



# Um Mapeamento Sistemático sobre Álgebra Relacional na Educação de Banco de Dados

## *A Systematic Mapping on Relational Algebra in Database Education*

Matheus Henrique Sechineli \*

André Luis Scherz†

### RESUMO

A Álgebra Relacional trata-se de um conjunto de operadores usados para realizar consultas em banco de dados relacional. Normalmente, esse conteúdo é tratado em disciplinas introdutória de banco de dados, sendo exposto na maioria dos casos de forma tradicional. Embora o estudo de Álgebra Relacional ajude a refinar a capacidade dos estudantes em escrever eficientes consultas de banco de dados, o ensino desse conteúdo ainda traz vários desafios a serem superados. Para entender, como a comunidade científica tem lidado com esses desafios, este trabalho tem como objetivo conduzir um mapeamento sistemático da literatura que aborde o ensino de Álgebra Relacional. Esse mapeamento resultou em trinta e sete estudos primários que foram publicados em diferentes conferências e periódicos. Esses trabalhos apresentam algumas abordagens de ensino e várias ferramentas que podem ser adotadas para apoiar o ensino-aprendizagem deste conteúdo. Com esses resultados, é possível identificar as principais iniciativas para melhorar a qualidade do ensino e apontar caminhos ainda a serem explorados.

**Palavras-chave:** Álgebra Relacional. Linguagem de Consulta de Banco de Dados. Mapeamento Sistemático.

### ABSTRACT

Relational Algebra is a set of operators used to perform queries in relational databases. Typically, this content is treated in introductory database courses, being exposed in most cases in a traditional way. Although studying Relational Algebra helps to refine students' ability to write efficient database queries, teaching this content still poses several challenges to be overcome. To understand how the scientific community has dealt with these challenges, this work presents a systematic mapping of the literature that addresses the education of Relational Algebra. This work resulted in thirty-seven primary studies that were published in different scientific conferences and journals. These primary studies present some teaching approaches and several tools that can be adopted to support the teaching-learning of this content. With these results, it is possible to identify the main initiatives to improve the quality of teaching and point out paths still to be explored.

**Keywords:** Relational Algebra. Database Query Language. Systematic Mapping.

## 1 INTRODUÇÃO

A Álgebra Relacional (AR) proposta por (CODD, 1970, 1972) é uma derivação da lógica de primeira ordem e da teoria de conjuntos. Ela tem sido usada como fundamento para desenvolvimento de linguagens de consulta em Banco de Dados. A Linguagem de Consulta Estruturada (SQL, do inglês *Structured Query Language*), por exemplo, é uma linguagem baseada na AR que tornou-se padrão nos diversos Sistemas de Banco de Dados. A

\* Departamento de Computação; ✉ [matheuss.2018@alunos.utfpr.edu.br](mailto:matheuss.2018@alunos.utfpr.edu.br); <https://orcid.org/0000-0002-6923-0141>.

† Departamento de Computação; ✉ [andreluis@utfpr.edu.br](mailto:andreluis@utfpr.edu.br); <https://orcid.org/0000-0002-8328-7144>.



AR é formada por operadores primitivos como: projeção, seleção, união, diferença, produto cartesiano, e o renomear. Eles são combinados para operarem nas tuplas de um banco de dados considerando uma ou várias relações.

As disciplinas introdutórias de Banco de Dados na graduação abordam a AR seguindo normalmente o modelo tradicional de ensino em que os exercícios são resolvidos com papel e caneta. Neste modelo, não é possível testar as expressões escrita com a AR, dificultando até a correção devido a quantidade de possibilidades corretas para o mesmo problema, principalmente os mais complexos. É importante considerar que a AR trás um jeito novo de pensar de forma procedural e exercita a lógica lapidando as qualidades que um aluno precisa ter para aprender a escrever melhores consultas em linguagens como SQL. Desta forma, a questão que norteia este trabalho é: o que a comunidade acadêmica tem proposto para ensinar/aprender álgebra relacional?

