



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Vinicius Barbosa Polito Lopes

**Integração da Inteligência Artificial Generativa ao Problem-Based Learning no
Ensino Superior:** um Design Science para Apoio à Prática Docente

Recife

2026

Vinicius Barbosa Polito Lopes

**Integração da Inteligência Artificial Generativa ao Problem-Based Learning no
Ensino Superior:** um Design Science para Apoio à Prática Docente

Trabalho apresentado ao Programa de Pós-graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco, como requisito parcial para obtenção do grau de Mestre em Ciência da Computação.

Área de Concentração: Engenharia de Software

Orientador (a): Vinicius Cardoso Garcia

Coorientador (a): Jackson Raniel Florencio da Silva

Recife

2026

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Lopes, Vinicius Barbosa Polito.

Integração da Inteligência Artificial Generativa ao Problem-Based Learning no Ensino Superior: um Design Science para Apoio à Prática Docente / Vinicius Barbosa Polito Lopes. - Recife, 2026.

126f.: il.

Dissertação (Mestrado) - Universidade Federal de Pernambuco, Centro de Informática, Programa de Pós-Graduação em Ciência da Computação, 2026.

Orientação: Vinicius Cardoso Garcia.

Coorientação: Jackson Raniel Florencio da Silva.

Inclui Referências e Apêndices.

1. Generative Artificial Intelligence; 2. Problem-Based Learning; 3. Active Methodologies; 4. Higher Education; 5. Design Science Research. I. Garcia, Vinicius Cardoso. II. Silva, Jackson Raniel Florencio da. III. Título.

UFPE-Biblioteca Central

Vinicius Barbosa Polito Lopes

“Integração da Inteligência Artificial Generativa ao Problem-Based Learning no Ensino Superior: um Guideline e um Chatbot Tutor Baseados em Pesquisa-Ação”

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Ciência da Computação. Área de Concentração: Engenharia de Software e Linguagens de Programação.

Aprovado em: 30/01/2026.

BANCA EXAMINADORA

Prof. Dr. Geber Lisboa Ramalho
Centro de Informática/UFPE

Prof. Dr. Igor Scaliante Wiese
Universidade Tecnológica Federal do Paraná /
Campus Campo Mourão

Prof. Dr. Vinicius Cardoso Garcia
Centro de Informática / UFPE
(orientador)

AGRADECIMENTOS

Gostaria de expressar meu sincero e profundo agradecimento a todas as pessoas que, de forma direta ou indireta, contribuíram para a realização deste trabalho.

Conciliar uma rotina de estudos, trabalho e os desafios cotidianos da vida não é uma tarefa simples, demorei um pouco para perceber isso, e esse percurso só foi possível graças ao apoio de pessoas muito importantes.

Primeiramente, agradeço aos meus pais e meu irmão. À minha mãe, que atuou como professora acadêmica por grande parte de sua vida e sempre foi um exemplo, mostrando-me, desde cedo, a importância da educação e da academia. Ao meu pai, que, embora não tenha seguido uma carreira acadêmica tão extensa quanto a de minha mãe, possui o título de mestre e já atuou como docente, me mostrando o valor do conhecimento e da formação acadêmica. Ao meu irmão, que precisou lidar com meu estresse e foi paciente comigo durante este período.

Estendo meu agradecimento aos demais familiares e amigos, que compreenderam minha ausência em diversos momentos e eventos, enquanto eu priorizava o desenvolvimento e a escrita deste trabalho.

Agradeço a Geber, pelo incentivo para que eu ingressasse no mestrado e pela indicação do meu orientador, apoio fundamental no início desta trajetória.

Registro um agradecimento especial ao meu orientador, Vinicius Cardoso Garcia, e ao meu coorientador, Jackson Raniel Florencio da Silva. Vinicius, agradeço pelos puxões de orelha, pela paciência e pelo suporte ao longo de todo esse caminho. Reconheço que nem sempre fui um orientando fácil, mas sou extremamente grato pelo suporte e pelas contribuições. Jackson, agradeço pela disponibilidade, paciência, conversas e incentivos constantes, que foram essenciais durante o desenvolvimento deste trabalho.

Chegar até aqui só foi possível graças ao apoio de todos vocês.

A todos, meu muito obrigado.

RESUMO

A expansão recente de ferramentas de Inteligência Artificial Generativa (GenAI) no ensino superior tem ampliado seu uso por estudantes e intensificado desafios pedagógicos, éticos e institucionais relacionados ao planejamento e à condução de atividades de ensino. Nesse contexto, esta dissertação investiga como a GenAI pode dar suporte a professores universitários na adoção de metodologias ativas, com foco no *Problem-Based Learning* (PBL), tendo como objetivo apoiar a prática docente por meio de orientações pedagógicas estruturadas para o planejamento e a condução de atividades educacionais. Para isso, a pesquisa adota a abordagem de Design Science Research, organizada em um ciclo de (1) diagnóstico, (2) planejamento, (3) intervenção, (4) avaliação e (5) reflexão. Como resultado, foram concebidos e desenvolvidos dois artefatos educacionais: (i) um *guideline* voltado a orientar docentes na condução de atividades de PBL com suporte de GenAI e (ii) um chatbot tutor destinado ao apoio pedagógico ao professor. A avaliação dos artefatos foi realizada por meio de revisão por especialistas, combinando análise qualitativa de feedback e uma etapa automatizada baseada em *LLM-as-a-Judge*, com critérios explícitos e verificações adicionais para mitigar vieses e instabilidades de julgamento. Os resultados indicam que o *guideline* contribui para a clareza e a aplicabilidade das orientações propostas, enquanto o chatbot oferece apoio operacional e consultivo ao docente. Conclui-se que a combinação de orientação pedagógica estruturada e suporte conversacional baseado em conhecimento pode reduzir barreiras de adoção e favorecer práticas mais transparentes e responsáveis no uso da GenAI no ensino superior.

Palavras-chaves: Inteligência Artificial Generativa; Problem-Based Learning; Metodologias Ativas; Ensino Superior; Design Science Research.

ABSTRACT

The recent expansion of Generative Artificial Intelligence (GenAI) tools in higher education has increased their use by students and intensified pedagogical, ethical, and institutional challenges related to the planning and delivery of teaching activities. In this context, this dissertation investigates how GenAI can support university instructors in the adoption of active learning methodologies, with a focus on Problem-Based Learning (PBL), aiming to support teaching practice through structured pedagogical guidance for the planning and facilitation of educational activities. To this end, the study adopts a Design Science Research, organized into iterative cycles of (1) diagnosis, (2) planning, (3) intervention, (4) evaluation, and (5) reflection. As a result, two educational artifacts were designed and developed: (i) a guideline aimed at supporting instructors in the facilitation of PBL activities with GenAI support, and (ii) a tutor chatbot intended to provide pedagogical support to instructors. The artifacts were evaluated through expert review, combining qualitative feedback analysis with an automated evaluation stage based on the LLM-as-a-Judge approach, using explicit criteria and additional checks to mitigate bias and judgment instability. The results indicate that the guideline contributes to the clarity and applicability of the proposed guidance, while the chatbot provides operational and consultative support to instructors. It is concluded that the combination of structured pedagogical guidance and knowledge-based conversational support can reduce adoption barriers and foster more transparent and responsible practices in the use of GenAI in higher education.

Keywords: Generative Artificial Intelligence; Problem-Based Learning; Active Learning Methodologies; Higher Education; Design Science Research.

LISTA DE FIGURAS

Figura 1 – Ciclo da pesquisa-ação	25
Figura 2 – Visão geral do desenho experimental adotado na avaliação do chatbot tutor, incluindo embaralhamento das perguntas, cegamento das respostas, avaliação por LLM-as-a-Judge, desceçamento, controle de consistência e auditoria humana.	38
Figura 3 – Ranking de priorização dos desafios da inteligência artificial segundo os docentes participantes.	48

LISTA DE CÓDIGOS

Código Fonte 1 – Arquivo de configuração utilizado no experimento comparativo com LLM-as-a-Judge	60
Código Fonte 2 – Exemplo de registro individual de julgamento no arquivo judgments.jsonl	70
Código Fonte 3 – Exemplo de linha consolidada do arquivo results.csv	70

LISTA DE QUADROS

Quadro 1 – Síntese da etapa de intervenção da pesquisa	33
Quadro 2 – Relação entre os eixos teóricos considerados no planejamento da ação e as questões práticas que orientaram sua definição.	49
Quadro 3 – Etapas do ciclo do Problem-Based Learning adotadas no guideline	51
Quadro 4 – Categorias funcionais de ferramentas de GenAI utilizadas no guideline . .	52
Quadro 5 – Escala ordinal de aderência pedagógica utilizada para avaliar a relevância das categorias de ferramentas em cada etapa do PBL	53
Quadro 6 – Estrutura padrão utilizada no detalhamento de cada uma das sete etapas do PBL com GenAI	54
Quadro 7 – Estrutura geral do guideline de PBL integrado à GenAI	55
Quadro 8 – Arquitetura do guideline em camadas: etapas, níveis e ferramentas	55
Quadro 9 – Níveis de implementação do PBL utilizados no guideline	56
Quadro 10 – Síntese dos registros de consistência e instabilidade dos julgamentos automáticos	69

LISTA DE TABELAS

Tabela 1 – Estrutura PICOC utilizada na RLR.	27
Tabela 3 – Estudos analisados na RLR quanto à abordagem de desafios associados à integração da inteligência artificial no ensino superior.	47
Tabela 4 – Dimensões pedagógicas que estruturam o guideline de integração do PBL com GenAI	49
Tabela 5 – Matriz de aderência pedagógica entre etapas do PBL e categorias funcionais de ferramentas de GenAI (escala Likert 1–5)	54
Tabela 6 – Síntese dos resultados da avaliação do guideline por especialistas (N = 3) .	58
Tabela 7 – Perguntas utilizadas como prompts no experimento comparativo com LLM-as-a-Judge	66
Tabela 8 – Critérios utilizados pelo LLM-as-a-Judge para o julgamento comparativo das respostas	68

SUMÁRIO

1	INTRODUÇÃO	12
1.1	MOTIVAÇÃO	14
1.2	OBJETIVOS	15
1.2.1	Objetivo Geral	15
1.2.2	Objetivos Específicos	15
1.3	ESCOPO	15
1.4	ESTRUTURA DO TRABALHO	16
2	FUNDAMENTAÇÃO TEÓRICA	17
2.1	GENAI NO ENSINO SUPERIOR	17
2.2	DESAFIOS PEDAGÓGICOS DA INTELIGÊNCIA ARTIFICIAL GENERATIVA	18
2.3	METODOLOGIAS ATIVAS E PROBLEM-BASED LEARNING	19
2.4	INTEGRAÇÃO ENTRE PBL E GENAI	21
2.5	ARTEFATOS EDUCACIONAIS COMO ESTRATÉGIA DE INTERVENÇÃO	22
3	METODOLOGIA	25
3.1	DIAGNÓSTICO DO PROBLEMA	25
3.1.1	Rapid Literature Review como instrumento de diagnóstico	26
3.1.2	Seminário com docentes universitários como estratégia exploratória de apoio ao diagnóstico	28
3.2	PLANEJAMENTO DA AÇÃO	29
3.3	INTERVENÇÃO	32
3.4	AVALIAÇÃO	34
3.4.1	Survey com Especialistas	34
3.4.2	Avaliação experimental do chatbot tutor	35
3.4.2.1	<i>Desenho experimental e tratamentos</i>	35
3.4.2.2	<i>LLM-as-a-Judge e justificativa metodológica</i>	39
3.4.2.3	<i>Critérios de avaliação e métricas analisadas</i>	39
3.4.2.4	<i>Mitigação de vieses e auditoria humana</i>	40
3.4.2.5	<i>Rastreabilidade e reprodutibilidade</i>	41
3.5	REFLEXÃO	41
4	PROJETO E CONSTRUÇÃO DOS ARTEFATOS	42

4.1	VISÃO GERAL DOS ARTEFATOS	42
4.2	CONSTRUÇÃO DO GUIDELINE	43
4.3	CONSTRUÇÃO DO CHATBOT	44
4.4	SÍNTESE DOS ARTEFATOS	45
5	RESULTADOS	46
5.1	RESULTADOS DO DIAGNÓSTICO	46
5.2	RESULTADOS DO PLANEJAMENTO DA AÇÃO	48
5.3	RESULTADOS DA INTERVENÇÃO	49
5.3.1	Desenvolvimento iterativo do guideline de PBL integrado à GenAI .	50
5.3.2	Chatbot tutor de apoio pedagógico	56
5.4	RESULTADOS DA AVALIAÇÃO DOS ARTEFATOS	57
5.4.1	Resultados da avaliação do guideline por especialistas	57
5.4.2	Resultados do experimento comparativo com LLM-as-a-Judge . . .	60
5.4.3	Configuração experimental e parâmetros controlados	60
5.4.4	Prompts e cobertura do contexto pedagógico	62
5.4.5	Crerios de julgamento e estrutura dos escores	66
5.4.6	Consistência dos julgamentos e tratamento de instabilidades	68
5.4.7	Artefatos de resultados gerados e rastreabilidade	69
5.4.8	Síntese dos resultados do experimento	71
5.5	SÍNTESE DOS RESULTADOS	72
6	DISCUSSÃO	73
6.1	RETOMADA DOS OBJETIVOS DA PESQUISA E NATUREZA DOS ACHADOS	73
6.2	DISCUSSÃO DO GUIDELINE COMO ARTEFATO PEDAGÓGICO MEDIADOR	74
6.3	O CHATBOT TUTOR COMO INSTRUMENTO DE APOIO E EXPERIMENTAÇÃO	76
6.4	IMPLICAÇÕES PEDAGÓGICAS E INSTITUCIONAIS	78
6.5	REFLEXÕES FINAIS	79
7	CONCLUSÃO	81
7.1	SÍNTESE DOS RESULTADOS	81
7.2	CONTRIBUIÇÕES DA PESQUISA	82
7.3	LIMITAÇÕES DA PESQUISA	82

7.4	TRABALHOS FUTUROS	83
	REFERÊNCIAS	84
	APÊNDICE A – PROTOCOLO DE RAPID LITERATURE REVIEW	
	(RLR)	90
	APÊNDICE B – GUIDELINE	93

1 INTRODUÇÃO

Ao final de novembro de 2022, a empresa e laboratório de pesquisa estadunidense OpenAI lançou um serviço denominado ChatGPT (OpenAI, 2022). O lançamento deste sistema representou um marco significativo no campo da inteligência artificial (IA) ao alcançar aproximadamente 100 milhões de usuários em um período de dois meses. Esse ritmo de adoção foi excepcional quando comparado a outros produtos tecnológicos amplamente difundidos, como o TikTok, que levou nove meses para atingir a mesma marca, e o Instagram, que necessitou de dois anos para alcançar esse número de usuários (TEAM, 2025; MILMO; AGENCY, 2023; HU, 2023). Esse rápido crescimento pode ser atribuído a diversos fatores, sendo o principal deles a democratização do acesso à IA, anteriormente restrita a especialistas e grandes organizações.

A disponibilização de uma ferramenta avançada, treinada com grande volume de informações públicas da internet e de uso intuitivo, transformou significativamente a relação das pessoas com sistemas baseados em IA. Essa mudança decorre do fato de que a OpenAI não disponibilizou apenas uma aplicação final para interação, mas também tornou acessível o modelo que sustenta o sistema. Tal decisão permitiu que pesquisadores, empresas e desenvolvedores tivessem acesso aos Large Language Models (LLMs) que antes eram exclusivos de grandes corporações, devido ao elevado custo de obtenção de dados e de treinamento necessário para esse tipo de tecnologia.

Essa abertura desencadeou transformações profundas em diversos setores. Áreas como marketing (CHUMA; OLIVEIRA, 2023; NGO, 2024), recursos humanos (BUDHWAR et al., 2023; ZARIFHONARVAR, 2023) e saúde (WEN; WANG, 2023) passaram a incorporar o uso de sistemas baseados em *Generative Artificial Intelligence* (GenAI) em diferentes processos e estratégias. No contexto educacional, o impacto foi ainda mais expressivo (LO, 2023; DEMPERE et al., 2023), uma vez que instituições, estudantes e professores se depararam com novas possibilidades e desafios relacionados ao uso pedagógico e ético dessas ferramentas.

A rápida disseminação de ferramentas de GenAI encontrou um cenário particular no ensino superior. No Brasil, diferentemente dos professores da educação básica, cuja formação é estruturada com foco na preparação didático-pedagógica, grande parte dos docentes universitários inicia suas atividades sem passarem por qualquer processo sistemático de formação pedagógica (GIL, 2000). Essa realidade está relacionada à valorização da experiência profissional como principal critério para o exercício da atividade docente, o que contribui para

a percepção de que o domínio técnico da área seria suficiente para garantir a qualidade do ensino (GIL, 2000). Como consequência, processos institucionais de seleção e desenvolvimento profissional tendem a priorizar competências técnicas, enquanto aspectos pedagógicos acabam ficando em segundo plano.

Essa ausência de formação pedagógica sistemática entre professores universitários impõe barreiras significativas à incorporação crítica de tecnologias como GenAI no cotidiano educacional (TUBELLA; MORA-CANTALLOPS; NIEVES, 2024). Ao mesmo tempo em que tais ferramentas oferecem potenciais transformadores para o ensino-aprendizagem, sua adoção requer compreensão didática, reflexão ética e alinhamento com os objetivos pedagógicos, independentemente da formação prévia do docente. No contexto da GenAI, cuja evolução é acelerada e de amplo impacto, a ausência de diretrizes explícitas pode fazer com que a mesma ferramenta atue tanto como apoio quanto como fator de fragilização do processo formativo.

Diante desse contexto de transformações digitais aceleradas no ensino superior, intensificadas pela presença crescente de ferramentas de GenAI, torna-se essencial fortalecer práticas pedagógicas que promovam participação ativa, autonomia intelectual e, sobretudo, desenvolvimento do pensamento crítico. As metodologias ativas assumem um papel central nesse processo por favorecerem a construção do conhecimento a partir da investigação, da colaboração e da análise criteriosa de informações (GIL, 2000; TAJRA et al., 2021). Essas práticas tornam o estudante protagonista do próprio percurso formativo e estimulam processos reflexivos que são fundamentais em um contexto educacional marcado pela presença crescente de ferramentas de GenAI.

Entre essas metodologias, destaca-se o *Problem Based Learning* (PBL), que se organiza em torno da análise de situações complexas e da condução de processos investigativos pelos estudantes. O PBL incentiva a formulação de hipóteses, a busca por evidências, a argumentação fundamentada e a reflexão contínua, todos elementos essenciais para o desenvolvimento do pensamento crítico (WALSH, 2005; GIL, 2000). Essas características tornam o PBL especialmente pertinente em um cenário no qual ferramentas de GenAI influenciam a produção e a circulação de informações, exigindo que docentes e estudantes adotem uma postura investigativa e criteriosa. Nesse sentido, a convergência entre GenAI e PBL evidencia a necessidade de orientar professores universitários na articulação entre práticas pedagógicas investigativas e tecnologias inteligentes, de modo a promover aprendizagens ativas, éticas e alinhadas às demandas contemporâneas do ensino superior.

1.1 MOTIVAÇÃO

A motivação para este trabalho decorre da constatação de que a rápida disseminação de ferramentas de GenAI no ensino superior tem impactado diretamente a prática docente, especialmente nos processos de planejamento e condução das atividades pedagógicas. Embora essas tecnologias sejam amplamente utilizadas por estudantes e reconhecidas por professores como potencialmente úteis, sua incorporação ao ensino superior tem ocorrido, em muitos casos, de forma pouco estruturada, sem articulação clara com objetivos pedagógicos ou metodologias de ensino consolidadas (SHAHZAD et al., 2025; ISMAIL et al., 2023). Esse cenário é agravado pela ausência de diretrizes institucionais que orientem o uso educacional da GenAI de maneira segura, crítica e alinhada às finalidades formativas do ensino superior.

Nesse contexto, a adoção de metodologias ativas torna-se particularmente relevante, uma vez que essas abordagens privilegiam a participação ativa dos estudantes, a investigação, a colaboração e o desenvolvimento do pensamento crítico. Entre elas, o Problem-Based Learning destaca-se por estruturar o processo de ensino-aprendizagem em torno de problemas complexos, exigindo mediação pedagógica cuidadosa e decisões didáticas intencionais por parte do professor. A presença crescente da GenAI nesse cenário amplia tanto as possibilidades de apoio ao trabalho docente quanto os riscos de uso inadequado da tecnologia, podendo comprometer princípios centrais do PBL, como a autonomia discente e o caráter investigativo da aprendizagem, caso não haja orientação pedagógica explícita.

Diante dessa lacuna, este trabalho é motivado pela necessidade de oferecer suporte estruturado aos professores universitários no uso da GenAI no contexto do PBL, de forma pedagógica, reflexiva e responsável. Em vez de propor soluções centradas em ferramentas específicas ou aplicações isoladas, a pesquisa busca apoiar a tomada de decisão docente por meio da construção de artefatos educacionais que traduzam princípios pedagógicos em orientações acionáveis para o planejamento e a condução de atividades de aprendizagem ativa. Ao adotar uma abordagem de Design Science Research (DSR), o estudo parte do diagnóstico de desafios concretos enfrentados por docentes e avança para o desenvolvimento e a avaliação de estratégias de suporte à prática pedagógica, contribuindo para o uso mais consciente e pedagogicamente orientado da GenAI no ensino superior.

1.2 OBJETIVOS

1.2.1 Objetivo Geral

Investigar como a Inteligência Artificial Generativa pode dar suporte a docentes no planejamento e condução de atividades de ensino, na educação superior, utilizando a metodologia de Problem Based Learning.

1.2.2 Objetivos Específicos

- (O1) Compreender as percepções, necessidades e desafios dos professores universitários na aplicação da metodologia PBL durante o planejamento e condução das atividades, por meio de um diagnóstico composto por *Rapid Literature Review* (RLR) e por uma atividade exploratória com docentes.
- (O2) Analisar como a GenAI pode dar suporte ao trabalho docente nas diferentes etapas do PBL, identificando estratégias de apoio ao planejamento e à condução das atividades pedagógicas.
- (O3) Propor e avaliar formas de dar suporte à prática docente no PBL por meio da GenAI, com base nas contribuições dos professores participantes e nos resultados observados ao longo do DSR.

1.3 ESCOPO

O escopo da pesquisa concentra-se no uso da GenAI como forma de dar suporte ao trabalho docente no contexto do ensino superior, tomando o PBL como metodologia pedagógica de referência. O foco recai especificamente sobre as necessidades, percepções e desafios enfrentados por professores universitários no planejamento e na condução de atividades pedagógicas mediadas por GenAI. Nesse contexto, o estudo busca compreender como essas tecnologias podem apoiar a prática docente e propor orientações que auxiliem a tomada de decisão pedagógica em ambientes de aprendizagem ativa.

Não constitui escopo do estudo avaliar o impacto da GenAI no desempenho acadêmico dos estudantes, mensurar ganhos de aprendizagem, comparar resultados educacionais entre

turmas ou conduzir estudos longitudinais em sala de aula. Da mesma forma, a pesquisa não se propõe a desenvolver, comparar ou validar ferramentas específicas de GenAI sob critérios técnicos, computacionais ou de eficiência algorítmica, nem a analisar políticas institucionais de adoção tecnológica ou aspectos regulatórios mais amplos.

1.4 ESTRUTURA DO TRABALHO

Esta dissertação está estruturada em seis capítulos. O Capítulo 1 apresenta o contexto da pesquisa, a motivação, a pergunta de pesquisa, os objetivos e o escopo do estudo. O Capítulo 2 reúne a fundamentação teórica sobre GenAI, metodologias ativas e PBL, além de discutir os trabalhos relacionados e as lacunas identificadas na literatura. O Capítulo 3 descreve a abordagem metodológica adotada, contemplando a RLR, o DSR, o contexto da investigação, os participantes e os procedimentos empregados para coleta e análise dos dados. O Capítulo 4 apresenta os principais resultados obtidos ao longo da investigação, incluindo o desenvolvimento do *guideline* para o uso de GenAI no PBL e do *chatbot tutor* voltado ao apoio aos docentes. O Capítulo 5 dedica-se à discussão dos resultados à luz da literatura e do contexto estudado, destacando as contribuições do trabalho, suas implicações para a prática docente e os limites da pesquisa. Por fim, o Capítulo 6 sintetiza as principais conclusões, retoma a pergunta de pesquisa e os objetivos estabelecidos, apresenta as limitações do estudo e aponta possibilidades para investigações futuras.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 GENAI NO ENSINO SUPERIOR

A ascensão da GenAI tem provocado mudanças significativas no contexto do ensino superior, especialmente a partir da popularização de modelos baseados em grandes modelos de linguagem, como o ChatGPT. Essas tecnologias passaram a ser amplamente utilizadas por estudantes para auxiliar em atividades acadêmicas, como escrita, *brainstorming*, síntese de informações e resolução de problemas, o que tem intensificado o debate sobre seus impactos nos processos de ensino e aprendizagem (CHAN; HU, 2023; OGUNLEYE et al., 2024).

Estudos recentes indicam que estudantes universitários tendem a perceber a GenAI de forma predominantemente positiva, reconhecendo seu potencial para oferecer suporte personalizado à aprendizagem, auxiliar na organização de ideias e facilitar o acesso à informação. Ao mesmo tempo, são recorrentes preocupações relacionadas à precisão das respostas geradas, à possibilidade de dependência excessiva da tecnologia, a questões éticas e à integridade acadêmica (CHAN; HU, 2023). Esses achados evidenciam que o uso da GenAI no ensino superior já se encontra incorporado às práticas estudantis, independentemente da existência de diretrizes institucionais claras.

Do ponto de vista institucional e pedagógico, a literatura aponta que a introdução da GenAI tem impactado práticas de ensino, avaliação e formulação de políticas educacionais. Ferramentas como ChatGPT e sistemas de geração de imagens têm potencial para apoiar docentes no planejamento de atividades, na criação de materiais didáticos e na diversificação de estratégias pedagógicas, mas também levantam questionamentos sobre a validade das avaliações tradicionais e a necessidade de revisão de políticas acadêmicas (CHIU, 2024).

Apesar do crescimento acelerado das pesquisas sobre o uso da GenAI no ensino superior, ainda não há consenso sobre diretrizes pedagógicas consolidadas para sua adoção. A literatura destaca a ausência de orientações claras que auxiliem professores e instituições a integrar essas tecnologias de forma alinhada aos objetivos educacionais, especialmente no que se refere ao equilíbrio entre o apoio à aprendizagem e a preservação do desenvolvimento de competências cognitivas e críticas (OGUNLEYE et al., 2024).

Nesse cenário, autores defendem que a GenAI deve ser compreendida como uma ferramenta de apoio às práticas pedagógicas, e não como substituta do papel docente. Sua integração efetiva depende de mediação pedagógica, desenvolvimento profissional dos professores e alinhamento

com os valores e objetivos institucionais do ensino superior (EAGER; BRUNTON, 2023). Assim, a discussão sobre o uso da GenAI no contexto educacional sai do questionamento sobre sua adoção e vai para a necessidade de orientar sobre como essa adoção pode ocorrer de maneira consciente, ética e pedagogicamente fundamentada.

2.2 DESAFIOS PEDAGÓGICOS DA INTELIGÊNCIA ARTIFICIAL GENERATIVA

A incorporação da GenAI no ensino superior tem evidenciado desafios pedagógicos que ultrapassam aspectos meramente tecnológicos. Em particular, o uso disseminado de ferramentas como o ChatGPT tem gerado questionamentos sobre mudanças nas práticas de ensino e avaliação (BOWER et al., 2024). Ao mesmo tempo, docentes ainda não se sentem suficientemente preparados para orientar o uso dessas ferramentas em atividades pedagógicas, o que reforça a necessidade de desenvolvimento profissional específico para esse contexto (BRANDÃO; PEDRO; ZAGALO, 2024).

Um dos desafios mais recorrentes diz respeito à integridade acadêmica e à autoria do trabalho discente. A capacidade dos modelos de linguagem de produzir textos plausíveis e sofisticados intensificou preocupações relacionadas ao plágio e ao uso indevido de conteúdo gerado por IA (WILLIAMS, 2023). Nesse debate, a literatura também aponta limites importantes das ferramentas de detecção de textos gerados por IA; estudos têm concluído que, atualmente, é inviável detectar de forma confiável conteúdos produzidos com assistência de IA, em função de taxas elevadas de falsos positivos e negativos (FARRELLY; BAKER, 2023). Evidências recentes indicam que não há ferramentas capazes de distinguir, com alto grau de certeza, textos gerados por IA daqueles produzidos por humanos (BRANDÃO; PEDRO; ZAGALO, 2024), o que impacta diretamente na forma como instituições e docentes podem conduzir processos avaliativos.

Outro desafio pedagógico refere-se ao risco de dependência excessiva dessas tecnologias por parte dos estudantes. A personalização proporcionada pela GenAI pode levar estudantes a adotarem uma postura mais passiva, com maior dependência da ferramenta (KONG; YANG, 2024). Em situações de uso recorrente e pouco orientado, esse fenômeno pode ultrapassar a dependência comportamental e configurar um processo de enfraquecimento do desenvolvimento de habilidades cognitivas fundamentais, como argumentação, análise crítica e resolução autônoma de problemas. Além disso, discussões sobre o uso de *chatbots* na educação destacam a tensão entre promover autonomia discente e evitar impactos negativos na autoeficácia acadêmica, sobretudo quando há sobreuso da tecnologia (WILLIAMS, 2023). Nesse contexto,

coloca-se a questão sobre quais funções cognitivas podem ser delegadas à GenAI e em que medida essa delegação é formativamente adequada, especialmente em diferentes etapas do desenvolvimento acadêmico. Esse cenário reforça a importância de mediação pedagógica docente, orientando o uso da GenAI como suporte ao processo de aprendizagem, e não como substituto do esforço cognitivo.

As implicações éticas da GenAI também se consolidam como um eixo central na literatura sobre ensino superior. Entre os principais pontos discutidos estão os riscos associados ao tratamento de dados sensíveis de estudantes (WILLIAMS, 2023) e a possibilidade de vieses algorítmicos reproduzirem desigualdades e estereótipos (WILLIAMS, 2023). Revisões e discussões críticas também reforçam que sistemas de IA não operam como entidades neutras e que seu funcionamento depende de dados de treinamento e escolhas humanas, o que exige atenção a aspectos de justiça e equidade em práticas educacionais mediadas por IA (FARRELLY; BAKER, 2023). Nesse contexto, a inexistência de políticas claras e orientações consistentes pode ampliar a insegurança docente e levar a respostas institucionais fragmentadas (FARRELLY; BAKER, 2023).

Diante desse conjunto de desafios, a literatura converge na defesa de iniciativas de formação docente voltadas ao uso pedagógico de GenAI. Revisões integrativas sobre desenvolvimento profissional destacam que a capacitação de professores deve contemplar dimensões técnicas, pedagógicas e críticas para apoiar a integração da GenAI às práticas de ensino (BRANDÃO; PEDRO; ZAGALO, 2024). Além disso, evidências de pesquisas com docentes indicam que compreender as crenças e percepções de professores sobre a GenAI é relevante, pois tais aspectos influenciam a forma como mudanças no ensino e na avaliação tendem a ocorrer no ensino superior (BOWER et al., 2024).

2.3 METODOLOGIAS ATIVAS E PROBLEM-BASED LEARNING

As metodologias ativas têm sido discutidas no ensino superior como alternativas ao modelo tradicional centrado na exposição do professor, no qual o estudante tende a adotar uma posição predominantemente receptiva (DOGANI, 2023). Em contraste, abordagens de aprendizagem ativa propõem que o estudante participe do processo de aprendizagem, como figura central, por meio de atividades que exigem interação com o conteúdo, como discussão, resolução de problemas e trabalho colaborativo, deslocando o foco da aula para práticas de engajamento e reflexão (TAJRA et al., 2021; GIL, 2000). Nessa perspectiva, o professor deixa de atuar apenas como transmissor de informações e passa a assumir um papel de orientação, estruturando

tarefas que convidam o estudante a agir e refletir sobre o que foi feito e aprendido (TAJRA et al., 2021; GIL, 2000).

Além de sua fundamentação pedagógica, a adoção de metodologias ativas também é sustentada por evidências empíricas no contexto do ensino superior, especialmente em áreas de graduação com histórico de desigualdades de desempenho. Uma síntese de estudos empíricos conduzidos predominantemente em cursos das áreas de ciência, tecnologia, engenharia e matemática indica que intervenções baseadas em aprendizagem ativa estiveram associadas, em média, à redução de lacunas de desempenho e de aprovação entre estudantes de grupos sub-representados e seus pares. Os resultados reportados mostram-se sensíveis tanto à intensidade quanto ao tempo dedicado às atividades desenvolvidas em sala de aula (THEOBALD et al., 2020). Esses achados reforçam que a efetividade de metodologias ativas depende não apenas de sua adoção nominal, mas também do modo como são implementadas, particularmente no que se refere ao espaço real dado à participação do estudante e ao uso de atividades que demandem análise e tomada de decisão (THEOBALD et al., 2020).

Entre as metodologias ativas, o PBL destaca-se por estruturar a aprendizagem a partir de um problema que orienta e impulsiona o estudo. Nessa abordagem, o objetivo central não é simplesmente resolver o problema apresentado, mas utilizá-lo como base para identificar necessidades de aprendizagem, articular hipóteses iniciais, buscar informações relevantes e integrar conhecimentos ao longo do percurso (WALSH, 2005). Em resumo, o problema atua como eixo organizador da aprendizagem, favorecendo estudo autônomo e construção coletiva de entendimento em pequenos grupos (WALSH, 2005; GIL, 2000).

A literatura também enfatiza que a qualidade do problema e do contexto em que ele é apresentado influencia diretamente a profundidade do aprendizado. Problemas mais ricos e próximos de situações reais tendem a exigir integração de conhecimentos, tomada de decisão e argumentação, além de promover colaboração e maior envolvimento dos estudantes no processo (SMITH et al., 2022). Essa característica é particularmente relevante no ensino superior, onde a formação precisa contemplar competências que extrapolam a memorização de conteúdos e incluem raciocínio, comunicação e desenvolvimento de pensamento crítico (SMITH et al., 2022).

A implementação do PBL também depende do papel do professor como tutor ou facilitador. Em vez de conduzir a aula por exposição, o docente apoia a dinâmica do grupo por meio de perguntas, acompanhamento do progresso e estímulo à participação equilibrada, evitando substituir o processo de investigação por explicações prontas (WALSH, 2005). Esse equilíbrio é importante porque o tutor precisa, ao mesmo tempo, impedir que o grupo se desvie do foco e

preservar a autonomia discente, mantendo o trabalho orientado por objetivos de aprendizagem claros (WALSH, 2005).

