

M259 Visualizing Information

Winter 2022

Schedule, Tues-Thur 1:00-2:50pm

George Legrady, Instructor

Weihao Qiu, Yuxian Li, Teaching Assistants

George Legrady © 2022

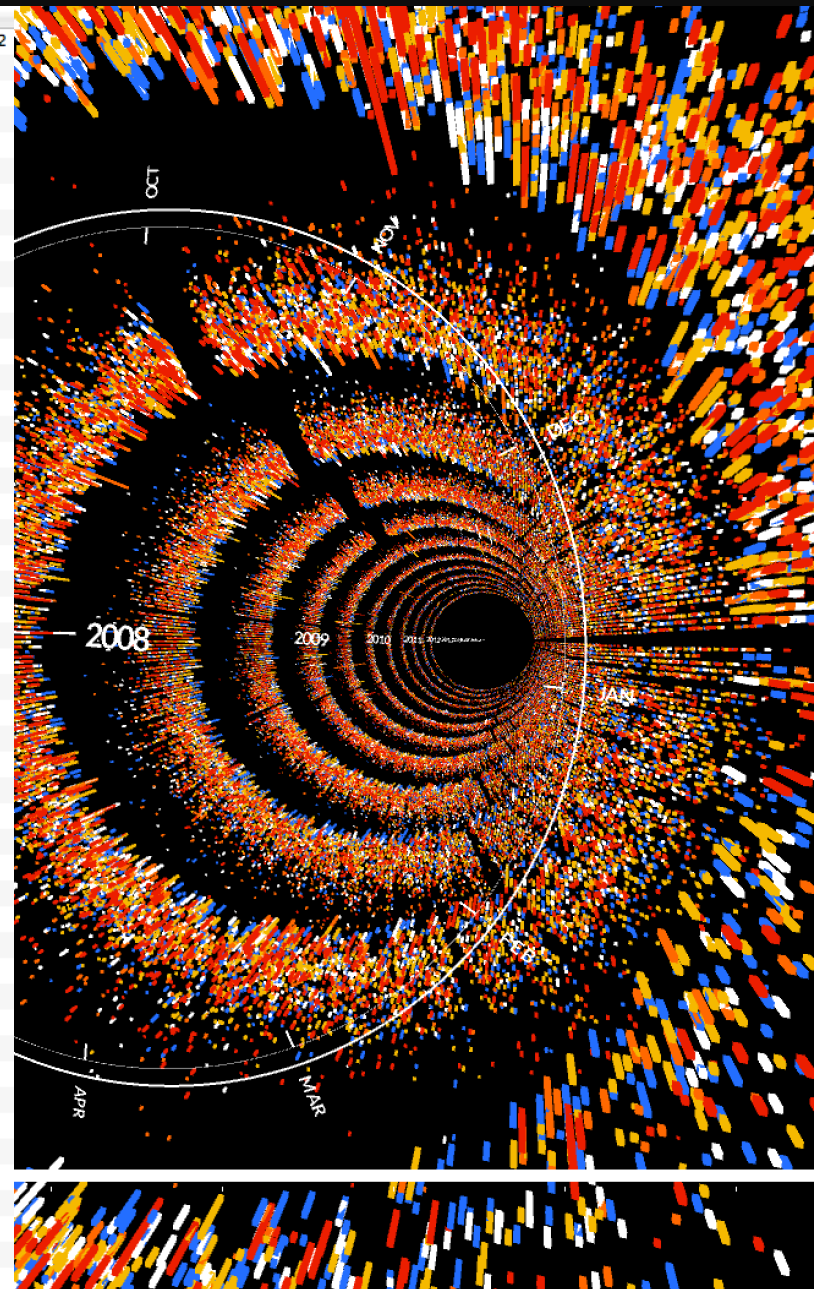
Experimental Visualization Lab

Media Arts & Technology

University of California, Santa Barbara

Course Focus: From Data to Visualization

collcode	itemtype	barcode	title	callNumber	deweyClass
nacd	accd	0010053294418	tigers have spoken	CD 782.421642 C2665T	782.421642
cafic	acbk	0010046075908	Heart on the line	FIC ARNOLD	NULL
nalpfc	acbk	0010047150288	Too hot to handle	FIC LOWELL	NULL
canf	acbk	0010027180271	Governing public schools new times new requirements	371.2 DANZBER 1992	371.2
cs9	acbk	0010002326972	great deception the inside story of how the Kremlin took over Cuba	972.91 M741G	972.91
canf	acbk	0010041599258	Logic or The right use of reason in the inquiry after truth with a va...	160 L793W 1996	160
cafic	acbk	0010050230746	Something about Emmaline	FIC BOYLE2005	NULL
canf	acbk	0010048238389	Understanding installation art from Duchamp to Holzer	709.04 R7277U 2003	709.04
cs9	acbk	0000103502175	Frank B Kellogg a biography	B K292 B	NULL
cs9	acbk	0010002324977	German raider Atlantis	940.953 F8517G	940.953
cafic	acbk	0010033394031	Zabelle	FIC KRICORI1998	NULL
canf	acbk	0000102981248	Reservation narrow gauge Omak Creek railroad Bow Arrow Short Li...	385.52097 L587R	385.52097
canf	acbk	0010004731641	place no one knew Glen Canyon on the Colorado	779 P833P2a	779
canf	acbk	0010002691342	Buddhism in translations passages selected from the Buddhist sacr...	294 W252B	294
canf	acbk	0010048100100	death in Washington Walter G Krivitsky and the Stalin terror	327.1247 K4592D 2003	327.1247
canf	acbk	0010054269096	101 diseases you dont want to get	614.5 P8715o 2005	614.5
nafic	acbk	0010040162363	Twilight in Texas	FIC THOMAS	NULL
nanf	acbk	0010041237933	Boundless healing meditation exercises to enlighten the mind and...	294.34435 T387B 2000	294.34435
canf	acbk	0010050859288	fragrance of faith the enlightened heart of Islam	297 R1294F 2004	297
nanf	acbk	0010053789672	How to reduce workplace conflict and stress how leaders and their...	658.1053 M327H 2005	658.1053
canf	acbk	0010034758416	home of the blizzard a true story of Antarctic survival	919.8904 MAWSON 1998	919.8904
nab	acbk	0010045761730	Lives of mothers daughters growing up with Alice Munro	B M9265M 2001	NULL
canf	acbk	0010054107684	Foxes in the henhouse how the Republicans stole the South and th...	324.70973 R29996J 2006	324.70973
canf	acbk	0010045523353	False intimacy understanding the struggle of sexual addiction	241.66 Sch196F 1997	241.66
canf	acbk	0010025511717	Lakota recollections of the Custer fight new sources of Indian milit...	973.82 LAKOTA 1991	973.82
cafic	acbk	0010020070768	Furors die a novel	FIC HOFFMAN	NULL
