

Data Visualization II

Jason H. Moore, Ph.D.

Edward Rose Professor of Informatics
Director, Institute for Biomedical Informatics
Perelman School of Medicine
University of Pennsylvania
Philadelphia, PA, USA

upibi.org

epistasis.org

jhmoore@upenn.edu

Golden Era of Biomedical Informatics

Moore and Holmes, *BioData Mining* (2016)

Why?

- Big data
- High-performance computing
- Powerful software
- Talented trainees
- Government recognition
- Industry recognition
- Patient recognition
- University investment
- PSB attendance booming!

Golden Era of Biomedical Informatics

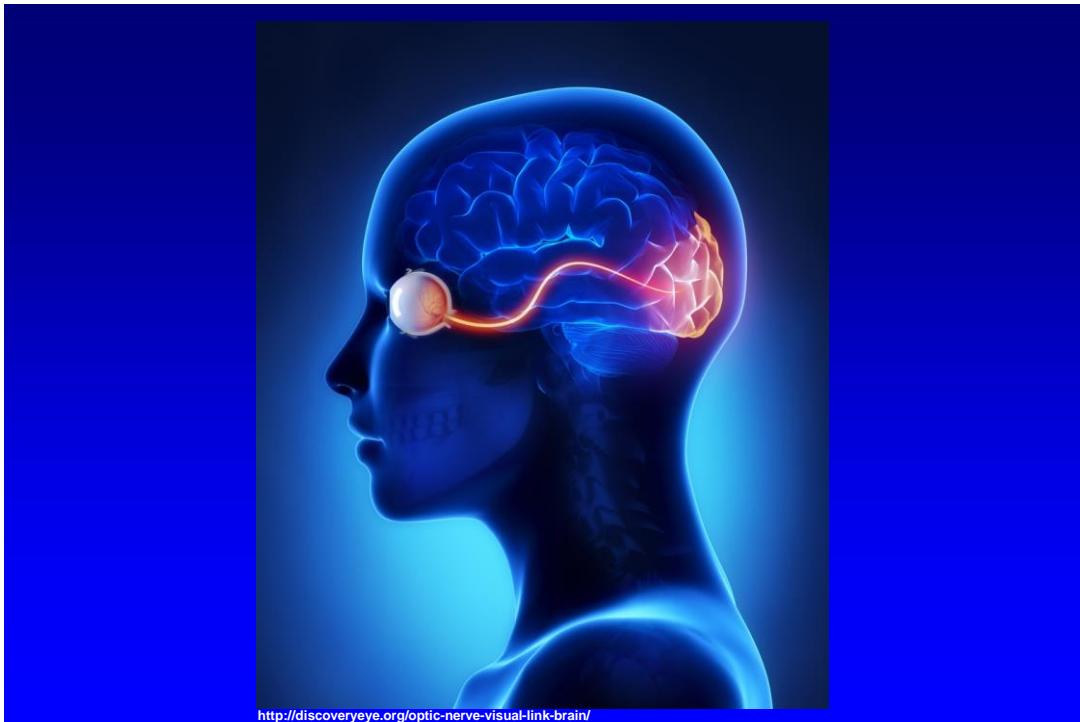
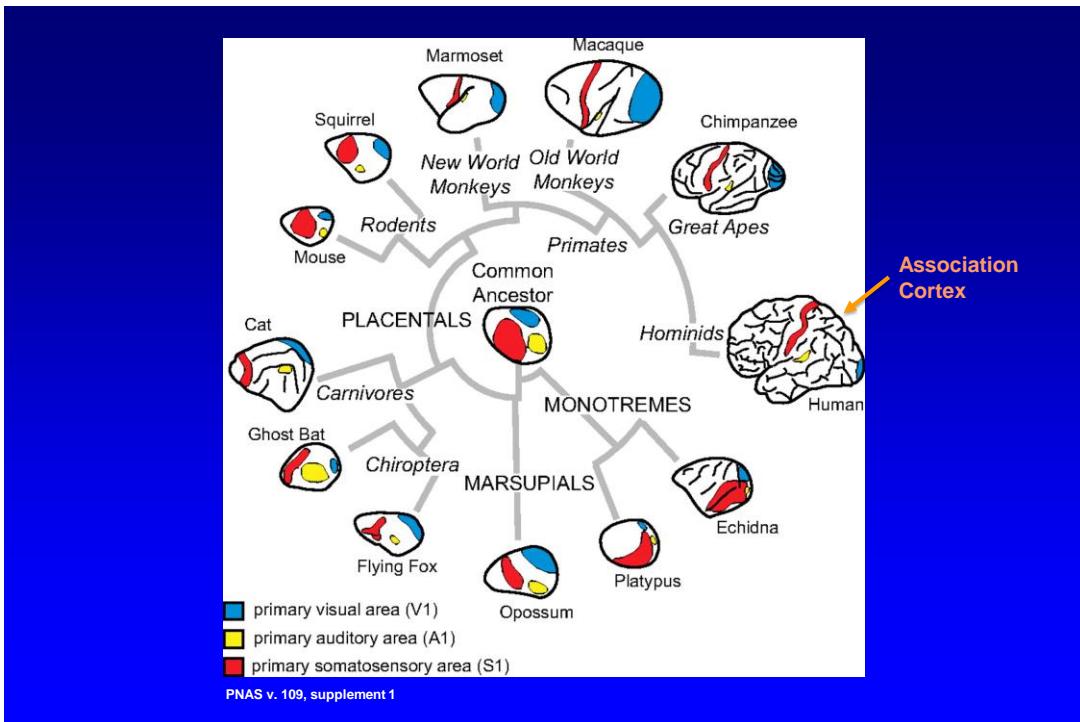
Moore and Holmes, *BioData Mining* (2016)

What next?

- Artificial intelligence
- Biomedical devices
- Data integration
- Data science
- Network science
- Informatician scientists
- No-boundary thinking
- Visual analytics



