS DE EVIDÊNCIAS
Q1	Docentes com atuação no setor externo	Projetos P&D financiados; consultorias; conselhos; serviços tecnológicos.
Q1	Currículo alinhado/flexível	Atas de revisão curricular; ementas orientadas a projetos; trilhas modulares.
Q1	Infraestrutura para validação	Lista/estado de equipamentos; convênios para ensaios; <i>hubs</i> e incubadoras.
Q2	Qualidade/maturidade de PTTs	IMI por PTT; relatórios de teste; bancas externas; diversidade tipológica.
Q2	Relevância dos TCCs	Demandas de parceiros; % TCC→PTT com IMI alto; licenciamentos; cartas de interesse.
Q2	Talento dos egressos	Empregabilidade; progressão; inovação; <i>feedback</i> de empregadores.
Q3	Impacto econômico/social	Estudos de caso; ganhos; prêmios; indicadores de sustentabilidade.
Q3	Transferência de conhecimento/tecnologia	Licenças; receitas; <i>spinoffs</i> ; capacitações; divulgação técnica.
Q3	Políticas públicas e desenvolvimento setorial	Adoções por órgãos; normas; pareceres com efeito decisório

Fonte: Elaborado pelos autores deste artigo

Os programas devem observar as normas dispostas na legislação e no documento referencial, aplicando as diretrizes específicas detalhadas, construídas em constante diálogo com a comunidade e aprovadas pelo Comitê de Avaliação da Área (Quadro 2).

Quadro 2 – Quesitos, definições e comentários sobre os questionários

QUESITOS/ITENS	PESO	DEFINIÇÕES E COMENTÁRIOS SOBRE OS QUESITOS/ITENS
1 – Governança & Conexão Estratégica	25%	<p>A finalidade deste quesito é verificar se o programa possui estrutura, pessoas e processos alinhados à sua missão aplicada e se há uma coerência entre sua missão, planejamento, corpo docente e parcerias. O peso de 25% reflete o entendimento de que uma governança eficaz é uma condição primordial para a geração de impacto.</p> <p>1.1 Experiência e engajamento do corpo docente com o setor externo (33% do Quesito): Avaliação da qualificação do corpo docente em termos de experiência prática, parcerias com a indústria, setor público e organizações sociais, e sua capacidade de aplicar o conhecimento em contextos reais.</p> <p>1.2 Alinhamento e flexibilidade curricular com participação de stakeholders (33% do Quesito): Verificação da adequação da estrutura curricular às demandas do mercado e da sociedade, com evidências de participação ativa de stakeholders (empresas, setor público, sociedade civil) no seu desenho e revisão periódica.</p> <p>1.3 Infraestrutura para prototipagem, ensaios e validação (34% do Quesito): Disponibilidade e adequação de laboratórios, equipamentos, ferramentas e ambientes que suportem o desenvolvimento, teste e validação de Produtos Técnico-Tecnológicos (PTTs) em condições próximas às de uso real.</p> <p>Evidências Auditáveis: Plano de ação anual com evidências, com metas SMART, orçamento e cronograma (atas de aprovação e relatórios de execução). Maturidade de dados (políticas, perfis de acesso, logs, dicionário de dados) (documentos institucionais e registros de auditoria). Participação de stakeholders (comitê externo, oficinas, consultas) (atas, listas de presença, pareceres e recomendações aplicadas). Transparência (relatórios públicos, dashboards) (links/PDFs publicados). Tempo de resposta ou % de ações concluídas no ciclo (registro consolidado em painel de indicadores).</p> <p>Rubrica Resumida (Nível 0-4): Nível 0: Processos ad hoc, sem estrutura definida para conexão estratégica. Nível 1: Planejamento incipiente e dados dispersos, com participação externa pontual. Nível 2: Governança parcial e participação episódica de stakeholders. Nível 3: Governança consistente com uso de painel de dados para decisões e participação ativa de stakeholders. Nível 4: Governança madura com revisão periódica, participação ativa e transparente de stakeholders, e compromisso claro com a missão aplicada.</p>
2 – Desenvolvimento de Soluções e Formação de Talentos	45%	<p>Este quesito se concentra no “coração” dos PPGs Profissionais: a capacidade de produzir soluções aplicadas (PTTs) e de formar talentos aptos a concebê-las, implementá-las e sustentá-las. O peso de 45% explicita a centralidade dos resultados implementados e da formação orientada ao mundo do trabalho. A finalidade é qualificar os PTTs e o processo formativo orientado ao impacto.</p> <p>2.1 Qualidade e maturidade dos PTTs (33% do Quesito): Avaliação dos Produtos Técnico-Tecnológicos desenvolvidos, considerando seu Índice de Maturidade para o Impacto (IMI), a diversidade tipológica, a participação de bancas externas e a realização de demonstrações/pilotos.</p> <p>Índice de Maturidade para o Impacto (IMI) para PTTs: O IMI é um instrumento fundamental para a avaliação da qualidade e maturidade dos PTTs. Ele expressa o grau em que um PTT está pronto para gerar impacto em seu contexto de uso, resultando de uma avaliação qualitativa com pesos e rubrica comum a todas as áreas, preservando suas especificidades. O IMI agrega os escores de seis critérios (C1-C6), ponderados por pesos adaptáveis a cada área.</p>

