San José State University
Department of Computer Science
Fall Semester 2022
CS 257 – Database System Principals, Section 2
Syllabus

Course and Contact Information

Course Title:	**Database System Principals**
Instructor:	**Professor Ahmed Ezzat, Ph.D.** **e-Mail:** Ahmed.Ezzat@sjsu.edu
Date:	**Fall 2022 (51076)**
Course Number:	**CS 257**
Credit Hours:	**3 Credit Hours**
Classroom:	MacQuarrie Hall, room 233
Office Hour:	Duncan Hall, room 282
Class Schedule:	**Every Tuesday + Thursday: 7:30pm – 8:45pm**
Office hour Schedule	**Every Tuesday (in DH-282): 6:20pm – 7:20pm** **Every Wednesday (Online): 12:30pm – 1:30pm**
Class Recording:	**All classes are recorded (audio + video) by the school and all recordings will be available on Camino (Class recordings folder) in less than 24 hours.**

COVID-19 and Monkeypox Safety Training

Students registered for a College of Science (CoS) class with an in-person component should view the CoS COVID-19 and Monkeypox Training slides for updated CoS, SJSU, county, state and federal information and guidelines, and more information can be found on the SJSU Health Advisories website. By working together to follow these safety practices, we can keep our college safer. Failure to follow safety practice(s) outlined in the training, the SJSU Health Advisories website, or instructions from instructors, TAs or CoS Safety Staff may result in dismissal from CoS buildings, facilities or field sites. Updates will be implemented as changes occur (and posted to the same links).

Grader: Shashwat Kadam **Email**: Shashwat.kadam@sjsu.edu

Text Book: **- Database Systems: The Complete Book,** Second Edition, Hector Garcia-Molina, Jeffrey Ullman, and Jennifer Widom, Pearson & Prentice Hall, 2009, ISBN-13: 978-0-13-187325-4.

Other Reading: Refer to the end of the Syllabus for list of Online useful articles to read.

Course Description: This course covers advanced topics in Database system design and implementation. The course is intended to help the students understand the following advanced topics including: Storage and Indexes, Query processing, Query Optimization, Concurrency Control, Transaction Management, Recovery, Data Warehouse, Parallel and Distributed database, NoSQL data models, and comparison between SQL and

NoSQL..

The course is structured around three goals:

1. Cover advanced database topics from the text book.
2. Cover an overview of active research topics in the database area from published papers.
3. **Homework:** are individual assignments.
4. **Projects:**
    a. **Individual Project:** each student will create an RDBMS personal development environment, e.g., using their own laptop or an alternative.
    b. **Group Project:** each group will do regular group project and Final group project. These 2 projects are meant to help you in two main areas:
        i. Design RDBMS schema and implement a shopping cart application (Group Project-2)
        ii. Become familiar with executing SQL on Hadoop using Hive, and Good understanding and compare the query optimization material in the class against the query plan generated by a mature database such as Oracle (Final Group Project)
5. **Group Research Paper:** Deeper Group Research by the students in any of the topics covered at high-level in item (2) and using optionally the list of publications supplied at the end of the Syllabus. Toward that goal, the students will be divided into set of small groups. Each group will select topic from the advanced research topics listed below. Each group will write a formal paper about their assigned topics and gives a presentation to the rest of the class describing: the problem being investigated, what others have done, and the group own views and perspective. List of publications in the relevant topics are provided at the end of this syllabus. The students are encouraged to research additional published papers on their topic.
6. **Tests**: The midterm will cover the lecture material covered in the class, and the research paper and the Final group project and the group project presentation represent the final exam.

