

# Candidacy Exam

## Textbook

- Ramakrishnan and Gehrke: *Database Management Systems (3rd edition)*

## Information Retrieval

1. Modern Information Retrieval: A Brief Overview; Singhal, IEEE Data Engineering Bulletin, December 2001
2. Introduction to Information Retrieval; Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze, Cambridge University Press, 2008. Chapter 1 (“Boolean Retrieval”), Chapter 6 (“Scoring, Term Weighting, and the Vector Space Model”), Chapter 7 (“Computing Scores in a Complete Search System”), and Chapter 11 (“Probabilistic Information Retrieval”).
3. The Anatomy of a Large-Scale Hypertextual Web Search Engine; Brin and Page, Proceedings of the WWW Conference, 1998

## Query Processing

1. Multidimensional Access Methods; Gaede and Ganther, ACM Computing Surveys, Vol. 30, No. 2, Pages 170-231, 1998
2. Access Path Selection in a Relational Database Management System; Selinger et al., RDS2, pp. 84-95
3. An Overview of Query Optimization in Relational Systems; Chaudhuri, PODS 1998
4. Everything You Always Wanted to Know About Compiled and Vectorized Queries But Were Afraid to Ask; Kersten et al; VLDB 2018
5. Neumann, Thomas. "Efficiently compiling efficient query plans for modern hardware." *Proceedings of the VLDB Endowment* 4.9 (2011): 539-550.
6. Goetz Graefe and William J McKenna. The volcano optimizer generator: Extensibility and efficient search. In *Proceedings of IEEE 9th international conference on data engineering*, pages 209–218. IEEE, 1993.
7. Graefe, Goetz. *Volcano: An Extensible and Parallel Query Evaluation System*. Oregon Graduate Center, 1989.
8. Column-Stores vs. Row-Stores: How Different Are They Really? Abadi DJ, Madden SR, Hachem N. SIGMOD 2008

## Transaction Management: Concurrency and Recovery

1. ARIES: A Transaction Recovery Method Supporting Fine-Granularity Locking and Partial Rollbacks Using Write-Ahead Logging; Mohan et al., RDS2, pp. 243-277
2. OLTP Through the Looking Glass, and What We Found There; Stavros Harizopoulos, et Al., SIGMOD Conference 2008

## Data Warehousing, Mining, Views and Models

1. An Overview of Business Intelligence Technology; Chaudhuri, et al., Communications of the ACM, Vol. 54 No. 8, Pages 88-98, August 2011.
2. Data Cube: A Relational Aggregation Operator Generalizing Group-by, Cross-Tab, and Sub Totals; Gray et al., Data Mining and Knowledge Discovery 1(1), 1997, pp. 29-53
3. Answering Queries Using Views: A Survey; Halevy, The VLDB Journal, Vol. 10, No. 4, Pages 270-294, 2001
4. BlinkDB: Queries with Bounded Errors and Bounded Response Times on Very Large Data; Agarwal; EuroSys 2011

## System Projects

1. The snowflake elastic data warehouse. Dageville, Benoit, et al. *Proceedings of the 2016 International Conference on Management of Data*. 2016.
2. C-store: a column-oriented DBMS; Stonebraker, M., Abadi, D. J., Batkin, A., Chen, X., Cherniack, M., Ferreira, M., ... & Zdonik, S. (2018).

# Visualization Interface

1. SUPPLE: automatically generating user interfaces. Gajos, Krzysztof, and Daniel S. Weld. *Proceedings of the 9th international conference on Intelligent user interfaces*. 2004
2. Zhang, Qianrui, et al. "Mining precision interfaces from query logs." *Proceedings of the 2019 International Conference on Management of Data* 2019.
3. Lux: Always-on Visualization Recommendations for Exploratory Data Science. Lee, Doris Jung-Lin, et al. in SIGMOD 2021.
4. Eichmann, Philipp, et al. "Idebench: A benchmark for interactive data exploration." *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*. 2020.
5. Moritz, Dominik, Bill Howe, and Jeffrey Heer. "Falcon: Balancing interactive latency and resolution sensitivity for scalable linked visualizations." *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019.