

## **Redesenho da Avaliação de Reação em Cursos de Formação Judiciária: uma abordagem inovadora com suporte de inteligência artificial**

Joriza Magalhães Pinheiro, Tribunal de Justiça do Estado do Ceará

Vanessa Maria Quariguasy Pereira Veras, Tribunal de Justiça do Estado do Ceará

Francisco Anastácio Cavalcante Neto, Tribunal de Justiça do Estado do Ceará

Welkey Costa do Carmo, Tribunal de Justiça do Estado do Ceará

Iago Capistrano Sá, Tribunal de Justiça do Estado do Ceará

Tema de interesse: Inovações, inteligência artificial e tecnologias de informação e comunicação em Sistemas de Justiça

### **RESUMO**

Este relato tem por objetivo descrever o redesenho inovador do sistema de avaliação de reação de cursos na Escola de Magistratura do um Tribunal de Justiça do Estado do Ceará. A justificativa reside na insuficiência do modelo anterior, baseado em escalas Likert, que gerava dados majoritariamente positivos porém pouco acionáveis para melhorias, criando um ciclo de feedback inócuo.

A inovação implementada é um sistema integrado que combina: (1) um formulário simplificado com três perguntas abertas (“Que bom...”, “Que pena...”, “Que tal...”), cujo preenchimento é estimulado ainda em sala de aula; (2) uma estratégia de comunicação para engajamento; e (3) um assistente de IA seguro, baseado no gerenciador Ollama e no modelo Gemma 2, hospedado na infraestrutura interna do tribunal, que analisa as respostas textuais e gera relatórios analíticos automáticos com insights de melhora, alertas de assédio e uma visão geral do sentimento da turma.

O projeto encontra-se em estágio final de implementação, com previsão de início da operação para setembro de 2025. Como resultados esperados, almeja-se um aumento significativo no engajamento e na qualidade do feedback, a geração de insights acionáveis para instrutores e coordenadores, e a criação de um ciclo de melhoria contínua transparente. A prática assegura a confidencialidade dos dados (com processamento interno) e cria um ciclo de feedback eficaz, onde a análise automatizada fornece à coordenação da Esmec evidências concretas para orientar decisões pedagógicas, otimizar cursos e demonstrar valor aos participantes.



## Introdução

O Poder Judiciário brasileiro, notadamente em sua esfera estadual, opera em um contexto de extrema complexidade e demanda social crescente. A formação e o desenvolvimento contínuo de magistrados e servidores – a Formação Judiciária – emergem, portanto, como pilares estratégicos para a garantia de uma prestação jurisdicional eficiente, moderna e alinhada aos anseios da sociedade. Esta formação transcende a mera transmissão de conhecimento técnico-jurídico; deve capacitar os operadores do direito a lidar com desafios multifacetados, que vão desde a gestão de processos e pessoas até questões sensíveis envolvendo direitos humanos, inovações tecnológicas e a complexidade de novas leis. Neste cenário, a constante inovação nos métodos de ensino-aprendizagem e, crucialmente, na avaliação destes programas formativos, deixa de ser uma opção e se torna um imperativo para a manutenção da qualidade e da legitimidade institucional.

Dentro do arcabouço da avaliação de treinamentos, o modelo proposto por Donald Kirkpatrick em 1959 permanece como um dos mais influentes e utilizados globalmente. Seu primeiro nível, a avaliação de reação, é dedicado a mensurar a satisfação, o engajamento e a percepção de relevância dos participantes em relação ao curso. Kirkpatrick defendia que uma reação positiva é um pré-requisito necessário, ainda que não suficiente, para a aprendizagem efetiva (KIRKPATRICK, 1959). Trata-se do primeiro filtro de qualidade, capaz de fornecer insights imediatos sobre a experiência do aluno, identificando pontos fortes a serem mantidos e gaps críticos a serem sanados em edições futuras. É a ferramenta mais ágil para ajustes finos e para a construção de uma cultura de melhoria contínua dentro da escola judicial.

Contudo, as limitações históricas desse nível são amplamente documentadas na literatura. A avaliação de reação, quando operacionalizada de forma tradicional por meio de questionários quantitativos com escalas de Likert, frequentemente gera um fenômeno conhecido como "viés de positividade" ou "efeito halo". Os participantes, por uma série de fatores que vão desde a simpatia pelo instrutor até a falta de tempo para refletir criticamente, tendem a atribuir notas máximas de forma automatizada. O resultado é uma enxurrada de dados quantitativos positivos que, embora confortáveis estatisticamente, são pobres em significado e pouco úteis para direcionar melhorias concretas. Como apontam Nawaz, Khushnood e Wisal (2022), em um meta-estudo abrangente, a correlação entre uma reação positiva e a efetiva transferência de



aprendizagem para o comportamento (nível 3) ou para resultados organizacionais (nível 4) é fraca e não significativa. O modelo, em sua aplicação prática, frequentemente falha em conectar os processos formativos com resultados institucionais e sociais tangíveis.