Course Learning Outcome:

Students are expected to learn the following:

- SQL Storage management architecture (both row and column-oriented) and the performance relevant issues,
- SQL Query processing architecture and issues with SQL runtime,
- SQL Query optimization and the issues relevant to query plan quality,

- Concurrency control; goals and approach,
- Transaction management and ACID properties,
- Database recovery; goal and approach,
- Parallel and Distributed database architecture and issues,
- **Overview to current active research topics** to help the students pursue deeper research in any of the topics covered; topics listed at the end of the syllabus. Students are expected to select a topic, survey published papers and write paper/presentation about the issue being addressed, how it is being addressed, and their own perspective.
- **Group Project for an advanced application to database (multi-tier application)** to help the students going through the steps for creating then using a database , i.e., E/R diagram design, schema normalization, database population, and finally querying the database.
- **Overview of NoSQL Data Models** and cover few examples of each
- Understand the tradeoff between SQL and NoSQL databases
- **Hands-on experience** with using SQL on Hadoop and how advanced optimizers translate SQL statement into Query plan.

Due Dates and Lateness

Homework, individual project, and Group projects, and Group Research paper: in all these assignments, each member will lose 10% of the assignment grade for each day late, and after 5 days assignments will not be accepted. All Homework, projects and Research paper are available in the Design Center Services.

Pre-requisite: *CS 157B or equivalent.*

Classroom Protocol:

The classroom include lecture, group discussion, extensive reading and information seeking, presentation, and structured written assignments. Collegial approaches are encouraged between students, and between student and faculty, so that learning is a joint endeavor.

This class will be conducted in an atmosphere of mutual respect. I encourage your active participation and welcome reasoned debate. However, your language and conduct during the class period must demonstrate respect for everyone's race, gender identity, or expression, sexuality, culture, beliefs, and abilities.

Course Requirements

Class participation is essential in understanding the course concepts and relies on being prepared for class. Extensive reading and writing are an expectation in this course. All written papers will demonstrate the format of the Publication Manual of the American Psychological Association (7th ed.), and will include title page, body of paper, references, and appendices if appropriate.

Grading Policy

Grading Scale: refer to the Grading System table below.
Student can ask/discuss about any grade only within 7-days from the date they received the grade.

□	**Homework Assignment (2 HWs)**	**(2*10)% = 20%%**
□	**P1 (individual) + P2 (Group)**	**P1 (6%) + P2 (9%) = 15%**
□	**Participation/bonus:**	**Additional optional 2%**
□	**Quizzes Total**	**5%**
□	**Midterm:**	**25%**
□	**Group Research Paper**	**15%**
□	**Final Exam (Group project):**	**20%**

Course Projects Assignment

The course include group-based project assignments and related documents must be handed in the classroom on due date (one copy per group). Familiarity with C language and Linux are expected for these projects.

- □ Project-1 (6 pts): Assignment is on Aug 25, 2022, and is due back on Sept 6, 2022.
- □ Final Project (20 pts): Assignment is on Aug 30, 2022, and is due back on Nov 24, 2022
- □ Research Paper (15 pts): Assignment is on Sept 1, 2022 and is due back on Nov 24, 2022
- □ HW1 (10 pts): Assignment is on Sept 6, 2022, and is due back on Sept 20, 2022.
- □ Project-2 (9 pts): Assignment is on Sept 27, 2022, and is due back on Oct 20, 2022.
- □ HW2 (10 pts): Assignment is on Nov 3, 2022, and is due back on Nov 17, 2022.

Due Dates and Lateness

Group members will lose 10% of the project grade for each day delay, and after 5 days, projects will not be accepted. Project descriptions are available on the Design Center Servers.

Classroom Protocol:

The classroom include lecture, group discussion, extensive reading and information seeking, presentation, and structured written assignments. Collegial approaches are encouraged between students, and between student and faculty, so that learning is a joint endeavor.

This class will be conducted in an atmosphere of mutual respect. I encourage your active participation and welcome reasoned debate. However, your language and conduct during the class period must demonstrate respect for everyone's race, gender identity, or expression, sexuality, culture, beliefs, and abilities.

Exams

There will be no make-up exams. If a student misses and exam without a legitimate excuse or advanced approval from the professor, a grade of zero will be recorded.

