Ways of seeing data: A survey of fields of visualization

Gordon Kindlmann

glk@uchicago.edu

Nov 19, 2012

Part of the talk series "Show and Tell: Visualizing the Life of the Mind"
http://rcc.uchicago.edu/news/show_and_tell_abstracts.html

THE UNIVERSITY OF CHICAGO
COMPUTER SCIENCE

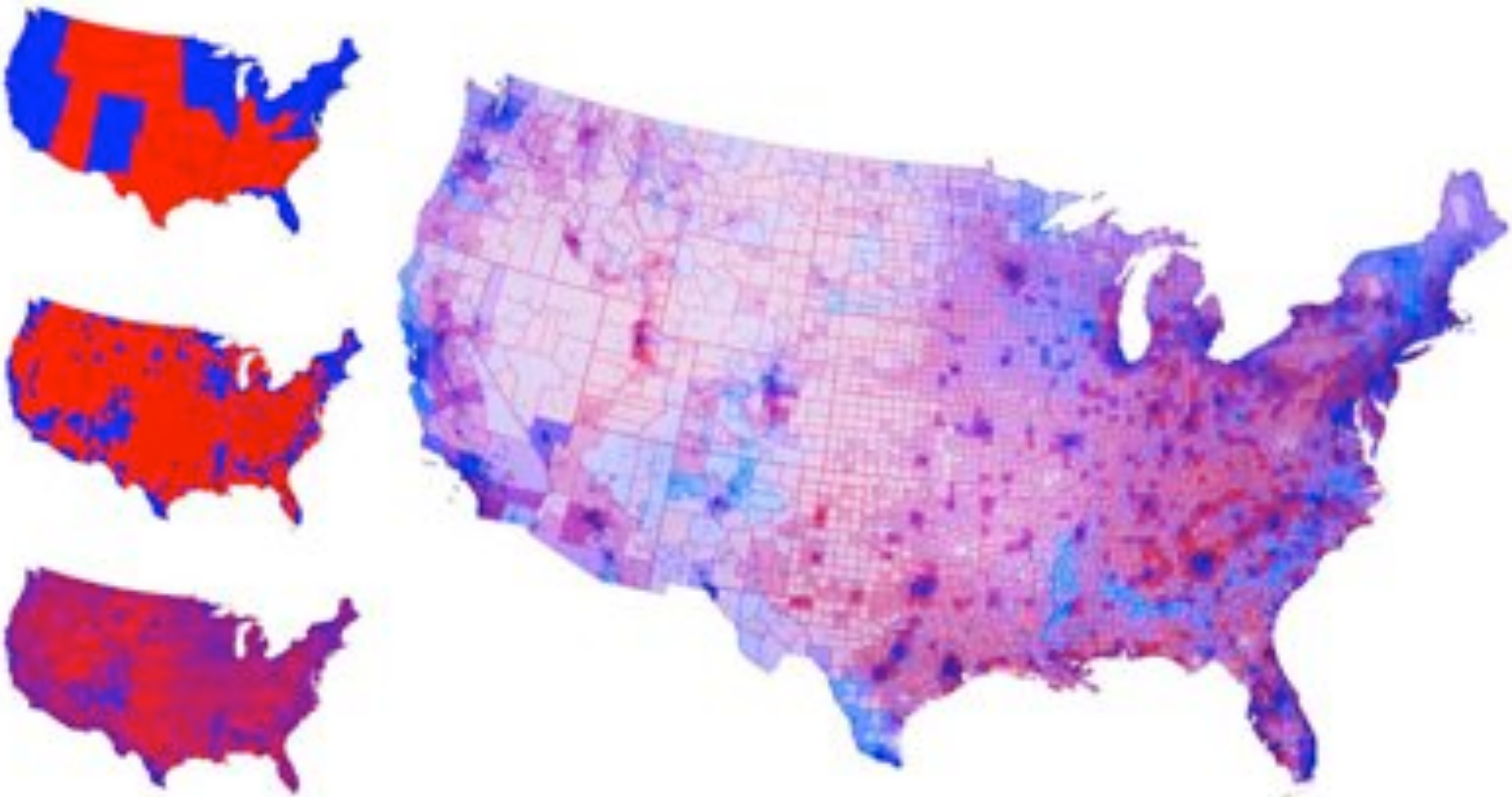
Computation Institute

2012 Presidential Election

REPUBLICAN | DEMOCRAT

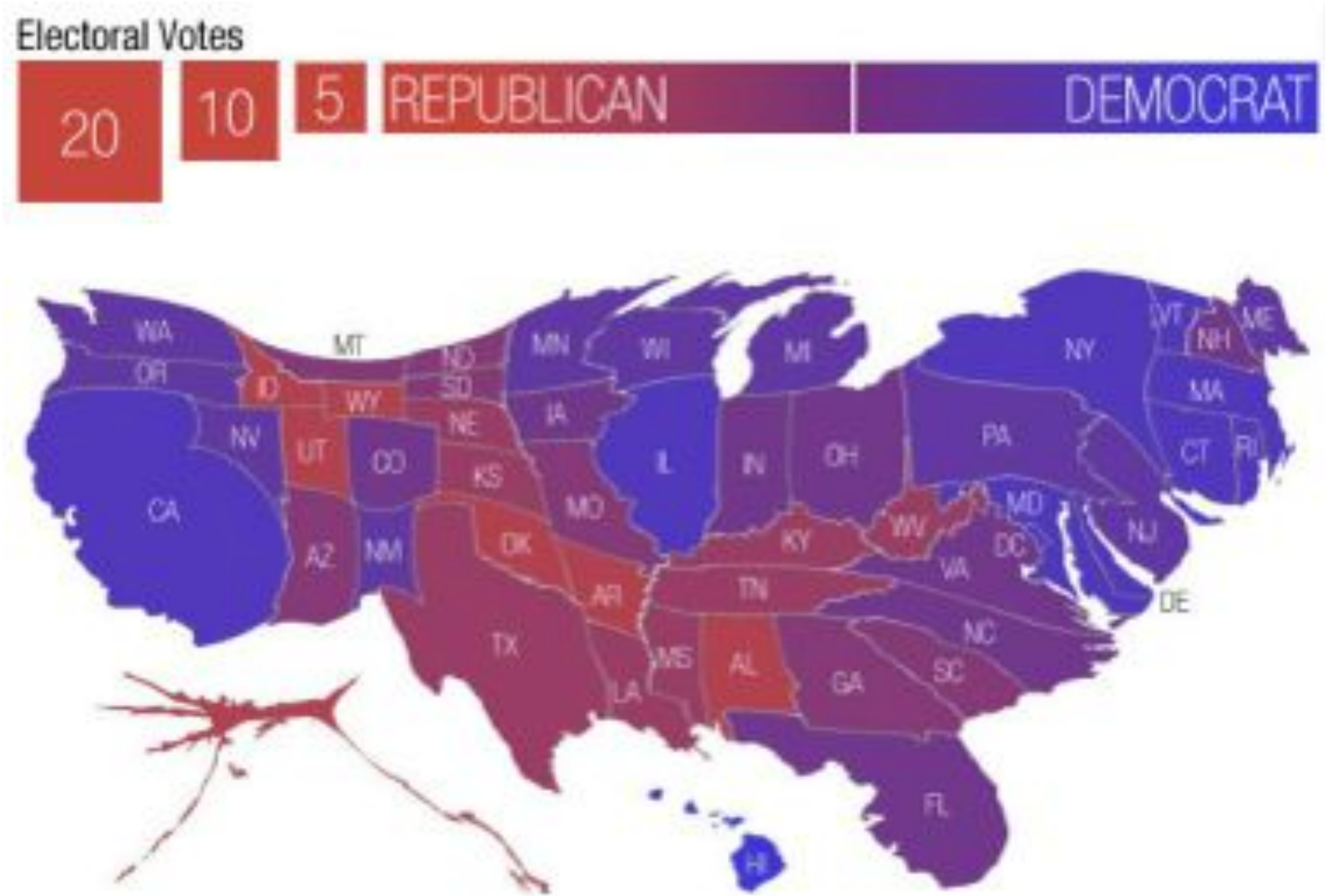
http://www.npr.org/blogs/itsallpolitics/2012/11/01/163632378/a-campaign-map-morphed-by-money

2012 Presidential Election



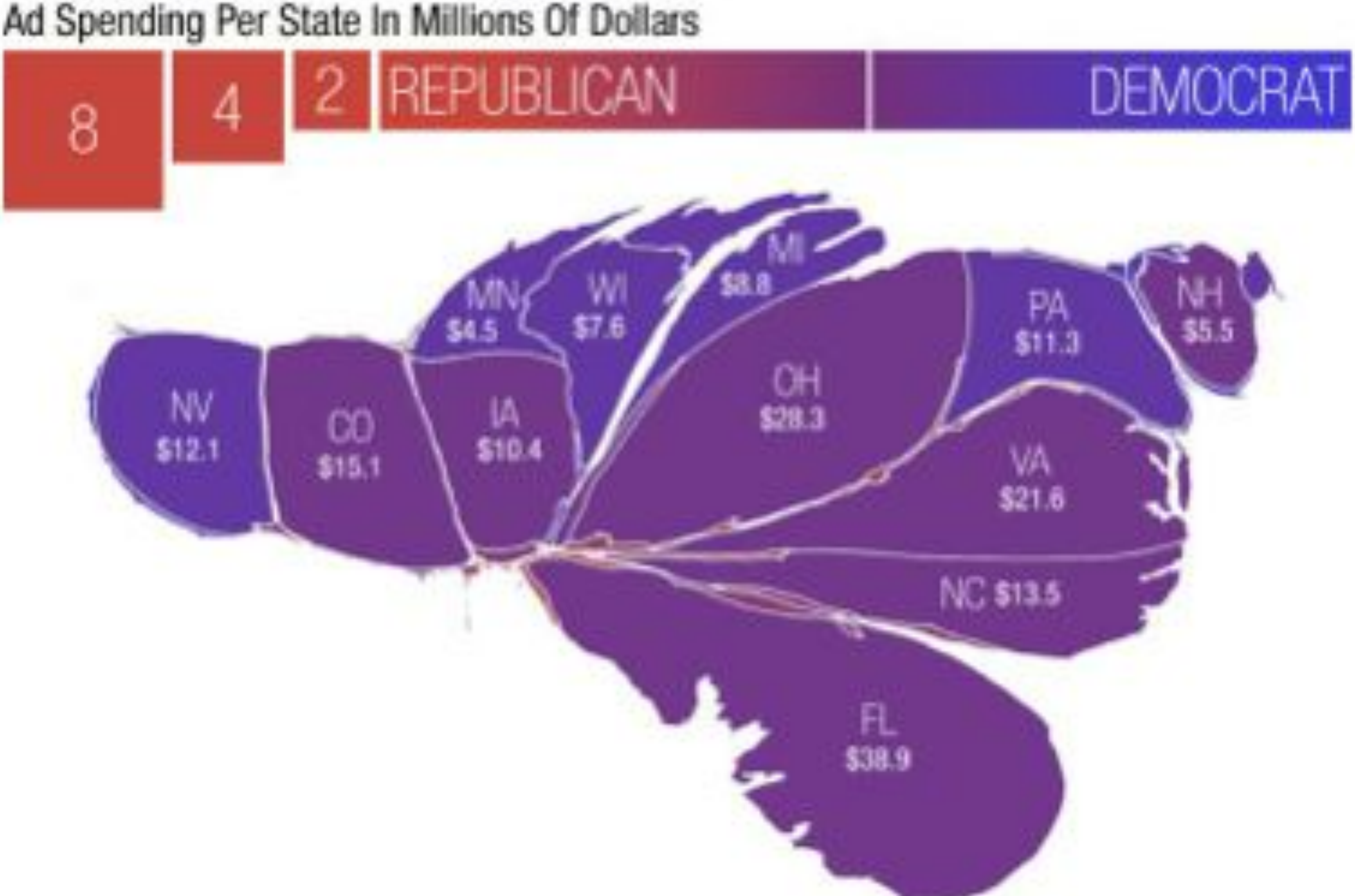
http://gizmodo.com/5960290/this-is-the-real-political-map-of-america-hint-we-are-not-that-divided

2012 Presidential Election



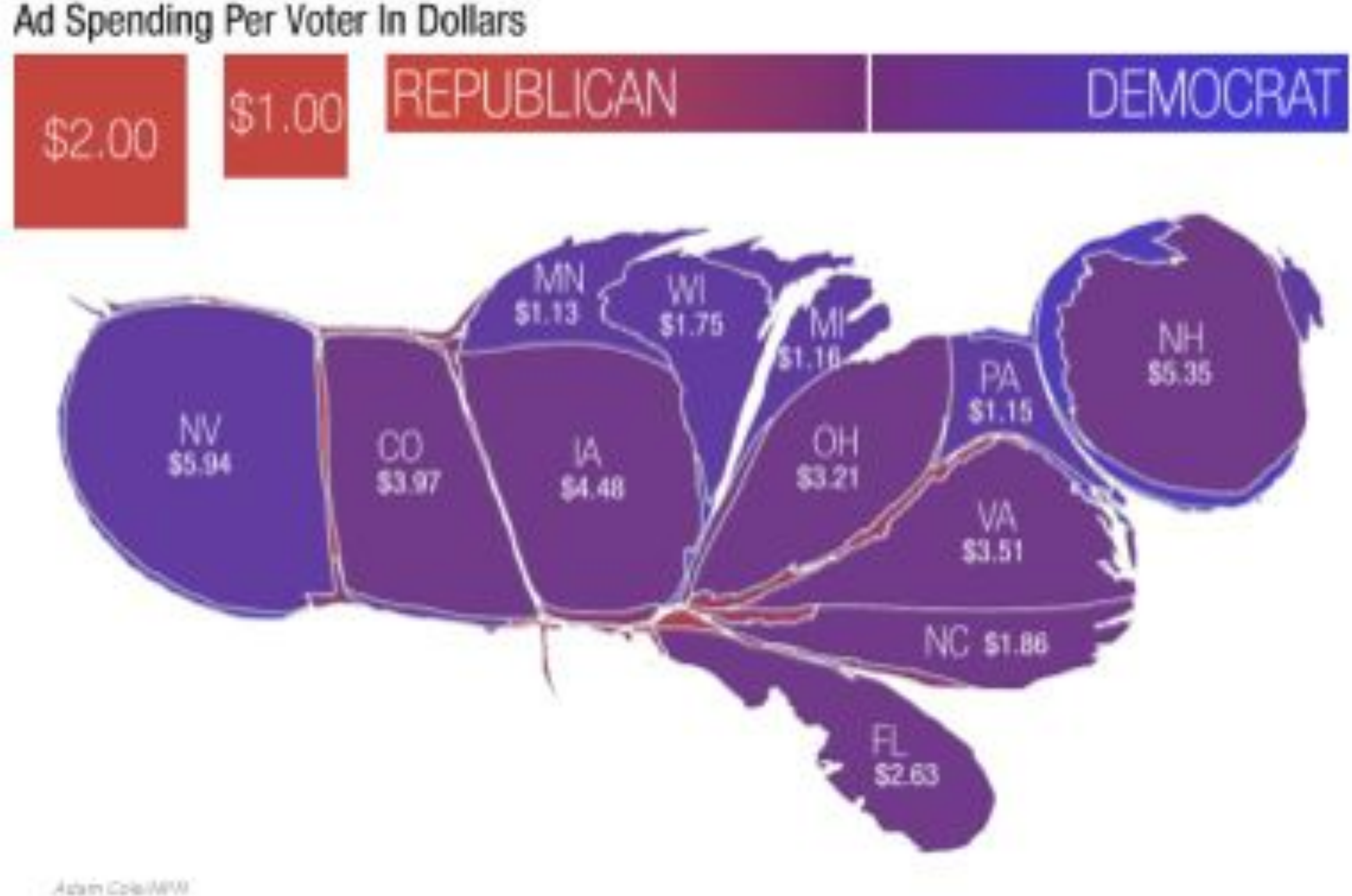
http://www.npr.org/blogs/itsallpolitics/2012/11/01/163632378/a-campaign-map-morphed-by-money

2012 Presidential Election



http://www.npr.org/blogs/itsallpolitics/2012/11/01/163632378/a-campaign-map-morphed-by-money

2012 Presidential Election



http://www.npr.org/blogs/itsallpolitics/2012/11/01/163632378/a-campaign-map-morphed-by-money

Clarifying distortions

Tube map from 1908

http://www.20thcenturylondon.org.uk/beck-henry-harry

http://briankerr.wordpress.com/2009/06/08/connections/
http://en.wikipedia.org/wiki/Harry_Beck

Clarifying distortions

Harry Beck 1933

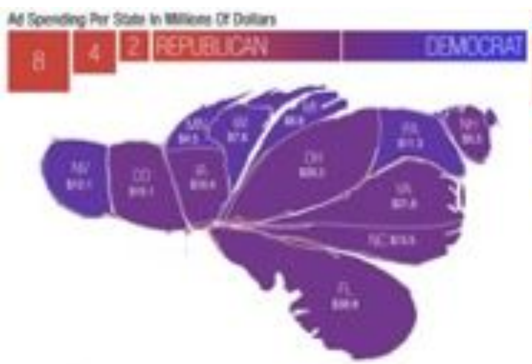
http://briankerr.wordpress.com/2009/06/08/connections/
http://en.wikipedia.org/wiki/Harry_Beck

Clarifying distortions

Clarifying distortions

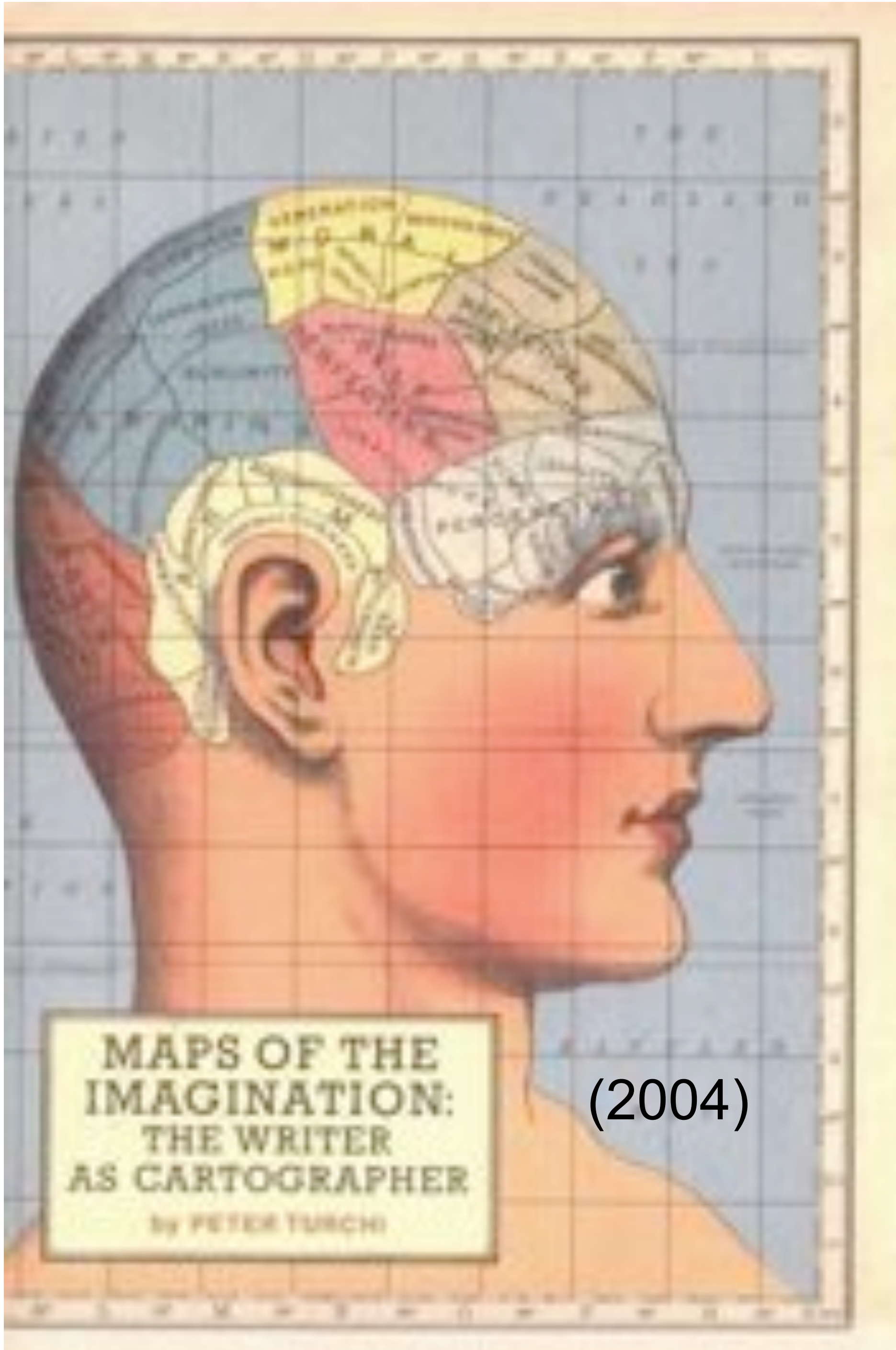
Joachim Böttger, Ulrik Brandes, Oliver Deussen, Hendrik Ziezold,
"**Map Warping for the Annotation of Metro Maps**"
IEEE Computer Graphics and Applications, 28(5):56-65, 2008