Do ponto de vista dos resultados educacionais, há evidências de que intervenções baseadas em PBL podem contribuir para o desenvolvimento do pensamento crítico no ensino superior. Uma meta-análise identificou um efeito global positivo do PBL na aquisição do pensamento crítico, indicando que, em conjunto, os estudos apontam ganhos associados à adoção dessa metodologia (LIU; PÁSZTOR, 2022). Assim, ao articular problemas, investigação e reflexão, o PBL se consolida como uma abordagem coerente com objetivos formativos voltados a competências de ordem superior.

Por fim, a centralidade da investigação, da troca entre pares e do estudo autônomo torna o PBL particularmente sensível à presença de tecnologias que fornecem respostas prontas, sem mediação. Em contextos permeados por ferramentas capazes de produzir textos e soluções rapidamente, torna-se necessário orientar sobre como essas tecnologias podem ser incorporadas sem reduzir o esforço investigativo e o papel formativo do problema, preservando a lógica de aprendizagem ativa que sustenta o PBL.

2.4 INTEGRAÇÃO ENTRE PBL E GENAI

A integração entre o PBL e a GenAI tem ganhado relevância no ensino superior, em parte porque *chatbots* de GenAI tornam viável a oferta de suporte conversacional sob demanda, aproximando-se da ideia de tutoria personalizada (KESTIN et al., 2025). Ao mesmo tempo, estudos indicam que ainda existe incerteza sobre como empregar essa tecnologia de modo educacionalmente efetivo quando comparada a práticas pedagógicas consolidadas (KESTIN et al., 2025). Assim, discutir integração, no contexto do PBL, implica analisar como a GenAI pode apoiar etapas investigativas e reflexivas sem deslocar o estudante para um papel passivo.

Do ponto de vista pedagógico, pesquisas em contextos de ensino de STEM indicam que a GenAI pode atuar como ferramenta de apoio ao aprendizado conceitual, oferecendo *feedback* personalizado e auxiliando na identificação e correção de concepções equivocadas (FATHI et al., 2025). Essas contribuições tendem a ser mais efetivas quando a interação com a GenAI é integrada a atividades estruturadas, como tarefas pré-aula, avaliações formativas e resolução de problemas em sala (FATHI et al., 2025). Em ambientes de aprendizagem ativa, a GenAI também é discutida como um recurso capaz de fortalecer processos de reflexão sobre o próprio aprendizado e o desenvolvimento de competências tecnológicas, desde que incorporada de

forma intencional ao desenho pedagógico (ELSAYARY, 2024).

Entretanto, a integração entre GenAI e PBL exige cautela, pois há evidências de que o uso intensivo dessas ferramentas pode reorganizar o processo de aprendizagem em torno da tecnologia, reduzindo a participação do estudante em processos metacognitivos. Esse fenômeno tem sido descrito como *metacognitive laziness*, caracterizando uma dependência da ferramenta que pode comprometer a autorregulação e o esforço cognitivo do aprendiz (FAN et al., 2025). Esse risco é particularmente relevante no PBL, cuja efetividade depende da investigação ativa e da identificação de lacunas de conhecimento ao longo do processo.

Diante desse cenário, a mediação docente assume papel central na integração da GenAI ao PBL. Cabe ao professor planejar quando e como a ferramenta será utilizada, formulando tarefas e perguntas que estimulem análise, comparação de alternativas e justificativa de decisões, em vez da simples aceitação de respostas geradas automaticamente (KESTIN et al., 2025). Essa necessidade de planejamento reforça que a efetividade da GenAI em contextos educacionais está associada à sua integração orientada por princípios pedagógicos claros, e não ao uso espontâneo da tecnologia.

Por fim, a literatura também discute o uso da GenAI no apoio ao desenho de avaliações mais autênticas e alinhadas aos objetivos de aprendizagem. Ferramentas baseadas em GenAI podem auxiliar professores na elaboração de atividades avaliativas e no alinhamento entre objetivos, atividades e critérios de avaliação, bem como oferecer análises e sugestões preliminares. Contudo, tais sistemas devem atuar apenas como instrumentos de apoio informativo, cabendo ao docente a tomada de decisões pedagógicas críticas e a responsabilidade final sobre o processo avaliativo (SALINAS-NAVARRO et al., 2024). Em síntese, os estudos indicam que a integração entre PBL e GenAI é promissora quando orientada por diretrizes explícitas, capazes de preservar o caráter investigativo da metodologia e mitigar riscos associados à dependência cognitiva (FAN et al., 2025; FATHI et al., 2025).

2.5 ARTEFATOS EDUCACIONAIS COMO ESTRATÉGIA DE INTERVENÇÃO

Em pesquisas educacionais orientadas à intervenção, a construção de artefatos educacionais tem sido compreendida como uma estratégia para transformar decisões pedagógicas em recursos concretos, passíveis de uso e avaliação em contextos reais. Estudos sobre qualidade das práticas formativas destacam que intervenções educacionais eficazes demandam não apenas princípios teóricos, mas também critérios, estruturas de apoio e instrumentos que orientem sua

implementação (WIESE; HATLEVIK; DAZA, 2024). Nessa perspectiva, os artefatos atuam como mediadores entre a teoria e a prática docente, permitindo que escolhas pedagógicas sejam explicitadas, testadas e aprimoradas ao longo do processo.

Artefatos educacionais podem assumir diferentes formatos, como guidelines, modelos de apoio, frameworks, instrumentos avaliativos ou ferramentas digitais, com o objetivo de auxiliar professores no planejamento, na condução e na avaliação de experiências de aprendizagem. Revisões sistemáticas sobre inovações digitais no ensino superior indicam que a efetividade dessas iniciativas depende fortemente do desenho pedagógico que sustenta o uso da tecnologia e da clareza com que esse desenho é documentado e compartilhado (CHOI-LUNDBERG et al., 2023). Ao explicitar racionalidades, etapas e condições de aplicação, artefatos educacionais favorecem a análise crítica, a adaptação a diferentes contextos e a transferência de práticas entre docentes.

No contexto da integração da GenAI ao ensino superior, a literatura aponta que a ausência de orientações claras tende a gerar cautela, insegurança e respostas fragmentadas por parte dos docentes. Análises de políticas institucionais e recursos educacionais relacionados ao uso de ferramentas como o ChatGPT indicam que professores frequentemente percebem lacunas em termos de diretrizes pedagógicas, formação e suporte institucional (WANG et al., 2024). Nesse cenário, a criação de artefatos como guidelines e materiais estruturados pode contribuir para explicitar objetivos, delimitar usos apropriados e apoiar uma adoção mais consciente da GenAI em práticas pedagógicas.

Além de guias pedagógicos, instrumentos e modelos conceituais também podem operar como artefatos de intervenção ao oferecerem meios para diagnosticar práticas e orientar processos de mudança. No campo da avaliação digital no ensino superior, por exemplo, estudos têm proposto instrumentos validados para mapear conhecimentos e percepções docentes, destacando seu papel no apoio a transformações pedagógicas diante do avanço de tecnologias digitais (VIBERG et al., 2024). Esses instrumentos evidenciam que artefatos educacionais não se limitam a produtos finais, mas funcionam como estruturas que organizam decisões pedagógicas e sustentam processos reflexivos.

Dessa forma, a utilização de artefatos educacionais como estratégia de intervenção justifica-se por oferecer aos docentes suportes concretos para mediar a incorporação da GenAI sem descaracterizar princípios pedagógicos fundamentais. Ao materializar orientações, critérios e exemplos de uso alinhados a objetivos educacionais, tais artefatos podem reduzir a insegurança docente, favorecer a mediação pedagógica e apoiar a adoção de práticas mais

coerentes e avaliáveis no ensino superior (WIESE; HATLEVIK; DAZA, 2024; CHOI-LUNDBERG et al., 2023).

3 METODOLOGIA

Esta pesquisa adota a abordagem metodológica de DSR, inspirada nos ciclos iterativos de uma pesquisa-ação (FILIPPO; ROQUE; PEDROSA, 2018) compostos pelas etapas de (1) diagnóstico, (2) planejamento da ação, (3) intervenção, (4) avaliação e (5) reflexão, sendo a etapa de intervenção o desenvolvimento dos artefatos educacionais. Essa escolha metodológica se justifica pela natureza prática e colaborativa da investigação, cujo objetivo é compreender e intervir nos desafios enfrentados por professores universitários diante do uso crescente de inteligência artificial generativa por estudantes, especialmente após o lançamento do ChatGPT (OpenAI, 2022).

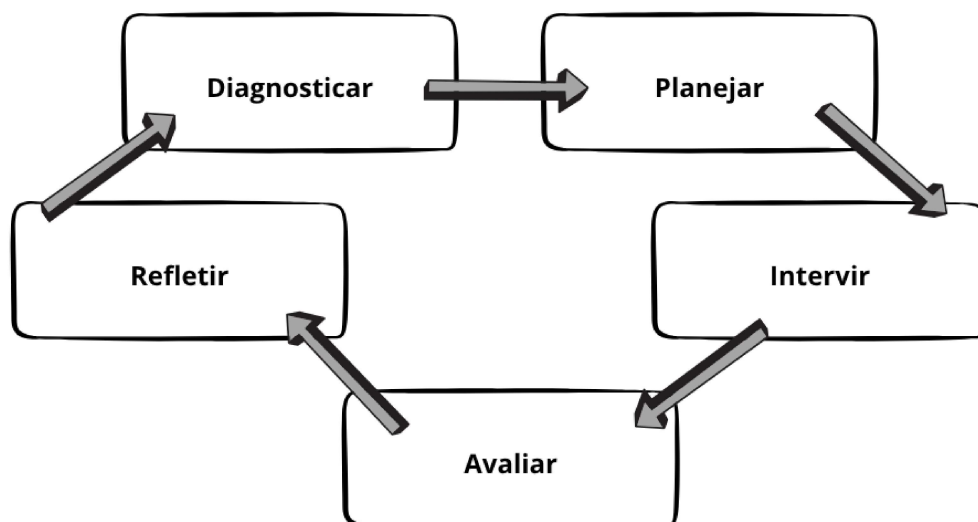


Figura 1 – Ciclo da pesquisa-ação

Fonte: Adaptado de (FILIPPO; ROQUE; PEDROSA, 2018)).

3.1 DIAGNÓSTICO DO PROBLEMA

A investigação iniciou-se com a etapa de **diagnóstico**, cujo objetivo foi identificar os principais desafios associados à presença da Inteligência Artificial Generativa (GenAI) no contexto educacional universitário, em consonância com o objetivo específico O1 desta pesquisa. Essa etapa foi estruturada em duas estratégias complementares: (i) a condução de uma *Rapid Literature Review* (RLR), voltada ao mapeamento sistematizado de desafios reportados na literatura recente, e (ii) a realização de um seminário exploratório com docentes universitários,

com a finalidade de validar, contextualizar e priorizar esses desafios a partir da prática pedagógica.

3.1.1 Rapid Literature Review como instrumento de diagnóstico

A RLR foi conduzida com base no método proposto por (CARTAXO; PINTO; SOARES, 2020), sendo escolhida por sua natureza pragmática, ágil e orientada à solução de problemas reais. Essa abordagem é particularmente adequada quando há necessidade de produzir evidências teóricas em curto prazo, com foco na aplicabilidade em contextos de pesquisa aplicada, como o presente estudo. Diferentemente das revisões sistemáticas tradicionais, a RLR admite flexibilizações metodológicas documentadas, como o uso de número reduzido de bases de dados ou análise individual, sem comprometer a transparência científica (GANANN; CILISKA; THOMAS, 2010). A RLR foi estruturada em três fases principais: (1) planejamento, (2) execução e (3) síntese dos resultados.

1. **Planejamento**, que envolveu a formulação da pergunta de pesquisa, seleção da base de dados, elaboração da estratégia de busca, e definição dos critérios de inclusão e exclusão;
2. **Execução**, com a busca, triagem e extração dos dados dos estudos selecionados, priorizando publicações recentes e com foco empírico;
3. **Síntese dos resultados**, com o agrupamento dos principais desafios encontrados e a apresentação dos achados de forma objetiva e útil para profissionais da educação superior.

Na etapa de *planejamento*, foi elaborada a pergunta de pesquisa que norteou o processo de revisão: *Quais são os principais desafios enfrentados por professores universitários na adoção da inteligência artificial com fins pedagógicos?* A construção dessa questão seguiu o modelo PICOC (Population, Intervention, Comparison, Outcome, Context), conforme diretrizes de (KEELE et al., 2007), como mostra a Tabela 1.

A busca foi realizada nas bases *Google Scholar* e *Scopus*, utilizando operadores booleanos e os termos: ("Artificial Intelligence"OR "AI"OR "LLM") AND ("Higher Education"OR "University Professors") AND ("Challenges"OR "Barriers"OR "Adoption"). Os critérios de inclusão priorizaram estudos publicados entre 2023 e 2025, com foco no ensino superior e que abordassem diretamente barreiras à adoção de IA por docentes. Foram excluídos

Componente	Descrição
Population (P)	Professores universitários
Intervention (I)	Uso de IA para fins pedagógicos
Comparison (C)	Práticas tradicionais de ensino sem IA
Outcome (O)	Identificação dos principais desafios na adoção da IA
Context (C)	Ensino superior

Tabela 1 – Estrutura PICOC utilizada na RLR.

artigos sobre educação básica, abordagens excessivamente técnicas e estudos desconectados de práticas pedagógicas.

O processo de seleção dos estudos ocorreu de forma incremental, em consonância com a natureza exploratória e pragmática da Rapid Literature Review. Inicialmente, foram identificados aproximadamente 50 artigos potencialmente relevantes a partir da aplicação da estratégia de busca definida nas bases selecionadas.

Na primeira etapa de triagem, baseada na leitura dos títulos e resumos, foram selecionados 34 estudos que apresentavam aderência temática ao escopo da pesquisa, contemplando o uso da inteligência artificial no ensino superior e a discussão de desafios enfrentados por professores universitários. Os estudos excluídos nessa etapa não abordavam diretamente o contexto educacional universitário ou tratavam a inteligência artificial sob uma perspectiva estritamente técnica.

Em uma segunda etapa, foi realizada a leitura das seções de metodologia e conclusões dos 34 estudos selecionados, com o objetivo de avaliar a relevância empírica e pedagógica das investigações. Como resultado dessa análise, 27 estudos foram considerados adequados para uma leitura aprofundada, compondo o corpus analítico da revisão. Esses estudos apresentavam descrição metodológica clara, discussão consistente dos desafios associados à adoção da inteligência artificial no ensino superior e relação direta com a prática docente.

Durante a etapa de extração de dados, foram sistematicamente coletados, a partir dos 27 estudos analisados, três elementos centrais: (1) os desafios relatados na adoção da inteligência artificial por professores universitários, (2) as soluções ou recomendações propostas para enfrentar esses desafios e (3) as metodologias de pesquisa empregadas, incluindo abordagens qualitativas, quantitativas, revisões teóricas e estudos de caso.

Por fim, na fase de *síntese*, os dados relativos aos desafios identificados nos estudos analisados foram submetidos à análise de conteúdo, na modalidade análise temática, conforme proposta por (BARDIN, 1977). As unidades de registro associadas aos desafios da adoção da

inteligência artificial no ensino superior foram codificadas e agrupadas em categorias temáticas.

3.1.2 Seminário com docentes universitários como estratégia exploratória de apoio ao diagnóstico

Com base nos resultados obtidos na Rapid Literature Review, foi realizada uma atividade de caráter exploratório na forma de um seminário com docentes universitários, com o objetivo de validar, complementar e priorizar os desafios identificados na literatura. Essa atividade integrou a etapa de diagnóstico da DSR e foi concebida como uma estratégia de escuta ativa e diálogo profissional, sem caráter de estudo empírico formal.

O seminário ocorreu a partir de um convite institucional para uma apresentação sobre inteligência artificial generativa a professores universitários, reunindo docentes de diferentes áreas do ensino superior. A atividade teve formato expositivo-dialogado, com apresentação dos desafios mapeados na revisão da literatura, seguida de momentos de manifestação espontânea dos participantes acerca de suas percepções e experiências profissionais. As contribuições foram consideradas de forma agregada e não identificável, sendo utilizadas exclusivamente para aprofundamento interpretativo do diagnóstico inicial.

Como parte da dinâmica do seminário, foi utilizada a ferramenta Mentimeter para apoiar a priorização individual dos desafios previamente identificados, permitindo a manifestação anônima das percepções dos participantes. A ferramenta foi empregada sem qualquer forma de identificação pessoal, e os dados gerados foram tratados de maneira agregada e descritiva, com a finalidade exclusiva de subsidiar o diagnóstico e orientar a etapa de planejamento da ação.

Adicionalmente, as discussões coletivas foram registradas por meio de gravação de áudio, previamente informada aos participantes, com o objetivo de apoiar a sistematização reflexiva das contribuições apresentadas. O pesquisador realizou a escuta integral das gravações, empregando um procedimento de organização interpretativa das percepções manifestadas, voltado à identificação de recorrências e pontos de convergência entre os relatos, sempre de forma agregada e não identificável. Esse tratamento teve caráter exclusivamente auxiliar e reflexivo, não envolvendo codificação sistemática, categorização formal ou aplicação de técnicas de análise qualitativa estruturadas. Esses registros foram utilizados apenas como suporte para a organização das percepções manifestadas de forma coletiva, sem identificação individual, não sendo tratados como dados empíricos centrais da pesquisa.

Em consonância com a Resolução CNS nº 510/2016, a atividade se enquadra como pesquisa

de opinião pública com participantes não identificados e como aprofundamento teórico de situações emergentes da prática profissional, não implicando riscos adicionais aos participantes nem exposição de dados sensíveis. Dessa forma, os registros produzidos no seminário foram utilizados exclusivamente como subsídio qualitativo para orientar o planejamento da ação e o desenvolvimento da proposta de intervenção.

De forma a explicitar a natureza metodológica dos elementos utilizados nesta etapa, destaca-se que o presente estudo distingue diferentes níveis de evidência e insumos analíticos. A Rapid Literature Review constitui a principal fonte de dados teóricos sistematizados do diagnóstico, oferecendo evidências documentadas na literatura recente. A atividade exploratória realizada por meio do seminário com docentes universitários não foi tratada como produção de dados empíricos formais, mas como subsídio interpretativo e reflexivo, destinado a contextualizar, complementar e priorizar os desafios identificados na literatura. As percepções coletadas nesse momento foram utilizadas de forma agregada, sem categorização sistemática, servindo como apoio à tomada de decisão na etapa de planejamento da ação. Essa distinção visa garantir clareza metodológica quanto ao estatuto dos elementos analisados, em consonância com os princípios da DSR.

3.2 PLANEJAMENTO DA AÇÃO

A etapa de diagnóstico permitiu identificar e priorizar desafios pedagógicos, éticos e institucionais associados ao uso da inteligência artificial generativa no ensino superior, a partir de evidências da literatura e de um subsídio exploratório oriundo da escuta de docentes. Esses insumos forneceram a base interpretativa para a etapa seguinte da DSR, na qual se torna necessário definir uma proposta de intervenção coerente com o contexto investigado.

Dessa forma, o planejamento da ação é compreendido como o momento em que os achados do diagnóstico são convertidos em decisões de projeto, delimitando objetivos operacionais, princípios pedagógicos, formato da intervenção e critérios de qualidade que orientarão sua implementação. Em vez de propor uma solução genérica, essa etapa buscou estabelecer uma relação explícita entre os desafios identificados no diagnóstico e as decisões pedagógicas adotadas, como a definição do PBL como eixo estruturante e a escolha por artefatos educacionais de apoio à prática docente.

No contexto desta pesquisa, o pesquisador assumiu um papel ativo na sistematização desses insumos e na definição da intervenção. Tal atuação não visa neutralidade absoluta, mas

intencionalidade reflexiva, reconhecendo o envolvimento do pesquisador como parte constitutiva do processo investigativo. Para mitigar possíveis vieses, foram adotadas estratégias como revisão por pares, validação do *guideline* por especialistas externos e explicitação dos critérios utilizados nas avaliações.

A etapa de **planejamento da ação** teve como objetivo definir, de forma fundamentada e coerente com os achados do diagnóstico, a proposta de intervenção a ser desenvolvida nos ciclos subsequentes da pesquisa. Essa etapa buscou transformar os desafios identificados tanto na RLR quanto na atividade exploratória com docentes universitários em diretrizes práticas para a construção de uma solução educacional aplicável ao contexto do ensino superior.

O planejamento da ação foi conduzido como um processo reflexivo, no qual evidências da literatura e percepções da prática docente foram articuladas para orientar decisões pedagógicas concretas. Nesse contexto, optou-se por adotar uma metodologia ativa de ensino que permitisse integrar o uso da GenAI sem comprometer os objetivos educacionais, tomando o PBL como eixo estruturante da proposta pedagógica.

A escolha do PBL como eixo estruturante não decorre da suposição de que outras metodologias ativas seriam inadequadas ao contexto da GenAI, mas do entendimento de que o PBL enfatiza, de forma particularmente explícita, a investigação orientada por problemas, a construção coletiva do conhecimento e o protagonismo discente. Tais características tornam essa abordagem especialmente sensível tanto aos potenciais quanto aos riscos da GenAI, uma vez que a tecnologia pode fortalecer o processo investigativo ou, se utilizada de forma acrítica, esvaziá-lo. Assim, o PBL foi adotado como recorte metodológico capaz de tornar mais visível e analisável a tensão entre apoio tecnológico e desenvolvimento do pensamento crítico.

Com base nesse enquadramento, o planejamento da ação foi estruturado em eixos pedagógicos complementares, utilizados para orientar as decisões da intervenção. Esses eixos foram definidos a partir da integração entre os desafios identificados no diagnóstico e as demandas práticas da atuação docente no contexto do PBL com apoio da GenAI. Em conjunto, eles orientam o planejamento e a condução de atividades pedagógicas mediadas por inteligência artificial no ensino superior.

O eixo **(1) metodologias ativas de ensino, com ênfase no PBL**, refere-se à necessidade de estruturar aulas que valorizem investigação, autonomia e protagonismo discente em um cenário no qual os estudantes já dispõem de ferramentas de inteligência artificial generativa. A escolha do PBL responde diretamente ao desafio de evitar que a GenAI seja utilizada apenas como atalho para respostas prontas, deslocando o foco da aprendizagem para o processo

investigativo.

O eixo **(2) competências digitais docentes** diz respeito ao que o professor precisa saber e desenvolver para conduzir atividades pedagógicas mediadas por tecnologia de forma intencional. Esse eixo considera que a integração da GenAI ao ensino exige do docente mais do que domínio técnico das ferramentas, envolvendo também a capacidade de tomar decisões pedagógicas conscientes sobre quando, como e por que utilizá-las.

O eixo **(3) avaliação formativa em contextos mediados por tecnologia** está relacionado à necessidade de acompanhar o processo de aprendizagem dos estudantes em atividades que envolvem o uso da GenAI. Esse eixo orienta a adoção de estratégias avaliativas que valorizem o percurso investigativo, a participação e a reflexão dos alunos, e não apenas a análise de produtos finais, que podem ser facilmente influenciados pelo uso de ferramentas automatizadas.

Por fim, o eixo **(4) uso pedagógico da inteligência artificial no ensino superior** refere-se à reflexão sobre como empregar a GenAI como ferramenta de apoio ao ensino sem distorcer os objetivos educacionais das atividades propostas. Esse eixo orienta decisões relacionadas aos limites, cuidados e possibilidades do uso da IA, garantindo que a tecnologia atue como suporte ao processo de aprendizagem e não como substituta do raciocínio dos estudantes.

A partir desses eixos, definiu-se que a intervenção assumiria a forma de dois artefatos complementares: um *guideline* e um chatbot tutor. O *guideline* foi concebido como um material de apoio estruturado, cujo objetivo é orientar o professor universitário na implementação de atividades de PBL integradas à GenAI, oferecendo referências conceituais, exemplos práticos e critérios de uso pedagógico da tecnologia. Sua estrutura busca responder às principais dúvidas identificadas no diagnóstico, especialmente relacionadas à organização das aulas, à condução das etapas do PBL e à avaliação dos estudantes.

O chatbot tutor, por sua vez, foi planejado como um artefato complementar ao *guideline*, com a função de apoiar o professor de forma mais imediata e contextualizada. Enquanto o *guideline* oferece uma visão estruturada e aprofundada da proposta pedagógica, o chatbot permite esclarecer dúvidas pontuais, sugerir caminhos e apoiar decisões ao longo do planejamento e da condução das atividades. A combinação desses dois artefatos visa atender a diferentes necessidades dos docentes, conciliando aprofundamento conceitual e suporte prático no cotidiano do ensino.

O desenho desses artefatos foi guiado por critérios de clareza, aplicabilidade pedagógica e

alinhamento com os eixos definidos no planejamento. Buscou-se evitar soluções excessivamente prescritivas ou dependentes de ferramentas específicas, priorizando orientações flexíveis que possam ser adaptadas a diferentes contextos institucionais, áreas de conhecimento e níveis de ensino.

Ressalta-se que a etapa de planejamento da ação não teve como finalidade produzir resultados empíricos, mas estabelecer as bases teóricas, metodológicas e pedagógicas necessárias para a implementação da intervenção. O desenvolvimento dos artefatos desenvolvidos, bem como a análise de sua contribuição para a prática docente, são abordadas nas etapas subsequentes de intervenção, avaliação e reflexão, conforme o ciclo da DSR adotado neste estudo.

3.3 INTERVENÇÃO

A etapa de intervenção, nesta pesquisa, foi compreendida como a construção de artefatos educacionais, e não como a aplicação direta de uma intervenção em contexto de sala de aula. Assim, ela correspondeu à materialização das decisões tomadas durante o planejamento da ação, resultando no desenvolvimento de dois artefatos educacionais complementares: um *guideline* para apoiar professores universitários na implementação do PBL integrado à GenAI e um chatbot tutor concebido como ferramenta de suporte pedagógico. Ambos os artefatos foram desenvolvidos com o objetivo de responder às dimensões pedagógicas identificadas no diagnóstico e sistematizadas nos eixos teóricos definidos no planejamento da ação.

Importante destacar que a intervenção não foi realizada no contexto de uma disciplina específica previamente estruturada em PBL, nem envolveu a criação e aplicação longitudinal de uma nova disciplina. Os artefatos desenvolvidos foram concebidos como instrumentos de apoio ao planejamento e à condução de atividades pedagógicas, podendo ser utilizados tanto para aprimorar disciplinas já existentes quanto para apoiar a criação de novas experiências estruturadas em PBL no ensino superior.

O *guideline* constituiu o principal artefato da intervenção. Ele foi concebido como um material orientador, de caráter prático, destinado a auxiliar professores universitários na integração do PBL com recursos de GenAI de forma pedagógica, crítica e intencional. Sua estrutura foi diretamente orientada pelos eixos definidos no planejamento da ação, contemplando orientações sobre a organização das atividades de ensino, o papel do professor como tutor no PBL, o uso responsável da GenAI ao longo das etapas da metodologia e estratégias de acompanhamento da aprendizagem.

O *guideline* não possui caráter prescritivo, buscando oferecer referenciais claros, exemplos e possibilidades de adaptação a diferentes contextos institucionais, áreas de conhecimento e níveis de ensino. O desenho do material priorizou critérios de clareza, aplicabilidade pedagógica e alinhamento com os princípios do PBL, evitando dependência de ferramentas específicas ou soluções rigidamente definidas.

Paralelamente ao *guideline*, foi desenvolvido um chatbot tutor com o objetivo de oferecer suporte complementar e mais imediato ao professor durante o planejamento e a condução de atividades de PBL integradas à GenAI. Enquanto o *guideline* atua como um artefato estruturado de consulta, o chatbot foi concebido como uma ferramenta interativa, capaz de responder a dúvidas pontuais, sugerir caminhos pedagógicos e apoiar a tomada de decisão docente de forma contextualizada.

O desenvolvimento do chatbot foi orientado pelos mesmos eixos teóricos que fundamentaram a construção do *guideline*, buscando refletir princípios do PBL e promover o uso pedagógico da GenAI como ferramenta de apoio ao ensino. O escopo do chatbot foi delimitado de modo a evitar respostas prescritivas ou substitutivas do julgamento docente, reforçando seu papel como suporte consultivo e não como agente decisório.

A avaliação dos artefatos desenvolvidos é apresentada na seção seguinte, na qual são descritos os procedimentos, instrumentos e critérios adotados para analisar sua clareza, aplicabilidade pedagógica e aderência aos princípios definidos no planejamento da ação.

Artefato	Objetivo	Forma de avaliação	Crítérios analisados
<i>Guideline</i> de PBL integrado à GenAI	Apoiar professores universitários na integração pedagógica, crítica e intencional do PBL com ferramentas de GenAI	Avaliação por especialistas em PBL e metodologias ativas, por meio de um Survey	Clareza do conteúdo, organização, aplicabilidade pedagógica, aderência aos princípios do PBL e utilidade no contexto docente
Chatbot tutor de apoio pedagógico	Oferecer suporte complementar e interativo ao professor durante o planejamento e a condução de atividades de PBL com GenAI	Experimento controlado conduzido pelo pesquisador, utilizando a abordagem <i>LLM as a Judge</i> , com comparação a sistemas de propósito geral	Aderência ao contexto do PBL, alinhamento aos princípios pedagógicos definidos no planejamento e contextualização das respostas às necessidades docentes

Quadro 1 – Síntese da etapa de intervenção da pesquisa

3.4 AVALIAÇÃO

A etapa de avaliação teve como objetivo analisar os artefatos educacionais desenvolvidos na intervenção, considerando sua qualidade pedagógica, aderência aos princípios definidos no planejamento da ação e comportamento em contextos controlados de uso. Em consonância com a abordagem da DSR, a avaliação foi concebida como um processo analítico e reflexivo, e não como uma mensuração de desempenho em larga escala.

A avaliação foi conduzida por meio de duas estratégias complementares. A primeira consistiu na validação pedagógica do *guideline* por especialistas em PBL e metodologias ativas, visando analisar sua clareza, aplicabilidade e coerência pedagógica. A segunda correspondeu a um experimento controlado para avaliação do chatbot tutor, utilizando a abordagem *LLM-as-a-Judge*, com o objetivo de comparar seu comportamento em diferentes configurações de uso, (1) com acesso ao *guideline*, e (2) sem acesso ao *guideline*.

3.4.1 Survey com Especialistas

A avaliação do *guideline* foi realizada por três especialistas do ensino superior, sendo dois docentes da área de Administração e um docente da área de Sistemas de Informação. Todos possuem experiência na condução de atividades acadêmicas e familiaridade com metodologias ativas, incluindo o Problem-Based Learning, o que contribuiu para uma análise qualificada da clareza, aplicabilidade e aderência pedagógica do artefato.

No caso do *guideline*, a avaliação assumiu caráter de validação pedagógica, sendo conduzida por especialistas com experiência em PBL e metodologias ativas no ensino superior. Essa estratégia foi adotada porque especialistas em PBL possuem experiência e embasamento teórico para avaliar, com maior rigor, a clareza, a aplicabilidade e a aderência pedagógica do material, mesmo diante das restrições temporais e institucionais que inviabilizaram sua aplicação em contextos reais de sala de aula. A coleta de dados ocorreu por meio de um formulário eletrônico, elaborado na plataforma Google Forms, contendo questões fechadas em escala Likert e questões abertas destinadas à obtenção de comentários, críticas e sugestões. Os critérios analisados incluíram a clareza do conteúdo, a organização do material, a aplicabilidade pedagógica, a aderência aos princípios do PBL e a utilidade do *guideline* no contexto da integração com a GenAI.

Todas as etapas da pesquisa que envolveram a participação de especialistas foram

previamente submetidas à apreciação ética, tendo sido aprovadas pelo Comitê de Ética em Pesquisa da instituição responsável. O estudo encontra-se devidamente registrado sob o Certificado de Apresentação de Apreciação Ética (CAAE) nº 92210425.3.0000.5208, em conformidade com as diretrizes éticas vigentes para pesquisas envolvendo seres humanos. A participação dos especialistas ocorreu de forma voluntária, mediante aceite do Termo de Consentimento Livre e Esclarecido, assegurando-se o anonimato, a confidencialidade das respostas e o uso exclusivo dos dados para fins acadêmicos.

A avaliação do *guideline* contou com a participação de 3 especialistas com experiência em PBL e metodologias ativas no ensino superior. As questões fechadas foram analisadas por meio de estatística descritiva, permitindo identificar tendências gerais quanto à clareza, organização, aplicabilidade pedagógica e aderência aos princípios do PBL. As respostas abertas foram examinadas a partir de uma análise de conteúdo em modalidade temática (BARDIN, 1977), com o objetivo de identificar recorrências, sugestões de aprimoramento e pontos de convergência entre as percepções dos avaliadores. Essa análise possibilitou uma compreensão qualitativa das potencialidades e limitações do material, servindo de base para a interpretação dos resultados apresentados no capítulo de resultados.

3.4.2 Avaliação experimental do chatbot tutor

3.4.2.1 Desenho experimental e tratamentos

A opção por um experimento controlado com julgamento automatizado, em vez de avaliação por docentes em uso prolongado, decorreu da necessidade de isolar o efeito do *guideline* sob condições experimentais replicáveis. Avaliações baseadas em participantes humanos introduziriam variabilidade interavaliador e demandariam um desenho longitudinal distinto do escopo deste estudo. Assim, o uso de *LLM as a Judge* foi adotado como estratégia de comparação controlada entre configurações do agente, preservando consistência nos critérios aplicados, sem substituir validações futuras em contexto real. Trata-se, portanto, de uma etapa inicial de avaliação controlada, que pode anteceder estudos posteriores com validação empírica junto a docentes em ambiente autêntico de uso.

A avaliação do chatbot tutor foi estruturada como um experimento controlado, considerando que avaliações com docentes baseadas em uso contínuo e prolongado do sistema demandariam condições de execução e acompanhamento que não se mostraram compatíveis com o recorte

temporal e operacional do estudo. Optou-se, portanto, pela realização de um experimento controlado, conduzido pelo próprio pesquisador, no qual duas configurações do mesmo agente baseado em modelo de linguagem foram comparadas. No tratamento T1, o agente operou sem acesso ao *guideline*; no tratamento T2, o agente operou com acesso ao *guideline*, incorporado ao *system prompt* em tempo de execução. O experimento caracteriza-se como um teste A/B pareado, no qual a unidade experimental foi cada situação-problema apresentada ao agente. Não houve público humano participante nessa etapa; o conjunto de prompts representa decisões pedagógicas simuladas derivadas do diagnóstico e do planejamento do *guideline*. O protocolo consistiu na execução independente das duas configurações (T1 e T2) para cada prompt, mantendo constantes modelo e parâmetros de geração, com randomização da ordem de apresentação e cegamento das respostas antes do julgamento comparativo. Em ambos os tratamentos, manteve-se constante o modelo subjacente do agente, baseado no modelo *gemini-3-flash-preview*, bem como os parâmetros de geração e a execução em sessões independentes, sem histórico conversacional, de modo a isolar o efeito do *guideline* como variável experimental.