QUESITOS/ITENS	PESO	DEFINIÇÕES E COMENTÁRIOS SOBRE OS QUESITOS/ITENS
2 – Desenvolvimento de Soluções e Formação de Talentos	45%	<p>Evidências Auditáveis:</p> <ul style="list-style-type: none"> - Criação/uso de <i>Fab Lab, living lab</i>, laboratório multiusuário; inventário de equipamentos; regulamento; agenda de uso; registros de projetos desenvolvidos. - Roadmap; TRL/MRL; relatórios de protótipo/piloto; testes/validação; backlog de demandas; lições aprendidas; critérios de escalabilidade. - Acordos de cooperação; planos de trabalho; entregas conjuntas; relatórios técnicos; comitês com parceiros; evidência de solução ancorada em ciência e tecnologia; - Depósitos de PI; softwares; manuais; produtos técnicos; licenças/contratos; registros de TT; vitrine tecnológica; casos de adoção por usuários. - Trilhas e cursos (extensão/profissionalizantes); certificações; metodologias (PBL, desafio); rubricas de competência; avaliações. <p>Rubrica Resumida do IMI (Nível 0 - 5):</p> <p>IMI de 0 -1: Sem maturidade IMI de 1,1 – 2: Maturidade baixa IMI de 2,1 – 3: Maturidade intermediária IMI de 3,1 – 4: Maturidade boa IMI de 4,1 – 5: Maturidade Alta</p>
3 – Impacto & Transformação do Território	30%	<p>A finalidade deste quesito é evidenciar o valor gerado pelos PPGs Profissionais fora da instituição. Ele mede a relevância territorial e setorial das entregas, ou seja, o quanto os projetos e egressos respondem a problemas reais, conectados a políticas públicas, cadeias produtivas e vocações locais/ regionais. Busca-se identificar resultados validados no mundo real e sua difusão.</p> <p>3.1 Impacto econômico e social direto (33% do Quesito): Avaliação de ganhos, eficiências, sustentabilidade e reconhecimento (prêmios) resultantes das soluções implementadas.</p> <p>3.2 Transferência de conhecimento e tecnologia (33% do Quesito): Análise de licenciamentos, criação de spin-offs, capacitações realizadas e divulgação técnica que demonstrem a efetiva transposição do conhecimento para a sociedade.</p> <p>3.3 Contribuição para políticas públicas e desenvolvimento setorial (34% do Quesito): Verificação da influência em normas, políticas e a adoção de soluções por órgãos governamentais ou setoriais.</p> <p>Evidências Auditáveis:</p> <p>Evidências de melhorias em políticas/serviços (ex.: redução de tempo/custo; aumento de cobertura) (laudos, dashboards públicos, atos normativos). Não é interessante dois parenteses abertos e fechados seguidamente, torna difícil de entender toda a frase.</p> <p>Parcerias com governos/empresas por eixo temático (acordos, atas, cronogramas). Casos de difusão/escala (municípios/unidades atendidas) (termos de cooperação, materiais de implementação).</p> <p>Comunicação cidadã (sumários, infográficos, linguagem acessível) (links, materiais, métricas de alcance).</p> <p>Rubrica Resumida (Nível 0 -4):</p> <p>Nível 0: Desconexão com agendas locais ou setoriais. Nível 1: Conexão episódica e sem resultados claros. Nível 2: Conexão recorrente com sinais de utilidade e reconhecimento inicial. Nível 3: Conexão sistemática com resultados documentados e impacto social/ econômico comprovado. Nível 4: Transformação territorial reconhecida, com difusão em escala e integração a políticas/planos, gerando valor público significativo.</p>

Fonte: Elaborado pelos autores deste artigo

A qualidade e a maturidade dos Produtos Técnico-Tecnológicos (PTTs) produzidos são avaliadas utilizando o Índice de Maturidade para o Impacto (IMI), que expressa o grau em que um PTT está pronto para gerar impacto em seu contexto de uso. O IMI resulta de uma avaliação qualitativa com pesos e rubrica comuns a todas as áreas, preservando suas especificidades, constituindo-se como instrumento fundamental para aferir a robustez e o potencial de transformação das soluções desenvolvidas pelos Programas Profissionais. O IMI é um indicador sintético (0 a 5) que qualifica a maturidade de impacto de cada PTT, permitindo monitoramento comparável e auditável. O objetivo não é substituir julgamento de especialistas, mas estruturar esse julgamento, tornar explícito o que conta como evidência e permitir análises agregadas do portfólio de cada programa.

O IMI é calculado como a soma ponderada de seis critérios (C1 a C6), com pesos ajustáveis por área e descritores orientados para evidências:

- 1) Robustez da Validação e Confiabilidade das Evidências (Peso: 25%), mede a solidez metodológica e a credibilidade das informações que sustentam o estágio de desenvolvimento e a aplicabilidade do PTT. Independente do PTT ser um *software*, um guia de manejo florestal ou uma técnica de performance teatral, a evidência de sua validação deve ser robusta.
- 2) Representatividade do Contexto de Aplicação (Peso: 20%), avalia o quão fiel e diversificado é o ambiente ou as condições nas quais o PTT foi demonstrado ou testado, em relação ao seu contexto de uso real. Se o PTT é uma intervenção social, o “ambiente” pode ser uma comunidade; se é uma metodologia de ensino, uma sala de aula real; se é uma nova técnica cirúrgica, um hospital.