Grading System:

Grading follow the standard distribution as shown below:

Score Range	Grade	GPA
95 – 100.00	A	4.0
90 - 94.9	A-	3.7
87 - 89.9	B+	3.3
84 - 86.9	B	3.0
80 - 83.9	B-	2.7
77 - 79.9	C+	2.3
73 - 76.9	C	2.0
70 - 72.9	C-	1.7
68 - 69.9	D+	1.3
62 - 67.9	D	1.0
60 -61.9	D-	0.7
Below 60	F	0.0

Course Outline

Week	Topic
1. Aug 23:	**Disk and Files, Records and Blocks organization, Storage (Ch 13) + Groups Creation + How we are going to work in the course**
1. Aug 25:	**Disk and Files, Records and Blocks organization, Storage (Contd.) + Project-1 Preview**
2. Aug 30:	**Disk and Files, Records and Blocks organization, Storage: Relational Indexes: B-tree, Hash table, Multi-dimensional indexes using hash and tree structures, Bitmap indexes (Ch 14) (Contd.) + Final Group Project Preview + Finalize Groups Membership**
2. Sept 1:	**Disk and Files, Records and Blocks organization, Storage (Contd.): Relational Indexes (Contd.) + Finalize Group Research Paper Preview**
3. Sept 6:	**Query Processing and Optimization (Ch 15 + 16) + HW1 handout/Preview + Return of P1**
3. Sept 8:	**Query Processing and Optimization (Contd.) + Finalize set of queries for Final Project**
4. Sept 13:	**Query Processing and Optimization (Contd.) + Finalize Group topics for Research Paper**
4. Sept 15:	**Query Processing and Optimization (Contd.)**
5. Sept 20:	**Concurrency Control (Ch 18) + Return of HW1**
5. Sept 22:	**Concurrency Control (Contd.)**
6. Sept 27:	**Concurrency Control (Contd.) + Project-2 Preview**
6. Sept 29:	**Transaction Management (Ch 19) + Quiz-1**
7. Oct 4:	**Transaction Management (Contd.) + Finalize queries for P2**
7. Oct 6:	**System Failure and Recovery (Ch 17) + Midterm Preview**
8. Oct 11:	**System Failure and Recovery (Contd.)**
8. Oct 13:	**Midterm (Closed book)**
9. Oct 18:	**Data Warehouse + OLAP + Data Mining (Ch 22)**
9. Oct 20:	**Data Warehouse + OLAP + Data Mining (Contd.) + Return of P2**
10. Oct 25:	**Data Warehouse + OLAP + Data Mining (Contd.)**
10. Oct 27:	**Parallel and Distributed Database (Ch 20)**
11. Nov 1:	**Parallel and Distributed Database (Contd.)**
11. Nov 3:	**Parallel and Distributed Database (Contd.) + HW2 Handout/Preview**
12. Nov 8:	**Parallel and Distributed Database (Contd.)**
12. Nov 10:	**IMDB Database + Column-oriented + Quiz-2**
13. Nov 15:	**IMDB Database + Column-oriented (Contd.)**

13. Nov 17:	**NoSQL KV-pair: Data Model and Examples + Return of HW2**
14. Nov 22:	**NoSQL KV-pair: Data Model and Examples (Contd.)**
14. Nov 24:	**NoSQL Document: Data Model and Examples +**
	Return Group Final project presentation and return research papers (both are due)
15. Nov 29:	**NoSQL Document: Data Model and Examples (Contd.)**
15. Dec 1:	**SQL vs. NoSQL**
16. Dec 6:	**Final (Final Project): Group PPT + Presentation**
16. Dec 8:	**Final (Final Project): Group PPT + Presentation (Contd.)**

Instruction Methods: In-Class Lecture

Exams

There will be no make-up exams. If a student misses and exam without a legitimate excuse or advanced approval from the professor, a grade of zero will be recorded. Attending the Final Project presentation is expected.

University Policies

General Expectations, Rights and Responsibilities of the Student

As members of the academic community, students accept both the rights and responsibilities incumbent upon all members of the institution. Students are encouraged to familiarize themselves with SJSU's policies and practices pertaining to the procedures to follow if and when questions or concerns about a class arise. See University Policy S90–5 at http://www.sjsu.edu/senate/docs/S90-5.pdf . More detailed information on a variety of related topics is available in the SJSU catalog, at http://info.sjsu.edu/web-dbgen/narr/catalog/rec-12234.12506.html. In general, it is recommended that students begin by seeking clarification or discussing concerns with their instructor. If such conversation is not possible, or if it does not serve to address the issue, it is recommended that the student contact the Department Chair as a next step.