Esta crítica à excessiva dependência de métricas quantitativas levou ao surgimento de outros modelos e adaptações. Jack Phillips, por exemplo, propôs um quinto nível (ROI - Retorno sobre o Investimento) para quantificar financeiramente o impacto do treinamento (PHILLIPS, 1996). Já Roger Kaufman direcionou o foco para a macroavaliação, argumentando que a avaliação deve medir não apenas o desempenho individual ou organizacional, mas também o impacto societal (KAUFMAN & KELLER, 1994). Essas perspectivas reforçam que a avaliação precisa ser um processo estratégico, gerador de evidências para a tomada de decisão, e não um mero ritual de final de curso.

No Tribunal de Justiça do Estado do Ceará, em análise, esta responsabilidade recai sobre a Escola de Magistratura (Esmec). Criada pela Lei nº 11.203 de 1986, ela atua na formação e treinamento de juízes e juízas, com sede em Fortaleza e núcleos em Crato, Iguatu, Sobral e Crateús. Sua missão é capacitar os integrantes do Poder Judiciário do Estado do Ceará para aprimorar a prestação jurisdicional e disseminar conhecimento.

A escola historicamente enfrentou esses mesmos desafios com o feedback relatados até aqui. Seu modelo de avaliação de reação, baseado em formulários longos com múltiplas afirmativas e escalas Likert, resultava em um paradoxo: índices de satisfação consistentemente superiores a 95%, mas com baixíssimo poder de diagnóstico para aperfeiçoamento pedagógico. As respostas, em sua grande maioria, não ofereciam feedback qualitativo substantivo, e o processo de análise manual de eventuais comentários textuais era inviável devido ao volume e à dispersão das informações. Além disso, o preenchimento era voluntário, o que resultava em uma taxa de resposta modesta, potencialmente enviesada (apenas os muito satisfeitos ou muito insatisfeitos se dedicavam a responder).

Esta combinação de feedback excessivamente positivo e pouco crítico, baixa utilidade para a melhoria real dos cursos e falta de engajamento dos participantes criou um cenário de estagnação. A avaliação cumpria uma função burocrática de certificação da atividade, mas falhava em seu propósito central: ser o motor inicial de um ciclo virtuoso de aprimoramento contínuo da formação judiciária. Foi essa lacuna entre a



coleta de dados e a geração de inteligência acionável que motivou a busca por uma abordagem inovadora. O redesenho do processo tornou-se urgente não apenas para modernizar um instrumento, mas para realinhar toda a prática avaliativa aos objetivos estratégicos do Tribunal, assegurando que a formação judicial seja um investimento com retorno claro em qualidade jurisdicional e, em última instância, em justiça para a sociedade.

### Caracterização da inovação

A inovação no setor público enfrenta obstáculos como rigidez processual e aversão ao risco. O projeto supera essas barreiras ao adotar uma governança de inovação que integra cocriação, agile e IA ética em ambiente controlado. Mais que uma ferramenta, a iniciativa representa uma política institucional de modernização, demonstrando como os órgãos podem transformar processos tradicionais em sistemas inteligentes e orientados por dados, assegurando excelência no serviço público.

O redesenho da avaliação de reação na Escola de Magistratura (Esmec) do Tribunal de Justiça do Estado do Ceará (TJCE) transcende a mera atualização de um formulário, configurando-se como um projeto estratégico integrado de gestão da qualidade educacional. Batizado provisoriamente de "Feedback Eficaz", esta iniciativa compreende um conjunto sistêmico de intervenções desenhadas para superar as limitações do modelo anterior e transformar a avaliação em um instrumento genuíno de melhoria contínua. O projeto teve seu início formal em maio de 2025, com previsão de implementação da solução completa até o final de setembro do mesmo ano.

A motivação central partiu de uma demanda da Comissão Permanente de Avaliação Institucional da Esmec (CPAI) (TJCE, 2025), que identificou a insuficiência dos dados gerados pelo modelo quantitativo tradicional para fundamentar decisões estratégicas. A solução concebida, portanto, foi arquitetada em quatro pilares interligados:

#### 1. Redesign da experiência do participante:

A substituição integral do questionário de múltiplas escalas Likert por um instrumento minimalista e qualitativo, composto por três perguntas abertas e indutoras:

"Que bom..." para capturar aspectos positivos e fortalecedores da experiência;

"Que pena..." para identificar frustrações, gaps e oportunidades de melhoria;

"Que tal..." para coletar sugestões concretas e ideias para o futuro.