nanf	acbk	0010042818186	devils dictionary of business monkey business high finance and lo...	330.0207 V895D 2005	330.0207
canf	acbk	0010041330571	Running with the Buffaloes a season inside with Mark Wetmore Ad...	796.42809 C7193L 2000	796.42809
canew	acbk	0010054076608	opened grave Sherlock Holmes investigates his ultimate case	FIC JAMES2006	NULL
canf	acbk	0010028375219	Transforming vision writers on art	810.80357 TRANSFO 1994	810.80357
cs9	acbk	0000102286903	De Shazer the Doolittle raider who turned missionary a true and th...	B D459W	NULL
nchol	jcbk	0010045871109	Jothams journey a storybook for Advent	J YTREEID	NULL
cs9o	acbk	0010001512887	Burri	B B942B	NULL
ncnew	jcbk	0010050556819	Lady in the water a bedtime story	E SHYAMAL	NULL
cs6ro	arbk	0010019836690	Bullard Arms	338.76834 J241B	338.76834
canf	acbk	0010038257472	Perfect bones a six point plan to promote healthy bones	616.716 L5785P 2000	616.716
nanf	acbk	0010050592780	Satisfaction the science of finding true fulfillment	155.9 B4581S 2005	155.9
nafic	acbk	0010022410574	Fly fishing tales literary bait by angling authors	FIC FLY FIS1994	NULL
nanf	acbk	0010037361580	Chariots of the gods unsolved mysteries of the past	001.94 DANIKEN 1999	001.94
canf	acbk	0010036046778	way of agape	241.4 MISSLER 1999	241.4
canf	acbk	0010026014141	Future of medicine toward a science of prevention based on ancle...	613 DUGLISS 1993	613
canf	acbk	0010045701462	Tanker operations a handbook for the person in charge PIC	623.88245 H8627T 2001	623.88245
caesi	bcbk	0010034398809	Angliiskii iazyk prosto o slozhnom prakticheskii kurs	RUSSIAN 428.24917 L576A	428.24917
canf	acbk	0010051680022	Gorgeous disaster the tragic story of Debra LaFave	364.153 L131L 2006	364.153
nacd	accd	0010040958539	Spirituals in concert	CD784.73 B322S	784.73
capf	acbk	0010025431536	Kathy and Mo show parallel lives	812.54 GAFFNEY 1992	812.54
nafic	acbk	0010046761242	revelation	FIC LITTLE	NULL
cacd	accd	0010046200415	peaceful Christmas	CD 782.21723 P3133	782.21723
canf	acbk	0010048571573	Tauntons family home idea book	728.37 St546T 2003	728.37



Data Science

Reviewed by [Jake Frankenfield and Caroline Banton](#) Updated Aug 27, 2019

What Is Data Science?

Data science provides meaningful information based on large amounts of complex data or [big data](#). Data science, or data-driven science, combines different fields of work in statistics and computation to interpret data for decision-making purposes.

Understanding Data Science

Data is drawn from different sectors, channels, and platforms including cell phones, social media, e-commerce sites, healthcare surveys, and Internet searches. The increase in the amount of data available opened the door to a new field of study based on big data—the massive data sets that contribute to the creation of better operational tools in all [sectors](#).