Dropping and Adding

Students are responsible for understanding the policies and procedures about add/drop, grade forgiveness, etc. Refer to the current semester's Catalog Policies section at http://info.sjsu.edu/static/catalog/policies.html. Add/drop deadlines can be found on the current academic year calendars document on the Academic Calendars webpage at http://www.sjsu.edu/provost/services/academic_calendars. The Late Drop Policy is available at http://www.sjsu.edu/aars/policies/latedrops/policy. Students should be aware of the current deadlines and penalties for dropping classes. Information about the latest changes and news is available at the Advising Hub at http://www.sjsu.edu/advising. Information about the latest changes and news is available at the Advising Hub at http://www.sjsu.edu/advising/.

Consent for Recording of Class and Public Sharing of Instructor Material

University Policy S12-7, http://www.sjsu.edu/senate/docs/S12-7.pdf, requires students to obtain instructor's permission to record the course. "Common courtesy and professional behavior dictate that you notify someone when you are recording him/her. You must obtain the instructor's permission to make audio or video recordings in this class. Such permission allows the recordings to be used for your private, study purposes only. The recordings are the intellectual property of the instructor; you have not been given any rights to reproduce or distribute the material." "Course material developed by the instructor is the intellectual property of the instructor and cannot be shared publicly without his/her approval. You may not publicly share or upload instructor generated

material for this course such as exam questions, lecture notes, or homework solutions without instructor consent."

Academic integrity
University Policy S12-7, http://www.sjsu.edu/senate/docs/S12-7.pdf, requires students to obtain instructor's permission to record the course. "Common courtesy and professional behavior dictate that you notify someone when you are recording him/her. You must obtain the instructor's permission to make audio or video recordings in this class. Such permission allows the recordings to be used for your private, study purposes only. The recordings are the intellectual property of the instructor; you have not been given any rights to reproduce or distribute the material." "Course material developed by the instructor is the intellectual property of the instructor and cannot be shared publicly without his/her approval. You may not publicly share or upload instructor generated material for this course such as exam questions, lecture notes, or homework solutions without instructor consent."

Campus Policy in Compliance with the American Disabilities Act
If you need course adaptations or accommodations because of a disability, or if you need to make special arrangements in case the building must be evacuated, please make an appointment with me as soon as possible, or see me during office hours. Presidential Directive 97-03 at http://www.sjsu.edu/president/docs/directives/PD_1997-03.pdf requires that students with disabilities requesting accommodations must register with the Accessible Education Center (AEC) at http://www.sjsu.edu/aec to establish a record of their disability. In 2013, the Disability Resource Center changed its name to be known as the Accessible Education Center, to incorporate a philosophy of accessible education for students with disabilities. The new name change reflects the broad scope of attention and support to SJSU students with disabilities and the University's continued advocacy and commitment to increasing accessibility and inclusivity on campus.

Expected Learning Outcome:

- Detailed understanding to relational database critical subsystems such as query processing and optimization techniques, concurrency control, transaction management and recovery.
- Detailed understanding to the different indexes support in RDBMS and how they are leveraged by the query optimizer
- Exposure to current advanced research topics related to RDBMS including Map/Reduce as a paradigm for distributed processing.
- Work with a small team to writing a research paper/presentation in one of selected topics – relevant published papers on these topics are listed below.
- Experience presenting your work to a relatively large technical audience.

Honor Code:	All students taking courses in the School of Engineering agree, individually and collectively, that they will not give or receive unpermitted aid in examinations or other course work that is to be used by the instructor as the basis of grading.
Attendance:	Required. All students are expected to attend all classes and the final at the official time. No alternate exams will be given – please plan accordingly before taking the class.