Para sintetizar métodos e ferramentas para ensinar/aprender AR, este trabalho tem como objetivo conduzir um mapeamento sistemático. Um Mapeamento Sistemático (MS) é um estudo secundário que tem como objetivo identificar e classificar o conteúdo relacionado com um tópico de pesquisa (KITCHENHAM; CHARTERS, 2007; PETERSEN; VAKKALANKA; KUZNIARZ, 2015). Esse corresponde a uma investigação ampla envolvendo estudos primários relacionados com o tópico de pesquisa específico, visando identificar as evidências disponíveis sobre esse tópico (FELIZARDO et al., 2017). Neste trabalho, 236 artigos coletados em diferentes indexadores foram submetidos a um rigoroso critério de seleção que resultou na identificação de trinta e sete estudos primários relevantes. Dentre os artigos relevantes, classificou-se os estudos que apresentam métodos/abordagens de ensino e aqueles que apresentam ferramentas educacionais. Com esses resultados, é possível identificar as principais iniciativas para melhorar a qualidade do ensino e apontar caminhos ainda a serem explorados.

Uma explicação sobre o método de pesquisa que inclui as etapas do MS é apresentada na Seção 2. Os resultados são apresentados na Seção 3. Por fim, a conclusão encontra-se na Seção 4.

## 2 MÉTODO

O MS é composto por processos metodológicos bem definidos que permitem sua reprodutibilidade. Um MS é composto de três fases: planejamento, condução e relatório. No planejamento, define-se as questões que motivaram a produção do MS e o protocolo de pesquisa que inclui: os critérios de inclusão e exclusão, as bases almeçadas, a estratégia de busca, a definição da *string* de busca, e os critérios de qualidade. Já na fase da condução, por meio da estratégia de busca, coleta-se os estudos primários a partir das bases de dados e indexadores de publicações científicas almeçados. Em seguida, ainda na condução, os critérios de seleção (inclusão e remoção) e de qualidade são aplicados aos estudos coletados para selecionar aqueles que são relevantes para a pesquisa. Por fim, a extração e síntese dos dados é realizada. O relatório é a última fase de um MS e descreve os resultados do trabalho para divulgação na comunidade científica.

O protocolo foi conduzido por duas pessoas sendo um pesquisador experiente e um pesquisador iniciante. Cada etapa da pesquisa como mostra a Figura 1 foi amplamente discutida e realizada por ambos pesquisadores. Por exemplo, na aplicação dos critérios de inclusão/exclusão ambos realizaram suas análises de forma independente. Em caso de discordância dos resultados, ambos pesquisadores discutiam novamente os critérios até um consenso ser alcançado. Essa discussão foi útil para o refinamento do protocolo e a equalização do conhecimento sobre o método entre ambos pesquisadores.

As seguintes questões de pesquisa descrevem o tópico de pesquisa abordado e delimitam o escopo deste

trabalho.

Q1 Quais são as iniciativas de pesquisa para o ensino e aprendizagem de Álgebra Relacional em cursos de banco de dados?

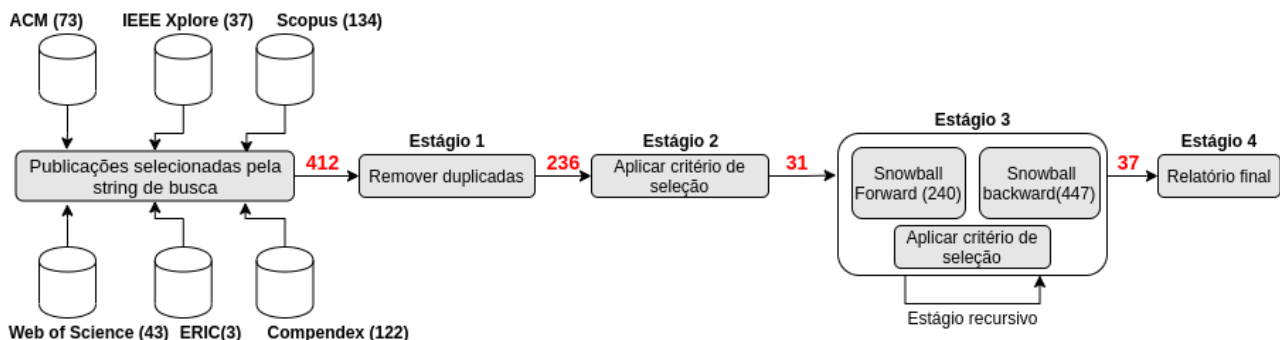
Q2 Em que locais e quando foram publicadas as iniciativas de pesquisa para o ensino e aprendizagem da Álgebra Relacional?

