Section 1: Week 1: Database Bibliography

Nate Bachmeier

TIM-8120: Distributed Systems

December 7th, 2019

North Central University

**Bibliography**

Data is the modern world's oil. It powers algorithms that can detect when we are falling into or out of love (Morris, 2016), secretly pregnant (Hill, 2012), and even topple national states (Cadwalladr, 2017). Oil begins its life as crude, and only through refining does it become much more. Similarly, data begins as sequences of unstructured bytes that Business Intelligence can unlock via statistical inferences (e.g., machine learning) and modeling techniques. Before formulating these predictions, a datastore needs to be chosen to host the data. Datastores come in various shapes and sizes to meet the needs of different volume, velocity, variety access pattern requirements (Taleb, Serhani, & Dssouli, 2018). Afterward, related facts and relationships transform into higher-order knowledge to enable decision processes.

## ****Data**** Storage ****Management**** in Cloud Environments (2019)

This is an article about different cloud databases (Yaser, Adel, & Rajkumar, 2017) and probably makes for a good opener.

Here is a related article A survey on data storage and placement methodologies for Cloud-Big Data ecosystem (Somnath Mazumdar1 , Daniel Seybold2 , Kyriakos Kritikos3\* and Yiannis Verginadis4)

## Data aggregation processes: a survey, a taxonomy, and design guidelines (2019)

The Data Aggregation Taxonomy (DAGGTAX) attempts to enumerate the different permutations of aggregation functions. These variations have (1) periodic, sporadic, or aperiodic timing requirements, as they (2) prepare, aggregate, and perform post-handling of data (Cai, Gallina, & Nyström, 2019).

## Big Data Quality: A Survey (2018)

Good data in, good aggregations out. It is important to measure your data quality upfront and often (Taleb, Serhani, & Dssouli, 2018). Read the article and put more content here.

* Database Technologies
  + Cloud Storage
  + Next Topic: A Graph Database of Yelp Dataset Challenge 2018 Krommuller
  + A Comprehensive Survey of OLAP: Recent Trends
  + <something no sql specific here>
* Migration
  + Data Storage Management in Cloud Environments
  + Document-Oriented Data Schema for Relational Database Migration to NoSQL
  + Bringing SQL databases to key-based NoSQL databases a canonical approach
* Business Intelligence
  + BI for Enterprise Systems
  + < -or- something BI specific here>
* BI Transformations
  + Aggregations
  + Big Data Quality