Impactos positivos e negativos da IA e dos LLMs no ensino de Banco de Dados

Daniel de Oliveira
danielcmo@ic.uff.br

Instituto de Computação – UFF
UFFeScience Research Group



Quem sou eu?



Daniel de Oliveira

Professor Associado do Instituto de Computação da UFF desde 2013

Mestre e doutor em Engenharia de Sistemas e Computação pela COPPE/UFRJ

200+

artigos científicos publicados em revistas e congressos

UFFeScience

grupo de pesquisa da UFF que ajudo a liderar

CNPq e FAPERJ

Bolsista de Produtividade e Cientista do Nosso Estado

Dados e IA

bancos de dados, proveniência, workflows e aprendizado de máquina

Professor de Princípios de Bancos de Dados e Bancos de Dados Não Convencionais — e pesquisador em Banco de Dados no UFFeScience

O aluno agora tem um assistente de BD

Ferramentas como ChatGPT, Claude, Gemini e Copilot já estão **no bolso de todo aluno**

Elas entendem o enunciado em português e devolvem SQL, diagramas e explicações **na hora**

A pergunta deixou de ser “o aluno vai usar?” e passou a ser “**como** ele vai usar?”

Modelar
esquemas ER e relacionais

Normalizar
dependências e decomposição

Gerar SQL
a partir de linguagem natural

Explicar
consultas passo a passo

Popular
dados de teste realistas

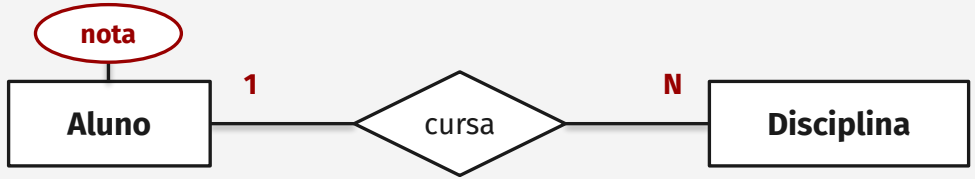
Otimizar
índices e reescritas

Boa parte dos exercícios clássicos de banco de dados é resolvida em segundos

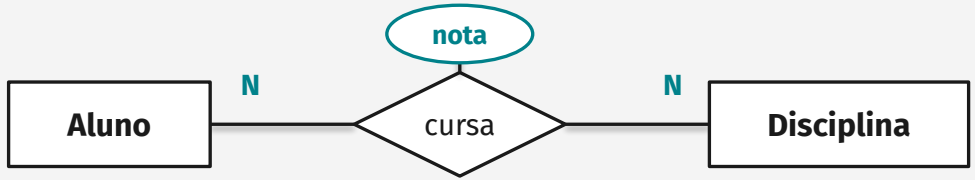
Na modelagem conceitual, a IA ainda erra muito

Enunciado: “alunos cursam disciplinas e recebem **nota** em cada”

Gerado pelo LLM: nota em Aluno, cardinalidade 1:N



Correto: N:N, nota como atributo do relacionamento



Erros frequentes

- **Cardinalidade** e participação trocadas
- **Atributo × entidade** (telefone, endereço)
- Atributos do **relacionamento** esquecidos
- **Entidades fracas** e especialização mal aplicadas
- **Regras de negócio** inventadas ou ignoradas
- Pula direto para **tabelas**

O LLM não pergunta, ele supõe! Modelagem conceitual é justamente perguntar!

E no SQL, o LLM erra com confiança

Pergunta: “quantos alunos há em cada disciplina, **incluindo as que não têm alunos**?”

Gerado pelo Claude: roda, mas está errado

```
SELECT d.nome, COUNT(*) AS total
FROM disciplina d
LEFT JOIN matricula m ON m.disciplina_id = d.id
GROUP BY d.nome;
```

Corrigido

```
SELECT d.id, d.nome, COUNT(m.aluno_id) AS total
FROM disciplina d
LEFT JOIN matricula m ON m.disciplina_id = d.id
GROUP BY d.id, d.nome;
```

Erros típicos

- Semântica de **NULL** e junções externas
- Junções que **duplicam linhas** antes de agregar
- **Colunas inventadas** que não existem no esquema
- Consultas certas, porém **lentas**
- Mistura de **dialetos** (PostgreSQL, MySQL, SQLite)

Só percebe o erro quem entende o modelo relacional

Os riscos para a aprendizagem

Ilusão de competência

A consulta roda, então parece certa, mesmo quando não está

Fundamentos pulados

Álgebra relacional e normalização viram “coisa que a IA faz”

Avaliação fragilizada

Listas feitas em casa deixam de medir o que o aluno sabe

Privacidade

Colar dados reais em chats externos pode violar a LGPD

Desigualdade

Quem pode pagar tem acesso a modelos melhores

Dependência

Sem a ferramenta, o aluno não sabe por onde começar

O problema não é usar IA! É usar IA no lugar de aprender! A IA virou a bengala do preguiçoso!

