

SISTEMA DE RECOMENDAÇÃO “ARENA DO SABER”: INTEGRAÇÃO DE MODELAGEM UML, PROCESSOS DE KDD e INTELIGÊNCIA ARTIFICIAL PARA A PERSONALIZAÇÃO PROATIVA DA APRENDIZAGEM

Marcos Antonio Gomes Xavier 

Universidade Federal do Rio de Janeiro - marcosxavier@ufri.br

Bruno Lopes Xavier 

Universidade Federal do Rio de Janeiro - blxavier@firjan.com.br

Sílvia Lima Gonçalves Araújo 

Universidade do Minho - silviauminho03@gmail.com

Marcio Sacramento de Oliveira 

Fundação Oswaldo Cruz - marcio.sacramento@fiocruz.br

Paulo Victor Rodrigues de Carvalho 

Universidade Federal do Rio de Janeiro - paulo.carvalho@ien.gov.br

Resumo: Este artigo descreve o desenvolvimento e a arquitetura técnica do Sistema de Recomendação (SR) da plataforma educacional Arena do Saber. O problema central reside na ineficácia de recomendações genéricas e na opacidade algorítmica "caixa-preta", que falham em engajar o estudante por não explicitar a base pedagógica das sugestões. A solução proposta é um motor proativo de "zero query", que integra processos de Descoberta de Conhecimento em Bases de Dados (KDD) e Inteligência Artificial (IA) para realizar cálculos não lineares em tempo real sobre o desempenho discente. Como pilar de sustentação e qualidade técnica, o projeto utiliza a Linguagem de Modelagem Unificada (UML) para simplificar conceitos complexos em estruturas claras, garantindo uma documentação padronizada e auditável através de diagramas de classe, atividades, sequência e estado. Na etapa de pré-processamento KDD, o sistema calcula indicadores técnicos fundamentais, como o Nível de Dificuldade do Questionário (NDQQ) e o Nível de Dificuldade de Acertos (NDAQ). A lógica de 8 passos permite ao motor identificar lacunas de conhecimento de forma autônoma e sugerir trajetórias personalizadas, ao mesmo tempo em que mitiga o efeito "caixa-preta" ao expor os critérios de decisão através do componente "Análise Passos Grid". Resultados preliminares indicam que essa transparência fortalece o vínculo de confiança do usuário e otimiza a gestão pedagógica, permitindo uma auditoria contínua e a curadoria de conteúdos sob demanda.

Palavras-chave: Sistema de Recomendação; Modelagem UML; Processos de KDD.

"ARENA DO SABER" RECOMMENDATION SYSTEM: INTEGRATION OF UML MODELING, KDD PROCESSES, AND ARTIFICIAL INTELLIGENCE FOR THE PROACTIVE PERSONALIZATION OF LEARNING

Abstract: This article describes the development and technical architecture of the Recommendation System (RS) of the Arena do Saber educational platform. The central problem lies in the ineffectiveness of generic recommendations and algorithmic opacity ("black box"), which fail to engage students by not clarifying the pedagogical basis of the

suggestions. The proposed solution is a proactive "zero query" engine that integrates Knowledge Discovery in Databases (KDD) processes and Artificial Intelligence to perform real-time non-linear calculations on student performance. As a pillar of technical support and quality, the project employs the Unified Modeling Language (UML) to simplify complex concepts into clear structures, ensuring standardized and auditable documentation through class, activity, sequence, and state diagrams. In the pre-processing stage (KDD), the system calculates fundamental technical indicators, such as the Questionnaire Difficulty Level (NDQQ) and the Difficulty Level of Correct Answers (NDAQ). The 8-step logic allows the engine to autonomously identify knowledge gaps and suggest personalized paths while mitigating the "black box" effect by exposing decision criteria through the Analise Passos Grid component. Preliminary results indicate that this transparency strengthens user trust and optimizes pedagogical management, enabling continuous auditing and on-demand content curation.

Keywords: Recommendation System; UML Modeling; KDD Processes.

Introdução

No cenário atual, a personalização da educação por meio de sistemas automatizados enfrenta o desafio da "caixa-preta", em que o estudante recebe sugestões sem compreender os critérios utilizados. Os Sistemas de Recomendação (SR) surgem como ferramentas para sugerir itens relevantes, auxiliando na decisão diante de grandes volumes de dados. A plataforma Arena do Saber diferencia-se por ser um sistema de "zero query", no qual a recomendação é proativa, baseada no histórico de desempenho do estudante e em uma análise transparente de 8 passos baseados em indicadores técnicos.

A arquitetura do motor de 8 passos exige uma coordenação precisa entre os artefatos de *software* para evitar inconsistências que comprometam a jornada do estudante.

Para que essa lógica não linear de 8 passos seja efetiva e compreensível, a robustez da engenharia de *software* por trás do motor é fundamental. Nesse contexto, a utilização da *Unified Modeling Language* (UML) - Linguagem de Modelagem Unificada justifica-se pela sua capacidade de atuar como uma ferramenta valiosa na especificação e documentação de sistemas, simplificando conceitos complexos em estruturas claras.

Para sustentar a complexidade do motor de 8 passos, a UML é adotada como padrão arquitetural, permitindo a especificação clara de requisitos e a redução de falhas estruturais desde as fases iniciais (Bernhardt *et al.*, 2023; Siqueira; Paulon; Guedes, 2019).

No Arena do Saber, essa modelagem colaborativa é essencial para sincronizar o fluxo de dados e mitigar conflitos nos modelos de domínio, assegurando que o processamento de indicadores como o NDQQ e o NDAQ seja traduzido em uma documentação auditável e livre de inconsistências.

Sob essa ótica, Cortiana Neto *et al.* (2019) destacam que a modelagem colaborativa visa reduzir atividades manuais e melhorar a tomada de decisão através da disseminação de novas ideias entre a equipe de desenvolvimento. Tal abordagem corrobora a necessidade de automação e sincronização observada na plataforma Arena do Saber para mitigar conflitos nos modelos de domínio.

Ademais, a inspeção rigorosa dos diagramas de classe é vital para o projeto, visto que inconsistências nesses modelos iniciais podem afetar diretamente a precisão dos cálculos do motor e o código final.

Dessa forma, este trabalho descreve o Sistema de Recomendação do Arena do Saber, demonstrando como uma fundação arquitetural sólida, documentada via UML, permite transformar dados brutos de interações em um guia prático, transparente e assertivo para a evolução acadêmica do discente.

Revisão de Literatura

Sistemas de Recomendação

Os sistemas de recomendação (SR) atuam como mediadores entre os ambientes informacionais e os interesses específicos dos sujeitos. De acordo com Alvarez *et al.* (2016), a integração desses sistemas à arquitetura da informação potencializa a encontrabilidade, facilitando a descoberta de dados pertinentes que o usuário poderia não localizar por meio de buscas tradicionais.

Nesse contexto, os sistemas de recomendação consolidam-se como soluções cruciais para mitigar o fenômeno da sobrecarga de informação (*information overload*). Ao apontar o usuário para itens novos e ainda não experimentados que sejam pertinentes à sua tarefa atual, o SR atua como um guia que navega por diversos tipos de conhecimento sobre os usuários e itens disponíveis, superando a barreira da exaustão informacional (Ricci; Rokach; Shapira, 2011).

No âmbito educacional, especificamente na área de *Technology Enhanced Learning* (TEL), essa mediação assume um papel estratégico na personalização do processo de aprendizagem. O objetivo central é apoiar estudantes e professores na identificação de recursos didáticos adequados em meio a uma oferta vasta e complexa de objetos de aprendizagem, garantindo que o discente navegue de forma guiada para atingir seus objetivos pedagógicos específicos (Drachsler *et al.*, 2015). O Sistema de Recomendação (SR) do Arena do Saber, tem uma abordagem baseada em conteúdo, onde mapeia o perfil do estudante

através de suas ações (notas e acertos), comparando-as com palavras-chave e níveis de dificuldade dos itens pedagógicos. Esta técnica garante a independência do usuário, permitindo recomendações mesmo que o estudante tenha um perfil "exótico" ou único.