Maps reflect conventions, choices, and priorities

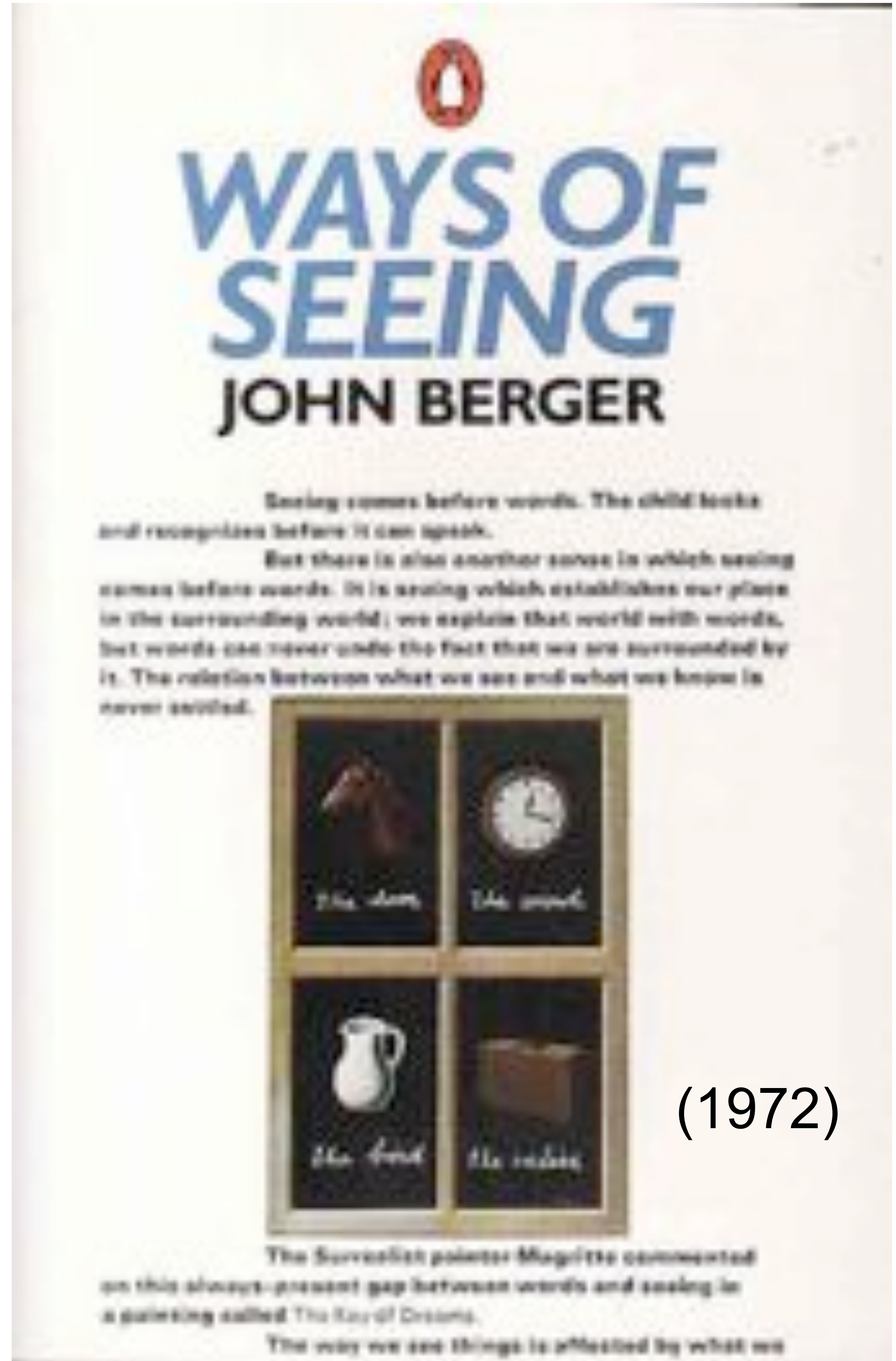


"A single map is but one of an indefinitely large number of maps that might be produced for the same situation or from the same data."
Mark Monmonier "How to Lie with Maps", 1991

Books that have influenced me



(2004)



(1972)

Fields of Visualization

Computer Science
Computer Graphics
Human-computer interaction
Perceptual Psychology

Statistics, Machine Learning

Calculus, Numerical Methods

Information Visualization

Scientific Visualization

Data Visualization

Info-graphics

Scientific Illustration

What is being visualized?

- Data = set of values (or datum) X
  - Spreadsheet: $\{X_i\}_{i=1..N}$; $X_i=(a_i,b_i,c_i,\ldots)$
    coordinates may be spatial or geographical
  - Function of time: X = F(t)
  - Function over 2D X = F(u,v) i.e. an image, or volume F(u,v,w), or 3D surface F(s,t)
  - Graph: X = (Vert,Edge) or (Vert,Arrow)
- Each X is a label or number (or vector of them)
- **Each different type (or flavor) of number has its own mathematical structure: "scales of measurement"**

Scales of measurement

SCIENCE

Vol. 103, No. 2684 Friday, June 7, 1946

On the Theory of Scales of Measurement

S. S. Stevens

Director, Psycho-Acoustic Laboratory, Harvard University

FOR SEVEN YEARS A COMMITTEE of the British Association for the Advancement of Science debated the problem of measurement. Appointed in 1932 to represent Section A (Mathematical and Physical Sciences) and Section J (Psychology), the committee was instructed to consider and report upon the possibility of "quantitative estimates of sensory events"—meaning simply: Is it possible to measure human sensation? Deliberation led only to disagreement, mainly about what is meant by the term measurement. An interim report in 1938 found one member complaining that his colleagues

by the formal (mathematical) properties of the scales. Furthermore—and this is of great concern to several of the sciences—the statistical manipulations that can legitimately be applied to empirical data depend upon the type of scale against which the data are ordered.

A CLASSIFICATION OF SCALES OF MEASUREMENT

Paraphrasing N. R. Campbell (Final Report, p. 340), we may say that measurement, in the broadest sense, is defined as the assignment of numerals to objects or events according to rules. The fact that numerals can be assigned under different rules leads

Stevens' 4 scales of measurements

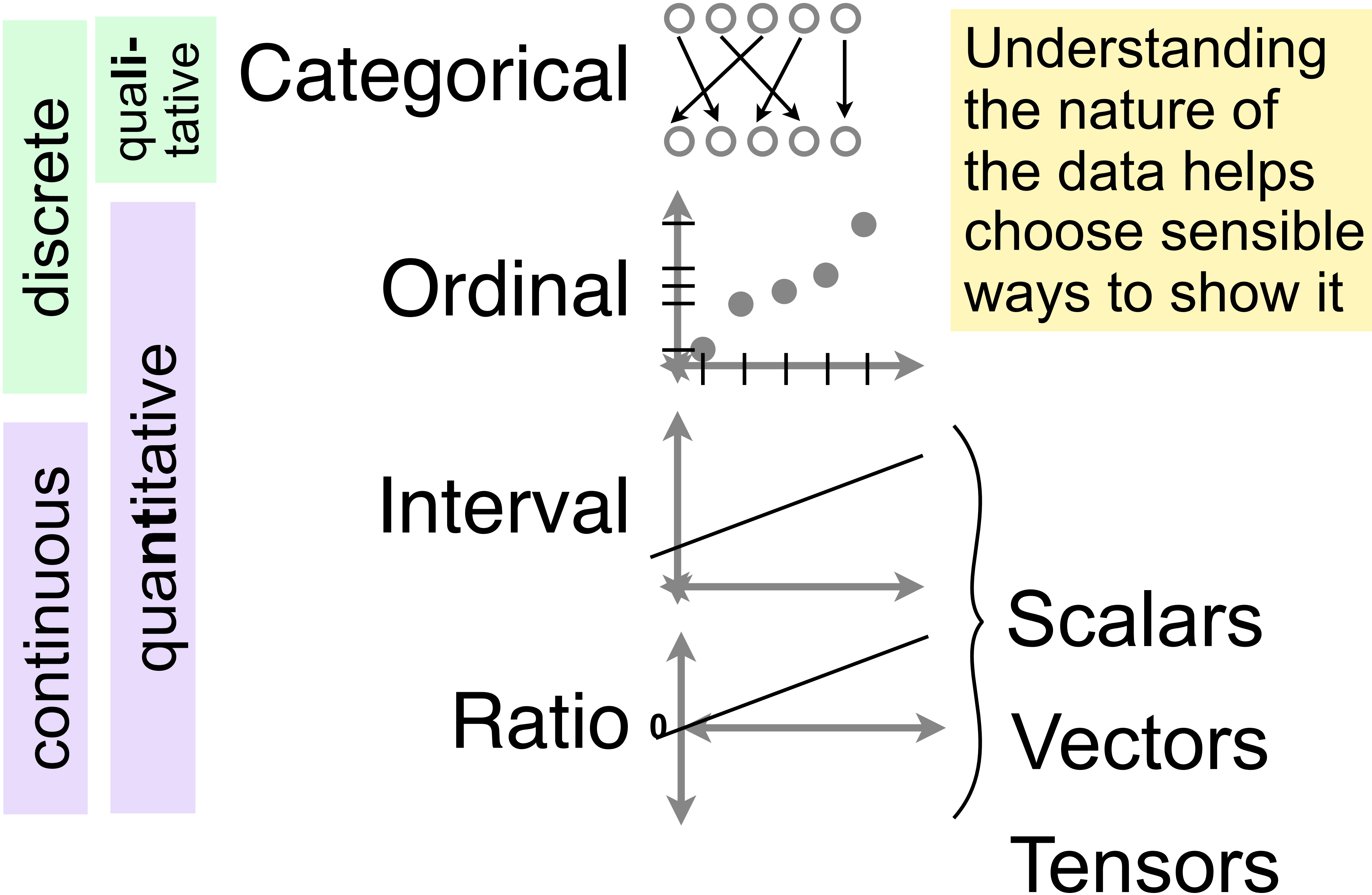
http://en.wikipedia.org/wiki/Level_of_measurement

Scale	Basic Empirical Operations	Mathematical Group Structure	Permissible Statistics (invariantive)
Nominal Categorical Qualitative	Determination of equality	*Permutation group* $x' = f(x)$ $f(x)$ means any one-to-one substitution	Number of cases Mode Contingency correlation
Ordinal	Determination of greater or less	*Isotonic group* $x' = f(x)$ $f(x)$ means any monotonic increasing function	Median Percentiles
Interval	Determination of equality of intervals or differences	*General linear group* $x' = ax + b$	Mean Standard deviation Rank-order correlation Product-moment correlation
Ratio	Determination of equality of ratios	*Similarity group* $x' = ax$	Coefficient of variation

Later scales specialize earlier scales

Some examples of these ...

The structure of data values



Symmetries and visualization

Symmetry of data: D = T(D)

- nominal: permutation
- ordinal: any monotonic function
- interval: affine transform: $y = ax + b$ $(a > 0)$
- ratio: $y = ax$ $(a > 0)$

Visualization: V(D)

- V maps from data D to image or image property
- **colormaps**: assignment of colors to values
- We will look at multiple examples of these

Keep in mind: V=V(D,P) for parameter set P

Symmetries and visualization

Basic idea: Visualization should reflect the symmetries and preserve structure in data

$$\mathbf{D = T(D) \Leftrightarrow V(D) \approx V(T(D))}$$

$$\mathbf{-D = T(D) \Leftrightarrow -V(D) \approx V(T(D))}$$

$$\mathbf{D_1 = D_2 \Leftrightarrow V(D_1) = V(D_2)}$$

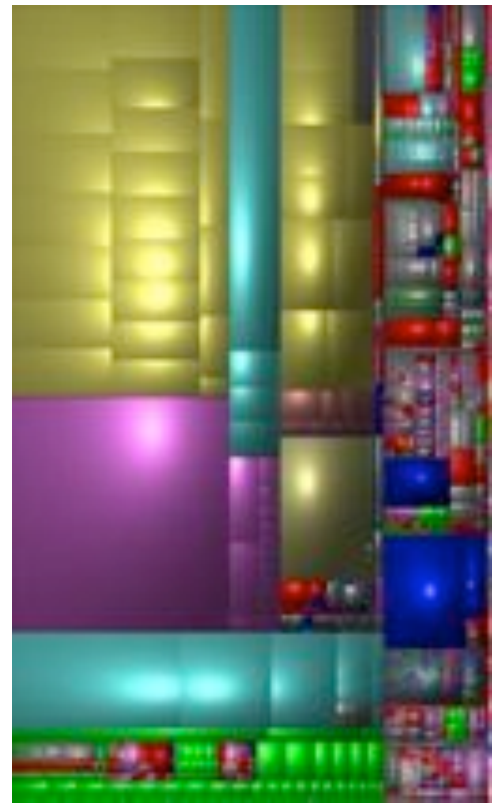
$$\mathbf{D_1 < D_2 \Leftrightarrow V(D_1) < V(D_2)}$$

Applies to computational representation of measurements, and to other quantities:

Debugging by symmetry

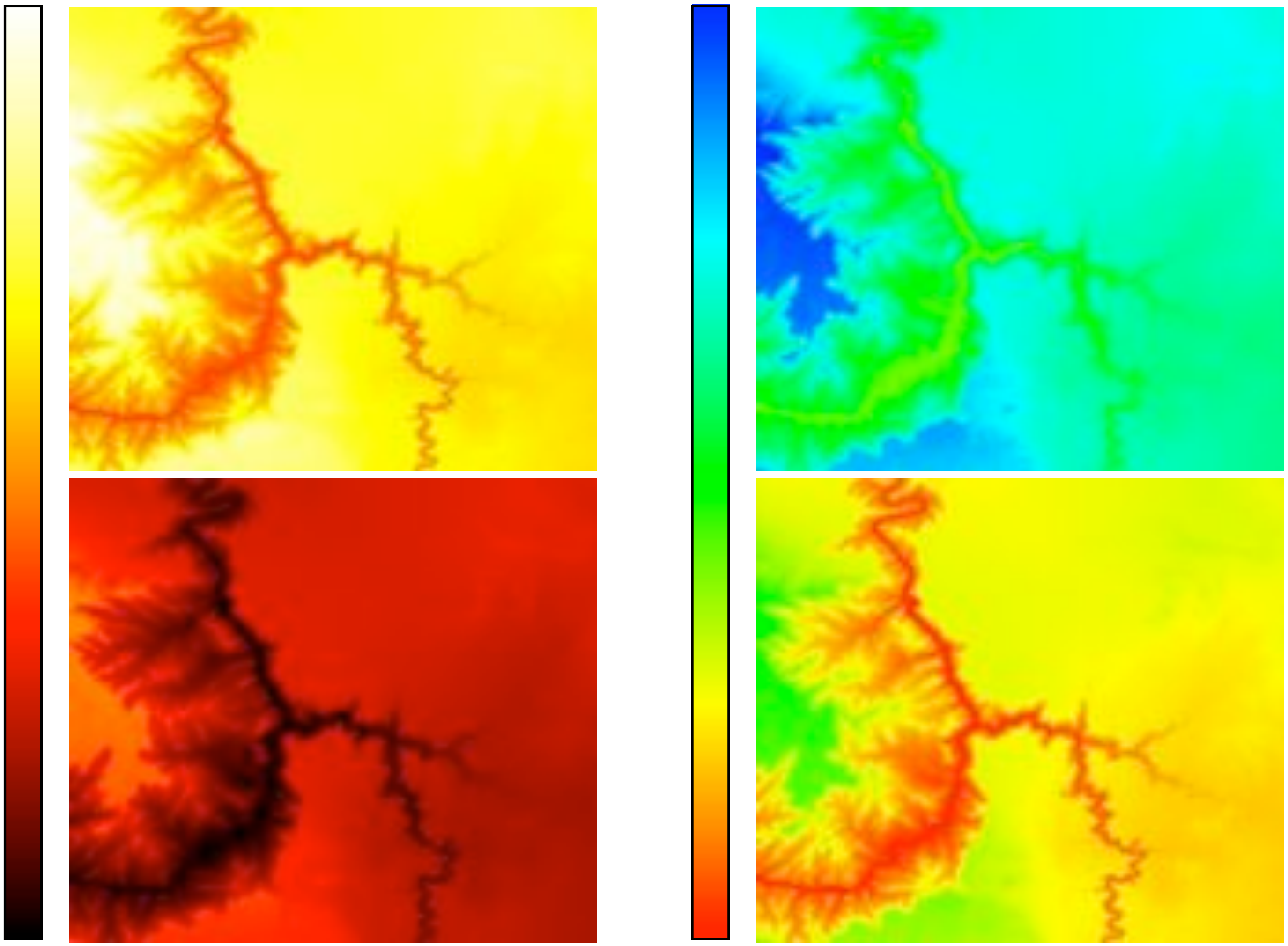
Colormap for categorical data

Coloring different types of files (MP3, JPG, TIFF, applications, etc)



Coloring (or effectiveness of coloring) of categorical values should be symmetric under permutation

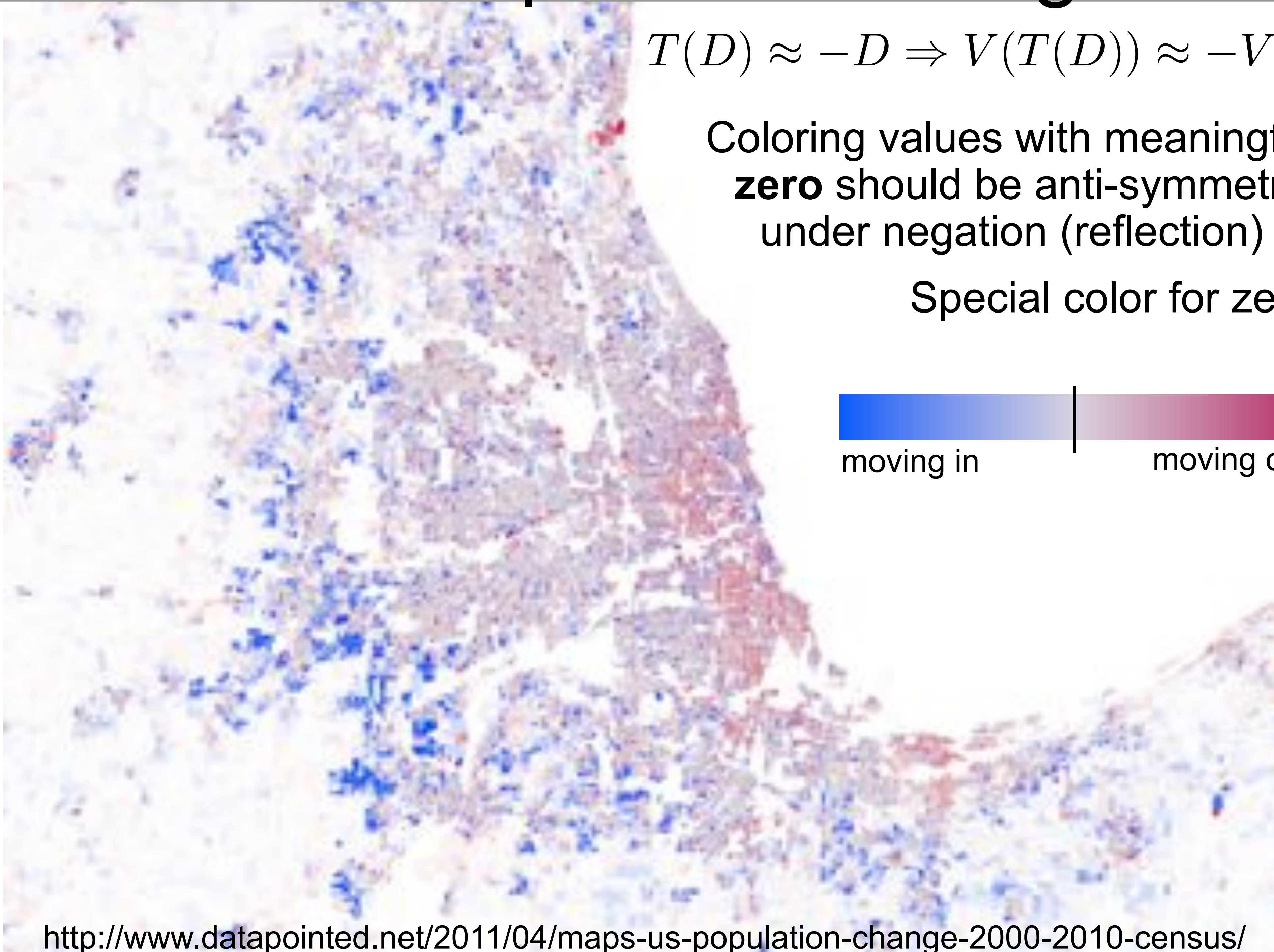
Colormap for interval data



$$V(D+c) \approx V(D) \quad V(D+c) \neq V(D)$$

Coloring of **interval** values should be symmetric under addition of constant (can convey intrinsic **ordering**)

