Computing

Yi Fang, PhD

Department of Computer Engineering
Santa Clara University

yfang@scu.edu
http://www.cse.scu.edu/~yfang/

This Talk

- Part I — Computing

- Part II — Computing at SCU

- Part III — Data Science

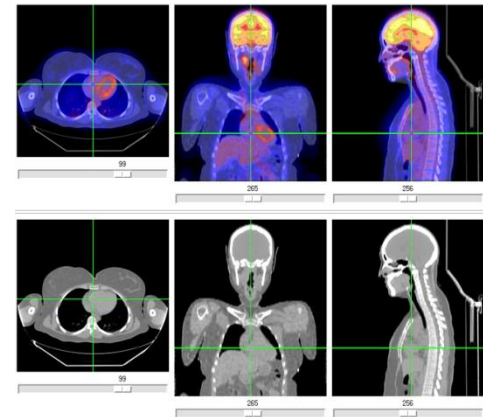- Part IV — Today's activity: PageRank

PART I — COMPUTING

What is Computing?

- Analysis, design and development of computer systems

- It is not just about programming

- It teaches you how to think more methodically and how to solve problems more effectively

Computing is everywhere!

What is Computing?
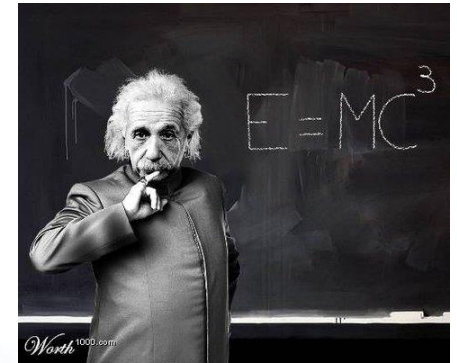
- Computing includes a variety of fields:
  - Mathematics
  - Computer science
  - Computer engineering
  - Information science
  - Electrical engineering

What is Computing?

- ## What is a computer professional?
  - Will I have to grow fuzzy hair?

- ## What does s/he do?
  - Will I have to sit in front of
    a computer all day?

- ## What kind of people will I work with?
  - Will I have to become a geek nerd?

What is Computing?

- FUN, COOL, and EXCITING
  - Cutting edge projects
  - Exciting and talented people
  - All over the world, in every sector
  - Significant impact on society and our planet

Why Study Computing?

- **Intellectually interesting**
  - Logical reasoning and mathematical thinking
  - Possible workings of the human mind

Why Study Computing?

- **Computing supports and links to most other areas of study**
    - Computing and neuroscientists – the brain
    - Computing and Biologists – Genome
    - Computing and Meteorologists – weather prediction
  - Future scientists require basic knowledge of Computing

Why Study Computing?

- **Computing teaches problem solving**

    - Decomposition, abstraction, modular design

    - Analysis and design are carefully reviewed

    - Always new methods being investigated

Why Study Computing?

- **Computing builds teamwork and leadership skills**
  - Plan, organize, control, lead complex projects
  - Learn to deal with mix of talents
  - Estimate and deal with risk

Why Study Computing?

- **Computing develops life-long learning skills ...** *"Change is the only constant"*

  – Promotes learning to learn

  "if GM had kept up with the technology like the computer industry has, we would all be driving $25.00 cars that got 1,000 miles to the gallon" – Bill Gates

  – Exponential growth makes many predictions look foolish

False Predictions

- *"I think there is a world market for maybe five computers"* -- Thomas J. Watson, founder and Chairman of IBM, 1943.

- *"Computers in the future may weigh no more than 1.5 tons"* -- *Popular Science*, 1949.

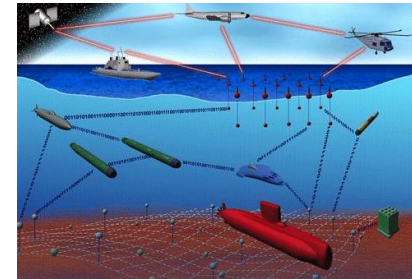- *"640K ought to be enough for anybody"* -- Bill Gates, 1981.
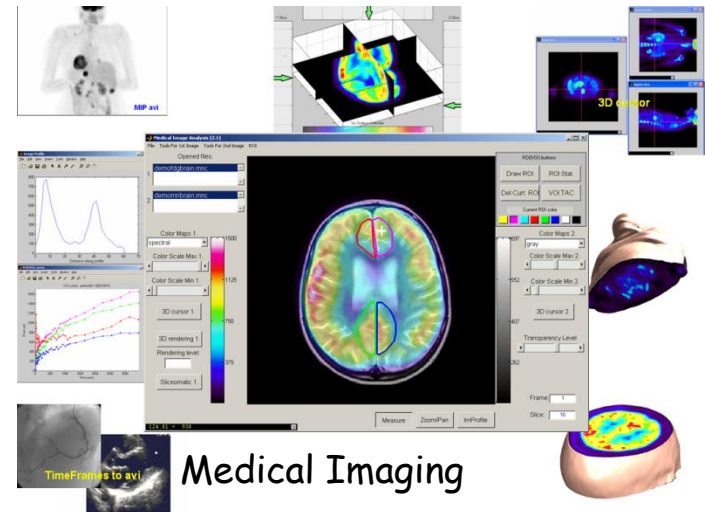
Future Applications