Após as questões de pesquisa, definiu-se a *string* de busca para ser utilizada nos repositórios/indexadores de artigos. Para isso, observou-se palavras chaves que poderiam representar pesquisas sobre ensino-aprendizado de AR. Com isso, chegou-se ao seguinte resultado: "relational algebra"AND (learn\* OR educat\* OR teach\* OR tutor\* OR train\* OR student\*). O termo *students* foi incorporado apenas num segundo momento em uma revisão do protocolo de pesquisa, pois em conduções iniciais percebeu-se que o termo era frequente e, ao ser incorporado na *string*, novos estudos relevantes seriam encontrados.

A *string* de busca foi aplicada nas bases: ACM, IEEE, Scopus, Web of Science, ERIC e Compendex. Estas bases são as mais relevantes para o escopo deste trabalho e apresentaram alta taxa de abrangência segundo (FELIZARDO et al., 2017) em pesquisas na Ciência da Computação.

A Figura 1 mostra que 412 artigos foram capturados nas buscas em todas as bases. Observe que o Compendex e o Scopus foram as bases que mais retornaram trabalhos. Em seguida, iniciou-se a remoção das duplicações. Desta forma, restaram 236 artigos para aplicação dos critérios de seleção.

Figura 1 – Etapas do processo



Fonte: Autoria própria (2021).

Para inclusão e remoção dos estudos aplicou-se rigorosamente os seguintes critérios de seleção: (i) escritos em língua inglesa, (ii) publicados em periódicos ou conferências com revisão de pares e (iii) abordar a temática de álgebra relacional na educação. Aplicando estes critérios nos 236 artigos do estágio 1, como mostrado na Figura 1, resultaram em 31 artigos.

Em seguida, com os 31 artigos selecionados aplicou-se a técnica de *snowball*. *Snowball backward* trata-se de uma estratégia para busca de artigos relevantes a partir das referências dos artigos selecionados. Em complemento, *snowball forward* refere-se a busca de artigos relevantes sobre a lista de citações dos artigos selecionados. Observa-se também que ambos são recursivos, uma vez que o processo se repete a cada novo artigo que atende os critérios de inclusão e exclusão. Como ilustrado na Figura 1, o estágio 3 trata-se do *snowball backward* e *forward*. Analisamos 240 artigos resultantes do *snowball forward* e 447 resultantes do *snowball backward*. Ao aplicar os critérios de inclusão e exclusão citados anteriormente, mais 6 artigos foram incluídos como estudos primários selecionados.



### 3 RESULTADOS

Nesta seção, sumariza-se as questões de pesquisa com base em nossos resultados. Destaca-se as principais contribuições, os locais e quando elas foram publicadas.

A Tabela 1 destaca os trabalhos classificados nesta categoria como abordagens de ensino. Em geral, eles apresentam metodologias para o ensino de AR (PS5, PS6, PS9, PS12, PS18, PS21, PS33 e PS37) e propõem alterações nas estruturas curriculares dos cursos (PS30).