Mural at Humboldt State University

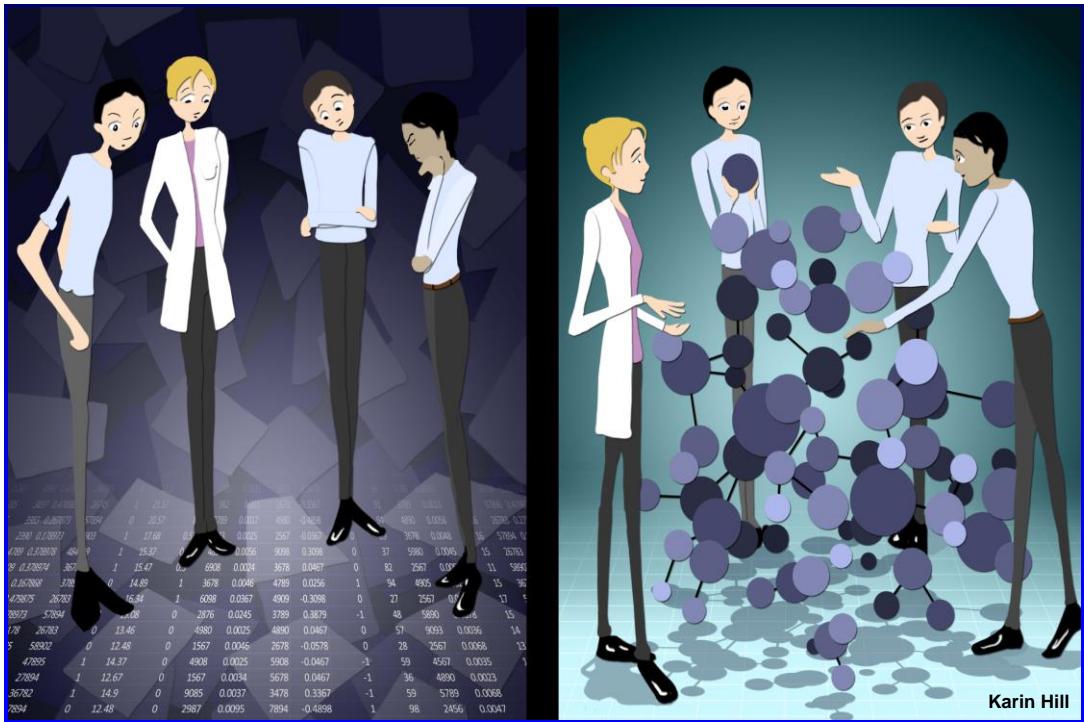


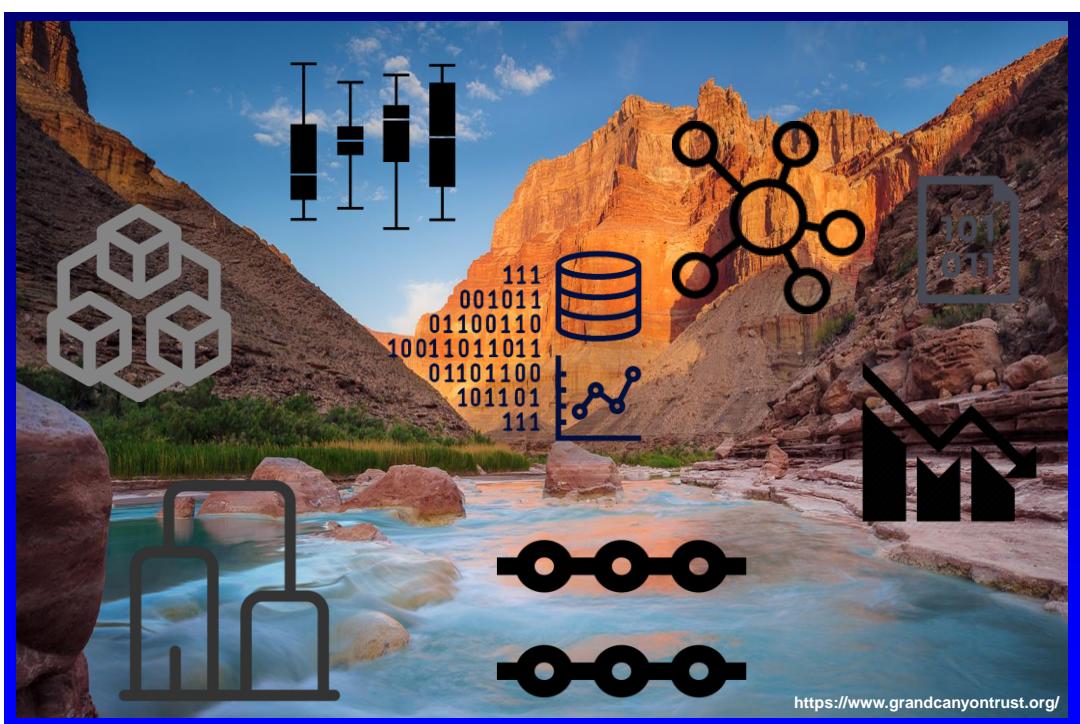
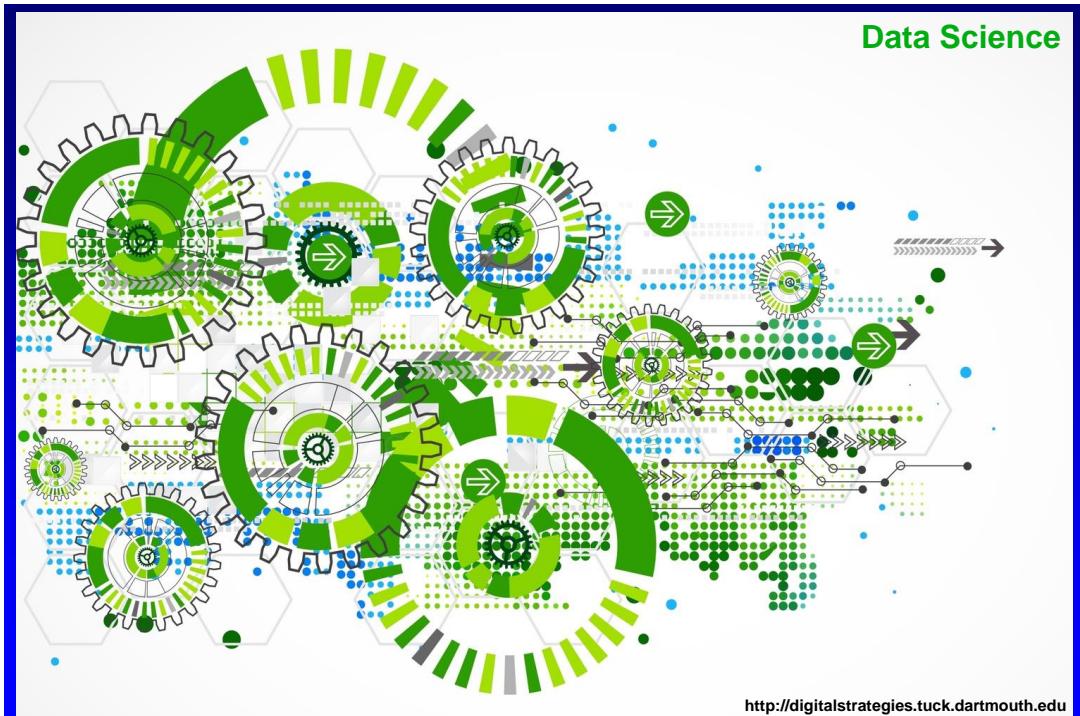


<https://www.grandcanyontrust.org/>



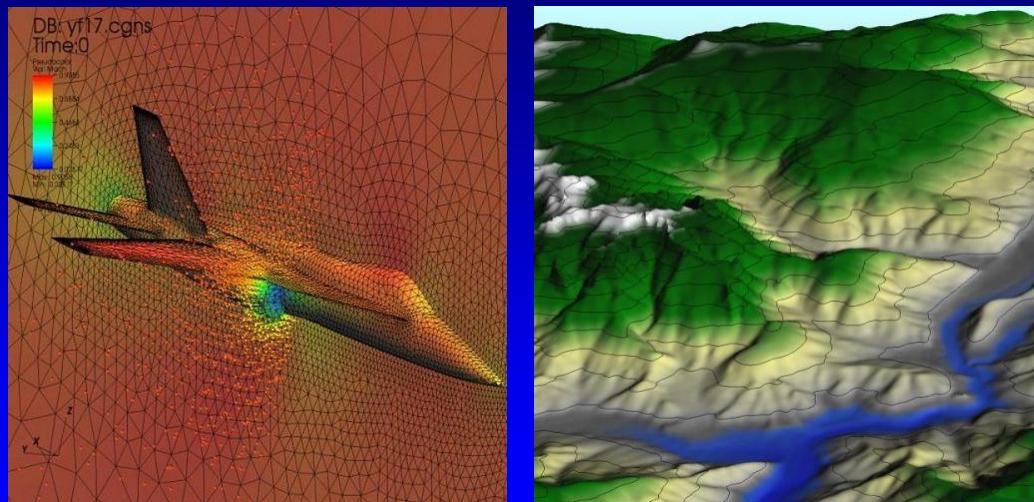
<https://blog.columbia.com/shenandoah-national-park-adventures/>



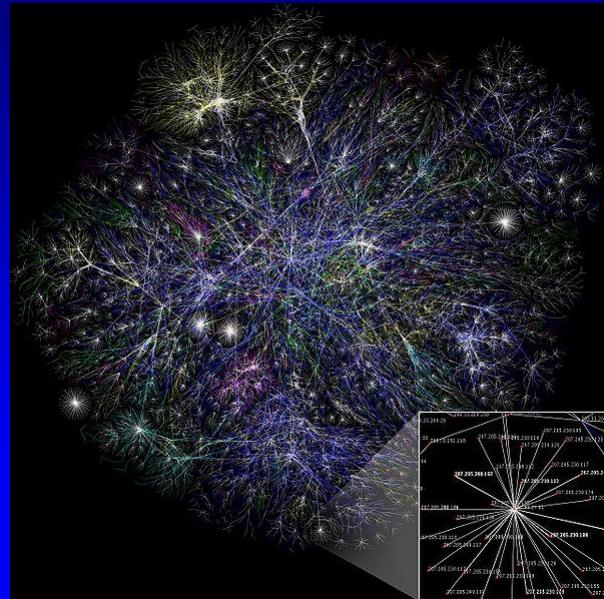


Definitions

Scientific Visualization



Information Visualization



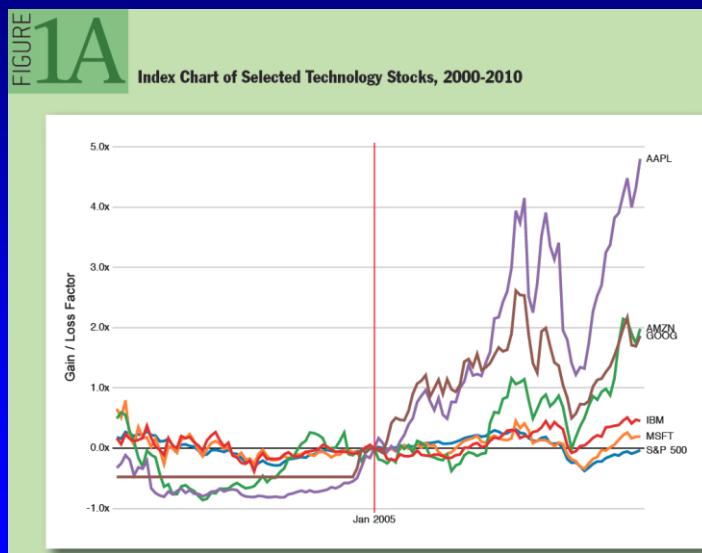
Visual Analytics



Information Visualization

An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

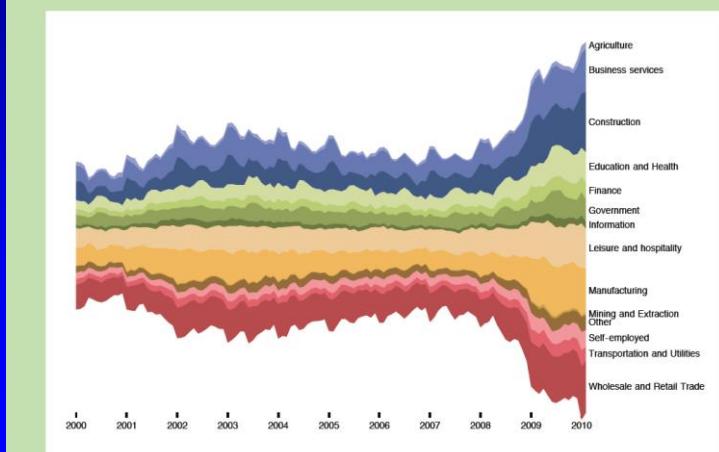


An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

FIGURE
1B

Stacked Graph of Unemployed U.S. Workers by Industry, 2000-2010

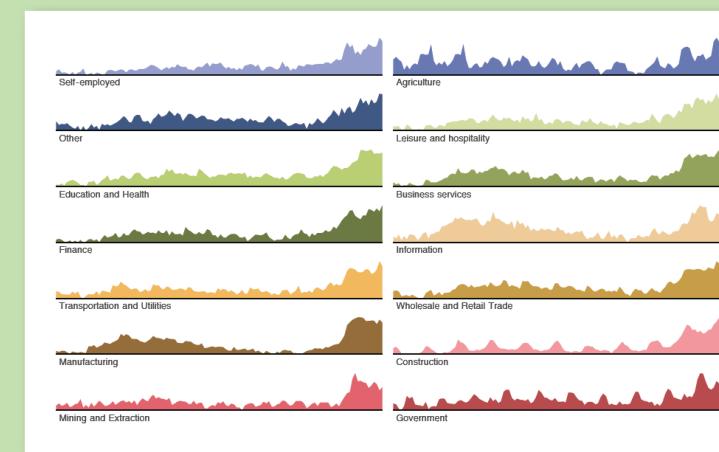


An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

FIGURE
1C

Small Multiples of Unemployed U.S. Workers Normalized by Industry, 2000-2010



An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010



An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010



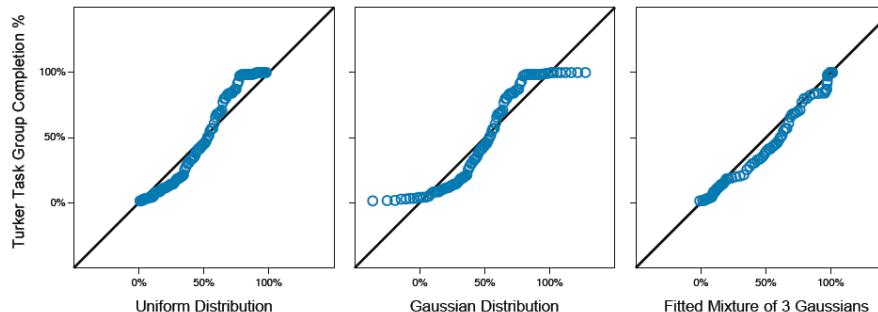
0	1	1	1	2	2	2	2	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	5	6	7	8	8	8	9
1	0	0	0	0	1	1	1	1	2	2	3	3	3	3	4	4	4	4	5	5	6	7	7	8	9	9	9	9	
2	0	0	1	1	1	5	7	8	9																				
3	0	0	1	2	3	3	3	4	6	6	8	8																	
4	0	0	1	1	1	1	3	3	4	5	5	5	6	7	8	9													
5	0	2	3	5	6	7	7	7	9																				
6	1	2	6	7	8	9	9	9																					
7	0	0	0	1	6	7	9																						
8	0	0	1	2	3	4	4	4	4	4	4	4	5	6	7	7	7	9											
9	1	3	3	5	7	8	8	8	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9		
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