Self-driving car

Personalized Healthcare

Transforming the nation's defense

Internet of Things

Medical Imaging

Comp. Science & Engineering

- Computer science
  - Often more mathematical
  - Computability theory
  - Algorithmic complexity
- Computer engineering
  - Often more hardware-oriented
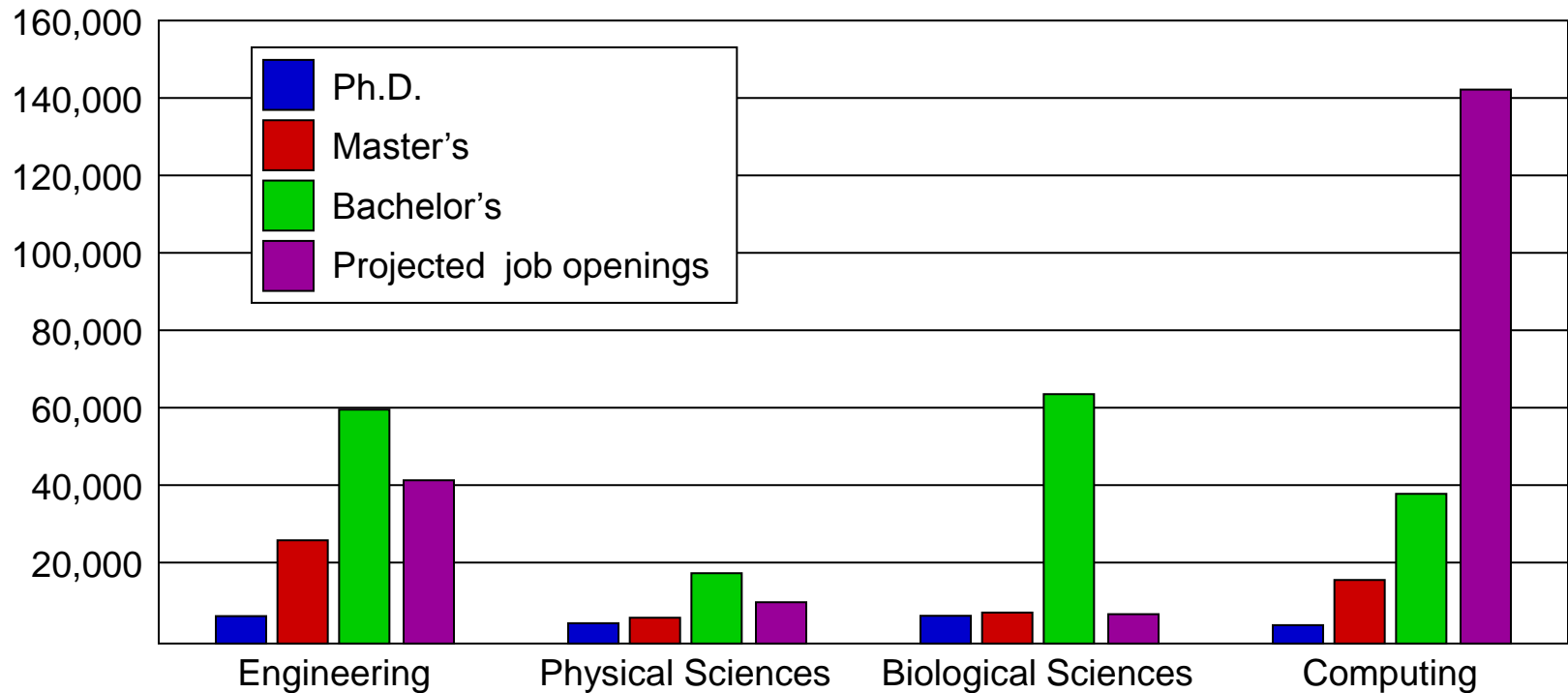 - Image and signal processing
  - Computer graphics

Career Opportunities

- System architect
- Network engineer
- Computer architect
- Software engineer
- Security specialist
- Game designer
- Test engineer
- Entrepreneur, musician, athlete, and more

Degree Production *vs.* Job Openings

Be Creative!