- 3) Abrangência e Consistência das Demonstrações (Peso: 20%), que mede a amplitude dos aspectos do PTT que foram demonstrados ou testados e a consistência dos resultados obtidos. Demonstrações podem incluir protótipos em funcionamento, aplicações piloto de metodologias ou evidências da efetividade de um protocolo em diferentes situações.
- 4) Validação por Stakeholders e Reconhecimento Externo (Peso: 20%), que mensura o grau de aceitação, reconhecimento e *feedback* positivo obtido de usuários finais, parceiros, especialistas, órgãos reguladores ou comunidades relevantes. Para um PTT da área de Artes, isso pode incluir o reconhecimento crítico e a recepção do público; para um PTT social, a aceitação pela comunidade-alvo e apoio de ONGs ou governo.
- 5) Potencial de Escalabilidade e Replicabilidade (Peso: 10%), que avalia a capacidade do PTT de ser expandido para atingir um público maior, ser adotado em outros contextos ou replicado por outras equipes. Uma metodologia de saúde, um produto de design ou um processo educacional também precisam ser escaláveis ou replicáveis para gerar impacto amplo.
- 6) Gestão de Riscos e Sustentabilidade (Peso: 5%), que permite a identificação e mitigação de riscos (técnicos, sociais, éticos, ambientais, financeiros) associados ao desenvolvimento e aplicação do PTT, bem como a consideração de sua viabilidade de longo prazo para manter o impacto.

As rubricas a seguir (Quadro 3) trazem descritores progressivos (0–5) que podem ser calibrados por área. Para cada critério, estão descritos a essência da progressão, os exemplos de evidências e os alertas sobre vieses comuns. Os níveis 0 – 1 representam estágios iniciais; 2 – 3 representam uso e validação emergentes; 4 – 5 denotam consolidação e potencial de escala.

Quadro 3 – Critérios e níveis dos Indicadores de Maturidade de Impacto (IMI)

CRITÉRIOS	NÍVEIS
C1 – 25% Robustez técnica/ metodológica	<ul style="list-style-type: none"> • Nível 0: A solução está no estágio de ideia não testada; não há método definido; evidências anedóticas. • Nível 1: Protótipo conceitual sem protocolo claro; testes informais; documentação incipiente. • Nível 2: Protocolo definido; primeiros testes controlados; indicadores de qualidade identificados. • Nível 3: Testes sistemáticos com amostra adequada; documentação reprodutível; revisão técnica interna. • Nível 4: Evidência robusta com análise estatística/engenharia; revisão por especialistas externos; aderência a normas. • Nível 5: Conjunto de evidências replicadas em múltiplos contextos; certificações pertinentes; documentação aditável.
C2 – 20% Representatividade do contexto	<ul style="list-style-type: none"> • Nível 0: Contexto fictício ou pouco representativo; usuário final não caracterizado. • Nível 1: Uso experimental sem caracterização de usuários ou variabilidade do ambiente. • Nível 2: Seleção de contexto representativo inicial; mapeamento de usuários e restrições. • Nível 3: Pilotos em ambiente típico; diversidade mínima de perfis e condições operacionais. • Nível 4: Demonstrações em múltiplos contextos típicos; análise de heterogeneidade de efeitos. • Nível 5: Evidências em rede (multiunidades/territórios/setores) com aprendizado comparativo.

CRITÉRIOS	NÍVEIS
C3 – 20% Demonstrações em uso	<ul style="list-style-type: none"> • Nível 0: Sem demonstração em uso; simulações ou protótipos virtuais apenas. • Nível 1: Demonstração restrita em ambiente controlado sem usuários reais. • Nível 2: PoC/MVP com usuários limitados; registro de uso inicial. • Nível 3: Demonstrações funcionais em serviços/organizações; logs de uso; treinamento inicial. • Nível 4: Operação assistida em escala limitada; métricas operacionais estáveis; documentação de suporte. • Nível 5: Operação sustentável; adoção regular; evidências longitudinalmente estáveis.
C4 – 20% Validação por stakeholders	<ul style="list-style-type: none"> • Nível 0: Não há validações externas; feedback informal. • Nível 1: Opiniões de usuários sem formalização. • Nível 2: Cartas de apoio de parceiros potenciais; reuniões registradas. • Nível 3: Parecer técnico/institucional; pilotos validados; recomendações de adoção. • Nível 4: Certificações, homologações ou pareceres de órgãos reguladores/gestores; contratos de uso. • Nível 5: Reconhecimento setorial; políticas públicas ou normas que incorporam a solução; prêmios qualificados.
C5 – 10% Escalabilidade/ Replicabilidade	<ul style="list-style-type: none"> • Nível 0: Sem plano de escala; dependência de peças-chave. • Nível 1: Ideia de escala sem estimativas de custo/benefício. • Nível 2: Plano preliminar de difusão; materiais de implantação iniciais. • Nível 3: Guia de replicação; requisitos de infraestrutura/competências; estimativa de custos. • Nível 4: Modelos de negócio/captura de valor; arranjos contratuais; suporte técnico; governança de dados. • Nível 5: Escalas efetivas em múltiplos locais; indicadores de custo total de propriedade; sustentabilidade comprovada.
C6 – 5% Riscos e Sustentabilidade	<ul style="list-style-type: none"> • Nível 0: Riscos não mapeados; conformidade legal ignorada. • Nível 1: Riscos listados genericamente; não há plano de mitigação. • Nível 2: Matriz de risco preliminar; mitigadores iniciais; análise de viabilidade ambiental/social. • Nível 3: Plano de risco revisado; testes de contingência; conformidade regulatória e LGPD (quando aplicável). • Nível 4: Monitoramento contínuo de riscos; auditorias periódicas; mecanismos de responsabilidade e transparência. • Nível 5: Governança de riscos integrada à operação; avaliação de ciclo de vida; relatórios públicos de sustentabilidade