FIGURE
2B

Q-Q Plots of Mechanical Turk Participation Rates

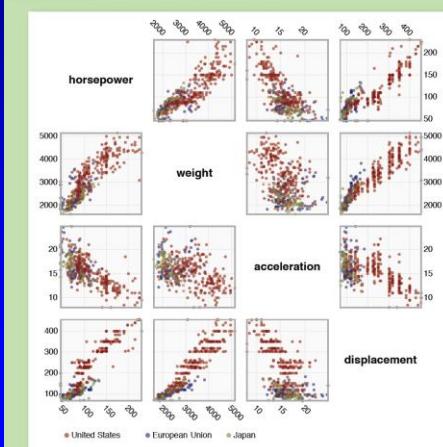


An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

FIGURE
2C

Scatter Plot Matrix of Automobile Data

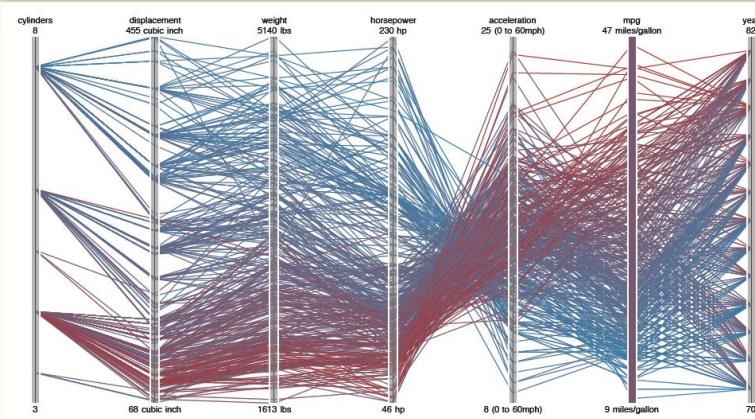


An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

FIGURE
2D

Parallel Coordinates of Automobile Data

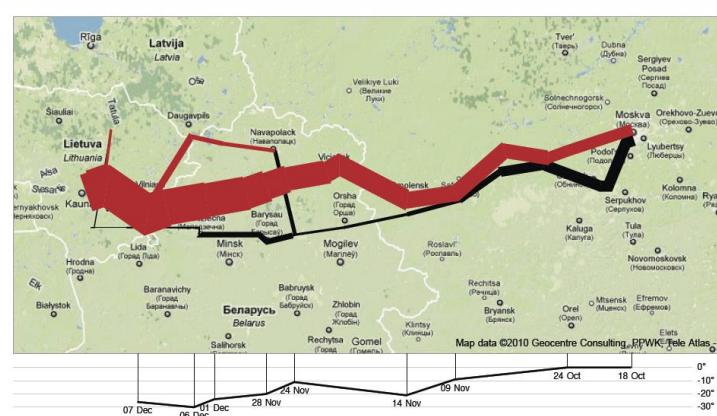


An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

FIGURE
3A

Flow Map of Napoleon's March on Moscow

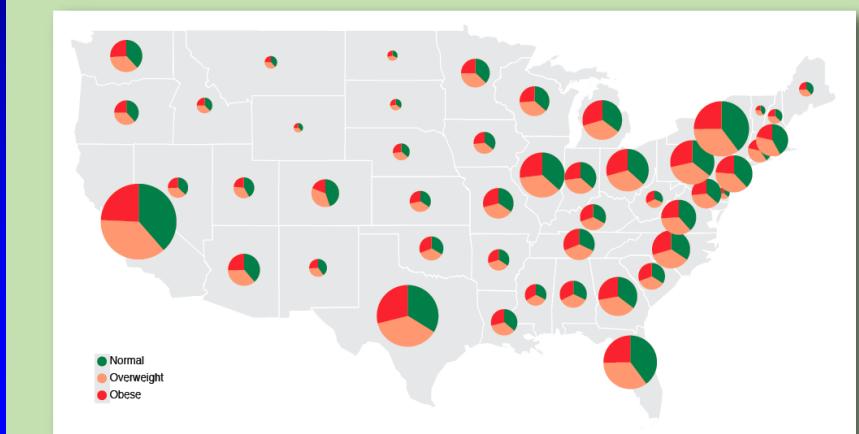


An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

FIGURE
3C

Graduated Symbol Map of Obesity in the U.S., 2008

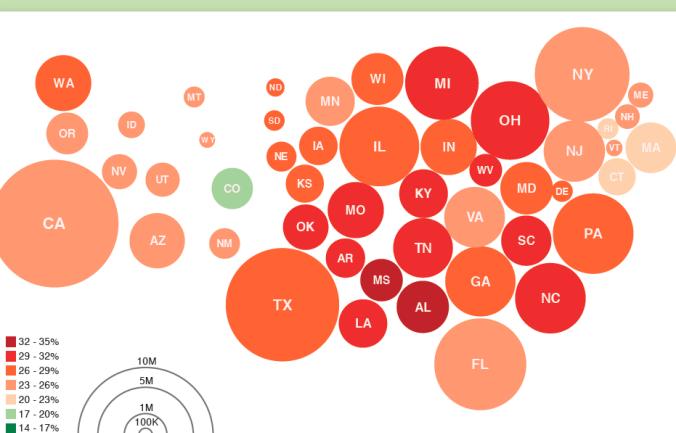


An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

FIGURE
3D

Dorling Cartogram of Obesity in the U.S., 2008

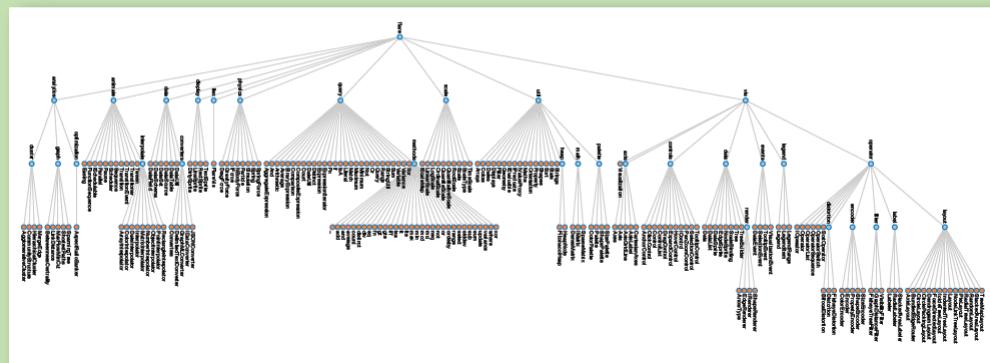


An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

FIGURE
4A

Radial Node-link Diagram of the Flare Package Hierarchy

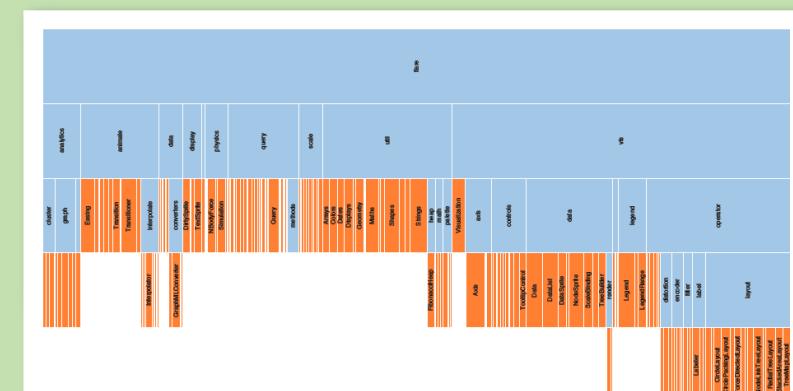


An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

FIGURE
4D

Icicle Tree Layout of the Flare Package Hierarchy

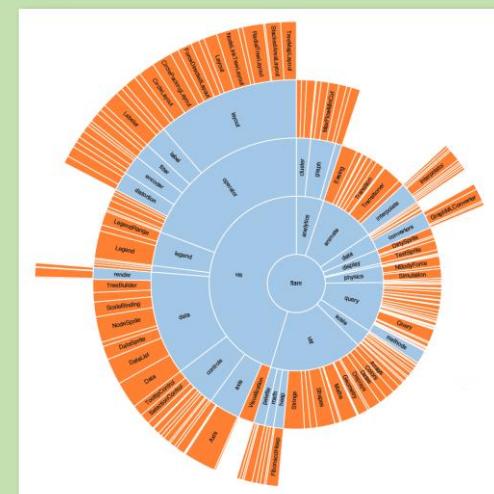


An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

FIGURE
4E

