PENNSTATE

DS 330: Visual Analytics for Data Sciences

IST

Welcome and Outline

- Today's agenda
  - Course introduction and policies
  - Key technologies used in this class

Teaching Team

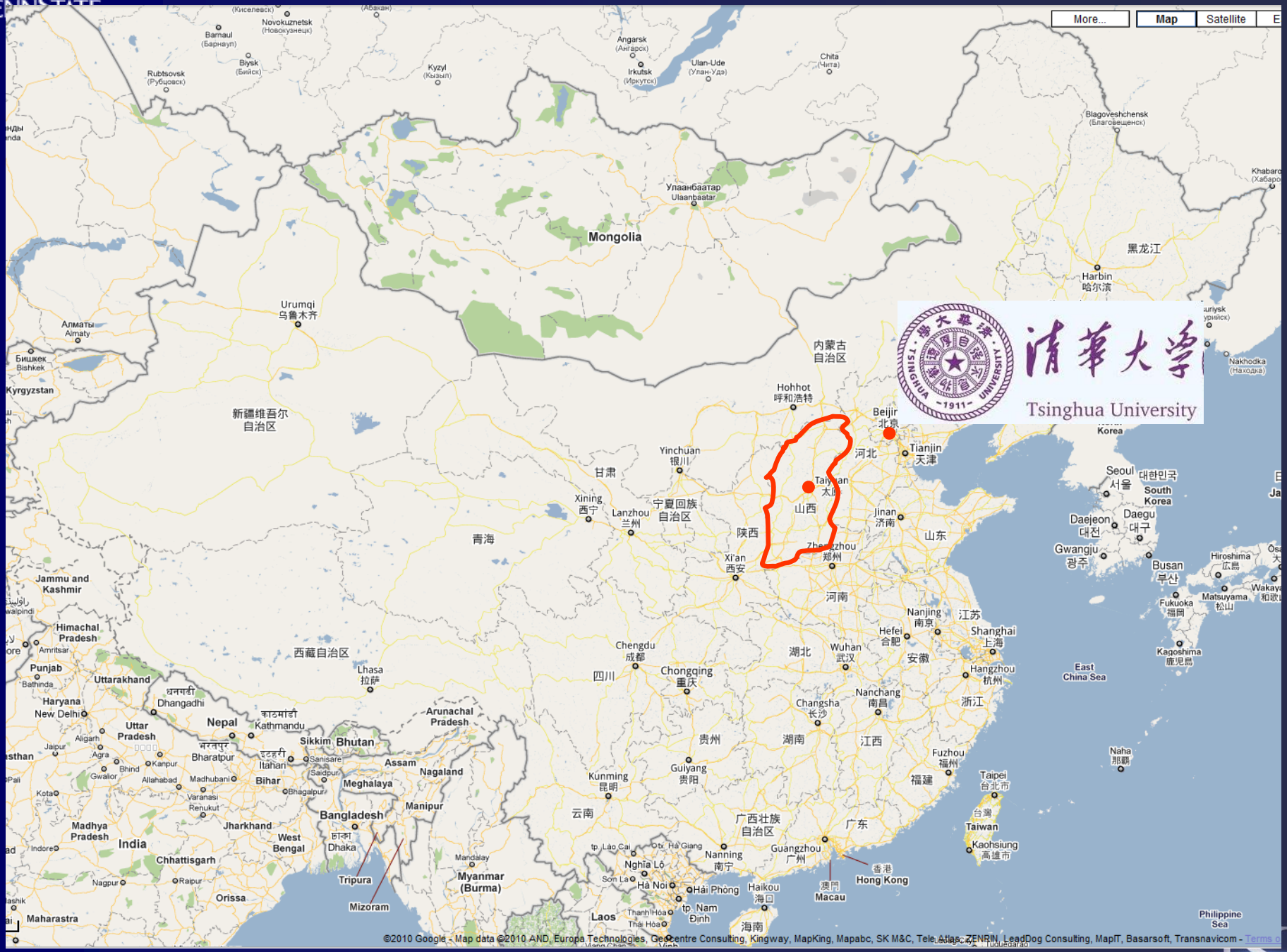
- Xiaolong "Luke" Zhang, instructor
- TA: Tianchun Wang
- LA: Mohamed Elmanzalawi

Self-Introduction

- Tianchun Wang

About Me

- Xiaolong "Luke" Zhang
  - Ph.D. in Information Science from University of Michigan
  - Research interest: human-computer interaction, information visualization and visual analytics