	 INSTITUTO BRASILEIRO DE ESTUDOS E PESQUISAS SOCIAIS	 Universidade de Brasília		 Universidade Potiguar
		 DGP DIREÇÃO-GERAL DA POLÍTICA DE JUSTIÇA	 Iluris Instituto de Investigação Interdisciplinar	 AJUS Administração da Justiça
	 GEJUD Grupo de Pesquisa Gestão, Desempenho e Efetividade do Judiciário	 InfoJus Núcleo de Pesquisa em Informação, Direito e Sociedade	 LIOrg LINGUAGEM, INSTITUIÇÕES E ORGANIZAÇÕES	

Para garantir a adesão e valorizar a participação, o preenchimento deste novo formulário foi inserido em uma estratégia de comunicação abrangente, que inclui uma nova identidade visual para a avaliação, como visto na *figura 1*, vídeos institucionais de estímulo à participação, peças publicitárias eletrônicas e impressas para divulgação em turmas presenciais e a distância (EAD), e parceria com a Assessoria de Comunicação do Tribunal para divulgar as melhorias implementadas a partir do feedback recebido.



Figura 1 - Banner de divulgação e tela de acesso

## 2. Processo colaborativo de ideação:

A concepção desta abordagem foi realizada de forma colaborativa em uma oficina de ideação conduzida pelo LabLuz – Laboratório de Inovação do TJCE, que teve como desafio a pergunta “Como dar utilidade às avaliações de reação dos seus cursos?”. O processo contou com a participação multiprofissional de instrutores, coordenadores pedagógicos da Esmec, representantes de Tecnologia da Informação (TI) e, crucialmente, alunos e ex-alunos, garantindo que a solução fosse cocriada com seus principais usuários.

## 3. Sistema inteligente e seguro de análise qualitativa:

	 INSTITUTO BRASILEIRO DE ESTUDOS E PESQUISAS SOCIAIS	 Universidade de Brasília		 Universidade Potiguar
		 DGP   DIREÇÃO-GERAL DA POLÍTICA DE JUSTIÇA	 Iluris Instituto de Investigação Interdisciplinar	 AJUS Administração da Justiça
	 GEJUD Grupo de Pesquisa Gestão, Desempenho e Efetividade do Judiciário	 InfoJus Núcleo de Pesquisa em Informação, Direito e Sociedade	 LIOrg LINGUAGEM, INSTITUIÇÕES E ORGANIZAÇÕES	



A migração para um modelo de respostas abertas criou o desafio da análise manual de um volume potencialmente grande de dados textuais subjetivos. Para superar esta barreira operacional sem perder a riqueza do feedback qualitativo, foi desenvolvida uma arquitetura de análise baseada em IA que prioriza a segurança e a confidencialidade dos dados.

Toda a solução é hospedada internamente na infraestrutura do TJCE. O sistema de avaliação ("AvaliaçãoEsmec") e o servidor de modelos de linguagem (Ollama) operam dentro do ambiente controlado do Tribunal. Nenhum dado sensível é externalizado para APIs ou servidores públicos, garantindo absoluta conformidade com as políticas de segurança e proteção de dados da instituição.

Quando uma turma finaliza suas respostas, um coordenador aciona manualmente a opção "Gerar Relatório". Este comando dispara um processo automatizado que executa, de forma sequencial, a geração de uma nuvem de palavras e uma série de chamadas configuráveis à API interna do Ollama.

Para cada chamada, a aplicação principal envia um prompt específico e altamente estruturado contendo o corpus textual das respostas. O Ollama, atuando como um gerenciador de modelos, processa essas solicitações utilizando o modelo selecionado para esta tarefa, o *Large Language Model* (LLM) Gemma 2, que é um tipo de inteligência artificial que, através do aprendizado profundo, é treinada com grandes quantidades de dados textuais para entender e gerar linguagem humana.

O objetivo de cada chamada à LLM é extrair informações específicas e altamente estruturadas. Os prompts são desenhados para instruir o modelo a realizar tarefas como: avaliar pontos positivos, elencar pontos negativos, detectar potenciais relatos de assédio ou discriminação (com base em uma lista de palavras-chave pré-definida) e realizar uma contagem sentimental básica. Criticamente, a LLM é instruída a retornar a análise de cada tarefa, automatizando a estruturação dos dados exclusivamente no formato *JavaScript Object Notation* (JSON), que é um formato de texto leve e legível por humanos, usado para armazenar e trocar dados de forma estruturada. Estes JSONs são então consolidados pela aplicação principal para gerar automaticamente um relatório analítico unificado por turma, contendo:

Análise de sentimento e contagem, ou seja, uma visão quantificada do feedback;

Nuvem de palavras para a visualização dos termos mais frequentes;



Sistema de alertas, sinalizando respostas que dispararam os filtros de detecção de assédio, permitindo um acionamento ágil e humanizado da coordenação.

Insights para melhoria em seções categorizadas com sugestões extraídas para a Escola, para o desenho do curso e para o desempenho dos instrutores, transformando feedback bruto em recomendações acionáveis.