The continually increasing access to data is possible due to advancements in technology and collection techniques. Individuals buying patterns and behavior can be monitored and predictions made based on the information gathered.

However, the ever-increasing data is unstructured and requires parsing for effective decision making. This process is complex and time-consuming for companies—hence, the emergence of data science.

Data science, or data-driven science, uses big data and machine learning to interpret data for decision-making purposes.

Data As Resource

- **Data capture**, much of it by autonomous recording systems, has become a **fundamental resource** in
- Science, environment, health/medical industry, government, also surveillance
- **prediction** and to assist in **decision-making** and **guiding policy**

Data Organization & Delivery Systems

- The *production, capture* and *exponential growth* of cloud-based storage of data have also resulted in extensive **research** and **analysis efforts** requiring methods by which
- **data is organized, analysed, managed, and**
- addresses
 - **intellectual property rights,**
 - **distribution infrastructure,**
 - **security,**
 - **longevity,**
 - **storage and compatibility**

Cluster of 30,000 patents between 2005-2014

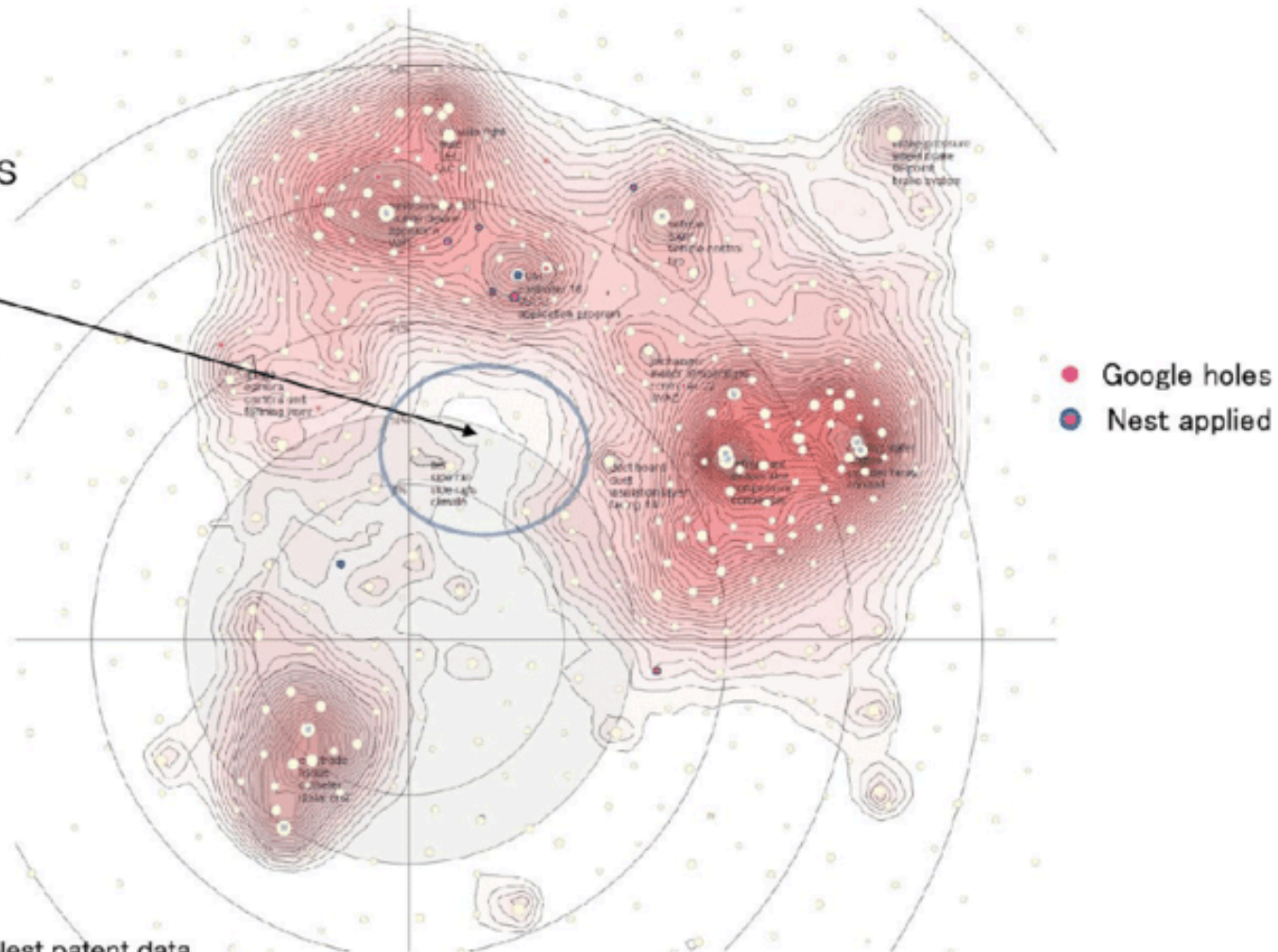
VALUENEX
intellectual innovator

Radar shows the past, present and future

METHOD AND APPARATUS
FOR RECHARGING
ELECTRIC VEHICLES...
Patent# 20110145141

Safety plug suitable for the
connection to a current
outlet... #20070139842

Recharge station, safety
plug...