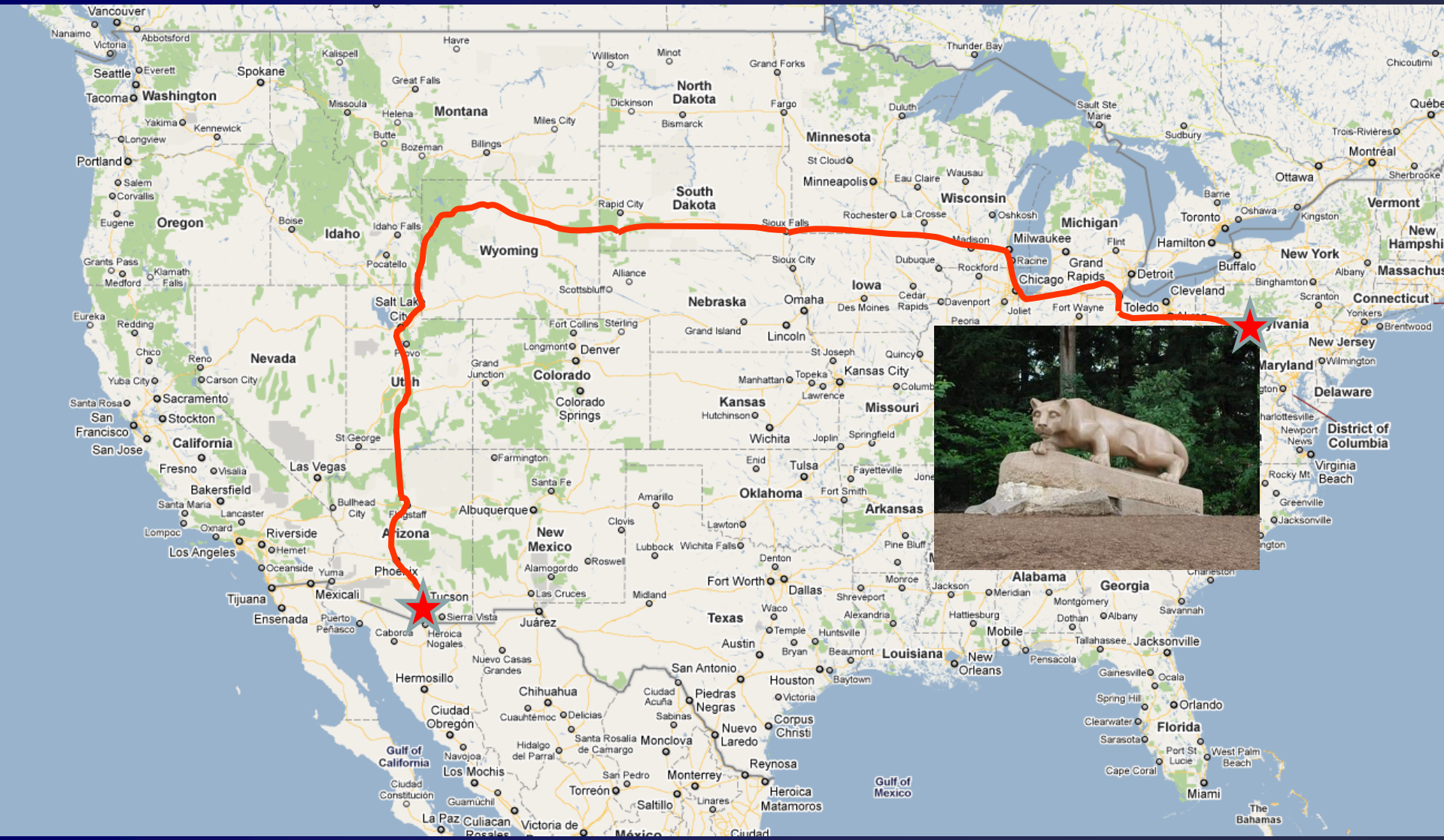

More...
Map
Satellite
Mongolia
Tsinghua University
清華大學
Urumqi
乌鲁木齐
Almaty
Алматы
Bishkek
Kyrgyzstan
新疆维吾尔
自治区
Hohhot
呼和浩特
内蒙古
自治区
Beijing
北京
Tianjin
天津
河北
Yinchuan
银川
甘肃
Xining
西宁
Lanzhou
兰州
宁夏回族
自治区
青海
陕西
山西
Taiyuan
太原
Zhengzhou
郑州
Xi'an
西安
山东
Jinan
济南
河南
Nanjing
南京
江苏
Hefei
合肥
Shanghai
上海
安徽
Wuhan
武汉
湖北
Chengdu
成都
四川
Chongqing
重庆
Changsha
长沙
Nanchang
南昌
浙江
Hangzhou
杭州
贵州
湖南
江西
Fuzhou
福州
Guiyang
贵阳
Kunming
昆明
云南
福建
Taipei
台北市
台湾
Taiwan
Kaohsiung
高雄市
广西壮族
自治区
广东
Guangzhou
广州
Nanning
南宁
香港
Hong Kong
Macau
澳門
Haikou
海口
海南
Hà Nội
Hải Phòng
Laos
Myanmar
(Burma)
Bangladesh
Dhaka
India
Nepal
Kathmandu
Bhutan
Sikkim
Assam
Nagaland
Meghalaya
Manipur
Arunachal
Pradesh
Tripura
Mizoram
Orissa
Chhattisgarh
Jharkhand
West
Bengal
Bihar
Madhya
Pradesh
Maharastra
Uttar
Pradesh
Uttarakhand
Haryana
New Delhi
Punjab
Himachal
Pradesh
Jammu and
Kashmir
西藏自治区
Lhasa
拉萨
Korea
Seoul
서울
대한민국
South
Korea
Daejeon
대전
Daegu
대구
Gwangju
광주
Busan
부산
Fukuoka
福岡
Hiroshima
広島
Matsuyama
松山
Kagoshima
鹿児島
East
China Sea
Philippine
Sea
Naha
那覇
©2010 Google - Map data ©2010 AND, Europa Technologies, Geocentre Consulting, Kingway, MapKing, Mapabc, SK M&C, Tele Atlas, ZENRIN, LeadDog Consulting, MapIT, Basarsoft, Transnavicom - Terms o

PENNSTATE

IST

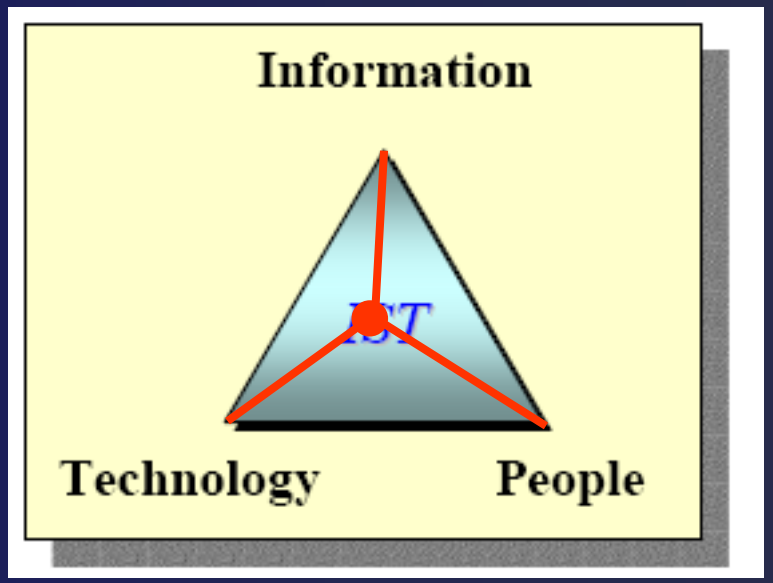
St George
Mesquite
Colorado City
Fredonia
Page
Kayenta
Kaibito
Shonto
160
Lukachukai
Many Farms
Tsaile
Grand Canyon National Park
Supai
Tonalea
Pinon
Tuba City
Chinle
Grand Canyon Village
Ganado
Fort Defiance
93
40
Williams
Flagstaff
191
Winslow
180
Holbrook
Cottonwood
Sedona
Big Park
Prescott
Camp Verde
Lake Havasu City
Parker Strip
Arizona
Snowflake
St Johns
Show Low
Pinetop-Lakeside
Eagar
Apache National Fores
Whiteriver
Quartzsite
Fortuna Foothills
Luis Río
Casas Adobes
Tucson
Safford
Duncan
Willcox
10
Three Points
Vail
Benson
Green Valley
Sonoita
Sells
19
Sierra Vista
Bisbee
Puerto Peñasco
Heroica
Agua Prieta