**Tabela 1 – Estudos Primários Selecionados**

ID	Título	Ano	Tipo
PS1	Teaching relational database concepts in a University Environment	1982	Ferramenta
PS2	An Academic Relational Algebra Database Management System	1990	Ferramenta
PS3	An Educational Tool for Formal Relational Database Query Languages	1993	Ferramenta
PS4	WinRDBI: A windows-based relational database educational tool	1997	Ferramenta
PS5	RELAX — the Relational Algebra Pocket Calculator Project	2000	Ferramenta
PS6	On making relational division comprehensible	2003	Abordagem
PS7	Virtura - A virtual tutor for relational algebra	2003	Ferramenta
PS8	A Computer-based Environment for the Study of Relational Query Languages	2003	Ferramenta
PS9	Gauging students' understanding through interactive lectures	2004	Abordagem
PS10	iDFQL – A query-based tool to help the teaching process of the relational algebra	2004	Ferramenta
PS11	Teaching relational algebra and relational calculus: A programming approach	2006	Ferramenta
PS12	Levels of abstraction in database query definition	2006	Abordagem
PS13	An automatic correction tool for relational algebra queries	2007	Ferramenta
PS14	A relational algebra query language for programming relational databases	2010	Ferramenta
PS15	Implementation of Relational Algebra Interpreter Using Another Query Language	2010	Ferramenta
PS16	Relational algebra programming with Microsoft access databases	2011	Ferramenta
PS17	A system to support teaching and learning relational database query languages and query processing	2011	Ferramenta
PS18	A Cooperative Learning-Based Strategy for Teaching Relational Algebra	2011	Abordagem
PS19	Relational Algebra Interpreter in Context of Query Languages	2011	Ferramenta
PS20	Relational algebra and SQL: Better together	2012	Ferramenta
PS21	Using a semi-realistic database to support a database course	2013	Abordagem
PS22	Support of relational algebra knowledge assessment	2013	Ferramenta
PS23	An Object-Oriented Relational Algebra Query Language	2013	Ferramenta
PS24	Learning Relational Algebra by Snapping Blocks	2014	Ferramenta
PS25	Relational Algebra Interpreter	2014	Ferramenta
PS26	DBSnap: Learning Database Queries by Snapping Blocks	2015	Ferramenta
PS27	Querying databases by snapping blocks	2015	Ferramenta
PS28	RAPT: Relational Algebra Parsing Tools	2015	Ferramenta
PS29	Automated Evaluation of Learner's Solutions Expressed in a Graphical Language: Application to the Relational Databases Domain	2015	Ferramenta
PS30	A strategy for improving the teaching of database management systems course: Redressing the balance of theory and practice	2016	Abordagem
PS31	TanQuery: A Tangible System for Relational Algebra Learning	2017	Ferramenta
PS32	DB-Learn: Studying Relational Algebra Concepts by Snapping Blocks	2018	Ferramenta
PS33	Explaining Wrong Relational Queries Using Small Examples	2019	Abordagem
PS34	RATest: Explaining Wrong Relational Queries Using Small Examples	2019	Ferramenta
PS35	Towards A Distributed Interactive Surface to Support Students with Hypoacusis in the Collaborative Learning of Relational Algebra	2019	Ferramenta
PS36	RelaX: A webbased execution and learning tool for relational algebra	2019	Ferramenta
PS37	Challenges with teaching and learning theoretical query languages	2020	Abordagem

Além desses trabalhos, destaca-se os artigos que propõem ferramentas para auxiliar no ensino e aprendizado



de AR. Dentre as ferramentas, classificou-se aquelas que são baseadas em WIMP (do inglês *Windows, Icons, Menu and Pointing device, Tokens*, Blocos e Textos). Ferramentas baseadas em WIMP são aquelas que as consultas são elaboradas com auxílio de estruturas gráficas como ícones, menus e janelas. Dentre nossos artigos desta categoria, destacam-se os PS4, PS7, PS10, PS13, PS17, PS36. Os estudos que usam ferramentas baseadas em *Token* são PS31 e PS35. *Tokens* são objetos físicos que permitem interagir com a interface digital por meio de câmeras que captam sua presença e função. A ideia é formar consultas com os *Tokens*. Ferramentas que utilizam blocos são aquelas que inspiraram-se em linguagens de programação por blocos em que cada bloco é uma operação primitiva e operações mais complexas podem ser construídas pelo encaixe dos blocos. Nesta categoria encontram-se as ferramentas PS24, PS26, PS27 e PS32. Por fim, ferramentas baseadas em texto são aquelas em que é necessário escrever textualmente as expressões algébricas. As ferramentas que usam esta abordagem são PS1, PS2, PS3, PS8, PS11, PS14, PS15, PS16, PS19, PS20, PS22, PS23, PS25, PS28, PS29, e PS34.

Os locais de publicação são sumarizados na Tabela 2. Observa-se que os 37 estudos relevantes estão distribuídos em vinte e oito locais, sendo que 22 são publicados em conferências e 6 estão publicados em periódicos científicos. Além disso, pode-se destacar que as conferências ISECOM, ITiCSE, SIGCSE, SIGMOD, TLAD, WCETE e o periódico SIGCSE Bulletin destacam-se por apresentar mais do que uma publicação.

**Tabela 2 – Conferências e periódicos dos artigos selecionados**

Local	Tipo	Qtde	Local	Tipo	Qtde	Local	Tipo	Qtde
ASEE	Conferência	1	ICCSA	Conferência	1	ITiCSE	Conferência	2
BNCOD	Conferência	1	ICDE	Conferência	1	JCSE	Periódico	1
BTW	Conferência	1	IJCTE	Periódico	1	JISE	Periódico	1
CISSE	Conferência	1	IJKM	Periódico	1	SEBD	Conferência	1
CONTIE	Conferência	1	IJMECS	Periódico	1	SIGCSE	Conferência	4
CSEDU	Conferência	1	IMCIC-ICSIT	Conferência	1	SIGCSE Bulletin	Periódico	2
DSDE	Conferência	1	INTERACCION	Conferência	1	SIGMOD	Conferência	2
FIE	Conferência	1	ISECOM	Conferência	3	TLAD	Conferência	2
ICAICTE	Conferência	1	ISITE	Conferência	1	WCETE	Conferência	1
ICCCNT	Conferência	1						