Fonte: Elaborado pelos autores deste artigo

A pontuação de cada critério utiliza escala ordinal de cinco níveis, com descritores progressivos. Para facilitar interpretação, propõe-se uma escala de maturidade de impacto: IMI de 0 a 1 (sem maturidade), 1 a 2 (maturidade baixa), 2,1 a 3 (maturidade intermediária), 3,1 a 4 (maturidade boa) e 4,1 a 5 (maturidade alta).

Os descritores devem ser calibrados por área. Por exemplo, na área da saúde, o critério de robustez pode exigir protocolo metodológico com ética aprovada e indicadores clínicos validados; em áreas de gestão e políticas públicas, robustez pode incluir desenho de intervenção, análise de processo e indicadores de desempenho organizacional. Essa calibragem é condição para comparabilidade justa, evitando isomorfismo indevido. O monitoramento utilizando essa sistemática, conjuga granularidade no nível do PTT e visão sistêmica no nível do programa de pós-graduação. O IMI individual permite identificar gargalos por critério e orientar melhorias. O IMI programático corresponde à média ponderada dos PTTs do período e pode incorporar ajustes por relevância estratégica ou validação externa robusta, desde que explicitados e auditáveis. Assim, propõe-se um

painel de indicadores essenciais com, no mínimo: IMI médio do programa; IMI por linha ou área; distribuição por quartis; porcentagem de PTTs com IMI maior ou igual a 3 e maior ou igual a 4; médias por critério (C1 a C6) e lacunas críticas. Recomenda-se incluir indicadores de desfechos externos quando disponíveis, como contratos, licenças, *spin-offs*, economia de custos, políticas influenciadas, treinamentos implantados e métricas de alcance de públicos específicos.

A proposta demanda um dossiê digital por PTT, contendo: síntese do problema e usuários, descrição da solução, evidências organizadas por critério, registros de uso e demonstrações, pareceres e validações, elementos de escalabilidade, matriz de riscos e próximos passos. O dossiê aumenta rastreabilidade, facilita auditoria e permite comunicação pública em linguagem técnica e executiva.

Em nível de programa, recomenda-se um ciclo inicial de implantação de 12 meses, incluindo: definição de descritores por área; capacitação de avaliadores; piloto interno; auditoria amostral; e elaboração de relatório anual

de impacto em versões técnica e executiva. Em âmbito sistêmico, a adoção na Capes exigiria um ciclo mais longo, com três fases: concepção (6 a 12 meses), piloto (12 a 18 meses) e ajustes e lançamento (18 a 24 meses), com seleção de programas, calibragem de avaliadores e revisão de pesos e descritores.

Para confiabilidade e mitigação de vieses, a governança deve incluir dupla codificação, cálculo de concordância interavaliadores, auditorias semestrais e trilhas de auditoria das evidências. Esses mecanismos atendem ao princípio de auditabilidade e reforçam a utilidade dos indicadores como instrumentos de gestão, e não como mera contabilidade.

A implementação da proposta de transformação avaliativa centrada no conceito de Maturidade para o Impacto (MPI) e no Índice de Maturidade para o Impacto (IMI) gera uma série de resultados articulados e desdobra implicações profundas para o sistema de pós-graduação profissional. O resultado mais imediato e estrutural é a redefinição do próprio objeto de avaliação. Em vez de uma ênfase desproporcional em *proxies* da produção acadêmica, como contagens de artigos e citações, prática criticada pelo Manifesto de Leiden (Hicks *et al.*, 2015) e pela Declaração de São Francisco (DORA, 2012), o foco avaliativo desloca-se integralmente para o ciclo de vida do Produto Técnico-Tecnológico (PTT) e sua trajetória de maturação. Isso significa que a “produtividade” de um programa passa a ser medida pela robustez do seu portfólio de soluções aplicadas, avaliado criteriosamente através dos seis critérios do IMI. Consequentemente, a proposta opera uma recalibragem prática e metodológica da estrutura atual da avaliação da Capes, que já tentava navegar essa complexidade através da diferenciação de pesos e subitens nas fichas por área. O modelo MPI/IMI oferece um arcabouço mais audacioso para essa diferenciação, transformando os “itens ampliados” das fichas atuais em eixos estruturantes com metodologia própria e auditável.