PENNSTATE

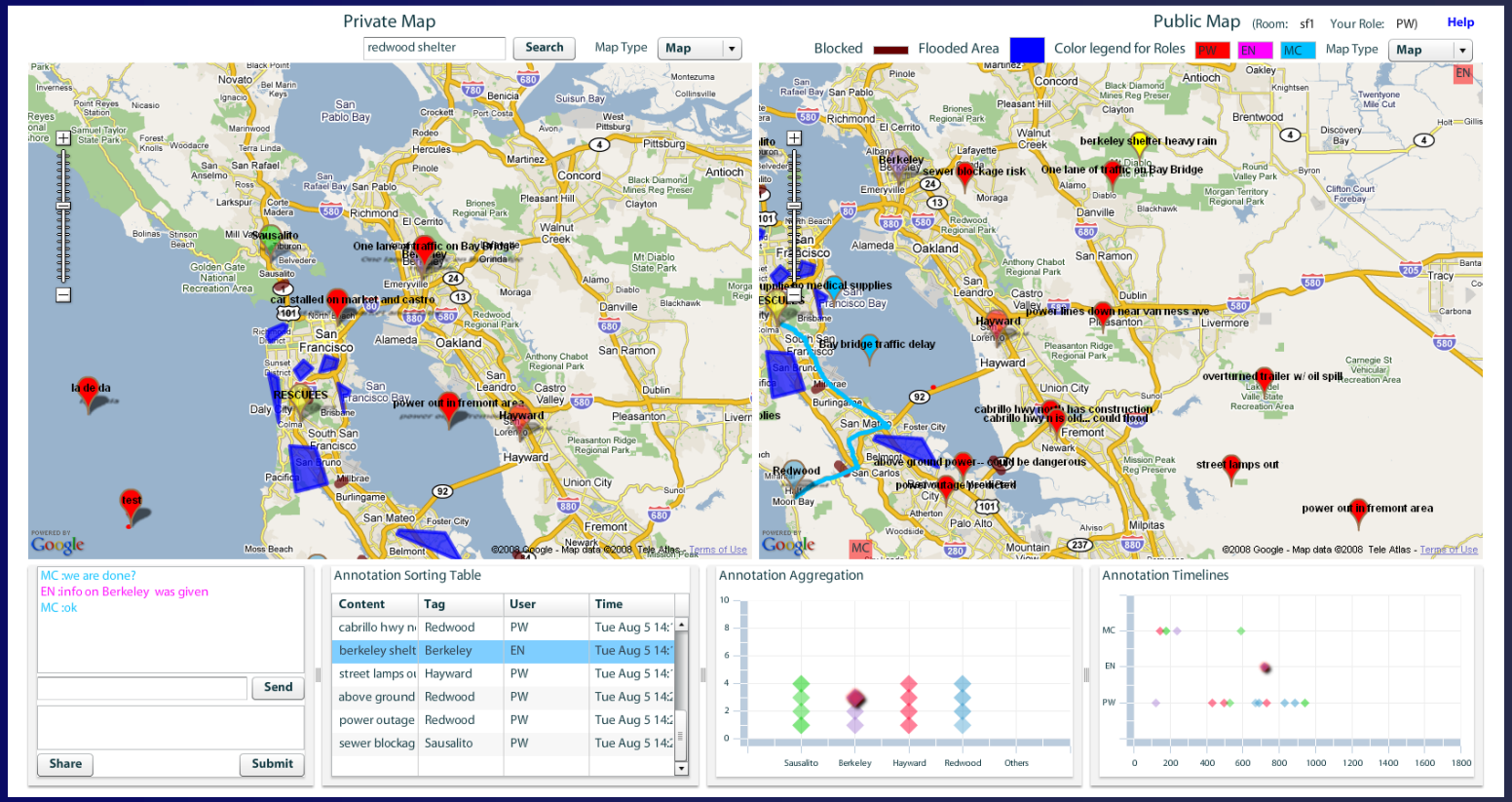


IST

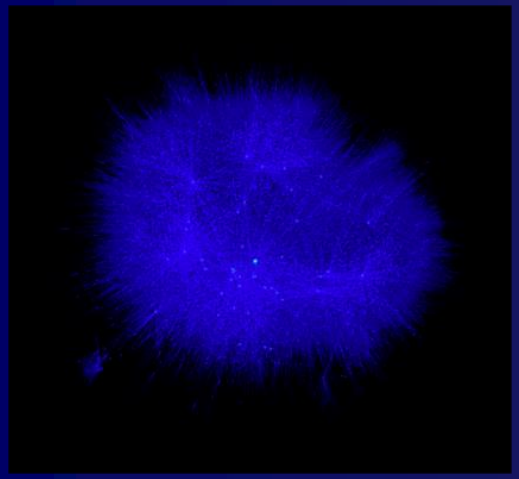
My Research in IST

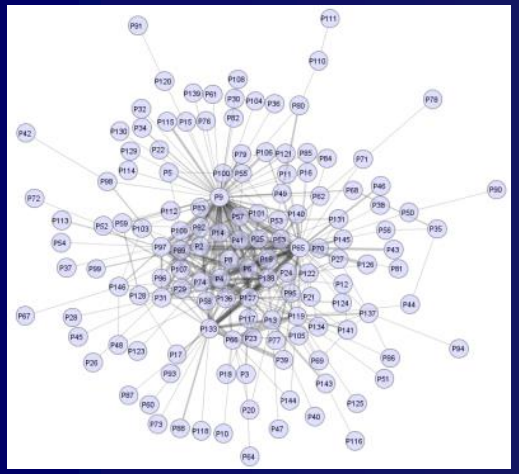