O experimento foi composto por 30 perguntas representativas de dúvidas realistas de professores universitários que pretendem planejar e conduzir atividades de PBL integradas à GenAI. As perguntas foram inicialmente geradas com apoio de um modelo de linguagem, a partir de instruções elaboradas pelo pesquisador que descreviam o contexto do problema investigado, os eixos pedagógicos definidos no planejamento da ação e situações plausíveis enfrentadas por docentes ao implementar PBL com GenAI. Após a geração inicial, o conjunto foi revisado e selecionado pelo pesquisador, com ajustes de redação e eliminação de redundâncias, buscando assegurar coerência com os objetivos do experimento e cobertura das diferentes dimensões pedagógicas consideradas. O conjunto final foi definido previamente à execução do experimento e mantido inalterado ao longo das avaliações. Essas perguntas foram definidas previamente e mantidas iguais durante toda a execução do experimento, com o objetivo de permitir uma comparação justa entre as duas configurações do agente e facilitar a repetição do estudo por outros pesquisadores. Para evitar que a ordem das perguntas influenciasse os resultados, a sequência de apresentação foi embaralhada por meio de um procedimento de randomização computacional, implementado em script pelo próprio pesquisador, com semente fixa para garantir reprodutibilidade. A ordem resultante foi registrada e preservada como parte dos artefatos do experimento. Além disso, para que a avaliação fosse imparcial, as duas respostas geradas para cada pergunta foram identificadas apenas como Resposta A e Resposta B, sem

indicar qual delas pertencia a cada configuração; essa informação foi guardada separadamente e utilizada apenas depois, no momento de organizar os resultados e calcular as métricas de comparação.

As 30 situações-problema utilizadas como prompts foram elaboradas pelo pesquisador a partir da triangulação entre os desafios identificados na revisão da literatura, as priorizações realizadas no seminário com docentes e as decisões pedagógicas explicitadas no planejamento do *guideline*, buscando representar dilemas plausíveis enfrentados por professores universitários no uso do PBL com GenAI. O conjunto de perguntas não foi submetido a validação estatística externa, mas passou por revisão interna quanto à clareza e coerência pedagógica, com o objetivo de garantir representatividade conceitual do domínio investigado. No experimento, não se buscou determinar uma resposta “correta” em sentido absoluto, mas comparar a qualidade pedagógica relativa entre as duas configurações do agente com base em uma rubrica previamente definida. A atribuição dos escores foi realizada pelo modelo configurado como *LLM-as-a-Judge*, enquanto o pesquisador atuou apenas na verificação de consistência dos julgamentos e na auditoria de casos instáveis conforme o protocolo experimental estabelecido.

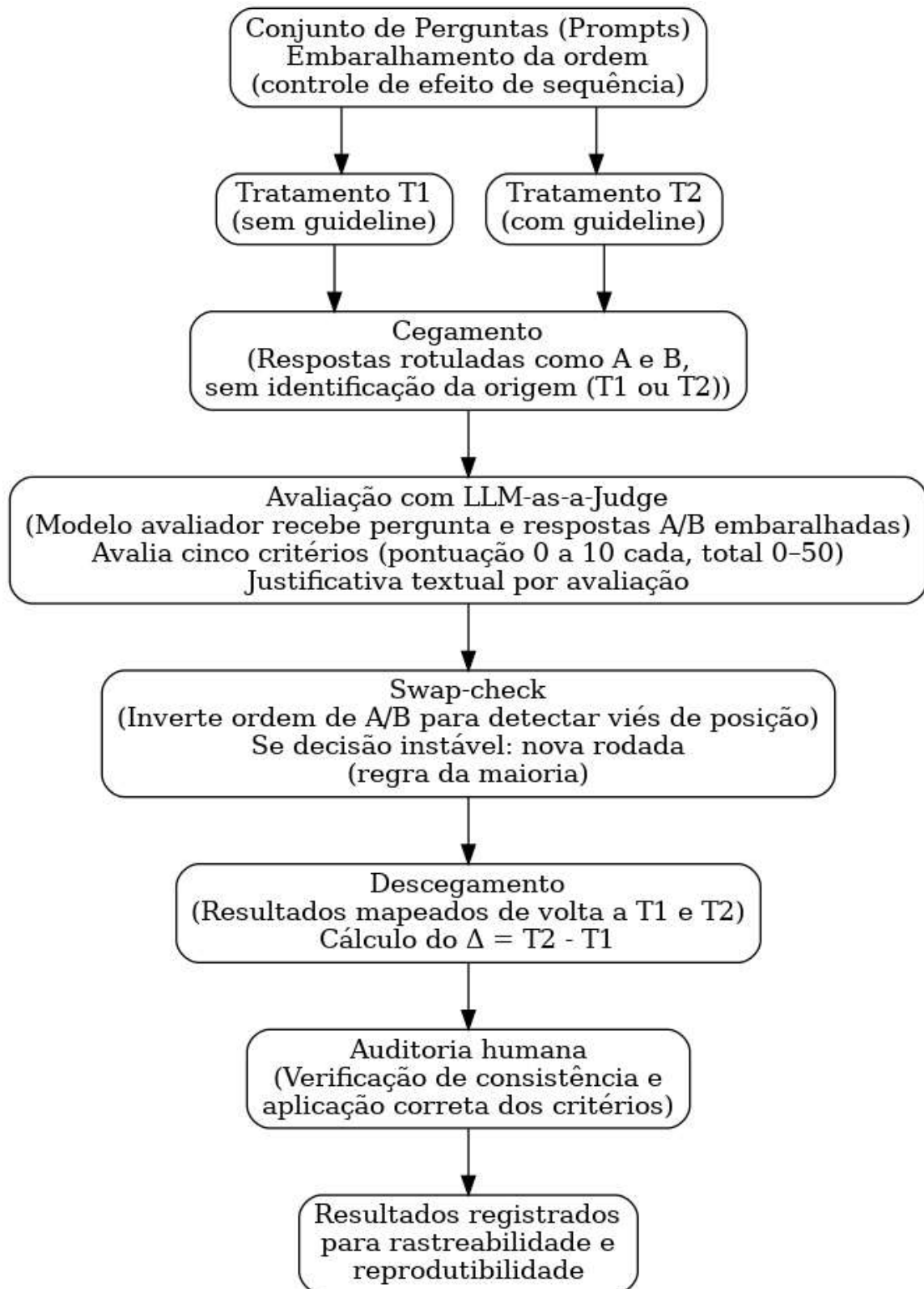


Figura 2 – Visão geral do desenho experimental adotado na avaliação do chatbot tutor, incluindo embaralhamento das perguntas, cegamento das respostas, avaliação por LLM-as-a-Judge, desceçamento, controle de consistência e auditoria humana.

3.4.2.2 *LLM-as-a-Judge e justificativa metodológica*

A comparação entre as respostas de T1 e T2 foi realizada por meio da abordagem *LLM-as-a-Judge* (LaaJ), na qual um modelo de linguagem distinto do agente atua como avaliador das respostas geradas com base em critérios de avaliação previamente definidos, aplicando notas para cada critério e produzindo um resultado comparativo. Essa decisão metodológica se fundamenta em evidências de que LLMs podem oferecer julgamentos úteis e reprodutíveis para tarefas de avaliação textual, sendo particularmente adequadas para comparações relativas entre sistemas quando o mesmo avaliador é utilizado de forma consistente, ainda que as notas absolutas possam refletir tendências próprias do avaliador. Essa perspectiva é discutida por (CHIANG; LEE, 2023), que analisam a viabilidade de LLMs como alternativa de avaliação em cenários onde o uso de avaliadores humanos é custoso ou limitado.

Ao mesmo tempo, reconhece-se que avaliadores baseados em LLM podem apresentar limitações, como variabilidade induzida por instruções e vieses sistemáticos, incluindo viés de posição em comparações pareadas. Trabalhos no contexto educacional discutem preocupações relacionadas à consistência e confiabilidade de feedback gerado e avaliado por LLMs (SEO et al., 2025). Adicionalmente, estudos recentes em Engenharia de Software reportam que avaliações comparativas podem se tornar inconsistentes quando a ordem das respostas é invertida, indicando a necessidade de mecanismos explícitos de controle quando se utiliza julgamento pareado, no qual duas respostas geradas para o mesmo prompt são avaliadas conjuntamente pelo mesmo julgador, sob os mesmos critérios, permitindo uma comparação direta entre as configurações avaliadas (WANG et al., 2025). Diante dessas limitações, o presente estudo adotou estratégias de mitigação descritas a seguir, interpretando os resultados como evidência exploratória e controlada, e não como substituição integral de validação humana em contexto real.

3.4.2.3 *Critérios de avaliação e métricas analisadas*

O avaliador recebeu, para cada prompt, o enunciado original e duas respostas (A e B), sendo instruído a retornar uma avaliação estruturada no formato JSON. O julgamento foi orientado por uma rubrica composta por cinco critérios pedagógicos definidos previamente: (1) aderência ao pedido, (2) qualidade pedagógica em PBL, (3) integração responsável da IA, (4) clareza e aplicabilidade e (5) completude com concisão. Esses critérios foram definidos com

base nos princípios do PBL e nas diretrizes de uso pedagógico da GenAI, conforme detalhado na seção de resultados.

A rubrica adotada atribuiu notas em escala de 0 a 10 para cinco critérios, além de uma justificativa textual, de modo a capturar dimensões relevantes para o objetivo do chatbot. Os critérios foram definidos para refletir o contexto docente e os eixos do planejamento da ação, contemplando a aderência ao contexto pedagógico do PBL, o alinhamento às etapas do PBL quando aplicável, a qualidade das orientações sobre o uso pedagógico da GenAI, a clareza e adequação comunicacional para docentes, e a completude e acionabilidade das recomendações sem prolixidade excessiva.

O escore total de cada resposta foi calculado como a soma dos cinco critérios, resultando em uma medida de 0 a 50. Após o julgamento, realizou-se o descegaamento, entendido neste estudo como a revelação controlada da correspondência entre as respostas rotuladas como A e B e as condições experimentais T1 e T2, calculando-se, para cada prompt, a diferença Δ , definida como $\Delta = score_{T2} - score_{T1}$. A métrica primária considerada para análise foi o Δ total, complementada por análises descritivas dos Δ por critério com finalidade diagnóstica. Para reduzir falhas de formatação, o julgador foi executado com validação do JSON e limite de tentativas, e casos em que não houve retorno válido foram tratados como perdas do julgamento automático, sendo excluídos da análise principal conforme regra definida no plano experimental.

3.4.2.4 Mitigação de vieses e auditoria humana

Considerando a possibilidade de *position bias* em julgamentos comparativos e a instabilidade sob inversão de ordem discutida na literatura (SEO et al., 2025; WANG et al., 2025), foi adotado um reforço anti-viés por inversão (*swap-check*), isto é, uma verificação adicional na qual a ordem de apresentação das respostas é trocada para identificar se o *LaaJ* tende a favorecer sistematicamente a primeira ou a segunda resposta. Para cada par de respostas, o *LaaJ* foi acionado duas vezes, alternando-se a ordem de apresentação entre A/B e B/A. Em seguida, as decisões obtidas nas duas ordens foram comparadas para verificar se a escolha do vencedor permanecia a mesma após a troca de posição das respostas. Quando as duas execuções divergiram quanto ao vencedor, o caso foi classificado como de baixa confiança e tratado por uma regra de decisão definida previamente para evitar ajustes influenciados pelos resultados, realizando-se uma rodada adicional de julgamento sob os mesmos parâmetros e determinando-se o vencedor final por maioria. Persistindo instabilidade após o número máximo de tentativas, o

caso foi registrado como empate técnico e encaminhado para auditoria humana.

A auditoria humana foi conduzida pelo pesquisador com o objetivo de verificar consistência interna entre justificativa e pontuações atribuídos pelo avaliador, bem como identificar falhas sistemáticas na aplicação dos critérios de avaliação em casos instáveis. Essa estratégia está em consonância com a recomendação de interpretar avaliações por LLM como evidência complementar e de empregar verificações adicionais quando o julgamento automático é utilizado em estudos com limitações de validação humana direta (CHIANG; LEE, 2023). O *judge prompt* foi mantido fixo e versionado para reduzir variações e assegurar rastreabilidade do processo de avaliação.

3.4.2.5 Rastreabilidade e reprodutibilidade

Todos os artefatos da execução do experimento foram registrados em diretórios de *run*, incluindo o conjunto de prompts, as sementes de randomização, o plano de cegamento A/B, as respostas do agente em T1 e T2, as saídas do julgador e os resultados agregados. Esse procedimento foi adotado para permitir replicação do experimento, auditoria do processamento e transparência na interpretação dos resultados. Os resultados obtidos nessa etapa são apresentados e discutidos nas seções subsequentes do trabalho.

3.5 REFLEXÃO

Nesta etapa, a reflexão é apresentada em nível metodológico, com o objetivo de explicitar como os resultados obtidos na intervenção foram interpretados à luz do ciclo de DSR, bem como registrar condicionantes e limitações que influenciaram o processo investigativo. A análise crítica aprofundada das implicações pedagógicas e teóricas dos achados é desenvolvida de forma mais abrangente no capítulo de Discussão.

Na metodologia, a reflexão constitui um momento central de revisão crítica do ciclo conduzido, no qual se examina a coerência entre os objetivos do estudo, as decisões tomadas ao longo do planejamento e os procedimentos adotados na intervenção e na avaliação. Assim, a reflexão metodológica não se limita à descrição de resultados, mas busca compreender em que medida as estratégias escolhidas produziram evidências compatíveis com o propósito da investigação e quais aspectos do desenho do estudo condicionam a leitura dos achados.

4 PROJETO E CONSTRUÇÃO DOS ARTEFATOS

4.1 VISÃO GERAL DOS ARTEFATOS

No contexto desta pesquisa, foram desenvolvidos dois artefatos com o objetivo de apoiar professores universitários na integração de ferramentas de GenAI em práticas pedagógicas baseadas em PBL: (i) um guideline estruturado e (ii) um chatbot tutor de apoio pedagógico, utilizado em caráter experimental.

O guideline foi concebido como o principal artefato da pesquisa, estruturado a partir das etapas do PBL, com o propósito de orientar o docente na identificação de oportunidades de uso da GenAI ao longo do processo de ensino-aprendizagem. Esse artefato busca traduzir conceitos abstratos em orientações práticas, apresentando exemplos de aplicação, sugestões de prompts e reflexões pedagógicas que auxiliam na tomada de decisão em sala de aula.

Complementarmente, o chatbot tutor foi desenvolvido como um agente conversacional com o objetivo de explorar, em um contexto controlado, formas de oferecer suporte dinâmico ao professor. Diferentemente do guideline, que possui caráter mais estático e estruturado, o chatbot permite interação em linguagem natural, possibilitando ao docente esclarecer dúvidas, explorar possibilidades de uso da GenAI e obter sugestões contextualizadas conforme suas necessidades específicas.

É importante destacar que o chatbot foi utilizado exclusivamente no contexto de um experimento conduzido nesta pesquisa, com a finalidade de avaliar seu potencial como mecanismo de apoio pedagógico. Dessa forma, seu papel não se configura como uma solução consolidada, mas como um artefato exploratório voltado à investigação de abordagens interativas de suporte ao docente.

Embora distintos em sua natureza e nível de maturidade, os dois artefatos foram concebidos de forma complementar. Enquanto o guideline fornece uma visão estruturada e sistemática das possibilidades de aplicação da GenAI no PBL, o chatbot atua como um mecanismo de apoio interativo, permitindo ao professor adaptar e aprofundar essas orientações em situações específicas. Em conjunto, os artefatos contribuem para a compreensão de diferentes formas de apoio à prática docente no uso de GenAI.

Dessa forma, os artefatos desenvolvidos nesta pesquisa não se configuram apenas como ferramentas isoladas, mas como diferentes instâncias de apoio à prática docente, voltadas à facilitação do uso crítico e contextualizado da GenAI no ensino superior.

4.2 CONSTRUÇÃO DO GUIDELINE

O guideline foi desenvolvido como o principal artefato desta pesquisa, com o objetivo de oferecer suporte estruturado a professores na integração de ferramentas de GenAI em práticas pedagógicas baseadas em PBL. Sua construção foi orientada pela necessidade de traduzir conceitos teóricos em orientações práticas, acessíveis e aplicáveis ao contexto do ensino superior.

A estrutura do guideline foi definida a partir das sete etapas do PBL (WALSH, 2005), de modo a alinhar o uso da GenAI ao fluxo pedagógico já consolidado dessa abordagem. Essa organização permite que o docente identifique, em cada etapa do processo, oportunidades específicas de utilização da GenAI como apoio às atividades de ensino e aprendizagem.

Cada etapa do guideline foi estruturada de forma padronizada, contendo quatro elementos principais: (i) o objetivo pedagógico da etapa, que descreve sua finalidade no contexto do PBL; (ii) um relato hipotético de aplicação, que exemplifica de forma prática como a GenAI pode ser utilizada; (iii) sugestões de prompts, que servem como ponto de partida para a interação com ferramentas de GenAI; e (iv) uma reflexão pedagógica, que estimula o docente a analisar criticamente o uso dessas tecnologias em sua prática.

Durante o processo de construção, optou-se por adotar uma abordagem baseada em categorias funcionais de uso da GenAI, em vez de ferramentas específicas. Essa decisão teve como objetivo garantir maior flexibilidade e longevidade ao artefato, evitando sua obsolescência diante da rápida evolução das tecnologias. Dessa forma, o guideline foca nas funções que a GenAI pode desempenhar no processo pedagógico, independentemente da ferramenta utilizada.

Adicionalmente, foi incorporada ao guideline uma escala de aderência pedagógica, utilizada para avaliar o nível de adequação entre as categorias funcionais da GenAI e as diferentes etapas do PBL. A definição dessa escala foi realizada a partir da interpretação e julgamento do autor, com base no referencial teórico adotado, sendo posteriormente revisada por pares no contexto desta pesquisa.

Essa escala contribui para orientar o docente na seleção de usos mais apropriados, considerando o contexto de cada etapa do processo de aprendizagem.

A construção do guideline ocorreu de forma iterativa, com refinamentos sucessivos a partir da análise teórica e da organização das possibilidades de uso da GenAI no contexto do PBL. Esse processo permitiu consolidar um artefato que equilibra estrutura e flexibilidade, oferecendo direcionamento ao docente sem impor um modelo prescritivo de aplicação.

Dessa forma, o guideline resultante se configura como um instrumento de apoio à

tomada de decisão pedagógica, auxiliando professores a incorporar a GenAI de maneira crítica, contextualizada e alinhada aos objetivos de aprendizagem.

4.3 CONSTRUÇÃO DO CHATBOT

O chatbot tutor de apoio pedagógico foi desenvolvido como um artefato experimental, com o objetivo de investigar o potencial de agentes conversacionais no suporte à integração de GenAI em práticas pedagógicas baseadas em PBL.

Diferentemente do guideline, que possui uma estrutura estática e organizada, o chatbot foi concebido como um sistema interativo, capaz de responder a solicitações em linguagem natural e oferecer suporte dinâmico ao docente. Seu papel consiste em auxiliar o professor na exploração de possibilidades de uso da GenAI, esclarecimento de dúvidas e geração de sugestões contextualizadas ao seu cenário de ensino.

A construção do chatbot foi orientada pelos mesmos princípios adotados no desenvolvimento do guideline, especialmente no que se refere ao alinhamento com as etapas do PBL e ao uso de categorias funcionais da GenAI. Dessa forma, o chatbot foi projetado para fornecer respostas que consideram o contexto pedagógico apresentado pelo usuário, mantendo coerência com a estrutura proposta no guideline.

Em termos de funcionamento, o chatbot utiliza um modelo de linguagem para interpretar as entradas do usuário e gerar respostas com base em instruções previamente definidas. Essas instruções foram elaboradas com o objetivo de direcionar o comportamento do sistema para um papel de tutor pedagógico, evitando respostas genéricas e incentivando reflexões sobre o uso da GenAI no ensino.

É importante destacar que o chatbot não foi desenvolvido como uma solução final ou produto consolidado, mas sim como um artefato exploratório utilizado em um experimento conduzido no contexto desta pesquisa. Especificamente, o chatbot foi utilizado para investigar se um agente enriquecido com o conhecimento estruturado do guideline seria capaz de produzir respostas mais alinhadas aos princípios do PBL, em comparação a abordagens não estruturadas.

Dessa forma, o experimento teve caráter exploratório, com o objetivo de analisar o impacto do guideline no comportamento do agente conversacional, e não de validar o chatbot como uma solução definitiva de apoio pedagógico.

4.4 SÍNTESE DOS ARTEFATOS

Os artefatos desenvolvidos nesta pesquisa foram concebidos com o objetivo comum de apoiar professores universitários na integração de GenAI em práticas pedagógicas baseadas em PBL, explorando diferentes formas de suporte à tomada de decisão docente.

O guideline se configura como o principal artefato, oferecendo uma estrutura organizada e sistemática que orienta o professor ao longo das etapas do PBL. Por meio de seus elementos (objetivos pedagógicos, exemplos de aplicação, sugestões de prompts e reflexões), o guideline fornece direcionamento prático sem impor um modelo rígido de uso, permitindo sua adaptação a diferentes contextos de ensino.

Por sua vez, o chatbot foi desenvolvido como um artefato experimental, com foco na exploração de abordagens interativas de apoio ao docente. Sua principal contribuição reside na possibilidade de oferecer suporte dinâmico e contextualizado, complementando a natureza mais estruturada do guideline por meio da interação em linguagem natural.

Dessa forma, os artefatos contribuem para o avanço da discussão sobre o uso de GenAI no ensino superior, oferecendo subsídios tanto práticos quanto exploratórios para a adoção dessas tecnologias em contextos pedagógicos baseados em PBL.

5 RESULTADOS

Este capítulo apresenta os resultados obtidos ao longo da pesquisa, organizados conforme as etapas do ciclo da DSR e alinhados aos objetivos específicos. Os resultados são apresentados de forma descritiva, sem interpretação aprofundada, que é desenvolvida no capítulo de Discussão.

5.1 RESULTADOS DO DIAGNÓSTICO

Os resultados do diagnóstico do problema foram obtidos a partir de duas fontes complementares de evidência: (i) uma RLR, conduzida com o objetivo de mapear desafios recorrentes associados à adoção da GenAI no ensino superior, e (ii) um seminário exploratório com docentes universitários, voltado à validação e priorização desses desafios a partir da prática pedagógica. Essa etapa desempenhou papel fundamental na compreensão do problema investigado e na orientação do planejamento da intervenção proposta.

A análise da literatura permitiu identificar sete categorias principais de desafios associados à adoção da Inteligência Artificial Generativa no ensino superior: (1) dependência excessiva da IA; (2) preocupações éticas e de privacidade; (3) falta de transparência e explicabilidade; (4) desigualdade no acesso a recursos tecnológicos; (5) ausência de formação docente; (6) barreiras culturais e linguísticas; e (7) impacto ambiental.

A frequência de ocorrência dessas categorias nos estudos analisados foi utilizada como recurso auxiliar para indicar a recorrência dos temas na literatura, permitindo estabelecer uma hierarquia inicial dos desafios mais frequentemente relatados. A Tabela 3 apresenta a sistematização dos desafios identificados na RLR e sua presença nos estudos analisados.

De forma complementar, os desafios sistematizados a partir da RLR foram aprofundados por meio de um seminário exploratório com docentes universitários, cujo objetivo foi priorizar os desafios identificados na literatura, com base na experiência prática dos participantes. A Figura 3 apresenta o ranking coletivamente construído durante a atividade.

O desafio mais prioritário para os docentes foi a dependência excessiva da IA por parte dos estudantes. Durante o seminário, foram relatadas dificuldades na avaliação das atividades acadêmicas, especialmente quanto à identificação da autoria e à presença de respostas que destoam do repertório conceitual usualmente demonstrado pelos alunos. Também foram mencionadas adaptações nos métodos avaliativos, como a adoção de avaliações orais e

Autor	Dependência excessiva da IA e redução da autonomia	Preocupações éticas e de privacidade	Falta de transparência e explicabilidade	Desigualdade no acesso a recursos tecnológicos	Falta de capacitação e habilidades específicas	Cultura e diversidade linguística	Impacto ambiental
(SHAHZAD et al., 2025)	Sim	Sim	Sim	Sim	Sim	Sim	Sim
(IRISH; GAZICA; BECERRA, 2025)	Sim	Sim	Sim	Não	Sim	Não	Não
(FENTA, 2025)	Sim	Sim	Sim	Sim	Sim	Não	Não
(GUIZANI et al., 2025)	Sim	Sim	Sim	Sim	Sim	Não	Sim
(SAAIDA, 2023)	Sim	Sim	Sim	Sim	Sim	Não	Não
(ISMAIL et al., 2023)	Não	Sim	Não	Não	Sim	Não	Não
(DAKAKNI; SAFA, 2023)	Sim	Sim	Sim	Sim	Sim	Sim	Não
(GRÁJEDA et al., 2024)	Sim	Sim	Sim	Sim	Sim	Não	Não
(PARI-BEDOYA; VARGAS-MURILLO; GUEVARA-SOTO, 2023)	Sim	Sim	Não	Não	Sim	Não	Não
(MICHEL-VILLARREAL et al., 2023)	Sim	Sim	Sim	Sim	Sim	Não	Não
(WALCZAK; CELLARY, 2023)	Sim	Sim	Sim	Sim	Sim	Sim	Sim
(AL-MUGHAIIRI; BHASKAR, 2024)	Sim	Sim	Sim	Sim	Sim	Sim	Não
(ZEB; ULLAH; KARIM, 2024)	Sim	Sim	Sim	Sim	Sim	Não	Não
(BAIG; YADEGARIDEHKORDI, 2025)	Sim	Sim	Sim	Sim	Sim	Não	Não
(YERALAN; LEE, 2023)	Sim	Sim	Sim	Não	Sim	Não	Não
(KHAN et al., 2025)	Sim	Sim	Sim	Sim	Sim	Sim	Sim
(TUBELLA; MORA-CANTALLOPS; NIEVES, 2024)	Sim	Sim	Sim	Não	Sim	Sim	Sim
(CORDERO; TORRES-ZAMBRANO; CORDERO-CASTILLO, 2024)	Sim	Sim	Sim	Sim	Sim	Sim	Não
(AN; YU; JAMES, 2025)	Sim	Sim	Sim	Não	Sim	Não	Não
(SOK; HENG, 2024)	Sim	Sim	Sim	Não	Sim	Não	Não
(ALOTAIBI; ALSHEHRI, 2023)	Sim	Sim	Não	Sim	Sim	Sim	Não
(RAHIMAN; KODIKAL, 2024)	Sim	Sim	Sim	Sim	Sim	Sim	Não
(KURTZ et al., 2024)	Sim	Sim	Sim	Sim	Sim	Não	Não
(CHAN; HU, 2023)	Sim	Sim	Sim	Sim	Sim	Não	Não
(RASUL et al., 2023)	Sim	Sim	Sim	Sim	Sim	Não	Não
(CUKUROVA; MIAO et al., 2024)	Sim	Sim	Sim	Sim	Sim	Sim	Sim
(DHAMIJA; DHAMIJA, 2025)	Sim	Sim	Não	Não	Sim	Não	Não
Total de estudos (Sim)	26	27	23	19	27	10	6

Tabela 3 – Estudos analisados na RLR quanto à abordagem de desafios associados à integração da inteligência artificial no ensino superior.



Figura 3 – Ranking de priorização dos desafios da inteligência artificial segundo os docentes participantes.

manuscritas.

Outro aspecto recorrente nos relatos dos docentes refere-se ao uso da inteligência artificial como principal fonte de informação pelos estudantes. Nesse contexto, foram destacadas preocupações relacionadas ao uso ético da IA, bem como a necessidade de diálogo transparente com os estudantes sobre os limites e as responsabilidades associados ao uso da IA no contexto acadêmico.

De forma integrada, o diagnóstico evidenciou que os desafios associados à adoção da Inteligência Artificial Generativa no ensino superior extrapolam os aspectos técnicos, envolvendo dimensões pedagógicas, éticas, institucionais e formativas. A recorrência desses desafios na literatura e sua priorização por docentes universitários indicaram a necessidade de uma intervenção orientada não apenas ao uso de ferramentas, mas também ao apoio à prática pedagógica e à tomada de decisão docente.

5.2 RESULTADOS DO PLANEJAMENTO DA AÇÃO

A fundamentação teórica dessa etapa foi organizada em quatro eixos principais: (1) metodologias ativas de ensino, com ênfase no PBL; (2) competências digitais docentes; (3) avaliação formativa em contextos mediados por tecnologia; e (4) uso pedagógico da inteligência artificial no ensino superior. Esses eixos não foram definidos previamente como categorias teóricas abstratas, mas decorrem da organização conceitual das principais decisões pedagógicas exigidas pelo contexto investigado.

Dessa forma, os eixos teóricos adotados sintetizam, de forma objetiva, as dimensões

Dimensão	Eixo do Guideline
1	Metodologias ativas de ensino, com ênfase no PBL
2	Competências digitais docentes
3	Avaliação formativa em contextos mediados por tecnologia
4	Uso pedagógico da inteligência artificial no ensino superior

Tabela 4 – Dimensões pedagógicas que estruturam o guideline de integração do PBL com GenAI

pedagógicas consideradas no planejamento da intervenção. A Tabela 2 explicita a relação entre cada eixo teórico e a questão prática que motivou sua adoção, tornando visível o encadeamento entre os desafios identificados no diagnóstico e as escolhas realizadas no planejamento da ação.

Eixo teórico considerado no planejamento	Questão prática que motivou sua adoção
Metodologias ativas de ensino (PBL)	Como estruturar atividades de ensino em um contexto no qual os estudantes já dispõem de ferramentas de GenAI?
Competências digitais docentes	O que o professor precisa saber e desenvolver para conduzir atividades pedagógicas mediadas por GenAI de forma crítica, intencional e alinhada aos objetivos educacionais?
Avaliação formativa em contextos mediados por GenAI	Como acompanhar o processo de aprendizagem e avaliar os estudantes, para além dos produtos finais, em atividades que envolvem o uso de GenAI?
Uso pedagógico da GenAI no ensino superior	Como empregar a GenAI como ferramenta de apoio pedagógico sem distorcer os objetivos educacionais das atividades propostas?

Quadro 2 – Relação entre os eixos teóricos considerados no planejamento da ação e as questões práticas que orientaram sua definição.

5.3 RESULTADOS DA INTERVENÇÃO

A intervenção resultou no desenvolvimento de dois artefatos educacionais complementares, concebidos para apoiar professores universitários na integração do PBL com ferramentas de GenAI. Os artefatos foram desenvolvidos de acordo com os eixos pedagógicos definidos no planejamento da ação.

5.3.1 Desenvolvimento iterativo do guideline de PBL integrado à GenAI

O desenvolvimento do guideline de PBL integrado à GenAI foi motivado diretamente pelos achados do diagnóstico, que indicaram a necessidade de um artefato educacional que apoiasse professores universitários na tomada de decisão pedagógica diante da incorporação crescente de ferramentas de GenAI por estudantes. À época da concepção do artefato, observou-se o surgimento de iniciativas e documentos orientadores voltados ao uso de IA na educação, a exemplo de guias amplos voltados à prática docente (MOLLICK, 2024; BOWEN; WATSON, 2025; PIMENTEL; CARVALHO, 2025). Contudo, no momento da concepção desta pesquisa, identificou-se uma lacuna, uma vez que não foram encontrados materiais orientadores que abordassem, de forma integrada e pedagogicamente fundamentada, o uso da GenAI no PBL, considerando as etapas do método, diferentes níveis de integração conforme o escopo da aplicação e a necessidade de um uso ético, reflexivo e consciente das ferramentas.

Diante dessa lacuna, a elaboração do guideline teve início com a retomada dos fundamentos do PBL, cuja origem está associada a um grupo de pesquisadores da McMaster University. Nesse contexto, (WALSH, 2005) foi responsável por sistematizar e popularizar a aplicação prática do método, organizando-o em um formato estruturado composto por sete fases. Esse formato foi adotado como referência neste estudo por oferecer um encadeamento operacional claro das etapas do PBL, adequado ao propósito de orientar professores universitários na integração pedagógica da GenAI, permitindo explicitar, fase a fase, as decisões pedagógicas envolvidas.

Etapas do PBL	Função pedagógica
Identificação do problema	Apresentar uma situação-problema aberta e contextualizada, estimulando a curiosidade e o engajamento inicial dos estudantes.
Exploração do conhecimento prévio	Ativar conhecimentos anteriores e experiências dos estudantes relacionadas ao problema proposto.
Geração de hipóteses	Estimular a formulação inicial de explicações, suposições e caminhos possíveis para resolução do problema.
Definição das questões de estudo	Delimitar questões investigativas que orientarão o estudo individual e coletivo.
Estudo individual	Promover a busca autônoma por informações e novos conhecimentos relevantes às questões formuladas.
Aplicação do conhecimento	Integrar e aplicar os conhecimentos adquiridos na análise e resolução do problema.
Avaliação e reflexão	Favorecer a metacognição, a avaliação do processo e a consolidação da aprendizagem.

Quadro 3 – Etapas do ciclo do Problem-Based Learning adotadas no guideline

A partir da definição do ciclo de referência do PBL, o desenvolvimento avançou para a questão central da intervenção, que consistiu em compreender como a GenAI pode dar suporte de modo intencional em cada etapa do PBL, preservando o protagonismo discente e o papel mediador do professor. As primeiras versões do artefato exploraram a possibilidade de associar diretamente etapas do PBL a ferramentas específicas de IA potencialmente úteis ao docente. Essa estratégia inicial, contudo, revelou uma limitação relevante, pois a rápida evolução do ecossistema de ferramentas, a heterogeneidade de serviços e a variabilidade de disponibilidade institucional tenderiam a tornar o guideline excessivamente dependente de produtos e exemplos pontuais, com risco de desvio de propósito e de perda de generalidade.

Como resposta a essa limitação, o desenvolvimento do guideline foi conduzido a um nível de abstração mais alto, priorizando categorias funcionais de ferramentas em vez de ferramentas individuais. Esse ajuste foi um ponto de inflexão no processo de construção do artefato, pois permitiu orientar a integração da GenAI com base em propósitos pedagógicos mais estáveis e transferíveis entre diferentes soluções tecnológicas. Nesse contexto, foi adotada uma taxonomia de ferramentas de IA aplicadas à educação (ALAFNAN; DISHARI; MOHDZUKI, 2024), que forneceu categorias com finalidades bem definidas e permitiu mapear a mesma ferramenta a mais de uma categoria, quando pertinente, sem comprometer a coerência do guideline.