#### 4. Fechamento do Ciclo de Melhoria:

A inovação se completa com o compromisso institucional de dar visibilidade aos resultados. A estratégia de comunicar publicamente as melhorias implementadas a partir das avaliações cria um ciclo virtuoso de feedback, demonstrando ao corpo discente que sua participação é valorizada e gera impactos tangíveis, o que, por sua vez, fortalece o engajamento nas avaliações futuras.

Em essência, o projeto "Feedback Eficaz" caracteriza-se não como uma ferramenta isolada, mas como uma metodologia integrada que combina design thinking, estratégia de engajamento e uma arquitetura de inteligência artificial segura e internalizada para fechar o ciclo da avaliação de reação, convertendo-a de um mero protocolo final em um motor de inovação pedagógica contínua.

#### Fundamentação teórica

O debate sobre os métodos de avaliação em treinamentos vem de longa data, centrado na dicotomia entre a mensuração quantitativa, padronizada e generalizável, e a análise qualitativa, rica em nuance e contexto, porém complexa de ser escalada (Sitzmann et al., 2008). As escalas Likert, ubíquas na avaliação de reação, são frequentemente criticadas por gerarem o que Paswan e Ganesh (2009) denominam de "viés de aquiescência" (a tendência de concordar com afirmações positivas) e por produzirem dados que, embora facilmente tabuláveis, são notoriamente pobres em explicar o "porquê" por trás das pontuações, oferecendo pouca utilidade para ações corretivas específicas (Brown, 2005).

Esta limitação do modelo puramente quantitativo levou a um crescente apelo por métodos mistos e abordagens qualitativas que capturem a voz autêntica do aprendiz. As respostas abertas fornecem dados narrativos que revelam nuances, contradições e sugestões concretas que as escalas numéricas obliteram (Saunders, 2012). No entanto, a análise manual desse conteúdo textual é historicamente inviável em larga escala devido ao tempo, custo e inconsistência da codificação humana, criando um paradoxo:

 Programa de Pós-Graduação em Administração UFPB	 IBEPES INSTITUTO BRASILEIRO DE ESTUDOS E PESQUISAS SOCIAIS	 Universidade de Brasília	 PROGRAMA DE PÓS-GRADUAÇÃO EM DIREITO	 Universidade Potiguar
 Centro Universitário	 1290 FACULDADE DE DIREITO UNIVERSIDADE DE COIMBRA	 DGP DIREÇÃO-GERAL DA POLÍTICA DE JUSTIÇA	 Iluris Instituto de Investigação Interdisciplinar	 AJUS Administração da Justiça
 Grupo de Pesquisa em Administração, Governo e Políticas Públicas do Poder Judiciário	 GEJUD Grupo de Pesquisa Gestão, Desempenho e Efetividade do Judiciário	 InfoJus Núcleo de Pesquisa em Informação, Direito e Sociedade	 LIOrg LINGUAGEM, INSTITUIÇÕES E ORGANIZAÇÕES	

a coleta de dados ricos é facilitada, mas sua análise torna-se um gargalo intransponível (Braun & Clarke, 2006).

É neste contexto que a Inteligência Artificial emerge como um disruptor catalítico. O campo de *Learning Analytics* tem avançado na aplicação de técnicas de Processamento de Linguagem Natural (PLN) para automatizar a análise de feedback educacional aberto. A análise de sentimento, em particular, tornou-se uma ferramenta valiosa para triar grandes volumes de texto, classificando opiniões em categorias como positivas, negativas e neutras (Pang & Lee, 2008). No entanto, a inovação proposta neste projeto vai além da simples análise de polaridade sentimental.

Trabalhos recentes demonstram a aplicação de modelos de linguagem para tarefas mais sofisticadas, como a extração de tópicos, a categorização temática automática e o resumo de feedback, transformando corpus textuais não estruturados em relatórios analíticos acionáveis (Kastrati et al., 2021; Wei et al., 2023). Este é um avanço crítico, pois desloca o foco da IA da simples mensuração para a geração de inteligência.

A autoridade na área, George Siemens, pioneiro em *Learning Analytics*, argumenta que o valor da IA na educação não está em substituir o julgamento pedagógico, mas em amplificar a capacidade humana de interpretação e ação (Siemens, 2013). Isto se alinha perfeitamente com a proposta aqui descrita: o assistente de IA não substitui o coordenador pedagógico; ele o empodera, filtrando o sinal do ruído e destacando os *insights* mais críticos e recorrentes de milhares de respostas textuais.

