Center of Excellence for Cyber Security, Data Analytics, and Cloud Computing at CSU

Sunnie Chung
Cleveland State University
Why Sudden Gaps in IT Workforce in Ohio in Data Analytics, Cyber Security and Cloud

- They are very new technologies to Computer Science in rise of Web Service on Internet (IoT)
- They were fast developed and fast evolving
- Research and Developments led by new Industries such as Google, Facebook, Yahoo.
- Academia was far behind in these areas
- The subjects are very complex and evolving unbelievably fast.
Center of
Data Analytics and
Cloud Computing
at CSU
Data Analytics includes

• Data Science
• Big Data Processing
• Data Mining
Data Science needs

- Computer Science
- Statistics
- Data Mining
- Big data Processing Systems
- Big Data Processing Skills and Techniques
Big Data and Cloud Computing

• Every CEO’s To Do List
• Defense Department
• NASA: Predict Natural Disaster like Earthquake
• NSA, Homeland Security: Predict and Prevent Terrorists’ Acts
• Internet start-ups
• Financial institutions
Data Analytics Examples

Any revenue driven critical business decisions

• Customer profiling
• Customer loyalty analysis

Cyber Security and Protection

• Fraud detection in Credit Card, ATM
• Intrusion Detection in Systems with sensitive data
Data Analytics Curriculum Should Teach

- Math, Statistics and Databases
- Big Data Specific Processing Techniques
- Cloud Computing
- Massively Parallel Big Data Processing Systems
- Data Source Modeling
- Data Mining Strategies
- Data Driven solutions
Required Courses for Data Analytics

- Databases
- Big Data Processing Systems:
  - Parallel Data Warehouse and OLAP
  - Map Reduce
  - Hadoop Based MPP Systems
- Statistics
- Data Mining
Core Systems for Big Data Processing

Massively Parallel Processing Systems

• Parallel Data Warehouse Based Systems:
  • Oracle, Tera Data, Microsoft PDW, IBM
  • In Memory NEW SQL Systems

• Hadoop/MapReduce Based Systems: No SQL systems
  • Mongo DB
  • Pig Latin
  • Hbase
  • Hive
  • And So many Others

• Cloud: Big Data Processing Systems on Cloud
  • Google Cloud, Amazon Cloud, Microsoft Azure, Oracle, IBM
The Big Picture of Big Data Processing Systems

Center of Data Analytics at CSU in Computer Science

CIS 530 Database Systems and Modern Database Processing
CIS 611 : Parallel Data Warehouse and OLAP:
CIS 612 : Big Data and Parallel Data Processing Systems
   Hadoop/ MapReduce
   No SQL Systems
   VM(Virtual Machine)
   Cloud Computing
CIS 660: Data Mining – Web/Text Mining
EEC 525 Machine Learning
CIS 695 Practicum in Data Analytics and Big Data Processing
IS-MCIS 696: Social Networks Analysis
CIS 667: Bioinformatics
Center of Data Analytics at CSU in College of Science

• Math and Statistics

Certificates on Applied Predictive Modeling:
MTH 521 : Time Series Analysis
MTH 531 : Categorical Data Analysis
MTH 537 : Operation Research
MTH 567 : Applied Linear Models I
MTH 638 : Operation Research II
MTH 668 : Applied Linear Models II
MTH 675 : Applied Multivariate Statistics
Center of Data Analytics at CSU in Business College

- Certificates on Business Analytic
  For SAS Certificate with SAS Enterprise Miner Tool
  BUS 575: Introduction to Business Analytics
  BUS 600: Applied Business Analytics
  BUS 601: Managing Databases for Business Analytics
  BUS 602: Strategy for Business Analytics
  BUS 603: SAS for Data and Statistical Analysis
  BUS 604: Advanced Business Analytics I
  BUS 606: Practicum in Business Analytics
Proposed Master Of Data Analytics at CSU