Ao adotar o modelo de "*zero query*", o Arena do Saber transforma a navegação em uma experiência de descoberta proativa, guiando o estudante diretamente ao conteúdo necessário para sanar suas lacunas de conhecimento.

A escolha da filtragem baseada em conteúdo para o Arena do Saber sustenta-se na premissa de que o conteúdo textual e os atributos dos itens pedagógicos possuem informações intrínsecas capazes de caracterizar as necessidades do estudante.

Segundo Silva (2014), este tipo de dado não estruturado é amplamente utilizado em Mineração de Dados, garantindo a independência do usuário, o que permite que novos materiais didáticos sejam recomendados sem a necessidade de avaliações prévias por outros estudantes. Essa abordagem é essencial para sistemas educacionais que buscam evitar que itens recém-incluídos fiquem invisíveis ao público.

O motor de recomendação incorpora variáveis contextuais (sistemas sensíveis ao contexto), como o Nível de Dificuldade do Questionário (NDQQ), o Nível de Dificuldade de Acertos (NDAQ) e o histórico de tentativas, para prever a utilidade de um novo conteúdo.

O seu processo segue as etapas de *Knowledge Discovery in Databases* (KDD) - Descoberta de Conhecimento em Bases de Dados, pré-processamento dos dados de acerto, análise por meio de classificação em tipos de recomendação e interpretação visual para o estudante.

Segundo Steiner *et al.* (2006) o conhecimento descoberto deve possuir três propriedades fundamentais: ser o mais preciso possível, ser compreensível para humanos e ser interessante, útil ou novo.

De acordo com o modelo de Fayyad *et al.* (1996 *apud* Steiner *et al.*, 2006), o processo de KDD é composto por cinco etapas interdependentes: Seleção - Escolha dos dados relevantes para o problema; Pré-processamento e Limpeza - Tratamento de ruídos e de dados inconsistentes; Transformação - Conversão dos dados em formatos adequados à mineração; Mineração de Dados (*Data Mining*) - Aplicação de técnicas para extrair padrões; Interpretação e Avaliação - Análise dos resultados obtidos para validar o conhecimento.

As técnicas de mineração desempenham, majoritariamente, três tarefas: classificação, agrupamento e descoberta de regras de associação. Entre as principais técnicas citadas para o reconhecimento de padrões e classificação (Steiner *et al.*, 2006), destacam-se: Árvores de Decisão -

Notáveis pela alta compreensibilidade e facilidade de interpretação; Redes Neurais - Conhecidas pela acurácia e robustez ao lidar com ruídos, embora careçam de facilidade de explicação imediata (modelo "caixa-preta"); Algoritmos Genéticos e Metaheurísticas - Utilizados para otimização e descoberta de regras; Máquinas de Suporte de Vetores (SVM) e Métodos Estatísticos - como a regressão logística e análise discriminante linear.

Diagramas de UML

A Linguagem de Modelagem Unificada (UML) permanece como um padrão indispensável na Engenharia de Software contemporânea, sendo essencial para garantir a eficiência no desenvolvimento de sistemas de alta qualidade (Bernhardt *et al.*, 2023).

O uso de ferramentas de modelagem colaborativa visa reduzir atividades manuais e melhorar a tomada de decisão através da disseminação de novas ideias entre os membros da equipe. Tal abordagem corrobora a necessidade de automação e sincronização observada na plataforma para mitigar inconsistências nos modelos de domínio (Cortiana Neto *et al.*, 2019).

A verificação e validação de artefatos UML auxiliam na garantia da qualidade e reduzem custos, uma vez que erros que passariam para os próximos passos seriam identificados e corrigidos nos estágios iniciais do desenvolvimento (Siqueira; Paulon; Guedes, 2019).

A utilização da UML no projeto do Arena do Saber justifica-se pela sua capacidade de atuar como uma ferramenta valiosa na especificação e documentação de sistemas, simplificando conceitos complexos em estruturas claras.

A adoção dos diagramas de UML no projeto está alinhada com as práticas mais recorrentes na indústria e na academia. A UML emerge como uma linguagem gráfica padrão que abrange desde sistemas corporativos até aplicações complexas de tempo real, consolidando-se como um recurso crucial para orientar o processo de desenvolvimento e a comunicação eficaz entre os *stakeholders*.

Metodologia

A metodologia de processamento do motor de recomendação do Arena do Saber segue um ciclo de vida estruturado que mimetiza as etapas clássicas de Descoberta de Conhecimento em Bases de Dados (KDD), transformando *logs* - registros de interação em conhecimento pedagógico útil.

Para garantir a integridade desta arquitetura, a modelagem via UML é utilizada como ferramenta essencial para a especificação e documentação, simplificando conceitos complexos em estruturas claras e orientando o processo de desenvolvimento de alta qualidade. O fluxo operacional é detalhado através de dois artefatos dinâmicos principais:

1. Diagrama de Atividades: Este diagrama descreve o comportamento dinâmico do sistema desde o gatilho inicial até a entrega da orientação personalizada. Ele ilustra as fases de pré-processamento e análise do modelo KDD, onde o motor calcula em tempo real os indicadores NDQQ e NDAQ. A precisão deste fluxo é garantida por técnicas de inspeção que identificam erros e inconsistências precocemente, reduzindo custos de desenvolvimento e assegurando que as conclusões pedagógicas sejam extraídas de forma isenta de contradições.

A seguir são detalhadas as 5 etapas de Fayyad *et al.* (1996 *apud* Steiner *et al.*, 2006) no fluxo prático do Sistema de Recomendação Arena do Saber:

a. Seleção (Selection)

Esta etapa inicial consiste na escolha dos dados relevantes para o problema pedagógico. No Arena do Saber, a seleção ocorre no momento em que o estudante finaliza uma tentativa em um questionário, disparando a coleta dos logs de interação, que incluem o registro bruto de acertos, erros e o histórico do discente.

b. Pré-processamento e Limpeza (Pre-processing/Cleaning)

Nesta fase, tratam-se ruídos e dados inconsistentes para garantir a qualidade da análise. No fluxo prático, isso é ilustrado no Diagrama de Atividades, onde o sistema verifica a integridade dos dados de entrada e inicia o cálculo dos indicadores técnicos fundamentais, como o Nível de Dificuldade do Questionário (NDQQ) e o Nível de Dificuldade de Acertos (NDAQ).

c. Transformação (Transformation)

A transformação envolve a conversão dos dados brutos em formatos adequados para a mineração. No Arena do Saber, o motor converte os acertos e erros em variáveis contextuais e indicadores de desempenho, como o Nível de Dificuldade do Tema (NDT) e a Média das questões incorretas (NDQ), preparando o terreno para a lógica de decisão.

d. Mineração de Dados (Data Mining)

É a aplicação de técnicas para extrair padrões, como classificação ou descoberta de regras. No sistema, esta etapa corresponde à execução da lógica de 8 passos e aos cálculos não lineares realizados em tempo real. O motor classifica o desempenho do estudante em um

dos seis tipos de recomendação (Tipos 0 a 5), identificando autonomamente lacunas de conhecimento. Os seis tipos de recomendação são:

- Tipo 0 (Refazer): Indicado quando o estudante obtém um índice de acertos entre 60% e 89%. O objetivo é consolidar o conteúdo do questionário atual.
- Tipo 1 (Reforço por Questionários): Aplicado em casos que não se enquadram nas categorias de avanço ou revisão profunda, focando em sugestões de novos questionários do mesmo tema.
- Tipo 2 (Reforço por Questões): Semelhante ao tipo 1, mas a recomendação foca em questões específicas do mesmo tema para sanar lacunas pontuais.
- Tipo 3 (Avançar): Gerado quando o estudante atinge 100% de acertos em questionários que não são classificados como "Fáceis" (NDQQ $\neq$ Fácil), direcionando-o para conteúdos mais avançados.
- Tipos 4 e 5 (Revisão Fundamental): Disparados quando o desempenho é inferior a 30% de acertos. O sistema orienta o estudante a retomar conceitos fundamentais do tema antes de prosseguir.