Categoria funcional	Finalidade no guideline (descrição sintética)
Chatbots	Interação conversacional para explorar ideias, levantar hipóteses, formular perguntas e estimular reflexão.
Content Creation	Geração de textos estruturados (enunciados, roteiros, resumos e materiais de apoio) para apoiar planejamento e mediação.
Text Enhancement	Revisão e aprimoramento de textos (clareza, coesão, adequação ao público e refinamento do tom).
Video Creation	Geração/edição de vídeos para apoio didático (explicações, demonstrações, materiais complementares).
Image Generation	Produção de imagens e representações visuais para cenários, protótipos e elementos gráficos.
Note-taking / Meeting Transcription	Registro e transcrição de discussões e reuniões, favorecendo rastreabilidade e síntese.
Inbox Management	Apoio a triagem e organização de informações e mensagens, útil como suporte periférico ao planejamento docente.
Presentation Slides	Construção de apresentações e organização visual de resultados e produtos de aprendizagem.
Conduct Research	Apoio à busca e síntese de fontes e referências para investigação e aprofundamento teórico.

Quadro 4 – Categorias funcionais de ferramentas de GenAI utilizadas no guideline

A definição das categorias funcionais permitiu explicitar os diferentes tipos de suporte que a GenAI pode oferecer à prática docente no contexto do PBL. No entanto, a mera identificação dessas categorias não era suficiente para orientar o professor quanto a quando e em que medida cada tipo de ferramenta poderia contribuir pedagogicamente em cada etapa do método. Tornou-se, portanto, necessário estabelecer critérios que permitissem qualificar o grau de adequação pedagógica dessas categorias ao longo do ciclo do PBL, considerando os princípios da metodologia e o papel mediador do docente. Nesse sentido, a definição desses critérios foi conduzida com base na abordagem de DSR, orientando o processo de construção da solução a partir da definição estruturada de mecanismos de apoio à decisão pedagógica.

A escala apresentada no Quadro 5 foi elaborada especificamente para este estudo, inspirada na lógica de escalas do tipo Likert, mas adaptada ao contexto pedagógico do Problem-Based Learning. Diferentemente das escalas Likert tradicionais, voltadas à mensuração de atitudes ou níveis de concordância, a presente escala tem caráter ordinal e busca classificar o grau de aderência pedagógica de categorias de ferramentas em cada etapa do PBL. A definição dos níveis baseia-se em princípios centrais do PBL, especialmente no papel do tutor como facilitador da investigação, na promoção do raciocínio crítico, da formulação de hipóteses,

da colaboração e da reflexão sobre a aprendizagem, bem como na necessidade de evitar intervenções excessivamente diretivas ou centradas na oferta de respostas prontas, conforme discutido na literatura sobre PBL (WALSH, 2005). Dessa forma, a escala não tem como objetivo medir concordância subjetiva, mas oferecer um critério estruturado para análise pedagógica comparativa.

Nível	Nome	Critério objetivo
1	Prejudicial	Enfraquece a investigação, fornecendo respostas prontas ou bloqueando o debate crítico.
2	Dispersivo	Não contribui de forma significativa e pode desviar o grupo da análise investigativa.
3	Suporte opcional	Complementa a fase em alguns aspectos, mas sem impacto essencial no processo.
4	Apoio investigativo	Facilita a exploração, a formulação de hipóteses ou a organização coletiva da aprendizagem.
5	Núcleo pedagógico	Potencializa diretamente o raciocínio crítico, a colaboração e a reflexão da etapa.

Quadro 5 – Escala ordinal de aderência pedagógica utilizada para avaliar a relevância das categorias de ferramentas em cada etapa do PBL

A aplicação dessa escala ao mapeamento entre as sete etapas do PBL e as categorias funcionais de ferramentas de GenAI resultou na matriz de aderência apresentada na Tabela 5, na qual cada célula corresponde ao nível de aderência atribuído (1–5) para uma combinação específica entre etapa e categoria. Os valores atribuídos na matriz foram definidos a partir de julgamento analítico do pesquisador, fundamentado nos princípios do PBL, na revisão de literatura e nas evidências coletadas ao longo da pesquisa. No contexto de DSR, essa atribuição caracteriza-se como uma decisão de design do artefato, com caráter orientador e não prescritivo.

A matriz de aderência foi desenvolvida pelo autor, que também foi responsável pela atribuição dos níveis de aderência apresentados, avaliada por revisão de pares. A atribuição dos níveis de aderência foi realizada com base na escala ordinal explicitamente definida no Quadro 5, concebida como parte do design do guideline. Essa escala foi construída para qualificar, do ponto de vista pedagógico, o potencial de contribuição das categorias funcionais de ferramentas de GenAI em cada etapa do ciclo do PBL, distinguindo usos prejudiciais, pouco relevantes, complementares ou centrais para a investigação e a aprendizagem. A definição conceitual de cada nível apoia-se nos princípios do Problem-Based Learning e no papel do tutor como facilitador da investigação descritos por (WALSH, 2005), bem como na organização funcional das ferramentas proposta por (ALAFNAN; DISHARI; MOHDZUKI, 2024). A tabela foi submetida a

uma revisão formativa com o objetivo de verificar sua clareza, coerência interna e plausibilidade pedagógica, sem pretensão de validação empírica.

Etapa do PBL	Chatbots	Content Creation	Text Enhancement	Video Creation	Image Generation	Note-taking / Meeting Transcription	Inbox Management	Presentation Slides	Conduct Research
Etapa 1 – Identificação do problema	4	3	4	3	4	4	2	3	1
Etapa 2 – Exploração do conhecimento prévio	4	3	3	2	3	4	1	3	1
Etapa 3 – Geração de hipóteses	3	3	3	2	4	4	1	3	1
Etapa 4 – Definição das questões de estudo	4	3	4	1	2	4	1	3	1
Etapa 5 – Estudo individual	3	3	4	2	3	4	3	3	5
Etapa 6 – Aplicação do conhecimento	3	4	4	3	4	4	2	4	2
Etapa 7 – Avaliação e reflexão	4	3	4	3	3	4	1	3	1

Tabela 5 – Matriz de aderência pedagógica entre etapas do PBL e categorias funcionais de ferramentas de GenAI (escala Likert 1–5)

Ressalta-se que a função dessa matriz é inspiradora, e não prescritiva. Ela atua como um instrumento de apoio à tomada de decisão docente, permitindo que o professor se situe, reflita e escolha, de forma intencional, como integrar a GenAI em cada etapa do PBL, considerando o contexto específico de sua disciplina.

A tabela de aderência entre etapas do PBL e categorias de ferramentas passou a desempenhar papel estruturante no guideline, pois serviu como base para a redação do núcleo do documento, correspondente ao detalhamento das etapas do PBL com GenAI. Para cada uma das sete etapas, o guideline foi construído seguindo um padrão comum, com quatro componentes principais: (i) o objetivo pedagógico da etapa; (ii) orientações sobre como integrar a GenAI de forma coerente com o propósito da fase; (iii) um exemplo de *prompt* elaborado para apoiar o docente no uso de ferramentas representativas das categorias sugeridas; e (iv) uma reflexão crítica destinada a reforçar o uso consciente e ético da tecnologia, prevenindo dependência excessiva, respostas acríticas e substituição do julgamento pedagógico.

Componente em cada etapa	Função no guideline
Objetivo pedagógico da etapa	Explicitar a finalidade formativa da fase no ciclo do PBL.
Orientações de integração da GenAI	Indicar como empregar a GenAI de modo coerente com o propósito pedagógico da fase.
Exemplo de <i>prompt</i>	Oferecer um modelo prático para apoiar o docente no uso de ferramentas representativas das categorias sugeridas.
Reflexão crítica	Reforçar uso consciente e ético, prevenindo dependência, respostas acríticas e substituição do julgamento docente.

Quadro 6 – Estrutura padrão utilizada no detalhamento de cada uma das sete etapas do PBL com GenAI

A elaboração desse núcleo evidenciou a necessidade de organizar o documento de modo a apoiar a compreensão progressiva do leitor. Uma apresentação direta do detalhamento das

etapas, sem contextualização, tenderia a dificultar a apropriação do guideline, sobretudo por docentes com menor familiaridade com o PBL ou com o uso pedagógico da GenAI. Assim, o documento passou a ser estruturado em seções que preparam o leitor antes de chegar ao núcleo prático. Primeiramente, foi construída uma introdução destinada a contextualizar o propósito do guideline e o cenário educacional no qual ele se insere. Em seguida, foi incluída uma seção que explicita a estrutura do documento, orientando o leitor sobre como navegar pelo material por meio das camadas *Etapas*, *Níveis* e *Ferramentas*.

Seção do guideline	Finalidade pedagógica
Introdução	Contextualizar o cenário educacional, o surgimento da GenAI no ensino superior e o propósito do guideline.
Estrutura do guideline	Orientar o leitor quanto à organização do documento e à lógica de articulação entre suas seções.
Detalhamento das etapas do PBL com GenAI	Apoiar a aplicação prática do PBL integrado à GenAI, explicitando objetivos, possibilidades de uso, exemplos de prompts e reflexões críticas em cada etapa.
Níveis de implementação do PBL	Apresentar diferentes formas de implementação do PBL, considerando variações de escopo temporal, complexidade e maturidade docente.
Categorias de ferramentas de GenAI	Explicitar a lógica de categorização das ferramentas e sua relação com as etapas do PBL, apoiando a tomada de decisão docente.
Aspectos éticos do uso da GenAI	Reforçar princípios de uso consciente, responsável e alinhado aos objetivos educacionais.
Conclusão	Retomar os princípios orientadores do guideline e seu caráter não prescritivo e adaptável.

Quadro 7 – Estrutura geral do guideline de PBL integrado à GenAI

Camada	Pergunta-chave	Referencial	Onde é detalhada
Etapas	O que fazer?	(WALSH, 2005)	Seção de detalhamento das etapas
Níveis	Em que profundidade aplicar?	(RIBEIRO, 2022a)	Seção de níveis de implementação
Ferramentas	Com quais apoios tecnológicos?	(ALAFNAN; DISHARI; MOHDZUKI, 2024)	Seção de categorias e matriz de aderência

Quadro 8 – Arquitetura do guideline em camadas: etapas, níveis e ferramentas

Após essa preparação, o detalhamento das sete etapas do PBL com GenAI foi apresentado como seção central do guideline. Na sequência, foi incorporada uma seção dedicada a diferentes

formas de implementação do PBL, reconhecendo que a metodologia pode assumir múltiplas configurações em função de tempo, contexto institucional, maturidade da turma e objetivos da disciplina. Para apoiar essa adaptação prática, foram definidos níveis de implementação do PBL, fundamentados em literatura sobre variações de aplicação do método (RIBEIRO, 2022b), incluindo desde formatos mais curtos (por exemplo, uma aula ou semana) até formatos mais extensos (por exemplo, um semestre). Esses níveis foram concebidos para reduzir barreiras de adoção e permitir ao docente iniciar com escopos menores e evoluir conforme experiência e contexto.

Nível	Configuração do problema	Duração	Foco pedagógico
1	Problemas curtos, um por aula	Um encontro	Exercício rápido de ciclo completo
2	Um problema por semana	Duas aulas	Aprofundamento moderado
3	Problemas de média duração	2–4 semanas	Investigação com mini-ciclos sucessivos
4	Um grande problema ao longo do semestre	Semestre inteiro	Integração plena com múltiplas dimensões

Quadro 9 – Níveis de implementação do PBL utilizados no guideline

Em continuidade, foi incluída uma seção destinada a explicitar como as categorias de ferramentas foram alocadas às etapas do PBL, apresentando a tabela de aderência construída e discutindo a lógica que sustenta esse mapeamento, além de exemplos de ferramentas que poderiam ser enquadradas em cada categoria. Por fim, foi elaborada uma seção específica sobre aspectos éticos, reforçando limites, cuidados e responsabilidades no uso da GenAI no ensino superior, bem como a necessidade de transparência, preservação de objetivos educacionais e estímulo à autonomia discente. O documento é encerrado com uma seção de conclusão que retoma os princípios orientadores do guideline e seu caráter não prescritivo e adaptável. A partir desse ponto, o guideline foi submetido à avaliação pedagógica por especialistas externos, cujos resultados são apresentados na seção seguinte.

5.3.2 Chatbot tutor de apoio pedagógico

O chatbot tutor foi concebido como um artefato complementar ao guideline de PBL integrado à GenAI, com o objetivo de oferecer apoio operacional ao professor universitário na aplicação das orientações do documento em sua prática docente. Sua proposta está associada à ampliação da acessibilidade e da aplicabilidade do guideline, atuando como um mecanismo de

suporte para consultas pontuais e apoio à tomada de decisão pedagógica ao longo das etapas do PBL.

No contexto deste estudo, o foco da avaliação foi direcionado à análise do guideline enquanto artefato central, por meio de um survey com especialistas, permitindo uma apreciação estruturada quanto à sua qualidade, utilidade e aderência pedagógica. Esse direcionamento orientou a organização da pesquisa, posicionando o chatbot tutor como um artefato de suporte utilizado para explorar, em nível experimental, o impacto do guideline na mediação de respostas pedagógicas.

A partir dessa perspectiva, adotou-se uma abordagem baseada no uso de modelos de linguagem como avaliadores de respostas geradas por outros sistemas, conhecida como LLM-as-a-Judge, amplamente utilizada em cenários de avaliação comparativa por sua escalabilidade e reprodutibilidade (LI et al., 2024; SZYMANSKI et al., 2025; TAN et al., 2024). Nesse contexto, o chatbot tutor foi empregado como instrumento experimental para operacionalizar um experimento comparativo controlado.

O experimento consistiu na comparação do comportamento de um mesmo modelo de linguagem em duas condições distintas: uma sem apoio do guideline e outra com apoio explícito do guideline. Essa configuração permitiu analisar o efeito do guideline como elemento orientador, observando sua influência na qualidade das respostas geradas segundo critérios pedagógicos relacionados ao PBL e ao uso intencional da GenAI.

Dessa forma, o chatbot tutor desempenha, neste trabalho, um papel instrumental na materialização de diferentes configurações de uso do modelo de linguagem, viabilizando a análise do impacto do guideline em um ambiente controlado. Essa etapa experimental complementa a avaliação por especialistas, contribuindo para a produção de evidências sobre o potencial do guideline como mecanismo de apoio à prática docente mediada por GenAI.

5.4 RESULTADOS DA AVALIAÇÃO DOS ARTEFATOS

5.4.1 Resultados da avaliação do guideline por especialistas

A avaliação do guideline foi realizada por especialistas com experiência docente no ensino superior e familiaridade com metodologias ativas de ensino, incluindo o PBL em diferentes níveis de aprofundamento. O objetivo dessa etapa foi obter uma apreciação formativa do material, considerando sua clareza, organização, aplicabilidade pedagógica e alinhamento conceitual, a

partir de perspectivas docentes diversas. Essa diversidade de familiaridade com o PBL permitiu avaliar o guideline tanto sob a ótica de professores mais experientes na metodologia quanto de docentes em processo de apropriação do método.

Por se tratar de uma avaliação formativa conduzida com um número reduzido de especialistas, os resultados apresentados nesta seção têm caráter exploratório e visam subsidiar a análise da adequação pedagógica do guideline. A avaliação do guideline contou com a participação de três especialistas, todos com experiência docente no ensino superior e diferentes níveis de familiaridade prévia com a metodologia PBL. Os resultados das questões fechadas, respondidas em escala ordinal de cinco pontos, indicaram avaliações predominantemente positivas. A clareza do material apresentou média de 4,0, enquanto a organização da estrutura do guideline obteve média de 4,33. Os itens relacionados à aplicabilidade pedagógica e à integração da GenAI ao PBL apresentaram médias entre 4,33 e 4,67, sugerindo percepção consistente de utilidade prática. A avaliação geral do guideline e a recomendação a outros professores alcançaram média de 4,67 e mediana igual a 5,0, indicando elevado nível de aceitação entre os especialistas participantes.

Item avaliado	Média	Mediana	Min–Max
Clareza do guideline	4,00	4,0	3–5
Organização da estrutura	4,33	4,0	4–5
Contribuições práticas para uso de IA	4,33	4,0	4–5
Aplicabilidade no ensino superior	4,67	5,0	4–5
Facilitação da integração GenAI + PBL	4,33	4,0	4–5
Respeito aos princípios do PBL	4,33	4,0	4–5
Alinhamento dos exemplos às etapas do PBL	4,33	4,0	4–5
Adaptabilidade a outras metodologias ativas	4,33	4,0	4–5
Avaliação geral do guideline	4,67	5,0	4–5
Recomendação a outros professores	4,67	5,0	4–5

Tabela 6 – Síntese dos resultados da avaliação do guideline por especialistas (N = 3)

A Tabela 6 apresenta a síntese quantitativa dos resultados da avaliação do guideline pelos especialistas, incluindo média, mediana e intervalo das respostas para cada item avaliado.

De modo geral, as respostas às questões fechadas indicaram uma avaliação positiva do *guideline*. Os especialistas expressaram concordância quanto à organização do conteúdo, à aplicabilidade pedagógica do material no contexto do ensino superior e à sua aderência aos

princípios do PBL. Também foi registrada concordância quanto ao potencial do *guideline* como recurso de apoio à prática docente, bem como quanto à sua recomendação a outros professores interessados em integrar GenAI a metodologias ativas.

A clareza do conteúdo apresentou avaliações majoritariamente favoráveis, embora com maior variação em comparação aos demais critérios analisados. Enquanto alguns avaliadores indicaram concordância plena quanto à facilidade de compreensão do material, outros sinalizaram concordância parcial ou posição neutra, sugerindo que determinados trechos poderiam ser tornados mais explícitos ou simplificados. Esse resultado aponta para oportunidades de aprimoramento relacionadas à legibilidade e à adequação do nível de detalhamento para docentes com diferentes graus de familiaridade com GenAI.

A análise das respostas abertas permitiu aprofundar a compreensão dos julgamentos expressos nas questões fechadas. Um primeiro conjunto de comentários destacou positivamente a clareza conceitual e a organização do *guideline*, mencionando a estrutura lógica do documento, a progressão entre seções e a explicitação da relação entre o PBL e o uso pedagógico da GenAI. Também foram ressaltados o caráter prático do material e a presença de exemplos e orientações que podem apoiar o planejamento e a condução de atividades em sala de aula.

Outro conjunto de comentários enfatizou a utilidade do *guideline* como instrumento de apoio ao docente, ao mesmo tempo em que apontou a necessidade de adaptações conforme diferentes realidades institucionais e áreas do conhecimento. Foram mencionados aspectos como políticas internas sobre o uso de GenAI, disponibilidade de ferramentas, condições de acesso tecnológico e especificidades curriculares como fatores que podem influenciar a adoção do material. Essas observações reforçam que o *guideline* deve ser compreendido como um artefato flexível, passível de contextualização, e não como um modelo prescritivo único.

Por fim, os especialistas apresentaram sugestões de aprimoramento relacionadas principalmente à forma e à apresentação do material. Entre elas, destacam-se a recomendação de ampliar ou explicitar referências acadêmicas em trechos introdutórios, ajustes na extensão do documento, melhorias na apresentação visual e na legibilidade de figuras, bem como a padronização de termos e títulos. Tais sugestões fornecem subsídios diretos para o refinamento do *guideline* em versões futuras.

Em síntese, a avaliação por especialistas indica que o *guideline* foi percebido como claro, organizado e pedagogicamente aplicável, com alinhamento consistente aos princípios do PBL. Ao mesmo tempo, os comentários qualitativos evidenciam pontos de melhoria que contribuem para o amadurecimento do artefato, reforçando seu caráter formativo e iterativo no contexto

da DSR.

5.4.2 Resultados do experimento comparativo com LLM-as-a-Judge

Esta seção apresenta os resultados do experimento comparativo conduzido com base na abordagem *LLM-as-a-Judge*, empregado como estratégia complementar para investigar o efeito da incorporação do guideline de PBL integrado à GenAI sobre a qualidade das orientações geradas por um agente baseado em modelo de linguagem. O experimento foi estruturado a partir de duas configurações do mesmo agente, uma sem apoio do guideline e outra orientada explicitamente por esse documento, mantendo-se constantes o modelo subjacente e os parâmetros de geração, de modo que a presença do guideline constituiu a variável de interesse. As evidências empíricas reportadas nesta seção incluem o conjunto de prompts utilizados, os critérios de julgamento, os registros de consistência dos julgamentos e os artefatos estruturados gerados a partir do processo avaliativo.

5.4.3 Configuração experimental e parâmetros controlados

A parametrização do experimento foi definida por meio de um arquivo de configuração versionado, utilizado como instrumento empírico para assegurar rastreabilidade e reprodutibilidade do procedimento comparativo. Esse arquivo explicita a hipótese testada, a regra de decisão adotada, os modelos empregados na geração e no julgamento das respostas, bem como os mecanismos de controle de variabilidade e de mitigação de vieses associados à avaliação automática.

```
{
  "experiment_name": "pbl_guideline_ab",
  "hypothesis": "T2 improves response quality over T1",
  "decision_rule": {
    "require_judge_ok": true,
    "support_h1_if": {
      "delta_mean_gt": 0,
      "wins_t2_gt_t1": true
    }
  },
  "prompts_count": 30,
  "models": {
    "agent": {
      "name": "gemini-3-flash-preview",
```

```

    "temperature": 0.1,
    "top_p": 0.9,
    "max_output_tokens": 4096,
    "max_retries": 1
  },
  "judge": {
    "name": "gemini-3-flash-preview",
    "temperature": 0.1,
    "top_p": 1.0,
    "max_output_tokens": 4096,
    "max_retries": 2,
    "force_json_mode": true,
    "json_max_attempts": 3
  }
},
"ab_test": {
  "t1": "without_guideline",
  "t2": "with_guideline"
},
"stability_controls": {
  "swap_check": true,
  "swap_max_retries": 2,
  "instability_resolution": "score",
  "tie_tolerance": 0.2
},
"audit": {
  "fraction": 0.2
},
"seeds": {
  "prompt_order": 12345,
  "ab_assignment": 23456,
  "audit_sample": 34567
}
}

```

Código Fonte 1 – Arquivo de configuração utilizado no experimento comparativo com LLM-as-a-Judge

O parâmetro `experiment_name` identifica o experimento conduzido, enquanto o campo `hypothesis` explicita a hipótese comparativa central, segundo a qual a configuração orientada pelo guideline (T2) apresenta melhor qualidade de respostas do que a configuração sem guideline (T1). A estrutura `decision_rule` define as condições sob as quais a hipótese é considerada suportada, restringindo a análise a julgamentos válidos (`judge_ok=true`) e estabelecendo como critérios um valor médio positivo de Δ e um número de vitórias de T2 superior ao de T1.

O bloco `models` especifica os modelos de linguagem empregados tanto na geração das respostas quanto no julgamento comparativo, bem como seus parâmetros de decodificação. A

utilização de temperatura baixa visa reduzir variabilidade estocástica, enquanto os limites de saída e o número máximo de tentativas controlam o comportamento do modelo em situações de falha. No caso do LLM-as-a-Judge, o forçamento de saída estruturada em formato JSON e o número máximo de tentativas de geração foram definidos para garantir consistência na extração automática dos escores e das justificativas.

A distinção entre as configurações avaliadas é explicitada no bloco `ab_test`, que define T1 como a condição sem apoio do guideline e T2 como a condição orientada pelo guideline em modo operacional. Essa separação fundamenta o desenho comparativo pareado do experimento, no qual a presença do guideline constitui a única variável de interesse.

Os mecanismos de controle de estabilidade e mitigação de vieses estão descritos no bloco `stability_controls`. A ativação do procedimento de inversão (*swap-check*) busca reduzir o viés de posição decorrente da ordem de apresentação das respostas ao avaliador automático. Nos casos em que persistem inconsistências após as inversões previstas, a decisão é tomada com base nos escores atribuídos, considerando ainda uma tolerância global para empates, de modo a lidar com diferenças marginais no julgamento comparativo.

Por fim, o bloco `audit` define a fração de julgamentos selecionada para auditoria adicional, enquanto o conjunto de seeds fixa valores para o embaralhamento da ordem dos prompts, a atribuição A/B e a amostragem para auditoria. Essas definições contribuem para a reprodutibilidade do experimento e permitem rastrear decisões específicas tomadas ao longo do processo avaliativo.

5.4.4 Prompts e cobertura do contexto pedagógico

O conjunto de prompts utilizado no experimento foi concebido para representar situações-problema recorrentes enfrentadas por professores universitários na aplicação do PBL em contextos que envolvem o uso de GenAI. Esses prompts constituem a unidade empírica central do experimento comparativo, uma vez que cada julgamento realizado pelo avaliador automático se baseia diretamente nas respostas geradas a partir dessas situações.

As perguntas foram formuladas de modo a refletir decisões pedagógicas autênticas associadas às diferentes etapas do ciclo do PBL, incluindo a concepção e reformulação de problemas, a mediação da aprendizagem durante o trabalho em grupo, a definição de critérios e estratégias de avaliação, bem como a integração ética e pedagógica da GenAI ao processo de ensino-aprendizagem. Dessa forma, os prompts não se restringem a aspectos técnicos do

uso da IA, mas enfatizam dilemas didáticos e formativos próprios da prática docente em metodologias ativas.

Ao todo, foram utilizadas 30 situações-problema distintas, apresentadas ao agente em ambas as configurações avaliadas. Esse conjunto busca cobrir uma diversidade de contextos pedagógicos, contemplando desde decisões iniciais de planejamento do PBL até reflexões posteriores sobre avaliação, autoria discente, equidade e reformulação de atividades. A abrangência dos prompts visa reduzir o risco de avaliações enviesadas por cenários excessivamente específicos e ampliar a validade do experimento em relação a diferentes práticas e disciplinas do ensino superior.

A Tabela 7 apresenta integralmente as perguntas utilizadas como prompts no experimento. A explicitação literal dessas situações-problema permite ao leitor compreender o escopo pedagógico avaliado, bem como reforça a transparência e a reprodutibilidade do procedimento experimental, uma vez que o julgamento automatizado incide diretamente sobre as respostas geradas a partir desse conjunto de entradas.

ID	Pergunta utilizada no experimento
P01	Criei um problema para PBL, mas tenho receio de que ele esteja direcionando demais a solução. Como posso revisar esse problema para torná-lo mais aberto sem perder o foco nos objetivos da disciplina?
P02	Meu problema de PBL envolve várias tecnologias possíveis. Devo restringir as escolhas dos alunos ou deixar totalmente em aberto? Como decidir isso?
P03	Tenho dificuldade em transformar conteúdos teóricos do plano de ensino em um problema autêntico de PBL. Como fazer essa transposição?
P04	Durante o PBL, percebo que estou intervindo demais nas discussões dos grupos. Como posso ajustar minha atuação sem abandonar os alunos?
P05	Alguns alunos ficam esperando validação constante das decisões. Como responder a isso sem virar “fonte de confirmação”?
P06	Quando um grupo toma uma decisão tecnicamente fraca, devo intervir imediatamente ou deixar o erro acontecer como parte do aprendizado?

Continuação na próxima página

ID	Pergunta utilizada no experimento
P07	Estou inseguro em avaliar soluções muito diferentes entre si. Como definir critérios que valorizem o processo sem tornar a avaliação subjetiva demais?
P08	Como alinhar os critérios de avaliação do PBL com os objetivos de aprendizagem formais da disciplina?
P09	Os alunos questionam a avaliação dizendo que “não ficou claro o que era esperado”. Como posso tornar isso mais transparente no PBL?
P10	Quero permitir o uso de IA generativa no PBL, mas sem que ela resolva o problema pelos alunos. Como orientar esse uso de forma pedagógica?
P11	Em quais etapas do PBL o uso de IA tende a contribuir para a aprendizagem, e em quais pode atrapalhar?
P12	Como posso orientar os alunos a usar IA para refletir sobre o problema, e não apenas para gerar respostas?
P13	Após aplicar PBL com apoio de IA, quais aspectos da atividade devo analisar para saber se a abordagem funcionou pedagogicamente?
P14	Como identificar se os alunos realmente desenvolveram pensamento crítico no PBL ou apenas cumpriram as tarefas?
P15	Que sinais indicam que o problema de PBL precisa ser reformulado para próximas edições da disciplina?
P16	Meus alunos querem começar a “pesquisar no Google/IA” antes de discutir o problema em grupo. Como estruturar a primeira sessão do PBL para garantir que eles formulem hipóteses e questões próprias antes de consultar a IA?
P17	Tenho uma turma grande e poucos tutores. Como adaptar o PBL para 40–60 alunos sem perder acompanhamento do processo, e como a IA pode ajudar sem virar vigilância?

Continuação na próxima página

ID	Pergunta utilizada no experimento
P18	Como desenhar checkpoints (marcos) no PBL para reduzir procrastinação e garantir progressão, sem transformar a atividade em 'tarefas' desconectadas do problema?
P19	Quero avaliar colaboração de forma justa no PBL. Como coletar evidências do trabalho individual e do grupo (incluindo uso de IA) sem criar burocracia excessiva?
P20	Em um projeto PBL, parte dos alunos está copiando respostas da IA. Como redesenhar a tarefa e a avaliação para exigir autoria (raciocínio, evidências e decisões) em vez de texto bem escrito?
P21	Como orientar os alunos a documentarem o uso de IA (prompts, outputs, decisões) de forma ética e útil para aprendizagem, sem expor dados sensíveis?
P22	Quais são sinais de que um grupo está 'travado' no PBL (sem avançar na investigação) e que intervenções mínimas eu posso fazer para destravar sem dar a resposta?
P23	Como equilibrar 'liberdade' e 'andaimagem' no PBL? Tenho medo de orientar pouco e eles se perderem, mas também de orientar demais e matar a autonomia.
P24	Quero que os alunos usem IA para revisar a qualidade das próprias hipóteses e perguntas de pesquisa. Que tipo de prompt e critérios devo ensinar para evitar respostas superficiais?
P25	Como lidar com alucinações e erros da IA durante o PBL sem proibir o uso? Que rotinas de verificação e validação posso exigir?
P26	No PBL, alguns grupos escolhem soluções 'bonitas' mas inviáveis. Como incorporar critérios de viabilidade (tempo, custo, recursos) no processo e na rubrica sem engessar o projeto?

Continuação na próxima página

ID	Pergunta utilizada no experimento
P27	Como planejar uma sessão de 'síntese' no final do ciclo PBL para garantir integração do conhecimento (e não só apresentação), e como a IA pode apoiar a reflexão sem substituir o pensamento do aluno?
P28	Quero que os alunos comparem diferentes caminhos de solução (trade-offs) e justifiquem decisões. Que atividades e perguntas-guia posso usar para promover isso no PBL com apoio de IA?
P29	Como adaptar o PBL para uma disciplina mais teórica (pouca prática de laboratório), mantendo autenticidade do problema, e como usar IA para criar contextos realistas sem virar 'história inventada'?
P30	Como tratar questões de equidade no PBL com IA (alunos com diferentes níveis de acesso/habilidade)? Que combinações de regras, apoio e avaliação reduzem desigualdades?

Tabela 7 – Perguntas utilizadas como prompts no experimento comparativo com LLM-as-a-Judge

5.4.5 Critérios de julgamento e estrutura dos escores

A avaliação comparativa das respostas geradas pelas duas configurações do agente foi realizada por meio de um modelo de linguagem atuando como julgador, a partir de critérios pedagógicos explicitamente definidos. Esses critérios foram formulados com base nos princípios do Problem-Based Learning e nas diretrizes de uso intencional da Inteligência Artificial Generativa no ensino superior, de modo a orientar o julgamento para aspectos didáticos e formativos, e não apenas para a qualidade textual ou correção superficial das respostas.

Cada critério foi avaliado de forma independente, permitindo que o julgador atribuisse escores específicos a dimensões distintas da resposta, incluindo a aderência ao pedido formulado pelo professor, a qualidade pedagógica das recomendações à luz do PBL, a forma como a Inteligência Artificial é integrada de maneira realista e segura, a clareza e aplicabilidade das orientações e o equilíbrio entre completude e concisão. Essa estrutura visa reduzir avaliações excessivamente holísticas e favorecer julgamentos mais analíticos e rastreáveis.

Os escores foram atribuídos em escala ordinal de cinco pontos, variando de 1 a 5, em que

valores mais elevados indicam maior adequação pedagógica da resposta em relação ao critério avaliado. Para cada par de respostas correspondente a um mesmo prompt, o julgador atribuiu escores separadamente às configurações T1 e T2, possibilitando a comparação direta entre elas em cada dimensão considerada.

A partir dos escores por critério, foi calculada uma medida agregada de desempenho para cada resposta, obtida pela soma dos valores atribuídos nos diferentes critérios. A diferença entre as pontuações agregadas das duas configurações, expressa por Δ , constitui a principal medida quantitativa utilizada no experimento para identificar vantagens relativas entre T1 e T2 em cada situação-problema avaliada. Valores positivos de Δ indicam desempenho superior da configuração orientada pelo guideline, enquanto valores negativos indicam vantagem da configuração sem guideline.

Essa combinação de escores por critério e medida agregada permite analisar tanto o comportamento global do agente quanto padrões mais específicos associados a dimensões pedagógicas particulares. A explicitação dessa estrutura de julgamento contribui para a transparência do experimento e fundamenta a interpretação dos resultados apresentada nas subseções subsequentes.

A Tabela 8 sintetiza os critérios de julgamento utilizados no experimento, bem como a descrição operacional de cada critério e a escala de pontuação adotada.

Critério	Descrição operacional	Escala
Aderência ao pedido	Avalia se a resposta atende diretamente à solicitação formulada pelo professor, mantendo-se dentro do escopo da situação-problema apresentada, sem desvios temáticos ou generalizações indevidas.	1 a 5
Qualidade pedagógica em PBL	Verifica se as recomendações estão alinhadas às práticas do Problem-Based Learning, considerando papéis, etapas do método, natureza dos problemas e estratégias de facilitação do professor.	1 a 5
Integração com IA (realista e segura)	Analisa se o uso da Inteligência Artificial é proposto de forma responsável, com atenção a aspectos éticos, privacidade, vieses e limitações, bem como com orientações acionáveis e contextualizadas.	1 a 5
Clareza e aplicabilidade	Considera a clareza da linguagem, a organização da resposta e a presença de orientações práticas que possam ser aplicadas pelo professor em sua atuação pedagógica.	1 a 5
Compleitude sem prolixidade	Avalia se a resposta cobre os pontos pedagógicos relevantes de forma equilibrada, sem omissões importantes nem excesso de explicações desnecessárias.	1 a 5

Tabela 8 – Critérios utilizados pelo LLM-as-a-Judge para o julgamento comparativo das respostas

5.4.6 Consistência dos julgamentos e tratamento de instabilidades

Durante o julgamento comparativo das respostas geradas pelas duas configurações do agente, foram analisados 30 pares de respostas correspondentes às situações-problema apresentadas ao modelo. Em todos os casos, os julgamentos automáticos foram considerados válidos, atendendo à condição `judge_ok=true`, o que permitiu a inclusão integral do conjunto de prompts na análise dos resultados.