CIS 530: Database Systems and Modern Database Processing
CIS 611: Enterprise Database and Parallel Data Warehouse and OLAP
CIS 612: Big Data and Parallel Data Processing Systems
  Hadoop and MapReduce
  NoSQL Big Data Processing Systems
  Cloud and VM (Virtual Machine)
CIS 695: Practicum in Data Analytics and Big Data Processing
CIS 660: Data Mining and Text Mining
EEC 525 Machine Learning
IS-MCIS 696: Social Networks Analysis
BUS 603: SAS for Data and Statistical Analysis
BUS 604: Advanced Business Analytics I
MTH 531: Categorical Data Analysis for Predictive Model
CIS 667: Bioinformatics

Sunnie Chung Cleveland State University
Current Research/Publications at CSU

Sunnie Chung

- Research on Integrating Big Data Management Systems
- Research on Data Analytics for Intrusion Detection
- Research on Data Analytics for Machine Fault Detection
Collaboration with Local Industry in Data Analytics/Big Data at CSU

• **Explorys by IBM**
  - website:  [https://www.explorys.com/](https://www.explorys.com/)
  - Data Analytic/ Big Data Processing on Health and Wellness Data
  - Data Analytic for Cleveland Clinic, Metro Health

• **Progressive**
  - Big Data Processing on Auto Insurance: Hadoop Based MPP Systems

• **PNC**
  - Big Data Processing Systems on Financial Data

• **Parker**
Connection to Local Data Analytics/Big Data Group

- **Annual Data Analytics and Big Data Processing Workshop/Meetup at CSU**
  
  EECS Dept of CSU is hosting the Big Data meeting annually to connect our students to the local Big Data Industry
  
  Big Data Industry Presentations on Nov 16 at CSU

- **Data Scientist Group**
  
  Regular webinar on Advanced Data Analytic Topics
Data Analytics and Big Data Processing Workshop at CSU on Nov 16

- 226 people attended from Academia and industry.
- Presentations:
  - IBM Watson Research Team
  - Data Analytics Project by Explorys IBM on Health Data from Cleveland Clinic
  - Data Analytics Project by Progressive on SNAPSHOT: GPS records from Clients’s cars
  - Provided Direct Job/Internships opportunities to CSU Data Analytic Students in Engineering and Business College.
Connection with Local Industry for Coop, Internship, incubator

• **IT Talent Event by JabsOhio** for IT Workforce Development in NEOhio

• **Explorys by IBM**
  - Data Analytic/ Big Data Processing on Health and Wellness Data
  - Data Analytic for Cleveland Clinic, Metro Health

• **Progressive**
  - Big Data Processing on Auto Insurance: Hadoop Based MPP Systems

• **PNC**
  - Big Data

• **Parker**

• **More than 50 local company participation**
Pipelining with Center of Excellence at CSU

- Encourage the best high school students to pursue Data Analytics and Cyber Security Careers
- Encourage Undergraduates to enrol Graduate School
- Feed the Data Analytics and Cyber Security student intern programs
- Develop the Next Generation of Scientists and Researchers
- Promoting STEM at the high school level
  - A true “grass roots” movement

Sunnie Chung Cleveland State University
Comparison with Good Examples in Data Analytics Curriculums

Based on Top 10 Best Data Analytics Programs on www.masterindatascience.org

Good Balance of Core Subjects, Analytic Skills and Practicum

• North Western University
• Indiana University Bloomington
• Carnegie Mellon (NSF Chosen MidWest Region Big Data Program)
North Western University

MSIA 401 Statistical Methods for Data Mining
MSIA 431 Analytics for Big Data
MSIA 489 Industry Practicum
MSIA 490-21 Predictive Models for Credit Risk Management
MSIA 490-23 Healthcare Analytics
MSIA 490-25 Intro to Java Programming
MSIA 490-27 Social Networks Analysis
MSIA 490 Intro to Databases & Information Retrieval
MSIA 411 Data Visualization
MSIA 420 Predictive Analytics
MSIA 421 Data Mining
MSIA 430 Introduction to Data Warehousing and Workflow Management
MSIA 490-20 Text Analytics
MSIA 490-20 Topics in Analytics with Python
MSIA 440 Optimization and Heuristics