@2016 VALUENEX, Inc.

intellectual innovator

Making Sense of the Data

Big Data Requires Methods by which to Make Sense of the Data

- a) How to quickly sift through data
- b) How to organize the data
- c) How to filter the data

Visual Language is a practice-based knowledge domain

This course provides practical and conceptual skillsets

The intent is to allow for experimentation in-between the full spectrum of *functionality* and *play of the imagination*

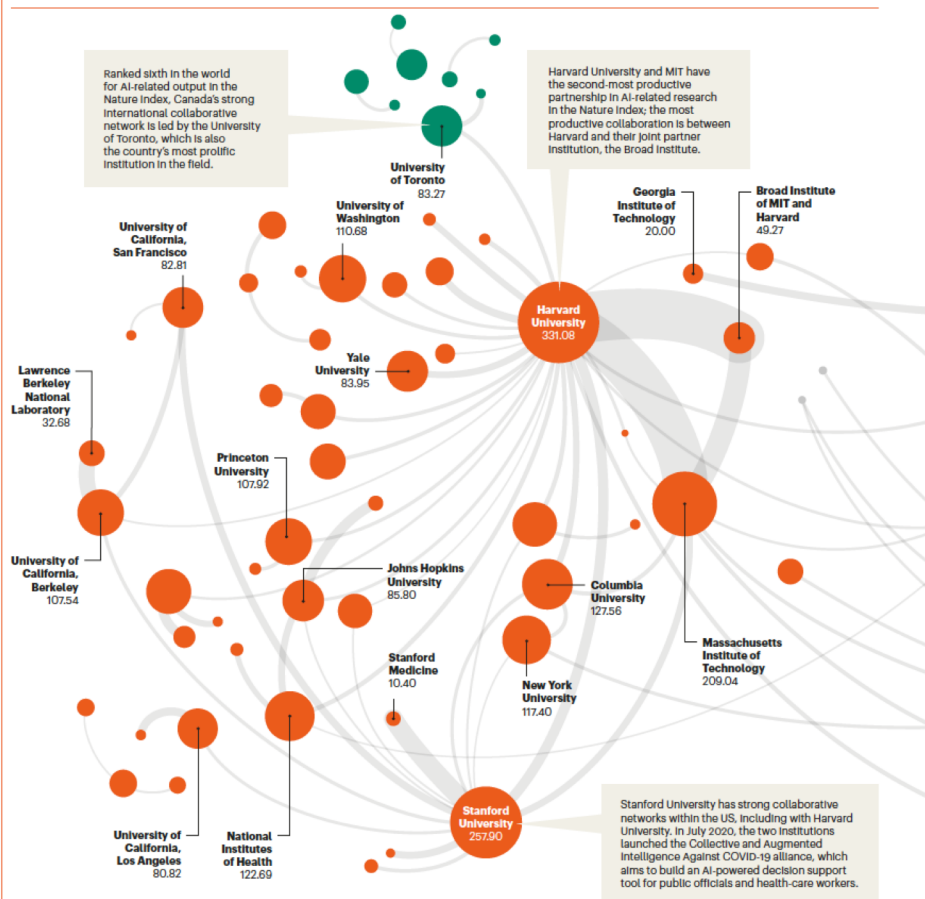
Collaborations on AI related papers in the Natural Sciences

Artificial intelligence

index

Network effect

Collaborations on AI-related papers in the natural-sciences journals tracked by the Nature Index reveal country strengths. Data analysis by Bo Wu. Infographic by Bec Crew and Tanner Maxwell.



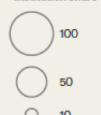
S112 | Nature | Vol 588 | 10 December 2020

© 2020 Springer Nature Limited. All rights reserved.

AI COLLABORATIONS

This graph comprises the top 200 collaborations among 146 institutions based on index data between 2015 and 2019, represented by lines and circles, respectively. Circles are sized according to each institution's share in artificial intelligence, and the thickness of the lines corresponds to a particular collaboration's total share between two institutions.