Make-up Work:	No. Presentations and papers are expected the same day as designated. All team members need to be present; no exceptions are allowed.
Resources:	All students are encouraged to read more from the additional reading list, as well as additional conference or Journals.
Revision Date:	7/16/2022

Advanced database Research Topics of Interest (For Group Presentations) List:

1. Column-Oriented Database:

- M. Stonebraker, D. Abadi, A. Batkin, et al. "C-Store: A Colum-oriented DBMS," Proceeding of the 31st VLDB Conference, 2005, http://db.csail.mit.edu/projects/cstore/vldb.pdf
- D. Abadi, S. Madden, and N. Hachem, "Column-Stores vs. Row-Stores: How Different are They Really?," SIGMOD'08, June, 2008, http://db.csail.mit.edu/projects/cstore/abadi-sigmod08.pdf
- D. Abadi, S. Madden, and M. Ferreira, "Integrating Compression and Execution in Column-Oriented Database Systems," SIGMOD 2006, http://db.csail.mit.edu/projects/cstore/abadisigmod06.pdf
- D. Abadi, et al., "SW-Store: a Vertically Partitioned DBMS for Semantic Web Data Management," VLDB 2009, http://delivery.acm.org/10.1145/1530000/1527469/778_2008_Article_125.pdf?key1=1527469&key2=9259629421&coll=portal&dl=ACM&CFID=69980356&CFTOKEN=94238114
- S. Harizopoulos, V. Liang, D. Abadi, and S. Madden, "Performance Tradeoffs in Read-Optimized Databases," VLDB Conference 2006, http://db.csail.mit.edu/projects/cstore/VLDB06.pdf
- Columnar Database and Data Warehouse: http://www.globaldataconsulting.net/articles/theory/columnar-databases-and-data-warehouse

2. In-Memory Database (IMDB):

- Bell Labs Home Page, " The Dali Main-Memory Storage Manager," http://www.bell-labs.com/project/dali/
- Jun Rao and Kenneth Ross, "Making B^+-Trees cache Conscious in Main Memory," Technical report, Columbia University, 1998, http://www.it.iitb.ac.in/~it603/Project/ref/cacheConsciousBTrees00.pdf
- Peter Boncz, Stefan Manegold, and Martin Kersten, "Database Architecture Optimized for the New Bottleneck: Memory Access," Proceeding of the 25th VLDB Conference, 1999, http://www.vldb.org/conf/1999/P5.pdf
- Tobin Lehman and Michael Carey, "A Study of Index Structures for Main Memory Database Management Systems," Proceeding of the 12th Int'l Conference on VLDB, 1986, http://www.vldb.org/conf/1986/P294.PDF
- Steve Graves, "In-Memory Database Systems," Linux Journal, September 2002, http://delivery.acm.org/10.1145/570000/566959/6133.html?key1=566959&key2=6548629421&coll=GUIDE&dl=GUIDE&CFID=69980356&CFTOKEN=94238114

3. Database in the Cloud:

- Brian Hayes, Cloud Computing," Communications of the ACM, 07/08, Vol. 51, No.7.
- Amazon SimpleDB™, http://aws.amazon.com/simpledb/#details
- F. Change, J. Dean, et al, "Bigtable: A Distributed Storage System for Structured Data," Google Inc., OSDI 2006, http://labs.google.com/papers/bigtable-osdi06.pdf
- MongoDB, "The best features of document database, key/values stores, and RDBMS in one," http://www.mongodb.org/display/DOCS/Home

- James Urquhart, "The Wisdom of Clouds: Microsoft and FathomDB target 'relational' Clouds," http://news.cnet.com/8301-19413_3-10184579-240.html
- Vertica, " Performance On-demand with Vertica Analytic Database for the Cloud," http://www.vertica.com/cloud?source=googleppc&_kk=cloud%20data&_kt=959f6441-46d6-4e8f-a115-52210102ca8e&gclid=CJuOgI2x8JsCFSRPagod-VnI_A
- Jason Massie's Blog, "Private Clouds and the RDBMS," http://jasonmassie.com/archive/2009/07/private-clouds-and-the-rdbms/
- A. Lenk, et al., "What's Inside the Cloud? An Architectural Map of the Cloud Landscape," CLOUD'09, Proceedings of the 2009 ICSE Workshop on Software Engineering Challenges of Cloud, http://portal.acm.org/results.cfm?query=Cloud&querydisp=Cloud&source_query=Owner%3AACM&start=21&slide=1&srt=score%20dsc&short=0&source_disp=&since_month=&since_year=&before_month=&before_year=&coll=portal&dl=ACM&termshow=matchall&range_query=&CFID=69980356&CFTOKEN=94238114
- Amazon SimpleDB: http://aws.amazon.com/simpledb/
- Amazon Elastic Compute Cloud (Amazon EC2): http://aws.amazon.com/ec2/
- Windows Azure Platform: http://www.microsoft.com/windowsazure/