Sunburst (Radial Space-filling) Layout of the Flare Package Hierarchy



An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

FIGURE
4F

Treemap Layout of the Flare Package Hierarchy

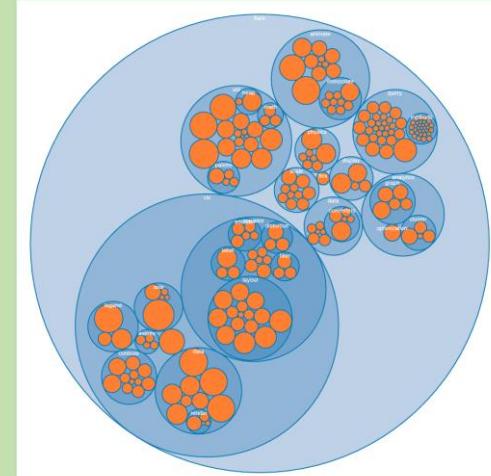


An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

FIGURE
4G

Nested Circles Layout of the Flare Package Hierarchy

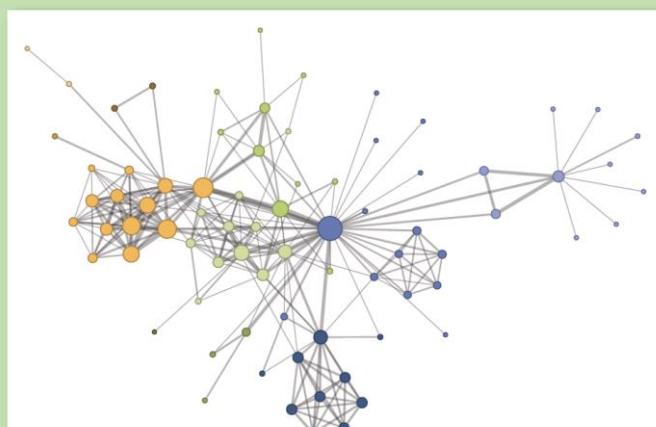


An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

FIGURE
5A

Force-directed Layout of *Les Misérables* Character Co-occurrences

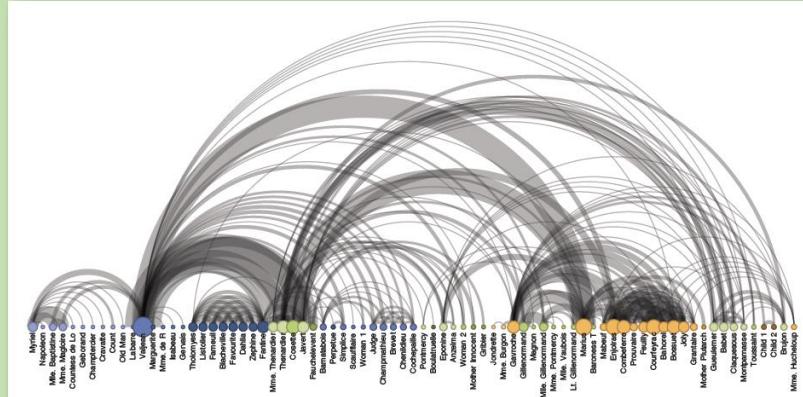


An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

FIGURE
5B

Arc Diagram of *Les Misérables* Character Co-occurrences

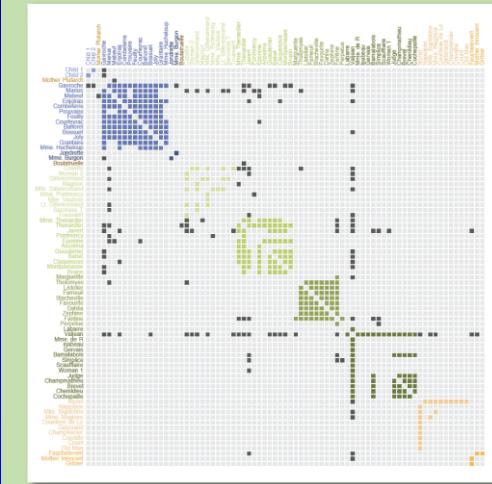


An Information Visualization Tour

Heer et al. *Communications of the ACM*, June 2010

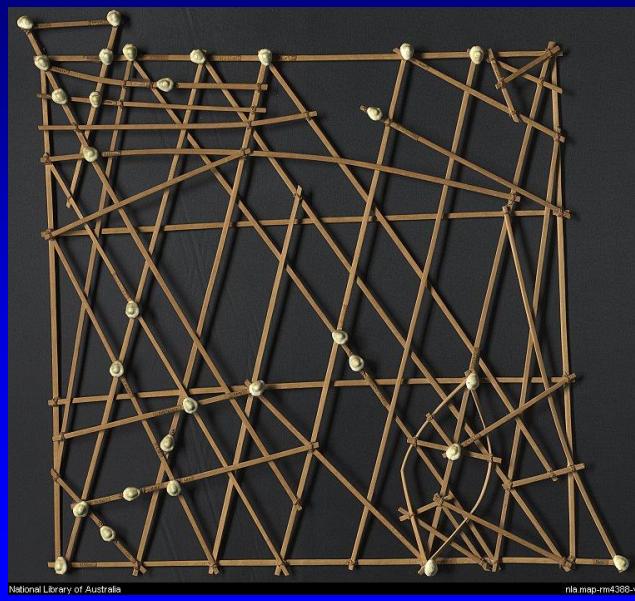
FIGURE
5C

Matrix View of *Les Misérables* Character Co-occurrences

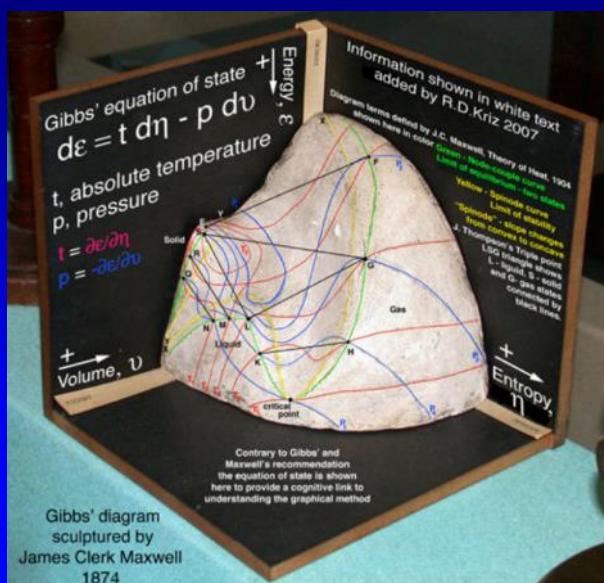


Physical Visualization

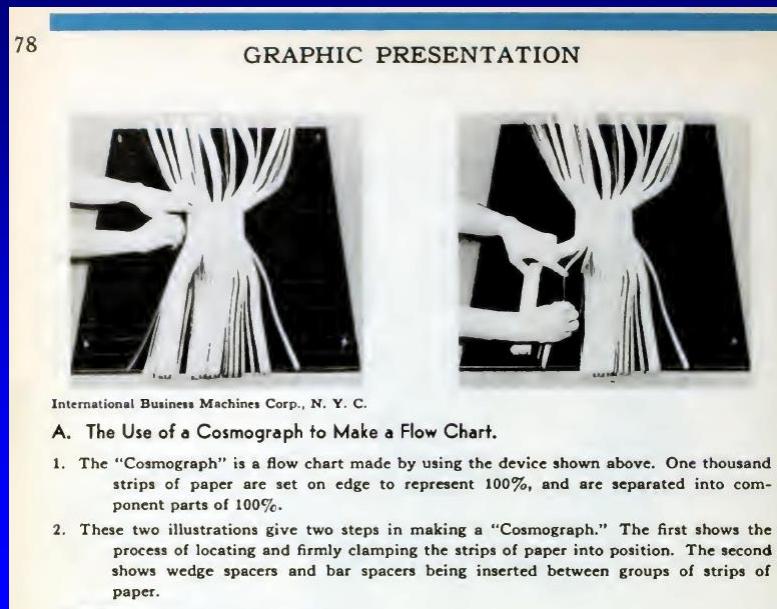
Marshall Island Stick Charts (1862)



Maxwell's thermodynamic surface (1874)



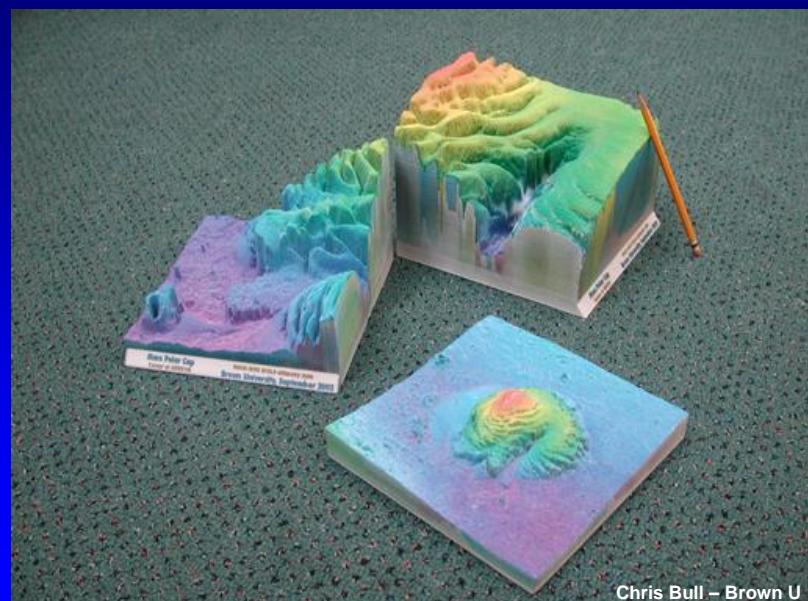
The Cosmograph – IBM (1933)



Bertin's Reorderable Matrices (1968)



3D-Printed Visualizations (2004)