- Computing is the only tech field in which you can create a product from scratch and commercialize it independently

PART II — COMPUTING AT SCU

Computing Degrees at SCU

- Undergraduate degrees
  - Computer science and engineering (CSE)
  - Web design and engineering (WDE)
  - Mathematics and computer science
- Graduate degrees
  - Computer science and engineering
  - Software engineering
- 5-year Master's program

Undergraduate CSE

- Combination of computer science and computer engineering
- Focuses on theoretical and practical aspects of computing
- Design and construction of both hardware and software systems
  - Computer networks, operating systems, algorithms, compilers, software engineering, embedded programming, Web programming, robotics, 3D animation

Undergraduate WDE

- New major started in 2009
  - One of the first such programs in the country
- Combines computing with other disciplines:
  - Graphic arts
  - Communication
  - Sociology
- What will these specialized graduates do?
  - Improve Web infrastructure
  - Develop interactive, multimedia content
  - Analyze the huge amount of information on the Web (Big data)
  - Understand the societal impact of the Web
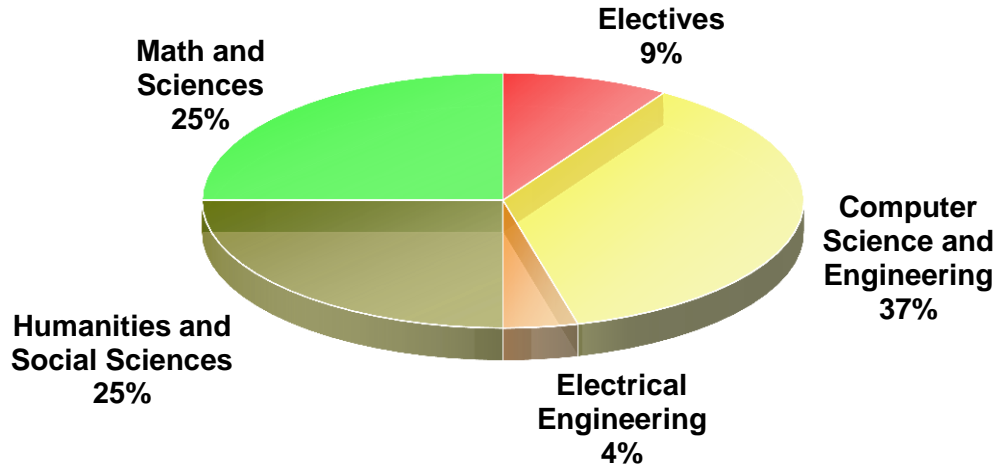
Where Will You Work?

- Recent graduates went to work for:
  - Cisco, Apple, Microsoft, IBM, Google, Facebook, Groupon, Amazon, Anritsu, F5 Networks
  - Starting salary range: $70K–$100K
- Recent graduates also continued their education:
  - Ph.D program at Berkeley, UCSD, etc.
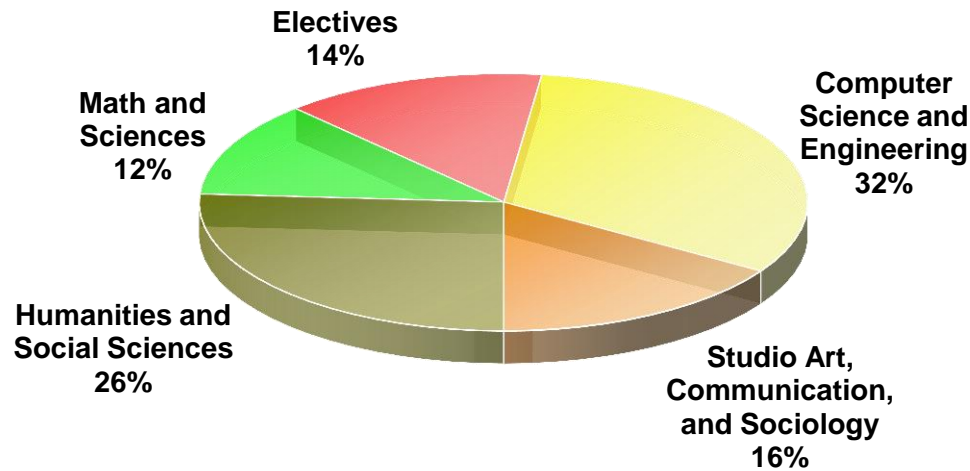 - M.S. programs at SCU, CMU, Stanford, etc.

Coursework

Computer Science and Engineering

- Electives 9%
- Computer Science and Engineering 37%
- Electrical Engineering 4%
- Humanities and Social Sciences 25%
- Math and Sciences 25%

Web Design and Engineering

- Electives 14%
- Computer Science and Engineering 32%
- Studio Art, Communication, and Sociology 16%
- Humanities and Social Sciences 26%
- Math and Sciences 12%

PART III — DATA SCIENCE

Data Science

Extraction of Knowledge from data

What is a Data Scientist?

Data scientists?