Ratio data: Population change



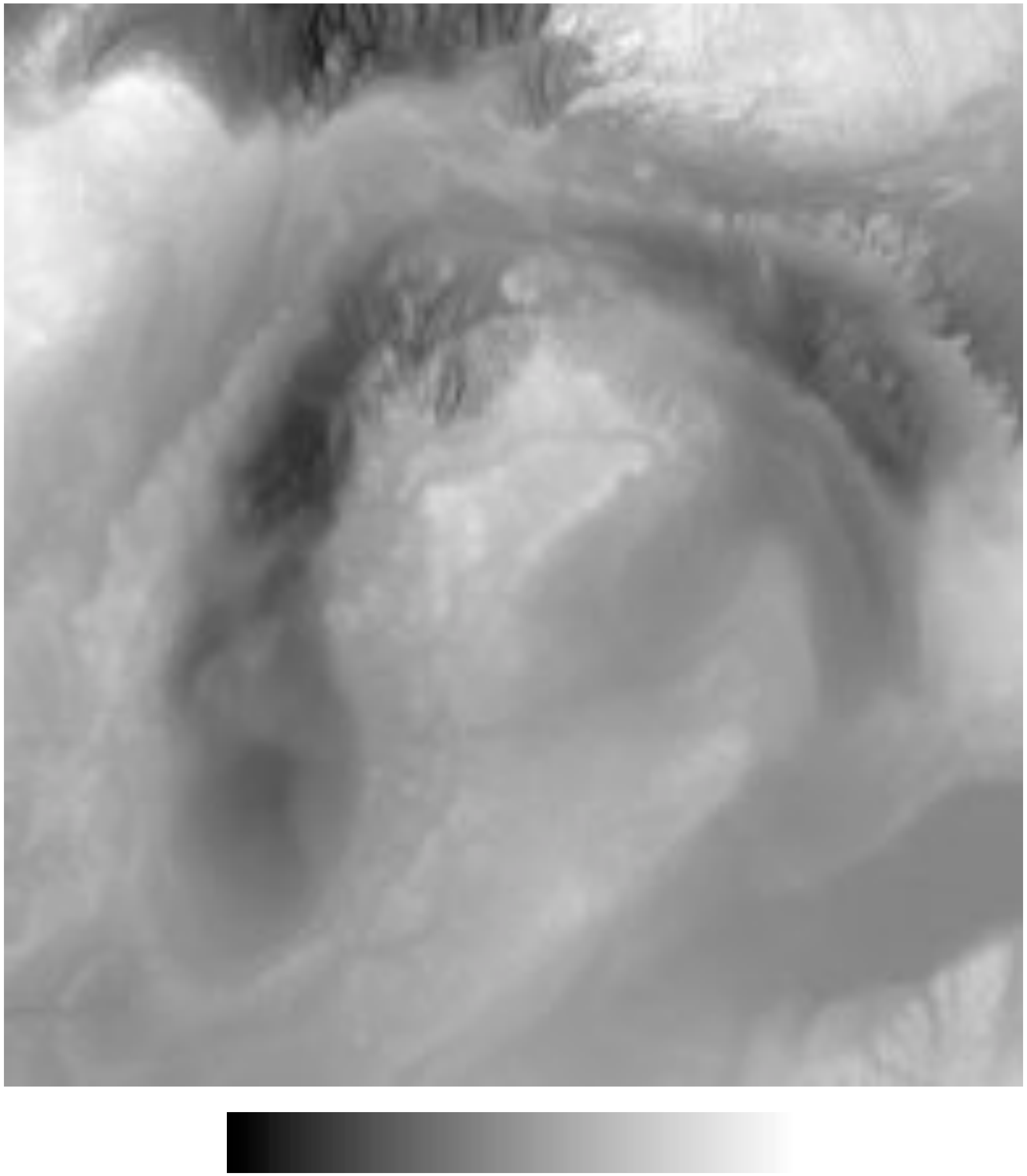
$$T(D) \approx -D \Rightarrow V(T(D)) \approx -V(D)$$

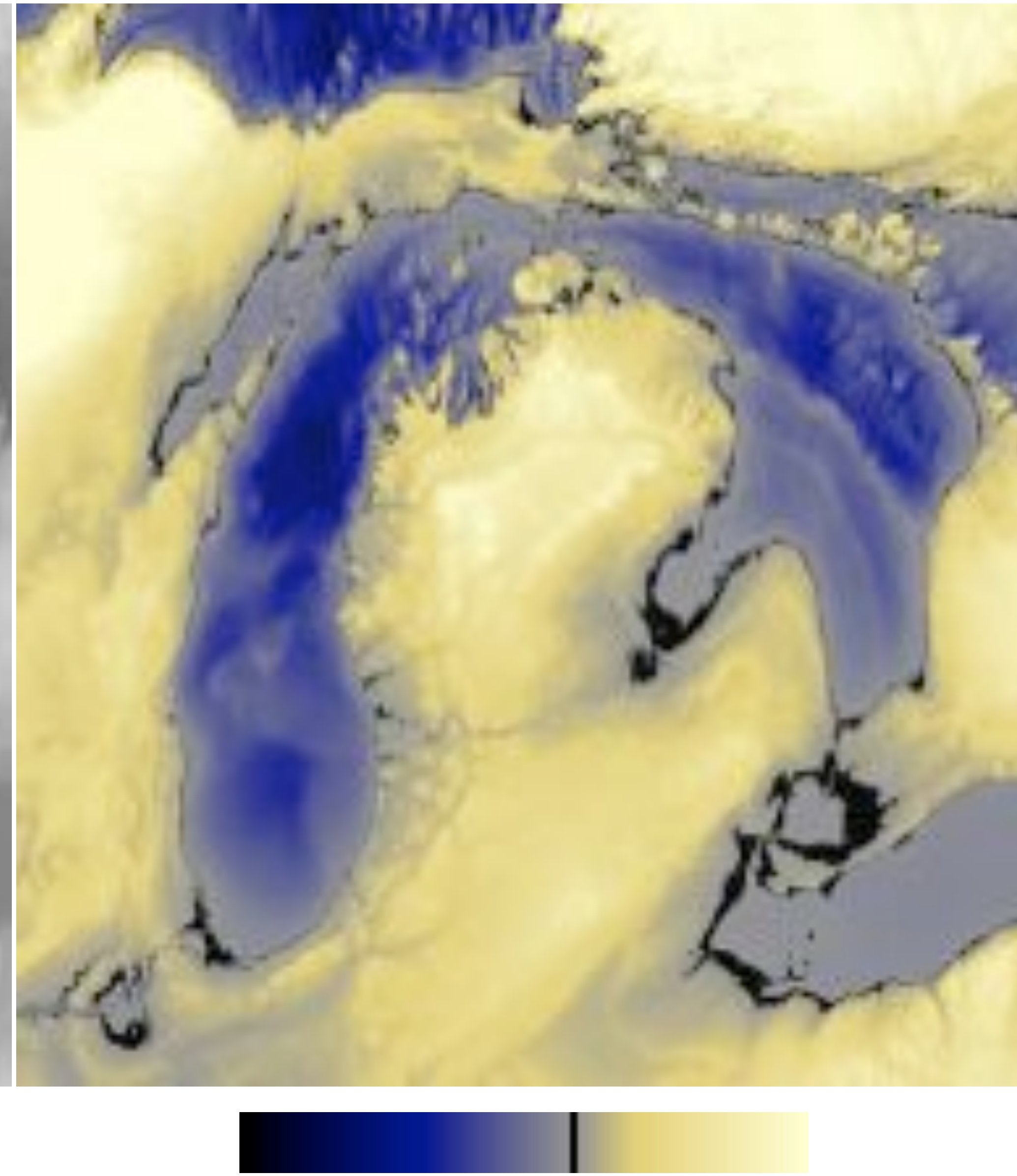
Coloring values with meaningful **zero** should be anti-symmetric under negation (reflection) ⇒

Special color for zero

http://www.datapointed.net/2011/04/maps-us-population-change-2000-2010-census/

Value of showing isocontours



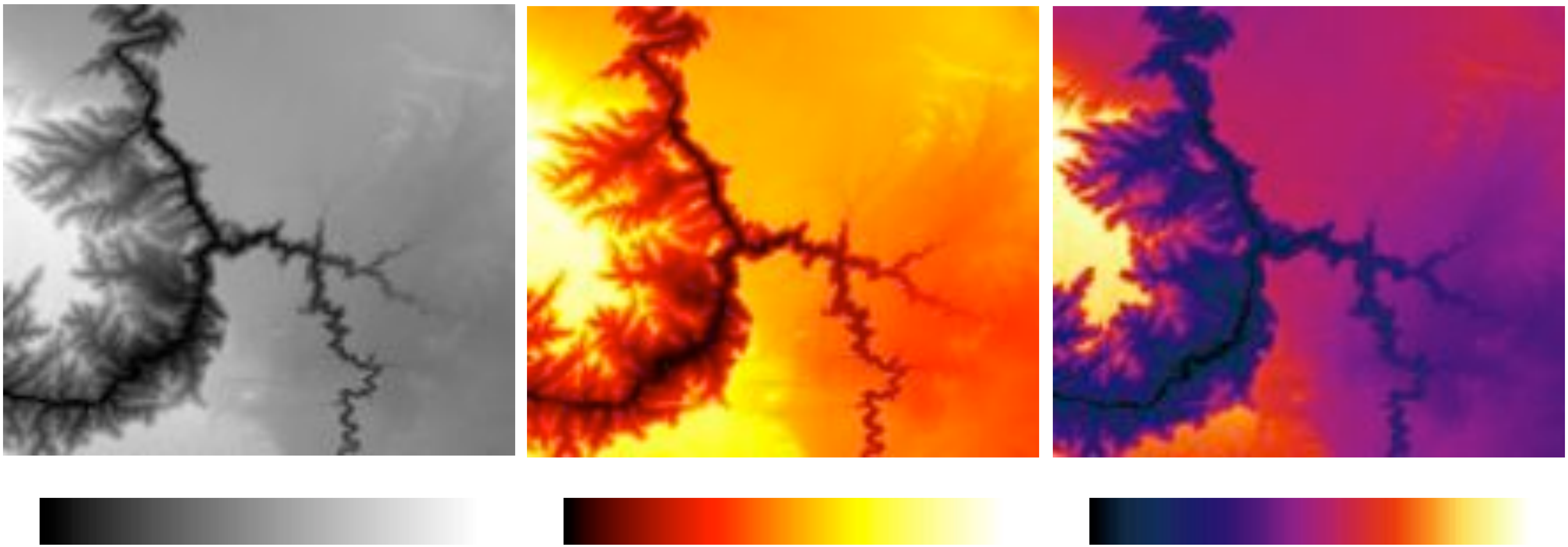


What colormap to use?

- Data symmetries can inform colormap choices
- Should know which “scale” (of 4) your data is
  - Is there ordering? (is the order crucial?)
  - Is there a zero? (is showing sign of values crucial?)
- Two common tasks
  - Picking off values at particular locations
    - (in nominal data: identifying category)
  - Showing over-all form: depends on **ordering**
  - Or some balance of both
- What attributes of color are ordered?

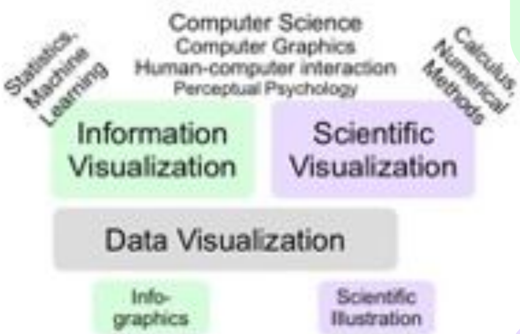
Ordering of color axes

Strategy: combine hue and luminance



- Shows basic height relationships
  - And helps for distant comparisons
- Common technique in scientific visualization
- Drawback: luminance not free for shading

Fields of Visualization related to data



Information Visualization

$\{X_i\}_{i=1..N};\ X_i=(a_i,b_i,c_i,\ldots)$

$X = F(t)$

$X = (Vert, Edge)$

$X = F(u,v)$

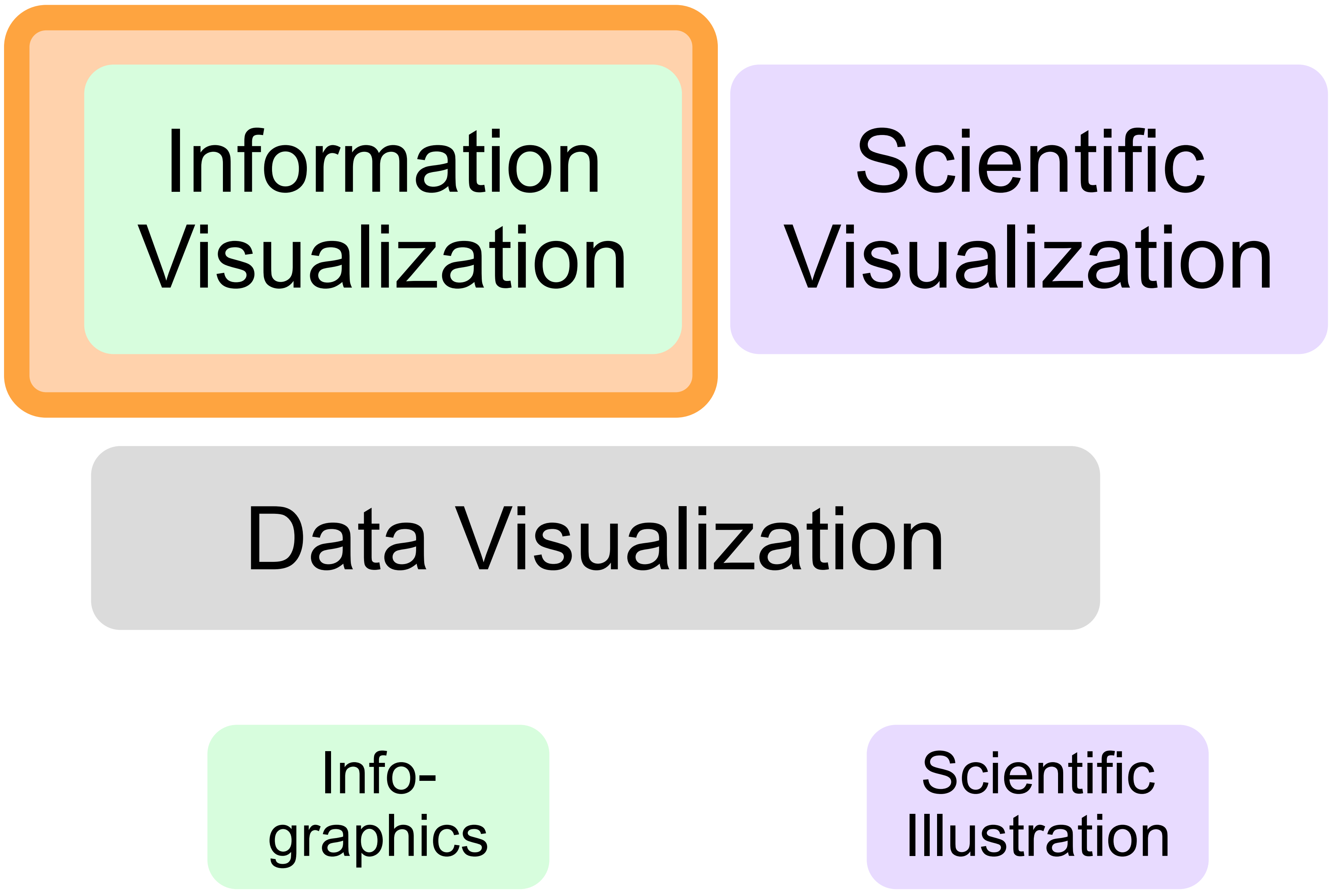
Scientific Visualization

$X = F(u,v,w)$

$X = F(s,t)$ (3D surface)

X : vectors, tensors

Fields of Visualization

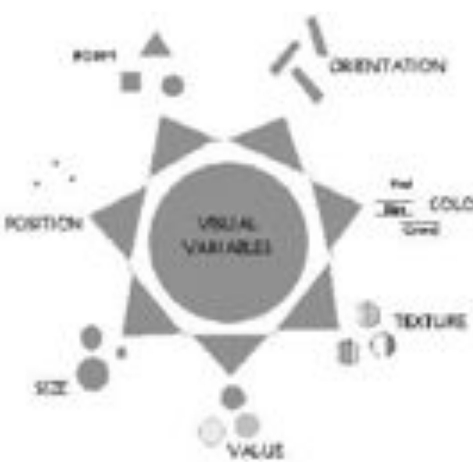


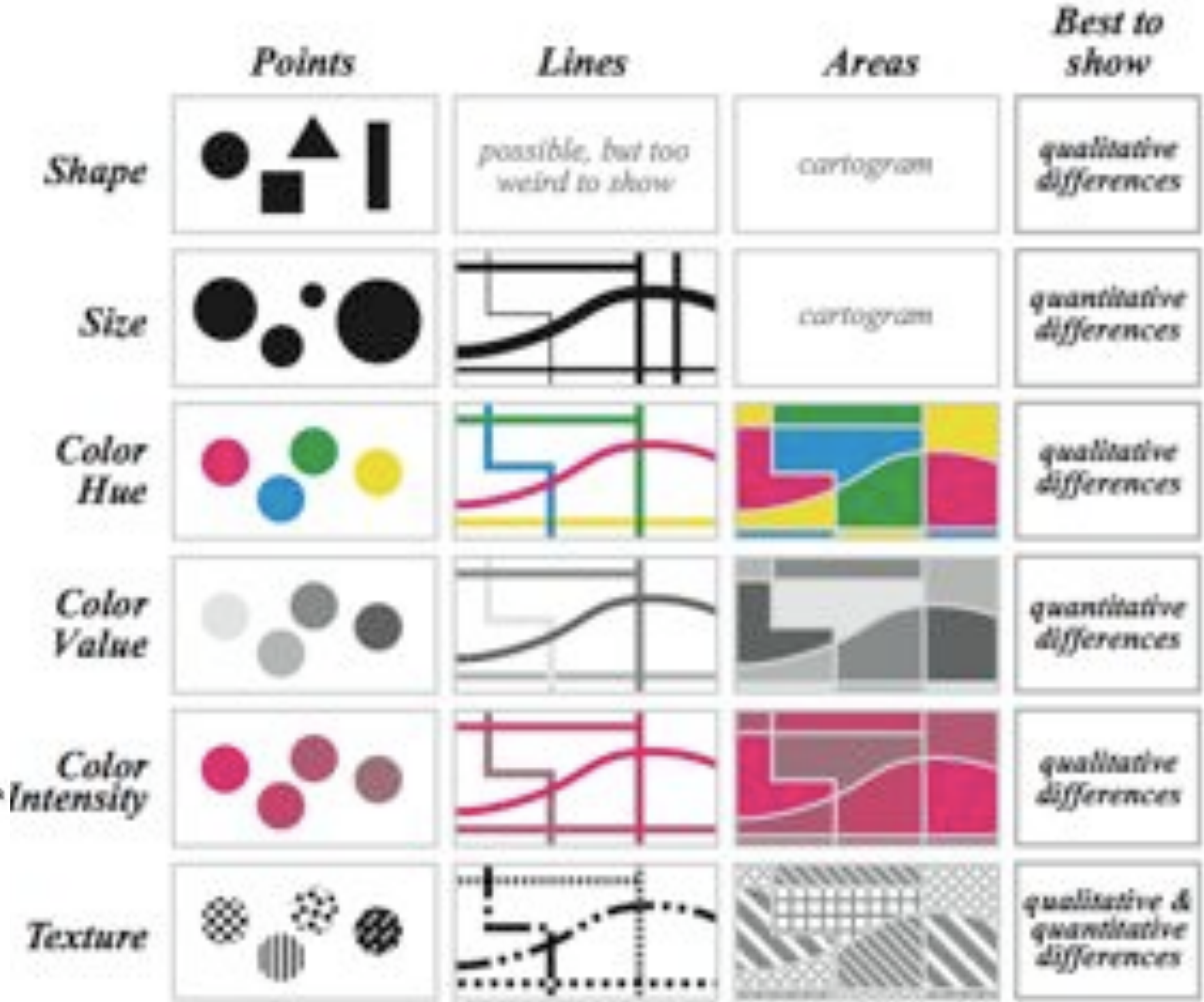
How to visually encode information?