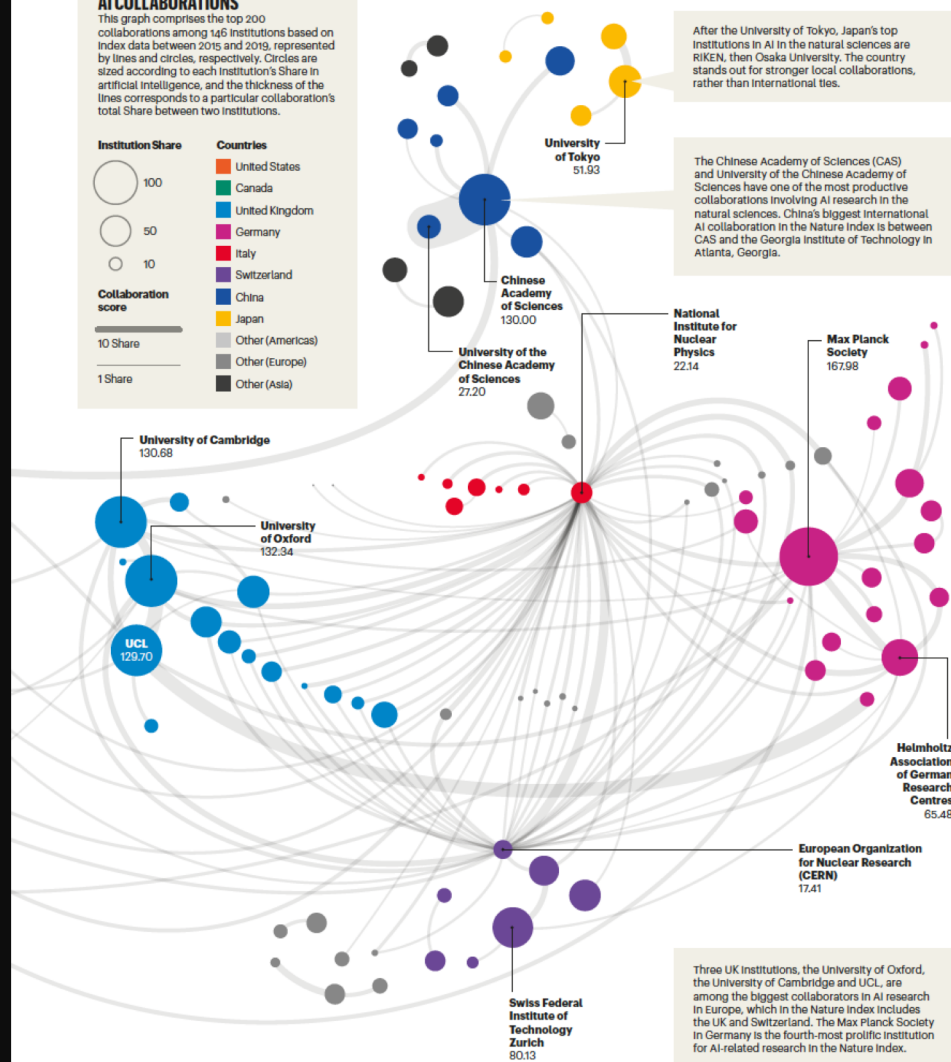
Institution Share



Collaboration score



Countries



Three UK institutions, the University of Oxford, the University of Cambridge and UCL, are among the biggest collaborators in AI research in Europe, which in the Nature Index includes the UK and Switzerland. The Max Planck Society in Germany is the fourth-most prolific institution for AI-related research in the Nature Index.

SOURCE: NATURE INDEX

Nature | Vol 588 | 10 December 2020 | S113

© 2020 Springer Nature Limited. All rights reserved.

The course is informed by

- **An overview of the techniques of Visual Language**
- **But also an appreciation of alternative explorations of *using data*, or alternative *applications of data***

- The course also embraces concepts from
- **Culture Analytics:** What does it mean to represent culture by data, and
- **Visual Analytics:** What are the methods by which visualizing can enable human-information discourse

Data Analytics

Reviewed by [Jake Frankfield](#) Updated Apr 27, 2019

What Is Data Analytics?

Data analytics is the science of analyzing raw data in order to make conclusions about that information. Many of the techniques and processes of data analytics have been automated into mechanical processes and [algorithms](#) that work over raw data for human consumption.