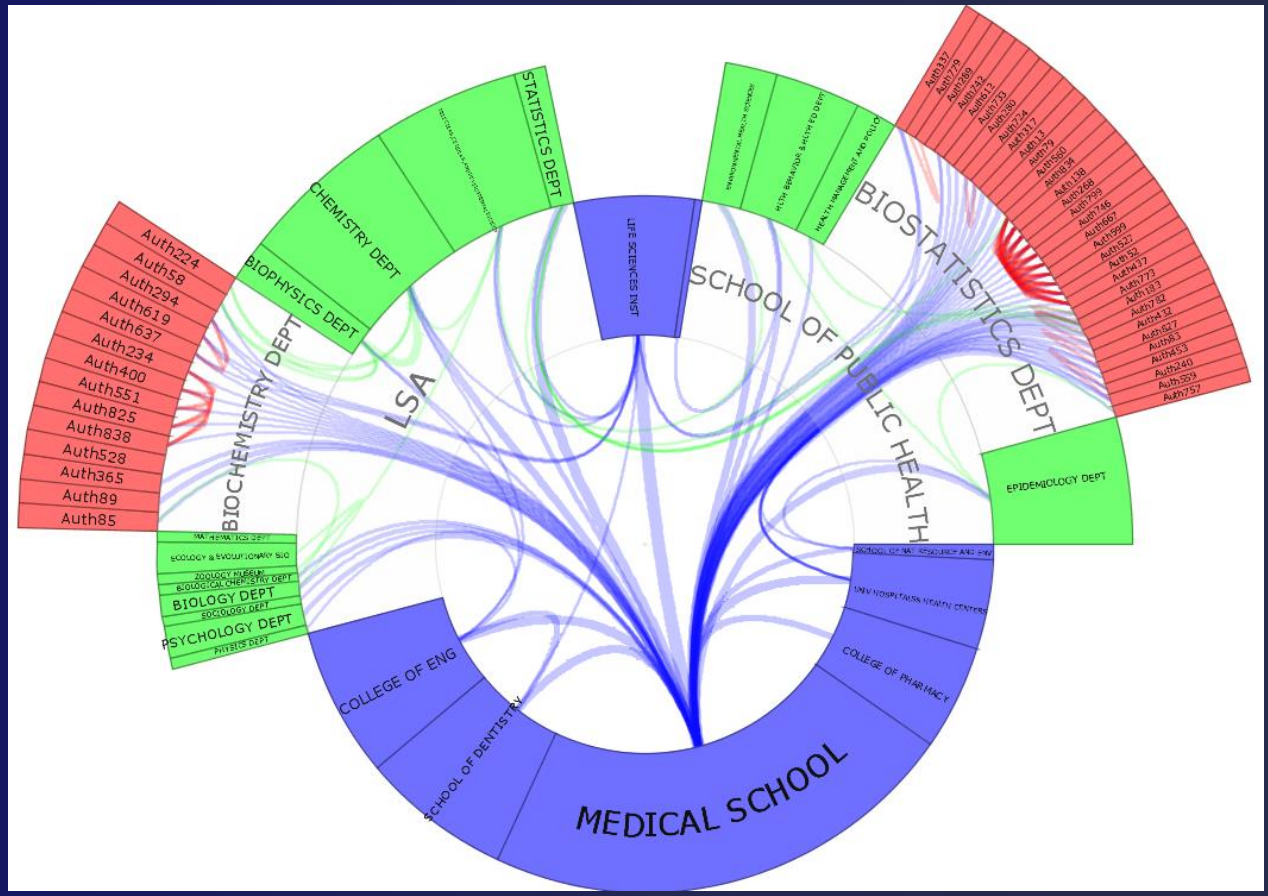
Geospatial Information Visualization in Decision Making



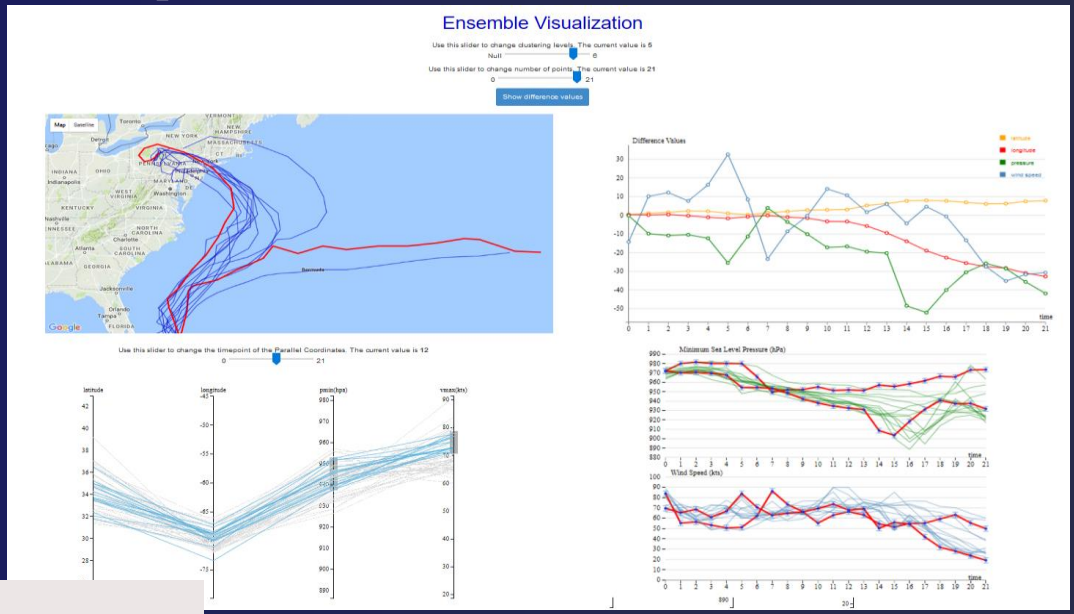
Social Network Visualization and Analysis