Jacques Bertin: French cartographer

1918 – 2010

“Semiology of Graphics”

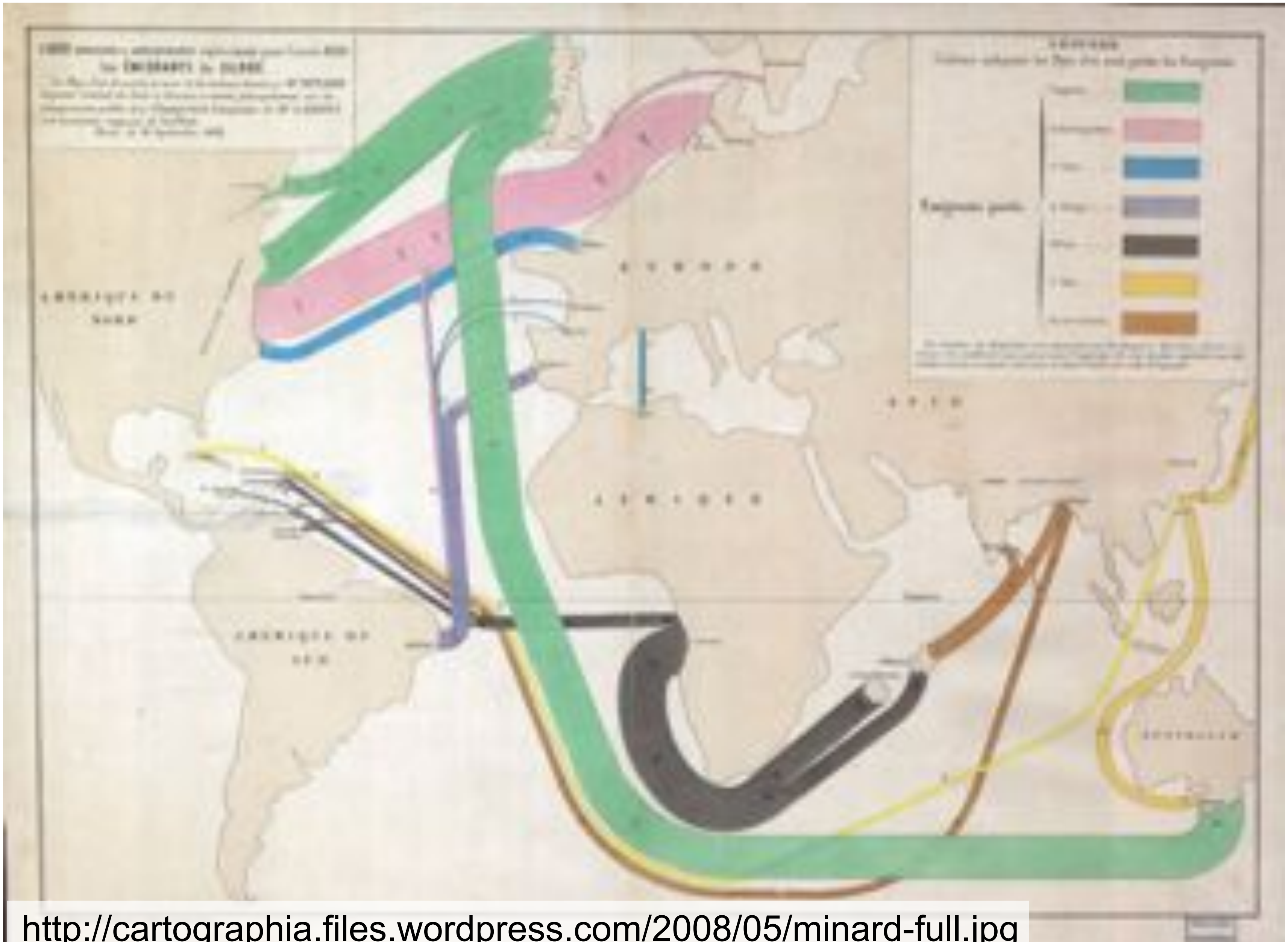




http://understandinggraphics.com/visualizations/information-display-tips/

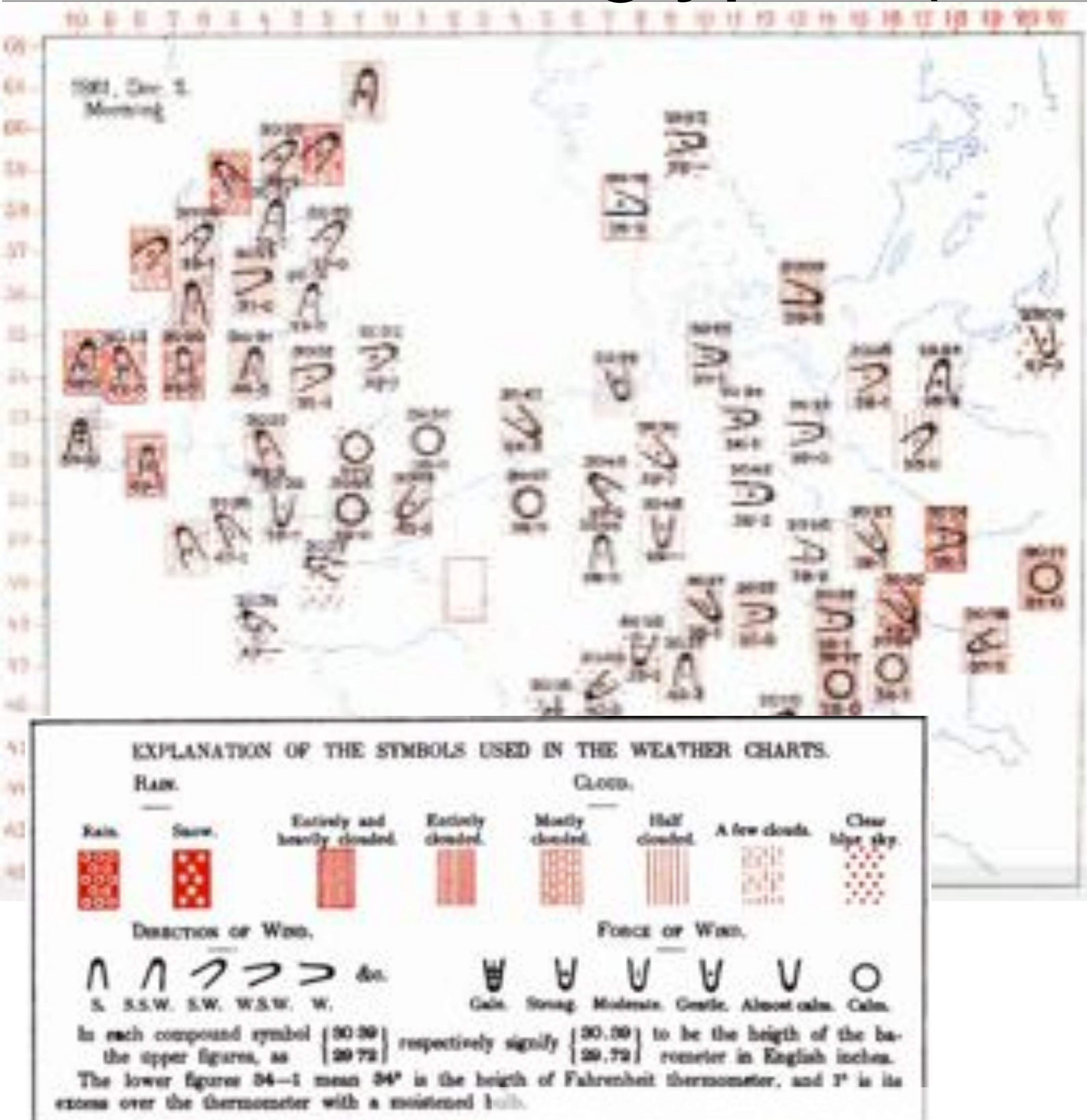
Charles Joseph Minard (1781 – 1870)

Amount and destinations of emigrants from Europe, Africa, China, and South Asia for the year 1858



http://cartographia.files.wordpress.com/2008/05/minard-full.jpg

Multi-variate glyphs (or icon)



Francis Galton
(1822 –1911)

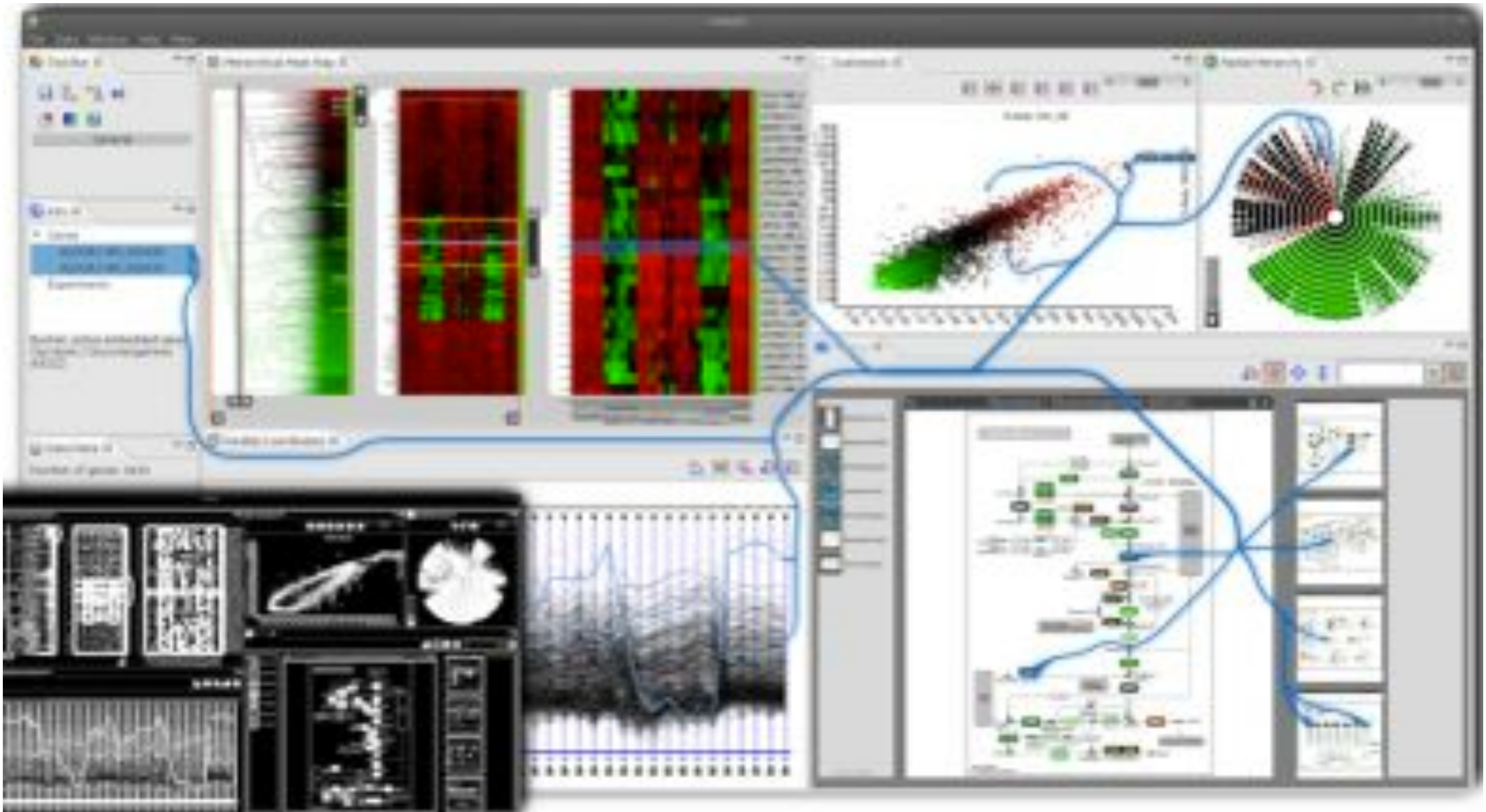
Glyphs shows wind force and direction, cloud, precipitation

Aided in discovery of anti-cyclones

Learn more at Michael Friendly's "Milestones Project": http://www.datavis.ca/milestones/

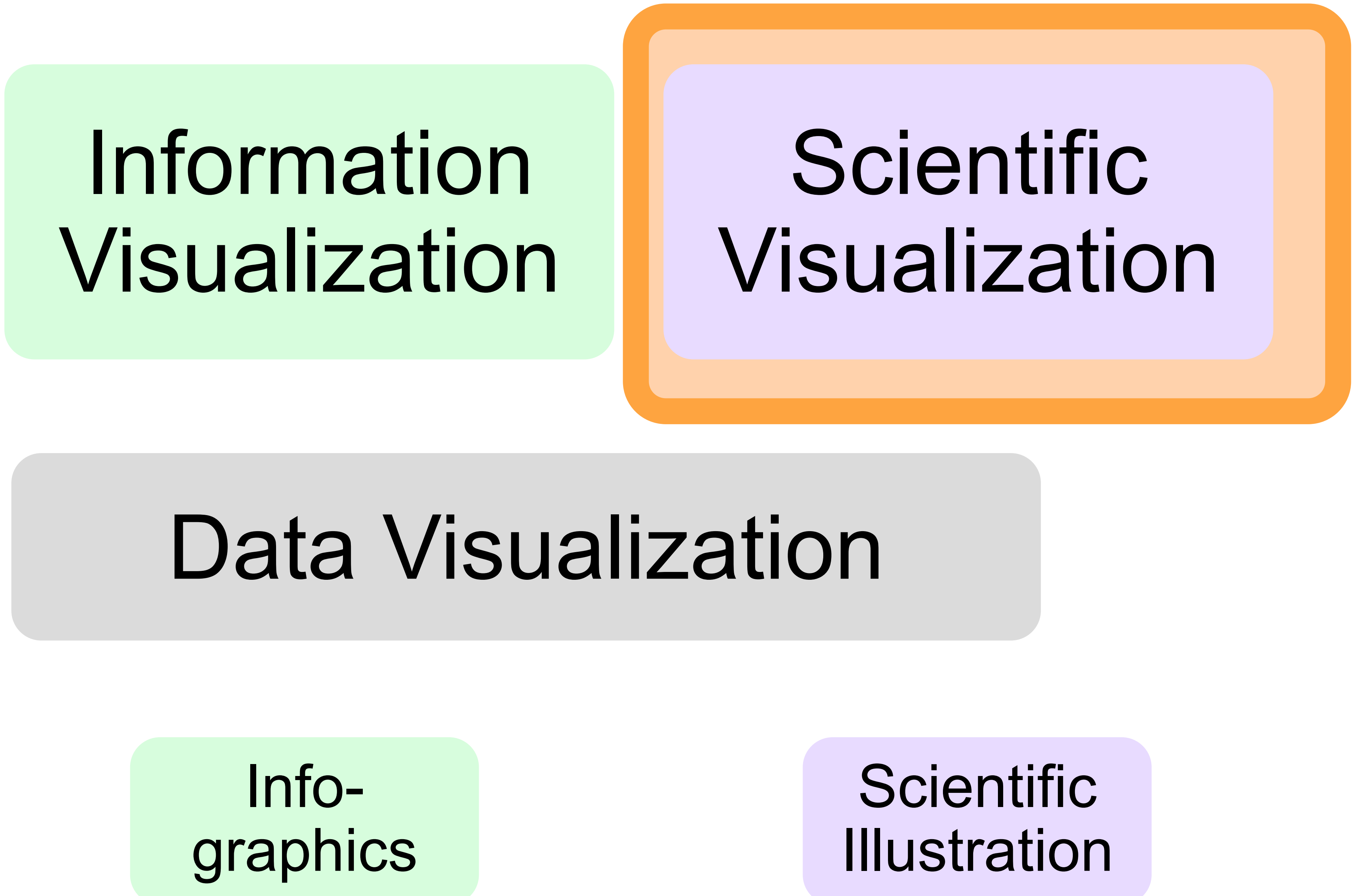
images from:
http://books.google.com/books?id=cFznAAAAMAAJ&pg=PA82&lpg=PA82&dq=%22explanation+of+the+symbols+used+in+the+weather+charts%22&source=bl&ots=pweJga3MXr&sig=xWXxTpCM9efg_7Y8XAnnKKIMRcs&hl=en&sa=X&ei=Ab-rUOixJMn8rQH31IHQDA&ved=0CCsQ6AEwAA#v=onepage&q=%22explanation%20of%20the%20symbols%20used%20in%20the%20weather%20charts%22&f=false

Linking multiple views



Markus Steinberger, Manuela Waldner, Marc Streit, Alexander Lex, and Dieter Schmalstieg. **Context-Preserving Visual Links.** IEEE Transactions on Visualization and Computer Graphics 17(12):2249-2258, 2011

Fields of Visualization



Main Data types

Scalar (e.g. Temperature, Pressure, Density, Intensity)

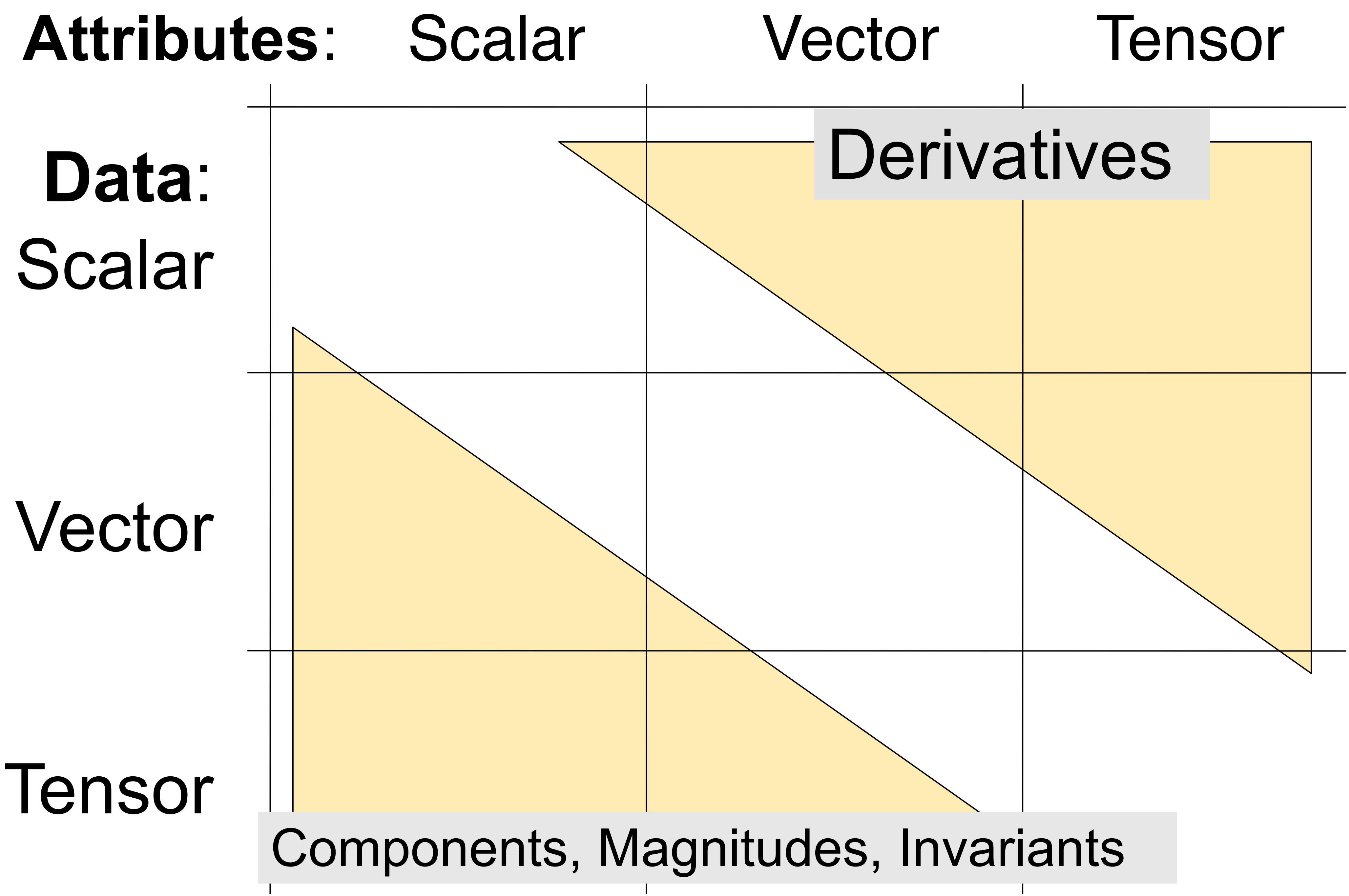
Vector (e.g. Velocity of flow, displacement in deformation)

Tensor (e.g. Jacobian of flow, Diffusion, Stress)

Data types vs attributes

Data: \ **Attributes**:	Scalar	Vector	Tensor
Scalar			
Vector			
Tensor			

Data types vs attributes



Data types vs attributes

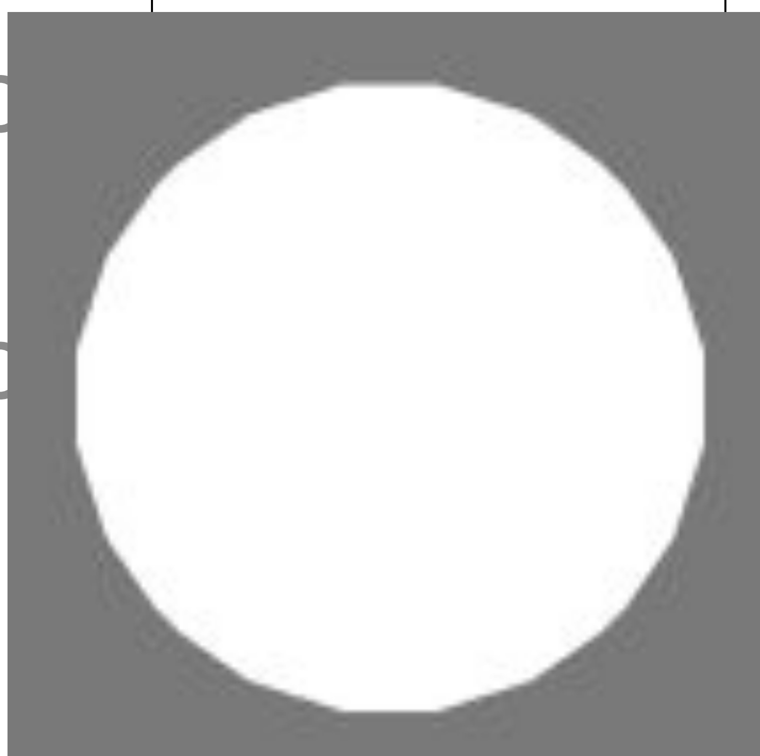
Data: \ **Attributes**:	Scalar	Vector	Tensor
Scalar		Gradient	Hessian, curvature
Vector	Magnitude (e.g. speed)		Jacobian
Tensor	Eigenvalues, Invariants	Eigenvectors	