Essas recomendações são persistidas no banco de dados e exibidas ao usuário através do componente Analise Passos Grid, que expõe a lógica técnica por trás de cada escolha para garantir a transparência do sistema.

Caso o motor identifique uma necessidade de reforço ou revisão para a qual não existam materiais cadastrados, ele registra automaticamente uma solicitação de conteúdo para a equipe docente.

e. Interpretação e Avaliação (Interpretation/Evaluation)

A fase final valida se o conhecimento descoberto é interessante, útil e compreensível para o usuário. No Arena do Saber, isso é materializado através do componente "Analise Passos Grid", que expõe visualmente os critérios de decisão ao estudante, mitigando o efeito "caixa-preta" e permitindo que ele valide a trajetória de estudo sugerida.

2. Diagrama de Sequência: Para compreender a dinâmica cronológica das interações, este artefato detalha o fluxo de mensagens entre o estudante, o motor e o banco de dados. Ele mapeia a transição dos dados de entrada até a persistência dos resultados, processo vital para a transparência do sistema. Ao expor a ordem das operações, o diagrama auxilia na mitigação do efeito "caixa-preta", permitindo que o estudante valide a lógica que sustenta sua jornada de aprendizagem.

A execução do motor de 8 passos, portanto, não é apenas uma sequência de regras fixas, mas um processo de modelagem colaborativa que visa reduzir atividades manuais e melhorar a tomada de decisão pedagógica. Através desta fundamentação técnica e gráfica, o sistema consegue identificar lacunas de conhecimento de forma autônoma, garantindo que o conhecimento descoberto seja preciso, compreensível e útil para o usuário final.

Detalhamento da Lógica de 8 Passos do Motor de Recomendação

O motor de recomendação proativo da Arena do Saber processa o desempenho discente através de uma análise estruturada em oito indicadores técnicos, que fundamentam a transição entre o processamento de dados brutos e a geração de uma orientação pedagógica transparente. Esta lógica é persistida em formato JSONB no banco de dados e exposta ao usuário por meio do componente Analise Passos Grid, detalhando as seguintes métricas:

Nível de Dificuldade do Questionário (NDQQ): Representa a média da dificuldade teórica (NDQ) de todas as questões que compõem o questionário aplicado.

Nível de Dificuldade de Acertos (NDAQ): Um índice empírico e global que reflete o percentual de acertos de todos os usuários da plataforma para aquele conjunto de itens, servindo como termômetro de complexidade real.

Índice de Acertos: O desempenho percentual individual do estudante na tentativa atual, classificado em faixas que variam de "Revisão Fundamental" (abaixo de 30%) a "Avançar Nível" (acima de 90%).

Incorretas: Contagem absoluta e proporcional de questões erradas, utilizada para identificar o volume de lacunas imediatas.

Nível de Dificuldade do Tema (NDT): Média dos NDQs de todas as questões vinculadas ao tema pedagógico, permitindo avaliar se a dificuldade é pontual ou estrutural no tópico.

NDQ Média (Incorretas): Calcula a dificuldade teórica média especificamente das questões que o estudante errou, refinando o diagnóstico sobre o nível de desafio que causou a falha.

Total de Alunos Respondentes (TAR): Métrica de confiabilidade estatística que indica o volume da amostra (n). Quando o TAR é inferior a 100, o sistema sinaliza visualmente que os indicadores NDAQ e NDQ Média possuem menor precisão estatística.

Tipo de Questão: Analisa metadados sobre a natureza das questões (fechadas ou abertas) e a quantidade de alternativas, influenciando o cálculo de probabilidade de acerto casual.

Essa estrutura permite que o motor execute uma lógica de decisão não linear: se o desempenho for insatisfatório em temas de alta dificuldade (NDT/NDQQ ≠ Fácil), o sistema prioriza o "Reforço" ou "Revisão"; se houver domínio pleno (>90% de acertos), o fluxo direciona para o "Avanço de Nível"⁹. Caso o motor identifique uma necessidade de conteúdo não disponível no acervo, ele registra automaticamente uma "Solicitação de Conteúdo" para curadoria docente.

Apresentação de Resultados

A apresentação e análise dos resultados obtidos no Sistema de Recomendação (SR) da Plataforma Arena do Saber representam a fase final e crítica do processo de KDD. Conforme estabelecido na literatura, para que o conhecimento extraído de grandes bases de dados seja considerado valioso, ele deve possuir as propriedades de ser correto, interessante e, sobretudo, compreensível por usuários humanos.

Neste sentido, a utilização de diagramas UML como ferramenta de especificação permitiu que conceitos complexos fossem simplificados em estruturas claras, garantindo a eficiência e a alta qualidade dos resultados apresentados.

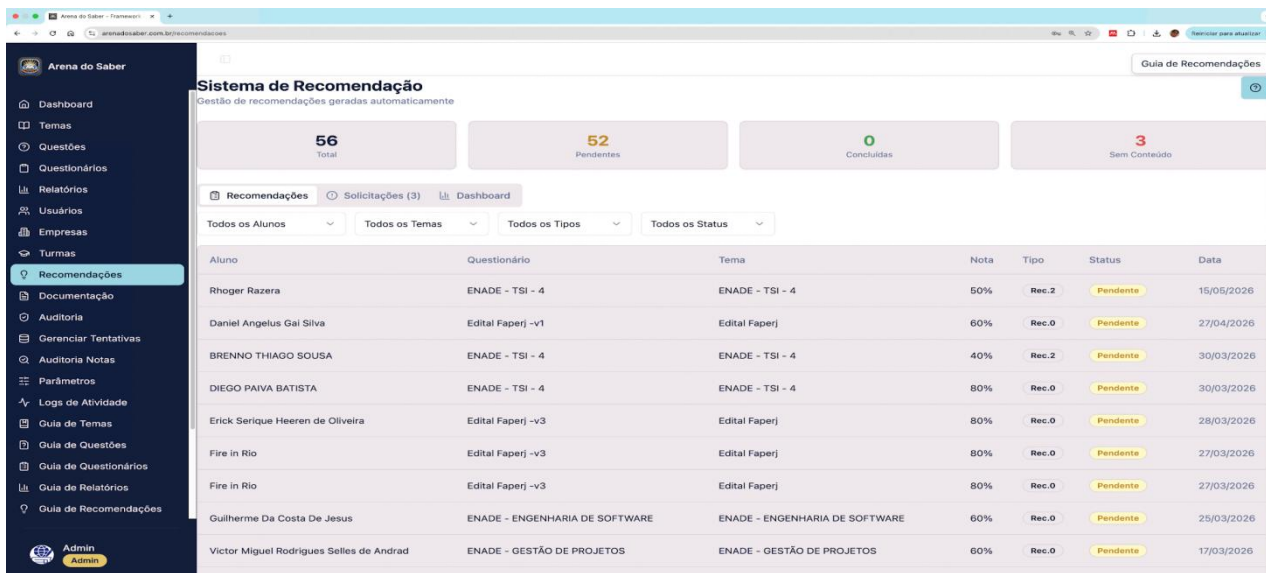
Os principais achados e funcionalidades validadas são detalhados a seguir:

- **Eficácia da Lógica Proativa:** O motor transforma registros brutos de interação em diagnósticos pedagógicos acionáveis de forma autônoma (*modelo zero query*). A eficácia desta abordagem é monitorada através de métricas quantitativas de desempenho, que incluem o volume de recomendações geradas, o tempo médio de resposta para os cálculos não lineares em tempo real e a taxa de engajamento (conversão de sugestões do estado 'pendente' para 'em andamento'). A precisão na identificação de lacunas é validada preliminarmente pela frequência de solicitações automáticas de conteúdo e pela consistência dos indicadores frente ao volume da amostra (TAR - Total de Alunos Respondentes), permitindo que o sistema supere a rigidez de modelos lineares com base em evidências empíricas de utilidade e performance.
- **Transparência e Mitigação da "Caixa-Preta":** Um resultado central foi a validação do componente "Análise Passos Grid". Ao expor os indicadores técnicos (NDQQ, NDAQ, TAR etc.) por meio de uma interface auditável, o sistema mitigou o efeito de "caixa-preta" típico de algoritmos de IA tradicionais. Essa visibilidade fortalece o vínculo de confiança, pois o

estudante passa a aceitar as sugestões para compreender como seu desempenho influencia a ação proposta.

- **Otimização da Gestão Pedagógica:** A inspeção rigorosa dos modelos de dados garante que o sistema seja capaz de gerenciar não apenas o que o estudante deve estudar, mas o que a plataforma precisa criar. A geração de solicitações automáticas de conteúdo quando o motor detectar lacunas permite uma curadoria sob demanda, reduzindo drasticamente o trabalho manual e os custos operacionais da equipe docente.
- **Colaboração e Qualidade Técnica:** A utilização de práticas de modelagem colaborativa durante o desenvolvimento, reflete em um sistema com menor número de conflitos técnicos e maior produtividade. A aplicação de técnicas de inspeção nos diagramas de classe assegura que erros estruturais fossem corrigidos nos estágios iniciais, resultando em um código final robusto que sustenta a precisão dos indicadores de desempenho apresentados nas interfaces gerenciais nas Abas: Recomendações, Solicitações e Dashboard, ilustrada na Figura 1.

Figura 1 - Captura da Tela da Aba Recomendações - Perfil - Administrador e Professor



Sistema de Recomendação
Gestão de recomendações geradas automaticamente

56 Total 52 Pendentes 0 Concluídas 3 Sem Conteúdo

Recomendações Solicitações (3) Dashboard

Todos os Alunos Todos os Temas Todos os Tipos Todos os Status

Aluno	Questionário	Tema	Nota	Tipo	Status	Data
Rhoger Razera	ENADE - TSI - 4	ENADE - TSI - 4	50%	Rec.2	Pendente	15/05/2026
Daniel Angelus Gai Silva	Editais Faperj -v1	Editais Faperj	60%	Rec.0	Pendente	27/04/2026
BRENNO THIAGO SOUSA	ENADE - TSI - 4	ENADE - TSI - 4	40%	Rec.2	Pendente	30/03/2026
DIEGO PAIVA BATISTA	ENADE - TSI - 4	ENADE - TSI - 4	80%	Rec.0	Pendente	30/03/2026
Erick Serique Heeren de Oliveira	Editais Faperj -v3	Editais Faperj	80%	Rec.0	Pendente	28/03/2026
Fire in Rio	Editais Faperj -v3	Editais Faperj	80%	Rec.0	Pendente	27/03/2026
Fire in Rio	Editais Faperj -v3	Editais Faperj	80%	Rec.0	Pendente	27/03/2026
Guilherme Da Costa De Jesus	ENADE - ENGENHARIA DE SOFTWARE	ENADE - ENGENHARIA DE SOFTWARE	60%	Rec.0	Pendente	25/03/2026
Victor Miguel Rodrigues Selles de Andrad	ENADE - GESTÃO DE PROJETOS	ENADE - GESTÃO DE PROJETOS	60%	Rec.0	Pendente	17/03/2026

Fonte: Próprio autor (2026).

A fundamentação técnica padronizada via UML não apenas orienta o desenvolvimento, mas torna-se um padrão para a análise e auditoria contínua do sucesso acadêmico dos estudantes.

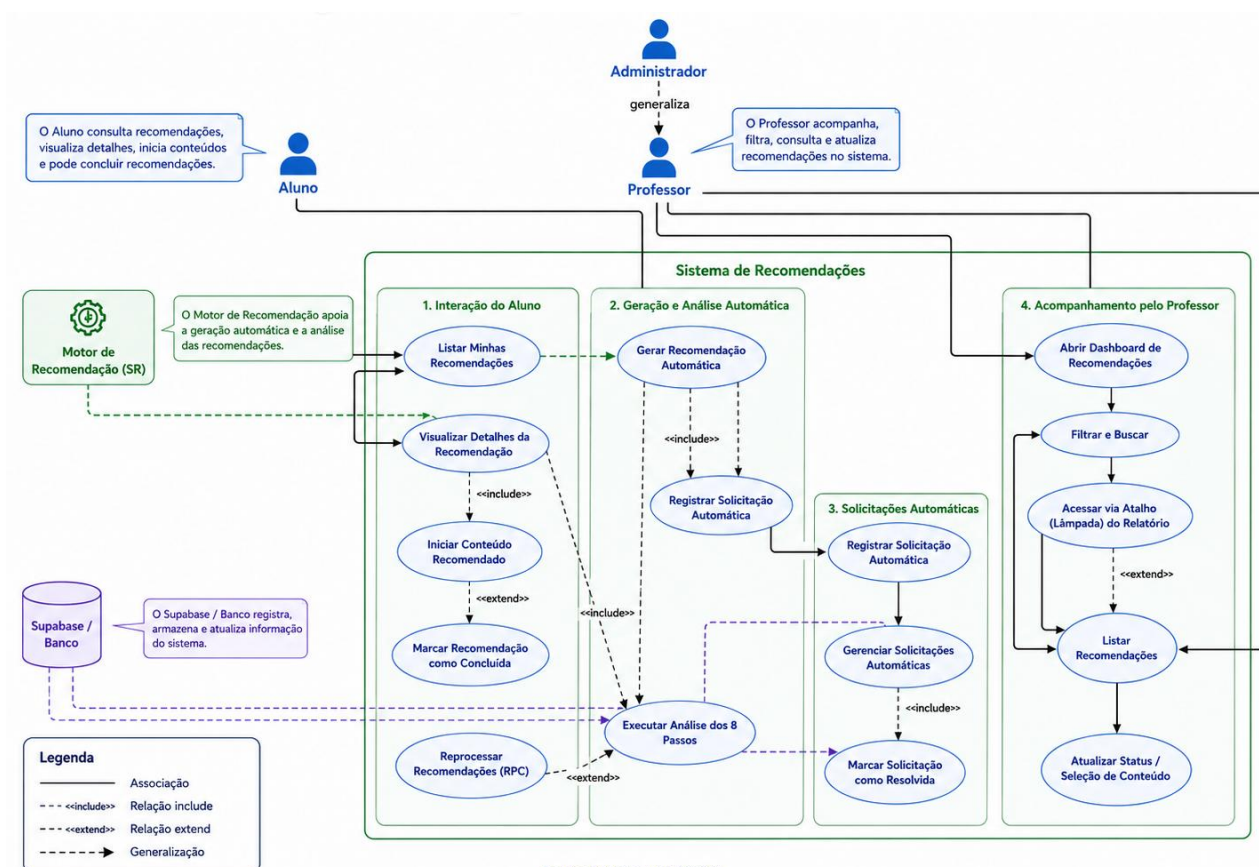
Diagramas UML de Recomendações Arena do Saber

Diagrama de UC

Para detalhar a arquitetura funcional e as interações entre os diversos atores que compõem o ecossistema da plataforma Arena do Saber, é apresentado o Diagrama de Casos de Uso (UC) (ilustrado na Figura 2). Este artefato UML permite visualizar como o motor de recomendação proativo se integra às rotinas de estudantes, professores e administradores, mapeando desde a execução automática da análise de 8 passos até a gestão de solicitações de novos conteúdos pedagógicos.