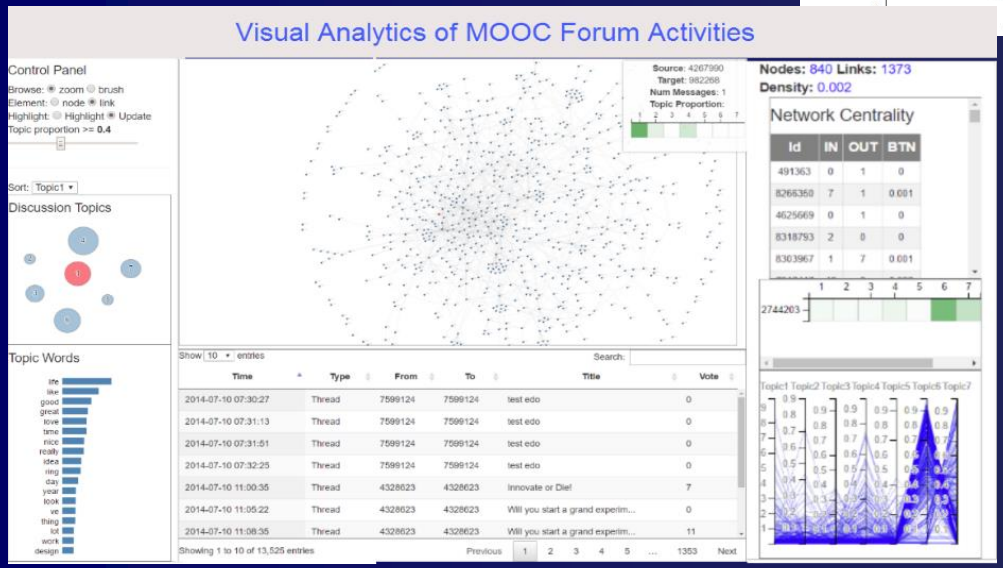






Visual Analytics Systems





My Teaching

- IST 210, 220, 489H
- DS 330
- SRA 468

- IST 501, 504, 505, 526

PENNSTATE

What Is This Course About?

IST

From
http://magazine.jhsph.edu

PENNSTATE

We Have Computers?!

IST

Data
Computer
Decision

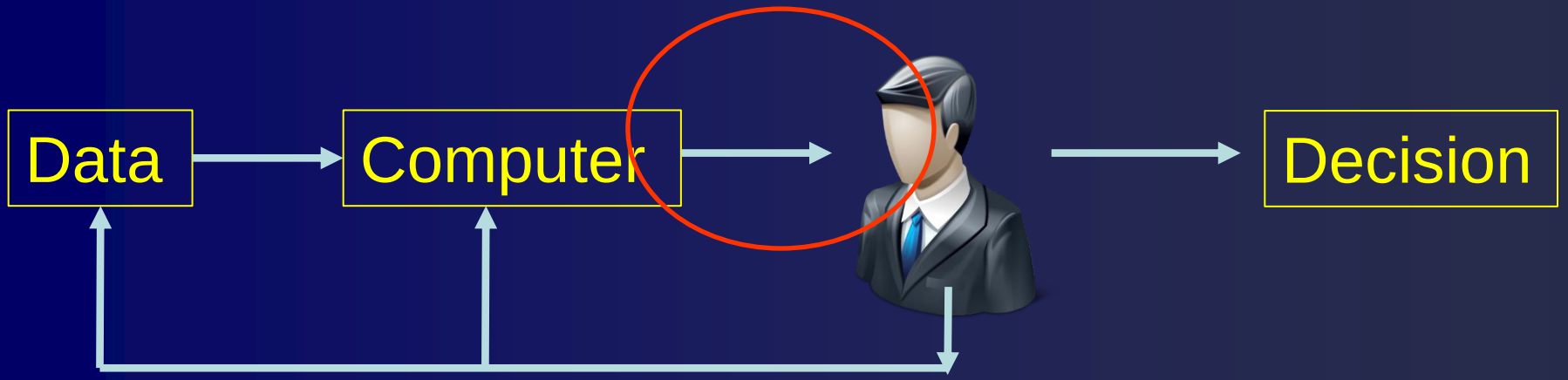
From
Person of Interest

A Scene from POI

- https://www.youtube.com/watch?v=RYAatPs7J_o

How Much Should We Trust Machines and Algorithms?

Human-in-the-Loop



Visualization becomes more and more important to data analysis.

NORAD: North American Aerospace Defense Command

PENNSTATE

This Course Helps You Know

- What visualization designs/tools/systems to use for specific data types and tasks.
- How to design/implement a visual analytics system.
- How to prepare data for visualization.

IST

PENNSTATE

Objectives

- Discuss the importance of information visualization and visual analytics;
- Understand the basic theories of visual perception that guide visualization design;
- Choose appropriate visualization designs based on data types and tasks; and
- Build visual analytics tools with current technologies, such as D3, JavaScript, Python, etc.