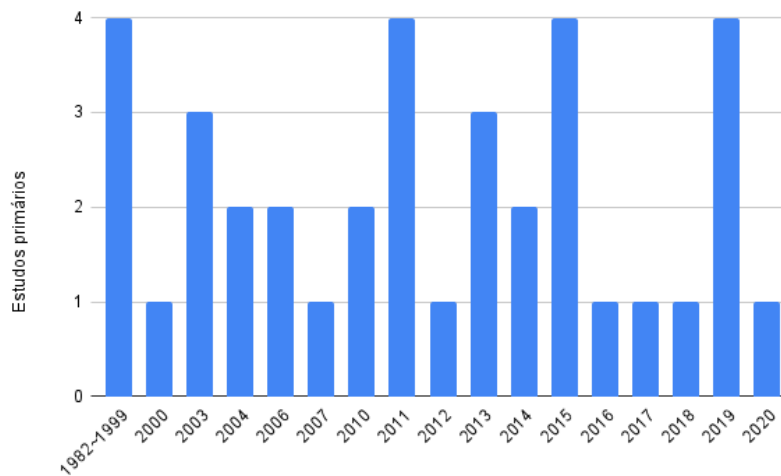
A Figura 2 ilustra a distribuição das publicações ao longo dos anos. Observe que este trabalho não limitou a pesquisa apenas aos anos mais recentes. Isso resultou em alguns trabalhos antigos, anteriores a 2000, que julga-se ser importante para a pesquisa, pois evidência que o ensino de AR tem sido pesquisado há vários anos. Ainda, sobre a Figura 2, observa-se que aproximadamente 60% dos artigos relevantes foram publicados nos últimos dez anos. Além disso, cinco trabalhos foram publicados nos últimos dois anos evidenciando que a pesquisa continua sendo atual e importante para a comunidade acadêmica.

#### 4 CONCLUSÃO

Tradicionalmente, o conteúdo de AR é abordado em disciplinas introdutórias de banco de dados. Entretanto, por seu fundamento teórico e matemático, os alunos costumam ter dificuldade de entendimento. Neste trabalho, buscou-se realizar um MS pra conhecer ferramentas e métodos que possam auxiliar no ensino-aprendizagem da AR. O MS resultou em trinta e sete estudos primários que destacam abordagens de ensino e ferramentas de apoio, principalmente, para validar as expressões escritas pelos alunos. Com esse mapeamento, pode-se entender como a pesquisa em educação de AR evoluiu ao longo dos anos e as principais contribuições. Para pesquisas



Figura 2 – Distribuição ao longo do tempo dos artigos publicados



Fonte: Autoria própria (2021).

futuras, almeja-se aplicar as ferramentas propostas com alunos em sala de aula para verificar a efetividade do uso para o ensino de AR.

## AGRADECIMENTOS

O presente trabalho foi realizado com o apoio da Fundação Araucária FA - Paraná/Brasil e da Universidade Tecnológica Federal do Paraná/Brasil.

## REFERÊNCIAS

- CODD, E. F. A Relational Model of Data for Large Shared Data Banks. **Communications of the ACM**, New York, NY, USA, v. 13, n. 6, p. 377–387, jun. 1970. ISSN 0001-0782. DOI: [10.1145/362384.362685](https://doi.org/10.1145/362384.362685).
- CODD, E. F. Relational Completeness of Data Base Sublanguages. In: PROCEEDINGS of 6th Courant Computer Science Symposium. New York, NY, USA: [s.n.], 1972. (Data base System), p. 65–98. ISBN 0-13-196741-X.
- FELIZARDO, Katia Romero et al. **Revisão Sistemática da Literatura em Engenharia de Software: teoria e prática**. 1. ed. [S.l.]: Elsevier, 2017.
- KITCHENHAM, Barbara; CHARTERS, Stuart. **Guidelines for performing systematic literature reviews in software engineering**. [S.l.], 2007.
- PETERSEN, Kai; VAKKALANKA, Sairam; KUZNIARZ, Ludwik. Guidelines for conducting systematic mapping studies in software engineering: An update. **Information and Software Technology**, Elsevier, v. 64, p. 1–18, 2015.