4. SQL and Stream Processing:

- E. Liarou, et al., "Exploiting the Power of Relational Databases for Efficient Stream Processing," EDBT '09, Proceedings of the 12th Int'l Conf on Extending Database Technology, http://delivery.acm.org/10.1145/1520000/1516398/p323-liarou.pdf?key1=1516398&key2=5799629421&coll=portal&dl=ACM&CFID=69980356&CFTOKEN=94238114
- M. Gaber, et al., "Mining Data Streams: A Review," SIGMOD 2005, http://delivery.acm.org/10.1145/1090000/1083789/p18-gaber.pdf?key1=1083789&key2=8510729421&coll=portal&dl=ACM&CFID=69980356&CFTOKEN=94238114
- M. Stonebraker, et al., "The 8 Requirements of Real-Time Stream Processing," Technical Report, Brown University, http://www.cs.brown.edu/research/db/publications/8rulesSigRec.pdf
- Stanford STREAM Home Page: http://infolab.stanford.edu/stream/
- The Aurora Project at Brown University, http://www.cs.brown.edu/research/aurora/
- D. Carney, et al., "Monitoring Streams – A New Class of Data Management Applications," VLDB 2002, http://www.cs.brown.edu/research/aurora/vldb02.pdf
- S. Chandrasekaran, J. Hollerstein, et al., "TelegraphCQ: Continuous Dataflow Processing for an Uncertain World," Proceedings of 2003 CIDR Conference, http://www.eecs.berkeley.edu/~franklin/Papers/TCQcidr03.pdf

5. Map Reduce & SQL Processing:

- Jeffery Dean and Sanjay Ghemawat, "MapReduce: Simplified Data Processing on Large Clusters, USENIX, 2004, http://labs.google.com/papers/mapreduce-osdi04.pdf
- Stanford University, " Generalizing Map-Reduce: The Computational Model Map-Reduce-Like Algorithms: Computing Joins," http://www.stanford.edu/class/cs345a/map-reduce2.ppt
- H. yang, A. dasdan, R. Hsiao, and D. Parker, "Map-Reduce-Merge: Simplified Relational Data Processing on Large Clusters," SIGMOD '07, June 2007, http://delivery.acm.org/10.1145/1250000/1247602/p1029-yang.pdf?key1=1247602&key2=4670539421&coll=GUIDE&dl=GUIDE&CFID=47469407&CFTOKEN=99969073

- Vertica, "The Database Column: MapReduce II," http://databasecolumn.vertica.com/2008/01/mapreduce-continued.html
- F. Afrati, and J. Ullman, "A New Computation Model for Rack-Based Computing," PODS '09, ACM.
- R. Chaiken, "SCOPE: Easy and Efficient Parallel Processing of Massive Data Sets," VLDB 2008, http://delivery.acm.org/10.1145/1460000/1454166/p1265-chaiken.pdf?key1=1454166&key2=4530729421&coll=portal&dl=ACM&CFID=69980356&CFTOKEN=94238114
- B. Cooper, et al., "PNUTS: Yahoo!'s Hosted Data Serving Platform," VLDB 2008, http://www.brianfrankcooper.net/pubs/pnuts.pdf
- M. Stonebraker, D. DeWitt, et al., "Compare MapReduce to RDBMS," http://www.dbms2.com/2009/04/14/stonebraker-dewitt-et-al-compare-mapreduce-to-dbms/
- David Dewitt, "MapReduce: A Major Step Backwards," http://www.databasecolumn.com/2008/01/mapreduce-a-major-step-back.html
- Jeff Dean, Sanjay Ghemawat, MapReduce: Simplified Data Processing on Large Clusters: http://labs.google.com/papers/mapreduce-osdi04-slides/index.htmlhttp://labs.google.com/papers/mapreduce-osdi04-slides/index.html