IST

Class Components

- Lectures
  - Theories, concepts, design guidelines, etc.
- Labs and exercises
  - Current technologies related visual analytics
- Assignments
  - Individual, team
- Team project
  - Implementation of visualization tools for data analysis
- Self-learning components
  - Development of your technical skills

Lectures

- Theories, concept, design guidelines, etc.
  - Visual perception
  - Visualization design
  - Visual analytics
- Technical issues
  - Web programming: JavaScript
  - Data processing: Python
  - Data query: SQL

Lab/Exercise

- Programming
  - Working on your assignments
- Learning about visual analytic tools

Assignments

- Individual assignments
  - Two essays
    - On the roles of visualization in visual analytics
  - Seven programming assignments
    - Visualization of common data types
- Team assignments
  - Project-oriented

Project

- Implementation of a set of visualization tools with specific analytical goals
  - Find a data set by your own
    - Approved by me
- Deliveries
  - Four progress reports
    - Description, data pre-processing, visualization design, implementation
  - One final report
  - Project presentation

Self-Learning

- Help to develop your own skills
  - HTML
  - CSS
  - JavaScript
  - JSON
  - Python
- Modules from linkedin.com/learning
  - Penn State subscribes its services.

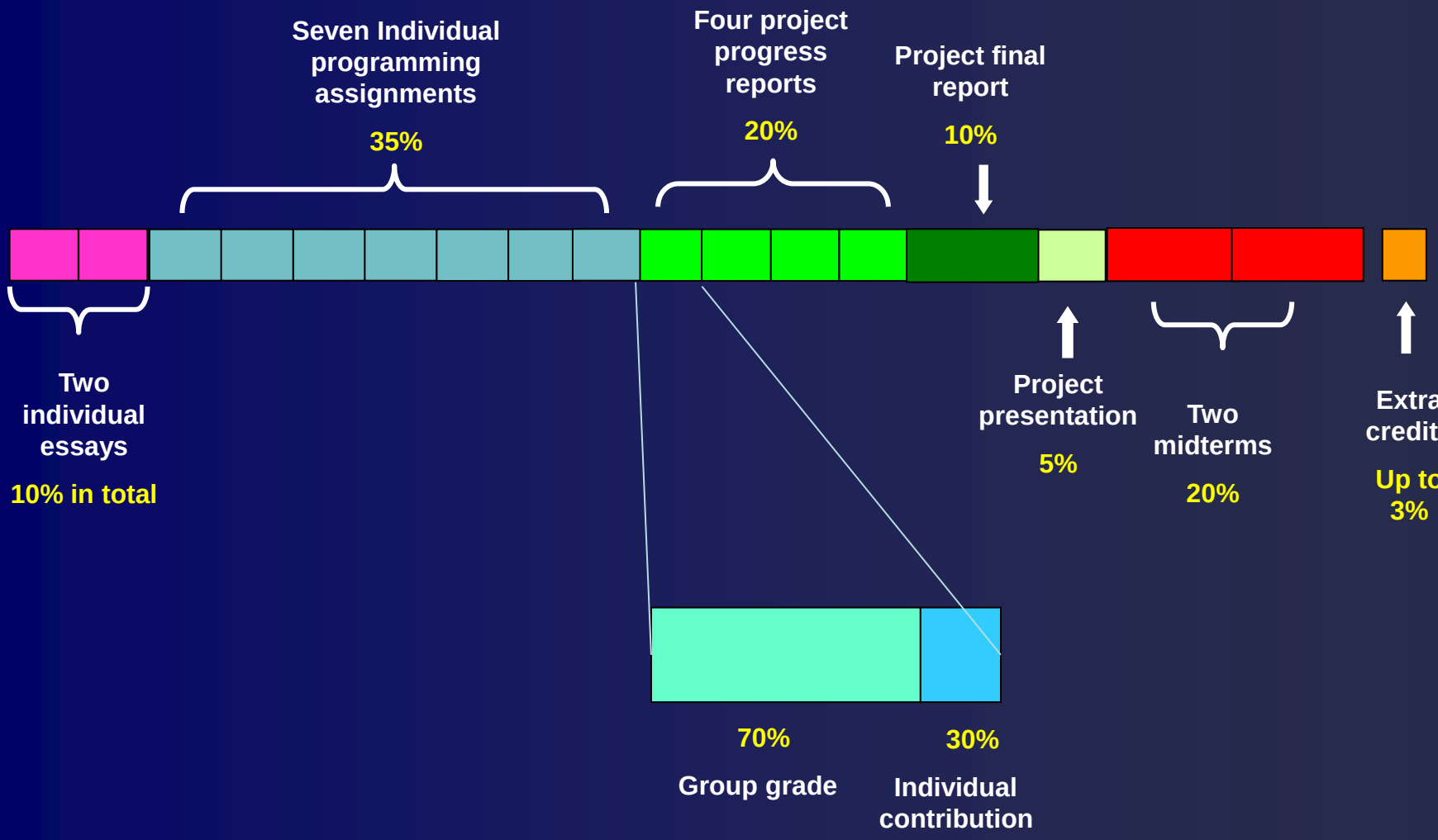
PENNSTATE

Course Structure

- Tuesday
  - Lectures
  - In-class exercise
- Thursday
  - Lab/exercise
  - Team project times

IST

Evaluation Components



Grading Policy

A	90 and above
A-	85-90
B+	80-85
B	75-80
B-	70-75
C+	65-70
C	60-65
D	50-60
F	below 50