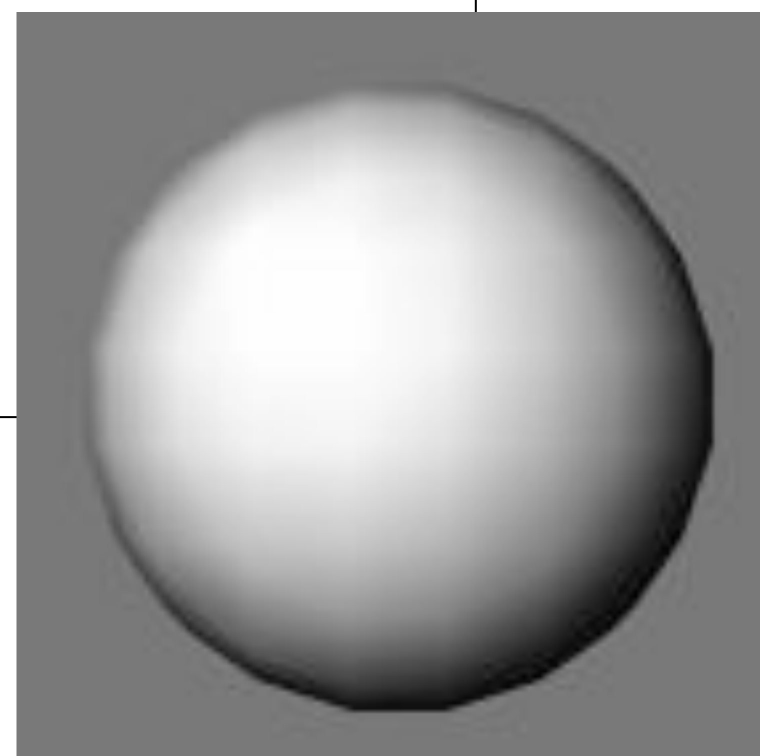
Derivatives

Components, Magnitudes, Invariants

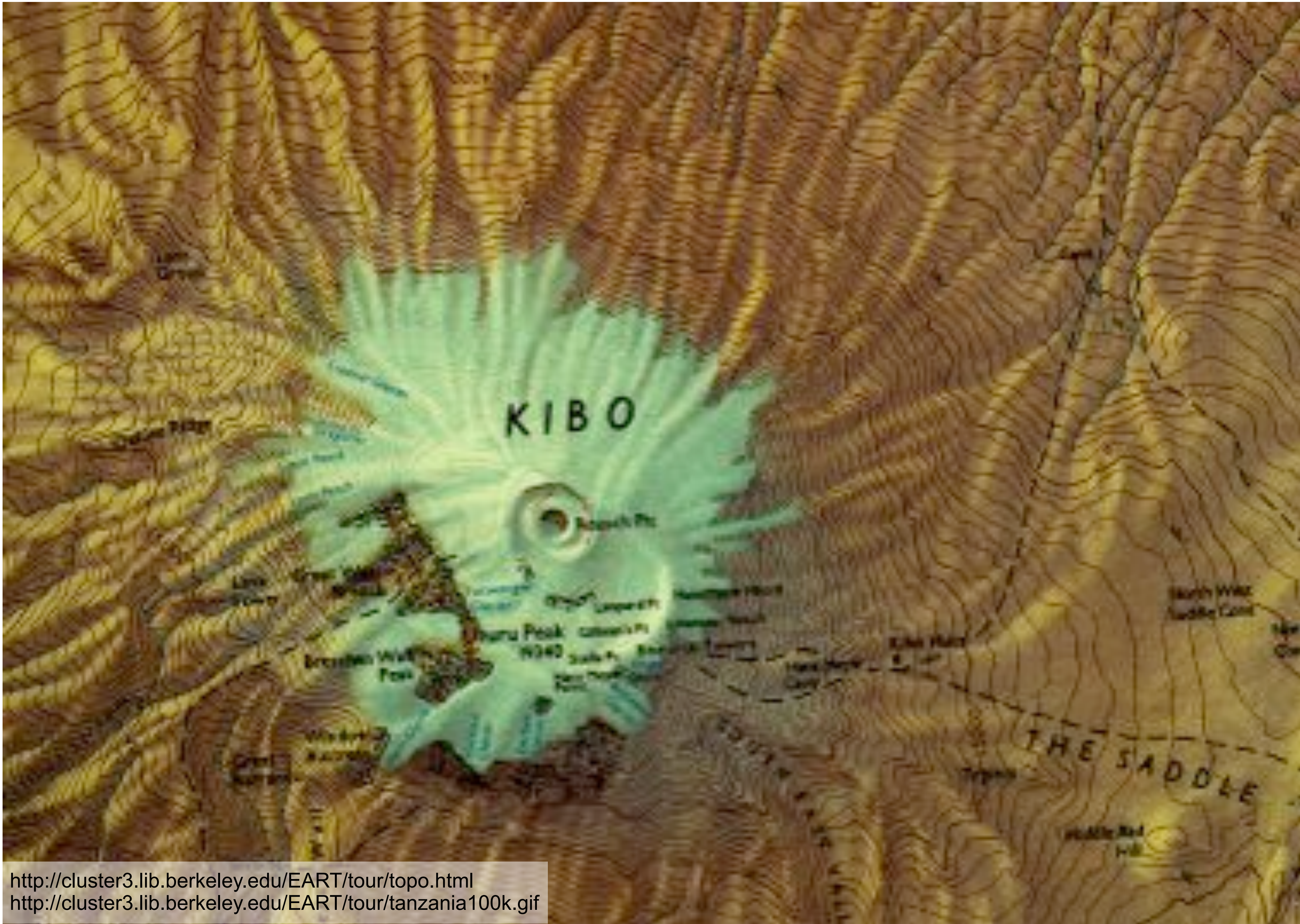
Simple example

Attributes:	Scalar	Vector	Tensor
Data: Scalar	Data itself: Isocontours	Gradient: Surface shading	
Vecto			
Tenso			



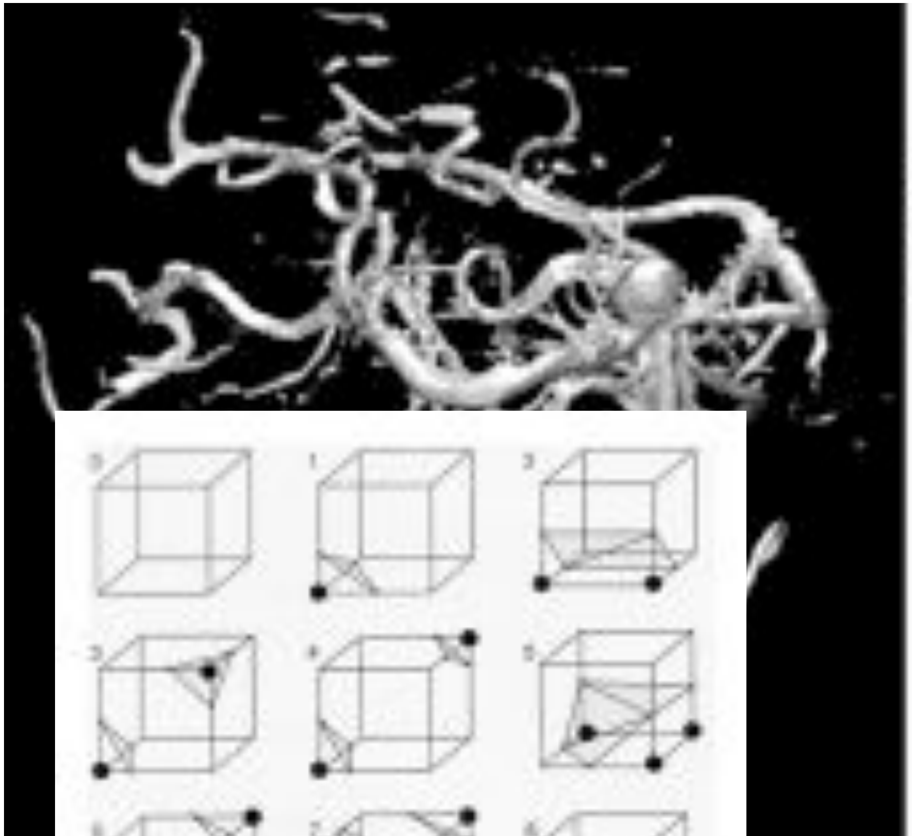


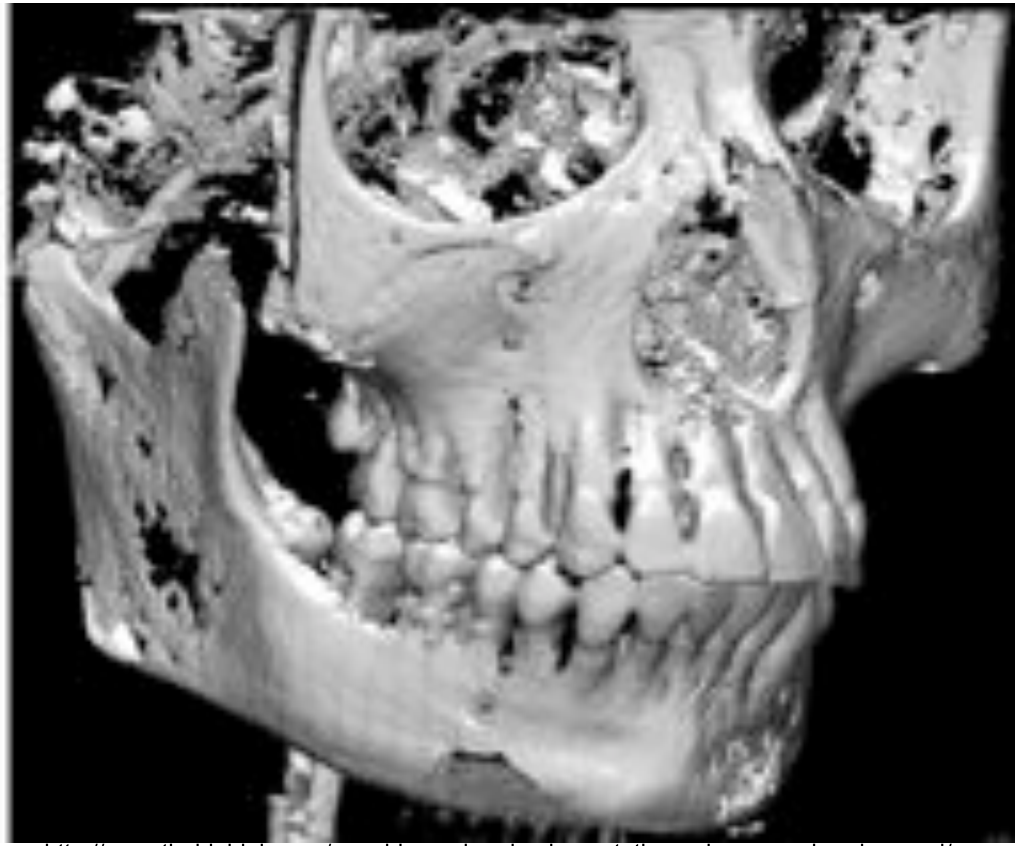
Isocontours on a topo map

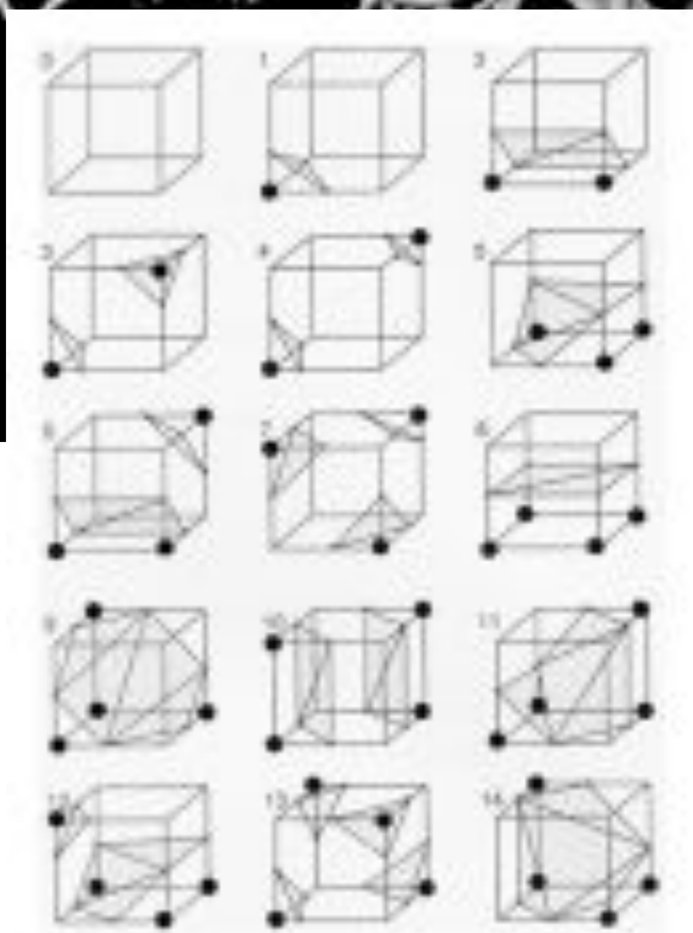


http://cluster3.lib.berkeley.edu/EART/tour/topo.html
http://cluster3.lib.berkeley.edu/EART/tour/tanzania100k.gif

3D Isocontours



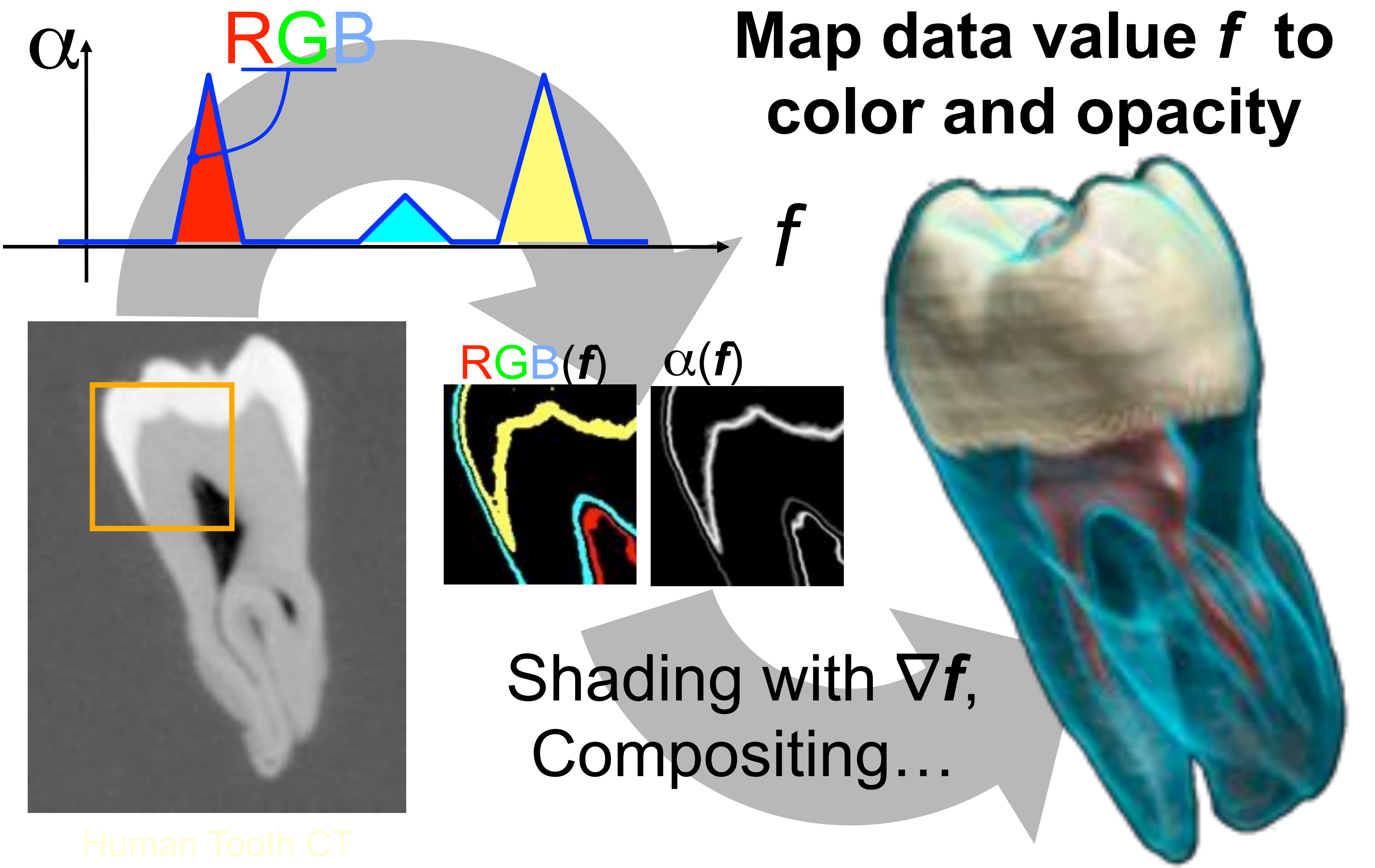




http://www.thebigblob.com/marching-cubes-implementation-using-opencl-and-opengl/

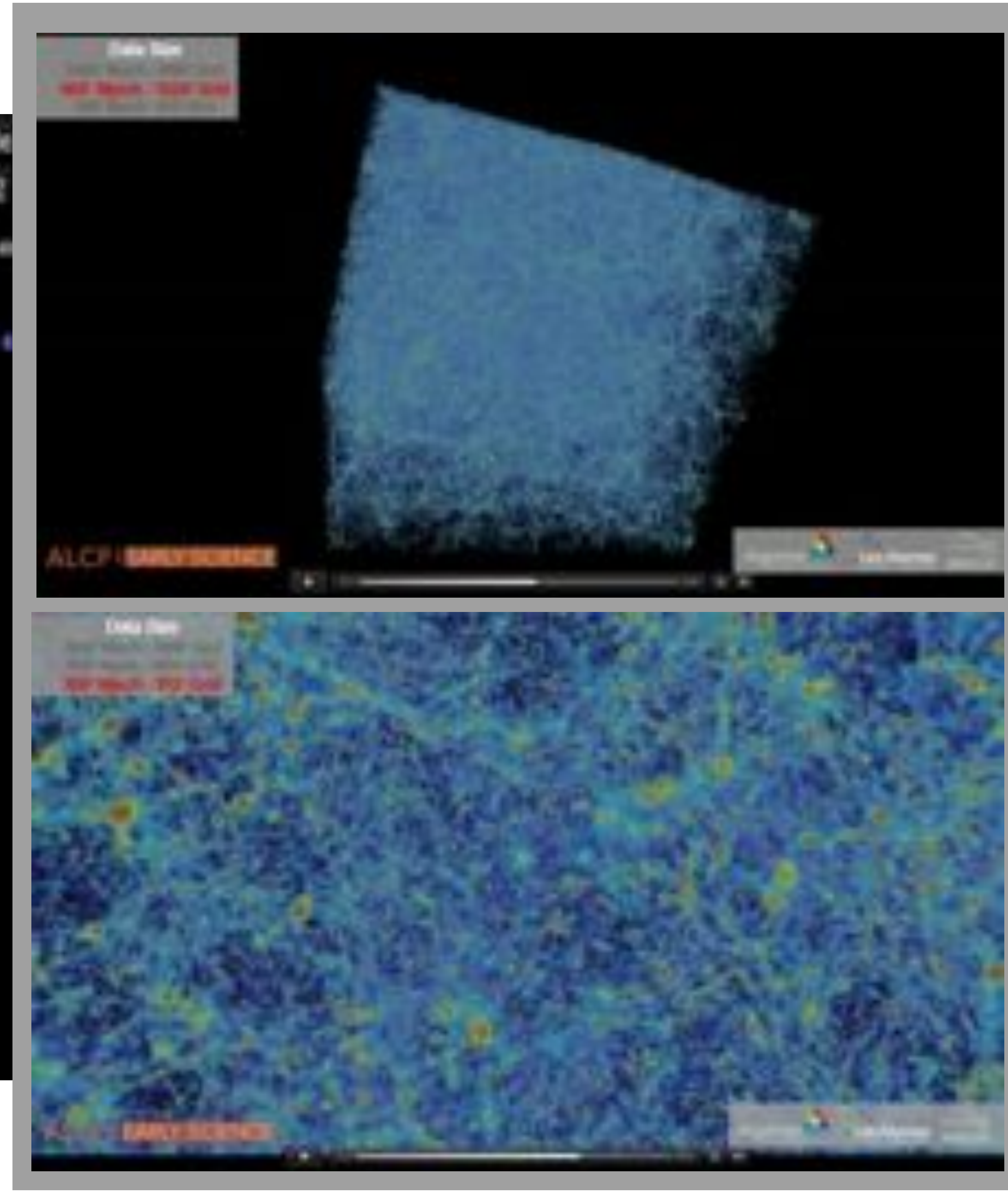
William E. Lorensen and William E. Cline, Marching Cubes: **A High Resolution 3D Surface Construction Algorithm.** Computer Graphics, 1987, 21, 163-169

Direct Volume Rendering



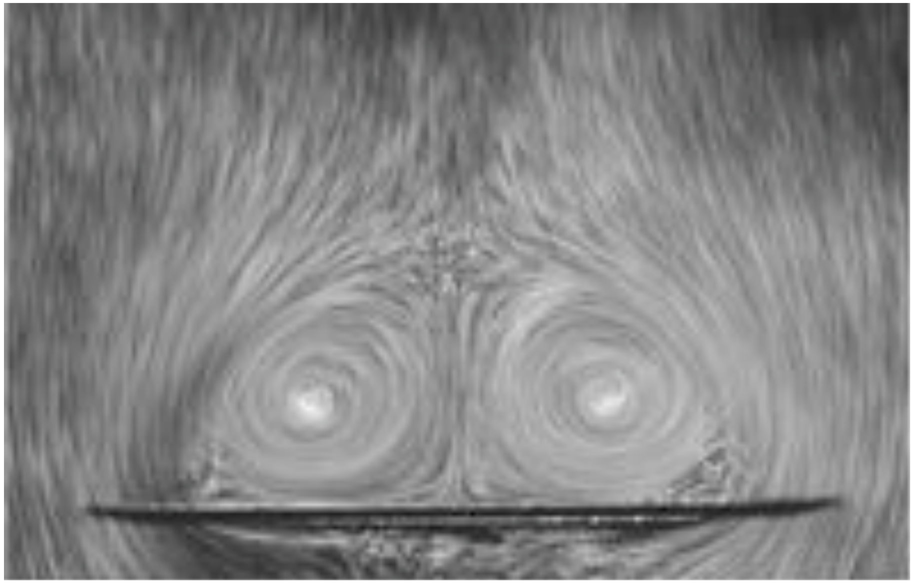
Volume rendering applications

https://commons.wikimedia.org/wiki/
File:Fetal_yawning_4D_ultrasound_ecografia_4D_
Dr._Wolfgang_Moroder.theora.ogv