Esse redirecionamento do foco produz, por sua vez, o alinhamento estratégico dos incentivos internos dos programas, combatendo diretamente o efeito comportamental negativo onde “a medida vira meta”. Ao instituir o IMI como métrica central, o modelo cria incentivos concretos para que a governança do programa priorize ações alinhadas à sua missão aplicada: investir em infraestrutura de validação, fomentar parcerias estratégicas duradouras, estruturar o acompanhamento pós-defesa das dissertações e teses e revisar currículos para incorporar competências em gestão aplicada. Esse realinhamento reflete-se diretamente na experiência formativa, que se torna mais densa e situada. O sucesso do egresso passa a ser aferido por sua capacidade efetiva de gerar inovação em seu setor, um dado que o próprio programa será levado a monitorar, atendendo ao princípio da aprendizagem contínua por meio da avaliação.

A operacionalização do modelo, com sua exigência de dossiês digitais e evidências auditáveis para cada critério do IMI, gera um resultado tangível: a sistematização de um acervo robusto de comprovação de impacto. Este acervo transcende sua função avaliativa para se tornar a base de uma comunicação estratégica de valor, assegurando transparência para a sociedade e *accountability* sobre o retorno do investimento público. A “vitrine digital de PTTs” baseada em evidências concretas responde ao apelo por uma avaliação contextualizada, que considere o território e os arranjos institucionais específicos.

No nível da análise comparativa, a aplicação do IMI permite uma sofisticação que vai muito além da geração de um *ranking* simplista. O resultado analítico mais rico é a possibilidade de traçar o perfil de maturidade de cada programa em um gráfico de radar, revelando padrões específicos de desempenho nos seis critérios. Um programa pode ter excelência em pesquisa aplicada (C1, C2 altos), mas lacunas em transferência (C4, C5 baixos), enquanto outro pode mostrar o perfil inverso. Esta análise permite diagnósticos precisos e ações corretivas direcionadas, transformando a avaliação em uma ferramenta de gestão. Da mesma forma, a comparação longitudinal do perfil de um mesmo programa ao longo dos quadriênios permite verificar se as ações implementadas resultaram em um aumento consistente da maturidade, efetivando o ciclo de melhoria contínua.

A discussão, no entanto, deve também considerar os desafios inerentes à implementação. A calibragem fina dos descritores das rubricas do IMI para cada uma das áreas da Capes é um trabalho colossal. A mudança da cultura avaliativa exigirá programas robustos de capacitação. A carga administrativa precisa ser balanceada com o princípio da proporcionalidade, para não onerar excessivamente programas menores. Em sua essência, seguindo a recomendação de Santos (2020) sobre o uso responsável de métricas, a proposta se posiciona não como uma substituição, mas como uma complementação necessária ao ecossistema de avaliação, oferecendo uma gramática específica e legítima para os Programas Profissionais. Ao fazer isso, o modelo MPI/IMI tem o potencial de fortalecer a identidade dos PPs e assegurar que a pesquisa aplicada se traduza efetivamente em inovação e transformação do território, gerando o valor público que é sua razão de ser.

A proposta MPI/IMI se insere em um movimento de reorientação da avaliação para reconhecer diversidade de produtos e trajetórias de impacto. Em programas profissionais, essa reorientação é especialmente necessária, porque parte relevante do valor gerado se materializa em soluções que não se convertem imediatamente em artigos: protocolos implementados, sistemas em operação, normativas influenciadas, tecnologias sociais adotadas, métodos de gestão incorporados e práticas profissionais

transformadas. Um sistema avaliativo que não captura essas evidências tende a induzir deslocamentos de esforço para outputs mais facilmente mensuráveis, mesmo quando menos aderentes à missão do programa.

O IMI busca equilibrar comparabilidade e especificidade. Ao manter critérios comuns e permitir pesos e descritores ajustáveis por área, o índice cria uma estrutura transversal sem impor homogeneização. Essa arquitetura pode reduzir ruídos ao tornar explícito o que conta como evidência, e ao aproximar avaliadores, programas e stakeholders de uma linguagem comum. Ainda assim, a experiência internacional alerta para riscos de fetichização do número. Índices sintéticos podem ser usados como atalho decisório ou como meta em si, produzindo distorções. A resposta é coerente com métricas responsáveis: o indicador deve apoiar, e não substituir, julgamento qualificado; e deve ser interpretado em contexto, com triangulação de evidências.

Outra limitação é a assimetria de capacidade institucional. Programas com redes de parceria consolidadas, maior infraestrutura e experiência em documentação podem produzir dossiês mais robustos, independentemente de diferença real de impacto. Por isso, o roteiro de implantação propõe aprendizado incremental, guias por área, capacitação e auditoria, além de mecanismos de comparabilidade que considerem contexto e vulnerabilidades. Também é necessário cuidado ético: validações externas devem ser verificáveis e livres de conflito de interesse, e impactos reportados precisam ser sustentados por evidências proporcionais.

