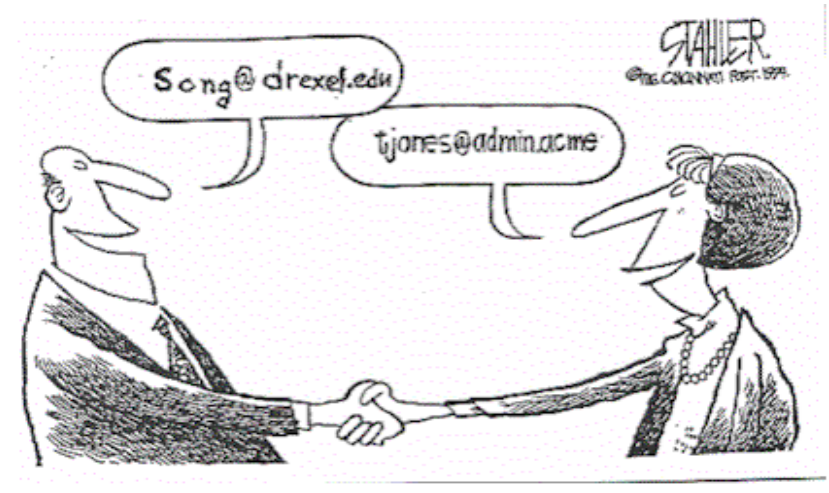
BIG DATA: TECHNOLOGIES AND APPLICATIONS

1-1 Course Overview

Il-Yeol Song, Ph.D.
College of Computing & Informatics
Drexel University
Philadelphia, PA 19104

DREXEL UNIVERSITY College of Computing & Informatics

Hello!



DREXEL UNIVERSITY College of Computing & Informatics

What are common here?

- **Google**
- **Facebook**
- **Amazon**
- **Yahoo**
- **Ebay**
- **Walmart**
- **Sears**
- **Expedia.com**
- **CNN**
- **SAP**
- **MSNBC.COM**
- **Netflix**
- **Oracle**
- **Teradata**
- **Chase Bank**

What are common here?

- Class list, transcripts, receipts from stores, phone bills, drivers' licenses, salary history, flight schedules, etc.
- Stock price and history
- A list of DNAs and their biological characteristics
- Credit card fraud detection
- What are 10 best risks?
- Which patients are responding to our therapy?
- Young men buy beer on Friday nights when they buy diapers
- In a retail chain, potato chip purchases were accompanied by a soda purchase in half the cases. That figure increases to 75% when there is a marketing promotion.
- Blue Cross found some providers had superior treatment success rates for some fatal diseases.
- Obama had an 80.9% chance of winning in 2012 election

Purposes of the Course

- A gentle introduction to big data **for non-technical students** with no programming knowledge
- Provide an overview on **big data technologies** and **their applications** to real-world projects
- Understand how **the data-driven paradigm** and big data technologies can be used to create innovative projects and **revolutionize our business environment as well as our society**.
- Helps you understand the current **state-of-the-practice of Big Data technologies** and use cases
- Provides hands-on experience on **Watson Analytics** and **Tableau**

DREXEL UNIVERSITY College of Computing & Informatics

Week 1: Introduction to Big Data

- **Monday (27 June)**
  - Course overview
  - Understanding Big Data
- **Tuesday (28 June)**
  - Evolution of database Technologies
  - Data models
  - Relational databases
- **Wednesday 29 June)**
  - SQL
- **Thursday (30 June)**
  - Data warehouses
  - OLAP. OLTP,
  - Business Intelligence,
  - Big Data Analytics
  - Data-Driven Paradigm
  - Big Data Use Cases
- **Friday (1 July)**
  - Data Science
  - Data Analytics Lifecycle

DREXEL UNIVERSITY College of Computing & Informatics

Week 2: Big Data Technologies

- **Monday (4 July)**
  - Landscape of Big Data Technologies
  - Big Data Processing Architecture
  - Hadoop and its ecosystems
- **Tuesday (5 July)**
  - Spark
- **Wednesday (6 July)**
  - Mid-term Exam and Review
  - NoSQL Databases
  - NewSQL Databases
- **Thursday (14 July)**
  - Visualization
  - Tableau
  - Presenting and communicating Analytics Project
- **Friday (15 July)**
  - Final Exam and Review
  - Term Project Presentation
  - Big Data Trends, Opportunities, and Challenges

DREXEL UNIVERSITY College of Computing & Informatics

Week 3: Big Data Analytics

- **Monday (11 July)**
  - Guide to big data analytics tools, trends and best practices:
  - Python programming Language
- **Tuesday (12 July)**
  - Predictive Analytics
- **Wednesday (13 July)**
  - Cloud-based Analytics
  - Watson Analytics
- **Thursday (7 July)**
  - Cloud Computing
  - In-Memory Databases
  - Data Virtualization
  - Big Data Warehousing
- **Friday (8 July)**
  - Internet of Things (IOT)
  - Smart Aging
  - Smart City

DREXEL UNIVERSITY College of Computing & Informatics

Grading Components & Dates

- Exam 1: 25%
- Exam 2: 25%
- Attendance: 10%
- Group project proposal: 10%
- Group project report or presentation: 30%

- Exam 1: Wednesday, July 6
- Exam 2: Friday, July 15
- Group Project proposal: Due Monday, July 4
- Group Project Report: Due Friday, July 15
- Group Project Presentation: Friday (July 15)

DREXEL UNIVERSITY College of Computing & Informatics

Instructor

- Professor, College of Computing & Informatics, Drexel University, Philadelphia, PA, USA
- PhD in CS, LSU, Baton Rouge, USA, 1988
- Named an **ACM Distinguished Scientist** in 2013
- Elected as an **ER Fellow** in recognition of contributions to conceptual modeling community
- Received **Peter Chen Award in Conceptual Modeling** in 2015
- Four teaching awards from Drexel (1991, 2000, 2001, 2011) including **Lindback Distinguished Teaching Award** (2001)
- Affiliated Professor of CS Dept, KAIST, Korea
- Co-Editor-in-Chief, Journal of Computing Science & Engineering

DREXEL UNIVERSITY College of Computing & Informatics

My Big Data Journey

- **Research Topics**
  - Conceptual Modeling
  - Data Warehousing
- **Deputy Director, NSF-Sponsored Research Center on Visual & Decision Informatics (CVDI), 2012-2014.**
  - Worked with 13 different companies
  - Managed 12 big data projects
- **Delivered keynote speeches on Big Data**
  - The First Asia-Pacific iSchool Conference in 2014
  - ACM SAC 2015 conference
  - ER2015 Conference
  - ALIEP 2016 Conference
- **Designed Data Science curriculum at Drexel**
  - PhD specialization in Data Science
  - BS in Data Science