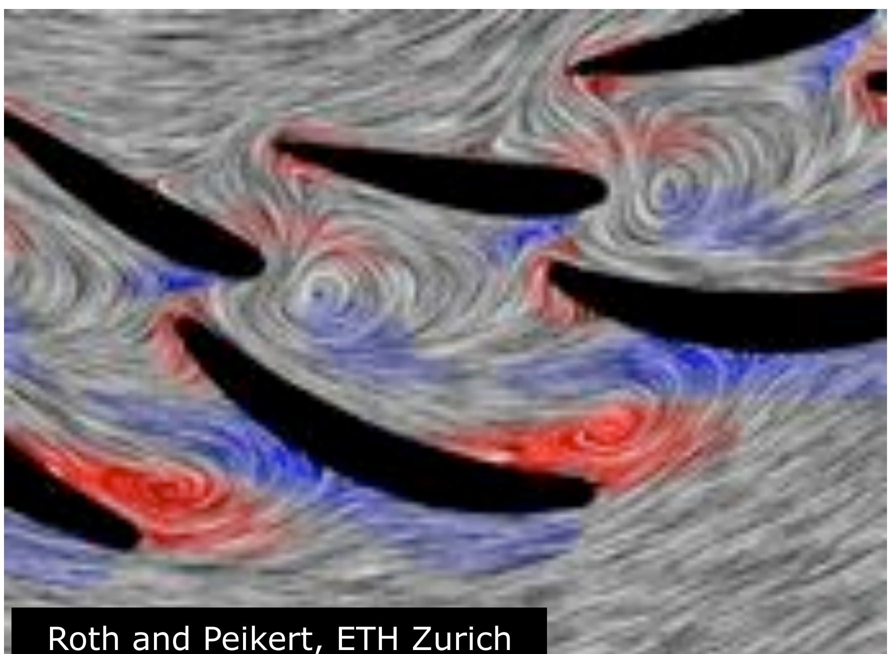
http://www.popsci.com/technology/article/2012-11/video-largest-most-hi-res-cosmological-simulations-known-universe

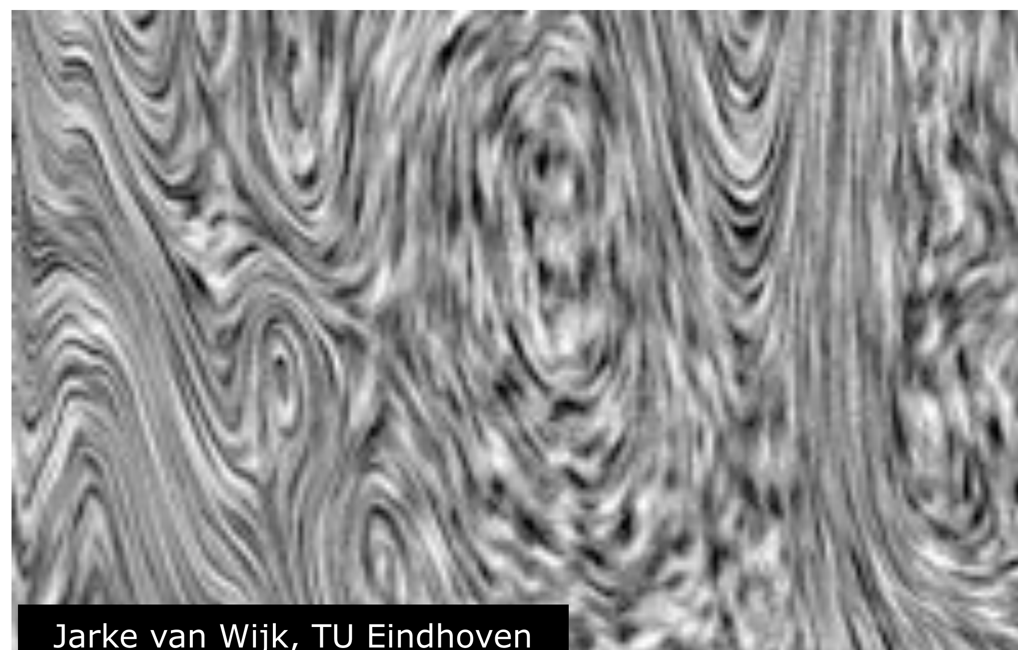
Vector Field Visualization



Experimental Fluid Dynamics

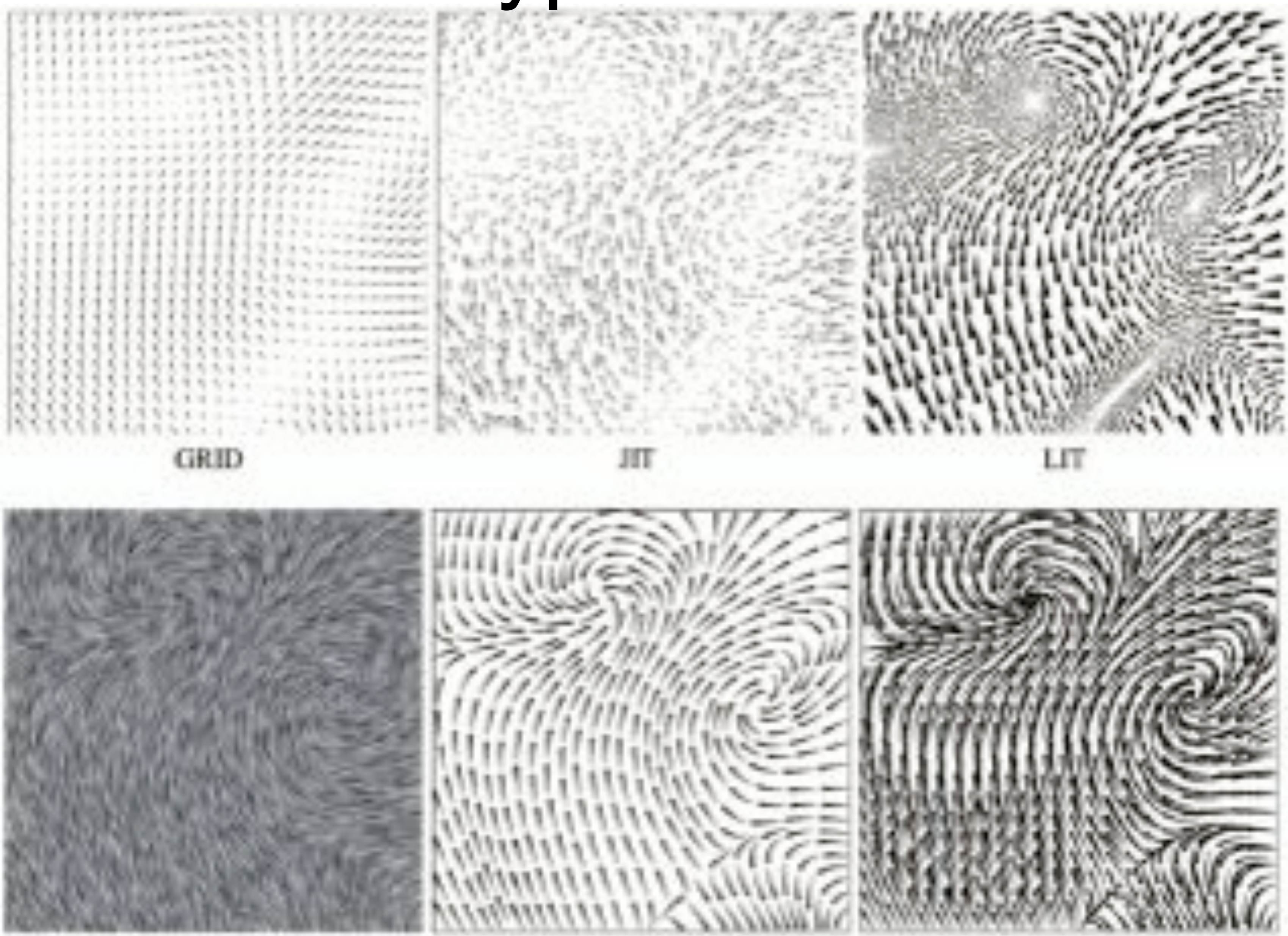
Visualization for understanding vortices, optimizing aerodynamics and turbine efficiency





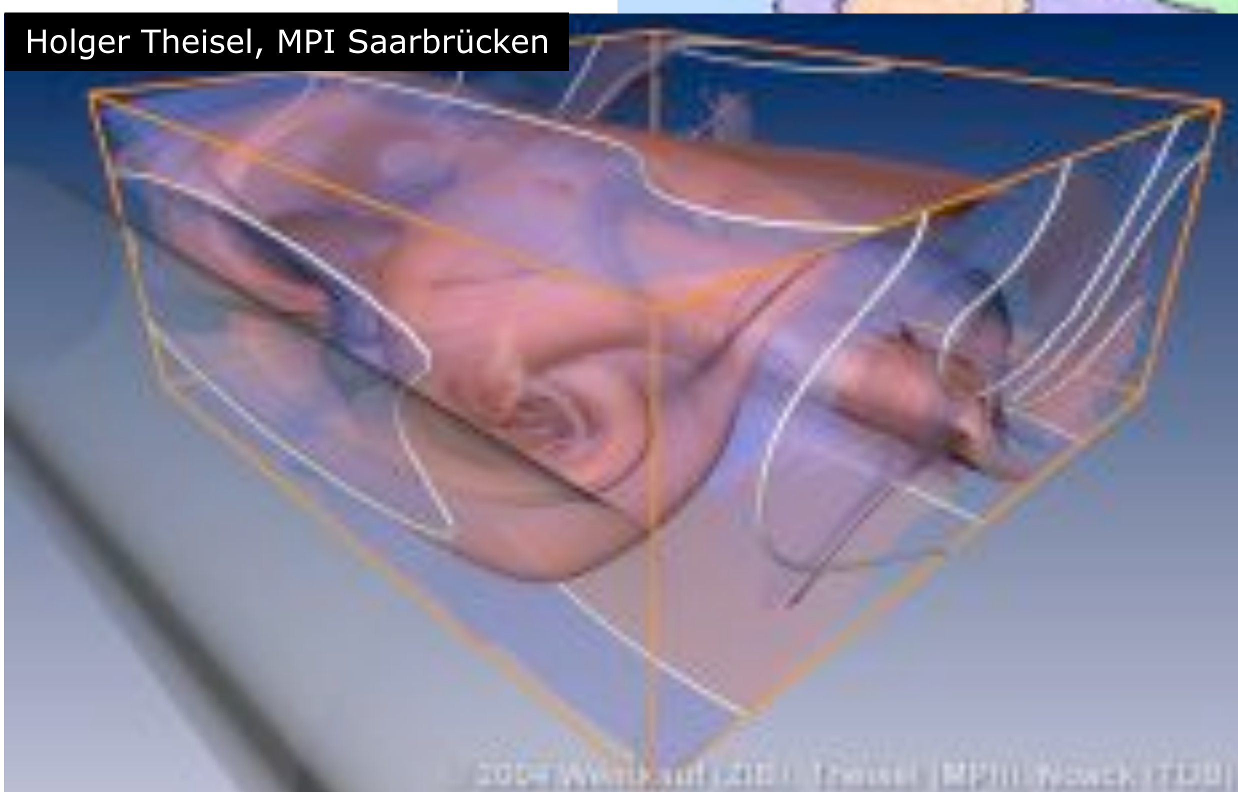
Flow visualization by Line Integral Convolution

Vector Field Glyphs



David H. Laidlaw, Robert M. Kirby, Cullen D. Jackson, J. Scott Davidson, Timothy S. Miller, Marco da Silva, William H. Warren, and Michael J. Tarr. **Comparing 2D Vector Field Visualization: Methods: A User Study**. Tranactions on Visualization and Computer Graphics 11(1):59-70, 2005

Vector Field Topology ~ River basins



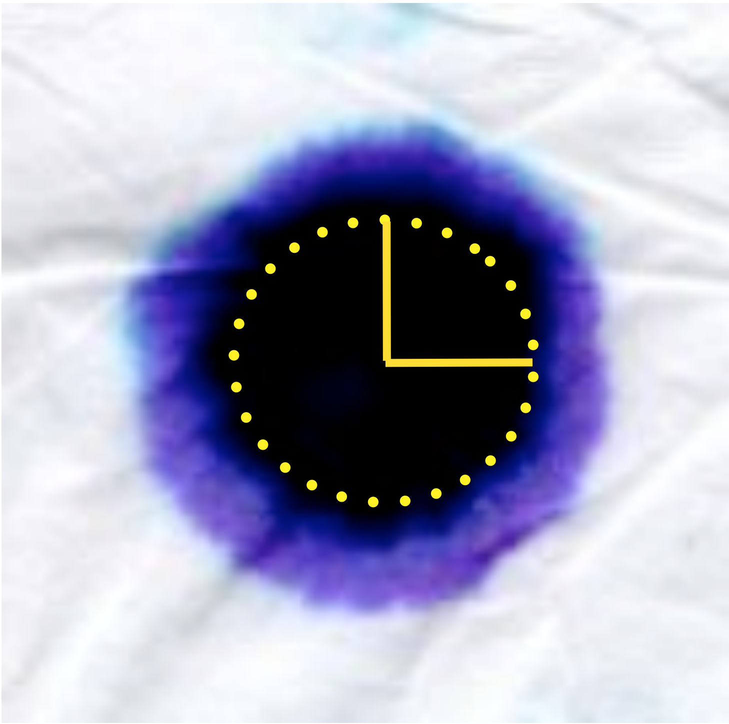
http://www.esrl.noaa.gov/psd/climateinfo/drought.html

What do tensors describe?

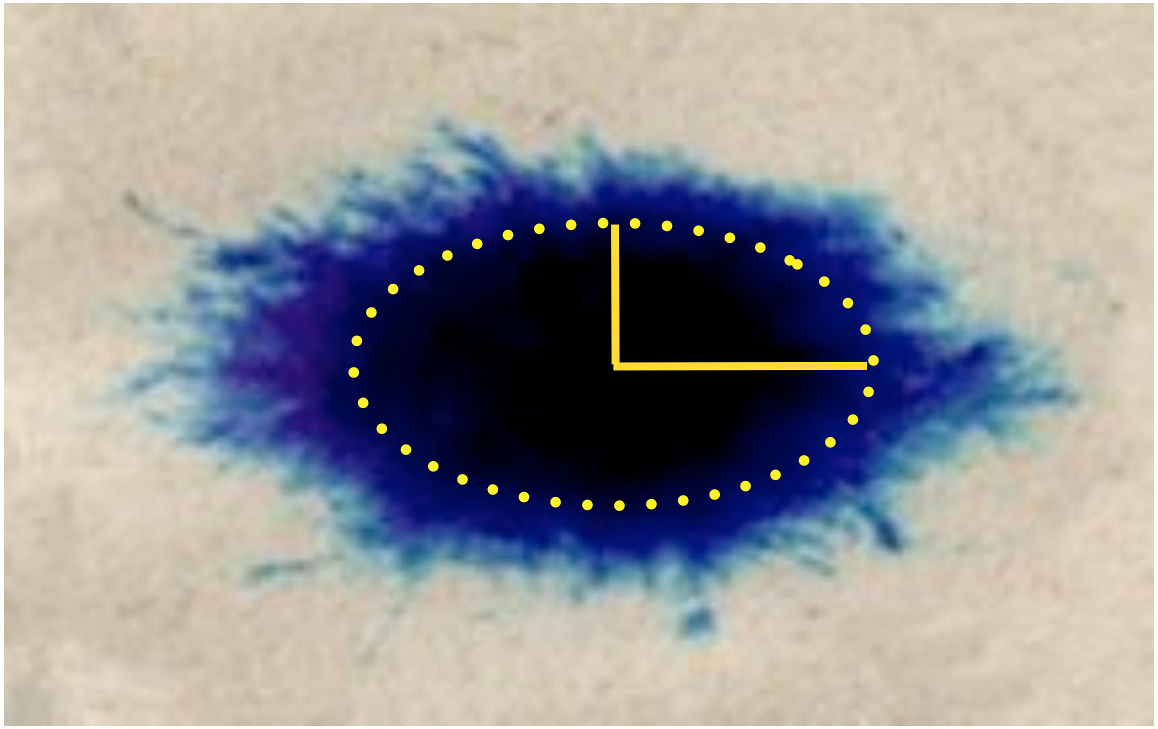
Anisotropy: (adj. anisotropic) how un-sphere-like

- Related to the variance of the eigenvalues
- More generally: "different in different directions"
- Isotropy (adj. isotropic) like a sphere, all eigenvalues equal