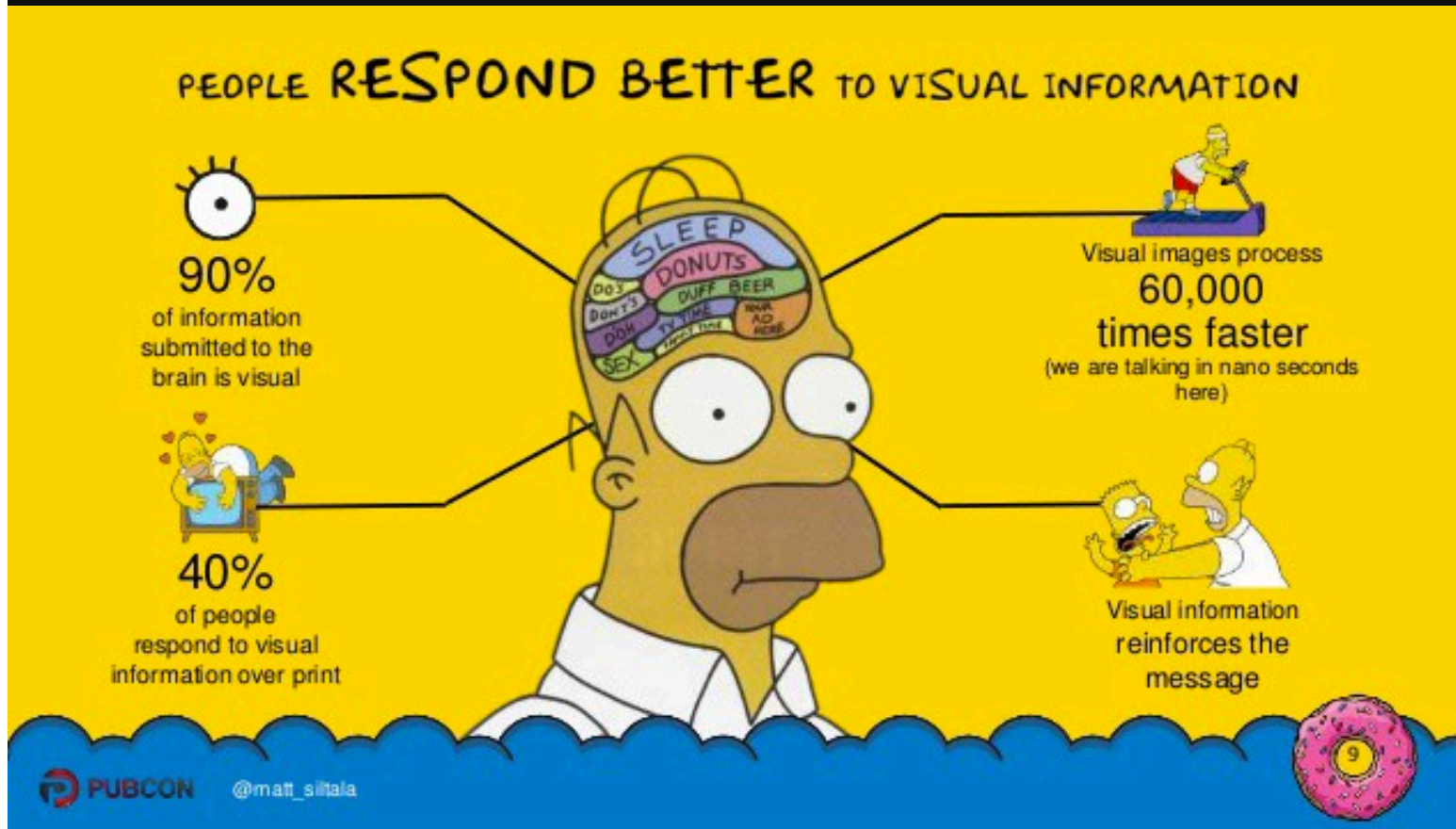
Data analytics techniques can reveal trends and metrics that would otherwise be lost in the mass of information. This information can then be used to optimize processes to increase the overall efficiency of a business or system.

Understanding Data Analytics

Data analytics is a broad term that encompasses many diverse types of data analysis. Any type of information can be subjected to data analytics techniques to get insight that can be used to improve things.

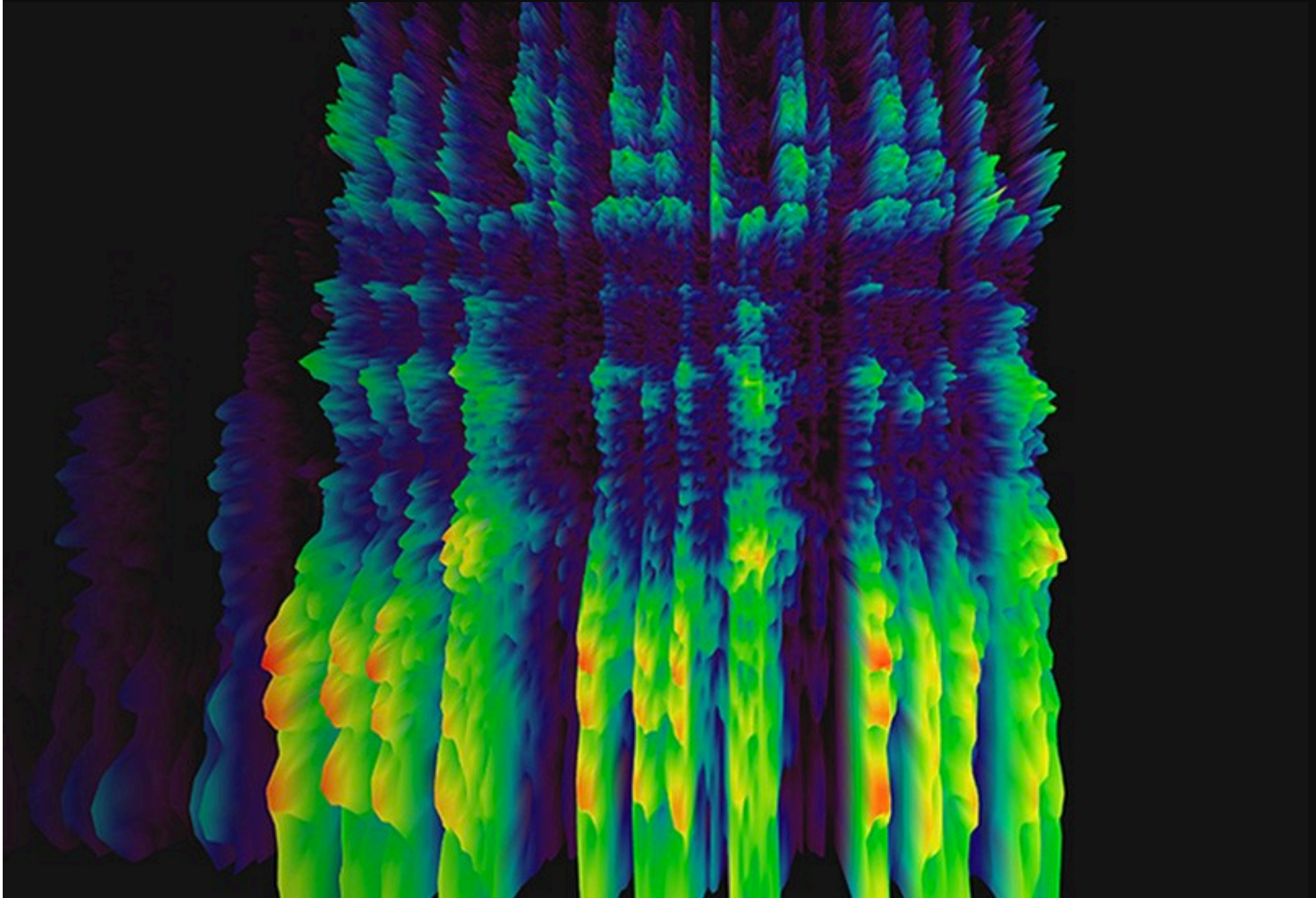
For example, [manufacturing](#) companies often record the runtime, downtime, and work queue for various machines and then analyze the data to better plan the workloads so the machines operate closer to peak capacity.