Portanto, este projeto situa-se na vanguarda de dois movimentos contemporâneos: 1) a ressignificação da avaliação de reação, reposicionando-a de um instrumento de mensuração de satisfação para uma ferramenta de inteligência qualitativa; e 2) a adoção de uma IA ética e centrada no humano, onde a tecnologia não é um fim em si mesma, mas um meio para escalar a análise de dados qualitativos, permitindo que a instituição ouça efetivamente cada voz e aja com agilidade e precisão para melhorar continuamente sua oferta educacional. A implementação de uma arquitetura interna e segura, como a aqui descrita, responde ainda a preocupações éticas e de privacidade levantadas por autores como Zawacki-Richter et al. (2019), críticas sobre o uso de plataformas externas com dados sensíveis de aprendizagem.

### Planejamento, desenvolvimento e implementação

O estudo se iniciou pela identificação do problema e solicitação (março de 2025). A Comissão Permanente de Avaliação Institucional da Esmec (CEPAI) formalizou a

 Programa de Pós-Graduação em Administração UFPB	 IBEPES INSTITUTO BRASILEIRO DE ESTUDOS E PESQUISAS SOCIAIS	 Universidade de Brasília	 PPGD	 Universidade Potiguar
 IESB Centro Universitário	 1290 FACULDADE DE DIREITO UNIVERSIDADE DE COIMBRA	 DGP DIREÇÃO-GERAL DA POLÍTICA DE JUSTIÇA	 Iluris Instituto de Investigação Interdisciplinar	 AJUS Administração da Justiça
 GPJus Grupo de Pesquisa em Administração, Governo e Políticas Públicas do Poder Judiciário	 GEJUD Grupo de Pesquisa Gestão, Desempenho e Efetividade do Judiciário	 InfoJus Núcleo de Pesquisa em Informação, Direito e Sociedade	 LIOrg Linguagem, Instituições e Organizações	



demanda por uma reformulação do sistema de avaliação, diagnosticando que os dados das escalas Likert não sustentavam decisões efetivas de melhoria pedagógica.

Em seguida se deu a estruturação e ideação (Abril de 2025). Utilizou-se o Canvas do Experimento, relatado por Carmo & Sá (2024) e visto na figura 2, para mapear hipóteses, definir indicadores-chave (como o número de incursões geradas a partir do feedback) e planejar a intervenção de forma visual. A ideação foi conduzida pelo LabLuz com participantes multidisciplinares (coordenação pedagógica, instrutores, TI, alunos e ex-alunos). Essa fase foi crucial para cocriar as soluções, resultando no conceito das três perguntas abertas e na visão inicial do assistente de IA.

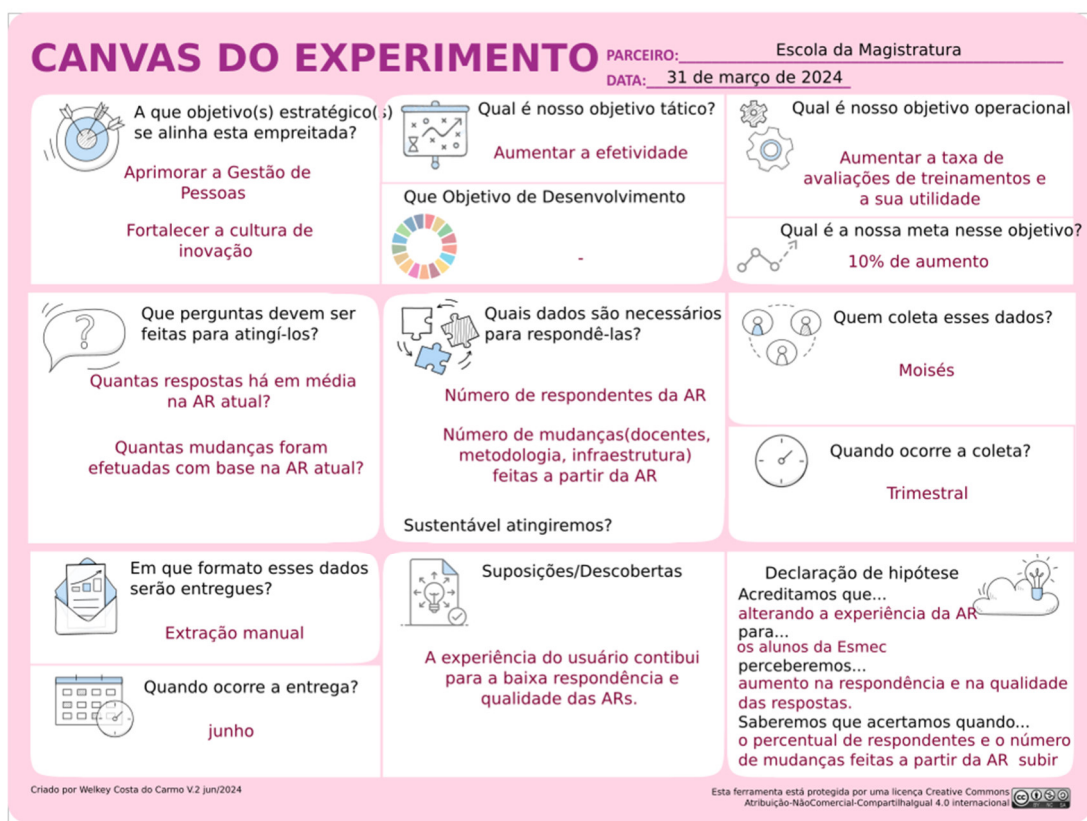


Figura 2- Canvas do experimento

O desenvolvimento e integração ocorreram de maio a julho de 2025. Nesta fase foram executados paralelamente:



A construção do novo formulário digital com as três perguntas abertas e a integração com o sistema de emissão de certificados e o desenvolvimento da arquitetura segura de análise. O laboratório de inovação implementou o Ollama localmente na infraestrutura do TJCE, integrou-o via API à aplicação de avaliação e desenvolveu os prompts estruturados para o modelo Gemma 2, com foco na geração de respostas em JSON, como visto na figura 3, abaixo.

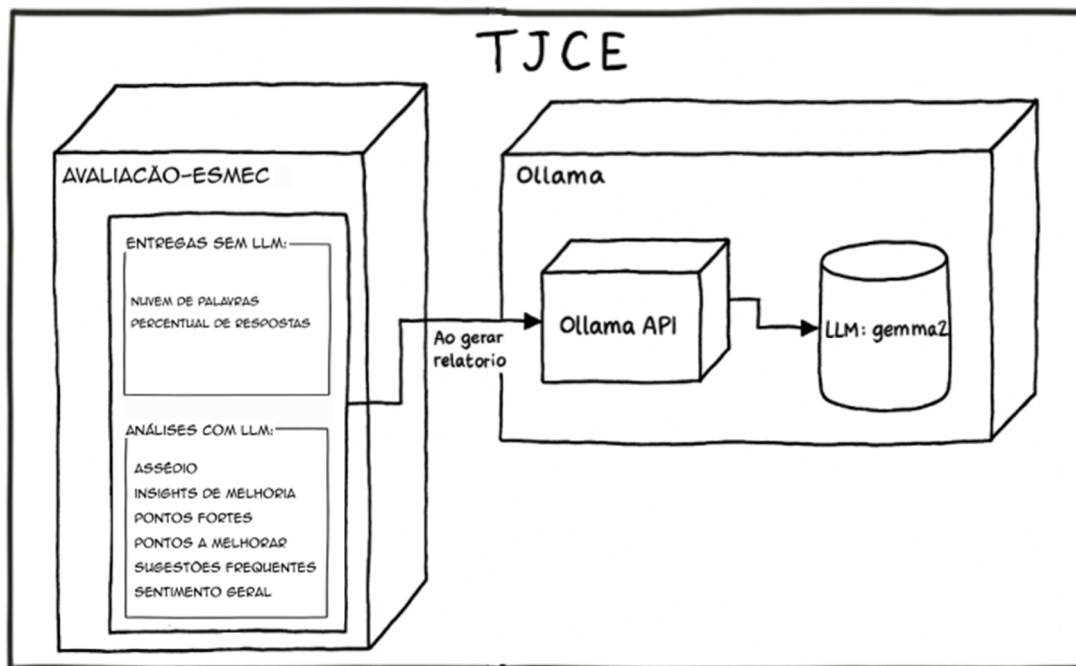


Figura 3- Arquitetura da solução

O próximo passo foi a validação e ajustes (Agosto de 2025). Realização de testes internos e uma reunião de validação com a equipe do projeto para refinamento dos prompts, do fluxo de geração de relatórios e da interface final.

A implementação se deu no final de agosto de 2025 e a etapa em curso é o lançamento oficial do novo sistema, acompanhado pela campanha de comunicação (vídeos, banners, materiais informativos) para engajar os participantes desde o primeiro momento em seus cursos.

## Métodos e Técnicas Empregados

Gestão: Canvas do Experimento para estruturar hipóteses e métricas.



**Cocriação:** Oficina de ideação com Design Thinking envolvendo stakeholders diversos.

**Desenvolvimento:** Metodologias ágeis para o desenvolvimento incremental do software e da integração com IA.

**Segurança:** Adoção de uma arquitetura de IA interna e confinada, utilizando o Ollama como orchestrator e o modelo open-source Gemma 2, garantindo que nenhum dado sensível deixe a infraestrutura do TJCE.

### Resultados alcançados

Considerando que a implementação está calendarizada para o final de setembro de 2025, os resultados apresentados são projeções e expectativas, baseadas na validação interna e em testes-piloto. A mensuração formal dos benefícios ocorrerá ao longo do último trimestre de 2025.

### Benefícios Esperados:

**Aumento da taxa de resposta:** espera-se elevar a taxa de participação nas avaliações de reação de uma média histórica de ~30% (modelo voluntário) para próximo de 100%, como consequência da campanha de comunicação.