- Not curved

PENNSTATE

Course Communication

IST

CANVAS

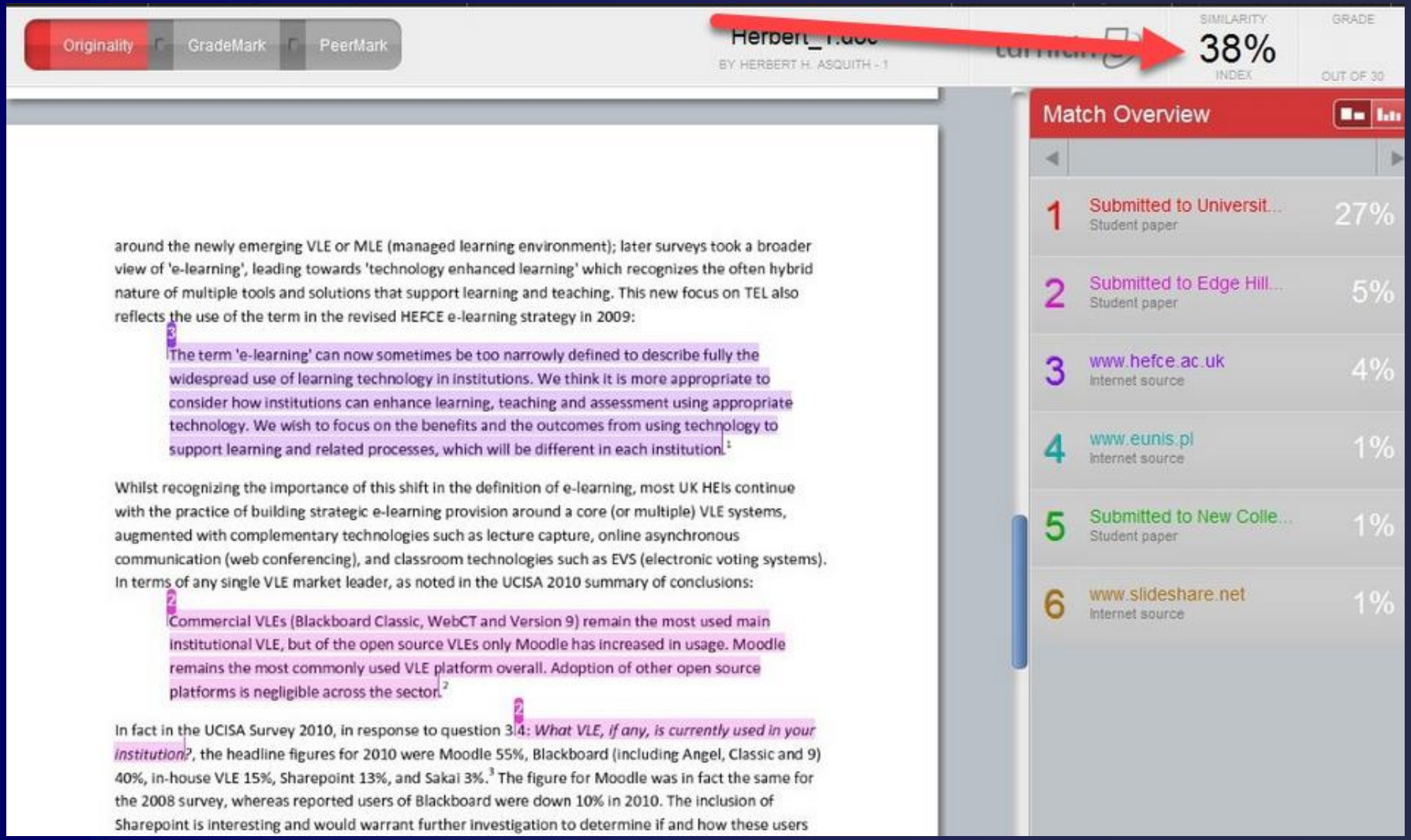
- Centralized repository of course materials
  - Syllabus, readings, slides, homework assignments, lab instructions, report instructions, etc.
- Assignment submission
  - All submission must be made online through CANVAS.

Office Hours

- Instructor
  - 1 – 2 pm on Monday
    - Or by appointment
  - E323 Westgate

Course Conduct

- All due dates are hard deadlines.
  - No extensions unless pre-approved by instructor.
- Attendance and participation
  - Involvement of class discussions
  - Let instructor know if you cannot come to the class in advance.
- Academic integrity
  - Group responsibility
  - Academic integrity
    - http://undergrad.psu.edu/aappm/G-9-academic-integrity.html
  - Writing assignment checked with Turnitin.


Originality
GradeMark
PeerMark
BY HERBERT H. ASQUITH - 1
SIMILARITY
38%
INDEX
GRADE
OUT OF 30
around the newly emerging VLE or MLE (managed learning environment); later surveys took a broader view of 'e-learning', leading towards 'technology enhanced learning' which recognizes the often hybrid nature of multiple tools and solutions that support learning and teaching. This new focus on TEL also reflects the use of the term in the revised HEFCE e-learning strategy in 2009:
The term 'e-learning' can now sometimes be too narrowly defined to describe fully the widespread use of learning technology in institutions. We think it is more appropriate to consider how institutions can enhance learning, teaching and assessment using appropriate technology. We wish to focus on the benefits and the outcomes from using technology to support learning and related processes, which will be different in each institution.[1]
Whilst recognizing the importance of this shift in the definition of e-learning, most UK HEIs continue with the practice of building strategic e-learning provision around a core (or multiple) VLE systems, augmented with complementary technologies such as lecture capture, online asynchronous communication (web conferencing), and classroom technologies such as EVS (electronic voting systems). In terms of any single VLE market leader, as noted in the UCISA 2010 summary of conclusions:
Commercial VLEs (Blackboard Classic, WebCT and Version 9) remain the most used main institutional VLE, but of the open source VLEs only Moodle has increased in usage. Moodle remains the most commonly used VLE platform overall. Adoption of other open source platforms is negligible across the sector.[2]
In fact in the UCISA Survey 2010, in response to question 3.4: *What VLE, if any, is currently used in your institution?*, the headline figures for 2010 were Moodle 55%, Blackboard (including Angel, Classic and 9) 40%, in-house VLE 15%, Sharepoint 13%, and Sakai 3%.[3] The figure for Moodle was in fact the same for the 2008 survey, whereas reported users of Blackboard were down 10% in 2010. The inclusion of Sharepoint is interesting and would warrant further investigation to determine if and how these users
Match Overview
1 Submitted to Universit... 27%
Student paper
2 Submitted to Edge Hill... 5%
Student paper
3 www.hefce.ac.uk 4%
Internet source
4 www.eunis.pl 1%
Internet source
5 Submitted to New Colle... 1%
Student paper
6 www.slideshare.net 1%
Internet source

Communication

- Check CANVAS frequently for
  - Latest update
  - Urgent announcement
- If having concerns, come to see me first.
  - University Policy

Some Tips

- Do readings and think through.
  - Visual analytics is not just about what tools to use.
    - Not all concepts can be taught through “entertaining” activities.
- Be active to link topics to real-world stories.
  - Keep track of up-to-date IT news and events
- Apply various learning strategies
  - Self-learning modules, learning from peers, etc.
- Take initiatives to sharpen your technical skills
- Communicate with TA, LA and instructor.