engineering

math and statistics

nerds

nerds

nerds

comp sci

nerds

hacking

awesome nerds
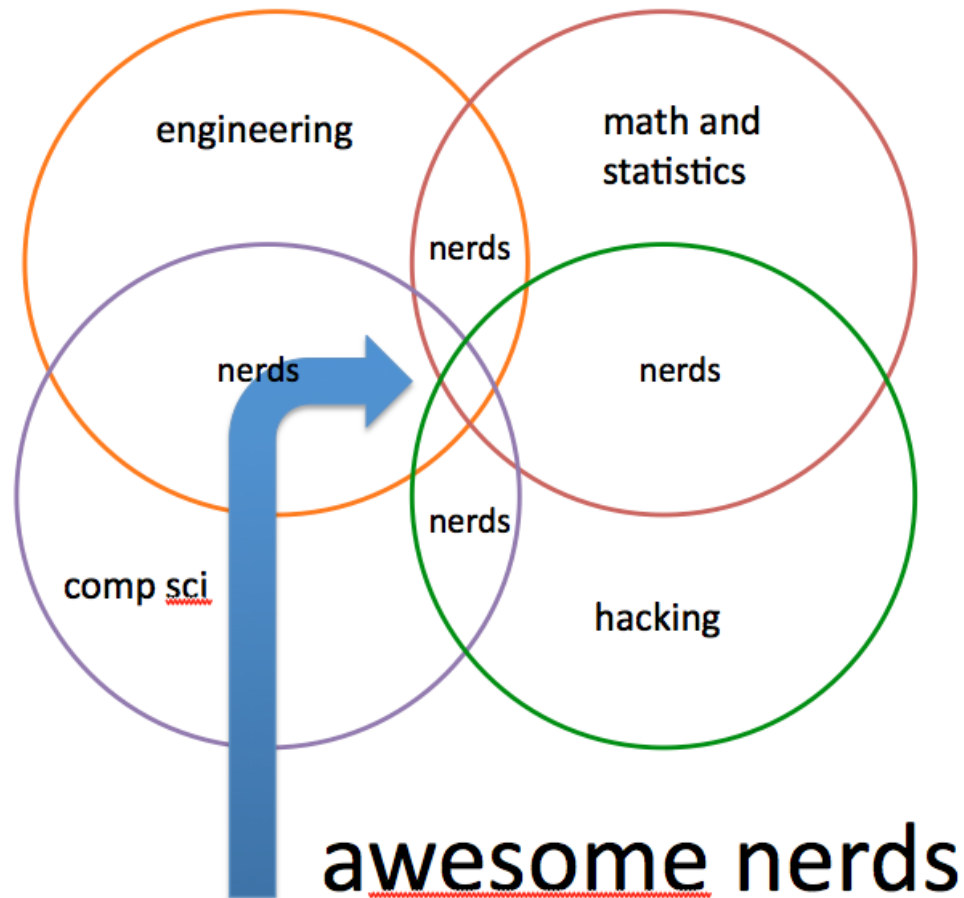
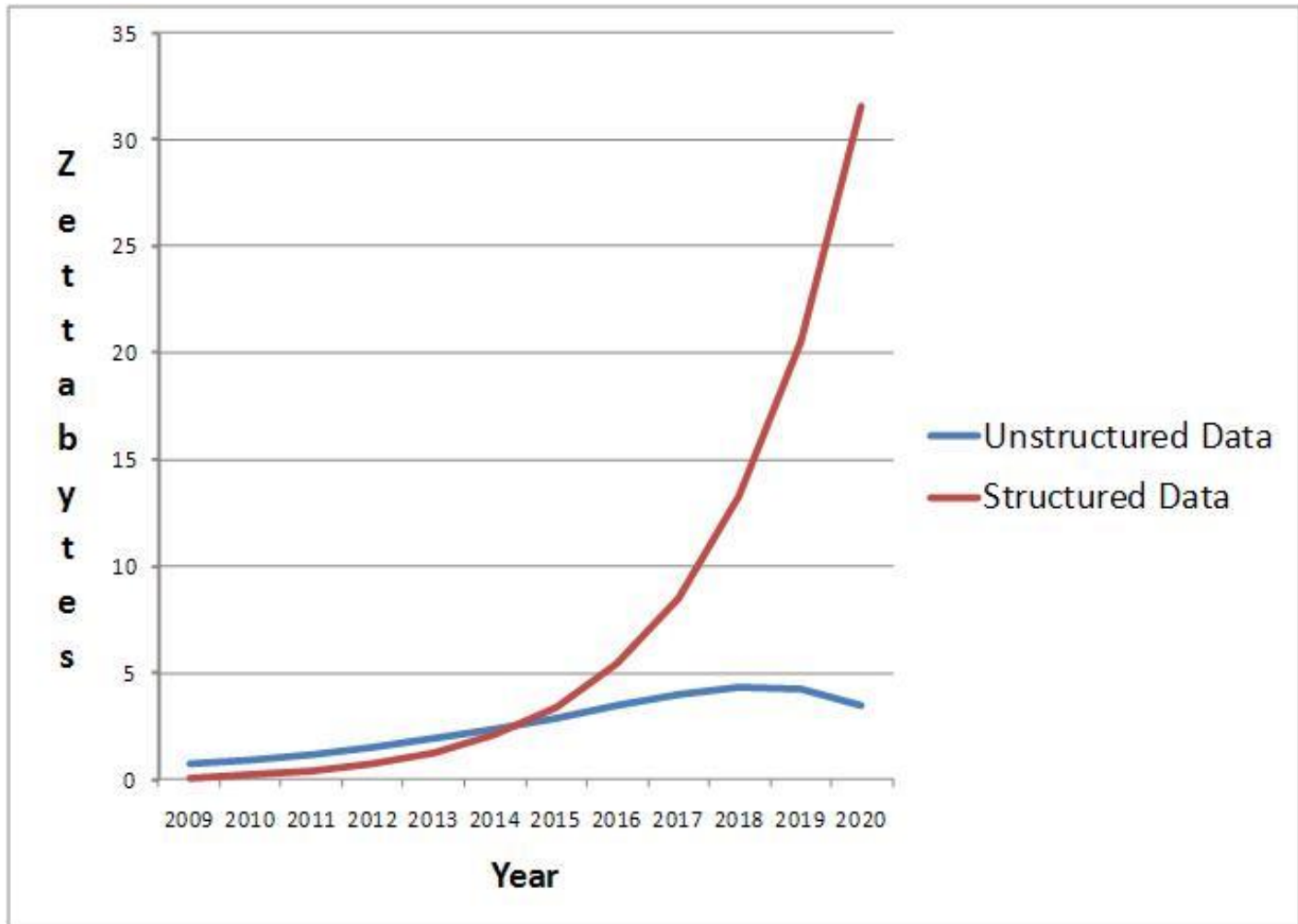# Unstructured Data