**Qualidade e profundidade do feedback:** projeta-se que as três perguntas abertas resultem em um volume de dados textuais pelo menos 5 vezes maior que o modelo anterior, com um aumento significativo na especificidade e no caráter acionável das críticas e sugestões.

**Eficiência operacional:** a expectativa é de que o tempo despendido pela coordenação para analisar o feedback de uma turma seja reduzido de várias horas (ou dias) de análise manual para questão de minutos, tempo necessário para revisar o relatório síntese gerado automaticamente pelo assistente de IA.

**Geração de ações de melhoria tangíveis:** antecipa-se um aumento mensurável no número de incursões e intervenções registradas pela Esmec (ex.: ajustes em ementas, substituição de material, feedback específico para instrutores) com base direta nas evidências extraídas dos relatórios.

### Plano de Mensuração e Métricas:

**Métrica quantitativa 1:** comparativo entre a taxa média de resposta histórica (2023-2024) e a taxa observada no 4º trimestre de 2025.

	 INSTITUTO BRASILEIRO DE ESTUDOS E PESQUISAS SOCIAIS			 Universidade Potiguar
		 DGP   DIREÇÃO-GERAL DA POLÍTICA DE JUSTIÇA	 Iluris Instituto de Investigação Interdisciplinar	
	 GEJUD Grupo de Pesquisa Gestão, Desempenho e Efetividade do Judiciário			

Métrica quantitativa 2: Número de ações de melhoria (incursões) implementadas pela Esmec com lastro nos relatórios do novo sistema, a ser comparado com o período anterior.

Evidência qualitativa: Análise de conteúdo dos relatórios gerados para verificar a presença de insights específicos, não genéricos, que possam direcionar melhorias.

Pesquisa de percepção: Aplicação de um breve questionário com os coordenadores pedagógicos e instrutores após a implementação para avaliar a utilidade percebida e a acionabilidade dos novos relatórios, comparando com o modelo anterior.

A implementação bem-sucedida desta inovação tem o potencial de posicionar a Esmec na vanguarda da gestão pedagógica no Judiciário brasileiro, oferecendo um modelo replicável de com tecnologias abertas e seguras para obter inteligência estratégica a partir da voz do usuário. Os resultados alcançados serão documentados em futuras publicações.

### Conclusões e recomendações

O projeto "Feedback Eficaz" representa uma ruptura significativa com os paradigmas tradicionais de avaliação de reação no poder judiciário. Seus objetivos primários – substituir um modelo quantitativo e burocrático por um qualitativo e acionável, e superar a barreira operacional da análise manual por meio de IA – foram integralmente alcançados em sua concepção e desenvolvimento. A solução técnica construída, segura e internalizada, atende diretamente à demanda por um ciclo de feedback mais ágil e significativo. No entanto, o sucesso final em atingir seus objetivos estratégicos de melhoria contínua só poderá ser definitivamente aferido após a implementação e consolidação da prática a partir de setembro de 2025.

Os principais resultados, até o momento, são de natureza processual e tecnológica: a cocriação de um novo instrumento de avaliação, o desenvolvimento de uma arquitetura de IA ética e confinada, e a criação de um fluxo de trabalho que transforma dados textuais brutos em relatórios estruturados. O sucesso na fase de ideação e desenvolvimento sugere uma alta probabilidade de que os benefícios esperados – engajamento massivo, insights profundos e eficiência operacional – se concretizem.

As perspectivas para a inovação são promissoras. A natureza *open-source* da combinação tecnológica (Ollama + Gemma 2) oferece flexibilidade para futuras adaptações, como o ajuste fino do modelo com os próprios dados da Esmec para



melhor capturar nuances do jargão jurídico-pedagógico, ou a integração com outros sistemas de gestão educacional do tribunal.

Recomendações para a institucionalização:

Para garantir que a inovação transcenda o status de projeto piloto e se torne uma prática permanente, recomenda-se realizar sessões regulares de treinamento com a equipe da coordenação pedagógica para interpretar e agir com base nos relatórios gerados pela IA, assegurando que o insight se transforme em ação.

Para além dos treinamentos, institucionalizar um procedimento operacional padrão (POP) que obrigue a revisão dos relatórios de cada turma e a documentação das decisões tomadas em resposta aos feedbacks, fechando o ciclo de forma transparente.

Também sugere-se manter e expandir a estratégia de comunicação, publicando trimestralmente um boletim mostre as principais melhorias implementadas a partir das avaliações, reforçando para os participantes o valor de sua contribuição.

Por último, recomenda-se estabelecer uma rotina para supervisionar a evolução da ferramenta, revisar a eficácia dos prompts, atualizar os modelos de LLM e garantir a conformidade contínua com as políticas de ética e segurança de dados.