Anisotropy example: water diffusion in newspaper

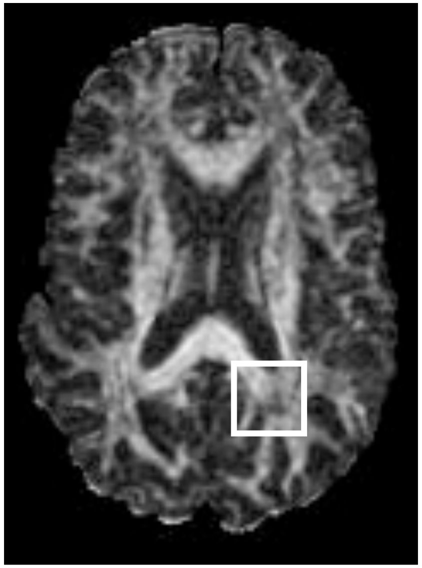


Kleenex: isotropic

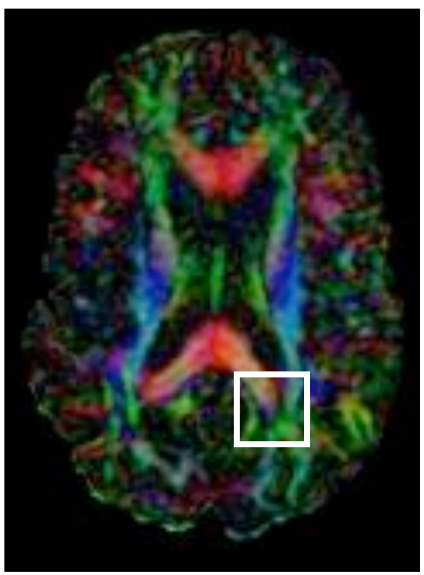


Newspaper: anisotropic

Images created by Gordon Kindlmann glk@uchicago.edu

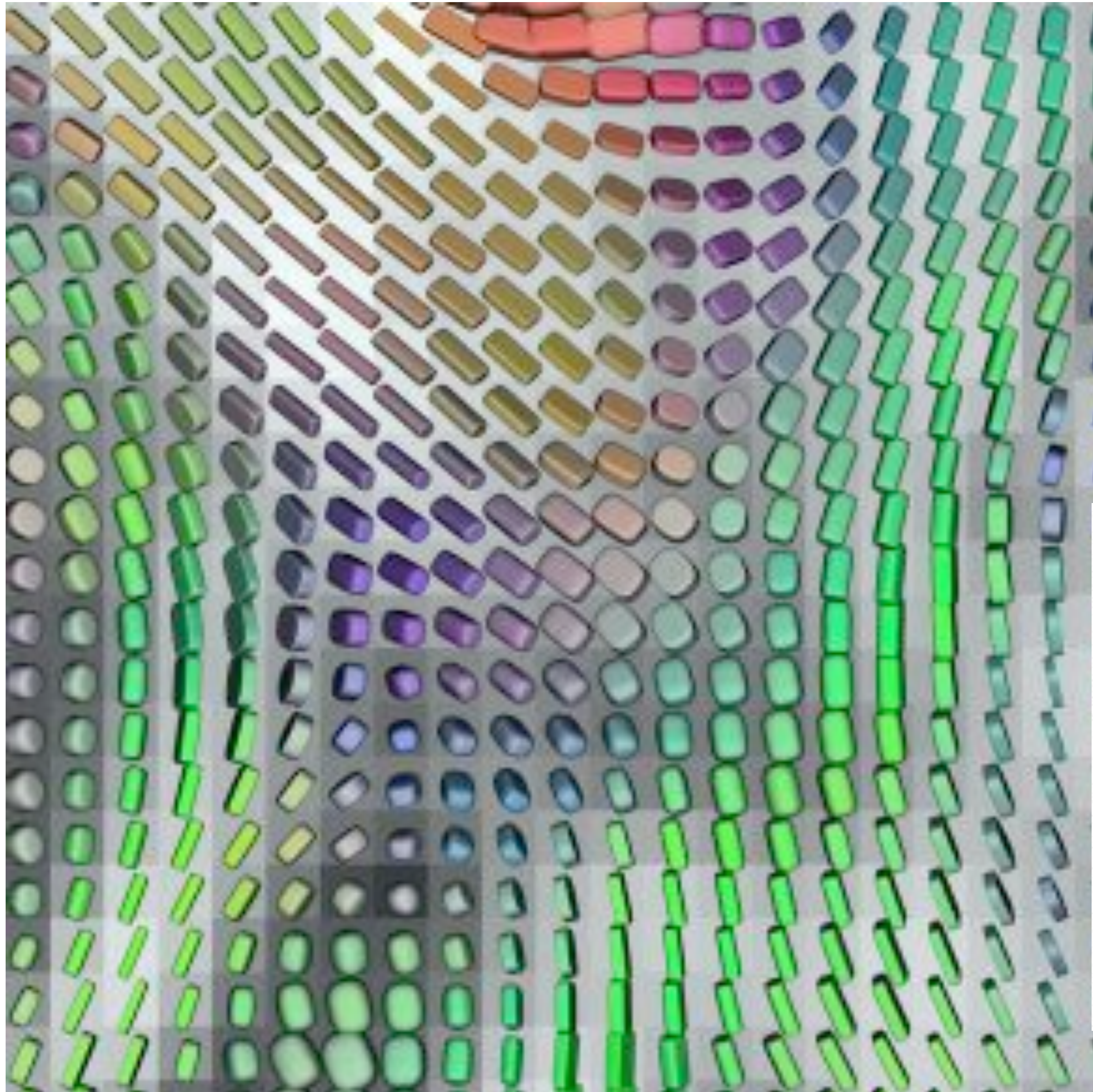


Grayscale: FA
Fractional Anisotropy

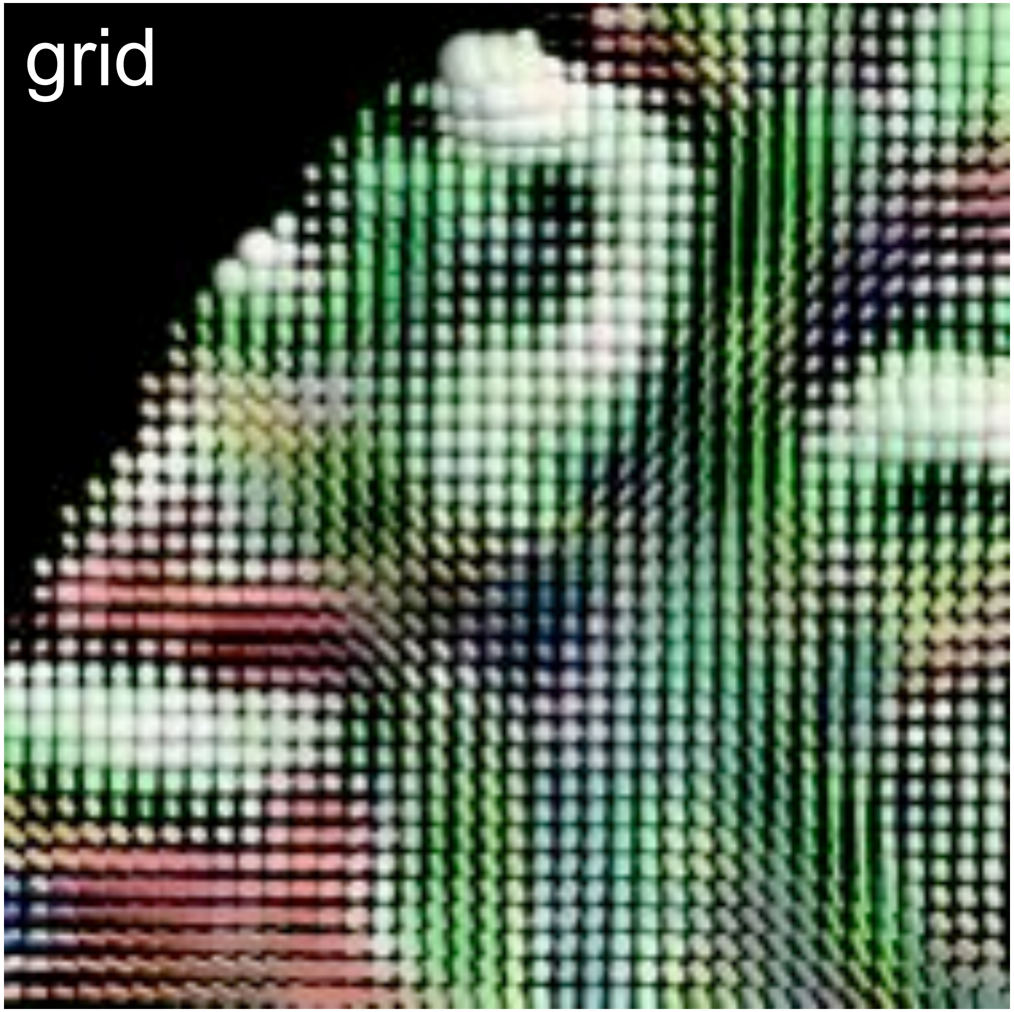


Color: RGB($\mathbf{v}_1$)
$\mathbf{v}_1$ principle eigenvector

Glyphs: full diffusion tensor →



Can break free of rectangular grid



Images created by Gordon Kindlmann glk@uchicago.edu

Gordon Kindlmann and Carl-Fredrik Westin. **Diffusion Tensor Visualization with Glyph Packing**. IEEE Trans. on Visualization and Computer Graphics, 12(5):1329-1335 2006

Images created by Gordon Kindlmann glk@uchicago.edu

Uncertainty visualization

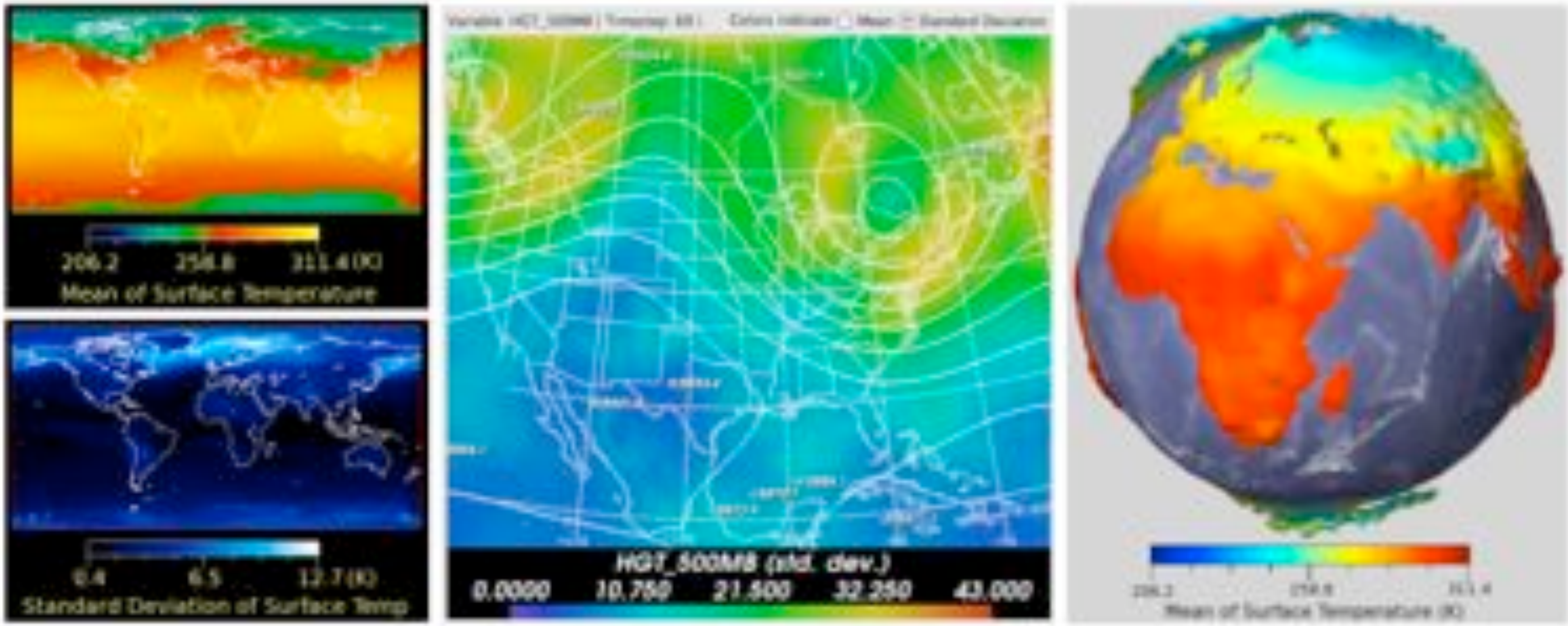
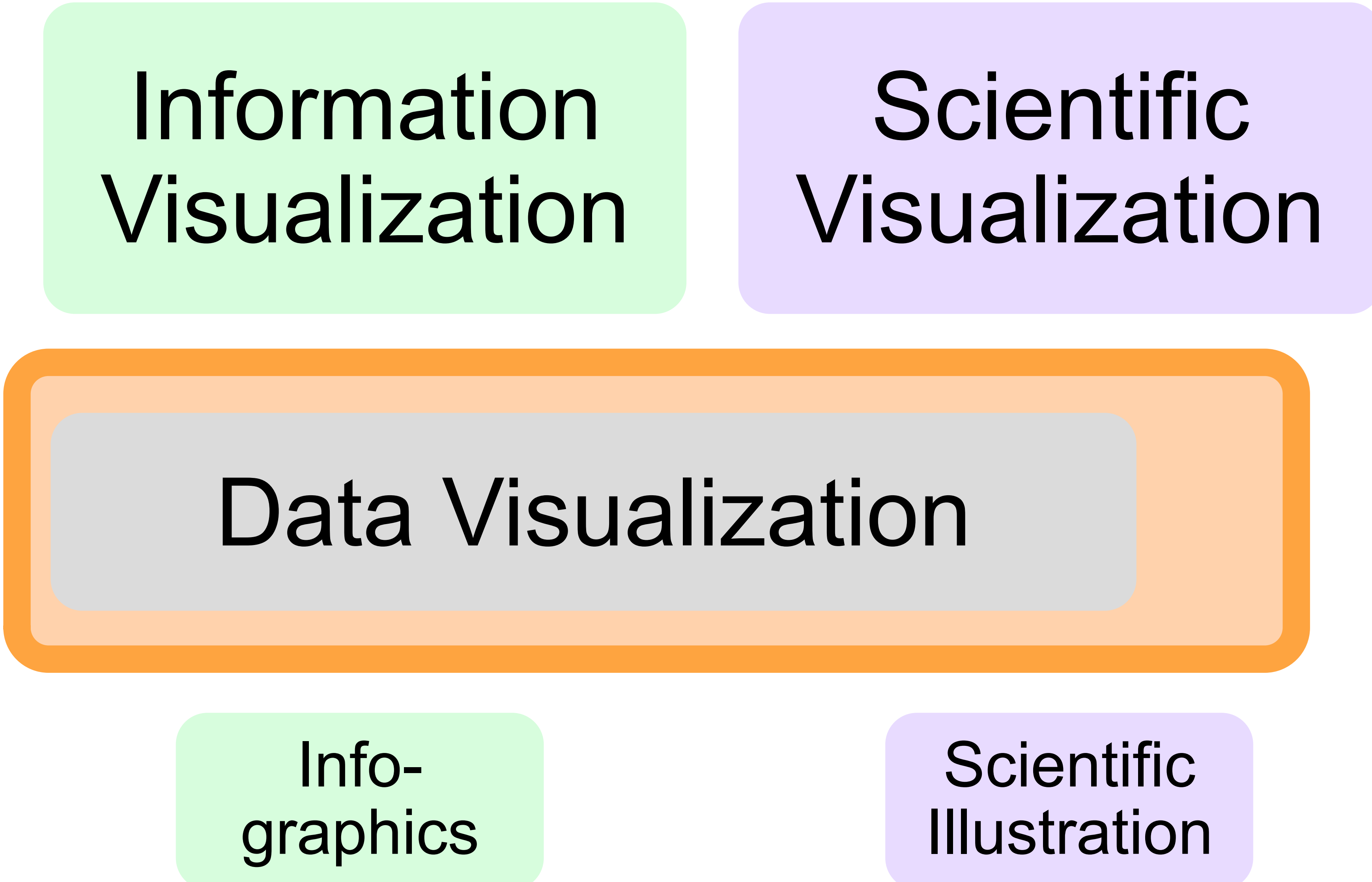


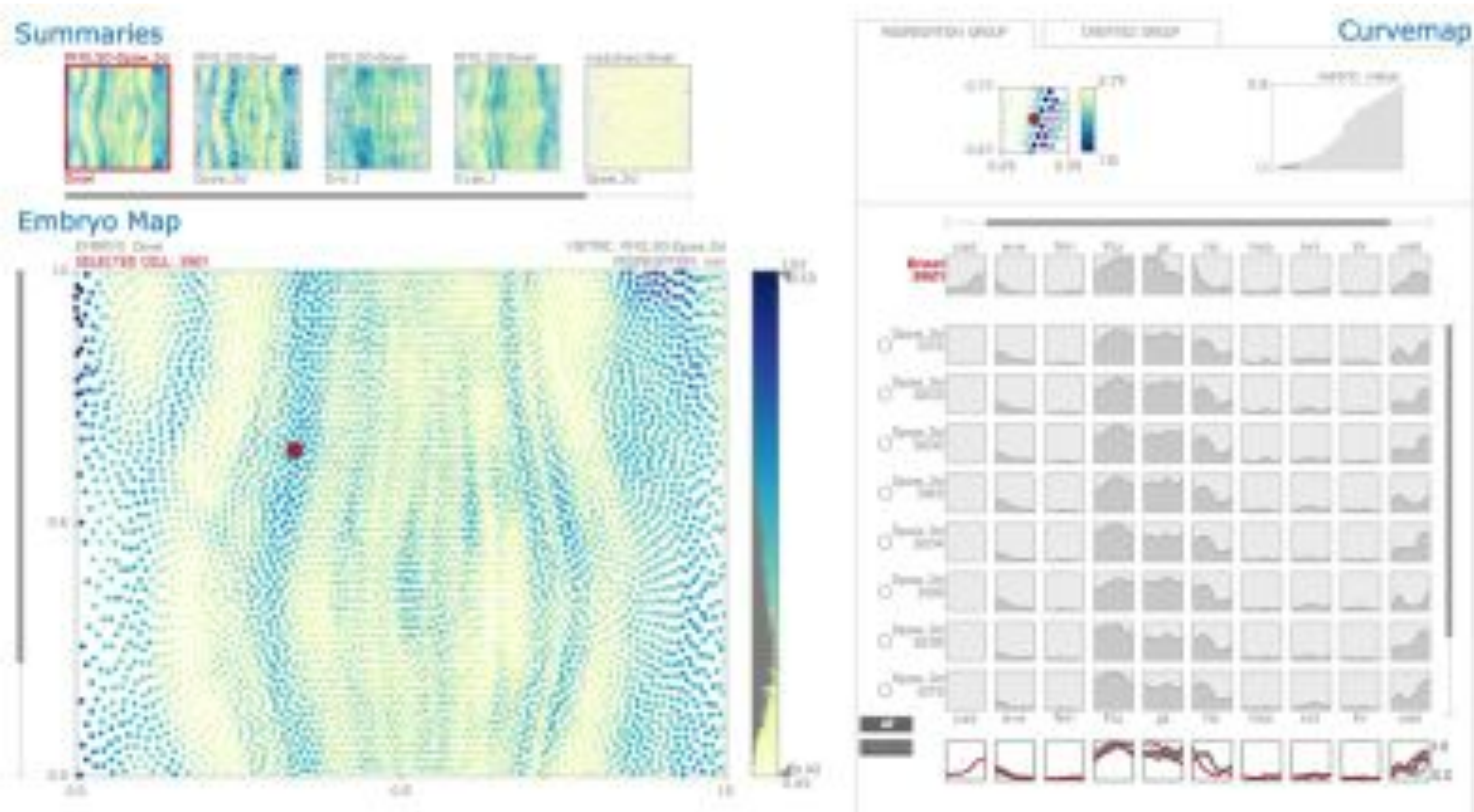
Figure 3. Visualizations of mean and standard deviation. (Left) Mean and standard deviation are visualized independently using color maps. (Center) Mean is presented through a color map, and standard deviation is shown as an overlaid contour. (Right) Standard deviation is mapped to a height field and mean is color mapped.

http://www.sci.utah.edu/~kpotter/library/papers/potter:2009:VUED/index.html

Fields of Visualization

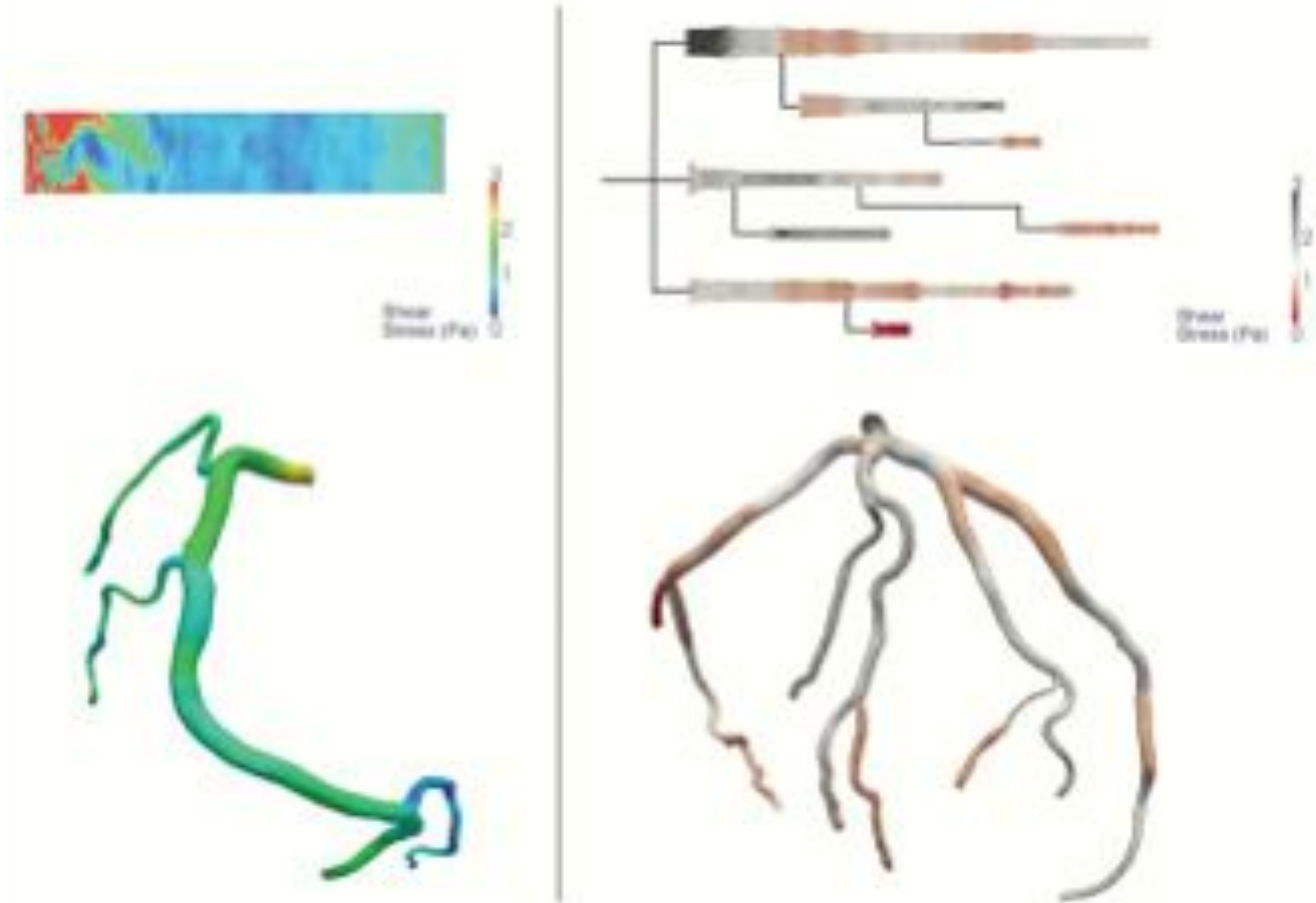


Gene expression data



Miriah Meyer, Tamara Munzner, Angela DePace, Hanspeter Pfister. **MulteeSum: A Tool for Comparative Spatial and Temporal Gene Exprssion Data** IEEE Transactions on Visualization and Computer Graphics (Proc InfoVis 2010), 16(6):908-917, 2010.

