# Why Manage Data

* data is very valuable; data collection is usually time consuming and hard
* Large amount of data and documents are being generated with high growth rate
* Multiple sources of data

# What is Data Management?

* Data management is the development, execution and supervision of plans, policies, programs and practices that *control, protect, deliver and enhance the value of data and information assets*

# Issues in Data Management

Data management Deals with issues:

* replication and persistence
* security

# Privacy versus Confidentiality

Privacy is (for our purposes) having control over how one shares oneself with others.

Security as the protection of data, preventing it from being improperly used

Confidentiality is information privacy, how information about an individual is treated and shared

* For many apps or services, you must accept their data sharing policies or you can’t use their services fully
* There could be an agent to interact in a narrative form with individual consumers
  + Are you willing to share your health data with company X?

# Social media and the loss of confidentiality

Target is predicting which women is pregnant from their purchases

many things can be predicted from Facebook “likes”

# Compliances and Regulations

* Ethics: the moral handling of data
* regulations to ensure that confidentiality is protected
* The process of ensuring you meet regulations is called compliance
* PCI (Payment Card Industry) standard
  + reduce credit card fraud
    - By placing specific regulations- information in encrypted
  + Companies who handle credit card have to with PCI standards
  + Audit (validation of compliance) is done annually

# Data Management and Data Science

* internet advertising: what implicit and explicit data is stored about a user
  + implicit: facebook like motorcycle, it will push motorcycle
  + explicit: something that is stated plainly-ur like in facebook
* medical informatics: for predicting fungal infections from nursing notes, the team needs to abide by confidentiality and security