Na avaliação vigente dos PPGs no Brasil, há avanços relevantes, mas a literatura tem sido consistente ao apontar um efeito colateral persistente: quando indicadores de produção e “sinais” bibliométricos se tornam o eixo organizador da qualidade, o sistema tende a induzir produtivismo e conformidade métrica, com perda de aderência entre avaliação, missão do programa e relevância social (Patrus; Dantas; Shigaki, 2015; Barata, 2019). Ao mesmo tempo, as diretrizes Capes 2025-2028 reconhecem explicitamente a necessidade de um olhar mais multidimensional, com maior ênfase em impacto para a sociedade, inovação e transferência, e propõem uma ficha estruturada por quesitos/itens/pesos (Capes, 2025a; Capes, 2025b). O desafio, portanto, não é apenas declarar novas dimensões, mas operacionalizá-las com critérios verificáveis e comparáveis, sem recair em proxies simplistas.

É nesse ponto que a proposta (MPI/IMI) ganha força em comparação ao status quo: ela desloca o foco de “*outputs* finais” para capacidade institucional de gerar, sustentar e demonstrar impacto, organizando a avaliação em dimensões causais e auditáveis (governança e conexão estratégica; desenvolvimento de soluções e formação; impacto e transformação territorial). Em termos internacionais, essa

arquitetura é convergente com recomendações centrais de avaliação responsável, que defendem contextualização, pluralidade de evidências e cautela com métricas de periódico como substitutas de qualidade (DORA, 2012; Hicks *et al.*, 2015; Wilsdon *et al.*, 2015). Além disso, ao explicitar trilhas de evidência, validação externa e julgamento por pares estruturado, o modelo se alinha ao movimento contemporâneo de reforma de avaliação, que busca reconhecer diversidade de contribuições, integridade, transparência e impacto além da academia (CoARA, 2022; Moher *et al.*, 2020).

Outro ponto forte é que a proposta aproxima a avaliação dos PPGs de práticas consolidadas em sistemas que incorporaram evidências corroboradas de impacto, como o REF no Reino Unido, que orienta o uso de testemunhos e documentação factual para sustentar estudos de caso (REF, 2021), e a experiência australiana de Engagement and Impact, que avaliou engajamento com usuários finais e impactos sociais, econômicos, ambientais e culturais (ARC, 2026). Na mesma direção, protocolos como o Strategy Evaluation Protocol holandês avaliam qualidade, relevância e viabilidade em ciclos longos, reforçando a ideia de maturidade e trajetória, e não apenas fotografia de resultados (VSNU; NWO; KNAW, 2021). Em síntese, o MPI/IMI fortalece a comparação com o status quo porque oferece um caminho metodológico para medir aquilo que a própria política avaliativa brasileira vem demandando: impacto com lastro, evidência verificável, e reconhecimento de missões e produtos diversos, sem reduzir qualidade a contagens.

Por fim, o MPI/IMI tem potencial de induzir melhorias internas. Ao evidenciar lacunas por critério, o painel orienta investimentos em robustez metodológica, validação externa, escalabilidade e gestão de riscos. Em nível sistêmico, a proposta contribui para dar materialidade a um discurso já presente na avaliação: reconhecer o papel dos PPs na inovação e no valor público. Para que isso ocorra, a adoção deve vir acompanhada de comunicação clara, guias, governança e processos que evitem captura por objetivos de curto prazo.

4 Considerações Finais

A transformação avaliativa proposta, centrada no conceito de Maturidade para o Impacto (MPI) e operacionalizada pelo Índice de Maturidade para o Impacto (IMI), não pretende suprimir a ciência acadêmica, mas complementar o repertório avaliativo com uma gramática própria dos Programas Profissionais. Essa gramática valoriza evidências de uso real, validação por *stakeholders* e resultados que transformam práticas, organizações e territórios.

O modelo oferece um vocabulário comum, auditável e adaptável às diferentes áreas do conhecimento, capaz de orientar decisões curriculares, fomentar parcerias estratégicas e direcionar investimentos. Ao alinhar os incentivos da avaliação com a missão aplicada dos PPs, espera-se promover uma produção de PTTs mais robusta, relevante e com maior potencial de geração de valor público. Com governança adequada, dados de qualidade e comunicação efetiva, a avaliação pode transcender seu papel de rito de passagem e se consolidar como um motor de aprendizagem e melhoria contínua para os Programas de Pós-Graduação Profissionais.

A consolidação da proposta demanda a implementação em fases, com perspectivas claras de curto, médio e longo prazo. Como agenda de desenvolvimento, recomenda-se a realização de pilotos interáreas para calibrar descritores e pesos, testar confiabilidade interavaliadores e investigar relações entre IMI e desfechos de médio prazo, como institucionalização, ganhos de eficiência, mudanças de política e trajetória de egressos. Uma avaliação mais justa e útil deve ser capaz de medir o que de fato importa para o país: soluções que funcionam, aprendizados que se acumulam e impactos que melhoram a vida das pessoas e a produtividade das organizações.