O diagrama ilustra o papel central do motor como um agente de automação que interage diretamente com a base de dados para persistir resultados e registrar lacunas, ao mesmo tempo em que destaca as funcionalidades de autogestão da aprendizagem oferecidas ao discente e as ferramentas de auditoria técnica e controle de fluxo disponibilizadas ao corpo docente e administrativo.

Figura 2 - Diagrama UC - SR Arena do Saber



Fonte: Próprio autor (2026).

Fonte: Próprio autor (2026).

Diagrama de Classe

Para representar a estrutura estática e a modelagem de dados que sustenta o motor de recomendação, é apresentado o Diagrama de Classe (Figura 3). Este artefato detalha as entidades fundamentais do ecossistema, como a Tentativa, que captura as ações e o desempenho bruto do estudante, e sua relação direta com a entidade Recomendação, na qual os resultados processados pela análise de 8 passos são persistidos.

O diagrama também mapeia a classe “Solicitação Automática”, que é disparada quando o sistema identifica lacunas pedagógicas, ou seja, quando o motor detecta a necessidade de reforço, mas não encontra conteúdos adequados já cadastrados.

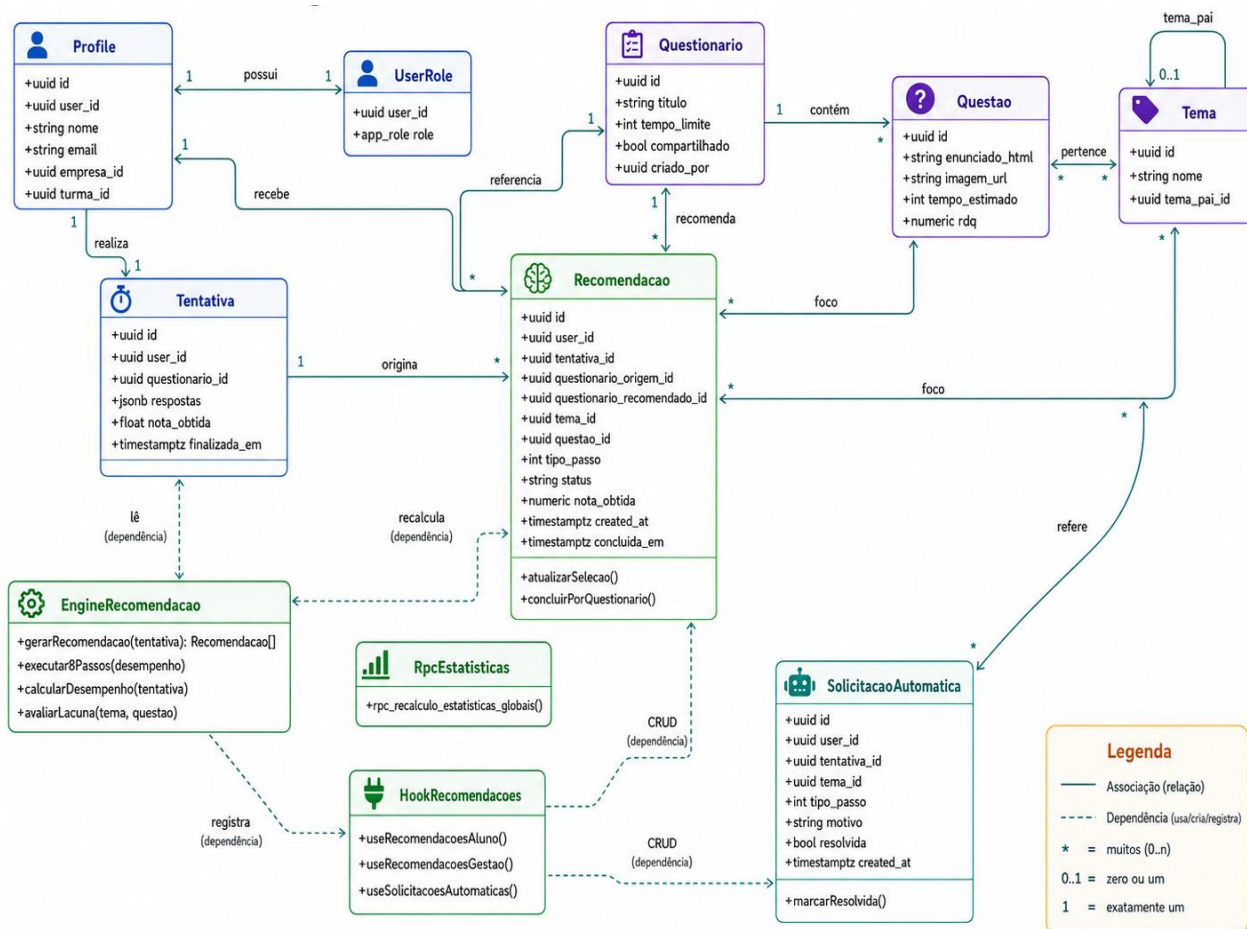
Essa modelagem garante que o sistema seja capaz de gerenciar não apenas o que o estudante deve estudar, mas também o que a plataforma precisa criar para atender à jornada de aprendizagem personalizada.

O Diagrama de Atividades ilustra o fluxo operacional e o encadeamento lógico das decisões tomadas pelo sistema. O processo é iniciado assim que o estudante finaliza uma tentativa no questionário, o que dispara o cálculo imediato de indicadores como NDQQ e NDAQ.

O diagrama destaca o ponto de decisão crítico em que o motor avalia a disponibilidade de recursos pedagógicos no banco de dados. Caso existam materiais compatíveis com o nível de desempenho identificado, o fluxo segue para a geração de uma recomendação específica (variando entre os Tipos 0 e 5); caso contrário, o sistema registra automaticamente uma lacuna de conteúdo para curadoria administrativa.

Esse fluxo mimetiza as etapas de KDD (Descoberta de Conhecimento), garantindo que cada *log* de interação seja transformado em um direcionamento assertivo e transparente ao discente.

Figura 3 - Diagrama de Classes - SR Arena do Saber



Fonte: Próprio autor (2026).

Diagrama de Atividades

Para detalhar o fluxo operacional e o encadeamento lógico das decisões tomadas pelo motor de recomendação proativo, é apresentado o Diagrama de Atividades (Figura 4).

Este artefato UML descreve o comportamento dinâmico do sistema desde o gatilho inicial, a finalização de um questionário pelo estudante, até a entrega da orientação personalizada.

O diagrama ilustra como o motor processa, em tempo real, os indicadores técnicos fundamentais, como o NDQQ e o NDAQ, mimetizando as etapas de pré-processamento e de análise do modelo KDD.

Um ponto de decisão crítico destacado neste fluxo é a verificação da disponibilidade de recursos pedagógicos: caso existam materiais compatíveis, o sistema gera uma recomendação classificada entre os Tipos 0 e 5; caso contrário, o motor registra de forma autônoma uma solicitação de conteúdo. Essa representação visual demonstra como a arquitetura de "zero query" transforma os logs de interação em caminhos assertivos para a jornada de aprendizagem.