Data Visualization Helps to Make Sense of Information



on right

Mapping Data: A spectrogram of the sound of a human voice



Course Format

Tues: Lectures, visualization analysis

Thur: Technical lab

Course Resources

WEBSITE:

<https://www.mat.ucsb.edu/~g.legrady/academic/courses/22w259/22w259.html>

Course syllabus, code samples, references, and project results

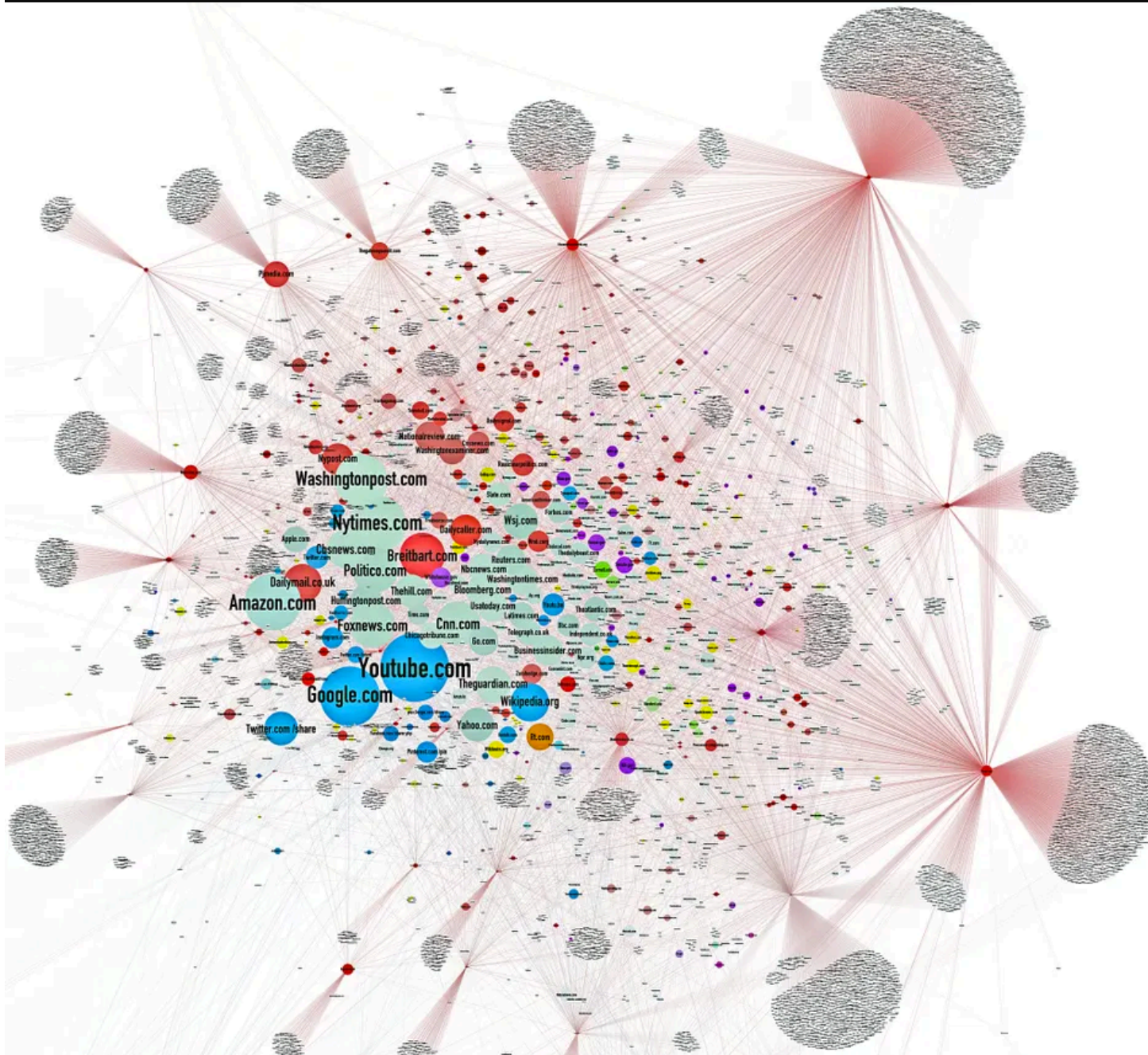
BLOG: (Student Forum)