5 Perspectivas Futuras

Como desdobramento natural desta proposta, as perspectivas futuras envolvem avançar da formulação conceitual para a validação empírica e a calibragem interáreas do MPI/IMI. Isso inclui pilotos com Programas Profissionais de diferentes áreas, testando a clareza dos descritores, a adequação dos pesos por campo e a consistência do processo de atribuição de escores. Um objetivo central dessa etapa é estimar a confiabilidade interavaliadores, identificar pontos de ambiguidade nas rubricas e produzir guias de evidências por tipologia de PTT, de modo a reduzir ruído e aumentar comparabilidade sem impor homogeneização indevida.

Em paralelo, recomenda-se uma agenda de pesquisa voltada para a validade do IMI como indicador de maturidade para impacto. Estudos longitudinais podem investigar a relação entre pontuações do IMI e desfechos observáveis de médio prazo, como institucionalização da solução, adoção por organizações parceiras, efeitos em políticas públicas, ganhos de eficiência, difusão territorial e trajetórias de egressos. Também é promissor explorar a complementaridade do IMI com abordagens internacionais correlatas, aprimorando a leitura de maturidade ao diferenciar “qualidade e robustez da solução” de “condições sociais e institucionais para adoção”, evitando interpretações lineares e simplificadoras do impacto.

Por fim, a implementação sustentável do MPI/IMI depende do fortalecimento de uma infraestrutura de evidências e governança de dados que torne o processo auditável e útil para gestão, e não apenas para prestação de contas. Isso implica padronizar dossiês digitais de PTT, definir trilhas de auditoria, critérios mínimos de validação externa e rotinas de aprendizagem institucional, com transparência e mitigação de conflitos de interesse. Uma perspectiva estratégica é consolidar painéis e relatórios públicos em linguagem técnica e executiva, capazes de comunicar valor público e orientar investimentos, ao mesmo tempo em que se desenham salvaguardas contra efeitos comportamentais, como a fetichização do número ou a produção de evidências apenas para “cumprir rubrica”, preservando o julgamento qualificado e contextual.

Agradecimentos

Este artigo é resultado do trabalho final do VI Curso de Atualização em Métricas de Desempenho Acadêmico e Comparações Internacionais (2025), promovido pela Universidade de São Paulo (USP), com apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp). Agradecemos aos dois revisores anônimos pelos comentários e sugestões, que contribuíram para o aprimoramento da proposta e do manuscrito.

Referências

- ARC – AUSTRALIAN RESEARCH COUNCIL. **Engagement and Impact Assessment**. [2026]. Disponível em: <https://www.arc.gov.au/evaluating-research/engagement-and-impact-ei>. Acesso em: 06 out. 2025
- BARATA, Rita Barradas. Mudanças necessárias na avaliação da pós-graduação brasileira. **Interface**, Botucatu, v. 23, e180635, 2019. DOI: <https://doi.org/10.1590/Interface.180635>. Disponível em: <https://www.scielo.br/j/icse/a/gBkWRwqC5svbVNL3R8QN4sx/>. Acesso em: 10 out. 2025.
- BERNSTEIN, Michael J. *et al.* The Societal Readiness Thinking Tool: A Practical Resource for Maturing the Societal Readiness of Research Projects. **Science and Engineering Ethics**, v. 28, art. 6, 2022. Disponível em: <https://link.springer.com/article/10.1007/s11948-021-00360-3>. Acesso em: 16 ago. 2025.
- CAPES – COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR. **Documentos do novo ciclo avaliativo 2025-2028**. Brasília, DF: Capes, 2025a. Disponível em: <https://www.gov.br/capes/pt-br/aceso-a-informacao/acoes-e-programas/avaliacao/sobre-a-avaliacao/documentos-do-novo-ciclo-avaliativo-2025-2028>. Acesso em: 13 jul. 2025.

CAPES – COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR. **Diretrizes comuns da Avaliação de Permanência dos Programas de Pós-Graduação *stricto sensu***: princípios e processos da Avaliação de Permanência do SNP (ciclo 2025-2028). Brasília, DF: Capes, 2025b. Disponível em: https://www.gov.br/capes/pt-br/centrais-de-conteudo/documentos/avaliacao/19052025_20250502_DocumentoReferencial_FICHA.pdf. Acesso em: 13 jul. 2025.

CoARA – COALITION FOR ADVANCING RESEARCH ASSESSMENT. **Agreement on Reforming Research Assessment**: the Agreement full text. 2022. Disponível em: <https://www.coara.org/agreement/the-agreement-full-text/>. Acesso em: 6 out. 2025.

DORA – DECLARAÇÃO DE SÃO FRANCISCO SOBRE AVALIAÇÃO DA PESQUISA. **Português Brasileiro**. San Francisco: DORA, 2012. Disponível em: <https://sfdora.org/read/read-the-declaration-portugues-brasileiro/>. Acesso em: 13 jul. 2025.

GHENO, Ediane Maria *et al.* Sistema de avaliação da Capes: indicadores e procedimentos de monitoramento e avaliação de desempenho. **Em Questão**, Porto Alegre, v. 25, n. 3, p. 184-213, 2019. DOI: 10.19132/1808-5245253.184-213.