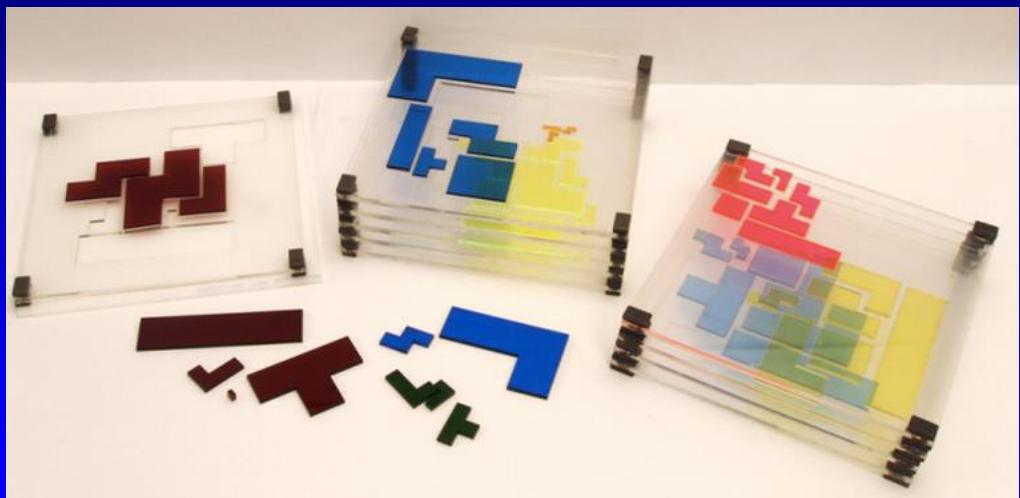
3-D Models (2005)



Unemployment rate plotted against inflation for 8 countries over 10 years

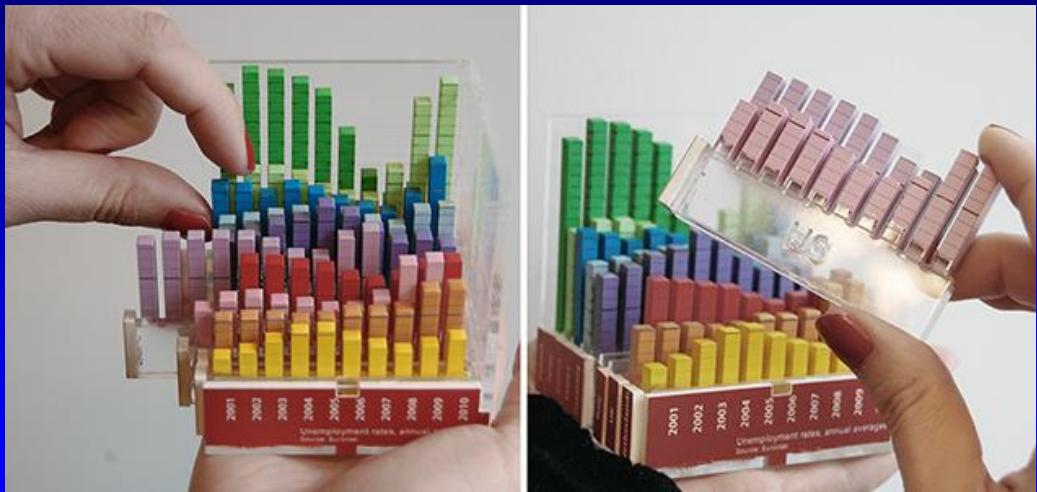
Tim Dwyer, Ph.D.

Data Sculptures (2010)



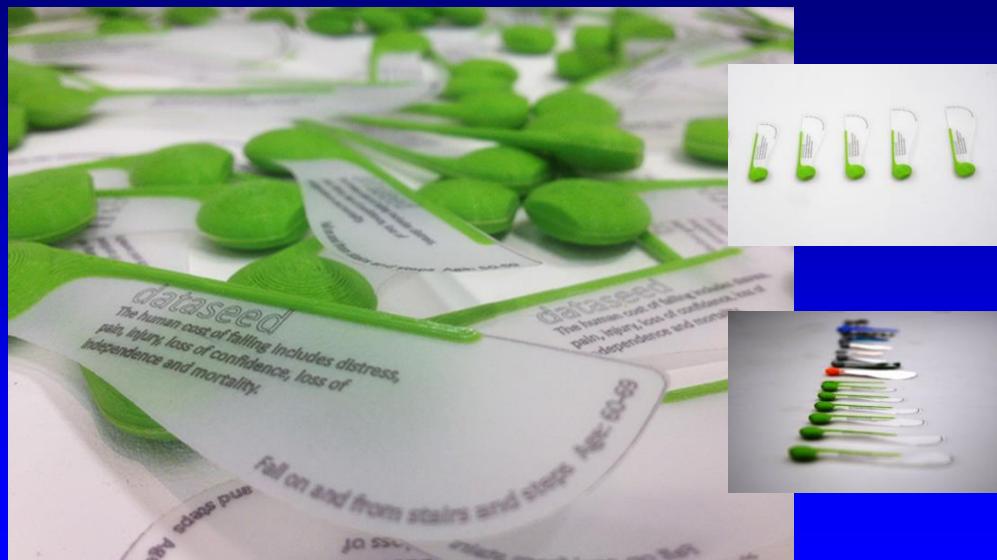
Andrew Vande Moere

Rearrangeable 3D Bar Chart (2013)



Yvonne Jansen and Pierre Dragicevic

Data Seeds (2015)



Nick Dulake and Ian Gwilt

dataphys.org/list/

List of Physical Visualizations and Related Artifacts

This is a chronological list of physical visualizations and related artifacts, maintained by Pierre Dragicevic and Yvonne Jansen. Thanks to Fanny Chevalier and our other contributors. Feel free to post a general comment or if you know of another interesting physical visualization, please submit it!

This list currently has 270 entries. Recent additions:

[Wearable Self](#)
[U.S. Unemployment Rate 1948-2015](#)
[EuroGums: Edible Population Pyramids](#)
[See more...](#)

Also show entries on related artifacts below

5500 BC – Mesopotamian Clay Tokens



The earliest data visualizations were likely physical: built by arranging stones or pebbles, and later, clay tokens. According to an eminent archaeologist (Schmandt-Besserat, 1999): "Whereas words consist of immaterial sounds, the tokens were concrete, solid, tangible artifacts, which could be handled, arranged and rearranged at will. For instance, the tokens could be ordered in special columns according to types of merchandise, entries and expenditures; donors [...]