Com o objetivo de mitigar possíveis efeitos de *position bias* associados à ordem de apresentação das respostas, o experimento incorporou um procedimento de verificação por inversão (*swap-check*). Esse procedimento foi acionado de forma sistemática, permitindo ao avaliador comparar as respostas com a ordem invertida sempre que aplicável. Ao longo do experimento, o *swap-check* foi executado em todos os julgamentos, resultando em um total de 35 rodadas de inversão distribuídas entre os 30 pares avaliados.

A análise dos registros de julgamento indicou que, na grande maioria dos casos, a preferência do avaliador manteve-se consistente após a inversão da ordem das respostas. Apenas um caso

apresentou instabilidade persistente, caracterizada pela ausência de concordância entre as rodadas de inversão. Nessa situação, conforme a política definida na configuração experimental, a decisão final foi tomada com base na comparação dos escores agregados atribuídos às respostas, respeitando-se a tolerância estabelecida para empates.

Os dados consolidados de consistência dos julgamentos evidenciam que o procedimento adotado foi suficiente para controlar instabilidades e tornar o processo avaliativo rastreável. Os registros referentes às rodadas de inversão, à identificação de instabilidades e às regras de decisão aplicadas foram preservados como parte dos artefatos do experimento, compondo a base empírica utilizada para a síntese dos resultados apresentada nas subseções subsequentes.

Indicador	Valor observado
Total de prompts avaliados	30
Julgamentos válidos (<code>judge_ok=true</code>)	30
Total de rodadas de swap-check	35
Casos com instabilidade persistente	1
Decisões por escore agregado	1
Casos com empate dentro da tolerância	0

Quadro 10 – Síntese dos registros de consistência e instabilidade dos julgamentos automáticos

5.4.7 Artefatos de resultados gerados e rastreabilidade

A execução do experimento comparativo resultou na geração de um conjunto de artefatos estruturados responsáveis por registrar, de forma explícita, as entradas do experimento, os julgamentos intermediários e os resultados consolidados do processo avaliativo. Esses artefatos constituem a base empírica do estudo e permitem tanto a inspeção detalhada dos julgamentos individuais quanto a rastreabilidade e a reprodutibilidade das análises realizadas.

O principal artefato gerado é o arquivo `judgments.jsonl`, que armazena os registros completos de julgamento para cada par de respostas avaliado. Cada entrada desse arquivo corresponde a uma situação-problema específica e contém informações sobre a validade do julgamento (`judge_ok`), os escores atribuídos por critério a cada configuração (T1 e T2), a decisão final do avaliador, a diferença agregada de desempenho (Δ), bem como dados relacionados à aplicação do procedimento de inversão de ordem (*swap-check*) e à estabilidade do julgamento. Esse artefato preserva a granularidade máxima do processo avaliativo, possibilitando a reconstrução integral das decisões tomadas pelo avaliador automático.

Para ilustrar a estrutura e o conteúdo desses registros, apresenta-se a seguir um exemplo representativo de uma entrada do arquivo `judgments.jsonl`.

```
{
  "prompt_id": "P10",
  "judge_ok": true,
  "winner": "T2",
  "scores": {
    "T1": {
      "aderencia_pedido": 3,
      "qualidade_pbl": 3,
      "integracao_ia": 2,
      "clareza_aplicabilidade": 3,
      "completude_concisao": 3
    },
    "T2": {
      "aderencia_pedido": 5,
      "qualidade_pbl": 5,
      "integracao_ia": 4,
      "clareza_aplicabilidade": 4,
      "completude_concisao": 4
    }
  },
  "delta": 7,
  "swap_check": true,
  "stable": true
}
```

Código Fonte 2 – Exemplo de registro individual de julgamento no arquivo `judgments.jsonl`

De forma complementar, o arquivo `results.csv` apresenta uma consolidação tabular dos julgamentos realizados, agregando os escores por configuração e sintetizando as métricas utilizadas na análise comparativa. Esse arquivo contém, para cada prompt, a identificação da configuração vencedora, as pontuações totais atribuídas a T1 e T2, o valor de Δ , além de indicadores relacionados à validade do julgamento e à aplicação do *swap-check*. O `results.csv` foi utilizado como base para a geração das sínteses quantitativas e dos quadros apresentados nesta seção.

A seguir, apresenta-se um exemplo representativo de uma linha do arquivo consolidado `results.csv`.

```
prompt_id, winner, score_t1, score_t2, delta, judge_ok, swap_check, stable
P10, T2, 14, 22, 8, true, true, true
```

Código Fonte 3 – Exemplo de linha consolidada do arquivo `results.csv`

Além desses arquivos centrais, o experimento gerou registros auxiliares associados à configuração e à execução do processo, incluindo arquivos de configuração versionados, logs de execução e subconjuntos de julgamentos selecionados para auditoria. Esses registros permitem verificar quais parâmetros estavam ativos em cada execução, bem como identificar os julgamentos incluídos ou excluídos da análise, reforçando a rastreabilidade temporal e técnica do experimento.

A explicitação dos artefatos produzidos e de exemplos representativos de seus registros delimita de forma clara o conjunto de evidências empíricas que fundamentam os resultados apresentados.

5.4.8 Síntese dos resultados do experimento

A partir do experimento comparativo conduzido com base na abordagem *LLM-as-a-Judge*, foi possível sintetizar os resultados observados considerando o conjunto de 30 situações-problema avaliadas. Todos os julgamentos realizados foram considerados válidos (`judge_ok=true`), permitindo a inclusão integral dos prompts na análise comparativa entre as duas configurações do agente.

No que se refere ao resultado global das comparações, a configuração orientada pelo guideline (T2) apresentou desempenho superior em 27 dos 30 prompts avaliados, enquanto a configuração sem guideline (T1) obteve desempenho superior em 3 casos. Não foram observados empates dentro da tolerância estabelecida. Esses resultados indicam uma predominância consistente de T2 nas avaliações realizadas pelo julgador automático.

A análise dos escores agregados evidenciou uma diferença média positiva de desempenho (Δ) em favor da configuração T2, com valor médio aproximado de 1,83 pontos e mediana igual a 2,0 pontos. Esses valores refletem a soma dos escores atribuídos aos critérios de julgamento definidos, considerando cada situação-problema de forma independente. A distribuição dos valores de Δ mostrou-se concentrada em valores positivos, com variações pontuais observadas ao longo do conjunto avaliado.

No que diz respeito à consistência dos julgamentos, o procedimento de verificação por inversão (*swap-check*) indicou elevada estabilidade nas decisões do avaliador automático. Apenas um caso apresentou instabilidade persistente após as inversões previstas, sendo resolvido com base nos escores agregados, conforme a regra de decisão definida na configuração experimental. Esse resultado sugere que os mecanismos adotados foram suficientes para mitigar efeitos de

viés de ordem e tornar o processo avaliativo rastreável.

Em conjunto, os resultados sintetizados nesta subseção caracterizam o comportamento observado das duas configurações do agente no experimento comparativo. As evidências apresentadas delimitam quantitativamente as diferenças de desempenho identificadas e fundamentam a análise interpretativa desenvolvida no capítulo de Discussão, no qual esses achados são examinados à luz dos objetivos da pesquisa e da literatura relacionada.

5.5 SÍNTESE DOS RESULTADOS

Os resultados da intervenção evidenciaram efeitos observáveis tanto na avaliação do guideline por especialistas quanto no experimento comparativo conduzido com base na abordagem *LLM-as-a-Judge*. Na avaliação por especialistas ($N = 3$), o guideline apresentou médias elevadas nos itens relacionados à clareza, organização, aplicabilidade pedagógica e utilidade geral, com valores predominantemente entre 4,33 e 4,67 em escala ordinal de cinco pontos. A avaliação geral do material e a recomendação a outros professores alcançaram mediana igual a 5,0, indicando elevada aceitação entre os participantes, ainda que acompanhada de sugestões pontuais de aprimoramento.

De forma complementar, o experimento comparativo realizado com 30 situações-problema indicou desempenho superior da configuração orientada pelo guideline (T2) em 27 dos 30 casos avaliados. A diferença agregada de desempenho (Δ) apresentou valor médio aproximado de 1,83 pontos e mediana igual a 2,0 pontos, considerando a soma dos escores atribuídos aos critérios de julgamento definidos. Todos os julgamentos foram considerados válidos, e os mecanismos de controle de consistência adotados mostraram elevada estabilidade, com apenas um caso de instabilidade persistente após aplicação do *swap-check*.

Em conjunto, os resultados apresentados neste capítulo caracterizam de forma descritiva e rastreável os efeitos observados da intervenção, fornecendo a base empírica para a análise interpretativa desenvolvida no capítulo de Discussão.

6 DISCUSSÃO

6.1 RETOMADA DOS OBJETIVOS DA PESQUISA E NATUREZA DOS ACHADOS

Esta pesquisa teve como objetivo geral investigar como a GenAI pode dar suporte à prática docente no ensino superior, tomando o PBL como referência pedagógica. Para atender a esse objetivo, o estudo foi orientado por três objetivos específicos: compreender os desafios enfrentados por docentes diante da presença crescente da GenAI no ensino superior (O1); analisar como a GenAI pode apoiar a prática docente nas diferentes etapas do PBL, com ênfase em decisões pedagógicas relacionadas ao planejamento e à mediação do processo investigativo (O2); e propor e avaliar formas de suporte à tomada de decisão docente por meio do desenvolvimento de artefatos educacionais (O3).

De modo geral, os objetivos O1 e O2 foram alcançados por meio do diagnóstico (RLR e atividade exploratória com docentes) e pela sistematização do apoio da GenAI às etapas do PBL, refletida no design do guideline. O objetivo O3 foi alcançado parcialmente, pois os artefatos foram desenvolvidos e avaliados por evidência formativa (revisão por especialistas) e evidência exploratória (experimento comparativo baseado em *LLM-as-a-Judge*), sem validação longitudinal em sala de aula ou mensuração de impacto direto na aprendizagem discente, o que limita inferências sobre eficácia educacional.

No diagnóstico, a análise temática identificou sete categorias de desafios associados à adoção da GenAI no ensino superior: (1) dependência excessiva da IA; (2) preocupações éticas e de privacidade; (3) falta de transparência e explicabilidade; (4) desigualdade no acesso a recursos tecnológicos; (5) ausência de formação docente; (6) barreiras culturais e linguísticas; e (7) impacto ambiental. Neste estudo, a intervenção abordou diretamente a dependência excessiva da IA e a ausência de formação docente, e incorporou princípios de uso responsável (ética, privacidade e transparência) de forma transversal. Desigualdade de acesso e impacto ambiental foram reconhecidos, mas permaneceram fora do escopo por envolverem fatores institucionais e de escala macro.

Assim, os achados discutidos a seguir devem ser interpretados como uma evidência situada, coerente com a natureza da DSR, e voltada ao suporte à prática docente e à qualidade de orientação pedagógica.

6.2 DISCUSSÃO DO GUIDELINE COMO ARTEFATO PEDAGÓGICO MEDIADOR

Os resultados obtidos ao longo da intervenção indicam que o guideline desenvolvido atua como um artefato pedagógico mediador ao traduzir princípios teóricos do PBL e cuidados associados ao uso da GenAI em orientações estruturadas para a prática docente. Em vez de prescrever procedimentos fixos ou recomendar ferramentas específicas, o guideline organiza decisões pedagógicas em torno das etapas do PBL, oferecendo ao professor referências para planejar, conduzir e acompanhar atividades de aprendizagem ativa mediadas por tecnologia.

Um elemento central desse papel mediador reside na adoção de uma taxonomia de ferramentas de GenAI, organizada por categorias funcionais e articulada às etapas do PBL. Ao mapear possibilidades de uso da GenAI conforme momentos distintos do ciclo do PBL, como compreensão do problema, investigação, síntese e reflexão, o guideline permite que o docente avalie a pertinência pedagógica do uso da tecnologia em cada etapa, evitando aplicações genéricas ou desalinhadas aos objetivos formativos. Essa organização contribui para reduzir a complexidade percebida na integração da GenAI ao ensino, sem deslocar a responsabilidade pedagógica do professor.

Quando comparado ao PBL tradicional, o guideline não altera a lógica central da metodologia, nem redefine suas etapas fundamentais. Problema, investigação, trabalho colaborativo, síntese e reflexão permanecem como elementos estruturantes do processo de aprendizagem. A contribuição do guideline consiste em explicitar como a GenAI pode atuar como recurso de apoio em cada uma dessas etapas, sem substituir práticas centrais do PBL, como a formulação de hipóteses, a busca ativa por informações e a negociação de significados em grupo. Nesse sentido, o artefato preserva a coerência com o PBL clássico, ao mesmo tempo em que responde a desafios contemporâneos associados à presença de tecnologias inteligentes no ensino superior.

Essa abordagem distingue o guideline desenvolvido neste estudo de modelos de PBL integrados à IA descritos na literatura, como o AI-PBL Framework proposto por (SYAHRI; INDRAWAN et al., 2025). Enquanto esse tipo de framework propõe a incorporação explícita de interações com sistemas de IA como parte da própria sintaxe do PBL, incluindo etapas dedicadas à formulação de instruções para a IA, o guideline aqui discutido adota uma postura mais flexível e menos prescritiva. Em vez de incorporar a IA como etapa obrigatória do método, o guideline oferece critérios para que o docente decida quando, como e se a GenAI deve ser utilizada, preservando maior autonomia pedagógica e favorecendo adaptação a diferentes

contextos institucionais e áreas de conhecimento.

Outro aspecto relevante da mediação pedagógica promovida pelo guideline refere-se à explicitação de princípios de uso ético e responsável da GenAI. As orientações enfatizam a tecnologia como recurso de apoio ao processo investigativo e reflexivo, e não como fonte de respostas prontas. Essa decisão pedagógica dialoga com preocupações recorrentes na literatura sobre dependência excessiva de tecnologias inteligentes e sobre a delegação indevida de processos cognitivos e metacognitivos aos sistemas de IA. Ao incorporar esses cuidados de forma transversal, o guideline contribui para preservar o protagonismo discente e o desenvolvimento do pensamento crítico, elementos centrais do PBL.

A avaliação do guideline por especialistas reforça essa interpretação. Os feedbacks indicam que o material apresenta clareza conceitual, organização coerente e aplicabilidade pedagógica, especialmente no apoio ao planejamento das atividades e na explicitação do papel do professor como mediador do processo de aprendizagem. Esses resultados sugerem que o guideline cumpre sua função ao tornar explícitas racionalidades pedagógicas que, em muitos contextos, permanecem implícitas ou dependentes exclusivamente da experiência prévia do docente, funcionando como um suporte à tomada de decisão pedagógica em cenários marcados pela rápida incorporação da GenAI.

Além disso, a natureza não prescritiva do guideline favorece sua adaptação a diferentes realidades institucionais, políticas de uso de tecnologia e culturas docentes. Ao evitar sequências rígidas de ações ou listas fechadas de ferramentas, o artefato mantém sua validade mesmo diante da evolução acelerada do ecossistema de GenAI, reforçando seu caráter formativo e orientador, em vez de normativo.

Entretanto, é importante reconhecer os limites do guideline como artefato pedagógico. Os resultados obtidos não permitem afirmar impactos diretos sobre a aprendizagem dos estudantes ou sobre a eficácia do PBL em contextos reais de sala de aula. O potencial mediador identificado refere-se ao apoio à prática docente, especialmente no planejamento, na reflexão e na tomada de decisão pedagógica, e não à mensuração de resultados educacionais.

Em síntese, a discussão dos resultados indica que o guideline constitui a principal contribuição desta pesquisa ao atuar como artefato pedagógico mediador entre teoria e prática. Seu valor reside menos na prescrição de usos da GenAI e mais na explicitação de critérios, categorias e possibilidades que auxiliam o professor a integrar intencionalmente a tecnologia ao PBL, preservando princípios da aprendizagem ativa, o protagonismo discente e a centralidade da mediação docente.

6.3 O CHATBOT TUTOR COMO INSTRUMENTO DE APOIO E EXPERIMENTAÇÃO

A partir da concepção do guideline como principal contribuição deste estudo, o chatbot tutor foi desenvolvido como um artefato complementar, com a função de operacionalizar esse referencial pedagógico em um formato conversacional e de permitir a exploração controlada de seu efeito sobre a qualidade das orientações produzidas por um agente baseado em modelos de linguagem. Assim, o chatbot não introduz um novo eixo conceitual à pesquisa, mas atua como extensão instrumental do guideline, tornando suas orientações acessíveis de forma contextualizada e sob demanda.

Nessa configuração, o chatbot funciona como uma interface de apoio à prática docente, permitindo que o professor formule questões específicas dizendo respeito ao planejamento, à condução ou à mediação de atividades de PBL, e receba respostas ancoradas nos critérios, categorias e cuidados explicitados no guideline. O valor do chatbot, portanto, não reside em sua autonomia ou em sua atuação como tutor discente, mas em sua capacidade de traduzir o conhecimento pedagógico estruturado do guideline em orientações situadas, preservando a centralidade do julgamento profissional do docente.

O experimento comparativo realizado com o chatbot teve como objetivo examinar, de forma exploratória, se a ancoragem explícita no guideline influencia a qualidade das orientações geradas pelo agente. Os resultados indicam um padrão consistente de melhoria relativa nas respostas produzidas quando o chatbot opera com acesso ao guideline, em comparação à condição em que esse referencial pedagógico não está presente. Essa diferença manifesta-se principalmente em aspectos relacionados à aderência ao pedido docente, à articulação com as etapas do PBL, à clareza e aplicabilidade das recomendações e à incorporação mais explícita de cuidados éticos no uso da GenAI.

Esses achados não devem ser interpretados como evidência de eficácia educacional em contextos reais de ensino, nem como validação do chatbot como produto educacional final. O experimento produz evidência comparativa sobre a qualidade da orientação pedagógica gerada em condições controladas, sugerindo que a explicitação de critérios e princípios pedagógicos, materializada no guideline, exerce influência sobre o comportamento do agente. Nesse sentido, o chatbot e o experimento funcionam como evidência indireta do valor do guideline enquanto base pedagógica explícita, e não como contribuição independente.

A interpretação dos resultados aponta para mecanismos plausíveis que ajudam a explicar esse padrão, sem assumir relações causais fortes. Entre esses mecanismos, destacam-se a

organização das orientações por etapas do PBL, a presença de recomendações acionáveis para a prática docente e a incorporação transversal de cuidados éticos e pedagógicos no uso da GenAI. Ao reduzir ambiguidades e tornar explícitas racionalidades pedagógicas que orientam a tomada de decisão do professor, o guideline fornece ao agente um enquadramento mais consistente para a geração de respostas alinhadas à aprendizagem ativa.

É igualmente necessário reconhecer os limites metodológicos do experimento. A abordagem baseada em *LLM-as-a-Judge* permite comparações sistemáticas em tarefas abertas, mas permanece sujeita a vieses e instabilidades inerentes a julgamentos automatizados, como sensibilidade à ordem de apresentação das respostas e preferências implícitas do modelo avaliador. As estratégias adotadas no protocolo experimental, incluindo cegamento, padronização do formato de julgamento e verificação de consistência, contribuem para mitigar esses riscos, mas não eliminam completamente suas limitações. Assim, os resultados devem ser compreendidos como evidência exploratória e complementar, adequada ao escopo e à natureza formativa da pesquisa.

Além disso, a adoção da abordagem *LLM-as-a-Judge* pode ser justificada por evidências recentes de que, em tarefas avaliativas bem delimitadas e com critérios explícitos, modelos de linguagem podem apresentar níveis de consistência e concordância próximos aos de avaliadores humanos (LI et al., 2024). Estudos em contexto educacional indicam que LLMs são particularmente adequados para atividades de avaliação formativa e comparativa, funcionando como instrumentos de triagem e apoio analítico quando o objetivo é comparar condições ou configurações distintas, e não substituir integralmente o julgamento humano (MENDONÇA; QUINTAL; MENDONÇA, 2025). Ao mesmo tempo, a literatura ressalta que a validade desse tipo de julgamento automatizado depende fortemente do desenho experimental, da explicitação dos critérios e da adoção de controles para mitigar vieses e instabilidades, reforçando que seus resultados devem ser interpretados como evidência exploratória e complementar, e não como prova conclusiva de eficácia educacional em contextos reais de ensino (LI et al., 2024).

Nesse enquadramento, o chatbot tutor cumpre seu papel no desenho do estudo ao reforçar, de forma empírica e controlada, a relevância do guideline como artefato pedagógico mediador. Em vez de deslocar o foco da contribuição central, o experimento com o chatbot evidencia que a explicitação pedagógica promovida pelo guideline constitui um elemento decisivo para a qualidade da orientação fornecida por agentes conversacionais voltados ao apoio à prática docente.

6.4 IMPLICAÇÕES PEDAGÓGICAS E INSTITUCIONAIS

Os achados desta pesquisa apontam implicações relevantes para a prática pedagógica de professores universitários e para instituições de ensino superior que enfrentam o desafio de integrar a GenAI a metodologias de aprendizagem ativa. Em ambos os casos, os resultados indicam que a simples disponibilidade de ferramentas baseadas em GenAI é insuficiente para garantir uso pedagógico adequado, reforçando a necessidade de mediação docente intencional e de referenciais explícitos que orientem a tomada de decisão em contextos de ensino.

Do ponto de vista da prática docente, o guideline desenvolvido oferece um suporte estruturado para a integração da GenAI ao PBL sem comprometer princípios centrais da aprendizagem ativa. Ao organizar possibilidades de uso da tecnologia por etapas do PBL e por categorias funcionais, o guideline auxilia o professor a decidir quando e como a GenAI pode apoiar o processo investigativo, a reflexão e a síntese, preservando o protagonismo discente e evitando usos que substituam etapas formativas do método. Nesse sentido, a implicação pedagógica central não está na adoção de ferramentas específicas, mas na explicitação de critérios pedagógicos que permitem ao docente exercer sua função mediadora de forma mais consciente e alinhada aos objetivos formativos.

Outra implicação pedagógica relevante refere-se ao uso responsável da GenAI. Os resultados sugerem que a incorporação explícita de cuidados éticos e pedagógicos, como a limitação do uso da IA como fonte de respostas prontas e o incentivo à reflexão crítica sobre suas saídas, contribui para mitigar riscos associados à dependência excessiva e à delegação indevida de processos cognitivos. Assim, o guideline pode apoiar docentes na construção de práticas mais transparentes e reflexivas, favorecendo abordagens avaliativas que valorizem processos, justificativas e participação ativa, em vez de apenas produtos finais.

No âmbito institucional, os achados indicam a importância de diretrizes pedagógicas que vão além de políticas genéricas de uso de tecnologia. A integração da GenAI no ensino superior demanda ações coordenadas que envolvam formação docente contínua, produção de materiais de apoio e definição de princípios orientadores que considerem especificidades disciplinares e contextuais. Nesse cenário, o guideline pode ser interpretado como um exemplo de artefato pedagógico que complementa políticas institucionais ao traduzir princípios abstratos em orientações acionáveis para a prática, sem assumir caráter normativo rígido.

As implicações institucionais também incluem a necessidade de repensar estratégias de avaliação e governança do uso da GenAI. Os resultados reforçam que abordagens avaliativas

centradas exclusivamente no produto tornam-se menos adequadas em contextos nos quais ferramentas inteligentes podem influenciar significativamente a produção discente. Nesse sentido, instituições são desafiadas a promover práticas avaliativas que enfatizem processos de aprendizagem, participação e reflexão, o que exige alinhamento entre diretrizes pedagógicas, formação docente e políticas institucionais mais amplas.

Por fim, no que se refere à pesquisa e ao desenvolvimento de artefatos educacionais, os achados destacam o valor de soluções que explicitam referenciais pedagógicos e que são avaliadas de forma formativa e exploratória. O uso do chatbot tutor e do experimento comparativo indica que agentes conversacionais tendem a produzir orientações mais alinhadas a objetivos educacionais quando ancorados em bases pedagógicas explícitas, reforçando a importância de integrar conhecimento pedagógico estruturado ao desenvolvimento de sistemas baseados em GenAI. Além disso, os resultados sugerem que avaliações automatizadas, quando utilizadas com cautela e transparência, podem desempenhar papel complementar em estudos exploratórios, sem substituir avaliações humanas ou evidências produzidas em contextos reais de ensino.

6.5 REFLEXÕES FINAIS

A interpretação dos resultados a partir do ciclo da DSR adotado neste estudo permite compreender a investigação como um processo iterativo orientado à prática docente, e não apenas como a produção de artefatos isolados. Os achados refletem decisões metodológicas e condicionantes institucionais específicas, devendo ser interpretados como evidência situada de um ciclo formativo e exploratório que buscou responder aos objetivos definidos para a pesquisa.

O diagnóstico inicial contribuiu diretamente para o atendimento do primeiro objetivo específico, ao identificar desafios pedagógicos, éticos e institucionais associados ao uso da Inteligência Artificial Generativa no ensino superior. Esses achados orientaram o planejamento da ação e fundamentaram a decisão de estruturar a intervenção em torno da construção de artefatos educacionais, entendidos como meios viáveis de apoio à prática docente no escopo temporal e institucional da pesquisa.

O segundo objetivo específico foi endereçado no desenvolvimento do guideline pedagógico, no qual os problemas identificados no diagnóstico foram traduzidos em critérios e orientações explícitas para a integração da GenAI às etapas do PBL. Esse movimento evidenciou o potencial do guideline como instrumento mediador, ao sistematizar princípios pedagógicos e apoiar a tomada de decisão docente sem assumir caráter prescritivo ou dependente de ferramentas

específicas.

O terceiro objetivo específico foi explorado de forma complementar por meio do desenvolvimento do chatbot tutor e da condução do experimento comparativo. A avaliação formativa, apoiada pela abordagem *LLM-as-a-Judge*, possibilitou refletir sobre o efeito da ancoragem pedagógica no comportamento do agente, oferecendo evidência exploratória sobre a relevância do conhecimento pedagógico estruturado para a qualidade da orientação fornecida, sem pretensão de inferir impactos educacionais diretos.

7 CONCLUSÃO

7.1 SÍNTESE DOS RESULTADOS

Esta pesquisa investigou como a GenAI pode dar suporte à prática docente no ensino superior, tomando o PBL como metodologia pedagógica de referência. Para isso, adotou-se a abordagem metodológica de DSR, inspirada nos ciclos iterativos de uma pesquisa-ação, estruturada em etapas de diagnóstico, planejamento, intervenção, avaliação e reflexão, com foco na compreensão dos desafios enfrentados por professores universitários e na proposição de formas de apoio à tomada de decisão pedagógica.

O diagnóstico inicial evidenciou que os desafios associados ao uso da GenAI no ensino superior extrapolam aspectos técnicos, envolvendo preocupações pedagógicas, éticas e institucionais, especialmente em contextos de metodologias ativas. Esses achados fundamentaram a escolha do PBL como eixo estruturante da intervenção e orientaram a definição de artefatos educacionais voltados ao apoio à prática docente.

No diagnóstico, a análise temática identificou sete categorias de desafios: dependência excessiva da IA; preocupações éticas e de privacidade; falta de transparência e explicabilidade; desigualdade de acesso; ausência de formação docente; barreiras culturais e linguísticas; e impacto ambiental. A intervenção tratou diretamente da dependência excessiva e da formação docente, incorporou princípios de uso responsável (ética, privacidade e transparência) de forma transversal, e reconheceu desigualdade de acesso e impacto ambiental como temas fora do escopo por demandarem ações institucionais e de escala macro.

Como resultado da intervenção, foram desenvolvidos dois artefatos complementares: um *guideline* pedagógico, concebido para orientar professores na integração intencional da GenAI às etapas do PBL, e um chatbot tutor, projetado como instrumento de apoio conversacional contextualizado. A avaliação do *guideline* por especialistas indicou clareza, organização e aplicabilidade pedagógica, enquanto o experimento comparativo com *LLM-as-a-Judge* sugeriu que a ancoragem do chatbot em conhecimento pedagógico estruturado contribui para respostas mais alinhadas aos princípios do PBL.

Em conjunto, os resultados indicam que a GenAI pode atuar como recurso de apoio ao planejamento, à mediação pedagógica e à reflexão docente, desde que integrada de forma intencional e orientada por princípios pedagógicos explícitos. Esses achados devem ser compreendidos como evidência formativa e exploratória, coerente com a natureza situada

da metodologia adotada.

7.2 CONTRIBUIÇÕES DA PESQUISA

As contribuições desta pesquisa podem ser analisadas sob três perspectivas complementares: pedagógica, metodológica e prática.

Do ponto de vista pedagógico, o estudo contribui ao explicitar caminhos para a integração da GenAI ao PBL sem descaracterizar os princípios da aprendizagem ativa. O *guideline* desenvolvido atua como artefato pedagógico mediador, ao traduzir conceitos teóricos e cuidados éticos em orientações estruturadas para a prática docente, reforçando o papel do professor como mediador do processo de aprendizagem.

No âmbito metodológico, a pesquisa contribui ao explorar o uso da abordagem *LLM-as-a-Judge* como estratégia de avaliação comparativa em contextos educacionais. Ao combinar critérios pedagógicos explícitos, controles de consistência e auditoria humana, o estudo amplia o repertório de métodos disponíveis para análise formativa de artefatos baseados em modelos de linguagem, destacando tanto seu potencial quanto seus limites.

Do ponto de vista prático, os artefatos desenvolvidos oferecem subsídios concretos para professores universitários e instituições de ensino superior que buscam orientar o uso responsável da GenAI. Ao priorizar estruturas flexíveis e não prescritivas, o estudo contribui para reduzir incertezas docentes e apoiar decisões pedagógicas em contextos marcados pela rápida disseminação de tecnologias inteligentes.

7.3 LIMITAÇÕES DA PESQUISA

As limitações deste estudo decorrem principalmente de condicionantes temporais e institucionais que afetaram o escopo de desenvolvimento e avaliação dos artefatos propostos. No contexto da abordagem de Design Science Research (DSR), a pesquisa concentrou-se na construção e avaliação inicial de artefatos educacionais, sem contemplar sua aplicação prolongada em contextos autênticos de ensino. A decisão de não realizar uma implementação em sala de aula foi influenciada pelo tempo necessário para obtenção da aprovação ética junto ao Comitê de Ética em Pesquisa da Universidade Federal de Pernambuco (CEP-UFPE), requisito para estudos envolvendo participação de docentes. Embora esse processo seja essencial para assegurar conformidade ética, ele reduziu a janela disponível para avaliações em

uso real. Em razão dessas restrições, optou-se por uma estratégia metodologicamente coerente com a DSR, fundamentada na avaliação de artefatos por especialistas e em experimento controlado para análise comparativa do chatbot tutor.

Além disso, a avaliação do *guideline* foi realizada com um número reduzido de especialistas, o que limita a amplitude de perspectivas consideradas e restringe inferências mais abrangentes sobre a aplicabilidade do material em diferentes áreas e realidades institucionais. De modo complementar, o experimento com *LLM-as-a-Judge* produziu evidência comparativa controlada sobre o efeito do *guideline* na qualidade das respostas do agente, porém não substitui uma avaliação de uso recorrente por docentes em situações autênticas de ensino, nem permite concluir sobre impactos diretos na aprendizagem discente. Assim, os achados devem ser interpretados como evidência inicial de avaliação do artefato, adequada a estágios iniciais de investigação em DSR, mas ainda insuficiente para demonstrar sua efetividade em contextos autênticos de uso.

7.4 TRABALHOS FUTUROS

A partir dos limites identificados, diversas possibilidades de trabalhos futuros podem ser delineadas. Uma direção relevante consiste na aplicação do *guideline* e do chatbot tutor em contextos reais de sala de aula, acompanhando o uso dos artefatos por docentes ao longo de um período prolongado, de modo a investigar impactos sobre práticas pedagógicas, dinâmicas de mediação e percepções docentes.

Outra possibilidade refere-se à ampliação da avaliação do *guideline* com um número maior e mais diverso de especialistas, incluindo docentes de diferentes áreas do conhecimento e instituições, o que permitiria analisar sua adaptabilidade e relevância em contextos variados do ensino superior.

No campo metodológico, estudos futuros podem explorar combinações entre julgamentos automatizados e avaliações humanas, investigando estratégias híbridas para análise de sistemas educacionais baseados em GenAI. Além disso, a adaptação do *guideline* e do chatbot para outras metodologias ativas pode contribuir para compreender a generalidade da abordagem proposta.

REFERÊNCIAS

- AL-MUGHAIIRI, H.; BHASKAR, P. Exploring the factors affecting the adoption ai techniques in higher education: insights from teachers' perspectives on chatgpt. *Journal of Research in Innovative Teaching & Learning*, Emerald Publishing Limited, 2024.
- ALAFNAN, M. A.; DISHARI, S.; MOHDZUKI, S. F. Developing soft skills in the artificial intelligence era: Communication, business writing, and composition skills. *Journal of Artificial Intelligence and Technology*, v. 4, n. 4, p. 305–317, 2024.
- ALOTAIBI, N. S.; ALSHEHRI, A. H. Prospers and obstacles in using artificial intelligence in saudi arabia higher education institutions—the potential of ai-based learning outcomes. *Sustainability*, MDPI, v. 15, n. 13, p. 10723, 2023.
- AN, Y.; YU, J. H.; JAMES, S. Investigating the higher education institutions' guidelines and policies regarding the use of generative ai in teaching, learning, research, and administration. *International Journal of Educational Technology in Higher Education*, Springer, v. 22, n. 1, p. 10, 2025.
- BAIG, M. I.; YADEGARIDEHKORDI, E. Factors influencing academic staff satisfaction and continuous usage of generative artificial intelligence (genai) in higher education. *International Journal of Educational Technology in Higher Education*, Springer, v. 22, n. 1, p. 5, 2025.
- BARDIN, L. Análise de conteúdo. *Lisboa: edições*, v. 70, p. 225, 1977.
- BOWEN, J. A.; WATSON, C. E. *Teaching with AI: A practical guide to a new era of human learning*. [S.l.]: JHU Press, 2025.
- BOWER, M.; TORRINGTON, J.; LAI, J. W.; PETOCZ, P.; ALFANO, M. How should we change teaching and assessment in response to increasingly powerful generative artificial intelligence? outcomes of the chatgpt teacher survey. *Education and Information Technologies*, Springer, v. 29, n. 12, p. 15403–15439, 2024.
- BRANDÃO, A.; PEDRO, L.; ZAGALO, N. Teacher professional development for a future with generative artificial intelligence—an integrative literature review. *Digital Education Review*, ERIC, v. 45, p. 151–157, 2024.
- BUDHWAR, P.; CHOWDHURY, S.; WOOD, G.; AGUINIS, H.; BAMBER, G. J.; BELTRAN, J. R.; BOSELIE, P.; COOKE, F. L.; DECKER, S.; DENISI, A.; DEY, P.; GUEST, D.; KNOBLICH, A. J.; MALIK, A.; PAAUWE, J.; PAPAGIANNIDIS, S.; PATEL, C.; PEREIRA, V.; REN, S.; ROGELBERG, S. C.; SAUNDERS, M. N. K.; TUNG, R. L.; VARMA, A. Human resource management in the age of generative artificial intelligence: Perspectives and research directions on chatgpt. *Human Resource Management Journal*, 2023.
- CARTAXO, B.; PINTO, G.; SOARES, S. Rapid reviews in software engineering. *Contemporary Empirical Methods in Software Engineering*, Springer, p. 357–384, 2020.
- CHAN, C. K. Y.; HU, W. Students' voices on generative ai: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, Springer, v. 20, n. 1, p. 43, 2023.
- CHIANG, C.-H.; LEE, H.-y. Can large language models be an alternative to human evaluations? *arXiv preprint arXiv:2305.01937*, 2023.

