

Introduction to Data Wrangling and Management

Reading

- Links to reading here

Notes

- Walkthrough of how to access the ACM resources
- Working with data / information has many facets.
 - Plan for what information to collect (*?)
 - Aggregation of data from various sources and formats (*)
 - Storage and updating of the data (*)
 - Access policies (*?)
 - Querying the data (*)
 - Visualizing information
 - Data Science Process Flowchart¹
- In-class activity: Plan a system for online evaluations at the college.
 - Who are our different clients?
 - What do they each need from the system?
 - What information do we need to store?
 - Where is that information coming from?
 - What information do the different clients need to have access to?
 - What are the logical entities that comprise that information? What are their relations?
 - What kind of access/service do we need to provide to our clients?
- In-class activity: Create application that uses the Twitter feed
 - What information might we get from Twitter?
 - What kind of information can we produce from that?
 - What do we need to store, and how?
 - What service would we offer?

¹https://en.wikipedia.org/wiki/Data_science#/media/File:Data_visualization_process_v1.png