- Documents
- Webpages
- Images
- Audio
- Video
- More…

Growth

Big Data

Any dataset where the size or speed of incoming data causes difficulties in processing

- Volume
- Velocity
- Variety

Law of Data

18 Months

the amount of time for digital data to double

Why do you care?

"Every single industry will be totally revolutionized by big data"

- Joe Tucci, EMC

Big Data Examples

- Google: > 100 PB; > 1T indexed URLs
- Facebook: 1 billion users; 40 billion photos
- YouTube: > 750 PB
- Twitter: > 55 billion tweets/year;

  > 150 million/day; 1700/second
- Text messages: 6.1 T/year; 876/person/year
- US cell calls: 2.2 T minutes/year;

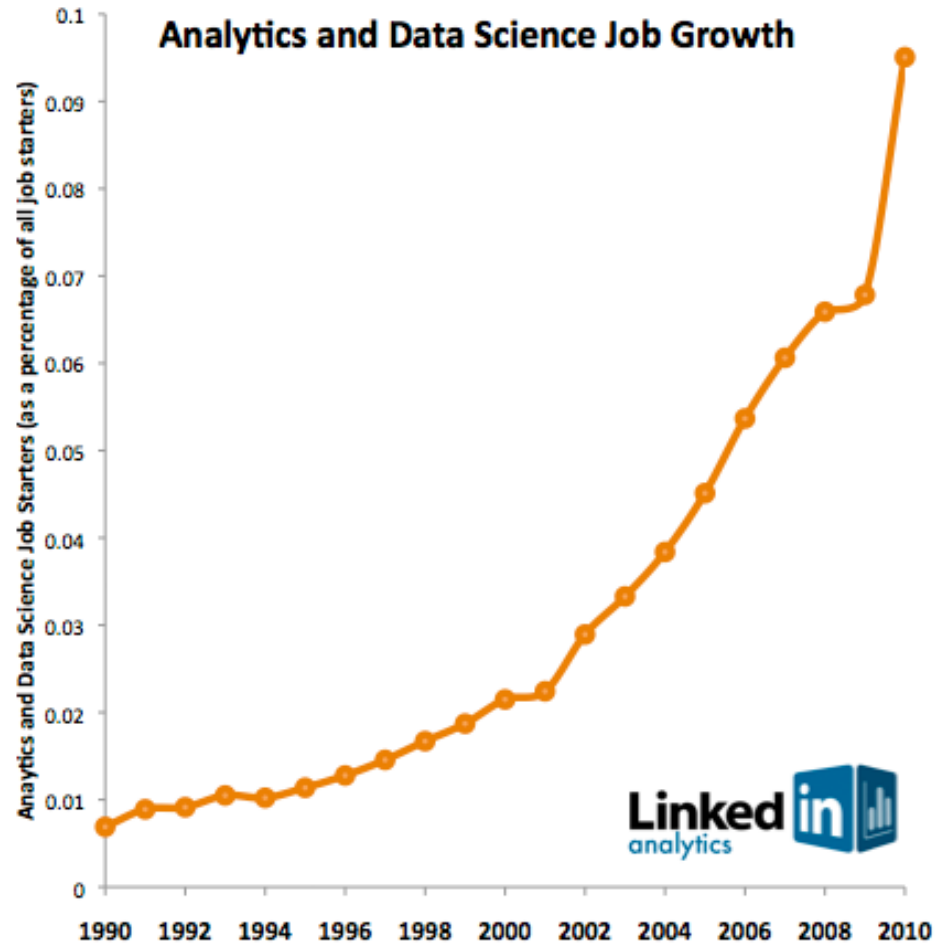 19 minutes/person/day

  ~ size of a YouTube

Driving Forces

Social

Mobile

Cloud

Sensors and The Internet of Things

Data Science Job Listing

Analytics and Data Science Job Growth

Data Scientist:
The Sexiest Job of the 21st Century

Meet the people who can coax treasure out of messy, unstructured data.
by Thomas H. Davenport and D.J. Patil

When Jonathan Goldman arrived for work in June 2006 at LinkedIn, the business networking site, the place still felt like a start-up. The company had just under 8 million accounts, and the number was growing quickly as existing members invited their friends and colleagues to join. But users weren't seeking out connections with the people who were already on the site at the rate executives had expected. Something was apparently missing in the social experience. As one LinkedIn manager put it, "It was like arriving at a conference reception and realizing you don't know anyone. So you just stand in the corner sipping your drink—and you probably leave early."