Sunnie Chung Cleveland State University
Indiana University Bloomington

• 2 years of Master of Data Science/Data Analytics or
• Hybrid: Master of Data Science and Computer Information Science
• Good balance of Courses on Core Subjects:
  Big Data Processing Application
  Advanced Database
  Advanced Algorithm
  Statistics
  Data Mining
  Security in Network System
  Information Visualization
  Cloud Computing
• Variety of good related Courses are available
Carnegie Mellon

- Business Intelligence & Data Analytics Curriculum:
- Prerequisite: OOP Programming Courses and 3 years Working Experience

95-703 Database Management 12
95-796 Statistics for IT Managers 6
95-710 Economic Analysis 6
95-797 Data Warehousing 6
94-806 Privacy in the Digital Age 6
95-868 Exploring and Visualizing Data 6
95-791 Data Mining 6
95-852 Analytics and Business Intelligence 6
95-866 Advanced Business Analytics 6
New Courses Needed for Master of Data Analytics at CSU

- Data Visualization
- Cloud Computing
Center of Cyber Security at CSU
Cyber Security Program Should Teach

- Building Secure Web Applications
- Detecting Software Vulnerabilities
- Database Security
- Secure Programming in the Cloud
- Network Security
- Secure Software Engineering
- Mitigating Software Vulnerabilities
Master of Cyber Security Includes

- Prevention and Protection Strategies in Cybersecurity
- Monitoring, Auditing, Intrusion Detection, Intrusion Prevention, and Penetration Testing
- Cyber Crime Investigation and Digital Forensics
- Human Aspects in Cybersecurity: Ethics, Legal Issues, and Psychology
Center of Cyber Security at CSU at Engineering College

- EEC 592 Hands-on Experience in Computer System Security
- EEC 688 Secure and Dependable Computing
- EEC 693 Network Security and Privacy I
- EEC 693 Network Security and Privacy II
- CIS 675, Information Security
- EEC 623 Software Quality Assurance
- EEC 693 Special topic: Secure Cloud Computing
Research on Cyber Security at CSU

• Ye Zhu:
  Active NSF grant on Password Security

• Ye Zhu, Chansu Yu, Haodong Wang:
  Active NSF grant on Security and Privacy in Mobile Device Encryption

• Sanchita Mal-Sarkar, Chansu Yu:
  Active NSF grant on Educational System for Software and Hardware security
Current Research Proposal for Cyber Security at CSU

- Chansu Yu, Haodong Wang
  - GPS Verification
- Sunnie Chung:
  - Research proposal on Data Analytics on Web Service logging data for Intrusion Detection/Prevention and Fraud Detection
  - Research Proposal on Cyber Database Security on Cloud by Querying Encrypted Database Without Decrypting
  - Research proposal on Building Querying System for Cyber Crime Investigation and Digital Forensics
New Courses Needed for Master of Cyber Security at CSU

- Prevention and Protection Strategies in Cybersecurity
- Monitoring, Auditing, Intrusion Detection, Intrusion Prevention, and Penetration Testing
- Cyber Crime Investigation and Digital Forensics
• 666: Data Protection-Privacy Challenges Posed by Internet & Cloud (includes cyber risk management)
• 622 Legal Responses to Terrorism
• 650 Int'l Aspects Intellect Prop
• 665 Computers & Law (predominantly criminal)
• 690 HIPAA and Privacy (health care info systems)
• 797 Info Technology & Law
• 882 Externship - U.S. Attorney
Connection to Community at CSU

- Cyber Security Consortium
- Dr. Julien Earls
- Dr. Chansu Yu
- Dr. Ye Zhu
What CSU Needs

• Faculty in these three areas:
  • Cyber Security
  • Data Analytics
  • Cloud Computing

• Support of the Existing Faculty for Research on:
  • Integrating Big Data Management Systems
  • Data Analytics for Intrusion Detection
  • **Cyber Database Security on Cloud**
  • Building Querying System for **Cyber Crime Investigation and Digital Forensics**

• Teaching and Research Labs for:
  • Data Analytics and Cloud Computing
  • Cyber Security and Cloud Computing