Virtual Visualization

Visualization and Virtual Reality

Second Life

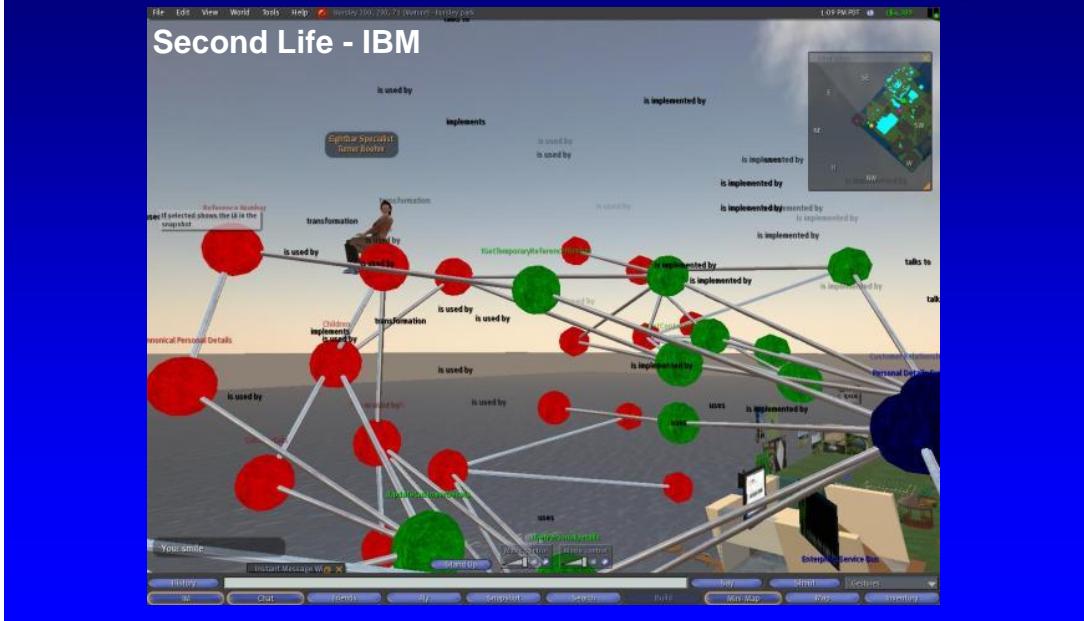


Visualization and Virtual Reality

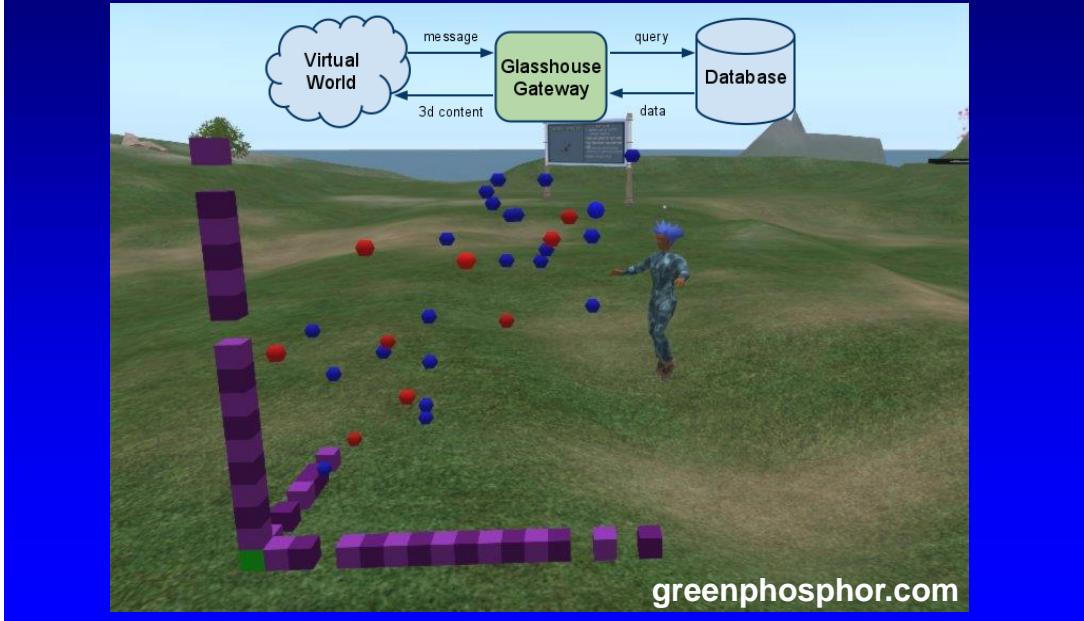
Second Life



Visualization and Virtual Reality



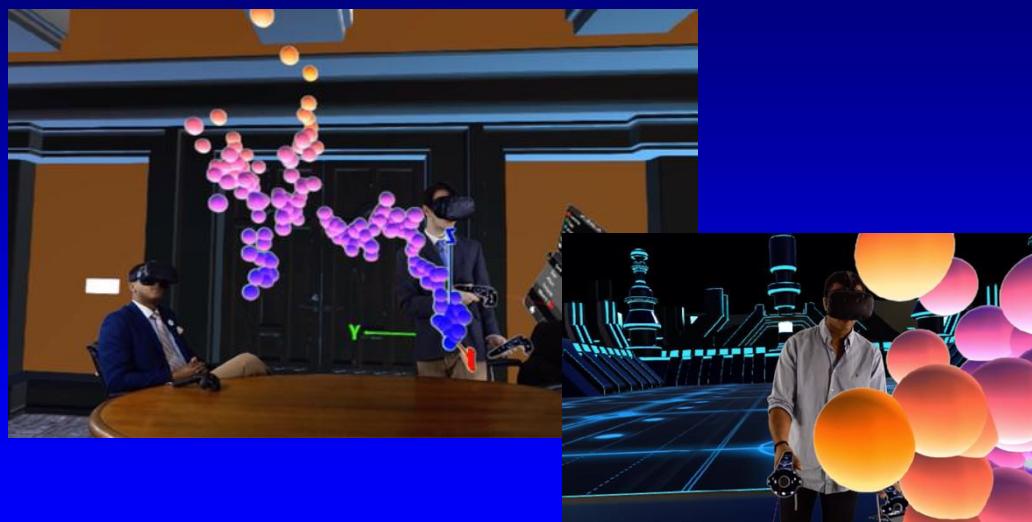
Visualization and Virtual Reality



Visualization and Virtual Reality



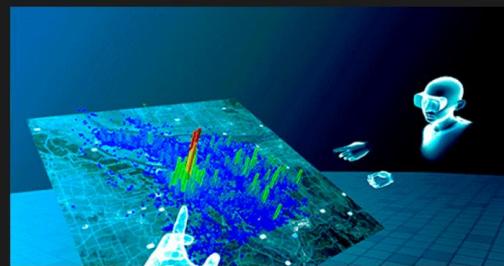
DatavizVR for HTC Vive



http://store.steampowered.com/app/551960/DatavizVR_Demo/

Virtualitics

We have built the first platform to merge **Artificial Intelligence**, **Big Data** and **Virtual/Augmented Reality**.



Fast, Reliable, Smart, Secure

Complete your tasks in a fraction of the time needed with traditional software and take advantage of smart routines.

Machine Learning and Natural Language

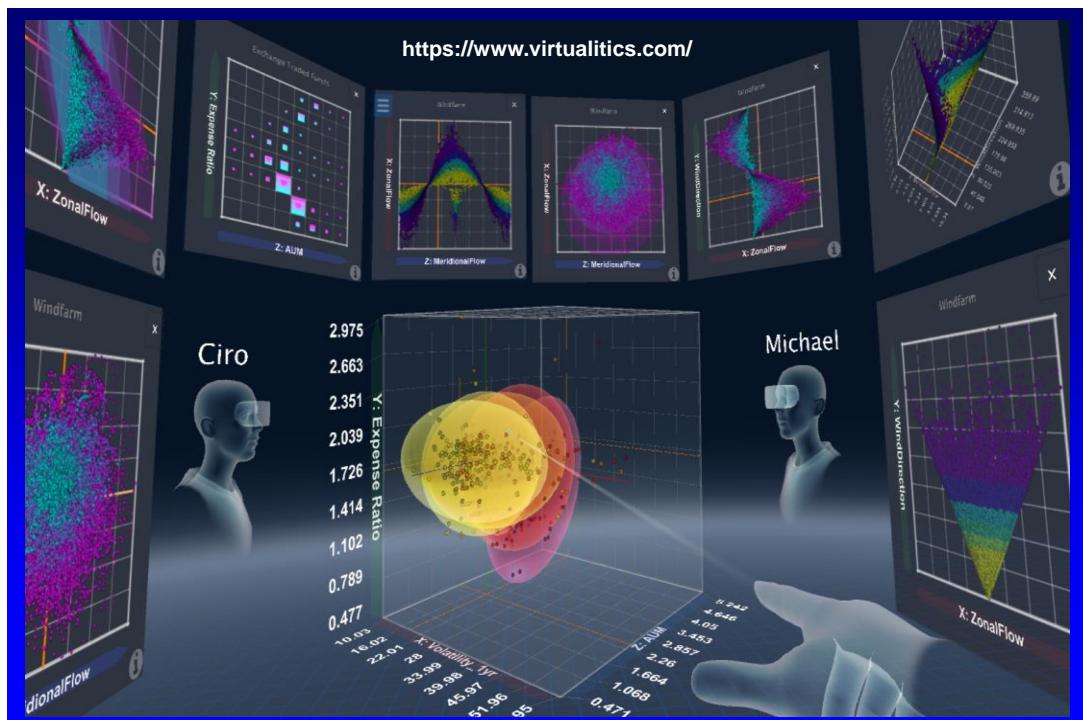
Smart mapping based on Machine Learning and Natural Language that summarizes the key patterns in the data.

Next generation data analytics

Visualize and understand your data as never before thanks to our innovative use of VR/AR, Machine Learning and Natural Language.

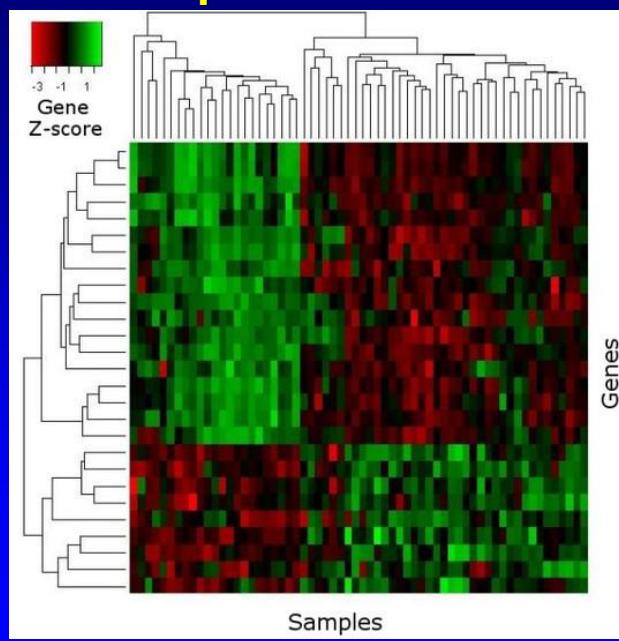
Shared Virtual Office (SVO)

Collaborative and fully customizable shared space where you can analyze data and present and discuss insights.