38

PART IV — TODAY'S ACTIVITY

The History of PageRank

- PageRank was developed by Larry Page (hence the name *Page*-Rank) and Sergey Brin.
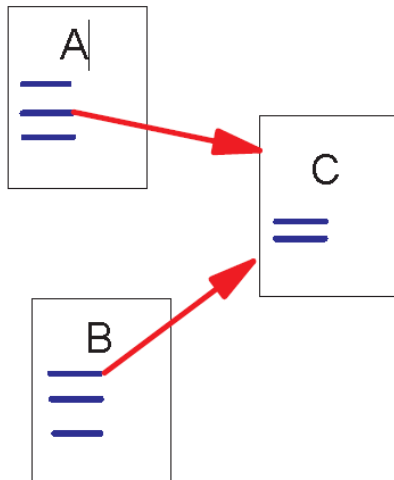

- It is first as part of a research project about a new kind of search engine. That project started in 1995 and led to a functional prototype in 1998.

- Shortly after, Page and Brin founded Google.

Link Structure of the Web

In-links and Out-links:
➢A and B are C's in-links
➢C is A and B's out-link

Assess the importance of a page based on links

Intuitively, a webpage is important if it has a lot of in-links.

What if a webpage has only one link coming from www.yahoo.com?

Intuition

"a page is important if <u>many important</u> pages <u>exclusively</u> link to it."

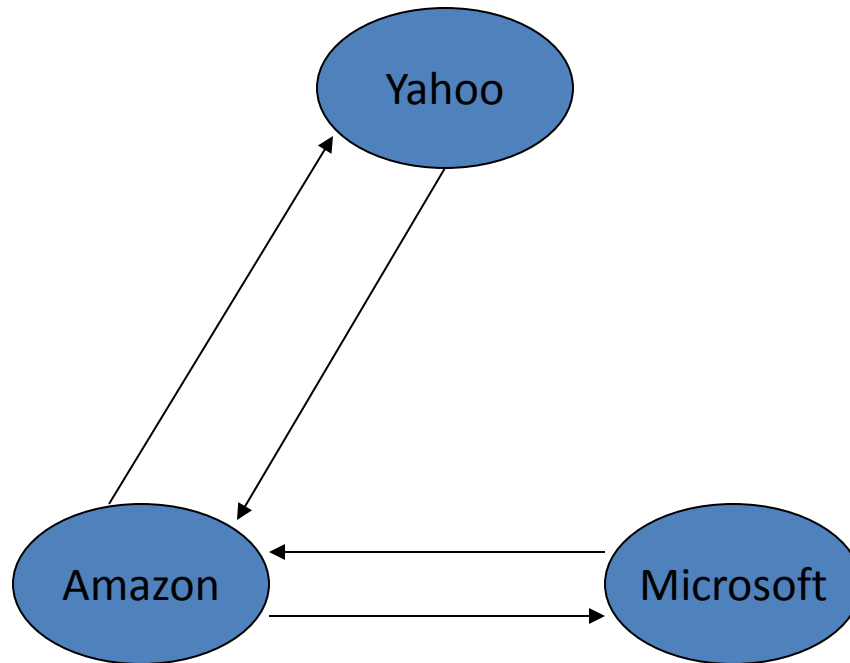# PageRank Calculation

- Page i's importance (or PageRank) is the sum of the votes from its inlinks

- Each link's vote is proportional to the importance (or Pagerank) of its source page