Communicate with Teaching Team

- Please let us know if you have any question.
  - Email, office hours
  - Using Canvas to send email.
- Come to see us in office hours.
  - You don’t have to have a specific question.

PENNSTATE

Questions?

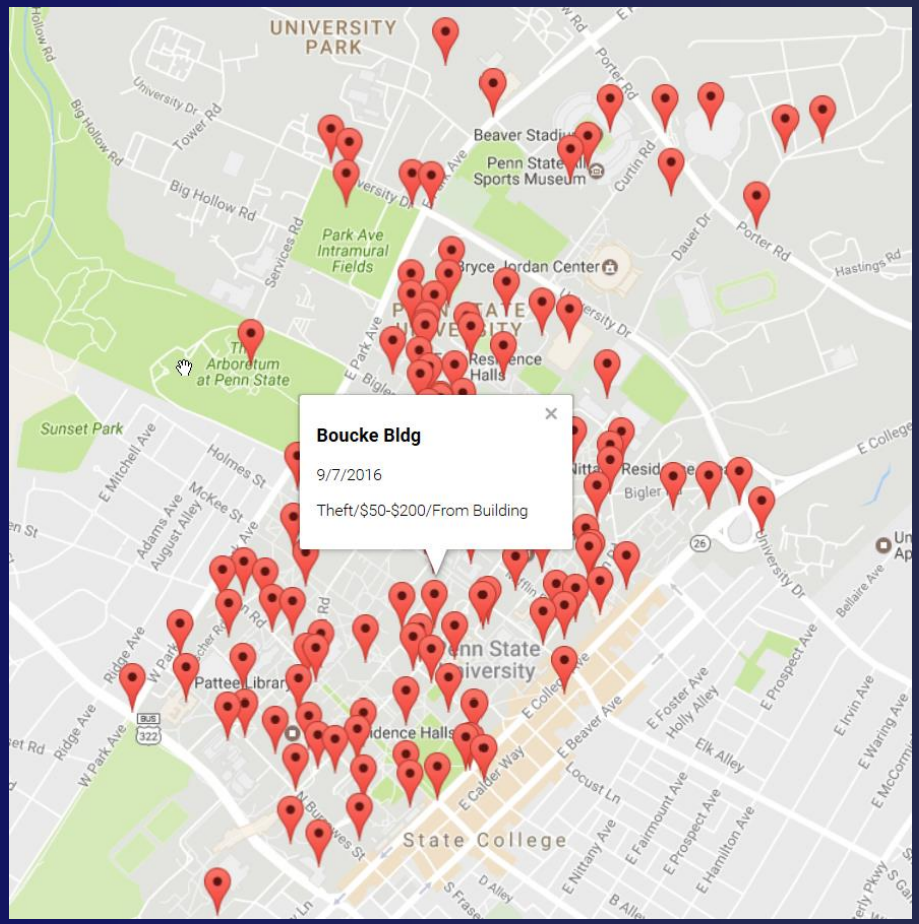
IST

Individual Essay Assignment 1

- Solve a puzzle
- Pay attention to the process and tools you used
- Write an essay about how you solve the puzzle
  - See instructions for what should be included.
- Due: Next Thursday before the class.

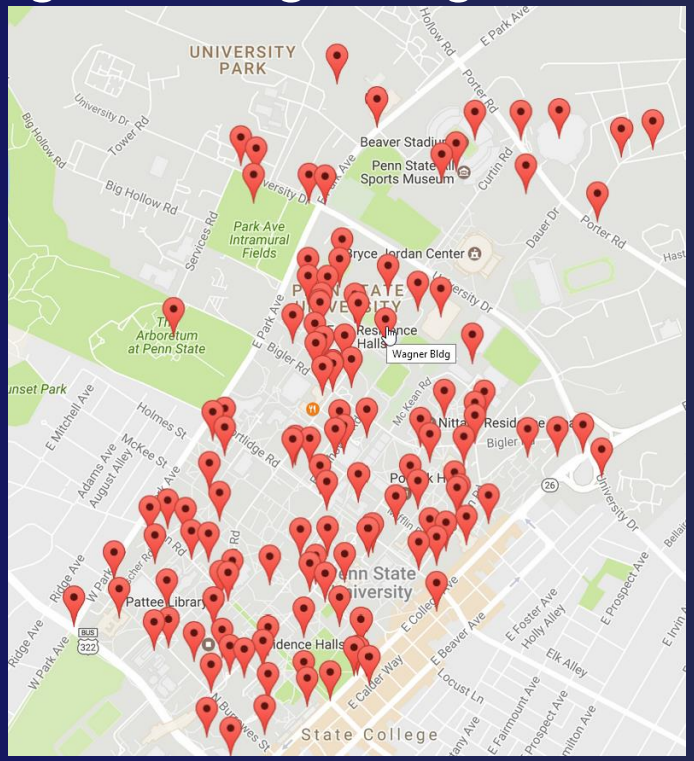
First Programming Assignment

- Penn State Crimes on Google Maps



What We Will Do on Thursday

- Lab/exercise: Google Maps with JavaScript
  - The 1st programming assignment



PENNSTATE

A Task Before Class on Thursday

- Read the website about Google Maps API key/authentication
  - https://developers.google.com/maps/documentation/javascript/get-api-key
  - Get a standard API user key!
  - Know how to use the key in JavaScript
  - Must have the API key ready before coming to the class on Wednesday

IST

What You Need to Do More After Today's Class

- Check CANVAS for the Readings and Self-Learning resources links
  - Readings and videos for Week 2
  - Start self-learning modules: HTML, CSS, JavaScript
- Work on individual assignment 1
- Think about whom you want to team with
  - Balanced skills: programming, web design, writing, etc.
  - Send instructor a message if you want to be a group with someone.
    - 4 members each group.