Post your concepts, sketches, share algorithms, visualization techniques here

PREVIOUS STUDENT WORKS: (vislab.mat.ucsb.edu/courses)

Reference projects with code and documentation

Spatial Map of Right Wing FAKE NEWS info system



https://www.theguardian.com/technology/2016/dec/04/google-democracy-truth-internet-search-facebook?utm_content=buffere55bd&utm_medium=social&utm_source=facebook.com&utm_campaign=buffer&fbclid=IwAR1pkm0ywt

Course Options

- A **self-directed data visualization project** where each student selects their data source and software at the start of the course and proceeds to realize an interactive visualization and its documentation on their own with regular presentation to the class
- **Requirements:**
 - 1) Has a multi-dimensional data to visualize
 - 2) Experience in programming
 - 3) Wants to gain experience in exploring visualization
- OR **follow the course syllabus**

Course Knowledge Acquisition

- **MySQL basics** to explore, identify and retrieve significant data from a dataset
- **Visualize data** through computer programming
- **Visual Language:** How form, color, space, timing, movement, etc. impact on content (this is the primary goal of the course)
- **Through visualization abstract** reveal patterns and relationships
- **Normalize data** to enhance legibility and coherence
- **Implement interactivity** within 3D volumetric visualization
- **Correlate 2** sets of data from diverse sources

An Interdisciplinary Approach

We want to integrate diverse expertises:

- **Analytical:** theoretical, cultural, information research (*social scientist*)
- **Technical:** Statistical aggregation and computational processes (*engineer, scientific*)
- **Aesthetic:** Visual design, expression (*artist, designer, architect*)

Course Schedule

Wk 1-2 Data Analytics, Knowledge Discovery

- Content Analysis with MySQL
- Discover unexpected, interesting patterns, anomalies in a large dataset

Wk 3-6 3D Interactive Visualization

- Frequency Data Mapping - Visual Language Basics
- Acquire visual language basics, visualize data in java-based Processing
- Design in 3D-interactive space/time, implement associative rule-mining and other algorithms, correlate with JSON

Wk 7-10 Student Defined Project

- Each student defines a project with their own data source, a project that builds on skills acquired through the two previous assignments

Course Data: Seattle Public Library Checkouts

- **A database of over 100 million datasets**
- **Data collected by the hour since September 2005**
- **Data is multivariate: numeric, sequential, semantic**
- **Multi-format: books, cds, dvds, etc.**
- **Broad topic range: Dewey Decimal Classification**

Project Components

Every project has conceptual, technical and aesthetic challenges:

- **MySQL Data discovery**
(What is an interesting query?)
- **3D Spatialization**
(Interactive, spatial visualization)
- **Design Decisions**
(Choices in how to organize the data have different degrees of effectiveness – or “interestingness”)

Software Resources

- MySQL open-source relational database management system
- MySQL Workbench to access the data
- Processing Java-based scripting language used by graphic designers
- JSON a text-based data-interchange format for data correlation

Why we are using the same data source

Seattle Public Library data for assignments
MySQL and 3D visualization so that:

- Results can be ***compared*** and ***shared***
- Learn quickly from previous projects

Final Project is student determined: Each student choose their data source for the final project

Data Science - Work Effort

- Collecting the data (19%)
- Cleaning and organizing data (60%)

Most data sets require extensive processing as they contain inaccuracies (noise)!!

-
- Building training sets (3%)
 - Mining data for patterns (9%)
 - Refining algorithms (4%)

-
- Visualization Follows

Intellectual Property, Copyright, etc.

- Data and lectures presented in the course is the property of the professor and the lab, so it needs to be credited
- You have copyright over your projects

But we **agree to share** – make sure to **give credit** if you use code from another source

`/* this section of code came from .....*/`

Fill out and sign forms...

**Lets get access to course software
and resources!**

George Legrady © 2021
Experimental Visualization Lab
Media Arts & Technology
University of California, Santa Barbara