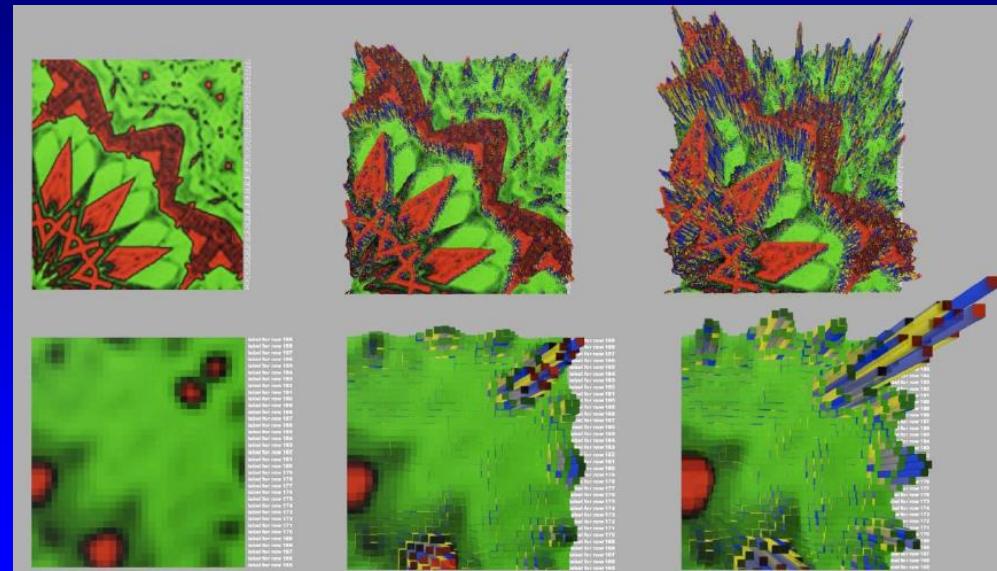


Harnessing the power of video game technology

2D Heat Map – Three Dimensions



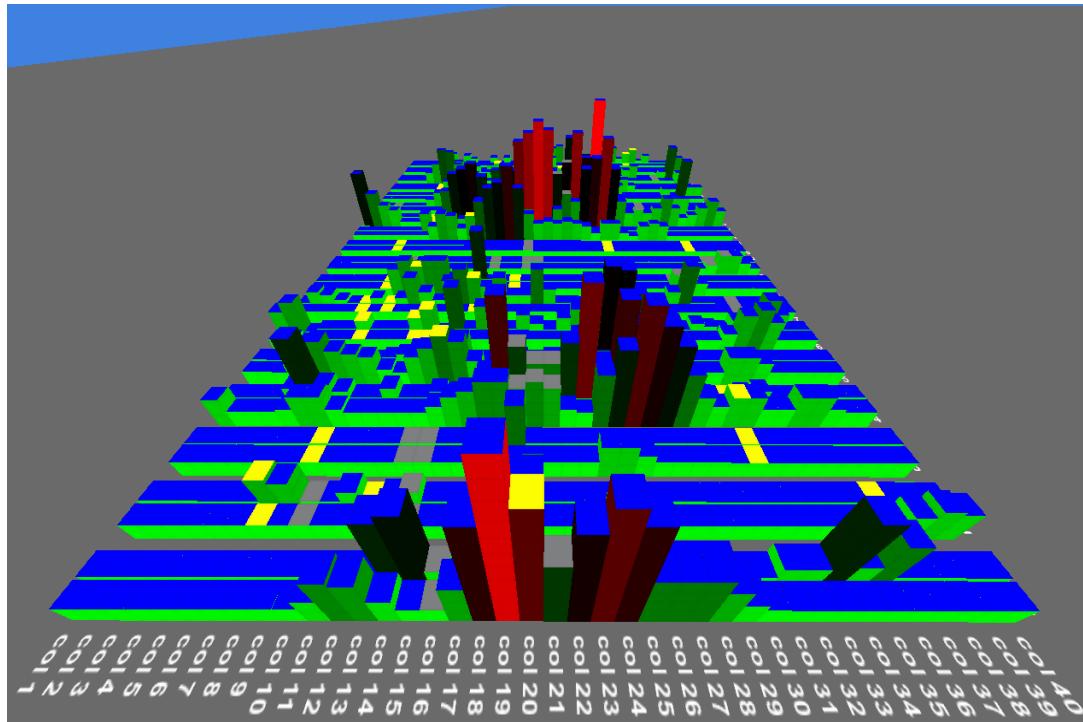
3D Heat Map – Five Dimensions

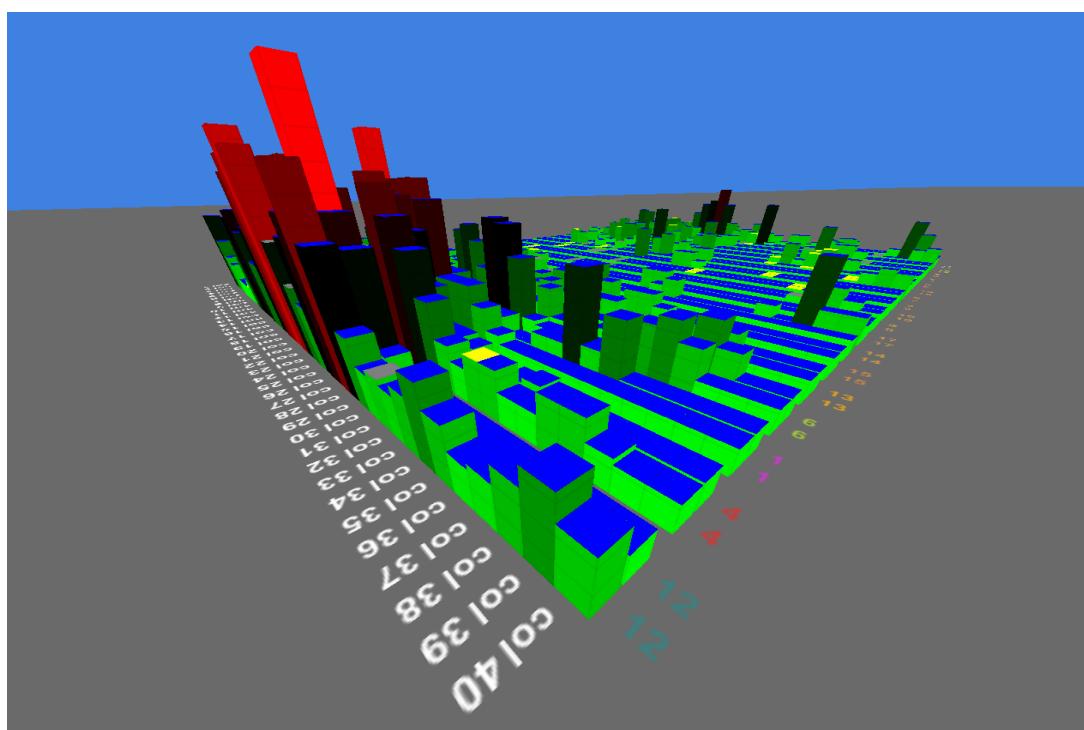
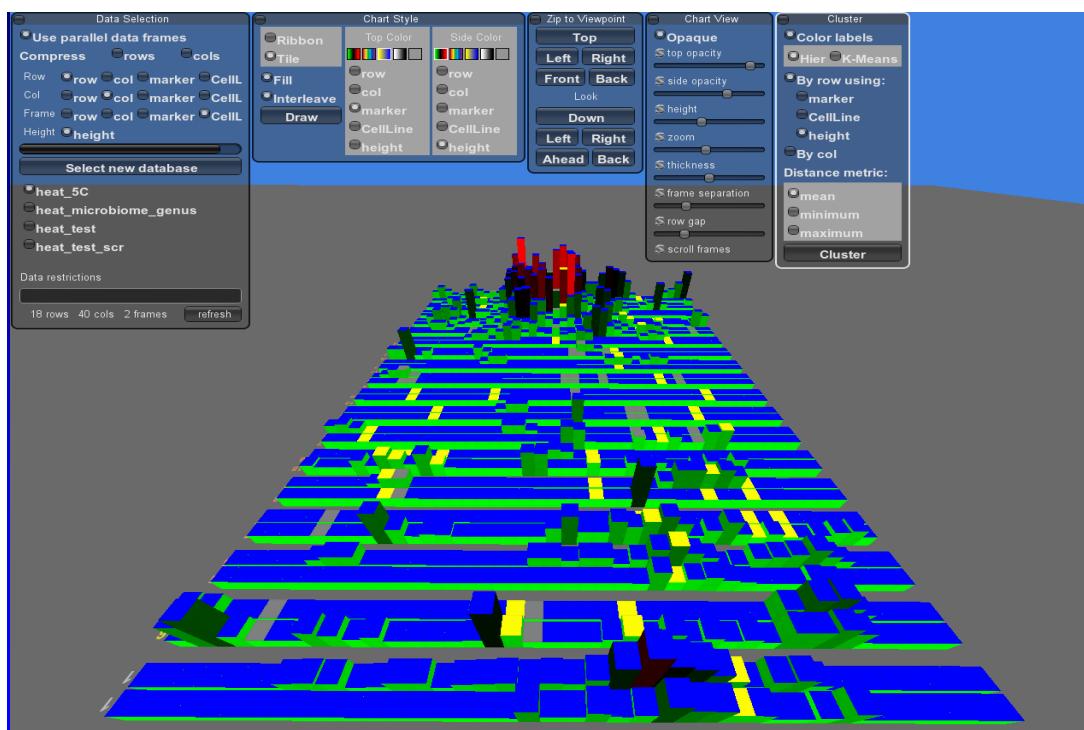


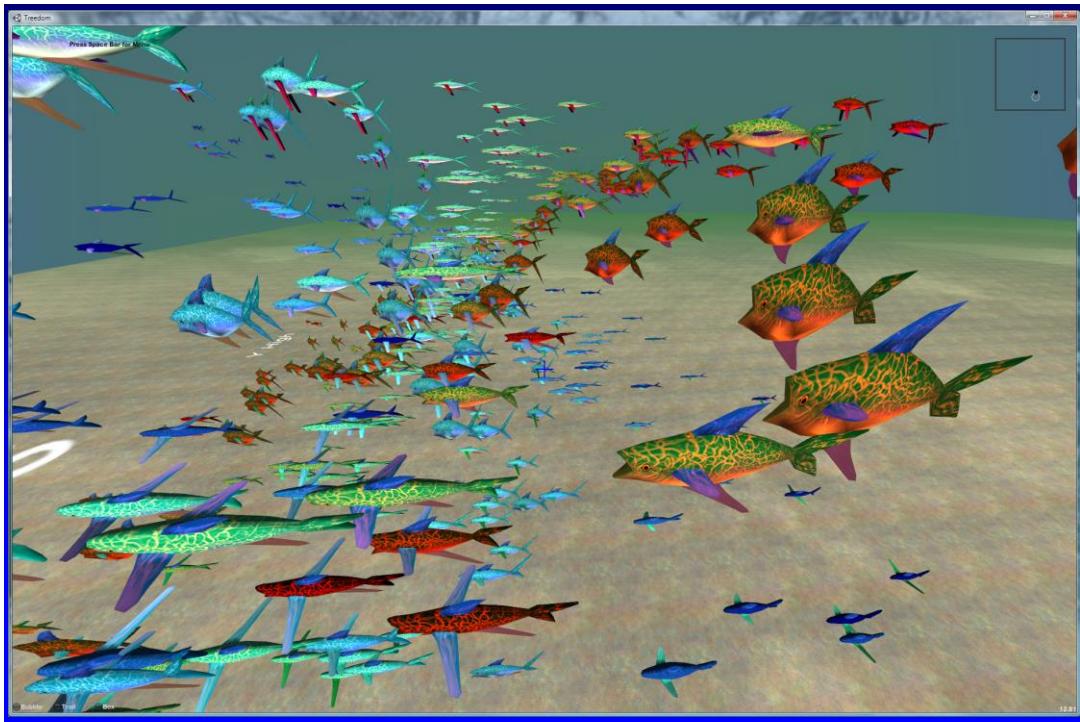
Visual Analytics using 3D Video Games – Unity 3D



3D Heat Map
<http://sourceforge.net/projects/dheatmap>
 Moore et al., *Pacific Symposium on Biocomputing* (2011)