Sugestões para Desenvolvimentos Futuros

O projeto estabelece uma base sólida para evoluções ainda mais sofisticadas, como:

**Análise longitudinal e preditiva:** Após a acumulação de dados suficientes, o sistema pode evoluir para identificar tendências de feedback ao longo do tempo, prever a satisfação de turmas futuras com base em características do curso ou do instrutor, e oferecer recomendações proativas de ajustes.

**Personalização do ensino:** Os dados qualitativos podem ser usados para criar perfis de aprendizagem, permitindo que a Esmec personalize recomendações de conteúdo e atividades para grupos específicos de magistrados e servidores.

**Expansão para outros níveis de Kirkpatrick:** A expertise adquirida pode ser aplicada para automatizar a análise de avaliações de aprendizagem (nível 2) e, com as adaptações necessárias, até mesmo de pesquisas de aplicação no trabalho (nível 3), criando um ecossistema integrado de avaliação educacional.

Em conclusão, o redesenho vai muito além de uma nova ferramenta. É a materialização de uma nova cultura de avaliação, mais crítica, ágil e orientada a dados. Sua

	 INSTITUTO BRASILEIRO DE ESTUDOS E PESQUISAS SOCIAIS			 Universidade Potiguar
		 DIREÇÃO-GERAL DA POLÍTICA DE JUSTIÇA	 Instituto de Investigação Interdisciplinar	 Administração da Justiça
	 Grupo de Pesquisa Gestão, Desempenho e Efetividade do Judiciário	 Núcleo de Pesquisa em Informação, Direito e Sociedade	 LINGUAGEM, INSTITUIÇÕES E ORGANIZAÇÕES	



implementação bem-sucedida tem o potencial de servir como um benchmark para outros tribunais e instituições de ensino, demonstrando de forma prática como a inteligência artificial generativa pode ser empregada de forma ética e eficaz para nutrir uma cultura de melhoria contínua e excelência na formação judiciária.

## Referências

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>

Brown, J. D. (2005). *Testing in language programs: A comprehensive guide to English language assessment*. McGraw-Hill.

Carmo, W. C., & Sá, I. J. C. (2024). O impacto das técnicas de linguagem simples e direito visual nas taxas de regressão de regime da execução penal: um experimento do LabLuz. In *Anais da X Enajus*. Enajus. Recuperado em 21 de setembro de 2025, de <https://enajus.org.br/anais/assets/papers/2024/sessao-2/o-impactodas-tecnicas-de-linguagem-simples-e-direito-visual-nas-taxas-de-regressaode-regime-da-execucaopenal-um-experimento-do-labluz.pdf>

Kastrati, Z., Dalipi, F., Imran, A. S., Pireva Nuci, K., & Wani, M. A. (2021). Sentiment analysis of students' feedback using NLP: A systematic review. *IEEE Access*, 9, 165668–165688. <https://doi.org/10.1109/ACCESS.2021.3135>

Kaufman, R., & Keller, J. M. (1994). Levels of evaluation: Beyond Kirkpatrick. *Human Resource Development Quarterly*, 5(4), 371–380. <https://doi.org/10.1002/hrdq.3920050408>

Kirkpatrick, D. L. (1959). Techniques for evaluating training programs. *Journal of the American Society of Training Directors*, 13(11), 3–9.

Nawaz, F., Khushnood, M., & Wisal, W. (2022). Kirkpatrick model and training effectiveness: A meta-analysis 1982 to 2021. *Business & Economic Review*, 14(2), 35–56. <https://doi.org/10.22547/BER/14.2.2>

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2), 1–135. <https://doi.org/10.1561/1500000001>

Paswan, A. K., & Ganesh, G. (2009). Higher education institutions: Satisfaction and loyalty among international students. *Journal of Marketing for Higher Education*, 19(1), 65–84. <https://doi.org/10.1080/0884124090290>

Phillips, J. J. (1996). ROI: The search for best practices. *Training & Development*, 50(2), 42–47.



Saunders, M. N. (2012). Choosing research participants. In G. Symon & C. Cassell (Eds.), *Qualitative organizational research: Core methods and current challenges* (pp. 35–52). Sage.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 57(10), 1380–1400. <https://doi.org/10.1177/0002764213498851>

Sitzmann, T., Brown, K. G., Casper, W. J., Ely, K., & Zimmerman, R. D. (2008). A review and meta-analysis of the nomological network of trainee reactions. *Journal of Applied Psychology*, 93(2), 280–295. <https://doi.org/10.1037/0021-9010.93.2.280>

Tribunal de Justiça do Estado do Ceará (TJCE). (2025). Comissão Permanente de Avaliação Institucional da ESME. Recuperado em 21 de setembro de 2025, de <https://www.tjce.jus.br/comissoes/comissao-permanente-de-avaliacao-institucional-da-esmec-cpai/>

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q. V., & Zhou, D. (2023). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 1–27. <https://doi.org/10.1186/s41239-019-0171-0>