HICKS, Diana *et al.* Bibliometrics: The Leiden Manifesto for research metrics. **Nature**, v. 520, n. 7548, p. 429-431, 2015. DOI: 10.1038/520429a.

IFD – INNOVATION FUND DENMARK. **Societal Readiness Levels (SRL) defined according to Innovation Fund Denmark**. Copenhagen: Innovation Fund Denmark, 2019. Disponível em: https://innovationsfonden.dk/sites/default/files/2019-03/societal_readiness_levels_-_srl.pdf. Acesso em: 13 jul. 2025.

MCMANUS Concepta *et al.* Assessment of the Brazilian postgraduate evaluation system. **Front. Educ.**, v. 7, p. 1036108, 2022. DOI: 10.3389/educ.2022.1036108.

MCMANUS Concepta; BAETA NEVES Abilio Afonso; CARVALHO, Carlos Henrique. Technical and artistic production by brazilian postgraduate courses with emphasis on Social Sciences and Humanities. **Revista NUPEM**, v. 15, n. 35, p. 179-204. 2023. DOI: <https://doi.org/10.33871/nupem.2023.15.35.179-204>.

MOHER, David *et al.* The Hong Kong Principles for assessing researchers: fostering research integrity. **PLOS Biology**, San Francisco, v. 18, n. 7, e3000737, 2020. DOI: <https://doi.org/10.1371/journal.pbio.3000737>.

OLIVEIRA, Maria Eliza Nogueira; ROTHEN, José Carlos. A política de avaliação da pós-graduação no Brasil: um estudo documental (1998-2018) a partir da área da Educação. **Cadernos de Pesquisa**, v. 29, n. 2, p. 411-439, 5. Jul. 2022. Disponível em: <https://periodicoseltronicos.ufma.br/index.php/cadernosdepesquisa/article/view/19497>. Acesso em: 10 out. 2025.

PATRUS, Roberto; DANTAS, Douglas Cabral; SHIGAKI, Helena Belintani. O produtivismo acadêmico e seus impactos na pós-graduação *stricto sensu*: uma ameaça à solidariedade entre pares? **Cadernos EBAPE.BR**, Rio de Janeiro, v. 13, n. 1, p. 1-18, 2015. Disponível em: <https://www.scielo.br/j/cebape/a/HL7xXqvSVnf43TjFfQ4NVwt/>. Acesso em: 10 out. 2025.

PEREZ, Olivia Cristina Avaliação em Disputa: A Reforma do Qualis e os Desafios para a Ciência. **Novos Debates**, v. 11, n. 1, 2025. DOI: <https://doi.org/10.48006/2358-0097/V11N1.E111011>.

REF – RESEARCH EXCELLENCE FRAMEWORK. **Additional guidance**: Guidance on providing testimonies for REF 2021 impact case studies. 2021. Disponível em: <https://2021.ref.ac.uk/guidance-and-criteria-on-submissions/guidance/additional-guidance/index.html>. Acesso em: 13 jul. 2025.

SANTOS, Solange Maria. Uso responsável de métricas e indicadores. **Portal Métricas**, São Paulo, 2020. Disponível em: <https://metricas.usp.br/uso-responsavel-de-metricas-e-indicadores/>. Acesso em: 10 out. 2025.

VSNU; NWO; KNAW. **Strategy Evaluation Protocol 2021-2027**. 2021. Disponível em: https://storage.know.nl/2022-06/SEP_2021-2027.pdf. Acesso em: 13 jul. 2025.

WILSDON, James *et al.* **The Metric Tide**: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. London: HEFCE, 2015. Disponível em: <https://www.ukri.org/wp-content/uploads/2021/12/RE-151221-TheMetricTideFullReport2015.pdf>. Acesso em: 13 jul. 2025.

Sobre os Autores

Andrea Viviana Waichman

E-mail: awaichman@gmail.com

ORCID: <https://orcid.org/0000-0002-6357-9019>

Doutora em Biologia de Água Doce e Pesca Interior pelo Instituto Nacional de Pesquisas da Amazônia.

Endereço profissional: Universidade Federal do Amazonas, Instituto de Ciências Biológicas, Departamento de Biologia, Av. General Rodrigo Otávio Jordão Ramos, n. 3.000, Manaus, AM. CEP: 69077-000.

Eduardo Meireles

E-mail: eduardo@eduardomeireles.com.br

ORCID: <https://orcid.org/0000-0002-6711-6572>

Doutor em Engenharia Urbana pela Universidade Federal de São Carlos.

Endereço profissional: Universidade Estadual de Minas Gerais, Unidade Frutal, Avenida Escócia, n. 1.001, Bairro Cidade das Águas, Frutal, MG. CEP: 38202-436.

Genildo Cavalcante Ferreira-Júnior

E-mail: genildo.junior@ifac.edu.br

ORCID: <https://orcid.org/0000-0002-7359-3738>

Doutor em Bioquímica e Biotecnologia pela Universidade Federal de Alagoas.

Endereço profissional: Instituto Federal do Acre, IFAC Reitoria, Via Chico Mendes, n. 3.084, Comara, Rio Branco, AC. CEP: 69906-310.