Diagrama de Sequência

Para compreender a dinâmica operacional e a ordem cronológica das interações no ecossistema Arena do Saber, é apresentado o Diagrama de Sequência (Figura 5).

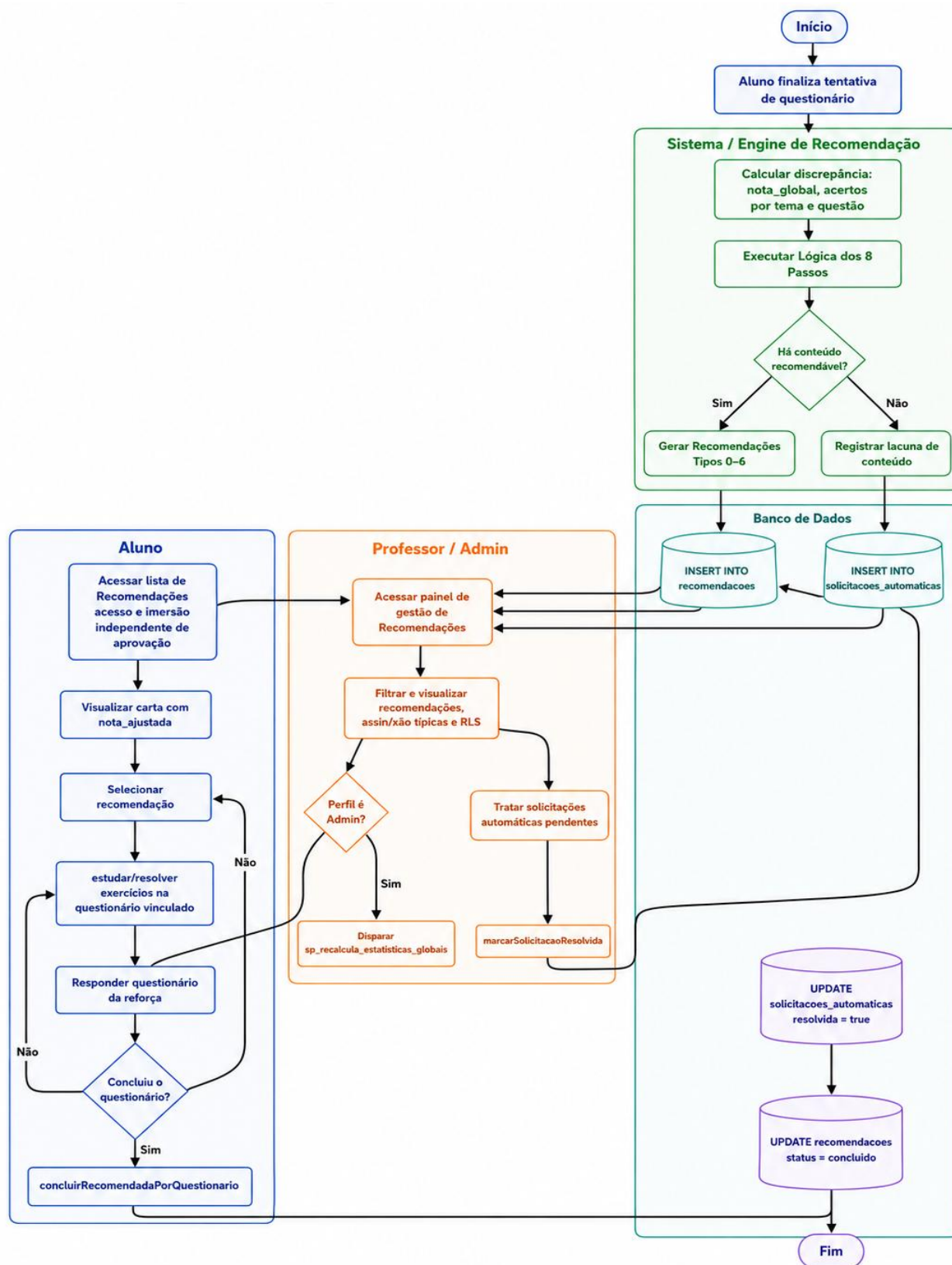
Este artefato detalha o fluxo de mensagens que ocorre desde o gatilho inicial, a finalização de um questionário pelo Estudante, até a entrega da recomendação personalizada.

O diagrama ilustra como o Motor de Recomendação atua como orquestrador central, interagindo com o Banco de Dados para realizar o mapeamento dos dados de entrada e o processamento de indicadores técnicos fundamentais, como o NDQQ e o NDAQ.

Através desta representação, observa-se a execução da lógica de 8 passos e a posterior persistência do resultado na tabela de recomendações.

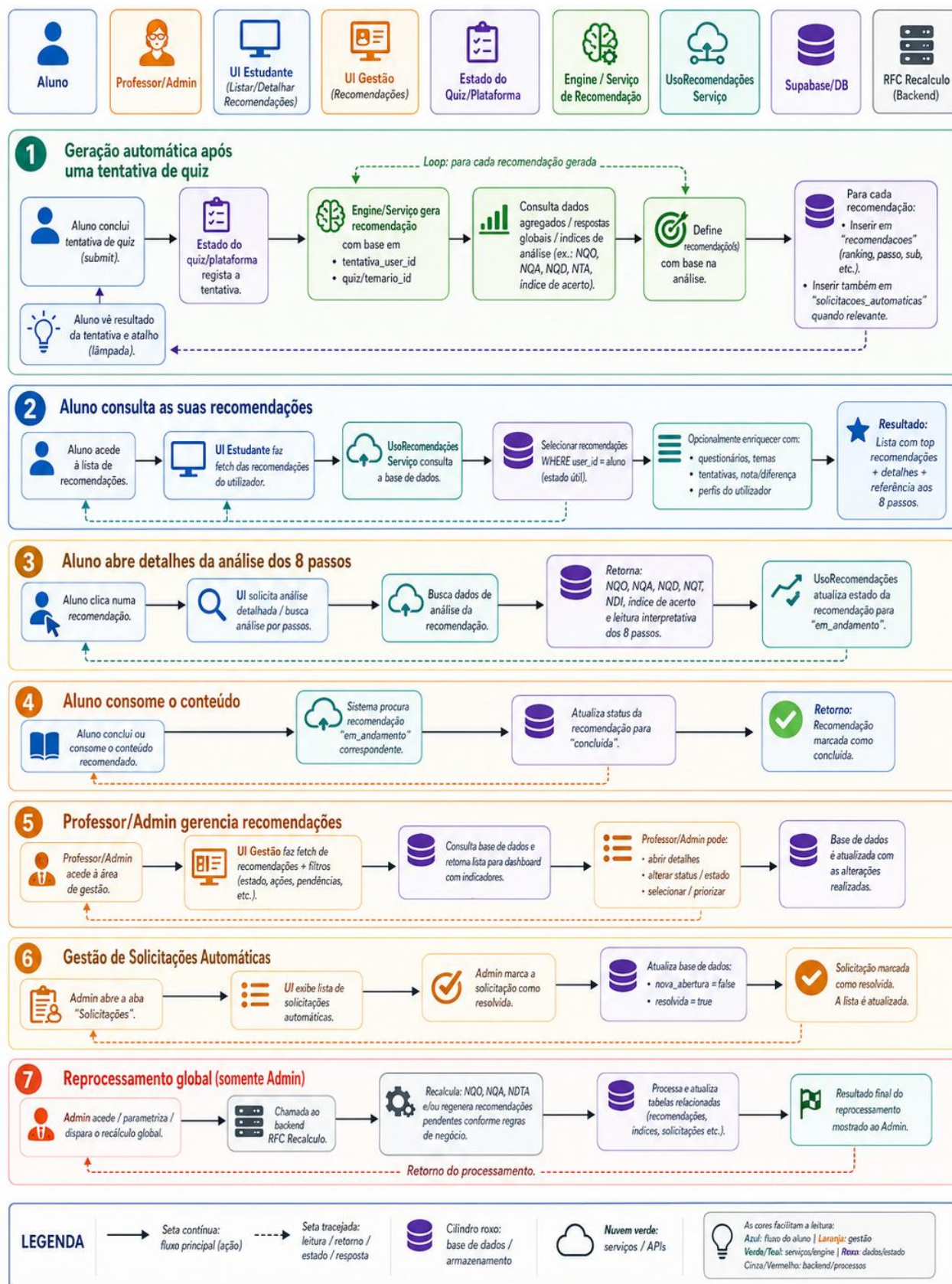
Esse fluxo culmina na atualização da interface visual, na qual o sistema expõe os critérios utilizados para mitigar o efeito "caixa-preta", garantindo que o estudante compreenda e valide a trajetória de estudo sugerida.