- If page j with PageRank $PR(j)$ has n outlinks, each link gets $PR(j)/n$ votes
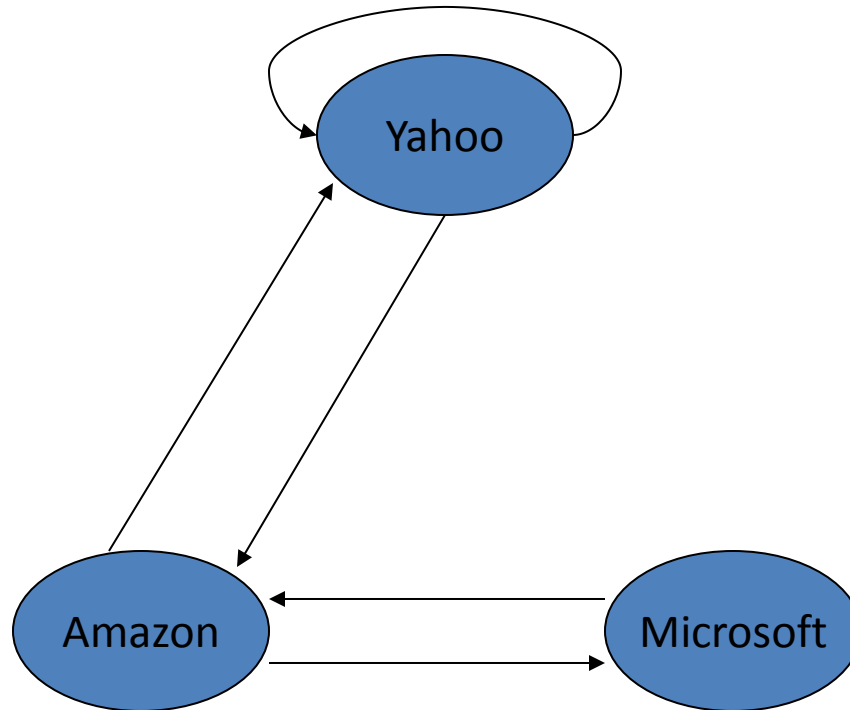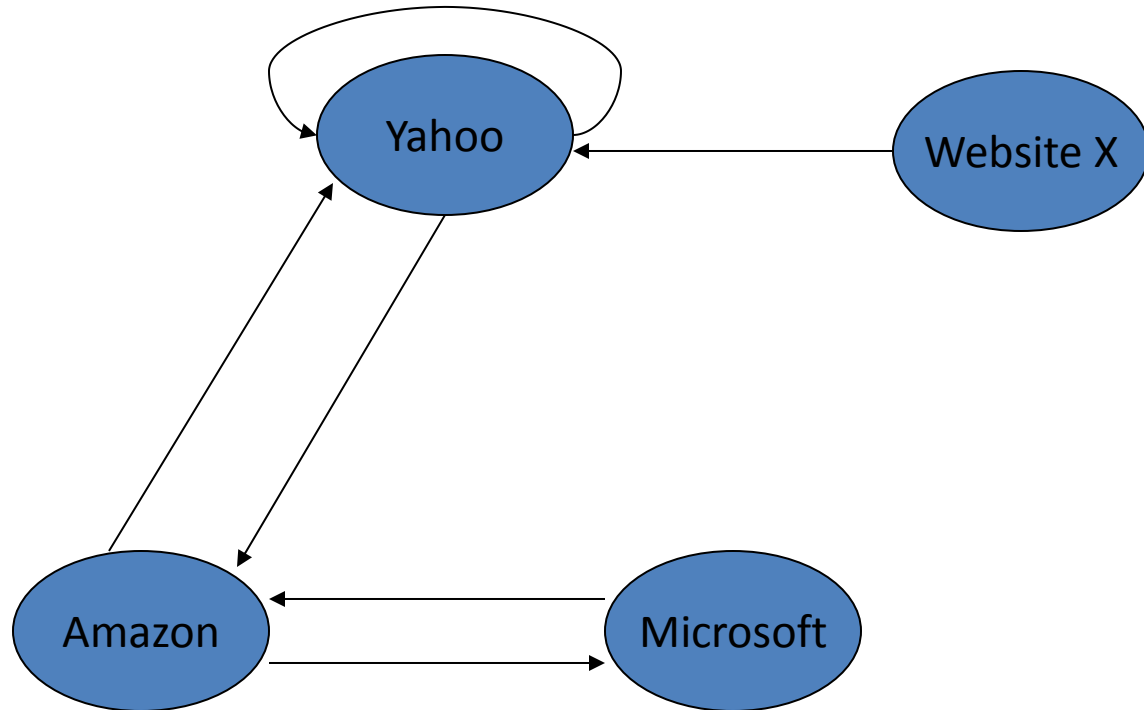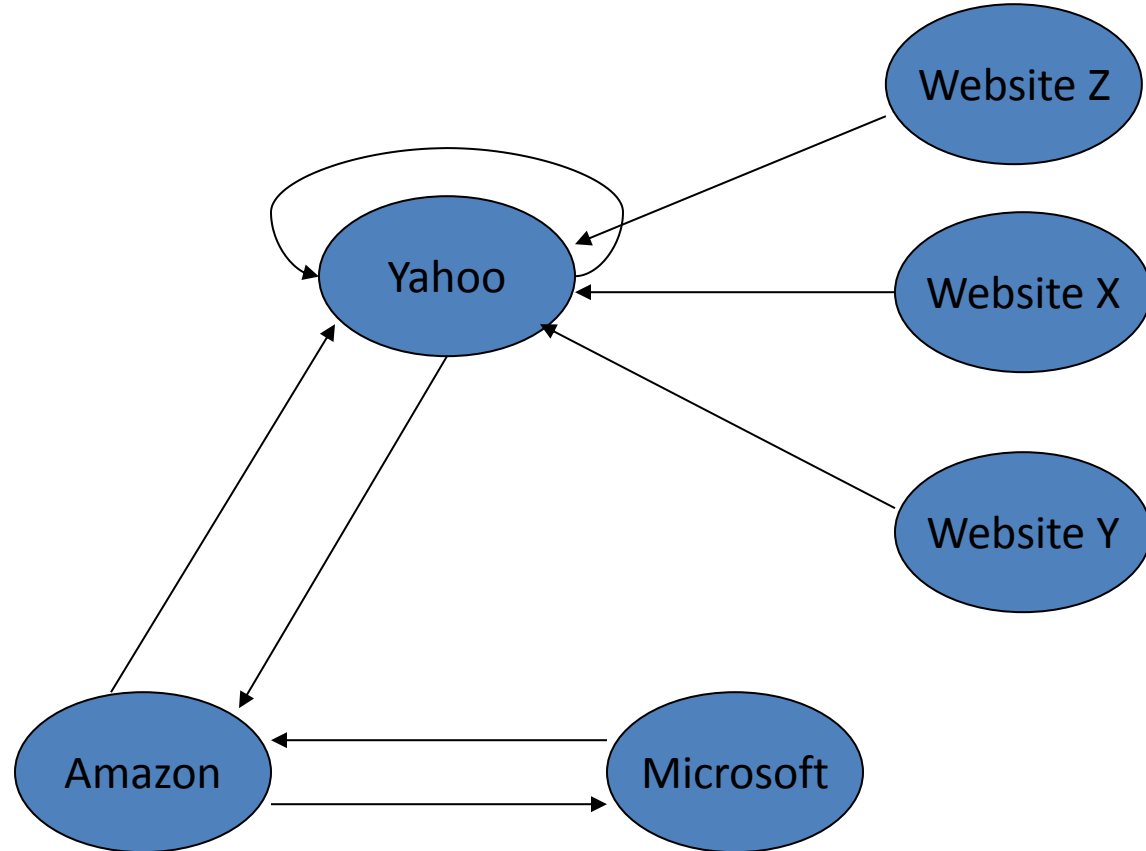
Example 1

