What have you learnt?

What Have You Learnt?

- Traditional vs. Modern Database Systems

- Essential skills for being a researcher

What Have You Learnt?

- Traditional vs. Modern Database Systems
  - Gain a much deeper understanding on the Big Data movement
  - Form your own opinion on what's novel about Big Data systems

- Essential skills for being a researcher

History

When	What
Early 1960 – Early 1970	The Navigational Database Empire
Mid 1970 – Mid 1980	The Database World War I
Mid 1980 – Early 2000	The Relational Database Empire
Mid 2000 – Now	The Database World War II

When	What
Early 1960 – Early 1970	The Navigational Database Empire
Mid 1970 – Mid 1980	The Database World War I
Mid 1980 – Early 2000	The Relational Database Empire
Mid 2000 – Now	The Database World War II

W 1/17	Background	Database Systems: Achievements and Opportunities (1990)
		The Asilomar Report on Database Research (1998)
F 1/19	Data Model	A Relational Model of Data for Large Shared Data Banks (1970) What Goes Around Comes Around (1960-1970, Sec I~IV only)

When	What
Early 1960 – Early 1970	The Navigational Database Empire
Mid 1970 – Mid 1980	The Database World War I
Mid 1980 – Early 2000	The Relational Database Empire
Mid 2000 – Now	The Database World War II

Traditional DBMS	A History and Evaluation of System R (1981) The Design of Postgres (1986)
	An Overview of Data Warehousing and OLAP Technology (1997) Data Lake and Parallel Database
Transaction Management	Granularity of Locks and Degrees of Consistency in a Shared Data Base (1976, Part 1 only) Granularity of Locks and Degrees of Consistency in a Shared Data Base (1976, Part 2 only)
	On Optimistic Methods for Concurrency Control (1981) Concurrency Control
Query Optimization	Access Path Selection in a Relational Database Management System (1979) The Volcano Optimizer Generator: Extensibility and Efficient Search (1993)
	Eddies: Continuously Adaptive Query Processing (2000) Guest Speaker: A Brief Overview of Query Optimization (Wentao Wu @ MSR)
Interactive Analytics	Data Cube: A Relational Aggregation Operator Generalizing Group-by, Cross-tab, and Sub-totals (1997) An Array-Based Algorithm for Simultaneous Multidimensional Aggregates (1997)

When	What
Early 1960 – Early 1970	The Navigational Database Empire
Mid 1970 – Mid 1980	The Database World War I
Mid 1980 – Early 2000	The Relational Database Empire
Mid 2000 – Now	The Database World War II

OLAP

Column Store	C-store: a column-oriented DBMS (2005)
	Dremel: Interactive Analysis Of Web-Scale Datasets (2010) Column-Stores vs. Row-Stores: How Different Are They Really? (2012)
MapReduce and Beyond	MapReduce: Simplified Data Processing on Large Clusters (2003) A Comparison of Approaches to Large-Scale Data Analysis (2009)
	Resilient Distributed Datasets: A Fault-tolerant Abstraction for In-memo Spark SQL: Relational Data Processing in Spark (2015)

OLTP

NoSQL	Chord: A Scalable Peer-to-peer Lookup Protocol for Internet Applicatior Bigtable: A Distributed Storage System for Structured Data (2006)
	Dynamo: Amazon's Highly Available Key-Value Store (2007) CAP Twelve Years Later: How the "Rules" Have Changed (2012)
NewSQL	OLTP Through the Looking Glass, and What We Found There (2008) Hekaton: SQL Server's Memory-optimized OLTP Engine (2013)
	Efficiently Compiling Efficient Query Plans for Modern Hardware (2011) Scalable SQL and NoSQL data stores (2010)

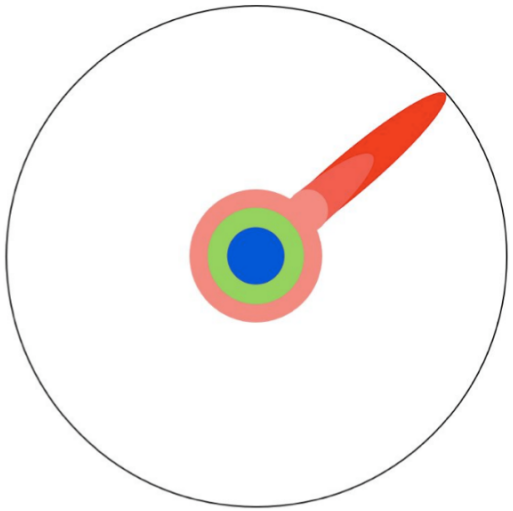
What Have You Learnt?

- Traditional vs. Modern Database Systems
  - Gain a much deeper understanding on the Big Data movement
  - Form your own opinion on what's novel about Big Data systems

- **Essential skills for being a researcher**

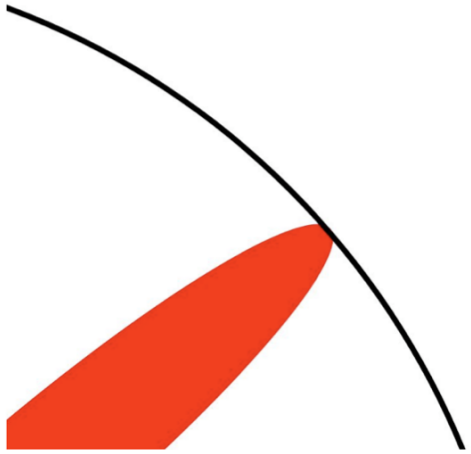
Essential Skills

Reading Papers

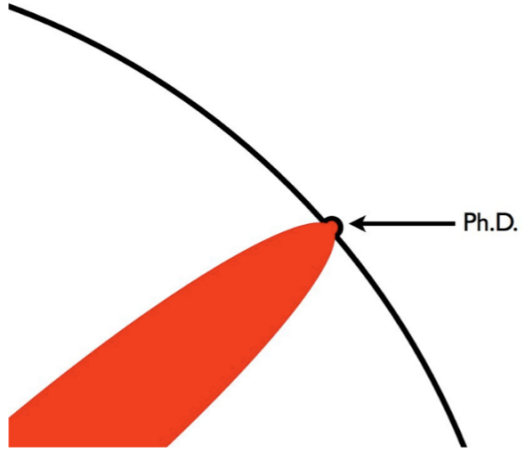


Critical Thinking

- Reviewing Papers
- Asking Questions
- ...



- Presentation
  - Giving Talks
  - ...



Great Achievements!

- You have read 16 + 18 = 34 papers
- You have asked 10+ questions
- You have written 30+ short reviews and 1 long review
- You have written a blog post
- You have given 1 talk

Not Finished Yet

- April 2nd
  - Long Review
- April 10th
  - **Teaching Evaluation**
- April 11th
  - Upload Posters (10 am)
  - Poster Session (2pm @ T9204 W&E)
- April 15th
  - Final Report (11:59 pm)