Figura 4 - Diagrama de Atividades - SR Arena do Saber



Fonte: Próprio autor (2026).

Figura 5 - Diagrama de Sequência - SR Arena do Saber



Fonte: Próprio autor (2026).

Diagrama de Estado

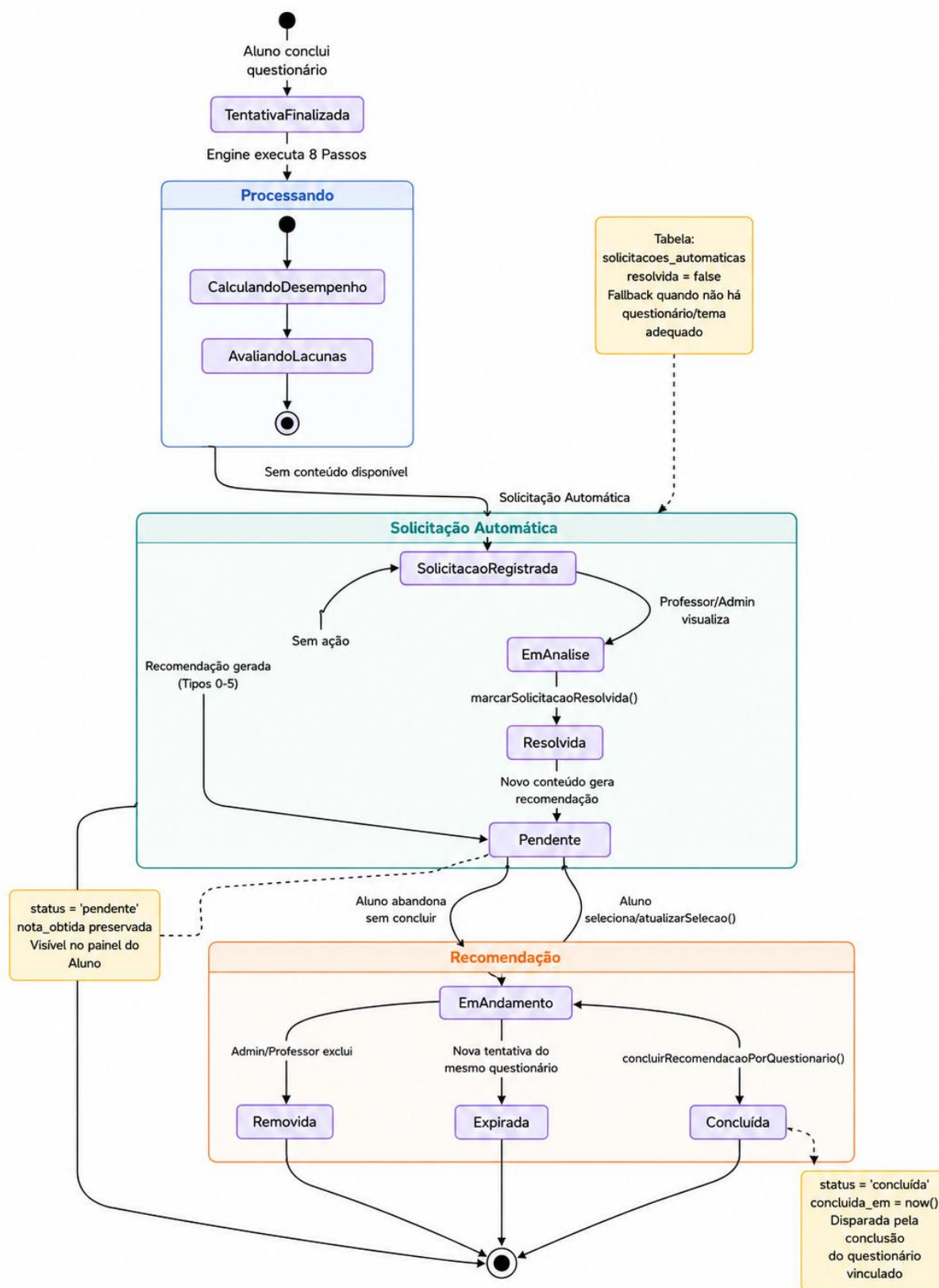
O Diagrama de Estado modela o comportamento dinâmico de uma recomendação, acompanhando sua evolução desde o gatilho inicial até o encerramento do ciclo de vida (ilustrado na Figura 6).

Este fluxo garante que o sistema mantenha a integridade da jornada de aprendizagem, gerenciando estados de erro, pendências e conclusões de forma automatizada. O ciclo é composto pelas seguintes fases principais:

1. **Processamento e Avaliação:** O ciclo inicia quando o estudante conclui um questionário, o que dispara o motor que alterna entre os estados de Calculando Desempenho e Avaliando Lacunas. Se não houver conteúdo disponível para o nível identificado, o sistema entra no estado de “Solicitação Registrada”, aguardando a curadoria pedagógica para transitar para “Em Análise” e, finalmente, “Resolvida”.
2. **Estado Pendente:** Uma vez gerada a recomendação (Tipos 0-5), ela assume o *status* 'Pendente', tornando-se visível no painel do estudante com a nota obtida preservada. Neste estágio, a recomendação aguarda a interação do usuário ou pode ser removida por um administrador ou professor.
3. **Execução e Engajamento:** Quando o estudante inicia o conteúdo sugerido, a recomendação passa para o estado “Em Andamento”. Caso o estudante abandone a atividade sem concluí-la, o sistema reverte o status para “Pendente”. Se uma nova tentativa do questionário original for iniciada antes da conclusão da sugestão, o estado atual é marcado como “Expirado”.
4. **Finalização:** O ciclo se encerra com sucesso no estado “Concluída”, disparado pela finalização do questionário vinculado, momento em que o sistema registra o “*timestamp*” exato da conclusão.

Esta modelagem permite que o sistema gerencie de forma inteligente a retenção e o progresso do estudante, garantindo que cada recomendação tenha um rastreamento claro e auditável do status do estudante.

Figura 6 - Diagrama de Estado - SR Arena do Saber



Fonte: Próprio autor (2026).

Considerações Finais

O Sistema de Recomendação da Plataforma Arena do Saber transforma diagnósticos pedagógicos em ações práticas, demonstrando que a integração de técnicas de *Data Mining* e filtragem baseada em conteúdo cria um ambiente de aprendizagem robusto e personalizado.

A principal contribuição deste trabalho reside na transição de sistemas tradicionais para um modelo de "zero query" que prioriza a transparência: ao expor os 8 passos da análise, o sistema mitiga o efeito de "caixa-preta" das IAs convencionais, construindo uma relação de confiança fundamental entre o estudante e a tecnologia educacional.