Weiss et al. *BioData Mining* (2015) 8:22
DOI 10.1186/s13040-015-0056-2

BioData Mining

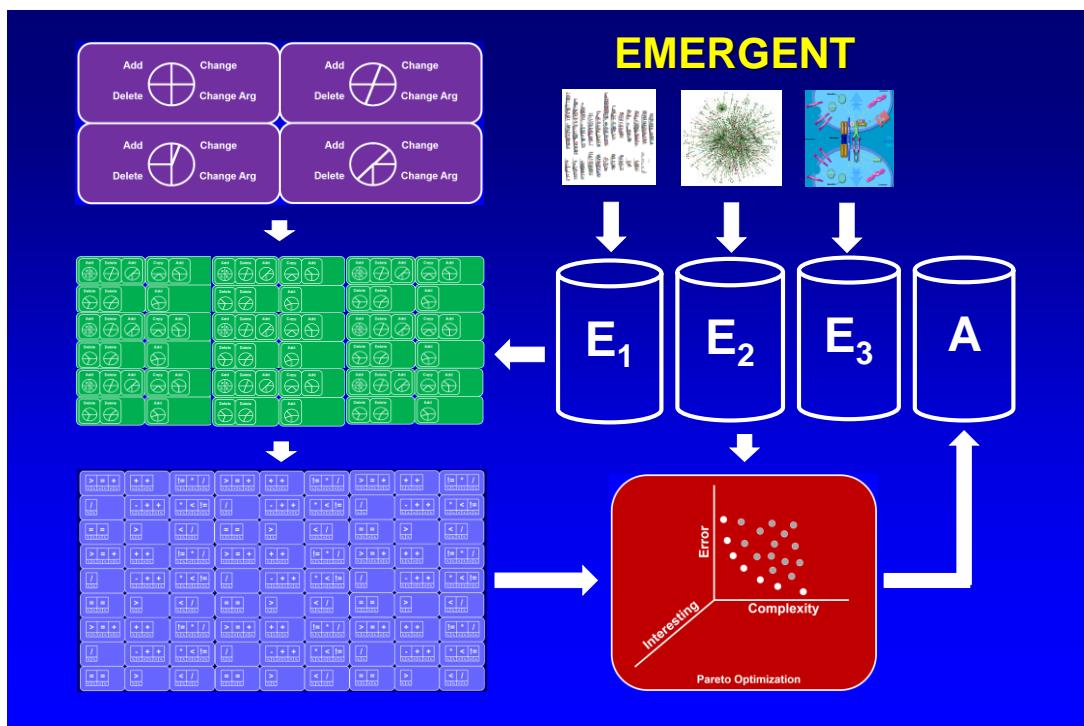
REVIEW **Open Access**

The role of visualization and 3-D printing in biological data mining

Talia L. Weiss¹, Amanda Zieselman¹, Douglas P. Hill¹, Solomon G. Diamond³, Li Shen⁴, Andrew J. Saykin⁴, Jason H. Moore^{1,2*} and for the Alzheimer's Disease Neuroimaging Initiative

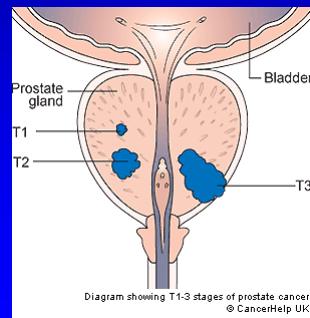
A 3D printed model of a complex network, likely a brain graph, composed of red connecting rods and white cube nodes. Each node is labeled with a number, such as 1, 2, 5, 6, 11, 13, 14, 18, 20, 25, 27, 30, 32, and 33. The model is displayed on a light-colored base plate with some text and numbers visible.

Application of visual analytics to genetics analysis using AI



Application to Prostate Cancer

- Dr. LaCreis Kidd (Univ. of Louisville, USA)
- 488 men with non-aggressive prostate cancer
- 687 men with aggressive prostate cancer
- Selected 219 SNPs
 - Biological filter
 - Apoptosis
 - Anti-oxidation
 - Carcinogen metabolism



© Jason H. Moore

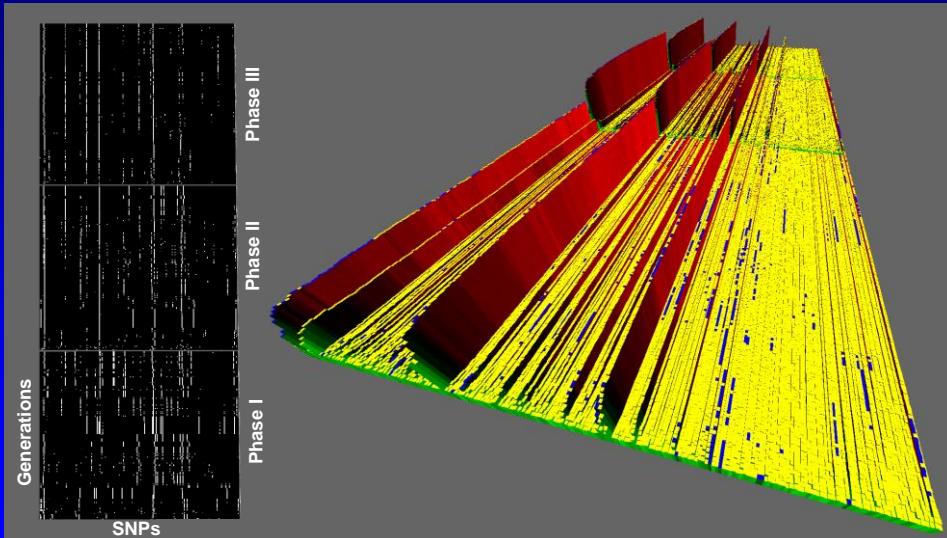
EMERGENT Parameters

- Function Set: +, -, *, %, <, <=, >, >=, ==, !=
- Fitness = Accuracy – α
- 18x18 grid of solutions
- 1000 gens/run
- 1000 runs/phase
- 3 Phases with cascading
- Expert knowledge from 3D visualization and HCI

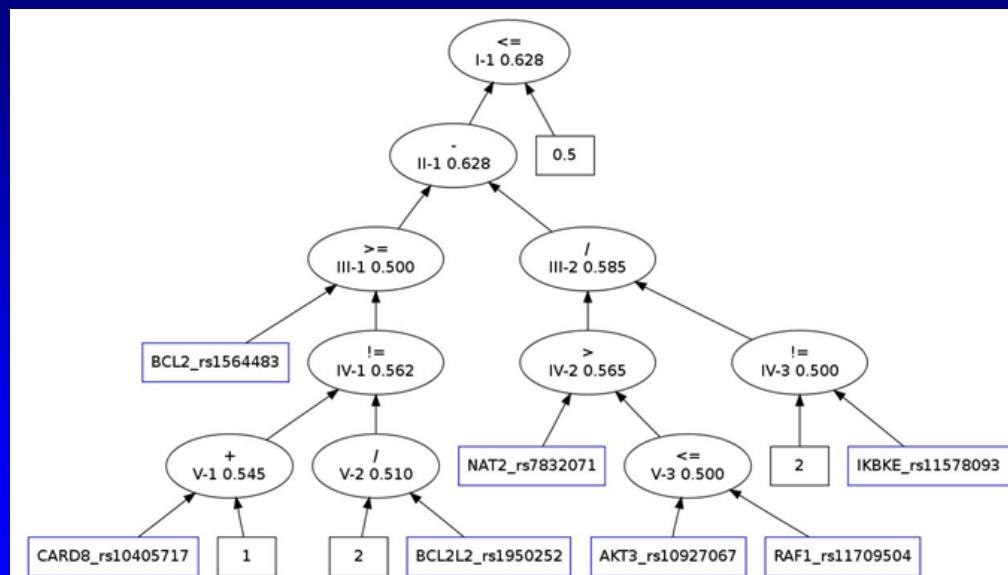
© Jason H. Moore

Visual Information as Expert Knowledge

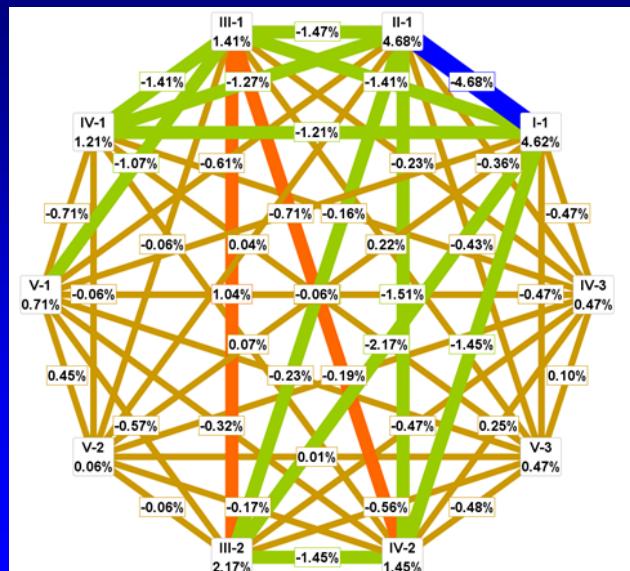
Moore et al., *Genetic Programming Theory and Practice* (2012)



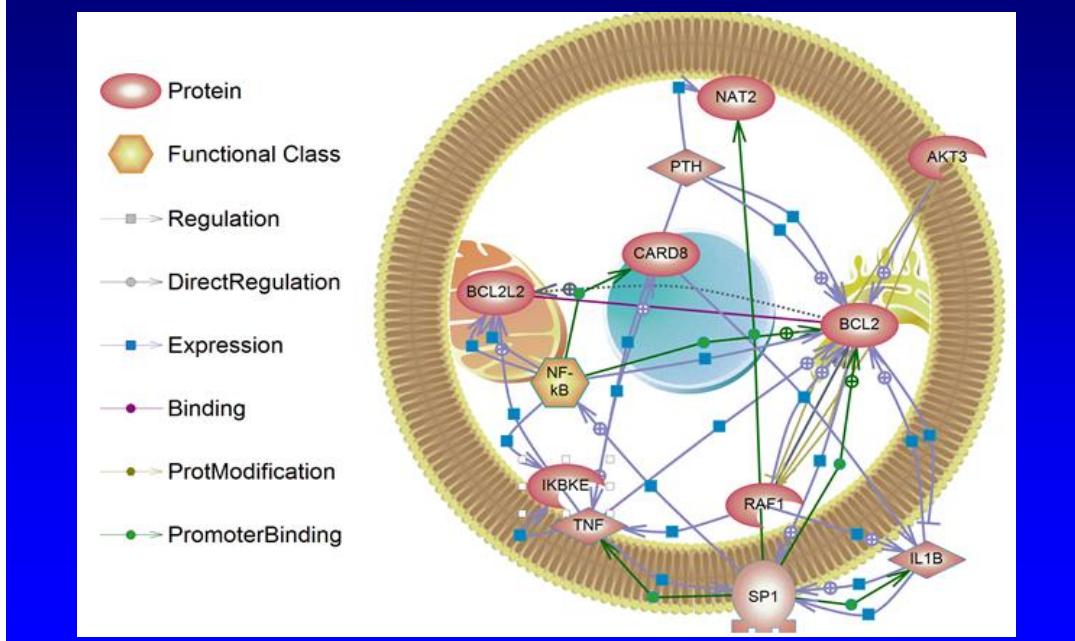
Best Model



Statistical Interpretation



Biological Interpretation



Visual Analytics at Penn



Penn IBI Idea Factory

Connecting Researchers with Ideas