O que perde e o que ganha peso no ensino

Perde peso

- Decorar sintaxe de SQL
- Escrever CRUD repetitivo
- Traduzir enunciado em consulta, linha a linha
- Exercícios com resposta única e pronta

Ganha peso

- **Especificar** o problema com precisão
- **Verificar e testar** o resultado
- **Modelagem** e decisões de projeto
- **Desempenho**: planos de execução e índices
- **Ética e privacidade** dos dados

O foco sai de escrever a consulta e vai para julgar se ela está certa

Novas práticas em sala de aula

1

Ache o erro

corrigir SQL gerado por LLM usando um banco de testes

2

Aluno × LLM

comparar soluções e justificar qual está certa e por quê

3

LLM como Tutor

assistente configurado para dar dicas, não respostas

4

Projeto real

dados do contexto do aluno e decisões documentadas

- Testes automáticos com **resultados esperados** tornam o erro do LLM visível para o aluno
- Usar **SGBDs diferentes** (SQLite, PostgreSQL, MariaDB) expõe a confusão de dialetos

O LLM vira objeto de estudo, não o atalho

Avaliação na era dos LLMs

Prova presencial sem IA	Fundamentos: álgebra relacional, normalização e leitura de consultas
Projeto com IA declarada	Uso permitido, com **histórico de prompts** e justificativas das escolhas
Defesa oral curta	O aluno **explica** o resultado que o modelo que entregou
Rubricas novas	Pontuam **verificação, testes e decisões**, não só o resultado final

Pedir o histórico de prompts é pedir a proveniência do trabalho

O próprio banco de dados também mudou

Bancos vetoriais

Busca por similaridade de embeddings (ex.: **pgvector**) entra no currículo

RAG

LLMs que consultam bases de dados e documentos para responder **com fontes**

Text-to-SQL

Interfaces em linguagem natural: **como funcionam e onde falham**

LLMs no SGBD

Recomendação de índices e ajuste de parâmetros; no IC-UFF, trabalhos tem avaliado índices sugeridos por LLMs

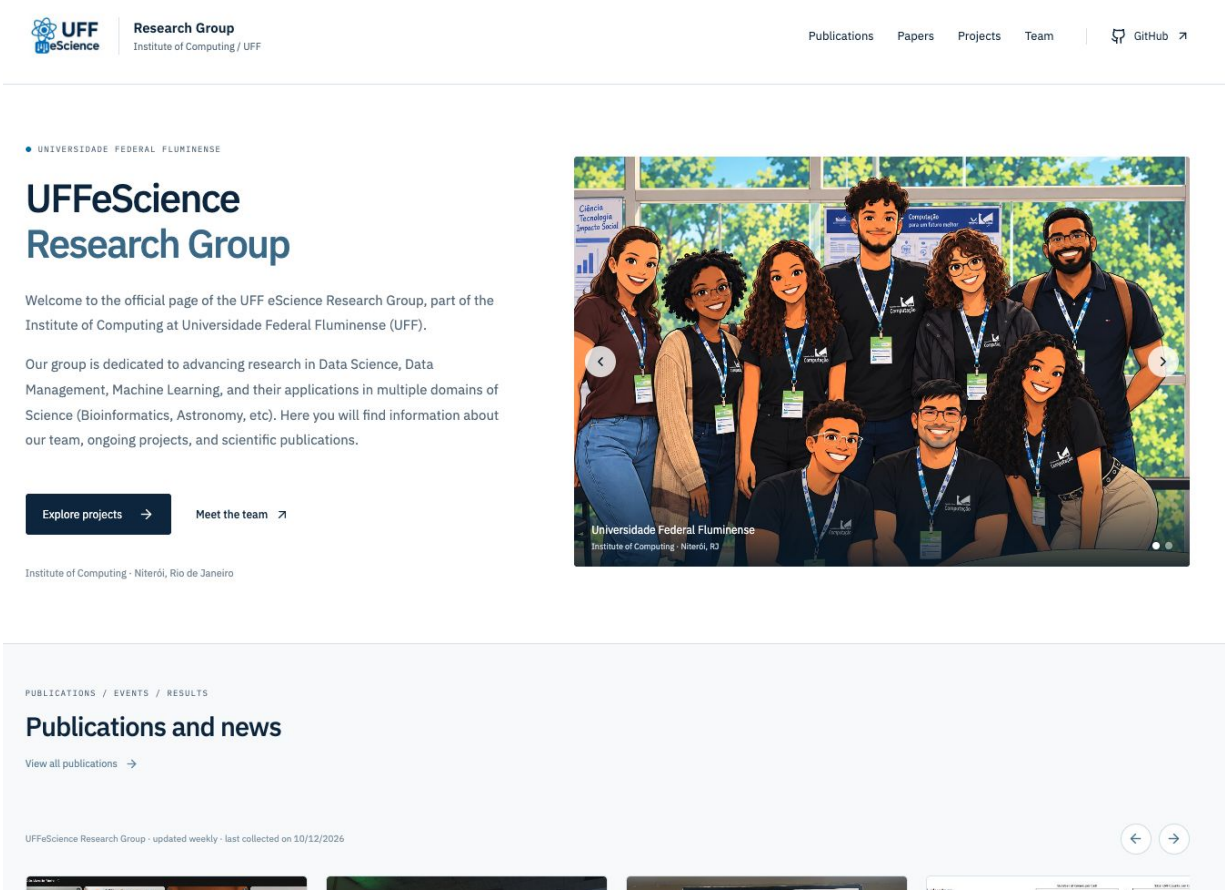
Ensinar banco de dados agora inclui integrar dados e modelos de linguagem

Qual a principal mensagem?

A IA não substitui entender dados: ela torna isso ainda mais importante



UFFeScience - http://uffescience.ic.uff.br/



Obrigado!



Instituto de Computação

UFF eScience

CEFET/RJ