CHIU, T. K. The impact of generative ai (genai) on practices, policies and research direction in education: A case of chatgpt and midjourney. *Interactive Learning Environments*, Taylor & Francis, v. 32, n. 10, p. 6187–6203, 2024.

CHOI-LUNDBERG, D. L.; BUTLER-HENDERSON, K.; HARMAN, K.; CRAWFORD, J. A systematic review of digital innovations in technology-enhanced learning designs in higher education. *Australasian Journal of Educational Technology*, v. 39, n. 3, p. 133–162, 2023.

CHUMA, E.; OLIVEIRA, G. G. de. Generative ai for business decision-making: A case of chatgpt. *Management Science and Business Decisions*, 2023.

CORDERO, J.; TORRES-ZAMBRANO, J.; CORDERO-CASTILLO, A. Integration of generative artificial intelligence in higher education: Best practices. *Education Sciences*, MDPI, v. 15, n. 1, p. 32, 2024.

CUKUROVA, M.; MIAO, F. et al. *AI competency framework for teachers*. [S.I.]: UNESCO Publishing, 2024.

DAKAKNI, D.; SAFA, N. Artificial intelligence in the I2 classroom: Implications and challenges on ethics and equity in higher education: A 21st century pandora's box. *Computers and Education: Artificial Intelligence*, Elsevier, v. 5, p. 100179, 2023.

DEMPERE, J. M.; MODUGU, K.; HESHAM, A.; RAMASAMY, L. K. The impact of chatgpt on higher education. *Frontiers in Education*, 2023.

DHAMIJA, A.; DHAMIJA, D. Understanding teachers' perspectives on chatgpt-generated assignments in higher education. *Journal of Interdisciplinary Studies in Education*, v. 14, n. 1, p. 38–62, 2025.

DOGANI, B. Active learning and effective teaching strategies. *International Journal of Advanced Natural Sciences and Engineering Researches*, v. 7, n. 4, p. 136–142, 2023.

EAGER, B.; BRUNTON, R. Prompting higher education towards ai-augmented teaching and learning practice. *Journal of university teaching and learning practice*, Open Access Publishing Association (OAPA) Launceston, Tasmania, v. 20, n. 5, p. 1–19, 2023.

ELSAIYARY, A. Integrating generative ai in active learning environments: enhancing metacognition and technological skills. *Journal of Systemics, Cybernetics and Informatics*, v. 22, n. 3, p. 34–37, 2024.

FAN, Y.; TANG, L.; LE, H.; SHEN, K.; TAN, S.; ZHAO, Y.; SHEN, Y.; LI, X.; GAŠEVIĆ, D. Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, Wiley Online Library, v. 56, n. 2, p. 489–530, 2025.

FARRELLY, T.; BAKER, N. Generative artificial intelligence: Implications and considerations for higher education practice. *Education Sciences*, Multidisciplinary Digital Publishing Institute, v. 13, n. 11, p. 1109, 2023.

FATHI, T. E.; SAAD, A.; LARHZIL, H.; LAMRI, D.; IBRAHMI, E. M. A. Integrating generative ai into stem education: enhancing conceptual understanding, addressing misconceptions, and assessing student acceptance. *Disciplinary and Interdisciplinary Science Education Research*, Springer, v. 7, n. 1, p. 6, 2025.

- FENTA, A. A. A review on enhancing education with ai: exploring the potential of chatgpt, bard, and generative ai. *Discover Education*, Springer, v. 4, n. 1, p. 38, 2025.
- FILIPPO, D.; ROQUE, G.; PEDROSA, S. Pesquisa-ação: possibilidades para a informática educativa. *Metodologia de Pesquisa Científica em Informática na Educação: Abordagem qualitativa de Pesquisa*, v. 3, 2018.
- GANANN, R.; CILISKA, D.; THOMAS, H. Expediting systematic reviews: methods and implications of rapid reviews. *Implementation Science*, Springer, v. 5, n. 1, p. 56, 2010.
- GIL, A. C. *Metodologia Do Ensino Superior*. [S.l.]: Editora Atlas SA, 2000.
- GRÁJEDA, A.; BURGOS, J.; CÓRDOVA, P.; SANJINÉS, A. Assessing student-perceived impact of using artificial intelligence tools: Construction of a synthetic index of application in higher education. *Cogent Education*, Taylor & Francis, v. 11, n. 1, p. 2287917, 2024.
- GUIZANI, S.; MAZHAR, T.; SHAHZAD, T.; AHMAD, W.; BIBI, A.; HAMAM, H. A systematic literature review to implement large language model in higher education: issues and solutions. *Discover Education*, Springer, v. 4, n. 1, p. 1–25, 2025.
- HU, K. *ChatGPT sets record for fastest-growing user base – analyst note*. 2023. Disponível em: <<https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>>.
- IRISH, A. L.; GAZICA, M. W.; BECERRA, V. A qualitative descriptive analysis on generative artificial intelligence: bridging the gap in pedagogy to prepare students for the workplace. *Discover Education*, Springer, v. 4, n. 1, p. 48, 2025.
- ISMAIL, F.; TAN, E.; RUDOLPH, J.; CRAWFORD, J.; TAN, S. Artificial intelligence in higher education. a protocol paper for a systematic literature review. *Journal of Applied Learning & Teaching*, Kaplan Business School Australia Sydney, NSW, v. 6, n. 2, p. 56–63, 2023.
- KEELE, S. et al. *Guidelines for performing systematic literature reviews in software engineering*. [S.l.], 2007.
- KESTIN, G.; MILLER, K.; KLALES, A. et al. *AI tutoring outperforms in-class active learning: an RCT introducing a novel research-based design in an authentic educational setting*. *Sci Rep* 15, 17458 (2025). 2025.
- KHAN, S.; MAZHAR, T.; SHAHZAD, T.; KHAN, M. A.; REHMAN, A. U.; SAEED, M. M.; HAMAM, H. Harnessing ai for sustainable higher education: ethical considerations, operational efficiency, and future directions. *Discover Sustainability*, Springer, v. 6, n. 1, p. 23, 2025.
- KONG, S.-C.; YANG, Y. A human-centered learning and teaching framework using generative artificial intelligence for self-regulated learning development through domain knowledge learning in k–12 settings. *IEEE Transactions on Learning Technologies*, IEEE, v. 17, p. 1562–1573, 2024.
- KURTZ, G.; AMZALAG, M.; SHAKED, N.; ZAGURI, Y.; KOHEN-VACS, D.; GAL, E.; ZAILER, G.; BARAK-MEDINA, E. Strategies for integrating generative ai into higher education: Navigating challenges and leveraging opportunities. *Education Sciences*, MDPI, v. 14, n. 5, p. 503, 2024.

LI, H.; DONG, Q.; CHEN, J.; SU, H.; ZHOU, Y.; AI, Q.; YE, Z.; LIU, Y. Llm-as-judges: a comprehensive survey on llm-based evaluation methods. *arXiv preprint arXiv:2412.05579*, 2024.

LIU, Y.; PÁSZTOR, A. Effects of problem-based learning instructional intervention on critical thinking in higher education: A meta-analysis. *Thinking Skills and Creativity*, Elsevier, v. 45, p. 101069, 2022.

LO, C. What is the impact of chatgpt on education? a rapid review of the literature. *Education Sciences*, 2023.

MENDONÇA, P. C.; QUINTAL, F.; MENDONÇA, F. Evaluating llms for automated scoring in formative assessments. *Applied Sciences*, MDPI, v. 15, n. 5, p. 2787, 2025.

MICHEL-VILLARREAL, R.; VILALTA-PERDOMO, E.; SALINAS-NAVARRO, D. E.; THIERRY-AGUILERA, R.; GERARDOU, F. S. Challenges and opportunities of generative ai for higher education as explained by chatgpt. *Education Sciences*, MDPI, v. 13, n. 9, p. 856, 2023.

MILMO, D.; AGENCY. *ChatGPT reaches 100 million users two months after launch*. 2023. Disponível em: <<https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app>>.

MOLLICK, E. *Co-intelligence: Living and working with AI*. [S.l.]: Penguin, 2024.

NGO, V. M. Does chatgpt change artificial intelligence-enabled marketing capability? social media investigation of public sentiment and usage. *Global Media and China*, v. 9, p. 101 – 125, 2024.

OGUNLEYE, B.; ZAKARIYYAH, K. I.; AJAO, O.; OLAYINKA, O.; SHARMA, H. A systematic review of generative ai for teaching and learning practice. *Education Sciences*, MDPI, v. 14, n. 6, p. 636, 2024.

OpenAI. *Introducing ChatGPT*. 2022. <<https://openai.com/index/chatgpt/>>. Acesso em: 23 mar. 2025.

PARI-BEDOYA, I. N. M. d. I. A.; VARGAS-MURILLO, A. R.; GUEVARA-SOTO, F. d. J. Challenges and opportunities of ai-assisted learning: a systematic literature review on the impact of chatgpt usage in higher education. Society for Research and Knowledge Management, 2023.

PIMENTEL, M.; CARVALHO, F. *IA generativa e educação: práticas e teorizações*. Porto Alegre, Brasil: Sociedade Brasileira de Computação, 2025. ISBN 978-85-7669-658-2. Disponível em: <https://www.researchgate.net/publication/397015920_IA_generativa_e_educacao_praticas_e_teorizacoes>.

RAHIMAN, H.; KODIKAL, R. *Revolutionizing education: Artificial intelligence empowered learning in higher education*. *Cogent Education*, 11 (1), 2293431. 2024.

RASUL, T.; NAIR, S.; KALENDRA, D.; ROBIN, M.; SANTINI, F. de O.; LADEIRA, W.; SUN, M.; DAY, I.; RATHER, R. A.; HEATHCOTE, L. The role of chatgpt in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning & Teaching*, Kaplan Business School Australia Sydney, NSW, v. 6, n. 1, p. 41–56, 2023.

RIBEIRO, L. R. de C. *Aprendizagem baseada em problemas (PBL): uma experiência no ensino superior*. [S.l.]: Edufscar, 2022.

RIBEIRO, L. R. de C. *Aprendizagem baseada em problemas (PBL): uma experiência no ensino superior*. [S.l.]: Edufscar, 2022.

SAAIDA, M. Ai-driven transformations in higher education: Opportunities and challenges. *International Journal of Educational Research and Studies*, v. 5, n. 1, p. 29–36, 2023.

SALINAS-NAVARRO, D. E.; VILALTA-PERDOMO, E.; MICHEL-VILLARREAL, R.; MONTESINOS, L. Using generative artificial intelligence tools to explain and enhance experiential learning for authentic assessment. *Education Sciences*, MDPI, v. 14, n. 1, p. 83, 2024.

SEO, H.; HWANG, T.; JUNG, J.; KANG, H.; NAMGOONG, H.; LEE, Y.; JUNG, S. Large language models as evaluators in education: Verification of feedback consistency and accuracy. *Applied Sciences (2076-3417)*, v. 15, n. 2, 2025.

SHAHZAD, T.; MAZHAR, T.; TARIQ, M. U.; AHMAD, W.; OUAHADA, K.; HAMAM, H. A comprehensive review of large language models: issues and solutions in learning environments. *Discover Sustainability*, Springer, v. 6, n. 1, p. 27, 2025.

SMITH, K.; MAYNARD, N.; BERRY, A.; STEPHENSON, T.; SPITERI, T.; CORRIGAN, D.; MANSFIELD, J.; ELLERTON, P.; SMITH, T. Principles of problem-based learning (pbl) in stem education: Using expert wisdom and research to frame educational practice. *Education Sciences*, MDPI, v. 12, n. 10, p. 728, 2022.

SOK, S.; HENG, K. Opportunities, challenges, and strategies for using chatgpt in higher education: A literature review. *Journal of Digital Educational Technology*, Bastas, v. 4, n. 1, p. ep2401, 2024.

SYAHRI, B.; INDRAWAN, E. et al. Ai-pbl framework: Innovative problem based learning model supported by artificial intelligence technology. *Data and Metadata*, AG Editor, v. 4, p. 1116–1116, 2025.

SZYMANSKI, A.; ZIEMS, N.; EICHER-MILLER, H. A.; LI, T. J.-J.; JIANG, M.; METOYER, R. A. Limitations of the llm-as-a-judge approach for evaluating llm outputs in expert knowledge tasks. In: *Proceedings of the 30th International Conference on Intelligent User Interfaces*. [S.l.: s.n.], 2025. p. 952–966.

TAJRA, S.; MONTEIRO, P.; CHAMON, E.; SANTANA, L.; RIBEIRO, J.; ARANHA, M. *Metodologias ativas e as tecnologias educacionais: conceitos e práticas*. Alta Books, 2021. ISBN 9786555202786. Disponível em: <<https://books.google.com.br/books?id=8T9IEAAQBAJ>>.

TAN, S.; ZHUANG, S.; MONTGOMERY, K.; TANG, W. Y.; CUADRON, A.; WANG, C.; POPA, R. A.; STOICA, I. Judgebench: A benchmark for evaluating llm-based judges. *arXiv preprint arXiv:2410.12784*, 2024.

TEAM, O. E. R. *OpenAI's new economic analysis*. 2025. Disponível em: <<https://openai.com/global-affairs/new-economic-analysis/>>.

THEOBALD, E. J.; HILL, M. J.; TRAN, E.; AGRAWAL, S.; ARROYO, E. N.; BEHLING, S.; CHAMBWE, N.; CINTRÓN, D. L.; COOPER, J. D.; DUNSTER, G. et al. Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, National Academy of Sciences, v. 117, n. 12, p. 6476–6483, 2020.

TUBELLA, A. A.; MORA-CANTALLOPS, M.; NIEVES, J. C. How to teach responsible ai in higher education: challenges and opportunities. *Ethics and Information Technology*, Springer, v. 26, n. 1, p. 3, 2024.

VIBERG, O.; MUTIMUKWE, C.; HRASTINSKI, S.; CERRATTO-PARGMAN, T.; LILLIESKÖLD, J. Exploring teachers' (future) digital assessment practices in higher education: Instrument and model development. *British journal of educational technology*, Wiley Online Library, v. 55, n. 6, p. 2597–2616, 2024.

WALCZAK, K.; CELLARY, W. Challenges for higher education in the era of widespread access to generative ai. *Economics and Business Review*, v. 9, n. 2, 2023.

WALSH, A. The tutor in problem based learning: a novice's guide. *Hamilton: McMaster University*, 2005.

WANG, H.; DANG, A.; WU, Z.; MAC, S. Generative ai in higher education: Seeing chatgpt through universities' policies, resources, and guidelines. *Computers and Education: Artificial Intelligence*, Elsevier, v. 7, p. 100326, 2024.

WANG, R.; GUO, J.; GAO, C.; FAN, G.; CHONG, C. Y.; XIA, X. Can llms replace human evaluators? an empirical study of llm-as-a-judge in software engineering. *Proceedings of the ACM on Software Engineering*, ACM New York, NY, USA, v. 2, n. ISSTA, p. 1955–1977, 2025.

WEN, J.; WANG, W. The future of chatgpt in academic research and publishing: A commentary for clinical and translational medicine. *Clinical and Translational Medicine*, v. 13, 2023.

WIESE, E.; HATLEVIK, I. K. R.; DAZA, V. How can universities ensure quality of practice in initial teacher education? *Teaching and Teacher Education*, Elsevier, v. 139, p. 104462, 2024.

WILLIAMS, R. *The ethical implications of using generative chatbots in higher education*. *Frontiers in Education*, 8, 1331607. 2023.

YERALAN, S.; LEE, L. A. Generative ai: Challenges to higher education. *Sustainable Engineering and Innovation*, Research and Development Academy, v. 5, n. 2, p. 107–116, 2023.

ZARIFHONARVAR, A. Economics of chatgpt: A labor market view on the occupational impact of artificial intelligence. *SSRN Electronic Journal*, 2023.

ZEB, A.; ULLAH, R.; KARIM, R. Exploring the role of chatgpt in higher education: opportunities, challenges and ethical considerations. *The International Journal of Information and Learning Technology*, Emerald Publishing Limited, v. 41, n. 1, p. 99–111, 2024.

APÊNDICE A – PROTOCOLO DE RAPID LITERATURE REVIEW (RLR)

PROBLEMA PRÁTICO

Apesar da crescente adoção da inteligência artificial (IA) no ensino superior, professores universitários enfrentam desafios significativos relacionados à integração pedagógica dessas tecnologias. Os desafios incluem aspectos éticos, metodológicos, institucionais e pedagógicos. Uma revisão rápida da literatura pode fornecer evidências úteis e atuais para auxiliar no direcionamento de ações práticas em curto prazo, como formação docente, elaboração de diretrizes e criação de metodologias de ensino.

OBJETIVO

Sintetizar de forma rápida e sistemática os principais desafios enfrentados por professores universitários na adoção de inteligência artificial no ensino superior com fins pedagógicos.

PERGUNTA DE PESQUISA

Quais são os principais desafios enfrentados por professores universitários na adoção de inteligência artificial para fins pedagógicos?

STAKEHOLDERS ENVOLVIDOS

- **Pesquisadores:** responsáveis por conduzir a RLR
- **Professores universitários:** potenciais usuários dos resultados

PLANEJAMENTO

- Bases de dados utilizadas: Google Scholar (Top 20) e Scopus (Top 10)
- Estratégia de busca: "Artificial Intelligence"OR "AI") AND ("Higher Education"OR "University Professors") AND ("Challenges"OR "Barriers"OR "Adoption")
- Período de busca: 2023–2025

- Critérios de inclusão:
 - Estudos focados no ensino superior
 - Estudos que tratam diretamente dos desafios enfrentados por docentes
- Critérios de exclusão:
 - Estudos sobre educação básica
 - Trabalhos excessivamente técnicos sem foco pedagógico
 - Estudos que não abordam a perspectiva docente

EXECUÇÃO

- Total de estudos analisados: 31 (21 do Google Scholar e 10 da Scopus)
- Seleção em três etapas:
 1. Leitura de título e resumo
 2. Leitura de introdução e conclusão
 3. Seleção final de 10 artigos
- Extração de dados:
 - Desafios relatados
 - Recomendações ou soluções
 - Metodologia utilizada

RELATO

- Tipo de síntese: narrativa e temática
- Categorias identificadas:
 - Dependência excessiva da IA por parte dos alunos
 - Redução da autonomia docente
 - Barreiras metodológicas e técnicas
 - Preocupações éticas e privacidade de dados
 - Desigualdade no acesso a recursos tecnológicos

- Produto final:
 - Artigo submetido ao SEMISH 2025
 - Seminário interativo com docentes
 - Priorização dos desafios via Mentimeter

LIMITAÇÕES

- Revisão realizada por um único revisor
- Uso de apenas duas bases de dados
- Amostra limitada de professores
- Revisão não exaustiva

APÊNDICE B – GUIDELINE

Guideline - Problem Based Learning Integrated with Artificial Inteligenced

Sumário

- 1. Introdução
 - 1.1 O Novo Contexto da Educação Universitária
 - 1.2 O Desafio Pedagógico
 - 1.2.1 A Necessidade de um Olhar Ético
 - 1.3 Por que o PBL é uma Resposta
 - 1.4 O Papel da GenAI no PBL
 - 1.5 Objetivo do Guideline
- 2. Estrutura do Guideline PBL com GenAI
 - 2.1 Etapas do PBL adaptadas de Walsh
 - 2.2 Níveis de implementação inspirados em Ribeiro
 - 2.3 Escala de aderência pedagógica das ferramentas
- 3. Detalhamento das Etapas do PBL com GenAI
 - Etapa 1 – Identificação do Problema
 - Etapa 2 – Exploração dos Conhecimentos Prévios
 - Etapa 3 – Levantamento de Hipóteses
 - Etapa 4 – Definição das Questões de Estudo
 - Etapa 5 – Estudo Individual
 - Etapa 6 – Retorno e Aplicação do Conhecimento
 - Etapa 7 – Avaliação e Reflexão
- 4. PBL e GenAI na prática: Atuação por Níveis
 - Nível 1 – Problemas curtos em uma aula
 - Nível 2 – Problemas semanais em múltiplas aulas
 - Nível 3 – Problemas de média duração (2 a 4 semanas)
 - Nível 4 – Um grande problema ao longo do semestre
- 5. Ferramentas de GenAI no Apoio às Etapas do PBL
 - Critérios de Aderência: Escala Likert aplicada ao PBL
 - Tabela de referência
 - Formas de acesso às ferramentas: Online e Local
- 6. Aspectos Éticos e Regulatórios
- 7. Considerações Finais

1. Introdução

1. O Novo Contexto da Educação Universitária

O lançamento do **ChatGPT**, em 2022, marcou o início de uma nova era para a educação. Em poucos meses, os *Large Language Models (LLMs)* deixaram de ser um recurso experimental para se tornarem **ferramentas de uso cotidiano**, acessíveis a qualquer aluno com conexão à internet. Esse fenômeno provocou uma verdadeira explosão de serviços de *Generative Artificial Intelligence (GenAI)*, oferecendo desde geração de textos e imagens até análise de dados e simulação de especialistas.

Esse acesso imediato e massivo concedeu ao estudante um poder inédito, a possibilidade de **produzir conteúdo sofisticado em segundos**, sem necessariamente dominar os fundamentos do tema. Por outro lado, gerou também uma mudança abrupta nas dinâmicas de aprendizagem, muitos alunos passaram a **delegar atividades inteiras às GenAI**, abrindo mão da reflexão crítica e da construção do raciocínio próprio.

Nesse processo, os professores universitários se viram diante de um dilema. Por um lado, reconheceram o potencial dessas ferramentas como apoio ao ensino; por outro, sentiram a **dificuldade de adaptar suas práticas** a um contexto em que o aluno já chega à sala de aula com acesso a uma tecnologia extremamente poderosa, como aconteceu com a disseminação dos celulares. Essa dificuldade se intensifica pelo fato de que, ao contrário de professores da educação básica, **a maioria dos docentes do ensino superior não recebeu formação pedagógica estruturada**, o que torna mais desafiadora a incorporação de metodologias ativas e o uso intencional de novas tecnologias em sua prática cotidiana.

2. O Desafio Pedagógico

Diante desse cenário, o desafio central não é decidir entre usar ou não usar GenAI em sala de aula, mas sim compreender como utilizá-la de forma pedagógica e intencional. A simples presença dessas ferramentas já é uma realidade inevitável; proibir seu uso seria tão ineficaz quanto tentar impedir os alunos de pesquisarem na internet.

O risco está em transformar a GenAI em um atalho para respostas prontas, esvaziando o processo de aprendizagem. Quando estudantes terceirizam integralmente suas tarefas, deixam de exercitar habilidades fundamentais como análise crítica, formulação de hipóteses, argumentação estruturada e colaboração. Essas competências são justamente as que o ensino superior deve fortalecer, pois são elas que diferenciam um profissional capaz de resolver problemas complexos de alguém que apenas repete soluções automáticas.

Para os professores, o desafio aparece em duas dimensões. A primeira é pedagógica: compreender em quais momentos a GenAI realmente enriquece a aprendizagem, como ao ajudar na criação de cenários de estudo ou na sugestão de hipóteses, e em quais momentos ela pode enfraquecê-la, oferecendo respostas automáticas que reduzem a reflexão dos alunos. A segunda é prática: encontrar formas de integrar a GenAI ao dia a dia sem transformar isso em mais uma sobrecarga de trabalho ou em um recurso dispersivo.

Sem uma orientação clara, há o risco de que a GenAI seja usada de maneira improvisada, criando confusão e insegurança em sala de aula. Por isso, torna-se urgente a criação de referenciais claros que ajudem o professor universitário a transformar a GenAI em um recurso pedagógico confiável, fortalecendo a autonomia dos alunos e garantindo que a aprendizagem se mantenha crítica, profunda e significativa.

2.1 A Necessidade de um Olhar Ético

O entusiasmo em torno da Inteligência Artificial (IA) não pode obscurecer os dilemas éticos que surgem quando a tecnologia entra na sala de aula. Professores e estudantes lidam diariamente com informações

peçoais, opiniões, produções escritas e registros de interação. Ao serem processados por ferramentas de IA, especialmente aquelas baseadas em serviços online, esses dados passam a circular em ambientes que fogem ao controle da instituição e do docente.

Isso significa que a integração da GenAI ao ensino superior precisa ser acompanhada por uma preocupação constante com **privacidade, consentimento e transparência**. Não se trata apenas de cumprir exigências legais, como a **Lei Geral de Proteção de Dados (LGPD)**, mas de cultivar uma cultura de responsabilidade acadêmica. Professores devem deixar claro quando e como a GenAI está sendo usada, informar os estudantes sobre os riscos e, sempre que possível, preferir soluções que mantenham maior controle sobre os dados.

Ao longo deste guideline, a dimensão ética reaparece em diferentes momentos. E, para consolidar esse aspecto, dedicamos uma seção específica ao final do documento, em que discutimos com mais profundidade os cuidados necessários para que o uso da GenAI no PBL seja não apenas criativo e inovador, mas também seguro e responsável.

3. Por que o PBL é uma Resposta

Existem diversas metodologias ativas que poderiam servir como base para a construção de um modelo pedagógico integrado à GenAI. Todas elas compartilham um princípio em comum: colocam o estudante em posição de maior protagonismo no processo de aprendizagem, estimulando-o a investigar, colaborar e aplicar o conhecimento de forma prática.

A escolha do **Problem Based Learning (PBL)** neste guideline não significa que seja a única alternativa possível, mas se justifica pelo papel central que o pensamento crítico ocupa em sua estrutura*. O PBL coloca os alunos diante de problemas complexos e abertos, em que não há uma única resposta correta. Nesses contextos, o mais importante não é o resultado final, mas o caminho investigativo ao levantar hipóteses, confrontar ideias, justificar decisões e sustentar raciocínios de forma consistente.

Essa característica se conecta de forma direta aos riscos trazidos pelo uso acrítico da GenAI. Enquanto modelos de linguagem oferecem soluções imediatas, o PBL valoriza o processo reflexivo que leva à construção do conhecimento. Assim, em vez de competir com a GenAI, o PBL cria espaço para que ela seja utilizada como recurso de apoio, sem retirar do estudante a responsabilidade pela análise crítica.

O professor, nesse cenário, atua como facilitador, desafiando hipóteses, instigando perguntas e garantindo que a investigação mantenha foco pedagógico. A GenAI pode complementar esse processo ao gerar exemplos de problemas, trazer múltiplas perspectivas ou fornecer materiais adicionais, mas sempre em função de uma experiência de aprendizagem que tem o raciocínio crítico como eixo central.

Dessa forma, o PBL se apresenta como uma resposta consistente ao desafio atual, embora não exclusiva. Outras metodologias ativas, como a Aprendizagem Baseada em Projetos (ABPj) ou a Sala de Aula Invertida, também podem oferecer caminhos valiosos para integrar GenAI de maneira pedagógica. O ponto essencial é que a GenAI deve estar a serviço de metodologias que valorizem a investigação, a autonomia e o pensamento crítico dos estudantes.

4. O Papel da GenAI no PBL

A integração da GenAI ao PBL deve ampliar o processo investigativo sem substituir o raciocínio humano. Na **identificação do problema**, a GenAI enriquece o contexto inicial ao propor variações e dados complementares, sem antecipar soluções. Durante a **exploração do conhecimento prévio**, ajuda a organizar

rapidamente ideias e a sugerir conexões e perguntas diagnósticas que orientam a conversa. Já na **geração de hipóteses**, oferece alternativas e contraexemplos plausíveis, cabendo ao grupo analisar a consistência de cada caminho.

Na **definição das questões de estudo**, a GenAI apoia o refinamento de perguntas, tornando-as investigáveis, delimitadas e alinhadas aos objetivos de aprendizagem. No **estudo individual**, atua como tutor auxiliar ao explicar conceitos e resumir materiais, sempre com validação de fontes e critérios de qualidade para evitar dependência acrítica. Na **aplicação do conhecimento**, contribui para organizar achados, comparar evidências e esboçar protótipos ou simulações, enquanto a síntese e a justificativa permanecem responsabilidade do grupo. Por fim, na **avaliação e reflexão**, a GenAI pode apoiar na organização das evidências de aprendizagem e no registro do percurso dos alunos, facilitando a visualização do progresso. Ainda assim, a interpretação dos resultados e a condução das reflexões permanecem responsabilidades do professor e dos próprios estudantes.

5. Objetivo do Guideline

Este guia tem como propósito oferecer uma referência prática para a integração da metodologia de ensino PBL com recursos de GenAI no ensino superior. A intenção não é substituir o papel do professor ou a autonomia do estudante, mas demonstrar de que forma a GenAI pode enriquecer cada etapa do processo, fortalecendo tanto a mediação docente quanto a participação ativa discente.

A proposta é que este material sirva como um ponto de partida para docentes que desejam explorar novas estratégias pedagógicas sem abrir mão do rigor acadêmico. Ao longo das seções seguintes, são apresentados exemplos, recomendações e caminhos possíveis para que a tecnologia seja incorporada de maneira crítica, responsável e alinhada aos objetivos formativos de cada curso.

2. Estrutura do Guideline PBL com GenAI

A proposta deste guideline se organiza em **três camadas complementares** que funcionam como um mapa para o professor universitário que deseja integrar o PBL com GenAI. A primeira camada responde à pergunta **“o que fazer?”**, apresentando as sete etapas do ciclo PBL, inspiradas em Walsh (2005) e reinterpretadas para contemplar o uso pedagógico da GenAI. A segunda camada responde à questão **“em que profundidade aplicar?”**, trazendo os níveis de implementação sugeridos por Ribeiro (2021). Por fim, a terceira camada responde **“com quais apoios tecnológicos contar?”**, propondo uma escala de aderência pedagógica que avalia o grau de relevância das ferramentas de GenAI para cada etapa do ciclo, inspirada na taxonomia de AIAfnan et al. (2024).

Essa arquitetura disponibiliza ao professor uma visão **processual** (as etapas do ciclo), **estrutural** (os níveis de implementação) e **tecnológica** (a escala de ferramentas) desse processo de ensino. Nos capítulos seguintes, cada uma dessas dimensões será explorada em maior profundidade, com exemplos, prompts e representações gráficas.

Camada	Pergunta-chave	Referencial	Onde será detalhada
Etapas	O que fazer?	Walsh (2005), adaptado ao uso de GenAI	Capítulo 3

Camada	Pergunta-chave	Referencial	Onde será detalhada
Níveis	Em que profundidade aplicar?	Ribeiro (2021);	Capítulo 4
Ferramentas	Com quais apoios tecnológicos?	AlAfnan et al. (2024), adaptado ao PBL	Capítulo 5

2.1 Etapas do PBL adaptadas de Walsh

O fio condutor deste guideline são as **sete etapas clássicas do ciclo PBL descrito por Walsh (2005)**, reinterpretadas aqui sob a perspectiva da GenAI. Essas etapas não são apresentadas como um modelo universal de PBL, mas como um **arcabouço específico deste documento**, voltado a integrar metodologias ativas com tecnologias emergentes.

Cada etapa será detalhada no Capítulo 3 com quatro elementos: (i) objetivo pedagógico, (ii) relato hipotético de aplicação, (iii) exemplos de prompts e (iv) reflexão docente.

Etapa	Descrição
1. Identificação do problema	O grupo lê e discute o caso, levantando impressões iniciais sem chegar a diagnósticos precipitados.
2. Exploração do conhecimento prévio	Os estudantes compartilham o que já sabem, acessam experiências anteriores e definem termos relevantes.
3. Geração de hipóteses	O grupo propõe hipóteses explicativas, buscando entender os mecanismos do problema.
4. Definição das questões de estudo	São formuladas perguntas focadas que ainda não podem ser respondidas com o conhecimento atual do grupo.
5. Estudo individual	Cada estudante pesquisa de forma autônoma, trazendo novos conhecimentos para a próxima sessão.
6. Aplicação do conhecimento	O grupo se reúne novamente, aplica os conhecimentos adquiridos ao problema e discute como eles mudam a compreensão inicial.
7. Avaliação e reflexão	Revisão do processo, feedback entre colegas e autoavaliação.

2.2 Níveis de implementação inspirados em Ribeiro

Saber as etapas é fundamental, mas não suficiente: também é preciso definir **em que profundidade o ciclo será aplicado dentro da disciplina**. Inspirados em Ribeiro (2021), este guideline organiza a implementação do PBL em quatro níveis progressivos.

Nível	Problema	Duração	Foco pedagógico
1	Problemas curtos, um por aula.	Um encontro	Exercício rápido de ciclo completo.
2	Um problema por semana.	Duas aulas	Aprofundamento moderado.

Nível	Problema	Duração	Foco pedagógico
3	Problemas que duram algumas semanas.	2–4 semanas	Investigação com mini-ciclos sucessivos.
4	Um grande problema ao longo do semestre.	Semestre inteiro	Integração plena com múltiplas dimensões.

Essa escala será detalhada no Capítulo 4, onde discutiremos como cada etapa de Walsh se adapta a esses diferentes níveis, sempre com sugestões de uso da GenAI.

2.3 Escala de aderência pedagógica das ferramentas

Por fim, a terceira camada do guideline diz respeito ao uso da tecnologia. Nem toda ferramenta de GenAI contribui igualmente para a aprendizagem investigativa. Para orientar o professor, propomos uma **escala de cinco níveis de aderência pedagógica**, que avalia a contribuição efetiva das ferramentas em cada etapa do PBL.

Nível	Nome	Critério objetivo
1	Prejudicial	Enfraquece a investigação, fornecendo respostas prontas ou bloqueando o debate crítico.
2	Dispersivo	Não contribui de forma significativa e pode desviar o grupo da análise investigativa.
3	Suporte opcional	Complementa a fase em alguns aspectos, mas sem impacto essencial no processo.
4	Apoio investigativo	Facilita a exploração, a formulação de hipóteses ou a organização coletiva da aprendizagem.
5	Núcleo pedagógico	Potencializa diretamente o raciocínio crítico, a colaboração e a reflexão da etapa.

Essa escala será aprofundada no Capítulo 5, junto com a apresentação da **taxonomia de ferramentas de IA** e a definição de um **kit indispensável para professores**.