Blood vessel visualization by straightening

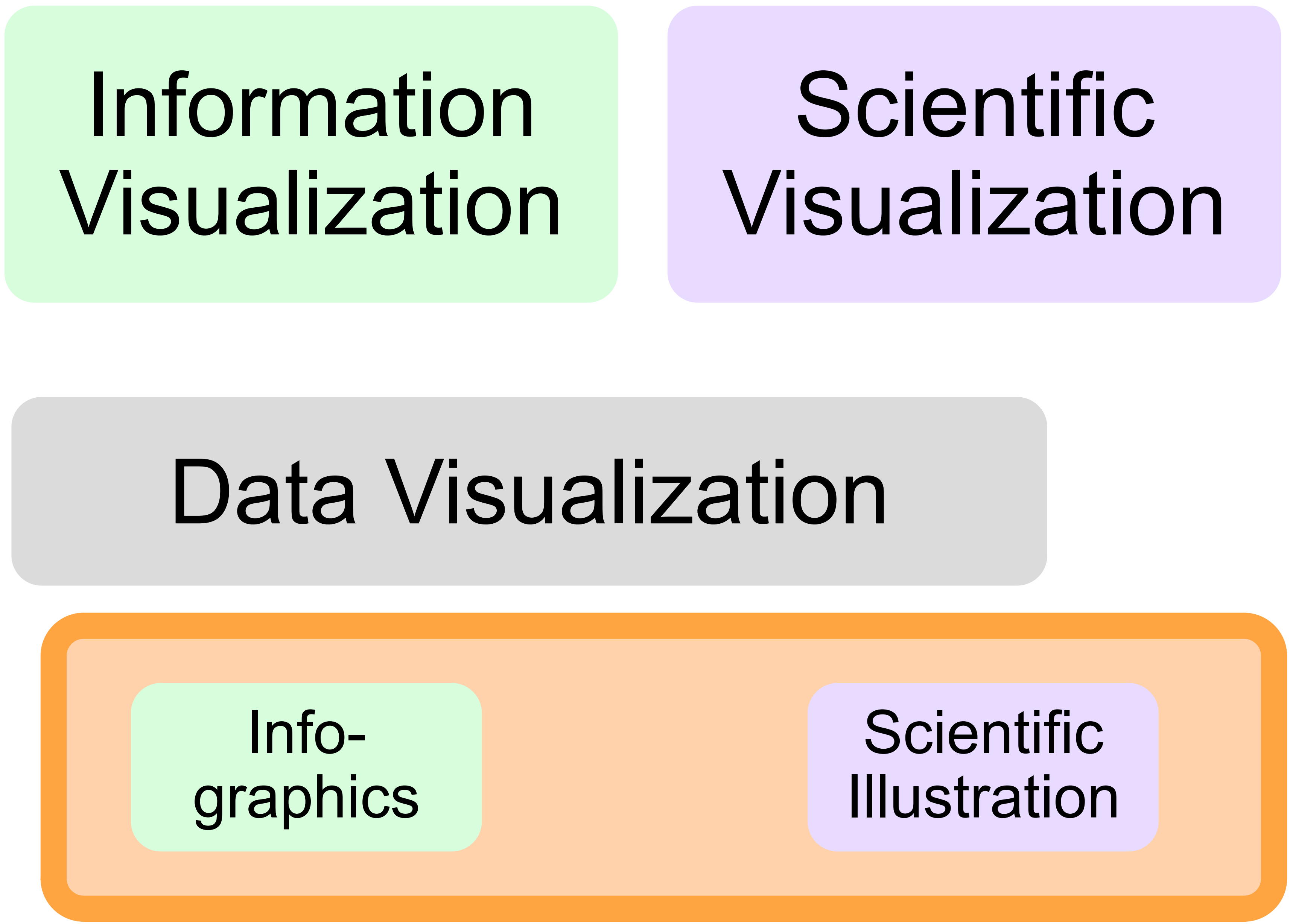


Michelle A. Borkin, Krzysztof Z. Gajos, Amanda Peters, Dimitrios Mitsouras, Simone Melchionna, Frank J. Rybicki, Charles L. Feldman, and Hanspeter Pfister. **Evaluation of Artery Visualizations for Heart Disease Diagnosis**. Proc Transactions on Visualization and Computer Graphics 17(2): 2479-2488 2011

River straightening

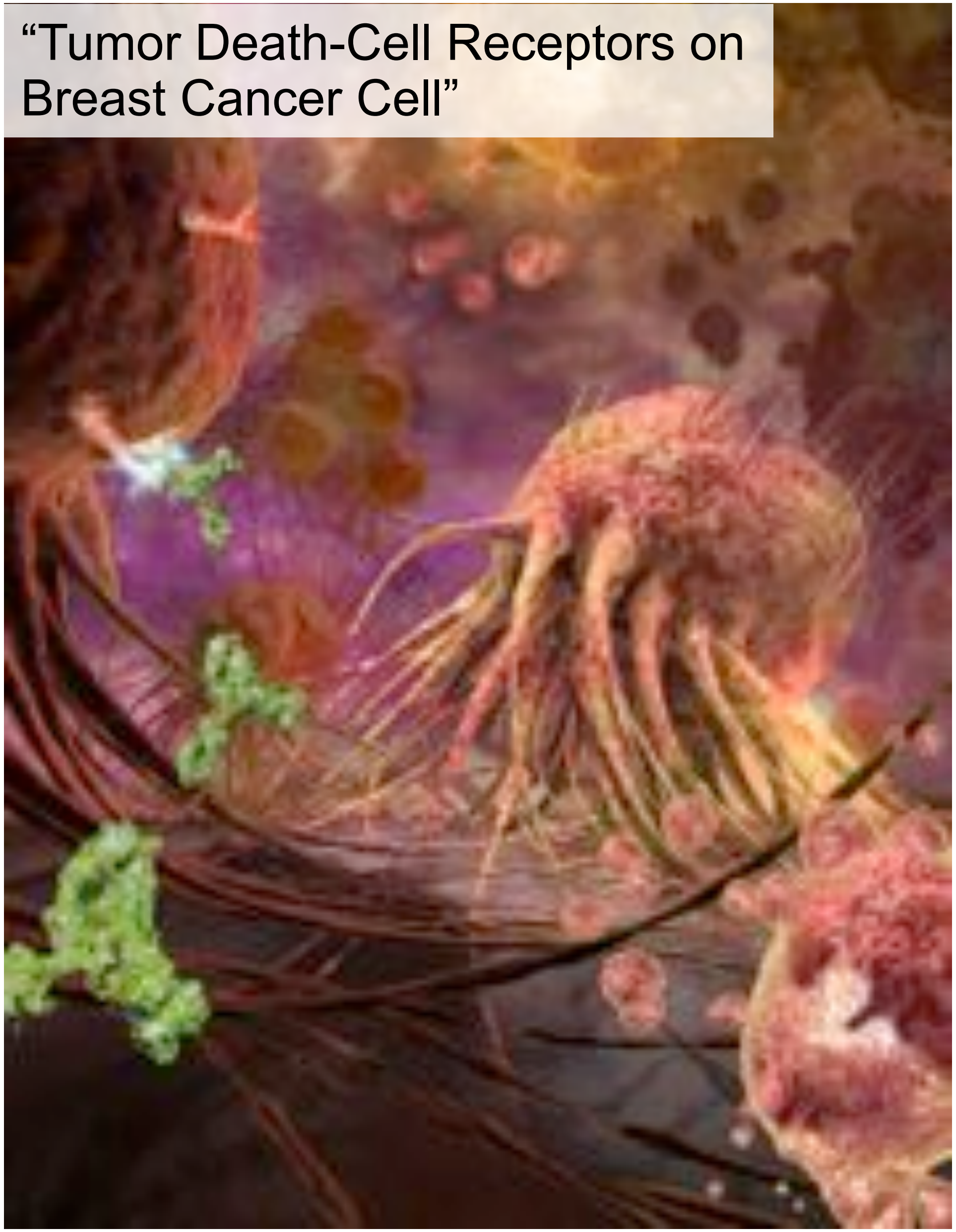
http://www.flickr.com/photos/85009674@N00/2931106116/sizes/l/

Fields of Visualization



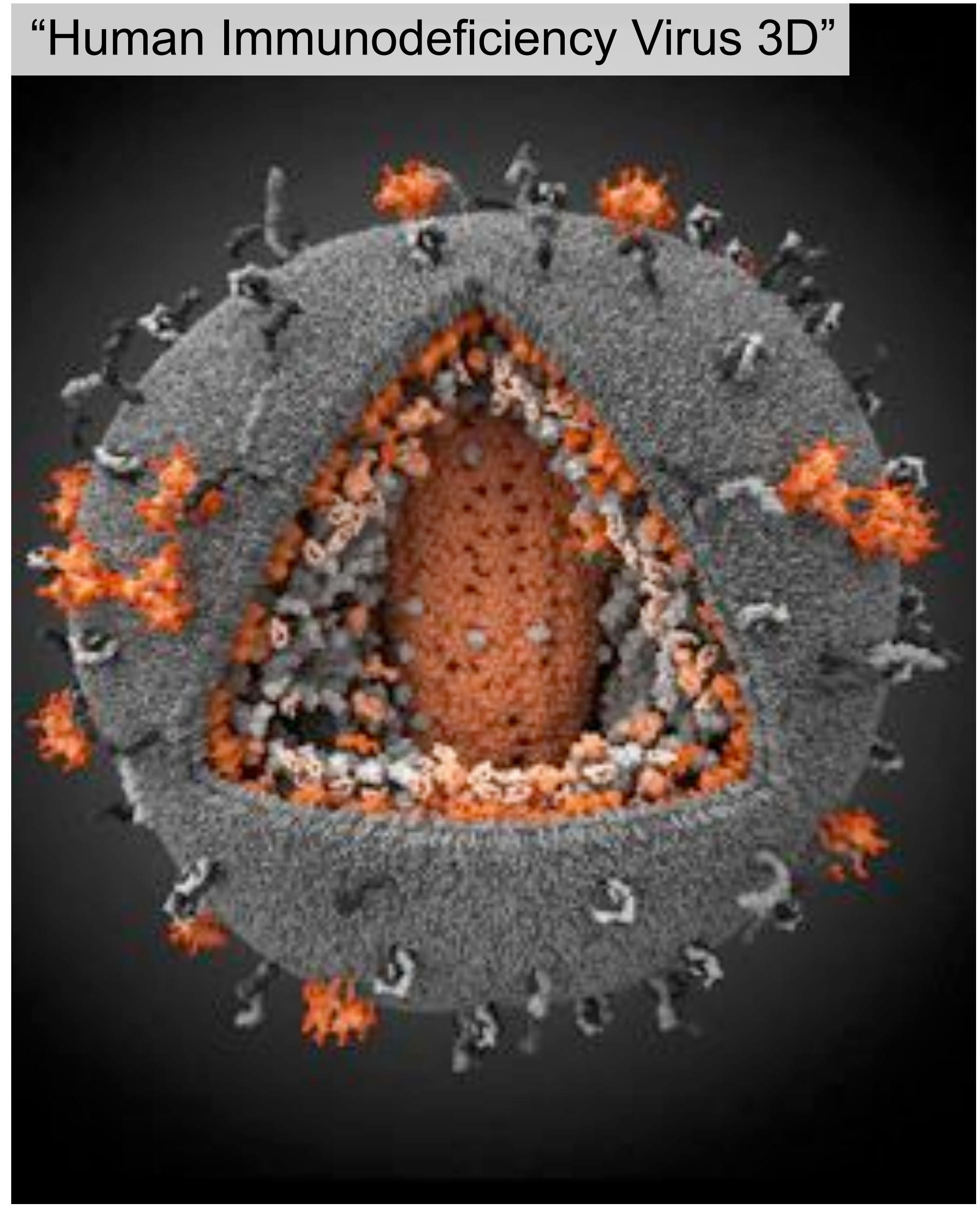
Scientific Illustration

http://www.nsf.gov/news/special_reports/scivis/challenge.jsp



“Tumor Death-Cell Receptors on Breast Cancer Cell”

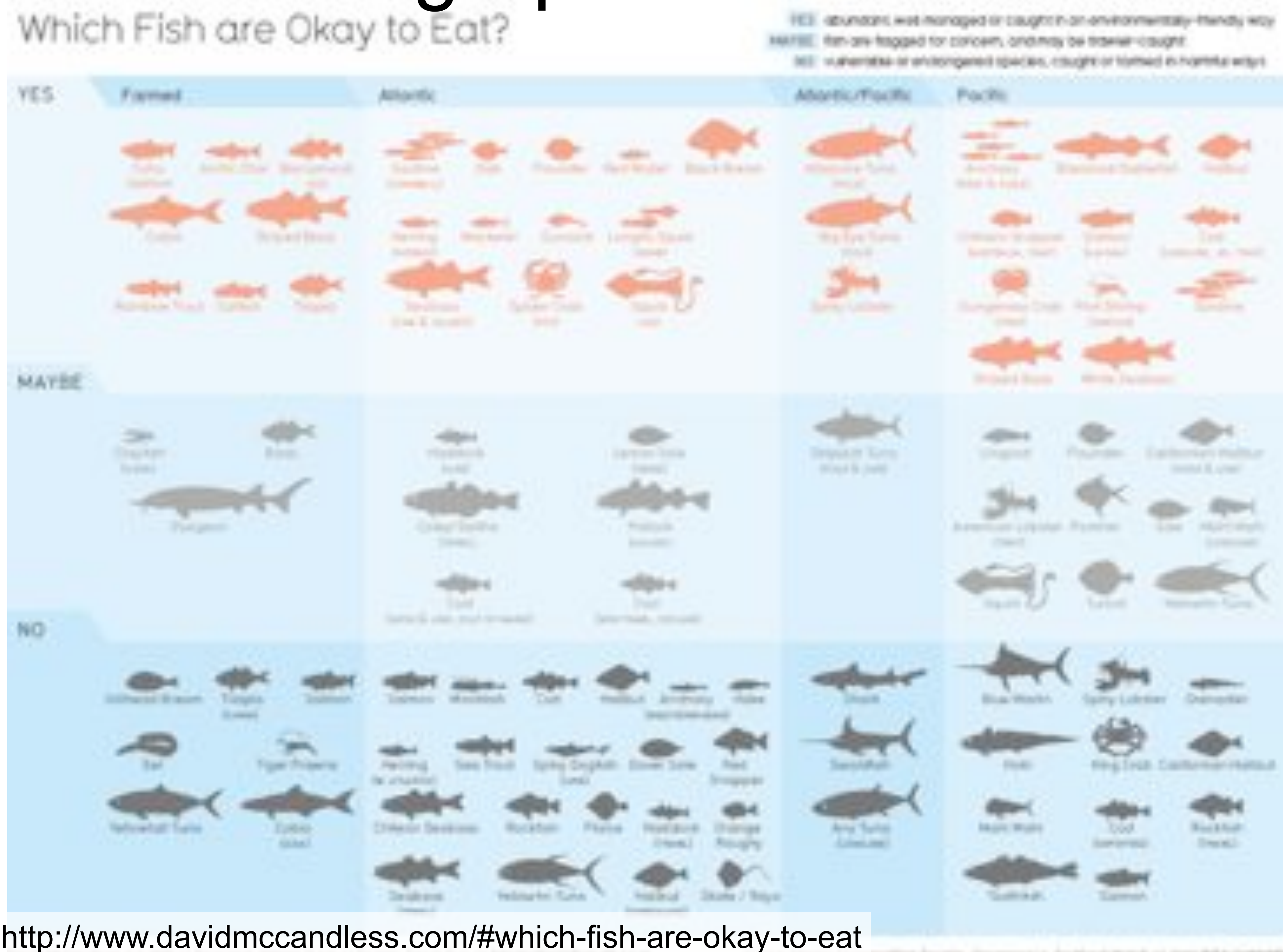
Emiko Paul, Quade Paul, Echo Medical Media; Ron Gamble, UAB Insight



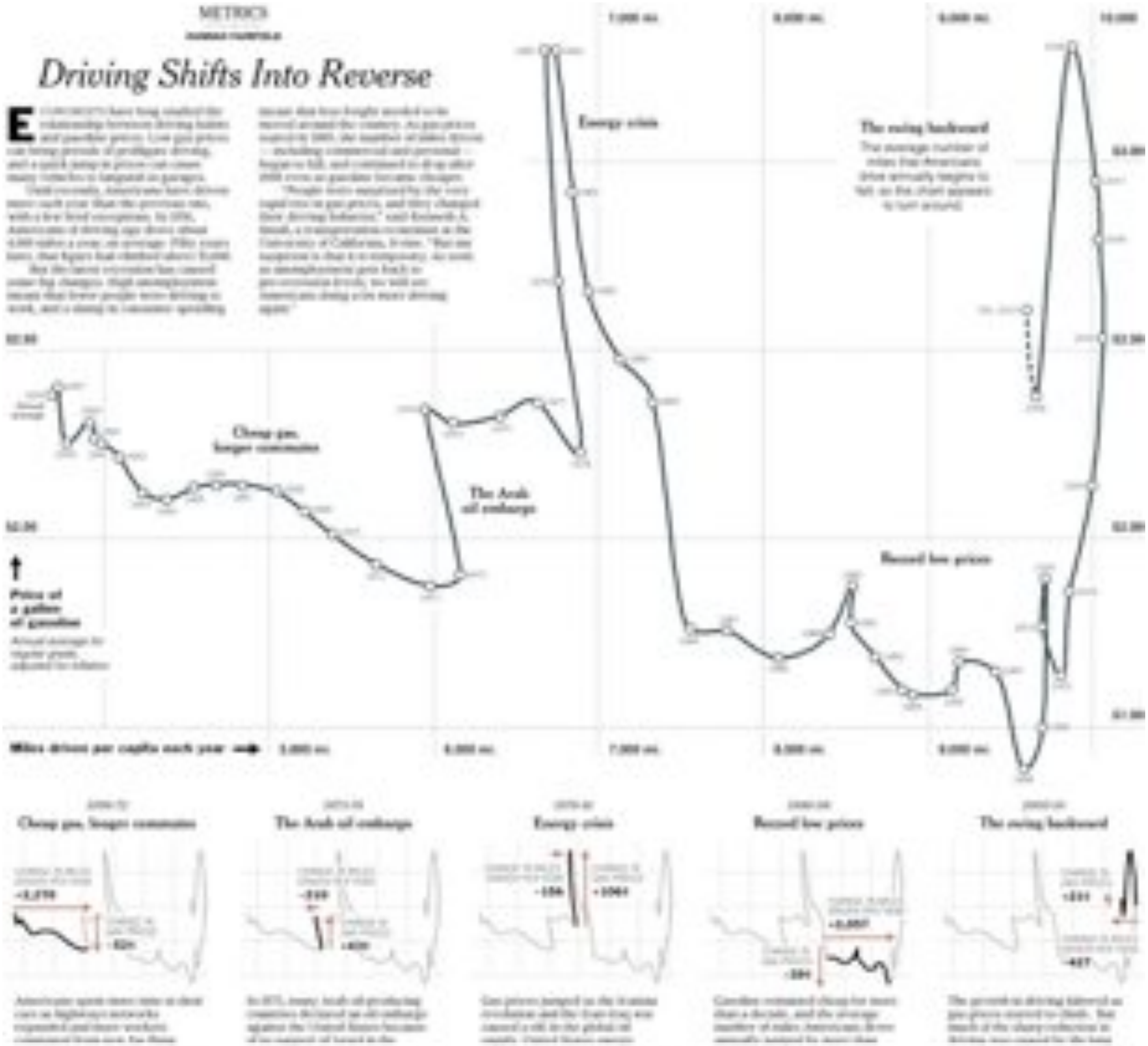
“Human Immunodeficiency Virus 3D”

Ivan Konstantinov, Yury Stefanov, Aleksander Kovalevsky, Yegor Voronin, Visual Science Company

Information graphics



http://www.davidmccandless.com/#which-fish-are-okay-to-eat



http://www.nytimes.com/imagepages/2010/05/02/business/02metrics.html

Explosion of “data science”

New job titles:

“data journalist”, “data designer”, “data scientist”

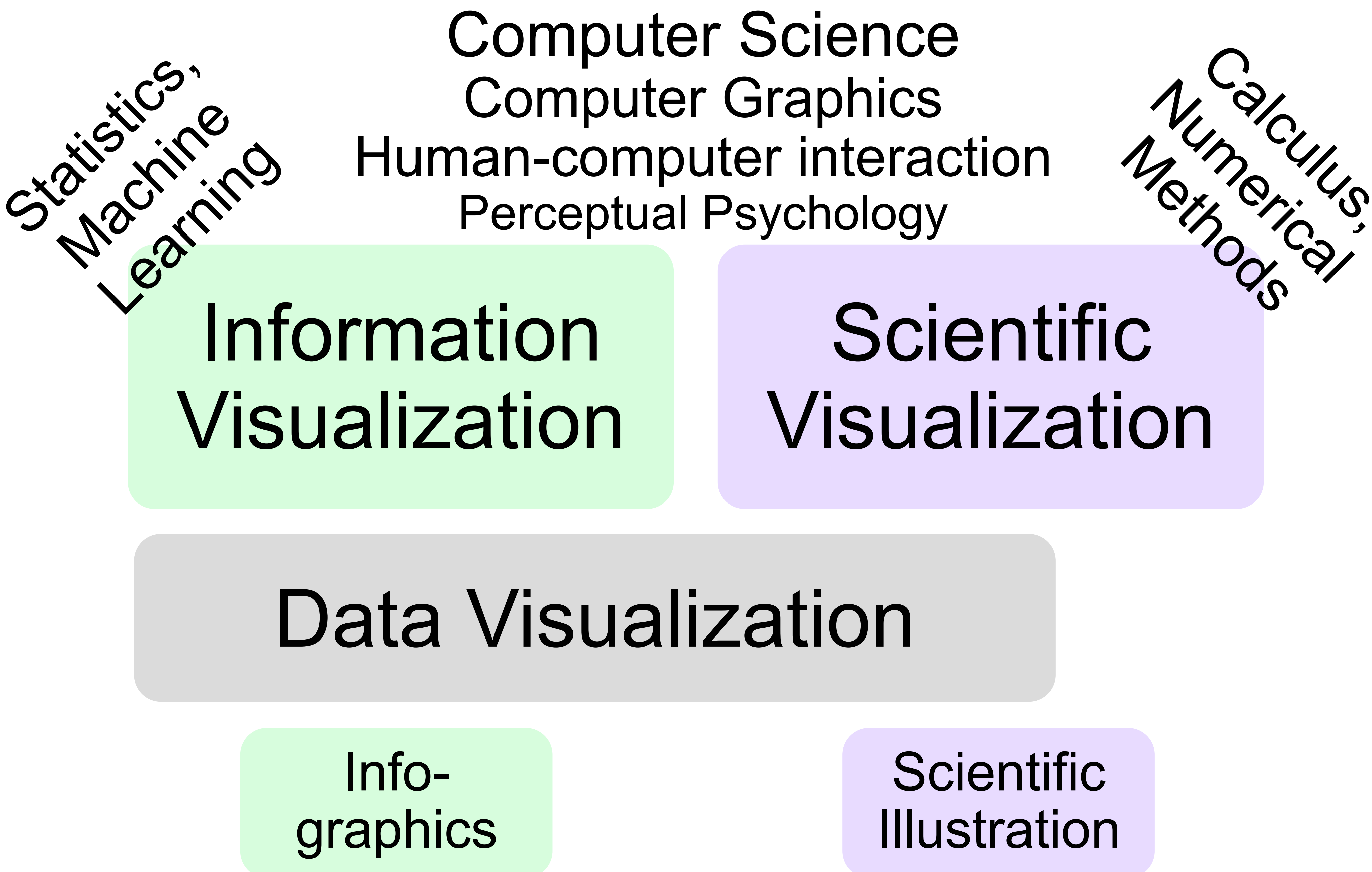
Many sites with pretty examples:

http://flowingdata.com
http://www.visualizing.org
http://www.datapointed.net
http://weloveinfographics.info
http://www.coolinfographics.com
http://ilovecharts.tumblr.com
http://fathom.info
http://infosthetics.com
http://visual.ly

Companies: http://www.tableausoftware.com

Competitions: http://www.kaggle.com

Fields of Visualization



Frederik de Wit
“Novissima et accuratissima totius Americae descriptio” (1670)
http://maps.bpl.org/id/m8711

World's Most Accurate Pie Chart

Thanks for your attention!