6. Solid State Devices and Database:

- S. Lee, et al., "A Case for Flash Memory SSD in Enterprise Database Applications," SIGMOD '08, ACM 2008, http://delivery.acm.org/10.1145/1380000/1376723/p1075-lee.pdf?key1=1376723&key2=8631539421&coll=GUIDE&dl=GUIDE&CFID=47470231&CFTOKEN=93521125
- Adam Eventhal, "Flash Storage Memory," Communications of the ACM, 07/08, Vol. 51, No.7, http://delivery.acm.org/10.1145/1370000/1364796/p47-leventhal.pdf?key1=1364796&key2=2941539421&coll=ACM&dl=ACM&CFID=47470386&CFTOKEN=92795940
- Information Week, "Scaling the Data Warehouse," http://www.informationweek.com/news/business_intelligence/warehouses/showArticle.jhtml?articleID=210900005, Oct. 2008.
- S. Lee, et al., "Advances in Flash Memory SSD Technology for Enterprise Database Applications," SIGMOD '09, http://delivery.acm.org/10.1145/1560000/1559937/p863-lee.pdf?key1=1559937&key2=7867629421&coll=GUIDE&dl=GUIDE&CFID=69980356&CFTOKEN=94238114
- S. Lee, et al., "A Case for Flash Memory in Enterprise Database Applications,",2008 ACM SIGMOD, Int'l Conference on Management Data, http://delivery.acm.org/10.1145/1380000/1376723/p1075-lee.pdf?key1=1376723&key2=9387629421&coll=GUIDE&dl=GUIDE&CFID=69980356&CFTOKEN=94238114
- J. Shin, et al., "ETL Design Exploration in Reconfigurable High-Performance SSD for Server Applications," ACM, Proceedings of 23rd ICS '09, http://delivery.acm.org/10.1145/1380000/1376723/p1075-lee.pdf?key1=1376723&key2=5338629421&coll=GUIDE&dl=GUIDE&CFID=69980356&CFTOKEN=94238114

7. Semantic Web and Data Integration:

- W3C, "Resource Description Framework (RDF)," http://www.w3.org/RDF/
- W3C, SPARQL Working Group Wiki," http://www.w3.org/2009/sparql/wiki/Main_Page
- W3C, "Web Ontology Language (OWL)," http://www.w3.org/2004/OWL/
- Rdf:about, "Comparing RDF to XML and RDBMS," http://www.rdfabout.com/comparisons.xpd

- D. Holland, et al. “Choosing a Data Model and Query Language for Provenance,” Technical report, Harvard University, http://www.eecs.harvard.edu/~kiran/pubs/ipaw08.pdf
- Nicole Alexander, et al., “An RDF Data Model for the Semantic Web,” Oracle, http://lambda.csail.mit.edu/~chet/papers/others/o/oracle/w3d_rdf_data_model.pdf
- J. Bailey, et al., “Semantic Web Query Languages,” Technical Report, University of Munich.
- Wikipedia.org, “Web Ontology Language (OWL),” http://en.wikipedia.org/wiki/Web_Ontology_Language
- Wikipedia.org, “SPARQL,” http://en.wikipedia.org/wiki/SPARQL
- OpenLink Software, “Mapping Relational Data to RDF with Virtuoso’s RDF Views,” http://www.openlinksw.com/virtuoso/Whitepapers/html/rdf_views/virtuoso_rdf_views_example.html
- Oracle, “Oracle Semantic Technologies Center,” http://www.oracle.com/technology/tech/semantic_technologies/index.html
- Wikipedia, “ Semantic web,” http://en.wikipedia.org/wiki/Semantic_Web