3. Detalhamento das Etapas do PBL com GenAI

Para tornar este guideline útil na prática docente, cada uma das sete etapas do PBL é apresentada de forma organizada e previsível. A estrutura foi pensada para equilibrar **clareza pedagógica** e **aplicabilidade imediata**, permitindo que o professor compreenda o propósito de cada momento e visualize como integrá-lo em sala de aula.

- **Objetivo pedagógico** → indica a função da etapa dentro do ciclo do PBL, ajudando o professor a manter o foco formativo.
- **Relato hipotético** → em terceira pessoa, mostra de forma ilustrativa como um professor **poderia aplicar** a etapa com o apoio de ferramentas de GenAI. Esse tom narrativo aproxima o guideline da realidade da sala de aula, sem impor um único caminho.

- **Exemplos de prompts** → apresentam comandos prontos (em inglês, com opção de saída em português), acompanhados de uma tabela analítica que explica a função pedagógica de cada instrução.
- **Reflexão docente** → sugere boas práticas, cuidados e alternativas possíveis, considerando diferentes contextos de ensino e níveis de acesso a tecnologias.

Essa organização busca **oferecer inspiração sem prescrição**: fornece modelos prontos para uso, mas abre espaço para que cada professor adapte as propostas conforme sua disciplina, perfil de turma e recursos disponíveis.

Etapa 1 – Identificação do Problema

Objetivo pedagógico Apresentar um problema inicial de forma clara, realista e instigante, de modo a despertar a curiosidade dos estudantes e levá-los a levantar questões preliminares. O enunciado deve conter ambiguidades suficientes para estimular a investigação, sem antecipar hipóteses ou soluções.

Relato hipotético Um professor de Engenharia de Dados poderia iniciar a experiência do PBL apresentando à turma um problema relacionado a uma fintech. Para preparar o material, ele recorreria a um **chatbot**, solicitando um enunciado em que clientes reclamam de divergências entre os relatórios exibidos no aplicativo e os resultados reais de suas carteiras de investimento.

Ao revisar o texto, o professor perceberia que a linguagem estava excessivamente técnica para estudantes de graduação. Nesse caso, poderia usar uma ferramenta de **text enhancement** para simplificar a redação, tornando-a mais acessível.

Além do texto, ele poderia decidir enriquecer a situação com um elemento visual, utilizando uma ferramenta de **image generation** para gerar uma tela simulada do aplicativo da fintech, mostrando os números divergentes. Esse recurso ajudaria a despertar a curiosidade e dar concretude ao caso.

Durante a discussão em sala, o professor poderia ativar uma ferramenta de **note-taking/transcription**, que transcreveria automaticamente as falas iniciais dos alunos. O resultado seria um registro textual pesquisável, preservando as primeiras impressões para consulta em etapas posteriores.

Exemplos de prompts

Prompt Estruturado (Geração do Cenário em Texto)

```
[PAPEL] Você é um especialista em Problem-Based Learning (PBL) e em [Área de Conhecimento].  
[CONTEXTO] Um professor está preparando uma atividade de PBL para estudantes de [Nível da Turma].  
[TAREFA] Criar um cenário de problema realista em que [Contexto do Problema].  
[RESTRICÇÕES]  
- Não fornecer soluções ou explicações.  
- Manter a descrição acessível ao nível da turma definido.  
[SAÍDA] Um cenário claro e envolvente em formato narrativo.
```

- [Área de Conhecimento] → área de conhecimento (ex.: Engenharia de Dados, Direito, Medicina)
- [Nível da Turma] → nível da turma (ex.: graduação, pós-graduação, curso técnico)

- [Contexto do Problema] → situação a ser simulada (ex.: inconsistências em relatórios financeiros de uma fintech)

Prompt Estruturado (Ajuste de Linguagem)

[PAPEL] Você é uma ferramenta de GenAI especializada em clareza e acessibilidade de texto.

[CONTEXTO] O seguinte enunciado de problema é destinado a estudantes de [Nível da Turma] na área de [Área de Conhecimento].

[TAREFA] Reescrever o texto para garantir clareza e acessibilidade, mantendo todas as ideias originais.

[RESTRIÇÕES]

- Não adicionar explicações ou soluções.
- Preservar a ambiguidade necessária para investigação.

[ENTRADA] [Cole aqui o enunciado original do problema]

[SAÍDA] Uma versão reescrita do enunciado do problema adaptada aos estudantes-alvo.

- [Nível da Turma] → nível da turma (ex.: graduação, pós-graduação, curso técnico)
- [Área de Conhecimento] → área de conhecimento (ex.: Engenharia de Dados, Direito, Medicina)
- [Cole aqui o enunciado original do problema] → enunciado que será simplificado

Prompt Estruturado (Geração de Imagem do Cenário)

[PAPEL] Você é uma ferramenta de GenAI especializada em [Estilo/Tipo de Imagem].

[CONTEXTO] Um professor está criando um recurso visual para apoiar uma sessão de PBL de [Nível da Turma] na área de [Área de Conhecimento].

[TAREFA] Gerar uma ilustração que represente [Tópico/Contexto].

[RESTRIÇÕES]

- A imagem deve apoiar o cenário do problema sem sugerir soluções ou causas.
- Manter o conteúdo realista e relevante para o contexto da sala de aula.
- Estilo: [Estilo/Tipo de Imagem].

[SAÍDA] Fornecer a ilustração gerada com uma legenda curta.

- [Estilo/Tipo de Imagem] → estilo desejado (ex.: painel realista, infográfico, esboço)
- [Nível da Turma] → nível da turma (ex.: graduação, pós-graduação, curso técnico)
- [Área de Conhecimento] → área de conhecimento (ex.: Engenharia de Dados, Direito, Medicina)
- [Tópico/Contexto] → situação do problema (ex.: inconsistências em relatórios de fintech)

Prompt Estruturado (Registro e Transcrição da Discussão)

[PAPEL] Você é um assistente de GenAI especializado em tomada de notas e transcrição.

[CONTEXTO] Um grupo de estudantes de [Nível da Turma] na área de [Área de Conhecimento] está discutindo as impressões iniciais sobre um problema de PBL.

[TAREFA] Transcrever e resumir as notas da discussão fornecidas em [Transcrição de

Entrada].

[RESTRIÇÕES]

- Destacar as impressões iniciais compartilhadas pelos estudantes.
- Extrair as principais perguntas levantadas sem alterar seu significado.
- Não adicionar interpretações ou respostas.

[ENTRADA] [Cole a transcrição ou notas da discussão]

[SAÍDA] Fornecer um resumo claro e estruturado das impressões e perguntas.

- [Nível da Turma] → nível da turma (ex.: graduação, pós-graduação, curso técnico)
- [Área de Conhecimento] → área de conhecimento (ex.: Engenharia de Dados, Administração, Medicina)
- [Transcrição de Entrada] → transcrição bruta ou notas da discussão

Reflexão docente Nesta primeira etapa, a função da GenAI é **amparar o professor na elaboração do problema e facilitar o registro das discussões iniciais**. É fundamental que o enunciado preserve pontos ambíguos, pois são essas lacunas que estimulam os alunos a questionar e investigar. A tentação de “embelezar demais” com imagens ou descrições excessivamente detalhadas deve ser evitada: o objetivo não é dar respostas, mas **abrir espaço para perguntas**. Em contextos com menor acesso tecnológico, um simples slide com dados fictícios ou uma narrativa oral já cumpre o mesmo papel.

Etapa 2 – Exploração dos Conhecimentos Prévios

Objetivo pedagógico Estimular os alunos a compartilhar e organizar coletivamente o que já sabem sobre o problema apresentado, identificando conceitos prévios, experiências relacionadas e possíveis lacunas de compreensão. Essa etapa é fundamental para construir uma base de partida comum, antes de avançar para hipóteses ou estudo individual.

Relato hipotético Após a apresentação do problema, o professor propõe que os alunos compartilhem o que já sabem sobre auditoria e integridade de dados. Essa etapa deve priorizar a troca direta entre os estudantes, valorizando a construção coletiva.

Para apoiar a atividade, o professor pode ativar uma ferramenta de **note-taking/transcription** que registra em tempo real tudo o que é dito. A transcrição gera um documento pesquisável, garantindo que nenhuma contribuição seja perdida.

Além disso, um **chatbot** pode ser usado de forma complementar, gerando um glossário com os principais conceitos que emergem da fala dos estudantes. Esse recurso funciona como apoio documental e ajuda a sistematizar o vocabulário da turma, tornando mais fácil a identificação de lacunas de conhecimento.

Exemplos de prompts

Prompt Estruturado (Tomada de notas/Transcrição – Registrar conhecimentos prévios)

[PAPEL] Você é um assistente de GenAI especializado em transcrição.

[CONTEXTO] Um grupo de estudantes de [Nível da Turma] na área de [Área de Conhecimento] está discutindo seus conhecimentos prévios sobre um problema de PBL.

[TAREFA] Transcrever a discussão em texto simples e pesquisável.

[RESTRIÇÕES]

- Registrar fielmente a fala sem resumir ou interpretar.
- Não adicionar explicações ou respostas.

[ENTRADA] [Arquivo de áudio ou transcrição bruta]
[SAÍDA] Uma transcrição em texto da discussão, preservando todas as contribuições.

- [Nível da Turma] → nível da turma
- [Área de Conhecimento] → área do conhecimento
- [Arquivo de áudio ou transcrição bruta] → gravação ou transcrição bruta

Prompt Estruturado (Chatbot – Glossário de termos relevantes)

[PAPEL] Você é um assistente de GenAI especializado em criação de glossários acadêmicos.
[CONTEXTO] Aqui está a transcrição dos conhecimentos prévios dos estudantes: [Cole a transcrição].
[TAREFA] Extrair os principais conceitos e gerar um glossário com definições curtas e claras.
[RESTRICÇÕES]
- Manter definições concisas e adequadas para estudantes de [Nível da Turma].
- Adicionar alguns termos relacionados que possam surgir naturalmente em discussões futuras.
[SAÍDA] Um glossário de termos-chave com definições.

- [Cole a transcrição] → transcrição registrada na discussão
- [Nível da Turma] → ex.: graduação, pós-graduação

Reflexão docente Nesta etapa, a GenAI atua como **apoio organizacional**, mas não deve substituir a interação entre os estudantes. O registro das falas e a geração de glossários ajudam a tornar visível o repertório inicial da turma, porém o mais importante é que os alunos **construam juntos** a base de conhecimentos e identifiquem coletivamente suas lacunas.

Etapa 3 – Levantamento de Hipóteses

Objetivo pedagógico Estimular os alunos a propor explicações iniciais e possíveis causas para o problema apresentado, construindo hipóteses que serão investigadas nas etapas seguintes. O foco é desenvolver a capacidade de questionamento, argumentação e raciocínio lógico, sem exigir respostas definitivas neste momento.

Relato hipotético Com base nas discussões iniciais, os alunos levantam hipóteses sobre as possíveis causas do problema: falhas em processos de ETL, inconsistências na integração de dados ou erros de cálculo nos relatórios.

Para tornar essas hipóteses mais concretas, o professor recorre a uma ferramenta de **image generation**, que cria diagramas ou fluxogramas representando visualmente os pontos levantados pela turma.

Enquanto a conversa ocorre, uma ferramenta de **note-taking/transcription** capta automaticamente todas as falas, transformando-as em texto pesquisável. Esse registro fiel garante que nenhuma hipótese formulada pelo grupo se perca no processo.

Exemplos de prompts

Prompt Estruturado (Geração de Imagem – Representação visual de hipóteses)

[PAPEL] Você é uma ferramenta de GenAI especializada em [Estilo/Tipo de Imagem].
[CONTEXTO] Um professor está apoiando uma sessão de PBL na área de [Área de Conhecimento] com estudantes de [Nível da Turma].
[TAREFA] Gerar um diagrama que represente as hipóteses propostas pelos estudantes sobre [Tópico/Contexto].
[RESTRICÇÕES]
- Representar o fluxo de dados e destacar pontos potenciais onde inconsistências podem ocorrer.
- Não fornecer explicações ou soluções.
- Manter a imagem simples, clara e focada na visualização.
- Estilo: [Estilo/Tipo de Imagem].
[SAÍDA] Uma representação visual das hipóteses com uma legenda curta.

- [Estilo/Tipo de Imagem] → ex.: fluxograma, esboço esquemático, infográfico
- [Área de Conhecimento] → ex.: Engenharia de Dados, Medicina
- [Nível da Turma] → ex.: graduação, pós-graduação
- [Tópico/Contexto] → ex.: inconsistências em relatórios financeiros de uma fintech

Prompt Estruturado (Tomada de notas/Transcrição – Transcrição de hipóteses)

[PAPEL] Você é um assistente de GenAI especializado em transcrição.
[CONTEXTO] Um grupo de estudantes de [Nível da Turma] na área de [Área de Conhecimento] está discutindo hipóteses para um problema de PBL.
[TAREFA] Transcrever a discussão em texto simples e pesquisável.
[RESTRICÇÕES]
- Capturar as hipóteses exatamente como foram faladas, sem interpretações ou avaliações.
- Não agrupar, categorizar ou reestruturar o conteúdo.
[ENTRADA] [Arquivo de áudio ou transcrição bruta]
[SAÍDA] Uma transcrição em texto bruto das hipóteses compartilhadas pelos estudantes.

- [Nível da Turma] → nível da turma (ex.: graduação, pós-graduação)
- [Área de Conhecimento] → área do conhecimento (ex.: Engenharia de Dados, Direito)
- [Arquivo de áudio ou transcrição bruta] → entrada de áudio ou transcrição bruta

Reflexão docente O levantamento de hipóteses pode gerar insegurança em alguns estudantes, que temem “errar”. Aqui, a GenAI deve ser vista apenas como apoio para **organizar e visualizar** as ideias, mas não para propor hipóteses no lugar dos alunos. O professor deve valorizar a diversidade de hipóteses, mesmo as improváveis, pois elas alimentam o processo investigativo. Diagramas e registros podem enriquecer a compreensão, mas não substituem a argumentação construída coletivamente.

Etapa 4 – Definição das Questões de Estudo

Objetivo pedagógico Transformar as hipóteses levantadas em **questões de investigação claras e viáveis**, que irão orientar os estudos individuais e coletivos. Essa fase ajuda os alunos a desenvolver a habilidade de formular perguntas relevantes, delimitando o escopo da investigação.

Relato hipotético O grupo precisa transformar as hipóteses em perguntas investigativas claras e viáveis. Para apoiar esse processo, o professor pode usar um **chatbot**, pedindo sugestões de formulações de perguntas abertas e instigantes que ajudem a guiar a investigação sem fornecer respostas prontas.

As perguntas elaboradas pelos alunos podem, em seguida, ser revisadas por meio de uma ferramenta de **text enhancement**, que refina a redação para deixá-las mais objetivas, acessíveis e bem estruturadas.

Enquanto isso, uma ferramenta de **note-taking/transcription** registra automaticamente a discussão da turma, gerando um texto pesquisável que documenta o raciocínio coletivo e assegura que nenhuma sugestão seja esquecida.

Exemplos de prompts

Prompt Estruturado (Chatbot – Transformar hipóteses em perguntas)

```
[PAPEL] Você é um assistente de GenAI especializado em criação de perguntas para PBL.  
[CONTEXTO] Estudantes de [Nível da Turma] estão trabalhando na área de [Área de Conhecimento].  
[ENTRADA] Hipóteses da Etapa 3: [Cole as hipóteses].  
[TAREFA] Transformar cada hipótese em uma ou mais perguntas de pesquisa abertas e claras.  
[RESTRICÇÕES]  
- As perguntas devem ser investigáveis (não podem ser sim/não).  
- Manter a linguagem acessível ao nível da turma definido.  
[SAÍDA] Uma lista de perguntas de pesquisa derivadas das hipóteses.
```

- **[Nível da Turma]** → nível da turma (ex.: graduação, pós-graduação)
- **[Área de Conhecimento]** → área de conhecimento (ex.: Engenharia de Dados, Medicina, Direito)
- **[Cole as hipóteses]** → hipóteses brutas registradas na Etapa 3

Prompt Estruturado (Melhoria de Texto – Refinar perguntas de pesquisa)

```
[PAPEL] Você é uma ferramenta de GenAI especializada em clareza e acessibilidade de texto.  
[CONTEXTO] As seguintes perguntas foram geradas por estudantes em uma sessão de PBL para [Nível da Turma] na área de [Área de Conhecimento].  
[TAREFA] Refinar cada pergunta para maior clareza, precisão e foco, sem alterar seu significado.  
[RESTRICÇÕES]  
- Não adicionar respostas ou explicações.  
- Preservar o caráter aberto das perguntas.  
[ENTRADA] [Cole aqui as perguntas geradas pelos estudantes]  
[SAÍDA] Uma lista revisada de perguntas de pesquisa claras, concisas e acessíveis.
```


- **[Nível da Turma]** → nível da turma (ex.: graduação, pós-graduação)
- **[Área de Conhecimento]** → área de conhecimento (ex.: Engenharia de Dados, Administração, Medicina)
- **[Cole aqui as perguntas geradas pelos estudantes]** → perguntas originais criadas pelos estudantes

Prompt Estruturado (Registro e Transcrição – Registrar perguntas de estudo)

```
[PAPEL] Você é um assistente de GenAI especializado em transcrição.  
[CONTEXTO] Um grupo de estudantes de [Nível da Turma] na área de [Área de  
Conhecimento] está definindo perguntas de estudo durante uma sessão de PBL.  
[TAREFA] Transcrever a discussão bruta em um texto simples e pesquisável.  
[RESTRIÇÕES]  
- Capturar exatamente o que os estudantes dizem ao propor perguntas.  
- Não reformular, resumir ou interpretar.  
[ENTRADA] [Arquivo de áudio ou transcrição bruta]  
[SAÍDA] Uma transcrição textual das perguntas de estudo conforme expressas pelos  
estudantes.
```

- **[Nível da Turma]** → nível da turma (ex.: graduação, pós-graduação)
- **[Área de Conhecimento]** → área de conhecimento (ex.: Engenharia de Dados, Direito, Medicina)
- **[Arquivo de áudio ou transcrição bruta]** → áudio original ou notas transcritas da discussão

Reflexão docente Esta é uma das etapas em que a mediação do professor é mais necessária: é comum os alunos formularem perguntas muito amplas ("Por que o sistema não funciona?") ou muito específicas ("O bug está na tabela X?"). O **Chatbot** pode apoiar com exemplos de boas perguntas, mas a formulação final deve vir dos alunos. O **Text Enhancement** garante clareza e foco, e o **Note-taking/Transcription** assegura que todas as contribuições sejam preservadas.

Etapa 5 – Estudo Individual

Objetivo pedagógico Estimular cada estudante a aprofundar-se nas questões de estudo definidas pelo grupo, pesquisando fontes, analisando informações e construindo compreensões próprias. Essa etapa fortalece a autonomia e garante diversidade de perspectivas na etapa seguinte.

Relato hipotético Cada aluno escolhe uma questão para aprofundar. Nesse momento, ferramentas de **conduct research** podem ser usadas para localizar fontes confiáveis, artigos acadêmicos e relatórios técnicos relacionados à investigação.

Os resumos produzidos individualmente são melhorados com **text enhancement**, que ajuda os alunos a refinar clareza, gramática e coesão textual.

Se desejarem, os alunos podem também ativar uma solução de **note-taking/transcription** para registrar suas próprias reflexões verbais enquanto estudam. O resultado é um arquivo textual pesquisável, que organiza suas falas em forma de notas pessoais para consulta posterior.

Exemplos de prompts

Prompt Estruturado (Conduzir Pesquisa – Resumir artigo ou documento)

[PAPEL] Você é uma ferramenta de GenAI especializada em suporte à pesquisa acadêmica.

[CONTEXTO] Estudantes de [Nível da Turma] estão investigando questões de estudo de uma sessão de PBL na área de [Área de Conhecimento].

[TAREFA] Resumir o seguinte documento: [Documento de Entrada].

[RESTRIÇÕES]

- Destacar apenas as partes mais relevantes para as questões de estudo da Etapa 4.
- Manter o resumo conciso, didático e adaptado ao nível acadêmico.

[ENTRADA] [Cole ou envie o documento]

[SAÍDA] Um resumo curto e focado, conectado às questões de estudo.

- [Documento de Entrada] → artigo, relatório ou material técnico
- [Nível da Turma] → ex.: graduação, pós-graduação
- [Área de Conhecimento] → ex.: Engenharia de Dados, Direito

Prompt Estruturado (Aprimoramento de Texto – Melhorar anotações dos estudantes)

[PAPEL] Você é uma ferramenta de GenAI especializada em clareza textual e escrita acadêmica.

[CONTEXTO] Estas são anotações escritas por um estudante de [Nível da Turma] na área de [Área de Conhecimento].

[TAREFA] Melhorar as anotações para torná-las mais claras, concisas e organizadas.

[RESTRIÇÕES]

- Não remover nenhuma ideia importante.
- Manter um tom adequado ao trabalho acadêmico.

[ENTRADA] [Cole aqui as anotações do estudante]

[SAÍDA] Uma versão mais clara e bem estruturada das anotações.

- [Cole aqui as anotações do estudante] → anotações originais do aluno
- [Nível da Turma] → nível da turma
- [Área de Conhecimento] → área de conhecimento

Prompt Estruturado (Tomada de Notas/Transcrição – Registrar reflexões pessoais)

[PAPEL] Você é um assistente de GenAI especializado em transcrição.

[CONTEXTO] Um estudante de [Nível da Turma] está registrando reflexões verbais enquanto estuda uma questão de PBL na área de [Área de Conhecimento].

[TAREFA] Transcrever o áudio bruto em texto simples e pesquisável.

[RESTRIÇÕES]

- Capturar a fala fielmente, sem resumos ou interpretações.
- Não reestruturar ou filtrar o conteúdo.

[ENTRADA] [Arquivo de áudio ou gravação ao vivo]

[SAÍDA] Uma transcrição em texto bruto das reflexões do estudante.

- [Arquivo de áudio ou gravação ao vivo] → gravação da fala do estudante

- [Nível da Turma] → nível da turma
- [Área de Conhecimento] → área de conhecimento

Reflexão docente O estudo individual é a etapa em que a GenAI pode ser especialmente útil, mas também a que mais exige vigilância crítica. Há o risco de o aluno simplesmente aceitar o que a ferramenta apresenta sem buscar confirmação em fontes confiáveis. O professor deve reforçar que a GenAI funciona como apoio, e não como substituto da pesquisa. Validar informações, cruzar referências e manter registros organizados são práticas essenciais para que a etapa resulte em conhecimento sólido e investigativo.

Etapa 6 – Retorno e Aplicação do Conhecimento

Objetivo pedagógico Promover a socialização do que foi estudado individualmente, integrando diferentes perspectivas para resolver o problema inicial. Essa etapa reforça a colaboração, o pensamento crítico e a capacidade de síntese coletiva.

Relato hipotético Ao retornarem ao grupo, os alunos integram suas descobertas em uma proposta coletiva de solução. Para estruturar o material, recorrem a uma ferramenta de **content creation**, que os ajuda a transformar os achados em um relatório ou documento final.

O texto é revisado com **text enhancement**, garantindo clareza e consistência. Elementos visuais são gerados por **image generation**, enriquecendo a apresentação com diagramas e gráficos. Por fim, uma ferramenta de **presentation slides** transforma o conteúdo em uma apresentação organizada e pronta para ser compartilhada.

Durante todo esse processo, uma ferramenta de **note-taking/transcription** permanece ativa, registrando automaticamente as discussões do grupo. Esse registro textual documenta como as decisões foram tomadas e como a solução foi construída, funcionando como memória coletiva da etapa.

Exemplos de prompts

Prompt Estruturado (Content Creation – relatório do grupo)

```
[PAPEL] Você é um assistente de GenAI especializado em escrita acadêmica.
[CONTEXTO] Estudantes de [Nível da Turma] compartilharam achados individuais em
[Área de Conhecimento].
[TAREFA] Integrar essas anotações em um relatório coerente que resuma a solução do
grupo para o problema de PBL.
[RESTRICÇÕES]
- Destacar concordâncias e divergências.
- Manter a estrutura clara, concisa e adequada para discussão em sala de aula.
[ENTRADA] [Cole aqui as anotações e resumos dos estudantes]
[SAÍDA] Um relatório estruturado do grupo.
```

- [Nível da Turma] → ex.: graduação, pós-graduação
- [Área de Conhecimento] → ex.: Engenharia de Dados, Medicina
- [Cole aqui as anotações e resumos dos estudantes] → insumos brutos dos alunos

Prompt Estruturado (Text Enhancement – Refinar relatório coletivo)

[PAPEL] Você é uma ferramenta de GenAI especializada em clareza e coerência textual.

[CONTEXTO] O rascunho a seguir foi escrito por estudantes de [Nível da Turma] durante uma sessão de PBL em [Área de Conhecimento].

[TAREFA] Revisar o texto para garantir clareza, fluxo lógico e estilo consistente.

[RESTRIÇÕES]

- Não adicionar novas informações ou interpretações.
- Manter as ideias e a estrutura originais.

[ENTRADA] [Cole aqui o rascunho do relatório]

[SAÍDA] Uma versão revisada e aprimorada do relatório.

- **[Nível da Turma]** → ex.: graduação, pós-graduação
- **[Área de Conhecimento]** → área de conhecimento
- **[Cole aqui o rascunho do relatório]** → versão preliminar criada pelos estudantes

Prompt Estruturado (Geração de Imagem – representação visual)

[PAPEL] Você é uma ferramenta de GenAI especializada em criação de diagramas visuais.

[CONTEXTO] Estudantes de [Nível da Turma] estão apresentando seus achados coletivos em [Área de Conhecimento].

[TAREFA] Gerar um diagrama que represente a solução sintetizada do grupo.

[RESTRIÇÕES]

- Basear o diagrama nos achados da Etapa 5 (estudo individual).
- Mostrar relações entre ideias sem sugerir novas causas ou soluções.
- Estilo: [Estilo/Tipo de Imagem].

[ENTRADA] [Cole aqui os achados sintetizados]

[SAÍDA] Um diagrama ou representação visual para apoiar a apresentação.

- **[Nível da Turma]** → ex.: graduação, pós-graduação
- **[Área de Conhecimento]** → área de conhecimento
- **[Estilo/Tipo de Imagem]** → ex.: diagrama, infográfico, fluxograma
- **[Cole aqui os achados sintetizados]** → síntese final do grupo

Prompt Estruturado (Apresentação – Minuta de slides)

[PAPEL] Você é uma ferramenta de GenAI especializada em criar apresentações educacionais.

[CONTEXTO] Estudantes de [Nível da Turma] estão se preparando para apresentar seus achados em [Área de Conhecimento].

[TAREFA] Criar uma apresentação preliminar com [Número de Slides] slides resumindo os principais achados do grupo.

[RESTRIÇÕES]

- Cada slide deve abordar uma ideia ou insight principal.
- Manter o conteúdo conciso, visualmente estruturado e adequado para apresentação em sala de aula.

[ENTRADA] [Cole aqui os achados sintetizados do grupo]
[SAÍDA] Uma minuta de apresentação estruturada em slides.

- [Nível da Turma] → ex.: graduação, pós-graduação
- [Área de Conhecimento] → área de conhecimento
- [Número de Slides] → quantidade desejada de slides (ex.: 5, 10)
- [Cole aqui os achados sintetizados do grupo] → conteúdo a ser apresentado

Prompt Estruturado (Registro/Transcrição – registrar discussão do grupo)

[PAPEL] Você é um assistente de GenAI especializado em transcrição.
[CONTEXTO] Um grupo de estudantes de [Nível da Turma] em [Área de Conhecimento] está discutindo seus achados coletivos na Etapa 6 de uma sessão de PBL.
[TAREFA] Transcrever a discussão bruta em texto simples e pesquisável.
[RESTRIÇÕES]
- Capturar as contribuições fielmente, sem reformular ou resumir.
- Não interpretar ou filtrar o conteúdo.
[ENTRADA] [Arquivo de áudio ou transcrição ao vivo]
[SAÍDA] Uma transcrição textual bruta da discussão do grupo.

- [Nível da Turma] → ex.: graduação, pós-graduação
- [Área de Conhecimento] → área de conhecimento
- [Arquivo de áudio ou transcrição ao vivo] → insumo de gravação ou texto bruto

Reflexão docente Essa é a etapa em que o professor mais deve enfatizar a **integração coletiva**. A GenAI pode ser um excelente apoio para organizar informações, gerar sínteses e até criar recursos visuais e apresentações, mas a curadoria do grupo e a mediação do docente são indispensáveis. O risco está em os alunos delegarem a síntese diretamente à GenAI sem discussão prévia. É importante reforçar que a ferramenta deve ajudar a **sistematizar o que o grupo construiu**, e não substituir o diálogo entre os estudantes.

Etapa 7 – Avaliação e Reflexão

Objetivo pedagógico Consolidar a aprendizagem por meio da avaliação do processo e da reflexão crítica dos alunos. Essa etapa garante que o grupo não apenas encontre respostas para o problema, mas também compreenda como construiu o conhecimento, quais competências desenvolveu e quais aspectos podem ser aprimorados.

Relato hipotético Na etapa final, os alunos compartilham reflexões sobre o que aprenderam, avaliando tanto o processo quanto os resultados. O professor pode recorrer a um **chatbot** para propor perguntas reflexivas que estimulem os estudantes a analisar não só as soluções alcançadas, mas também o caminho percorrido.

Cada estudante pode usar **text enhancement** para revisar seus textos reflexivos, garantindo clareza, coesão e tom acadêmico.

Enquanto a turma discute em voz alta suas percepções, uma ferramenta de **note-taking/transcription** transcreve as falas em tempo real, gerando um registro pesquisável da avaliação coletiva. Esse material serve

como fonte de feedback para os alunos e como subsídio para o professor melhorar futuras práticas de PBL apoiadas por GenAI.

Exemplos de prompts

Prompt Estruturado (Chatbot – Gerar perguntas reflexivas)

[PAPEL] Você é um assistente de GenAI especializado em aprendizagem reflexiva.
[CONTEXTO] Estudantes de [Nível da Turma] completaram um ciclo de PBL em [Área de Conhecimento].
[TAREFA] Gerar [Número de Perguntas] perguntas abertas de reflexão para orientar a avaliação dos estudantes.
[RESTRICÇÕES]
- As perguntas devem abordar resultados de aprendizagem, colaboração e melhoria do processo.
- Evitar perguntas de sim/não.
[SAÍDA] Uma lista de perguntas reflexivas.

- **[Número de Perguntas]** → quantidade desejada (ex.: 5)
- **[Nível da Turma]** → nível acadêmico (ex.: graduação, pós-graduação)
- **[Área de Conhecimento]** → área de conhecimento (ex.: Engenharia de Dados, Direito)

Prompt Estruturado (Aprimoramento de Texto – Melhorar escrita reflexiva)

[PAPEL] Você é uma ferramenta de GenAI especializada em clareza textual e escrita acadêmica.
[CONTEXTO] O texto reflexivo a seguir foi escrito por um estudante de [Nível da Turma] na área de [Área de Conhecimento].
[TAREFA] Melhorar a clareza, coerência e tom acadêmico mantendo todas as ideias originais.
[RESTRICÇÕES]
- Não remover ou adicionar novas ideias.
- Manter o texto conciso e com estilo reflexivo.
[ENTRADA] [Cole o texto reflexivo do estudante]
[SAÍDA] Um texto reflexivo revisado, adequado para trabalhos acadêmicos.

- **[Nível da Turma]** → nível acadêmico (ex.: graduação, pós-graduação)
- **[Área de Conhecimento]** → área de conhecimento (ex.: Engenharia de Dados, Direito)
- **[Cole o texto reflexivo do estudante]** → texto original produzido pelo aluno

Prompt Estruturado (Tomada de Notas/Transcrição – Comparar discussões iniciais e finais)

[PAPEL] Você é um assistente de GenAI especializado em análise de transcrições.
[CONTEXTO] Uma turma de estudantes de [Nível da Turma] na área de [Área de Conhecimento] discutiu o problema de PBL no Passo 1/2 e agora está refletindo no Passo 7.

[TAREFA] Comparar as transcrições, destacando mudanças de entendimento, evolução das perspectivas e evidências de crescimento colaborativo.

[RESTRIÇÕES]

- Não adicionar interpretações além do que está explicitamente nas transcrições.
- Manter a comparação estruturada e concisa.

[ENTRADA] [Cole as transcrições do Passo 1/2 e do Passo 7]

[SAÍDA] Uma comparação lado a lado mostrando crescimento e reflexão.

- **[Nível da Turma]** → nível acadêmico (ex.: graduação, pós-graduação)
- **[Área de Conhecimento]** → área de conhecimento (ex.: Engenharia de Dados, Administração)
- **[Cole as transcrições do Passo 1/2 e do Passo 7]** → transcrições brutas das discussões iniciais e finais

Reflexão docente A avaliação no PBL deve ser **formativa e reflexiva**, não apenas somativa. O **Chatbot** pode apoiar sugerindo perguntas que estimulem análise crítica, enquanto o **Text Enhancement** ajuda a refinar reflexões escritas e o **Note-taking/Transcription** assegura um registro fiel da discussão. O papel central continua sendo dos alunos: refletir sobre **o que aprenderam, como aprenderam e como podem evoluir** em futuras investigações.

4. PBL e GenAI na prática: Atuação por Níveis

Até aqui, vimos como o PBL pode ser detalhado etapa por etapa (Capítulo 3) e como existem diferentes níveis de implementação (Capítulo 2). Neste capítulo, unimos essas duas perspectivas para mostrar, de forma prática, **como o professor pode conduzir cada etapa das sete fases do PBL em diferentes níveis de aplicação**.

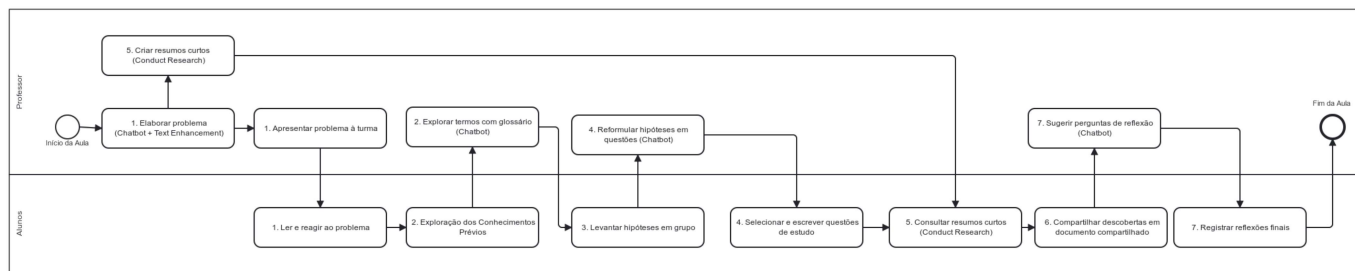
A proposta é simples:

- Para cada **nível de implementação**, descrevemos o que o **professor faz**, o que os **alunos fazem** e como a **GenAI pode auxiliar opcionalmente** em cada uma das sete etapas.
- O tempo de cada atividade também é indicado de forma aproximada, sem determinar a duração da aula, mas ajudando o professor a organizar o ritmo.
- O uso da GenAI nunca é obrigatório: o professor decide quando e como ela pode apoiar, sempre em função do objetivo pedagógico.