Para sustentar essa complexidade técnica, a adoção da Linguagem de Modelagem Unificada (UML) foi um pilar essencial no projeto por sua capacidade de atuar como uma ferramenta valiosa na especificação e documentação, simplificando conceitos complexos em estruturas claras e garantindo a eficiência no desenvolvimento de sistemas de alta qualidade.

O uso de diagramas como Caso de Uso, Classe, Atividades, Sequência e Estado permitiu uma visão padronizada que orientou o processo de desenvolvimento e a comunicação eficaz entre os *stakeholders*.

Além disso, a eficiência operacional do motor de 8 passos foi potencializada pelo uso de práticas de modelagem colaborativa, que visam reduzir atividades manuais e melhorar a tomada de decisão através da disseminação de novas ideias e sincronização de atividades.

Paralelamente, a aplicação de técnicas de inspeção e validação nos artefatos UML, especialmente nos diagramas de classe, garantiu a integridade do sistema ao identificar inconsistências precocemente, assegurando que os cálculos não lineares realizados em tempo real refletissem fielmente a lógica pedagógica pretendida.

Em suma, o projeto Arena do Saber demonstra que a evolução dos sistemas educacionais depende da união entre a personalização inteligente da IA e uma fundação arquitetural sólida e padronizada. Ao mimetizar o funcionamento do cérebro humano, amparado por uma documentação técnica auditável, o sistema oferece caminhos de estudo assertivos que escalam a personalização de forma autônoma.

Trabalhos futuros devem focar na validação empírica da transparência algorítmica por meio de protocolos com usuários e escalas de utilidade percebida, através de métricas quantitativas de recomendação - Recall@k, visando comprovar cientificamente a eficácia da plataforma na mitigação do efeito 'caixa-preta'.

Agradecimentos

A FAPERJ - Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, da qual autor principal deste artigo possui um vínculo (Editais FAPERJ Nº 24/2024, 25/2024 e 23/2025) por ser o proponente/orientador no projeto da pesquisa. O apoio da bolsa da FAPERJ para os bolsistas está sendo fundamental para estimular principalmente o envolvimento na implementação do projeto, em especial nas atividades atreladas ao processo de inovação. O presente trabalho também foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Referências

- ALVAREZ, Edgar Bisset *et al.* Os Sistemas de Recomendação, Arquitetura da Informação e a Encontrabilidade da Informação. **TransInformação**, Campinas, v. 28, n. 3, p. 275-286, set./dez. 2016. DOI: <http://dx.doi.org/10.1590/2318-08892016000300003>
- ARENA DO SABER. Site. Disponível em: <https://arenadosaber.com.br/>. Acesso em: mai. 2026.
- BERNHARDT, Gustavo Henrique *et al.* A utilização da linguagem de modelagem unificada (UML) na engenharia de software: uma avaliação da adoção e relevância. **Anais do 21º Encontro Científico Cultural Interinstitucional**, Cascavel: Centro Universitário FAG, 2023. ISSN 1980-7406. Disponível em: <https://www4.fag.edu.br/anais-2023/Anais-2023-48.pdf> Acesso em: mai. 2026.
- CONTRERAS, Bruno de Santana Braga; DIGIAMPIETRI, Luciano Antonio. Avanços nos sistemas de recomendação de artigos científicos: Uma revisão sistemática da literatura. **Revista Brasileira de Computação Aplicada**, Passo Fundo, v. 17, n. 2, p. 78-86, ago. 2025. DOI: <https://doi.org/10.5335/rbca.v17i2.16376>.
- CORREIA, Raissa C. **Revisão de técnicas de Ciência de Dados para Sistemas de Recomendação**. Campinas: UNICAMP, 2019. (Relatório Técnico - IC-PFG-19-40). Disponível em: <https://ic.unicamp.br/~reltech/PFG/2019/PFG-19-40.pdf> Acesso em: mai. 2026.
- CORTIANA NETO, Raul Antonio *et al.* Supporting collaborative modelling in UML class diagrams. **Revista Brasileira de Computação Aplicada**, [S. l.], v. 11, n. 3, p. 39–46, 2019. DOI: [10.5335/rbca.v11i3.8469](https://doi.org/10.5335/rbca.v11i3.8469). Disponível em: <https://ojs.upf.br/index.php/rbca/article/view/8469>. Acesso em: 23 maio. 2026.
- DRACHSLER, H. *et al.* Panorama of Recommender Systems to Support Learning. In: RICCI, F.; ROKACH, L.; SHAPIRA, B. (eds.). **Recommender Systems Handbook**. 2. ed. Boston, MA: Springer, 2015. p. 421-451. DOI: [10.1007/978-1-4899-7637-6_12](https://doi.org/10.1007/978-1-4899-7637-6_12)
- IFSULDEMINAS. **O uso da UML no design de sistemas orientados a objetos e ferramentas de modelagem**. JOSIF, [s. l.], 2026. Disponível em: <https://josif.ifsuldeminas.edu.br/ojs/index.php/anais/article/view/2790/2492> Acesso em: mai. 2026.
- OLIVEIRA, Danielle Sales de. **Sistema de recomendação baseado em filtragem colaborativa em uma base de dados de feedback implícito**. 2019. Trabalho de Graduação (Bacharelado em Ciência da Computação) – Centro de Informática, Universidade Federal de Pernambuco, Recife, 2019. Disponível em: https://www.cin.ufpe.br/~tg/2019-1/propostas_CC/prop_dso.pdf Acesso em: mai. 2026.

RICCI, F.; ROKACH, L.; SHAPIRA, B. (eds.). **Recommender Systems Handbook**. New York: Springer, 2011. DOI: [10.1007/978-0-387-85820-3_1](https://doi.org/10.1007/978-0-387-85820-3_1) Disponível em: https://www.cs.ubbcluj.ro/~gabis/DocDiplome/SistemeDeRecomandare/Recommender_systems_handbook.pdf Acesso em : jul.2026.

SANTOS, Pedro Henrique Do Amaral; BORGES, João Henrique Gião; FLORIAN, Fabiana. Estudo da integração de Inteligência Artificial (IA) em Sistemas de Recomendação (SR): revisão bibliográfica. **RevistaFT**, [s. l.], v. 29, n. 140, nov. 2024. DOI: 10.69849/revistaft/fa10202411101906 Disponível em: <https://revistaft.com.br/estudo-da-integracao-de-inteligencia-artificial-ia-em-sistemas-de-recomendacao-sr-revisao-bibliografica/> Acesso em: mai. 2026.

SIQUEIRA, Déborados Santos e; PAULON, Matheus Montanha; GUEDES, Gilleanes Thorwald Araujo. Técnicas de Inspeção para Diagramas de Classes UML: Uma Revisão Sistemática. In: ESCOLA REGIONAL DE ENGENHARIA DE SOFTWARE (ERES), 3. , 2019, Rio do Sul. **Anais [...]**. Porto Alegre: Sociedade Brasileira de Computação, 2019 . p. 41-48. Disponível em: <https://sol.sbc.org.br/index.php/eres/article/view/8494/8395> Acesso em: mai. 2026

SILVA, Rafael Glauber Nascimento e. **Sistema de Recomendação baseado em conteúdo textual**: avaliação e comparação. 2014. 122 f. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal da Bahia / Universidade Estadual de Feira de Santana, Salvador, 2014. Disponível em: https://repositorio.ufba.br/bitstream/ri/19281/1/dissertacao_mestrado_ciencia_computacao_rafael_glauber.pdf Acesso em: mai. 2026.

STEINER, M. T. A. *et al.* Abordagem de um problema médico por meio do processo de KDD com ênfase à análise exploratória dos dados. **Gestão & Produção**, [s. l.], v. 13, n. 2, p. 325–337, 2006. DOI: <https://doi.org/10.1590/S0104-530X2006000200013>



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)