Example 2
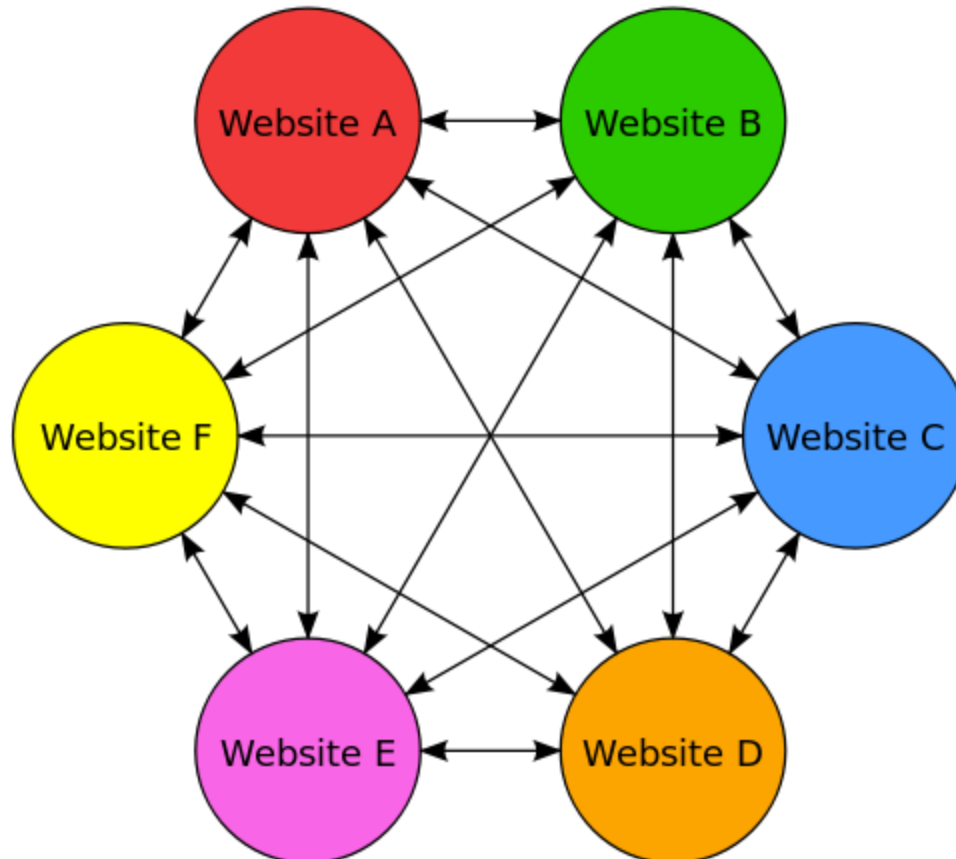
Example 3

Example 4

Search Engine Optimization

- Increase the visibility of your website in search engines by considering how PageRank works

- Link spam takes advantage of PageRank

Link Spam

A form of spamming trying to increasing the PageRank of member pages

Question

- How to find the important persons on Twitter by applying PageRank?

Summary

- Part I — Computing

- Part II — Computing at SCU

- Part III —Data Science

- Part IV —PageRank

- Computing fields are a lot of fun!