Nível 1 – Problemas curtos em uma aula

Contexto geral

O Nível 1 é voltado a professores que querem experimentar o PBL com GenAI em encontros únicos. Todo o ciclo das sete etapas acontece em uma única aula, geralmente de 60 a 120 minutos. O foco não é resolver o problema em profundidade, mas estimular a curiosidade investigativa, a colaboração e a reflexão rápida. Como o tempo é reduzido, a GenAI aparece como um recurso para economizar minutos preciosos, seja simplificando textos, criando materiais de apoio ou sugerindo perguntas de fechamento.



Roteiro de aula (referência: 90 minutos)

Na **Etapa 1 – Identificação do Problema** (cerca de 10 minutos), o professor elabora o enunciado com auxílio de um *chatbot* para gerar o cenário e de uma ferramenta de *text enhancement* para ajustar a linguagem. Ainda nesta etapa, pode-se usar uma ferramenta de *conduct research* para criar **resumos curtos de apoio**, que só serão consultados posteriormente pelos alunos. Depois, o problema é apresentado à turma, que lê e reage ao enunciado.

Na **Etapa 2 – Exploração dos Conhecimentos Prévios** (15 minutos), os alunos discutem em pequenos grupos o que já sabem sobre o tema. O professor pode apoiar essa fase com um *chatbot*, pedindo um glossário rápido de termos relacionados ao problema. Isso evita longas exposições iniciais e mantém o tempo da aula focado na troca entre os estudantes.

Na **Etapa 3 – Levantamento de Hipóteses** (15 minutos), os grupos propõem explicações iniciais para o problema. O professor conduz a discussão valorizando todas as contribuições, sem necessidade de validação nesse momento. O importante é registrar hipóteses diversas como potenciais caminhos investigativos.

Na **Etapa 4 – Definição das Questões de Estudo** (10 minutos), o professor usa um *chatbot* para ajudar a reformular as hipóteses em perguntas investigativas claras e abertas. Os alunos, em conjunto, escolhem quais perguntas seguirão adiante.

A **Etapa 5 – Estudo Individual** (15 minutos) acontece em sala, de forma breve. Nesse momento, os alunos acessam os **resumos curtos previamente criados pelo professor** com apoio de *conduct research*. O material fornece insumos mínimos para que todos avancem no raciocínio investigativo dentro do tempo limitado.

Na **Etapa 6 – Retorno e Aplicação do Conhecimento** (15 minutos), os alunos compartilham suas descobertas em um documento coletivo. O professor pode acompanhar e, se necessário, recorrer a um *chatbot* para propor questões adicionais que incentivem a análise crítica do material compartilhado.

Por fim, a **Etapa 7 – Avaliação e Reflexão** (10 minutos) é o momento de fechamento. Os alunos registram percepções rápidas, em papel, chat ou formulário online. O professor pode apoiar-se em um *chatbot* para gerar perguntas de reflexão que orientem esse encerramento, garantindo síntese e consciência do processo de aprendizagem.

Kit essencial de ferramentas

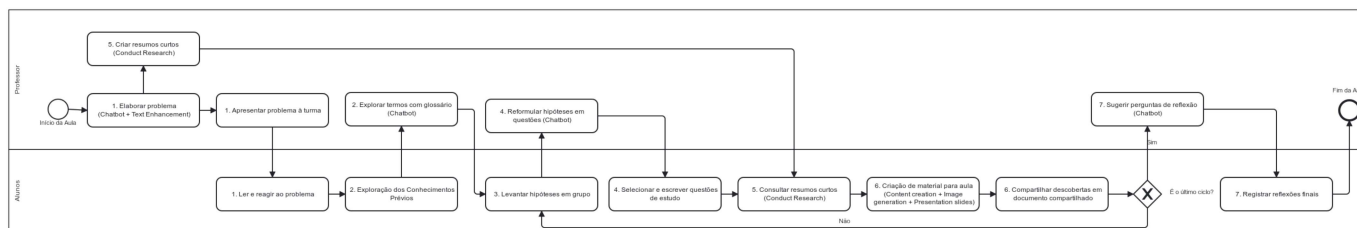
- ◆ **Chatbot** → indispensável para apoiar a criação do problema, a formulação de glossários rápidos, a reformulação de perguntas e a sugestão de reflexões finais.
- ◆ **Text enhancement** → assegura clareza nos enunciados e nas reflexões, evitando perda de tempo com ajustes de linguagem em sala.
- ◆ **Conduct research** → essencial para produzir resumos curtos que viabilizam o estudo individual em tempo reduzido.

- **Note-taking/transcription** → opcional, útil em turmas grandes para registrar falas, mas pode ser substituído por registros dos próprios alunos em documentos compartilhados.

Nível 2 – Problemas semanais em múltiplas aulas

Contexto geral

No Nível 2, um único problema é trabalhado ao longo de **uma semana e duas aulas**. O ciclo completo do PBL (etapas 1 a 7) se fecha dentro da semana: costuma-se avançar até a etapa 4 na primeira aula e concluir as etapas 5 a 7 na segunda. As mesmas ferramentas do Nível 1 continuam presentes, mas há fôlego para **qualificar produtos intermediários** (relatos, mapas, slides) e para que os alunos **pesquisem além dos resumos** preparados pelo professor.



Roteiro de aplicação (referência: 1 semana, 2 aulas)

Na **Etapa 1 – Identificação do Problema**, o professor prepara e apresenta o enunciado do caso. Um *chatbot* apoia a geração do cenário e o *text enhancement* ajusta a linguagem ao nível da turma. Ainda nesta etapa, o professor pode recorrer a *conduct research* para **criar resumos curtos de apoio** (textos-base sintéticos, com links ou referências) que serão disponibilizados aos alunos na fase de estudo.

Na **Etapa 2 – Exploração dos Conhecimentos Prévios**, os alunos discutem o que já sabem e o que precisam saber. Para acelerar a ativação do repertório, o professor pode usar um *chatbot* para produzir um **glossário rápido** de termos-chave ou propor **perguntas provocadoras** que orientem a conversa. O objetivo é consolidar uma base comum antes de avançar.

Na **Etapa 3 – Levantamento de Hipóteses**, os grupos formulam explicações iniciais para o problema e registram suas ideias. Quando fizer sentido, *image generation* pode apoiar a **visualização das hipóteses** (mapas, esquemas, rascunhos). O professor atua como moderador, incentivando diversidade de caminhos antes de qualquer validação.

Na **Etapa 4 – Definição das Questões de Estudo**, as hipóteses mais promissoras são transformadas em **perguntas investigativas claras e abertas**. Um *chatbot* ajuda a **reformular** e **tornar viáveis** essas questões (escopo, foco, critérios), fechando a primeira aula com **perguntas de estudo bem definidas**.

Na **Etapa 5 – Estudo Individual**, já na segunda aula, cada aluno consulta os **resumos curtos preparados pelo professor** e **amplia** essas leituras com suas próprias buscas, apoiadas por *conduct research*. Quando apropriado, *content creation* pode ajudar a **organizar um breve relatório parcial** (com fontes, achados e dúvidas).

Na **Etapa 6 – Retorno e Aplicação do Conhecimento**, os grupos socializam descobertas e conectam-nas às questões de estudo. Para comunicar melhor, podem lançar mão de *presentation slides* e, quando útil, de *image generation* para **diagramas** ou **protótipos rápidos**. O foco é integrar as evidências, comparar caminhos e verificar o que responde (ou não) às perguntas.

Por fim, na **Etapa 7 – Avaliação e Reflexão**, a turma realiza o **fechamento do ciclo semanal**. *Chatbots* podem sugerir **perguntas de meta-reflexão** (o que aprendemos? que evidências suportam nossas conclusões? que limitações restam?) e *text enhancement* contribui para **lapidar** sínteses finais. Se a turma usou registros contínuos, *note-taking/transcription* pode ajudar a recuperar trechos-chave das discussões para embasar a reflexão.

Kit essencial de ferramentas

- **Chatbot** → criação/ajuste do problema, glossários, reformulação de questões, prompts de meta-reflexão.
- **Text enhancement** → clareza nos enunciados, relatórios e sínteses finais.
- **Conduct research** → resumos curtos criados pelo professor e apoio às buscas próprias dos alunos.
- **Content creation** → organização de relatórios parciais e sínteses intermédias.
- **Image generation** → mapas/diagramas de hipóteses e elementos visuais de apoio.
- **Presentation slides** → comunicação estruturada dos resultados na etapa 6.
- **Note-taking/transcription** → opcional; mais útil aqui do que no Nível 1, pois ajuda a recuperar discussões entre uma aula e outra.

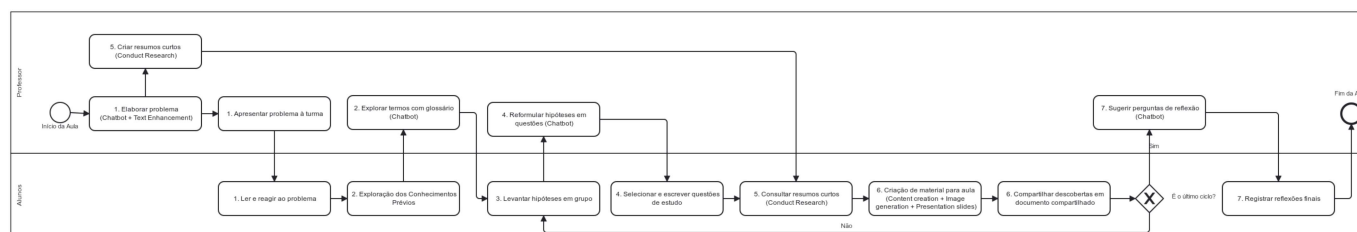
Nível 3 – Problemas de média duração (2 a 4 semanas)

Contexto geral

O Nível 3 é voltado a professores que desejam aprofundar a investigação em torno de um mesmo problema ao longo de **2 a 4 semanas**. Nesse formato, o processo deixa de ser uma experiência pontual ou semanal para se tornar uma **trajetória de pesquisa em mini-ciclos**. O problema inicial é suficientemente complexo para exigir revisões, novas hipóteses e diferentes caminhos de estudo, o que dá origem a **iteração contínua das etapas 3 a 6**.

A lógica é a seguinte: o primeiro ciclo começa com todas as sete etapas, mas, após o fechamento inicial, os alunos retomam hipóteses, redefinem questões e realizam novos estudos em mini-rodadas. Apenas ao final das semanas previstas, a **Etapa 7 – Avaliação e Reflexão** assume caráter global, consolidando todo o percurso.

Esse tempo ampliado permite usar um repertório mais rico de ferramentas de GenAI, incluindo **content creation, image generation e presentation slides**, que apoiam a produção de relatórios intermediários, protótipos e apresentações parciais.



Roteiro de aplicação (referência: 3 semanas, 2 aulas por semana)

Na **Etapa 1 – Identificação do Problema**, o professor apresenta um enunciado elaborado com apoio de *chatbot* para gerar cenários e de *text enhancement* para ajustar clareza e tom. Assim como nos níveis

anteriores, pode também preparar **resumos curtos de apoio** com *conduct research*, garantindo que os alunos tenham uma base inicial para começar a explorar o tema.

Na **Etapa 2 – Exploração dos Conhecimentos Prévios**, os alunos compartilham percepções e repertórios já existentes. O professor pode enriquecer esse processo com glossários ou perguntas provocadoras geradas por *chatbots*. Aqui, *note-taking/transcription* começa a ser mais valioso, pois os registros acumulados se tornam referência para os ciclos seguintes.

Na **Etapa 3 – Levantamento de Hipóteses**, os grupos levantam diferentes explicações para o problema. Com *image generation*, essas hipóteses podem ser transformadas em **mapas conceituais, diagramas de relações ou protótipos iniciais**, facilitando a visualização de alternativas. O professor atua para garantir diversidade de caminhos investigativos.

Na **Etapa 4 – Definição das Questões de Estudo**, as hipóteses são transformadas em perguntas investigativas com apoio de *chatbots*, que ajudam a tornar as questões mais abertas, claras e exequíveis. A turma define quais perguntas nortearão o primeiro ciclo de estudo.

Na **Etapa 5 – Estudo Individual**, os alunos aprofundam a busca por informações. Além dos resumos disponibilizados pelo professor, cada estudante realiza suas próprias investigações com *conduct research*. O maior tempo disponível permite integrar também *content creation*, por exemplo, na **produção de relatórios parciais**, fichamentos ou resumos semanais que organizem as evidências encontradas.

Na **Etapa 6 – Retorno e Aplicação do Conhecimento**, os grupos compartilham resultados, com apoio de *presentation slides* para estruturar apresentações e de *image generation* para enriquecer visualmente suas descobertas. As discussões coletivas alimentam a formulação de **novas hipóteses**, reiniciando o mini-ciclo: as etapas 3 a 6 são repetidas duas ou três vezes ao longo das semanas, sempre com acréscimos e refinamentos.

Por fim, na **Etapa 7 – Avaliação e Reflexão**, a turma realiza um **fechamento global**. *Chatbots* podem sugerir perguntas de meta-reflexão sobre a trajetória, enquanto *text enhancement* auxilia na síntese final dos relatórios ou apresentações comparativas entre hipóteses iniciais e finais. O uso de *note-taking/transcription* garante que reflexões anteriores não se percam e fortalece a visão de processo contínuo.

Kit essencial de ferramentas

- **Chatbot** → apoia a criação do problema, a elaboração de glossários, a reformulação de questões de estudo e a proposição de perguntas de meta-reflexão.
- **Text enhancement** → assegura clareza e consistência em relatórios parciais e finais.
- **Conduct research** → amplia as fontes de consulta e fornece resumos confiáveis para embasar a investigação.
- **Content creation** → permite a produção de relatórios intermediários, fichamentos e sínteses semanais.
- **Image generation** → transforma hipóteses e descobertas em diagramas, mapas e protótipos.
- **Presentation slides** → organiza apresentações intermediárias e finais em cada iteração.
- **Note-taking/transcription** → opcional, mas valioso aqui, pois os registros acumulados ajudam os alunos a acompanhar a evolução entre os ciclos.

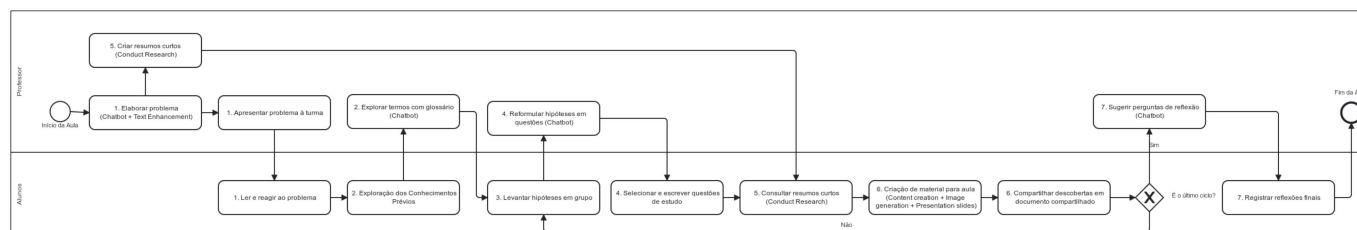
Nível 4 – Um grande problema ao longo do semestre

Contexto geral

O Nível 4 representa a aplicação mais profunda do PBL, em que **um único problema guia toda a disciplina durante o semestre**. Esse problema é amplo, interdisciplinar e aberto, exigindo que os alunos avancem em sucessivos ciclos de investigação e integração de conhecimentos. O professor atua como **curador e mediador contínuo**, ajudando os estudantes a manterem a coerência do processo e a articularem diferentes perspectivas ao longo do tempo.

A lógica é de **mini-ciclos sucessivos das etapas 3 a 6**, cada um durando em média duas a três semanas. No início, a turma passa pelas etapas 1 e 2 com mais calma, construindo uma base sólida de compreensão do problema. Ao longo do semestre, as iterações vão gerando relatórios parciais, protótipos e apresentações intermediárias. O processo se encerra com a **Etapa 7 – Avaliação e Reflexão**, que sintetiza a trajetória global.

Esse tempo estendido abre espaço para a utilização de **todo o repertório de ferramentas de GenAI**, desde chatbots e *text enhancement* até *conduct research*, *content creation*, *image generation* e *presentation slides*.



Roteiro de aplicação (referência: semestre completo, 12–16 semanas)

Na **Etapa 1 – Identificação do Problema**, o professor apresenta um desafio amplo e complexo, elaborado com apoio de *chatbots* para criação de cenários e *text enhancement* para assegurar clareza. Pode também recorrer a *image generation* para ilustrar o contexto e a *content creation* para produzir relatórios introdutórios que deem corpo ao problema. Essa fase pode ocupar de um a dois encontros, ou até se estender por mais tempo caso seja necessário enriquecer o material inicial.

Na **Etapa 2 – Exploração dos Conhecimentos Prévios**, os alunos compartilham repertórios individuais e coletivos. A turma pode construir **mapas conceituais** apoiados por *image generation*, enquanto um *chatbot* sugere glossários ou perguntas provocadoras. Ferramentas de *note-taking/transcription* registram a troca de ideias, criando um banco de referências que será consultado ao longo do semestre.

Na **Etapa 3 – Levantamento de Hipóteses**, os grupos começam a propor explicações para o problema. Diferente dos níveis anteriores, essa etapa é revisitada várias vezes, pois novas hipóteses emergem a cada mini-ciclo. *Image generation* e *content creation* são úteis aqui, para estruturar hipóteses em diagramas ou relatórios parciais que permitam acompanhar a evolução das ideias.

Na **Etapa 4 – Definição das Questões de Estudo**, o professor e os alunos transformam hipóteses em perguntas investigativas. *Chatbots* apoiam a formulação de perguntas claras e abertas, garantindo foco sem perder a complexidade. Essa etapa é repetida em cada mini-ciclo, pois as questões se refinam conforme novas descobertas são feitas.

Na **Etapa 5 – Estudo Individual**, os alunos mergulham em pesquisas mais robustas, utilizando *conduct research* para acessar fontes externas e *content creation* para elaborar fichamentos, relatórios semanais ou pequenos artigos. O tempo estendido permite que cada aluno se especialize em partes diferentes do problema, criando complementaridade entre os grupos.

Na **Etapa 6 – Retorno e Aplicação do Conhecimento**, os grupos apresentam suas descobertas intermediárias. Aqui, *presentation slides* ajudam a estruturar as exposições e *image generation* enriquece os protótipos e mapas conceituais. Ao longo do semestre, essas apresentações parciais funcionam como checkpoints, nos quais a turma reavalia hipóteses e redefine perguntas de estudo, iniciando novos mini-ciclos.

Por fim, na **Etapa 7 – Avaliação e Reflexão**, realizada no final do semestre, a turma compara hipóteses iniciais com as finais e sintetiza a trajetória de investigação. *Chatbots* podem sugerir perguntas de meta-reflexão que orientem a discussão coletiva, enquanto *text enhancement* apoia a produção do **relatório final integrador**. Essa etapa é breve em tempo, mas decisiva em termos pedagógicos, pois consolida o aprendizado e evidencia a evolução coletiva ao longo de todo o processo.

Kit essencial de ferramentas

- **Chatbot** → apoio contínuo, da criação de cenários e perguntas provocadoras até a meta-reflexão final.
- **Text enhancement** → garante clareza e consistência em relatórios parciais, apresentações e síntese final.
- **Conduct research** → expande o alcance investigativo com acesso a fontes externas e resumos de artigos.
- **Content creation** → viabiliza relatórios intermediários e finais, além de sínteses semanais que documentam a evolução da turma.
- **Image generation** → apoia a criação de mapas conceituais, diagramas e protótipos visuais ao longo dos mini-ciclos.
- **Presentation slides** → estrutura apresentações intermediárias e finais, garantindo clareza e organização.
- **Note-taking/transcription** → opcional, mas bastante útil em turmas grandes para registrar discussões e criar um histórico contínuo de ideias.

5. Ferramentas de GenAI no Apoio às Etapas do PBL

A seguir, apresentamos um **panorama de categorias de ferramentas de GenAI**, funcionando como um **manual de referência rápida**. Essa taxonomia organiza os recursos disponíveis, distingue entre soluções **online (SaaS)** e **locais (open-source)**, e indica em quais momentos do PBL cada tipo de ferramenta pode oferecer maior valor pedagógico.

O objetivo não é prescrever quais ferramentas devem ser usadas, mas ampliar o repertório dos docentes, mostrando alternativas que podem ser escolhidas conforme as necessidades de cada disciplina. Para que essa escolha não seja feita de maneira arbitrária, foi necessário definir **critérios de aderência pedagógica** que nos permitissem avaliar o papel real de cada tecnologia no processo investigativo do PBL.

Critérios de Aderência: Escala Likert aplicada ao PBL

Para realizar essa avaliação, utilizamos uma **escala Likert de 5 pontos**, aplicada em cada cruzamento entre **ferramenta** e **etapa do PBL**. A ideia não é apenas indicar *onde* uma ferramenta pode ser usada, mas **o quanto ela contribui pedagogicamente** para professores e alunos em cada fase.

Nível	Nome	Critério objetivo
-------	------	-------------------

Nível	Nome	Critério objetivo
1	Prejudicial	Enfraquece a investigação, fornecendo respostas prontas ou bloqueando o debate crítico.
2	Dispersivo	Não contribui de forma significativa e pode desviar o grupo da análise investigativa.
3	Suporte opcional	Complementa a fase em alguns aspectos, mas sem impacto essencial no processo.
4	Apoio investigativo	Facilita a exploração, a formulação de hipóteses ou a organização coletiva da aprendizagem.
5	Núcleo pedagógico	Potencializa diretamente o raciocínio crítico, a colaboração e a reflexão da etapa.

📖 Dessa forma, a tabela apresentada a seguir não é apenas descritiva: ela traduz os **resultados dessa análise de aderência**, destacando em quais etapas cada categoria de ferramenta obteve pontuação 4 ou 5 na escala.

Tabela de referência

Categoria de ferramentas	O que essas ferramentas fazem?	Fases do PBL com alta aderência (nível 4 ou 5 na escala Likert)
Chatbots	Plataformas interativas em linguagem natural que permitem uma conversa contínua. São úteis para explorar ideias em tempo real, propor hipóteses, responder perguntas, sugerir questões e estimular reflexões.	Etapa 1 – Identificação do problema (Professor) Etapa 2 – Exploração dos conhecimentos prévios (Professor) Etapa 4 – Definição das questões de estudo (Professor) Etapa 7 – Avaliação e reflexão (Professor e Aluno)
Content Creation	Ferramentas voltadas à produção de textos em formatos específicos e prontos para uso, como relatórios, resumos, artigos, propostas e campanhas.	Etapa 6 – Aplicação do conhecimento (Professor)

Categoria de ferramentas	O que essas ferramentas fazem?	Fases do PBL com alta aderência (nível 4 ou 5 na escala Likert)
Text Enhancement	Soluções destinadas a revisar e melhorar textos já produzidos. Realizam correções gramaticais, ajustes de clareza e coesão, refinam o tom e garantem adequação ao público.	Etapa 1 – Identificação do problema (Professor)
		Etapa 4 – Definição das questões de estudo (Professor e Aluno)
		Etapa 5 – Estudo individual (Aluno)
		Etapa 6 – Aplicação do conhecimento (Professor e Aluno)
		Etapa 7 – Avaliação e reflexão (Aluno)
Image Generation	Ferramentas que produzem imagens originais a partir de descrições em texto. São úteis para criar cenários, protótipos visuais e elementos gráficos.	Etapa 1 – Identificação do problema (Professor)
		Etapa 3 – Geração de hipóteses (Professor)
		Etapa 6 – Aplicação do conhecimento (Professor e Aluno)
Note-taking / Meeting Transcription	Aplicativos que capturam automaticamente falas em reuniões ou discussões e as transformam em texto pesquisável.	Etapas 1 a 7 (Professor/Aluno)
Presentation Slides	Plataformas que geram apresentações automaticamente a partir de textos ou tópicos fornecidos.	Etapa 6 – Aplicação do conhecimento (Professor e Aluno)
Conduct Research	Ferramentas que apoiam investigações acadêmicas e profissionais, localizando fontes confiáveis e resumindo artigos.	Etapa 5 – Estudo individual (Aluno)

📖 **Nota de referência** A taxonomia apresentada foi adaptada da proposta de [AlAfnan et al., 2024](#), publicada no *Journal of Artificial Intelligence and Technology*, que organiza aplicações de GenAI em categorias funcionais (*AI Applications – Functions*). Aqui, as categorias foram ajustadas para o **contexto do PBL universitário** e avaliadas segundo a **escala de aderência Likert** às etapas de Walsh, permitindo identificar onde as ferramentas têm maior relevância pedagógica.

Formas de acesso às ferramentas: Online e Local

- **Ferramentas online (SaaS):**

- Acesso imediato, basta criar conta.
- **Prós:** fácil de usar, sem necessidade de instalação.
- **Contras:** dependência de internet, risco de privacidade, custo de assinatura.

- **Ferramentas locais (open-source):**

- Requerem instalação e configuração inicial.
- **Prós:** garantem privacidade dos dados, podem rodar sem internet, maior controle do ambiente.
- **Contras:** exigem conhecimento técnico ou suporte de TI, podem precisar de hardware dedicado (ex.: GPU).

Como acessar e configurar as ferramentas de GenAI

Nesta seção, explicamos como os professores podem acessar os serviços de GenAI apresentados na taxonomia. Dividimos em **ferramentas online (SaaS)** e **ferramentas locais (open-source/self-hosted)**.

1. Chatbots

- **Online (SaaS):**

- **ChatGPT** (OpenAI) → acesso direto via navegador, mediante criação de conta.
- **Claude** (Anthropic) → disponível online, exige criação de conta.
- **Gemini** (Google) → acesso via navegador ou app móvel.

- **Local (open-source):**

- **LM Studio** → aplicativo desktop (Windows/Mac/Linux). O app roda localmente sem conta, mas para baixar modelos diretamente pela interface é necessário login. Alternativamente, os modelos podem ser baixados manualmente (ex.: Hugging Face) e carregados sem conta.
- **Ollama** roda modelos de linguagem localmente no terminal (Windows/Mac/Linux). Não exige conta para uso básico, mas se quiser habilitar a função de busca na internet, precisa de login.

2. Content Creation

- **Online (SaaS):**

- **Jasper** → criação de textos automatizados para marketing e relatórios; exige conta paga após período de teste.
- **Copy.ai** → gera resumos, rascunhos e textos criativos; precisa de conta.
- **Notion AI** → integrado ao Notion; requer conta e assinatura para uso completo.

- **Local (open-source):**

- **LM Studio** → aplicativo desktop (Windows/Mac/Linux). O app roda localmente sem conta, mas para baixar modelos diretamente pela interface é necessário login. Alternativamente, os modelos podem ser baixados manualmente (ex.: Hugging Face) e carregados sem conta.

- **Ollama** roda modelos de linguagem localmente no terminal (Windows/Mac/Linux). Não exige conta para uso básico, mas se quiser habilitar a função de busca na internet, precisa de login.

3. Text Enhancement

- **Online (SaaS):**
 - **Grammarly** → revisão ortográfica, gramatical e de clareza; exige conta.
 - **Wordtune** → reescreve trechos de texto em diferentes estilos; precisa de conta.
 - **DeepL Write** → focado em clareza e precisão da linguagem; requer conta para salvar textos.
- **Local (open-source):**
 - **LanguageTool** → corretor ortográfico e gramatical, disponível como software livre e offline.
 - **LM Studio** → aplicativo desktop (Windows/Mac/Linux). O app roda localmente sem conta, mas para baixar modelos diretamente pela interface é necessário login. Alternativamente, os modelos podem ser baixados manualmente (ex.: Hugging Face) e carregados sem conta.
 - **Ollama** → roda modelos de linguagem localmente no terminal (Windows/Mac/Linux). Não exige conta para uso básico, mas se quiser habilitar a função de busca na internet, precisa de login.

4. Video Creation

- **Online (SaaS):**
 - **Runway** → edição e criação de vídeos com GenAI, exige conta.
 - **Pika** → geração de vídeos a partir de texto e imagens; precisa de conta.
 - **Descript** → edição de vídeo e áudio baseada em transcrição; requer conta.
- **Local (open-source):**
 - **Stable Video Diffusion** → modelo open-source da Stability AI; exige GPU potente.
 - **Open-Sora** → projeto inspirado no Sora da OpenAI, para gerar vídeos com GenAI localmente.

5. Image Generation

- **Online (SaaS):**
 - **DALL-E 3** → integrado ao ChatGPT, exige conta.
 - **Midjourney** → roda via Discord, exige conta.
 - **Stable Diffusion Web** → interface web para testar Stable Diffusion, requer conta para alguns recursos.
- **Local (open-source):**
 - **Stable Diffusion** → modelo open-source de geração de imagens; roda localmente com GPU.
 - **ComfyUI** → interface modular para Stable Diffusion.
 - **Automatic1111** → interface popular para rodar Stable Diffusion localmente.

6. Note-taking / Meeting Transcription

- **Online (SaaS):**

- **Fireflies** → integração com Zoom, Meet e Teams para transcrição automática; exige conta.
- **Mem** → captura de notas inteligentes com GenAI; precisa de conta.
- **Local (open-source):**
 - **Whisper** → modelo open-source da OpenAI para transcrição em várias línguas; pode rodar em CPU ou GPU.
 - **Vosk** → ferramenta de reconhecimento de fala offline; roda em dispositivos modestos.
 - **Coqui STT** → alternativa livre para transcrição de fala em texto.

7. Inbox Management

- **Online (SaaS):**
 - **SaneBox** → organiza caixas de entrada automaticamente; exige conta.
 - **EmailTree** → prioriza mensagens e sugere respostas; precisa de conta.
- **Local (open-source):**
 - **Scripts personalizados com LLM local (LM Studio ou Ollama)** → podem ser integrados ao cliente de e-mail (ex.: Thunderbird) para classificação automática de mensagens.

8. Presentation Slides

- **Online (SaaS):**
 - **Beautiful.ai** → cria apresentações com design automático; exige conta.
 - **Decktopus** → gera apresentações rápidas; precisa de conta.
 - **Slidesgo** → oferece modelos prontos, com GenAI para sugestões; requer conta para downloads premium.
- **Local (open-source):**
 - **Marp** → cria apresentações a partir de Markdown.
 - **LaTeX Beamer** + LLM local → gera apresentações em LaTeX, útil para professores que já usam esse ecossistema.

9. Conduct Research

- **Online (SaaS):**
 - **Perplexity AI** → busca em múltiplas fontes com referências; exige conta para histórico e recursos avançados.
 - **Consensus** → pesquisa em artigos científicos; precisa de conta.
 - **Scite** → análise de citações acadêmicas; requer conta.
- **Local (open-source):**
 - **LM Studio** → aplicativo desktop (Windows/Mac/Linux). Ele roda modelos de linguagem localmente e pode ser usado para explorar documentos carregados pelo professor, como artigos

em PDF ou materiais da disciplina. Dessa forma, as respostas são baseadas nesses textos de referência, funcionando como um apoio para revisão de literatura em sala de aula.

As categorias de ferramentas apresentadas formam um repertório de apoio que pode ser mobilizado de acordo com cada etapa do PBL. Elas não substituem o processo pedagógico, mas oferecem alternativas que ampliam a investigação, facilitam a produção e tornam a aprendizagem mais dinâmica. O essencial é que sejam utilizadas de forma crítica, sempre alinhadas ao propósito formativo da metodologia.

Com essa visão, encerramos a exploração prática das possibilidades de integração entre GenAI e PBL e avançamos agora para uma reflexão conclusiva sobre o papel dessa integração no ensino superior contemporâneo.

6. Aspectos Éticos e Regulatórios

Se a inteligência artificial amplia as possibilidades pedagógicas, também expõe professores e estudantes a novas vulnerabilidades. Discutir ética não é um adendo burocrático, mas parte essencial da integração responsável da tecnologia ao PBL.

O primeiro ponto é a **privacidade**. Ferramentas online armazenam e processam dados em servidores externos, muitas vezes fora do país. Isso significa que anotações de aula, gravações de voz e até relatórios de estudantes podem ser compartilhados inadvertidamente. Sempre que possível, é recomendável adotar soluções **locais ou institucionais**, que mantenham o controle das informações no próprio ambiente acadêmico.

O segundo ponto é o **consentimento**. Quando a atividade envolver gravação, transcrição ou qualquer forma de coleta de dados dos alunos, é importante que eles saibam exatamente o que está acontecendo. Essa transparência fortalece a confiança e evita a sensação de vigilância.

Por fim, há a questão do **uso pedagógico responsável**. A GenAI não deve ser usada para substituir a reflexão, mas para instigá-la. Um recurso que oferece respostas prontas pode ser sedutor, mas empobrece o processo formativo. A proposta deste guideline é justamente diferenciar usos que ampliam a investigação daqueles que a esvaziam.

Ao trazer essas considerações, não pretendemos transformar cada professor em especialista em legislação, mas oferecer referências claras para que a inovação caminhe lado a lado com a proteção dos estudantes e com a integridade acadêmica. Afinal, usar GenAI em sala de aula não é apenas uma escolha técnica, mas também uma decisão ética.

7. Considerações Finais

A integração entre **Aprendizagem Baseada em Problemas (PBL)** e **Inteligência Artificial** não é apenas uma inovação metodológica, mas uma oportunidade de recolocar o estudante no centro do processo educativo. Ao longo deste guideline, mostramos como cada etapa do PBL pode ser enriquecida por recursos de GenAI, sem que isso comprometa a autonomia investigativa ou o rigor acadêmico.

O papel do professor permanece essencial. Ele é o mediador que garante foco, profundidade e criticidade, evitando que a GenAI se torne apenas um atalho para respostas prontas. Do lado do estudante, a GenAI deve ser vista como ferramenta de apoio à investigação, e não como substituto do esforço intelectual.

As ferramentas apresentadas não esgotam as possibilidades. Ao contrário, compõem um catálogo em constante evolução, que pode e deve ser adaptado às condições de cada curso, instituição e disciplina. O mais importante é que docentes se sintam encorajados a experimentar de forma intencional e crítica, observando os efeitos sobre a aprendizagem e ajustando o processo conforme necessário.

Em última instância, este guideline não pretende oferecer receitas prontas, mas mapas de possibilidades. Cabe a cada professor, junto de seus alunos, trilhar o caminho que melhor conecta o PBL, a GenAI e os objetivos formativos da educação superior contemporânea.